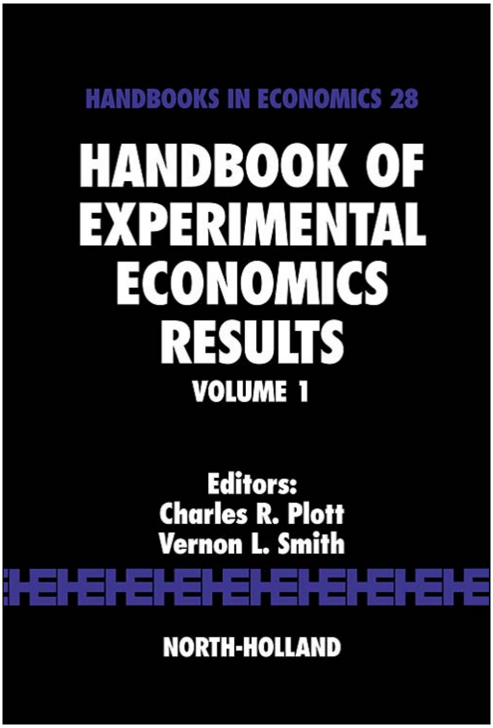


Nicole Bouchez and Daniel Friedman, Equilibrium Convergence in Normal Form Games. In: C.R. Plott and V.L. Smitt, editors, Handbook of Experimental Economics Results, Volume 1. Amsterdam: North-Holland, 2008, p. 472.
ISBN: 978-0-444-82642-8

Chapter 35

LEARNING TO FORECAST RATIONALLY

HUGH KELLEY and DANIEL FRIEDMAN

Economics Department, University of Santa Cruz, Santa Cruz, CA 95064, USA

1. Introduction

Economists routinely assume that all participants in the economy are rational forecasters who can correctly incorporate all available information when they form expectations of price and other variables that matter to them. Economists such as Marcet and Sargent (1989) point out that, when assessing the relevance of rational expectations models, researchers must ask whether or not repeat experience allows people to closely approximate rational forecasts. If people exhibit systematic departures from rational forecasts, then much of economic theory needs reconstruction.

The empirical literature on forecast rationality is surprisingly thin. Surveys of consumers, professional economists, and other market participants generally find that forecast errors have a non-zero mean, are correlated with other observable information, and follow an adaptive process (Camerer, 1995, pp. 609–611). Laboratory experiments with discrete forecasting tasks often indicate persistent biases (e.g., Grether, 1990). The most relevant previous experiment, Williams (1987, pp. 1–18) finds autocorrelated and adaptive errors when laboratory market participants forecast next period's market price. Other ties to existing literature can be found in Kelley (1998) and Kitzis et al. (1997).

The experiment described in this chapter isolates the forecasting process in two different stochastic individual choice tasks. The first task is based on Roll (1984), who finds that even in a very simple field financial market (Florida Orange Juice futures) where only two news variables are relevant (Florida weather hazard and competing supply, mainly from Brazil), the news can only account for a small fraction of the price variability. The second task is a variant of psychologists' standard discrete Medical Diagnosis task, e.g., Gluck and Bower (1988).

2. The Tasks

2.1. Orange Juice Forecasting (OJ)

In each trial, a subject views two continuous variables on her monitor: x_1 (called weather hazard) and x_2 (called Brazil supply), as in the upper-left corner of Figure 1. The values of x_1 and x_2 are independent random draws from the uniform distribution on (0, 100). The task is to forecast the dependent variable y (called orange juice futures price). The

Handbook of Experimental Economics Results, Volume 1

DOI: 10.1016/S1574-0722(07)00035-2

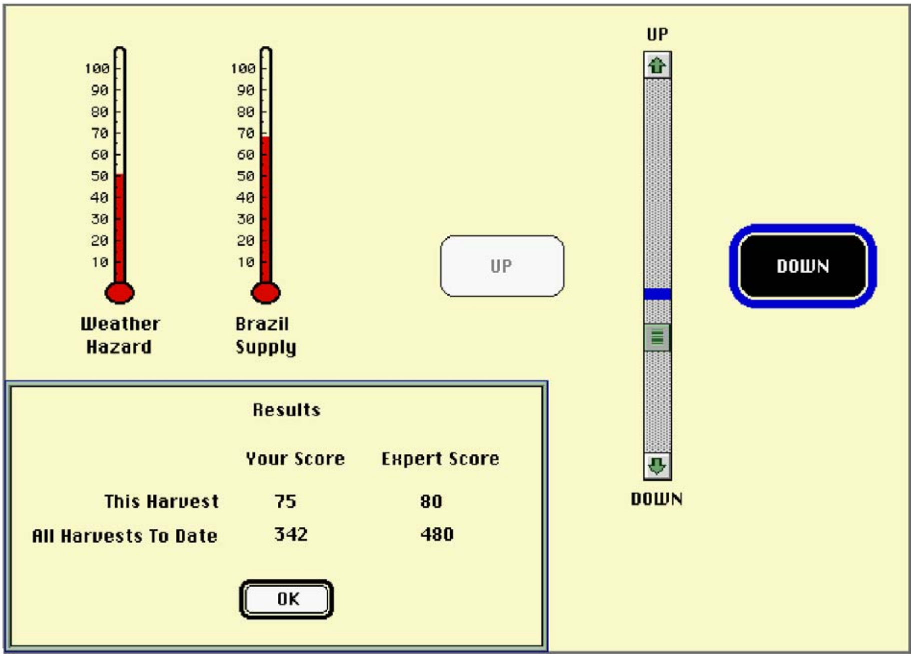


Figure 1. The values of x_1 (Weather Hazard) and x_2 (Brazil Supply) are random draws from the uniform distribution on (0, 100). The subject uses the slide bar on the right side of the display to enter her forecast. She uses the mouse to slide the small box up or down to enter her forecast. Then the realized value of the price appears on the same slide bar as a blue (dark) bar. Finally, in some treatments a score box is provided in the lower left of the screen. Score calculations are described in the text.

subject uses the slide bar on the right side of the monitor display to enter the forecast. The realized value of y appears on the same vertical display as the slide bar, as indicated in Figure 1. The realized value on trial t is

$$y_t = a_o + a_1 x_{1t} + a_2 x_{2t} + v e_t, \tag{1}$$

where the coefficients a_1 and a_2 are unknown to the subject (implicitly – they are the objects of learning) with baseline values 0.4 and -0.4, v is the noise amplitude (typically $v = 10$), e_t is an independent random variable drawn each trial from the uniform distribution on $(-1, 1)$, and a_o is the intercept used to center the data so that y_t falls at the middle of the vertical scale when $e_t = 0$ and $x_{1t} = x_{2t} = 50$.

In one treatment (called History), before entering her forecast the subject can view a summary of price outcomes from previous trials where values of x_1 and x_2 are close to current values. In another treatment (called Score), at the end of each trial the monitor displays a score S as shown at the bottom of Figure 1, computed from the continuous forecast c and the actual price p according to the quadratic scoring rule $S(c, p) =$

$80 - 280(c - p)^2$. [The "expert score" in Figure 1 uses the forecast c obtained from Equation (1) with e_t set to 0.]

Sessions last 480 trials, and so far 57 subjects have been tested. Most subjects' cumulative scores fall between 36,000 and 37,000, with a theoretical maximum score of $480 * 80 = 38{,}400$. Additional treatments include high noise ($v = 20$ instead of 10), asymmetric weights (e.g., $(a_1, a_2) = (0.24, -0.56)$ instead of $(0.40, -0.40)$) and structural breaks (e.g., weights shift from symmetric to asymmetric in trial 241). See Kelley (1998) for a more complete description of the task and treatments.

2.2. The Medical Diagnosis Task (MD)

The medical diagnosis task also is a stochastic individual choice task with 480 trials for each subject, two independent variables (the symptoms) and one dependent variable (the disease). The user interface is quite similar to that in the OJ task. However, the independent variables (temperature and blood pressure) are discrete with four possible values (high, medium high, medium low and low), and the dependent variable is binary (the disease is either Autochus or Burlosis). Conditional on the realized disease, the symptoms are independently drawn according to likelihoods unknown to the subject.

Subjects' continuous response c_t in trial t consists of naming the disease deemed more likely and indicating (with the slide bar) the degree of confidence. The response is coded as a continuous variable between 0 (completely confident that the disease is B) and 1 (completely confident that the disease is A). Kitzis et al. (1997) show that the true Bayesian relationship between symptoms and diseases can be very closely approximated by the linear equation

$$y_t = a_o + a_1 x_{1t} + a_2 x_{2t}, \tag{2}$$

where now y_t is the posterior log odds of disease A over disease B, and x_{it} is the discrete variable with values 1, 0.3, -0.3 and -1 respectively for high, medium high, medium low and low values of symptom $i = 1, 2$. The unknown coefficients a_1 and a_2 again are the implicit objects of learning; the true values are 1.39 and -2.30.

We tested 123 subjects in a 2×3 factorial design with the treatments History (vs No History) and Score (vs No Score and vs Score + Pay) with 20+ subjects in each cell.

3. Results

3.1. Rolling Regressions

Although a subject may think of the task in various idiosyncratic ways, the analyst can summarize the subject's beliefs by seeing how he responds to the current stimuli x_{it}. Moreover, the analyst can summarize the learning process by seeing how the subject's response to stimuli changes with experience.

Given Equations (1) and (2), learning thus can be seen in the changes over time in a subject's implicit subjective values of the coefficients a_1 and a_2. The data analysis reconstructs the implicit values of these coefficients and tracks their changes over time.

The reconstruction proceeds as follows. For the OJ task, take the subject's actual forecast c_t in trial t as the dependent variable, and take the actual values of x_{it} as independent variables. Then, run a rolling regression of c_t on the two independent variables over a moving window of 160 consecutive trials, incrementing the last trial T from 160 to 480. The procedure in the MD task is the same except that the dependent variable is the log odds of the continuous choice, $L(c_t) = \ln[(c_t + .01)/1.01 - c_t)]$; c_t is shifted by .01 away from 0 and 1 to avoid taking the log of zero. The intercept coefficient is constrained to its objective value in the results shown below in order to reduce clutter and to improve statistical efficiency.

Effective learning is indicated by rapid convergence of the coefficient estimates a_{iT} (as T increases) to the objective values a_i. Obstacles to learning are suggested by slow convergence, convergence to some other value, or divergence of the coefficient estimates. This empirical approach embodies some of the theoretical ideas on learning in Marcet and Sargent (1989) as explained in Kelley and Friedman (2002).

3.2. OJ Learning Curves

Figure 2 presents two examples. Top panel shows the simulated performance of a Marcet–Sargent econometrician who uses realized prices for all trials observed so far to estimate the coefficients a_1 and a_2 and then uses these coefficients in Equation (1) with $e_t = 0$ to forecast the current price. Learning seems immediate (within 160 trials). The R^2 for the first 160 trial window of data was 0.93 and ended at the same level, 0.93, for the last window.

Bottom panel shows that the actual subject who earned the top score came fairly close to the Marcet–Sargent ideal. The coefficient estimates indicate that he slightly overresponded to current symptoms throughout the session, but the overresponse was negligible by the last 160 trials. His R^2 for the first 160 trial window of data was 0.94 and increased to 0.96 by the last window. This high scoring subject is fairly representative; coefficient estimates for other subjects in most treatments sometimes indicate overresponse and sometimes underresponse, but on average are quite close to or slightly beyond objective values.

Table 1 summarizes the main departures from effective learning detected so far. Coefficient estimates indicating "Significant" under and overresponse by the end of the session ($T = 480$) are about equal in the baseline and asymmetric treatments, but underresponse is much more prevalent than overresponse in the structural break treatments. In the high noise treatment (amplitude $v = 20$ instead of $v = 10$), overresponse is much more common than underresponse. Figure 3 presents the corresponding histograms, which clearly show that coefficient estimates for subjects in the high noise environment tend strongly toward overresponse.

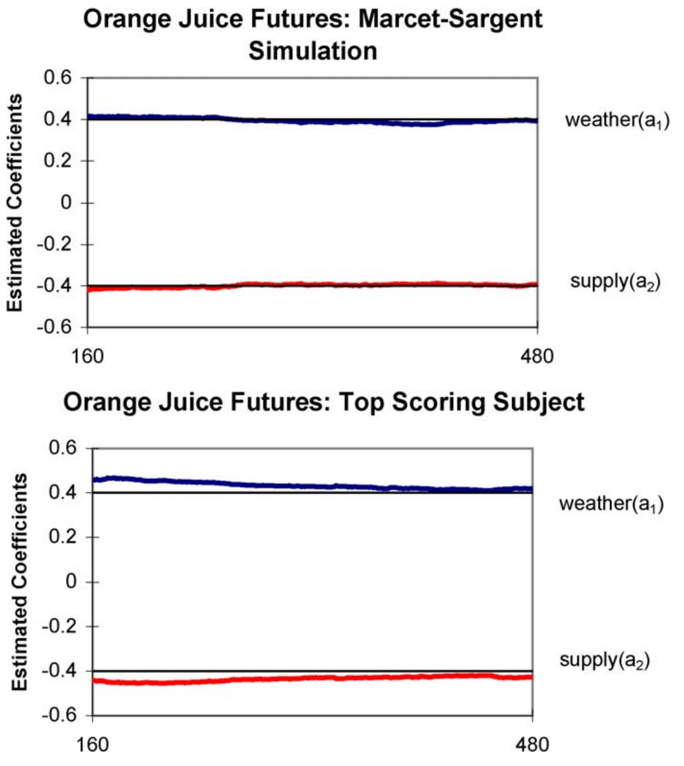


Figure 2. (Top) Coefficient estimates from rolling regression $y_t = a_1x_{1t} + a_2x_{2t} + ve_t$ on trials $\{T - 159, T - 158, \ldots, T\}$ for $T = 160$ to 480. The equation is estimated using the Marcet–Sargent Model to generate the forecasts y_t. (Bottom) Coefficient estimates from rolling regression $c_t = a_1x_{1t} + a_2x_{2t} + ve_t$ on trials $\{T - 159, T - 158, \ldots, T\}$ for $T = 160$ to 480. The equation is estimated using forecasts c_t from Subject 44.

Table 1
Over and under-response in Orange Juice forecasting

	Under response	Objective	Over response
Symmetric weights	20	11	21
Asymmetric weights	10	1	9
High noise	5	3	12
Structural break	13	2	7

Note: Coefficients a_1 and a_2 are estimated at $T = 480$ for all Ss for the equation $c_t = a_1x_{1t} + a_2x_{2t}$. Responses for each subject are classified as over or underresponse if the estimate differs from the objective value by more than $1.96 * \text{std error}$.

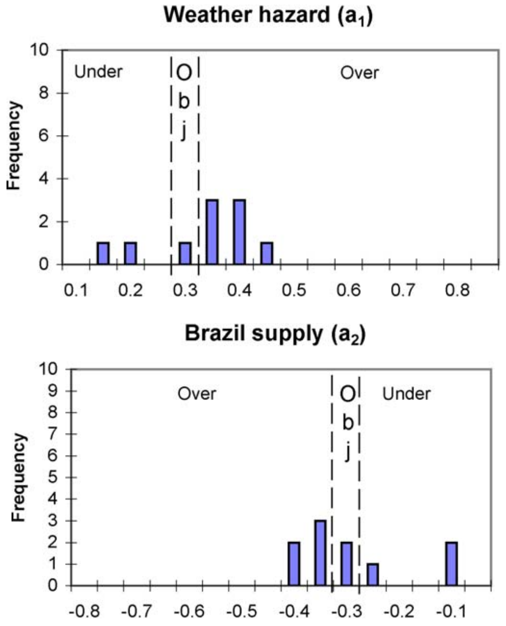


Figure 3. Distribution of final ($T = 480$) coefficient estimates in high noise treatment for the Orange Juice Futures experiment. "Obj" indicates estimates near objective values of (0.33, −0.33) respectively. "Under" (and "Over") refer to cases where the absolute value of the estimate is less (and more) than the objective value.

3.3. MD Learning Curves

Figure 4 provides two examples from the second experiment, Medical Diagnosis. Top panel simulates a Bayesian econometrician (the MD counterpart of Marcet–Sargent) who uses realized disease outcomes for all trials observed so far to estimate the coefficients a_1 and a_2 and then uses these coefficients in Equation (2) with $e_t = 0$ to predict the current disease. Ideal learning is a bit slower and more erratic than in the OJ task, but it still converges to the true values $a_1 = 1.39$ and $a_2 = -2.3$ quite rapidly.

Bottom panel shows that the actual subject who earned the top score differs noticeably from the Bayesian ideal. The coefficient estimates for this subject indicate persistent overresponse and are quite representative of the subject pool. The histograms in Figure 5 confirm that overresponse is indeed the prevailing bias in our MD data.

4. Discussion

Kelley (1998) reports several robustness checks. OJ specifications designed to capture prior beliefs and non-linear responses to news detected some transient effects in many subjects, but for the most part the final regression is indistinguishable from the basic

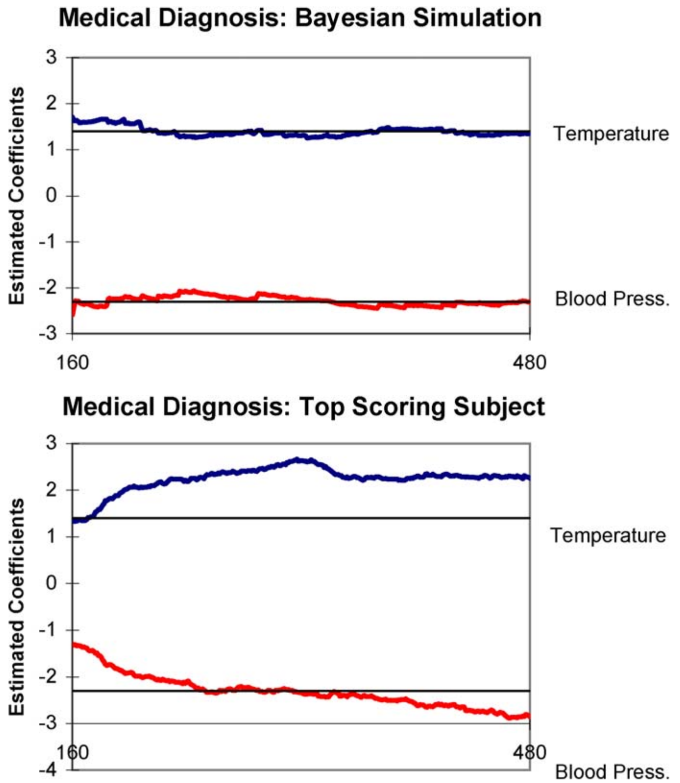


Figure 4. (Top) Coefficient estimates from rolling regression $y_t = a_1x_{1t} + a_2x_{2t}$ on trials $\{T - 159, T - 158, \ldots, T\}$ for $T = 160$ to 480. The equation is estimated using the Bayesian model to generate the forecasts y_t. (Bottom) Coefficient estimates from rolling regression $c_t = a_1x_{1t} + a_2x_{2t}$ on trials $\{T - 159, T - 158, \ldots, T\}$ for $T = 160$ to 480. The equation is estimated using the forecasts ct from Subject 28.

specification presented above. Eight parameter MD specifications that allow separate learning for each level of each symptom also converged roughly to the basic specification presented above, but we detect a general bias towards overresponding to the more informative symptom levels and underresponding to the less informative symptom levels.

We draw three conclusions from the data analysis. First, the rationality assumption is a good first approximation to subjects' forecasts at the end of 480 learning trials. Second, systematic biases towards under or overresponse can be detected in specific circumstances, e.g., overresponse in the noisier OJ environment. Third, more experiments are needed in a wider variety of tasks and environments in order to understand more fully when people can learn to forecast rationally. We anticipate that the rolling regressions and learning curves featured in this chapter will continue to be a useful tool in that research.

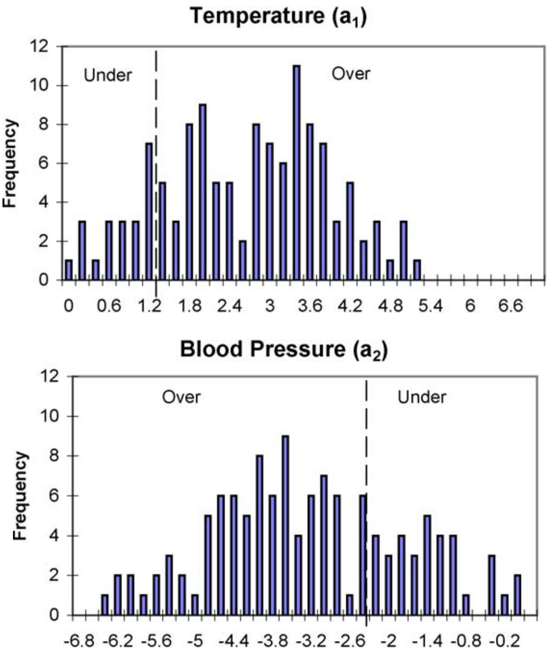


Figure 5. Distribution of final ($T = 480$) coefficient estimates for the Medical Diagnosis experiment. Objective values indicated by vertical dashed line. Objective values are (1.39) for Temperature and (−2.3) for Blood Pressure.

References

Camerer, C. (1995). "Individual decision making". In: Kagel, J.H., Roth, A.E. (Eds.), The Handbook of Experimental Economics. Princeton University Press, Princeton, NJ, pp. 587–703.

Gluck, M.A., Bower, G.H. (1988). "From conditioning to category learning: An adaptive network model". Journal of Experimental Psychology: General 117, 225–244.

Grether, D.M. (1990). "Testing Bayes rule and the representativeness heuristic: Some experimental evidence". Journal of Economic Behavior and Organization 17, 31–57.

Kelley, H. (1998). "Bounded rationality in the individual choice experiment". Unpublished thesis, Economics Department, University of California, Santa Cruz.

Kelley, H., Friedman, D. (2002). "Learning to forecast price". Economic Inquiry 40, 556–573.

Kitzis, S., Kelley, H., Berg, E., Massaro, D., Friedman, D. (1997). "Broadening the tests of learning models". Journal of Mathematical Psychology 42, 327–355.

Marcet, A., Sargent, T. (1989). "Convergence of least squares learning mechanisms in self referential linear stochastic models". Journal of Economic Theory 48, 337–368.

Roll, R. (1984). "Orange juice and weather". American Economic Review 74, 861–880.

Williams, A.W. (1987). "The formation of price forecasts in experimental markets". Journal of Money, Credit, and Banking 19, 1–18.

Chapter 53

EQUILIBRIUM CONVERGENCE IN NORMAL FORM GAMES

NICOLE BOUCHEZ and DANIEL FRIEDMAN

Economics Department University of California, Santa Cruz

In this chapter we examine convergence behavior in simple bimatrix games. We classify the possible types of simple games, pick interesting examples of each type, and summarize convergence behavior under various information and player matching protocols. See Friedman (1996), Cheung and Friedman (1997) and Bouchez (1997) for more complete descriptions of the experiments.

We begin with normal form games that have only two alternative strategies and a symmetric population of players. These games are defined by a 2×2 matrix $\mathbf{A} = ((a_{ij}))$ specifying the payoff to any player choosing strategy i when the opponent chooses j. Evolutionary game theory predicts that the direction of change in the fraction $p \in (0, 1)$ of players choosing the first strategy is given by the sign of the payoff differential $D(p) = (1, -1)A(p, 1-p)' = (1-p)a - pb$, where $a = a_{12} - a_{22}$ and $b = a_{21} - a_{11}$. When $D(p)$ is positive (the first strategy has the higher payoff) then $p < 1$ increases and the fraction $1 - p$ of players choosing the alternative strategy 2 decreases; the opposite is true when $D(p)$ is negative. The graph of $D(p)$ is a straight line with intercept a at $p = 0$ and value $-b$ at $p = 1$. Thus (apart from the degenerate case $a = b = 0$ in which a player is always indifferent between her two actions) each payoff matrix falls into one of three qualitatively different types as shown in Figure 1. In accordance with this classification scheme, we used single and two population games of all three types.

The next most complicated case is a single population of strategically identical players with three alternative actions. Here the payoff matrix $\mathbf{A}$ is 3×3 and the current state is a point in two-dimensional simplex $S = \{(p, q, 1-p-q) \in \mathbf{R}^3\colon\ p, q \geqslant 0,\ p+q \leqslant 1\}$. The classification of matrices becomes more complex as the edges of the simplex retain all three possibilities and the interior can be a sink, source, saddle or center. We use only a version of the "Hawk–Dove–Bourgeois" (HDB) game (see Figure 2 for an illustration).

1. Laboratory Procedures and Treatments

The experiments consist of 60–120 minutes laboratory sessions with 6 to 24 undergraduate subjects. Population size varies from 8 to 16 in the results presented here; Friedman (1996) finds strategic behavior contrary to the evolutionary assumption with population sizes of 6 and smaller. After instruction and a few practice periods, each session consists of 60–200 periods broken up into runs of 10 to 16 periods. Over 90% of the subjects earned between \$8 and \$32 per session.

Handbook of Experimental Economics Results, Volume 1

DOI: 10.1016/S1574-0722(07)00053-4

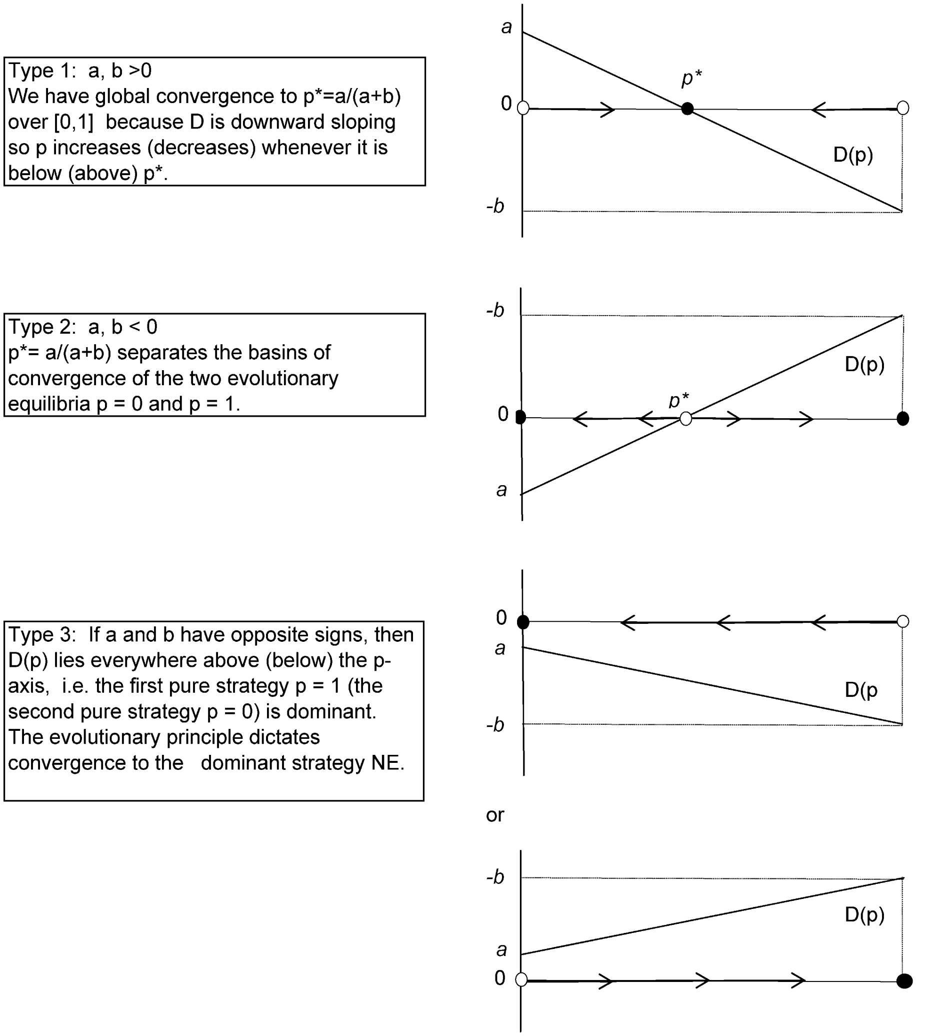


Figure 1. The classification of linear one-dimensional evolutionary games. A 2×2 matrix $A = (a_{ij})$ specifies the payoff to any player choosing strategy i when the opponent chooses j. The direction of change in the fraction p of players choosing the first strategy is the sign of the payoff differential $D(p) = (1, -1)A(p, 1 - p)' = (1 - p)a - pb$, where $a = a_{12} - a_{22}$ and $b = a_{21} - a_{11}$. The slope of $D(p)$ and location of the root $p^* = a/(a + b)$ of $D(p) = 0$ determine the type of the matrix A.

The treatments used were *random pairwise* (RP) and *mean matching* (MM) matching protocols, and the amount of historical information that appears on each player’s screen (Hist/No Hist).

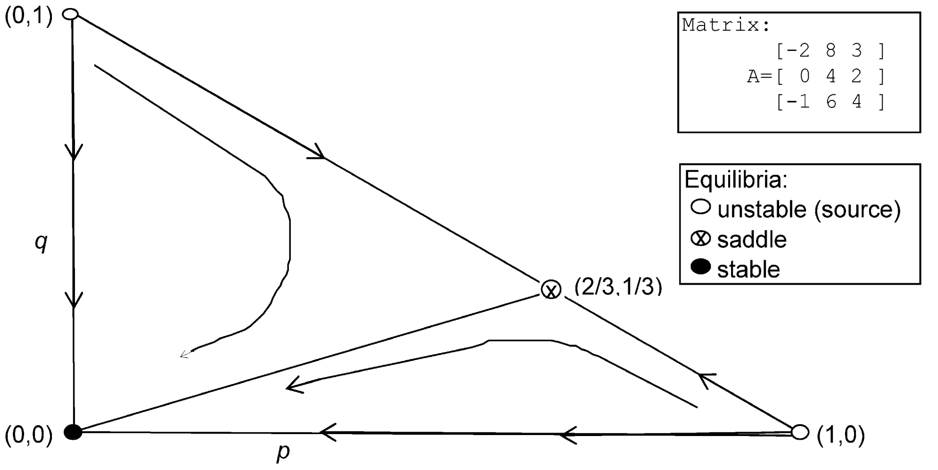


Figure 2. The Hawk–Dove–Bourgeois game. This version of the Hawk–Dome–Bourgeois game has a corner NE at $(p, q) = (0, 0)$ and an edge NE at $(2/3, 1/3)$. Under standard dynamics (e.g., replicator dynamics; see Weibull, 1995, for a simple exposition), the corner NE is an evolutionary equilibrium and the NE is a saddle point. The equations $d\mathbf{p}/dt = (1, 0, 0)'\mathbf{As} - (1/3, 1/3, 1/3)'\mathbf{As}$ and $d\mathbf{q}/dt = (0, 1, 0)'\mathbf{As} - (1/3, 1/3, 1/3)'\mathbf{As}$ characterize the dynamics in the interior of the simplex.

Under *random pairwise* (RP) the player had a single opponent randomly chosen each period. Players view their own (but not the opponent's) payoff matrix, and type "a" or "b" at the keyboard to indicate the choice of the first or second strategy. Under RP matching for payoff matrix $\mathbf{A}$, a player's choice of strategy $i = 1$ or 2, gives *expected* payoff $(2 - i, i - 1)\mathbf{A}(p, 1 - p)'$ when the fraction of potential opponents choosing strategy 1 is p. However, his *actual* payoff depends on the action taken by his actual opponent, and so has some variance around its expectation. The variance is eliminated in the alternative matching procedure, called *mean matching* (MM). Here each player is matched once against each possible opponent in each round and receives the average (mean) payoff over all his matches.

The other major treatment in our experiments is the amount of historical information that appears in the upper left box on each player's screen. In the minimum level, *No Hist*, the player receives no historical information other than what she could tabulate herself: her own action and actual payoff in previous periods. In the other level, *Hist*, the box also displays the previous periods' full distribution of choices in the opponent population.

All treatments are held constant within a run to test for convergence. Runs are separated by obvious changes in the player population and/or the payoff matrix, and the history box is erased at the beginning of a new run.

2. Results

We have collected more than 300 such runs and used various statistical tests as well as summary graphs to study convergence properties. Figures 3–5 show some of the summary graphs for both one, two population, and three choice games. The main findings, presented more fully in Friedman (1996), can be summarized as follows.

(1) Some behavioral equilibrium (BE) is typically achieved by the second half of a 10 to 16 period run. The operational definition of BE is that strategy selection is almost constant in each population in a given run (or half-run). "Almost constant" means that the mean absolute deviation from the median number of players choosing a given strategy is less than 1 player ("tight") or less than 2 players ("loose"). Overall we observe tight BE in over 70% of second half-runs and loose BE in over 98% of second half-runs. Tight BE was achieved most reliably in type 3 games (over 95% of all half-runs). By contrast, type 1 games achieved tight BE in only 55% of half runs, but achieved loose BE in 96%.

(2) BE typically coincides with a Nash equilibrium (NE), especially with those (called evolutionary equilibria (EE)) that evolutionary game theory identifies as stable. We operationalize NE by replacing the median number of players by the NE number. In second half-runs, for example, about 79% of the loose BE are loose NE, and 84% of those are loose EE. The only notable exception to this conclusion is that in type 2 runs the BE sometimes coincided with the non-EE mixed NE. A closer look at the graphs suggests that many of these cases actually represent slow divergence from the mixed NE, and many of the half-runs deemed BE but not NE seem to represent slow or incomplete convergence to an EE (a pure NE).

(3) Convergence to BE is faster in the mean-matching (MM) than in the random-pairwise (RP) treatment, and faster in the *Hist* treatment than in the *No Hist* treatment. In particular, the slow and incomplete convergence observed in type-1 games arises mainly in RP matching protocol and *No Hist* runs. The results from type-2, single population games and all two-population games support the same conclusion. There is, however, an interesting exception. The few instances of non-convergence in type-3 games arise more often under MM than under RP.

(4) Individual behavior at a mixed strategy BE is better explained by idiosyncratic "purification" strategies than by identical individual mixed strategies. In particular, in the simplest type 1 game, Hawk–Dove, we see persistent heterogeneity in which some players consistently pick the first ("Hawk") strategy and others consistently pick the other ("Dove") strategy.

(5) "Hawk–Dove–Bourgeois" is a 1-population 3-action game with a triangular state space and with one corner NE (an EE) with target area b^2 and one edge NE (not an EE) with target area $2b^2$. Only one session was explored in Friedman (1996). Additional data has been collected (Bouchez, 1997) and the results discussed here are for the combined data. Loose (tight) convergence was found to some BE in 41 (7) of 46 half-runs, loose (tight) convergence to the EE in 8 (3) half-runs,

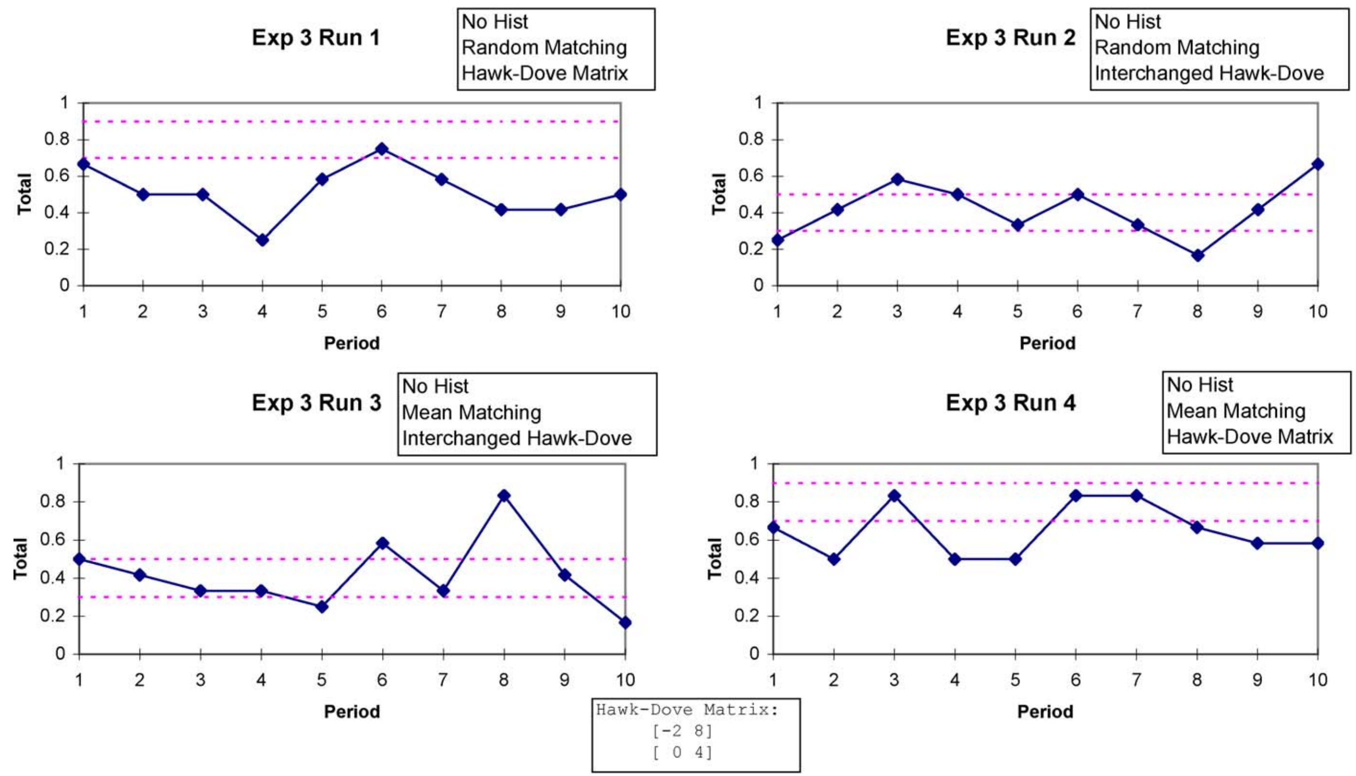


Figure 3. Single population sessions: Exp3 runs 1–4. These graphs chart the time path of pt in the first four runs of the first usable session. The type 1 payoff matrix here has unique mixed NE $p^* = 2/3 = 8/12$. That is, in NE 8 of 12 players choose the first strategy (or 4 of 12 when the matrix rows and columns are interchanged as in runs 2 and 3). The graphs show a tolerance of 1 player in the band around NE. The time paths in the first four runs suggest that the NE attracts states pt outside the tolerance band $p^* \pm 1/12$, but there is considerable behavioral noise so hits occur in only about 50% of the periods.

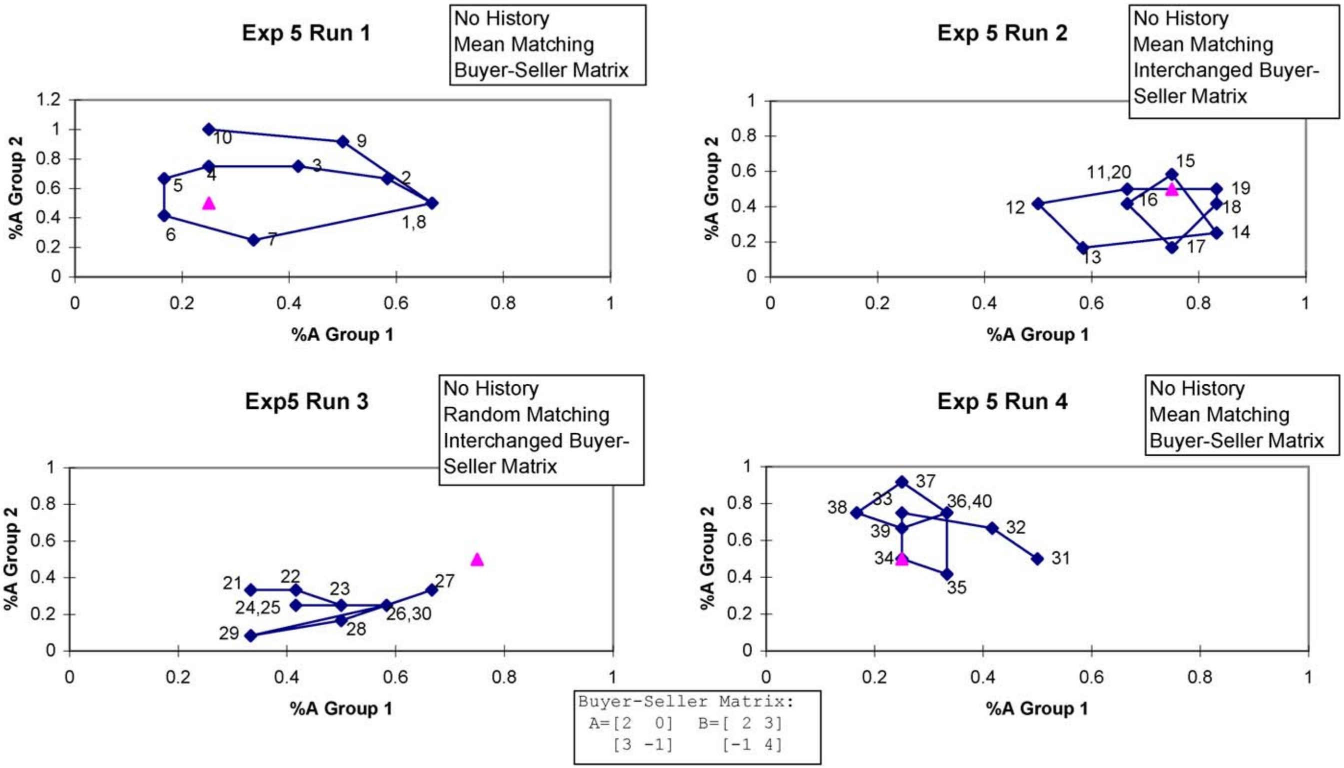


Figure 4. Two population sessions: Exp5 runs 1–4. Graphs of the behavior in the first four runs of exp5, the first 2-population session. All periods use the buyer–seller matrix or its interchange, so the unique NE (denoted by ▲) is at $(p, q) = (.25, .50)$ or, for the interchanged version, at $(.75, .50)$. The graphs show a 2-period moving average of the time path in the unit square. The graph for the first run looks like an unstable counterclockwise spiral diverging from the NE. The second run looks like a tidy counterclockwise double loop around the NE, neither converging nor diverging. The third run uses the RP matching protocol; at best there is a weak tendency to drift towards the NE. The fourth run reverts to MM and looks like a counterclockwise spiral possibly converging to the NE.

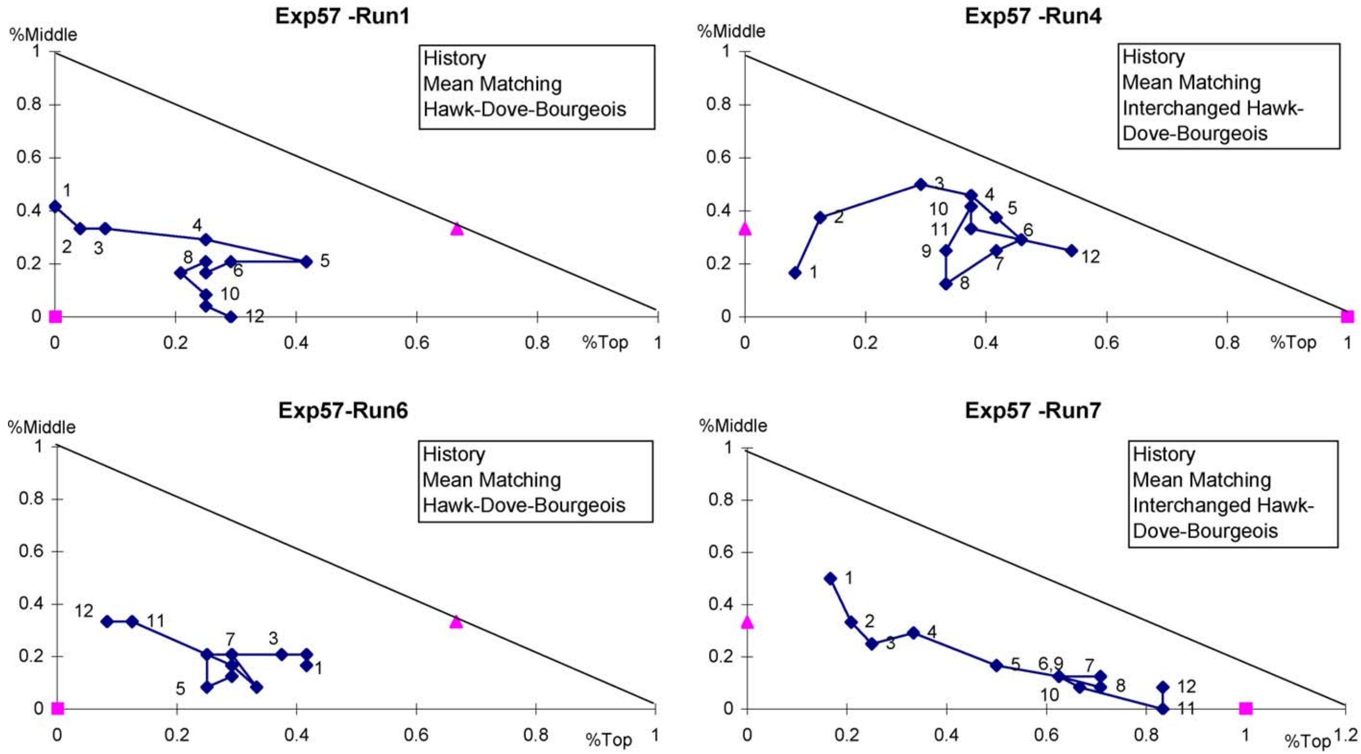


Figure 5. Single population three choice sessions: Exp57 runs 1, 4, 6, 7. Graphs of the behavior in the first four HDB runs of exp57. The two NE are at $(p, q) = (0, 0)$ and $(2/3, 1/3)$ (for the interchanged version, at $(1, 0)$ and $(0, 1/3)$) and are represented by ■ and ▲, respectively. The graphs show a 2-period moving average. The run 1 has loose convergence in the second half of the run as does run 7. Run 4 shows no convergence in either half. Run 6 has loose convergence in both the first and second halves of the experiment.

and no loose or tight convergence was found to the edge NE despite its larger area. The data are sparse but consistent with evolutionary game theory.

3. Discussion

For all three types of one-dimensional games and their two-dimensional analogues, the states reliably achieve a loose behavioral equilibrium (BE) even within the first half-run of 5 periods. Most of the loose BE are also tight BE, the main exceptions occurring in two-dimensional games with unique Nash equilibria (NE). Most BE coincide with NE, and most of the observed NE are indeed evolutionary equilibria (EE). In general, the "evolutionary" treatments of mean-matching (MM) and feedback (Hist) appear to improve convergence to EE. Thus the main tendencies of the data are consistent with evolutionary game theory.

The exceptions or boundaries to these main tendencies may be of special interest. Friedman (1996) shows that when group size is smaller than 6, players much more often appear willing to sacrifice current personal payoffs to increase group payoffs (and perhaps own later payoffs). Cooperative behavior (foregoing the dominant strategy) is sometimes observed in type 3 prisoner's dilemma sessions which have runs splitting the players into groups of size 2 or 4, and it is especially prevalent in the runs with the smaller groups. Such behavior is notably less frequent in sessions where the minimum group size remains above 6.

Perhaps the most surprising finding concerns another boundary for evolutionary game theory. An influential branch of the theory (Kandori, Mailath, and Rob, 1993, and Young, 1993) argues that in simple coordination (type 2) games with two pure strategy (corner) NE = EE and one interior NE, the "risk-dominant" corner EE is most likely to be observed because it has the larger basin of attraction, and indeed that *only* the risk-dominant EE will be observed in the relevant limiting case. Friedman (1996) shows that the data reviewed in this chapter support the contrary theoretical view of Bergin and Lipman (1995) that one can bias convergence towards the other ("payoff-dominant") EE by increasing the potential gains to cooperation. In some applications evolutionary game theory may have to be supplemented by a theory of trembles (or "mutations") that allows for forward-looking attempts to influence others' behavior.

References

Bergin, J., Lipman, B.L. (1995). "Evolution with state-dependent mutations". Draft manuscript, Queen's University Department of Economics.

Bouchez, Nicole (1997). "A study and application of learning in population games". Draft manuscript, University of California, Santa Cruz.

Cheung, Yin-Wong, Friedman, Daniel (1997). "Individual learning in normal form games: Some laboratory results". Games and Economic Behavior 19, 46–76.

Friedman, Daniel (1996). "Equilibrium in evolutionary games: Some experimental results". The Economic Journal 106 (434), 1–25.

Kandori, Michihiro, Mailath, George, Rob, Rafael (1993). "Learning, mutations, and long run equilibria in games". Econometrica 61 (1), 29–56.
Young, H. Peyton (1993). "The evolution of conventions". Econometrica 61 (1), 57–84.
Weibull, Jorgen W. (1995). "Evolutionary Game Theory". MIT Press, Cambridge, MA.

Chapter 33

A COMPARISON OF MARKET INSTITUTIONS

TIMOTHY N. CASON
Purdue University

DANIEL FRIEDMAN
University of California, Santa Cruz

This chapter summarizes a laboratory experiment comparing four different trading institutions, illustrated in Figure 1, in a random values environment with four single unit buyers and four single unit sellers. Figure 2 summarizes the main findings on efficiency: the continuous double auction (CDA) and high-frequency multiple-call market (MCM) are the most efficient institutions, followed by the low-frequency MCM and the single call market (SCM), with our implementation of uniform price double auction (UPDA) trailing the pack. Figures 3 and 4 summarize the main findings on transaction prices and volume. Price deviations from competitive equilibrium tend to be smallest in the SCM. Volume is highest in CDA, decreases with call frequency in MCM and SCM, and is lowest in UPDA. The body of this chapter describes the institutions, environment, performance measures, and results.

1. Market Institutions

The CDA is the richest trading institution in terms of within-period information feedback, trading opportunities and strategic complexity. The CDA sessions were conducted using multiple-unit double auction (MUDA) trading software (Plott, 1991), constrained to a single market and a single unit per trader. Every trader's screen displays the current market bid and ask. Buyers (sellers) are free to accept the market ask (bid) at any time, and the transaction is executed immediately. A transaction immediately removes both the market bid and ask. At any time during a trading period, buyers (sellers) who have not yet transacted are free to seize the market bid (ask) by posting a bid exceeding (ask below) the current market bid (ask). Traders perform their record-keeping by hand. Each period consists of 110 seconds of trading, which was sufficient for the typical 2- to 3-unit trading volume in our environment.

The SCM is the simplest trading institution in that it offers only one trading opportunity and minimal information within each trading period. The SCM sessions were conducted using a variant of the software employed in Friedman (1993) and several other studies. The SCM institution solicits a bid (or highest acceptable purchase price for a single unit) b_i from each buyer i and an ask (or lowest acceptable sale price) a_j

Handbook of Experimental Economics Results, Volume 1

DOI: 10.1016/S1574-0722(07)00033-9

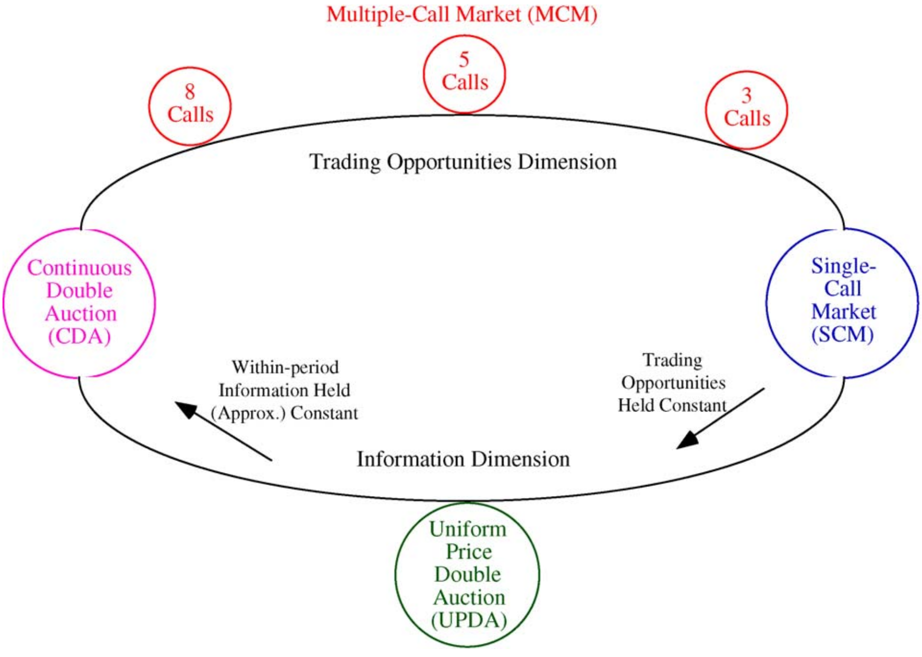


Figure 1. An overview of the trading institution comparison and experimental design.

from each seller j. The solicitations are simultaneous and private during the trading period. The demand revealed in $\{b_i\}$ and the supply revealed in $\{a_j\}$ then are cleared at a uniform equilibrium price p^*. With indivisible units, there often is an interval $[p_l, p_u]$ of market clearing prices, in which case the chosen price is $(1-k)p_l+kp_u$, where $k = 0$, $k = 0.5$ or $k = 1$. At the conclusion of the trading period traders observed p^* and the bids, asks, values, costs and profits of all other traders. Detailed analysis presented in Cason and Friedman (1997) indicates that trader behavior is generally insensitive to the pricing rule k, so for this institutional comparison we pool the data across the three k treatments.

As illustrated in Figure 1, the remaining two trading institutions link these polar institutions in two different ways. The MCM is just like the $k = 0.5$ SCM except that it has several calls (clearings) per period so traders have more trading opportunities. The number of calls is preannounced to be 3, 5, or 8 each period, with the length of the trading period is fixed at 120 seconds. Unaccepted bids and asks are automatically renewed for the next call, although traders can revise them at any time as in the CDA. After each call, all traders see whether their own offer was successful, and observe the transaction price (if any) as well as the best-rejected bid and best-rejected ask. At the end of the period subjects receive the same "full information" as in the SCM.

The UPDA market provides continuous information feedback as in the CDA, while limiting the number of trading opportunities to one per period as in the SCM. In the

UPDA, one call is held at the end of the trading period, but during the period traders submit and revise market bids and offers while observing changes in the possible terms of trade. A variety of information conditions are possible in the UPDA, many of which were explored in different environments from ours by McCabe, Rassenti, and Smith (1993) and by Friedman (1993). The choices made for this experiment reflect a desire to approximate the information conditions and environment of the CDA. Therefore, we provided subjects with the current "indicated market price" and the current best rejected bid and best rejected offer, which corresponds most closely to the CDA information concerning the current market bid and offer and available terms of trade. Moreover, the sessions allow traders to cancel bids and offers (which is allowed by many computerized CDA implementations including ours) and calls the market after 90 seconds. The UPDA sessions were conducted using the same market software (again appropriately modified) that was used in the SCM and MCM sessions. Once again, traders received complete information regarding the other traders values, costs, bids asks and profits at the conclusion of each period, and all sessions used the pricing rule $k = 0.5$.

2. Market Environment

To maintain comparability across institutions, we held constant several features of the environment. All sessions take place in a random values environment: in each trading period the buyers' redemption values and the sellers' costs are independently drawn from the uniform distribution with range [$0.00, $4.99]. This fact is publicly announced at the beginning of the session, and subjects have no other information regarding other subjects' drawn values during the trading period. The same sequences of drawn values were used in each session and across institutions to limit between-session variability. When the same subjects are brought back as experienced we employed a different set of random values, and these values were held constant across all experienced sessions.

All sessions used four buyers and four sellers, each with a trading capacity of only one unit. All inexperienced sessions ran 30 trading periods, and all experienced sessions ran 40 trading periods. In the 30-period inexperienced sessions, traders switched buyer and seller roles before period 9 and before period 25; in the 40-period experienced sessions, traders switched roles before period 11 and before period 31. This switch was preannounced, as was the number of buyers and sellers in each session. For each institution except SCM we conducted 3 sessions with inexperienced subjects and 2 sessions with experienced subjects. For the SCM we used three sessions (two employing inexperienced subjects) in each of the three k treatments.

3. Related Work

Several laboratory experiments have also compared these trading institutions in alternative environments. Smith et al. (1982) compares performance of the CDA to several variants of the SCM in a repetitive stationary environment. Price formation was more rapid and reliable in the CDA but a multiple-unit recontracting version of the SCM had equivalent allocational efficiency. Friedman and Ostroy (1995) find that both the CDA and the SCM eventually produce highly efficient allocations even when the induced values and costs are chosen to encourage strategic misrepresentation. McCabe, Rassenti, and Smith (1993) study UPDA in an environment with additive random shifts superimposed each period on otherwise repetitively stationary demand and supply schedules. Friedman (1993) examines another variant of the UPDA institution as well as the MCM institution in an asset market environment. These studies find that the best variants of UPDA and MCM are almost as reliable as the CDA in producing prices and allocations near competitive equilibrium. The interested reader should consult our previous work (Cason and Friedman, 1996, 1997) for more details of additional experimental treatments in the CDA and SCM environments, respectively.

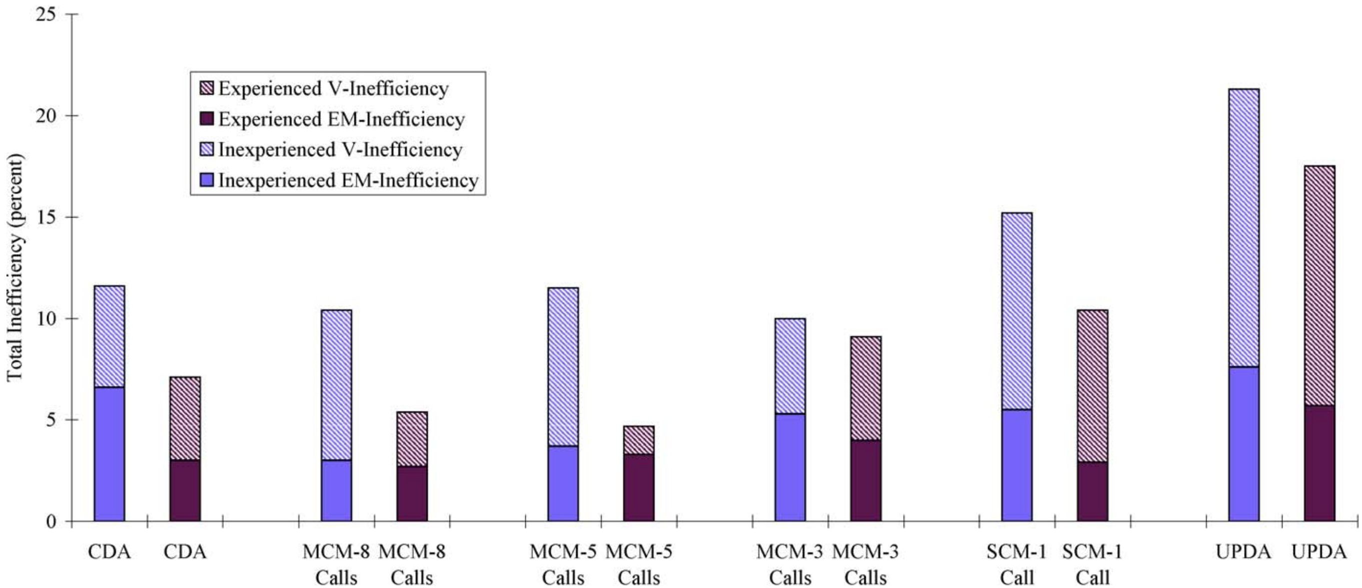
Two caveats are in order. First, we chose a thin, random-values environment because it provided the clearest view of the price formation process. It is not necessarily the most representative of important field environments, and the environmental robustness of our conclusions is not yet established. Second, several of the institutions have variants that may have different performance characteristics. In particular, McCabe, Rassenti, and Smith (1993) find that UPDA efficiency depends fairly sensitively on implementation details, and is not enhanced by two details we chose to maintain comparability to the CDA, viz., a fixed closing time and a two-side update rule.

4. Results

4.1. Market Efficiency

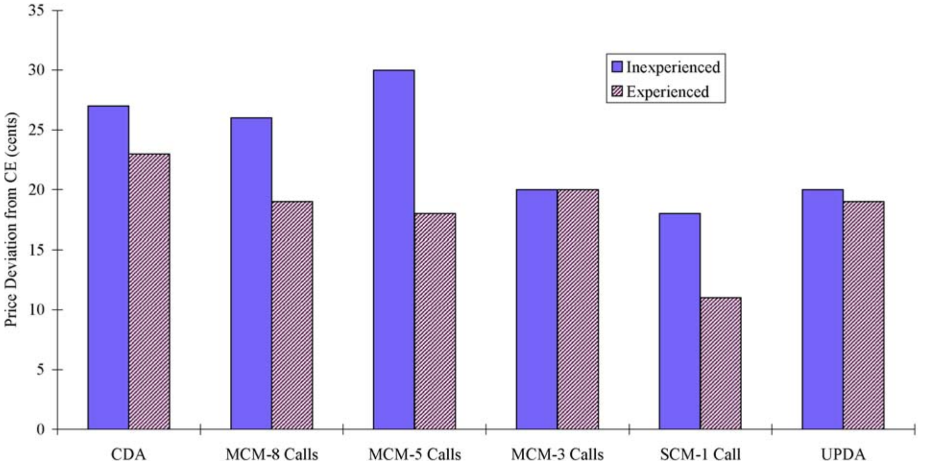
Define market *inefficiency* as the percentage of the maximum possible gains from exchange which traders fail to realize. Figure 2 summarizes the outcomes. Our most surprising finding is that UPDA is the most inefficient trading institution in this thin volume, random-values environment. In statistical tests based on a random effects model reported in Cason and Friedman (1999), we show that the UPDA inefficiency is significantly greater than the CDA inefficiency and is marginally significantly greater than the SCM inefficiency. UPDA inefficiency is also significantly greater than the inefficiency in all three MCM treatments.

Like UPDA, the SCM has only one transaction opportunity per period and has lower efficiency than the other institutions in both experience conditions. This provides some evidence that multiple trading opportunities are important to generate increases in efficiency. However, the differences between SCM efficiency and CDA and MCM efficiency are not statistically significant. Finally, note that the MCM institution generates efficient outcomes that compare favorably with (and are not statistically distinguishable from) the CDA outcomes, so it would appear that 3 to 5 calls per period are sufficient to generate market efficiency comparable to the CDA benchmark.



Note: CDA denotes continuous double auction; MCM denotes multiple call market; SCM denotes single call market; UPDA denotes uniform price double auction. Inefficiency falls with experience in all trading institutions. V-inefficiency arises when volume falls below the competitive equilibrium volume, and EM-inefficiency arises when extra-marginal units displace.

Figure 2. Trading inefficiency is greatest for the two institutions that permit only one transaction opportunity per period (UPDA and SCM), and inefficiency is lowest for the two institutions that permit multiple transaction opportunities per period (CDA and MCM).



Note: CDA denotes continuous double auction; MCM denotes multiple call market; SCM denotes single call market; UPDA denotes uniform price double auction.

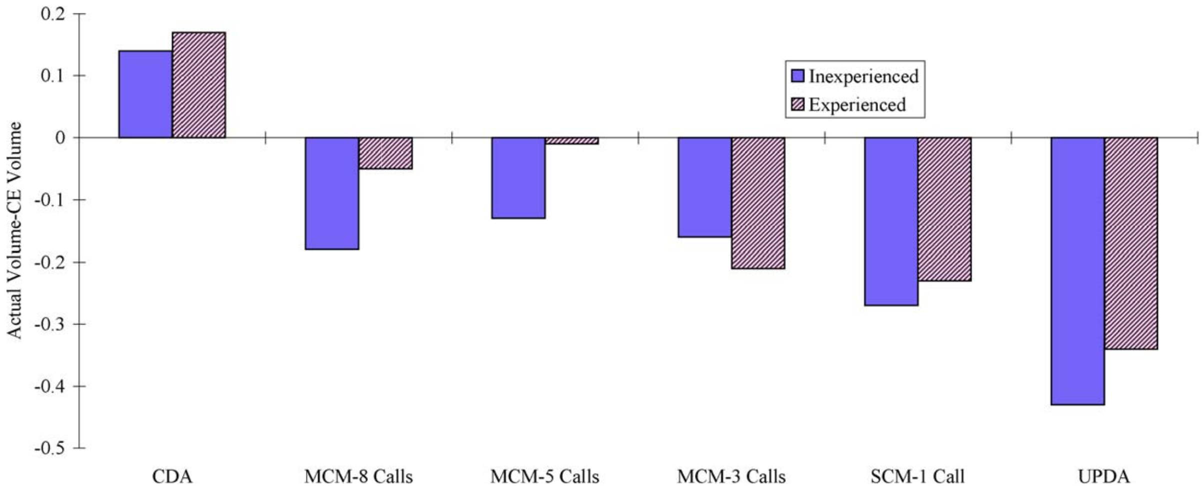
Figure 3. Mean absolute deviations from the competitive equilibrium are lowest in the SCM, in which all trades occur at one price; deviations tend to be greatest in the institutions that permit multiple transaction opportunities per period (CDA and MCM).

Efficiency can fall short of 100 percent if (a) traders with extra-marginal units transact (EM-inefficiency), or (b) profitable trades are not executed (low volume or V-inefficiency). The bars of Figure 2 distinguish the mix of V and EM inefficiency across institutions (see Cason and Friedman (1999) and Rust, Palmer, and Miller (1993) for details of this inefficiency decomposition). Both types of inefficiency are common, and in the inexperienced sessions the low volume efficiency losses exceed the extra-marginal efficiency losses in four of the six institutions. In all institutions the increase in efficiency due to experience generally occurs because of reductions in both types of inefficiency. In the 5-call and 8-call MCM, the reduction in V-inefficiency is quite pronounced. V-inefficiency is lowest for the CDA and MCM sessions, probably due to the multiple transaction opportunities permitted by these institutions.

4.2. Transaction Prices

The standard benchmark for price is the competitive equilibrium (CE), which equates true demand and true supply for the value and cost realization that period. In our thin, random values environment in nearly every period there exists a *range* of CE prices. Mean transaction prices in a period are within the CE range in less than one-half of the periods for all institutions. (The CDA and MCM final transaction prices, not shown here, are within the CE range no more often.)

Figure 3 presents the mean absolute deviation of average transaction prices (each period) from the nearest endpoint of the CE range. Of course, the price deviation is 0 if



Note: CDA denotes continuous double auction; MCM denotes multiple call market; SCM denotes single call market; UPDA denotes uniform price double auction.

Figure 4. Trading volume increases with experience, and exceeds the CE volume in the CDA; volume roughly declines with the number of transaction opportunities in the MCM, and is lowest in the UPDA.

average prices are within the CE range. Price deviations are smallest in the SCM, and are largest (at least for experienced traders) in the CDA. Statistically speaking (again, see Cason and Friedman (1999) for details), the CDA mean price deviation exceeds: (1) the UPDA mean price deviation; (2) the 3-call MCM mean price deviation; and (3) the SCM mean price deviation. The SCM mean price deviation is also lower than the 5-call MCM mean price deviation and the 8-call MCM mean price deviation. The other mean price deviations are not significantly different.

4.3. Transaction Volume

The competitive equilibrium (CE) also provides a benchmark for trading volume. CE volume ranges from 0 to 4 units but usually is 2 or 3. Figure 4 shows that actual transaction volume increases with experience in every institution, with the minor exception of the 3-call MCM. Volume is highest in the CDA, where it exceeds the CE benchmark. Volume falls below the CE benchmark in the other institutions and declines almost monotonically with the number of calls in MCM and SCM. It is lowest in UPDA. Cason and Friedman (1999) report that most of these differences are statistically significant.

5. Discussion

The comparison of four trading institutions in a thin market, random values environment supports the following general conclusions. First, trading efficiency in the uniform price double auction is lowest, and the single call market efficiency is second lowest. This suggests that multiple trading opportunities, as in the continuous double auction and the multiple call market, help generate high efficiency. Second, the primary source of efficiency losses in these (single opportunity) institutions is insufficient trading volume. Third, transaction prices are less accurate on average (in that they deviate more from competitive equilibrium levels) in the continuous double auction and multiple call market. Taken together, these results highlight a key tradeoff when the trading institution permits multiple transaction opportunities. Multiple transaction opportunities substantially reduce (low volume) inefficiency due to underrevelation of traders' true values and costs, but also reduce pricing accuracy because traders allow more noise when negotiating transactions.

References

Cason, T., Friedman, D. (1996). "Price formation in double auction markets". Journal of Economic Dynamics and Control 20, 1307–1337.

Cason, T., Friedman, D. (1997). "Price formation in single call markets". Econometrica 65, 311–345.

Cason, T., Friedman, D. (1999). "Price formation and exchange in thin markets: A laboratory comparison of institutions". In: de Antoni, E., Howitt, P., Leijonhufvud, A. (Eds.), Money, Markets and Method: Essays in Honour of Robert W. Clower. Edward Elgar, Cheltenham, UK, pp. 155–179.

Friedman, D. (1993). "How trading institutions affect financial market performance: Some laboratory evidence". Economic Inquiry 31, 410–435.

Friedman, D., Ostroy, J. (1995). "Competitivity in auction markets: An experimental and theoretical investigation". Economic Journal 105, 22–53.

McCabe, K., Rassenti, S., Smith, V. (1993). "Designing a uniform price double auction: An experimental evaluation". In: Friedman, D., Rust, J. (Eds.), The Double Auction Market. Addison–Wesley, Reading, MA, pp. 307–332.

Plott, C. (1991). "A computerized laboratory market system and research support systems for the multiple unit double auction". Social Science Working Paper 783, California Institute of Technology.

Rust, J., Palmer, R., Miller, J. (1993). "Behavior of trading automata in a computerized double auction market". In: Friedman, D., Rust, J. (Eds.), The Double Auction Market. Addison–Wesley, Reading, MA, pp. 155–198.

Smith, V., Williams, A., Bratton, W.K., Vannoni, M. (1982). "Competitive market institutions: Double auctions vs. sealed bid–offer auctions". American Economic Review 72, 58–77.

Chapter 13

THE MATCHING MARKET INSTITUTION

DANIEL FRIEDMAN and CHANGHUA RICH
University of California, Santa Cruz

The matching market (MM) institution is a two-sided auction procedure that collects bids from buyers and asks from sellers, and iteratively matches the highest remaining bid with the highest remaining ask less than or equal to it. Each buyer pays his bid price and the seller receives the bid price. Rich and Friedman (1998) and Rich (1996) document recent use of the MM by the Chinese Environmental Protection Agency and occasional use in various financial markets.[1]

The MM institution maximizes the number of agreeable matches (i.e., transactions) in a given set of bids and asks, and gives the seller the best agreeable price in each transaction. However, buyers and sellers in the MM have strong incentives to bid and ask strategically so as to underreveal willingness to transact (Rich and Friedman, 1998). Actual performance, in terms of efficiency as well as transaction volume and seller's surplus, therefore is a question of theoretical and practical interest.

In this chapter we summarize a first laboratory experiment comparing the MM institution to the natural alternative, the uniform price (UP) or single call market. The procedures and results are reported more fully in Rich and Friedman (1998). We find that, compared with UP, MM has lower efficiency, has about the same average volume but greater variability, and gives sellers a *smaller* fraction of the surplus.

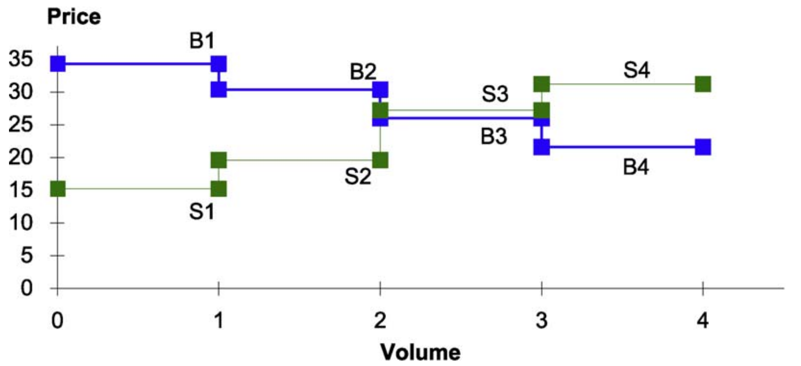
1. Experimental Procedures

The experiment consisted of ten sessions; five conducted by hand at Wuhan University (WU) in China and five conducted over a computer network at University of California, Santa Cruz (UCSC). Each session involved 16 or more trading periods, roughly half under MM and half under UP. Three sessions at each site used inexperienced subjects, four buyers and four sellers, and the other two sessions used experienced subjects, six buyers and four sellers. Figure 1 shows the induced value and cost parameters; note that experienced buyers received the higher value schedule B' in some periods. On average subjects earned about \$10 per hour at UCSC and 25 Yuan at WU, well in excess of average opportunity cost.

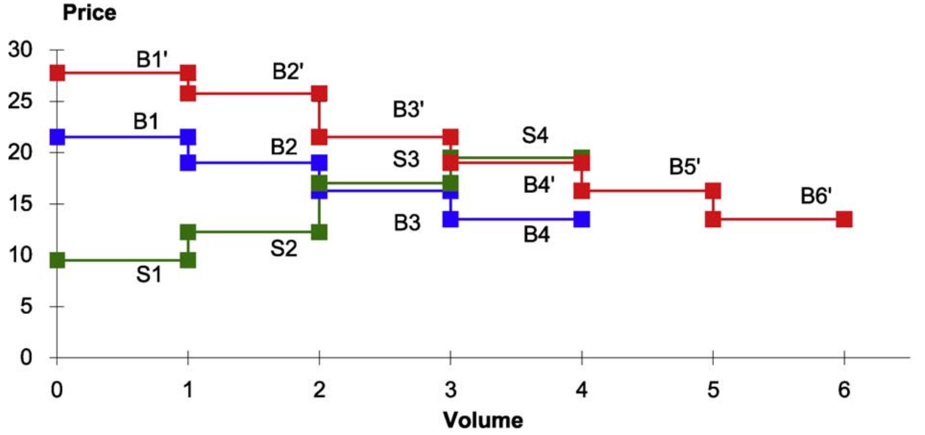
[1] A general definition is that a matching market is a mechanism for pairing individuals for mutual benefit. In one well-known special case, often referred to as the marriage market, utility is non-transferable and the value of a potential match depends nonlinearly on the characteristics of the two parties (Gale, 1968; Becker, 1973; Roth and Sotomayor, 1990; Shimer and Lones, 2000). Our MM is a polar case of transferable utility in which the value of a match is simply the difference between buyer's value and seller's cost.

Handbook of Experimental Economics Results, Volume 1

DOI: 10.1016/S1574-0722(07)00013-3



(A)



(B)

Figure 1. (A) Induced values and costs in inexperienced sessions. In sessions with inexperienced subjects, supply is induced each period using cost schedule $S1$–$S4$. Demand is induced each period using value schedule $B1$–$B4$, producing the competitive equilibrium (CE) price = 27.2, the CE volume = 2, and the total CE profit = 30 each period. (B) Induced values and costs in experienced sessions. In sessions with experienced subjects, supply is induced each period using cost schedule $S1$–$S4$. Demand is induced in some periods using value schedule $B1$–$B4$, producing CE price = 17.0, CE volume = 2, and CE total profit = 18.75. In other periods demand is induced using $B1'$–$B6'$, producing CE price = 19.5, CE volume = 3, and CE total profit = 36.25.

To illustrate the two market institutions, suppose that the bids of the four buyers are 4.60, 3.75, 2.25, and 1.50, and that sellers' asks are 6.00, 4.00, 2.81, and 1.40. In the MM, the highest ask would be rejected since it is above the highest bid, and the other three asks would be matched respectively with the three highest bids. The outcome is three transactions at prices 4.60, 3.75, and 2.25 with revealed surplus 4.60 − 4.00 = 0.60, 3.75 − 2.81 = 0.94, and 2.25 − 1.40 = 0.85. The actual surplus, of course, depends on the three buyers' true values and the three sellers' true costs; presumably the actual surplus exceeds the total revealed surplus of $0.60 + 0.94 + 0.85 = 2.39$.

The same bids and asks lead to a different outcome in the UP institution. Here the two highest bids are matched with the two lowest asks at the highest market clearing price, 3.75. In this case the two transactions yield a revealed surplus of $4.60 - 1.40 = 3.20$ plus $4.00 - 3.75 = 0.25$ for a total of 3.45. See the Cason and Friedman chapter in this Handbook and the citations therein for a complete description of the UP institution, referred to there as the "$k = 1$ SCM."

In each period of the experiment, the trading institution, MM or UP, was announced (and typically held constant for 4–8 periods) and each trader privately submitted a bid or ask. The collected bids and asks and resulting transactions then were publicly displayed; true values and costs remained private information.

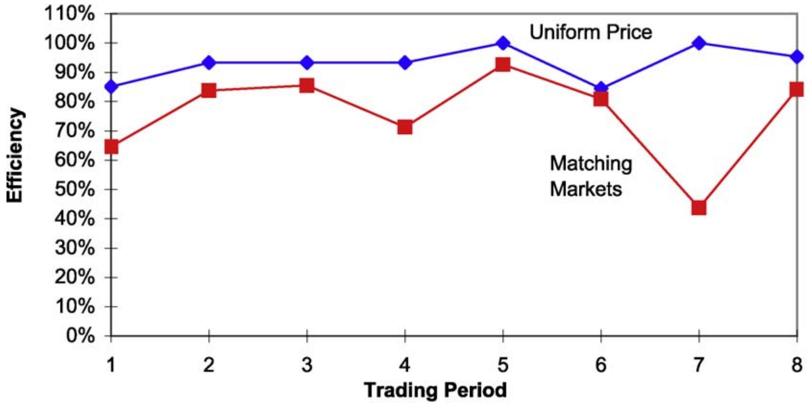
2. Results

We focus here on how the market institution (MM or UP) affects three performance variables: efficiency, trading volume, and surplus split. The variables are measured in 264 market periods in ten sessions, 133 MM periods and 131 UP periods. Efficiency is defined each period as the observed total profits as a fraction of the competitive equilibrium (CE) profit shown in Figure 1. The overall mean efficiency is 80 percent under the matching market compared to 92 percent under the uniform price market. Figure 2 shows the trends for the first eight periods averaged over all six inexperienced sessions in Panel A. In each of these periods the average efficiency is higher in UP than in MM. Panel B shows the corresponding averages over all four experienced sessions. Here the data are more erratic and in three periods the MM has higher efficiency, but overall UP is still clearly the more efficient market institution.

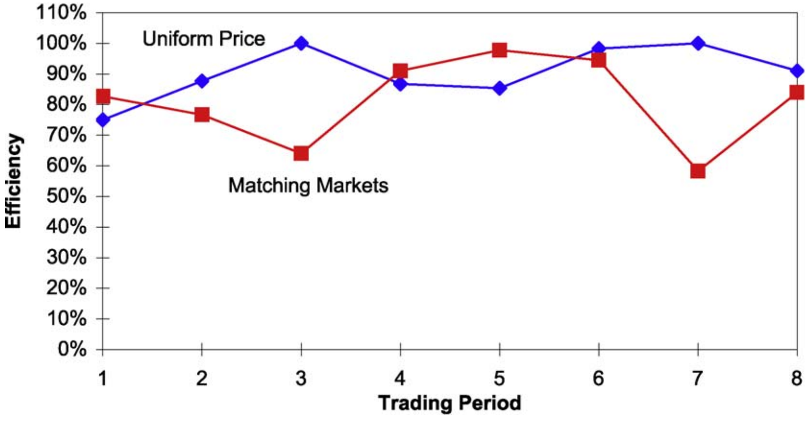
Volume deviation is defined in percentage terms each period as observed trading volume minus competitive equilibrium volume and then divided by CE volume. These values are quite small on average in both institutions: −0.6% in MM and −5% in UP. But volume is considerably more variable in MM, where the absolute value of volume deviation (VDVAB) averages 21% in MM and 5% in UP. That is, positive and negative deviations are both large in the MM but tend to offset each other, while positive deviations are extremely rare in UP and negative deviations are rather moderate. Figure 3 shows the trends in VDVAB. In each of the first eight periods the average absolute volume deviation in the six inexperienced sessions is larger in MM than in UP. Again the experienced data are more erratic – MM has smaller average absolute deviation in one period and equal in another. The overall conclusion is the same: the MM institution is less reliable than the UP in delivering the CE trading volume.

Why are the experienced session averages more erratic than the inexperienced?

Partly it is simply the difference in sample size, with four experienced versus six inexperienced sessions. The other part of the explanation is that both institutions perform erratically. In a stationary repetitive environment, traders in the UP institution tend to



(A)

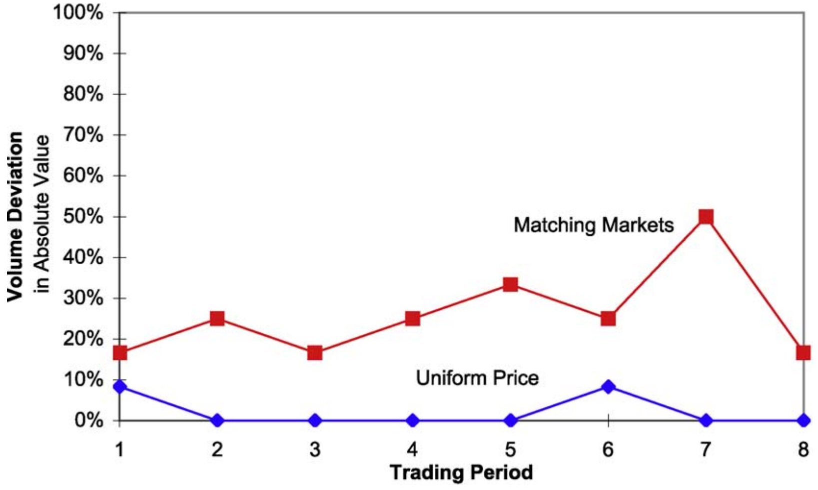


(B)

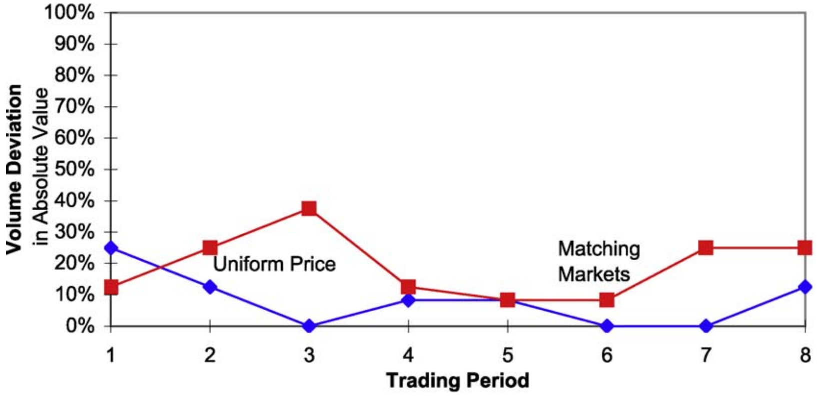
Figure 2. (A) Average efficiency in inexperienced sessions. The first eight periods are averaged over all six inexperienced sessions; see Figure 1 Panel A for induced values and costs. Note that average efficiency in every period is higher in the uniform price institution than in the matching institution. (B) Average efficiency in experienced sessions. The first eight periods are averaged over all four experienced sessions; see Figure 1 Panel B for induced values and costs. Note that overall UP is more efficient although the MM has slightly higher efficiency in three periods.

shade their bids and offers closer and closer to the anticipated clearing price, and occasionally they overshoot the price and fail to transact, with a drastic impact on efficiency and volume (Friedman and Ostroy, 1995; Smith et al., 1982). The UP data in Figures 2 and 3 reflect this strategic behavior, and also reflect the fact that the MM also encourages strategic behavior and produces even more erratic outcomes.

The final performance measure we consider is the surplus split. Define ratio of sellers' profit (RSP) as actual profits earned by all sellers in a given market period as a percent-



(A)



(B)

Figure 3. (A) Average volume deviation in inexperienced sessions. The first eight periods are averaged over all six inexperienced sessions; see Figure 1 Panel A for induced values and costs. Note that volume is considerably more variable in MM compared to UP. (B) Average volume deviation in experienced sessions. The first eight periods are averaged over all four experienced sessions; see Figure 1 Panel B for induced values and costs. Note that the MM is again less reliable than the UP in delivering the CE volume.

age of profits that sellers would earn in competitive equilibrium. Likewise, define RBP as the ratio of buyers' actual profit to CE profit. Overall, average RBP is 91.2% in MM versus 99.9% in UP, while average RSP is 72.9% in MM versus 86.6% in UP. Thus the profits of both buyers and sellers are notably lower in MM than in UP. To assess the relative loss (or surplus split *per se*) we consider the ratio of average RSP to average RBP.

The ratio is 79.9% in MM and 86.7% in UP. Thus the MM impairs sellers' relative, as well as absolute, profitability.

3. Discussion

Our findings offer little encouragement to advocates of the MM in field markets. Compared either to the competitive equilibrium (CE) theoretical benchmark or to the actual performance of the uniform price (UP) market institution, the MM institution fails to deliver on its main selling point: that it will generate higher trading volume. It does generate considerably more variable volume but average trading volume is quite close to the CE benchmark and to the UP average. A secondary selling point is that the MM will offer sellers a larger share of the surplus. However, our data show that sellers in the MM get a smaller slice of a smaller pie than in UP.

Rich and Friedman (1998) show that prices were also more variable in MM, even after averaging across transactions within a period. The main problem with the MM in practice is that it is much less efficient; loss of potential gains from trade averaged about 8% in UP and about 20% in MM. Rich and Friedman (1998) argue that the reason for the poor showing is that buyers substantially understate their willingness to pay and sellers understate their willingness to accept in both institutions but especially in the MM institution. These results serve as a caution to those who contemplate field use of the MM, and serve as a challenge to theorists who wish to construct general models of price formation.

References

Becker, Gary (1973). "A theory of marriage, Part I". Journal of Political Economy 81 (4), 813–846.

Friedman, Daniel, Ostroy, Joseph (1995). "Competitivity in auction markets: An experimental and theoretical investigation". The Economic Journal 105 (428), 22–53.

Gale, David (1968). "Optimal assignments in an ordered set: An application of matroid theory". Journal of Combinatorial Theory 4, 176–180.

Rich, Changhua Sun (1996). "An experimental investigation of two systems of emission permits trading". Dissertation chapter, University of California, Santa Cruz.

Rich, Changhua Sun, Friedman, Daniel (1998). "The matching market institution: A laboratory investigation". American Economic Review 88 (5), 1311–1322.

Roth, Alvin, Sotomayor, Marilda (1990). "Two-sided Matching: A Study in Game-Theoretic Modeling and Analysis". Cambridge University Press, New York.

Shimer, Robert, Lones, Smith (2000). "Assortative matching and search". Econometrica 68 (2), 343–369.

Smith, Vernon L., Williams, Arlington W., Bratton, Kenneth, Vannoni, Michael (1982). "Competitive market institutions: Double auctions versus sealed bid-offer auctions". American Economic Review 72 (1), 58–77.