Learning Competitive Equilibrium in Laboratory Exchange Economies

Sean Crockett*
Social and Information Sciences Laboratory
California Institute of Technology

February 2005

Abstract

Crockett, Spear and Sunder [2005b] propose an algorithm whereby boundedly rational agents with standard neoclassical preferences learn competitive equilibrium in a repeated static exchange economy. In this paper a laboratory market is instituted to examine the hypothesis that people are at least as sophisticated as these agents. The adopted market institution strongly restricts the space of agent actions, facilitating the identification of decision rules. Evidence for learning competitive equilibrium is mixed due to strong heterogeneity in decision-making. Some subjects clearly demonstrate the ability to learn across periods. However, a majority exhibit little evidence of learning, and many are, in fact, simply content to *satisfice*, though the opportunity to do better was fairly straightforward. The presence of satisficers permits non-competitive outcomes within this particular market institution, but I conjecture learners may stimulate convergence to competitive equilibrium in others.

Keywords: Disequilibrium, Learning, Pure Exchange, Zero Intelligence, Experimental Economics

JEL Classification Numbers: C91, C92, D50, D83

*Contact: scrockett@ist.caltech.edu Homepage: http://ist.caltech.edu/∼scrockett.
I gratefully acknowledge the advice and tutelage of John Duffy and Stephen Spear. I would also like to thank Bryan Routledge, Shyam Sunder, Dan Houser, Michael Peress, and conference/seminar participants at Indiana University, the University of Arizona, Carnegie Mellon University, and George Mason University for useful feedback. A very special thanks is owed to Allen Geary, Jr. for co-developing the software that made this research possible. For generous financial support, I thank the Tepper School Department of Economics, Stephen Spear, and the International Foundation for Research in Experimental Economics. Finally, thanks to the Pittsburgh Experimental Economics Laboratory for use of its great facilities. All mistakes are my own.

Introduction

Economics lacks a plausible, decentralized theory of competitive price formation. This fact is astounding, when one considers the extent to which economists rely on equilibrium analysis. The profession has mostly been content to take on faith that markets somehow impose equilibrium prices, given which, of course, equilibrium allocations become a trivial consequence of individual optimization.

Crockett, Spear and Sunder [2005b] (hereafter referenced as CSS) develop a theory of competitive price formation, the Negishi algorithm. CSS builds on the Gode, Spear and Sunder [2000] zero intelligence algorithm in which the random generation of bids and offers from agents' welfare-enhancing opportunity sets generates Pareto optimal allocations in a pure exchange economy. In CSS, agents are permitted to know if they are subsidizing others at such allocations and to veto subsidizing allocations,[1] restricting subsequent iterations of the algorithm to those trades that are both Pareto-improving and provide strictly smaller subsidies, and ultimately greater utility, for such agents. In this simple institution actions of minimally sophisticated agents based on local information will lead the market to approximate competitive equilibrium. The algorithm also addresses the behavioral critique of mathematically derived equilibria, namely, that cognitively-limited humans are unable to maximize.

The Negishi algorithm is an existence proof of sorts: CSS demonstrates the existence of an informationally decentralized institution and set of behaviorally plausible strategies such that competitive equilibrium will be obtained for $M \times \ell$ economies populated by agents with generic neoclassical preferences. This a step forward in the competitive price formation literature. However, the Negishi algorithm cannot be taken seriously as *descriptive* of price equilibration without testing the robustness of its assumptions. Do people always restrict themselves to utility-improving trades? Do they learn from past exchange to revise their expected terms of trade over time? If so, do they do so in a way consistent with CSS, or in some alternative way?

Market experiments are uniquely qualified on several dimensions to provide insight in answering these questions. By carefully controlling preferences and institutions, one may isolate key aspects of individual behavior and observe how such behavior influences aggregate outcomes. This paper reports the implementation of CSS economies in the laboratory, in order to assess the robustness of the Negishi algorithm's behavioral restrictions (i.e., zero intelligence plus subsidization constraint ratcheting) and suggest alternatives, if necessary.

The results of the experiments were not entirely anticipated but nevertheless strongly align along several stylized facts. Subjects were generally quite adept at exhausting gains from trade within the individually rational set, but evidence for learning competitive equilibrium is mixed due to a substantial range of heterogeneous decision-making. The behavior of some subjects was consistent with some learning process that ratchets utility gains over time (though not necessarily CSS learning). However, about 70% of subjects exhibited no evidence of learning

[1] Loosely put, an agent is said to be subsidizing others at some Pareto optimal allocation if, given prices defined by the normalized (common) utility gradient at this allocation, he could not afford to buy back his initial endowment.

across periods, and many were, in fact, simply content to *satisfice*, even though the opportunity to potentially do better was fairly straightforward.

The existence of so many non-learners implies the possibility that competitive prices may not be implemented in some institutional settings. In the market institution used in the experiments, subjects were presented with a sequence of random proposed reallocations during a period of exchange, where each subject could see only his own component of each proposal. He could then choose to accept or reject the reallocation of goods in question, with trade taking place only upon unanimous consent. Each period lasted five minutes, portfolios were "consumed" (i.e., exchanged for money) at the end of each period, and each new period began with everyone owning their original endowments. Twenty-five percent of the subjects were content to accept every utility-improving trade proposal throughout the session with few and unsystematic exceptions. Another 35% did systematically reject a fraction of small utility-improvements, but did not exhibit evidence of ratcheting expected utility upwards from one period to the next, and their period-to-period earnings fluctuated greatly. The final allocation in each period of economies populated with these two groups of subjects (about 60% in total) was typically Pareto optimal and established in 2-3 minutes of play. It should take no great leap of sophistication for these subjects to realize that periods are sufficiently long for them to become increasingly selective over time in hopes of "ratcheting up" their earnings across periods, but these subjects (and another 10% whose decisions are more difficult to characterize) do not do so.

Therefore, we observe evidence that people generally engage in activity that improves their condition, but some are much better than others in learning how to improve their condition dynamically. Camerer and Weigelt [1993] come to a similar conclusion with regards to sophistication heterogeneity in a 1-asset (plus cash) double auction experiment. Game theory admits a rich collection of boundedly rational learning processes.[2] The results from this experiment and others suggest that perhaps it is time for general equilibrium theory to investigate the cognitive limitations of at least some subset of agents in a meaningful way.

Related Literature

The standard textbook equilibration story is the *tatonnement* process. Léon Walras introduced the idea that attaining equilibrium could be modelled as a fictitious auctioneer announcing prices, then collecting orders from consumers who specify individual supply and demand at the announced prices. If the aggregate demand for a good exceeds its supply, its price is adjusted upward, and if supply exceeds demand, the price is adjusted downward. Walras reasoned that this procedure would cause the economy to eventually settle into equilibrium. However, Scarf [1960] demonstrated the existence of an open set of economies having a unique equilibrium which is unstable under the Walrasian *tatonnement*. The ensuing stability literature concluded that it is always possible to construct a *tatonnement* procedure specific to a given economy for which some competitive equilibrium would be stable [Saari and

[2] See Fudenberg and Levine [1998] for a survey of learning in game theory.

Simon 1978]. However, the procedure requires information not only about prices, but also about derivatives of excess demand (which is equivalent to knowing second-derivatives of all agents' utility functions) in order to coordinate price adjustment rates across markets, an unobserved phenomenon. Since the substantial information and coordination requirements clearly negate the benefits of decentralized markets, the results from this literature can only be interpreted as strengthening the negative implications of Scarf's example.

An alternative to *tatonnement* processes is to model the economy as a large, strategic Shapley and Shubik [1977] market game. In this model agents submit bids and offers on "trading posts" for each good. A bid for a good is expressed as a specific quantity of an intrinsically valued numeraire good, and an offer is specified as a quantity of the good in question. After bids and offers are collected, the price in each trading post is set equal to the ratio of the aggregate bid to the aggregate offer, and then the posted goods are reallocated at those prices. When agents have no market power (i.e., there are a continuum of agents), Walrasian equilibrium prices correspond to the Nash equilibrium of the underlying game in which agents take the bids and offers of other agents as given and choose their own bids and offers as best responses. Hence, in this model, the question of how the economy arrives at an equilibrium turns on the stability or instability of mechanisms for implementing the Nash equilibrium.

Chatterji and Ghosal [2004] examine a market game with a continuum of agents in which agents may trade in two goods. They show that an out-of-equilibrium adjustment mechanism based on rationalizability of observed bids and offers has features that closely resemble those of the Walrasian *tatonnement* in the sense that any competitive equilibrium stable under the Walrasian procedure will also be stable under their procedure. Of course, the restriction of this result to the case of two commodities limits its usefulness since it is well-known that in this setting, the Walrasian *tatonnement* will always converge to some competitive equilibrium. Ghosal and Morelli [2004] examine market games with any finite number of goods and agents where retrading is permitted, and demonstrate the existence of allocations on the Pareto frontier that can be approximated arbitrarily closely when agents (myopically) play a static Nash equilibrium at every round of retrading.[3]. However, the outcome of these retrading process are not necessarily competitive. Further, the Nash implementation mechanism is undefined.

Another study of market games takes a very different approach. Temzelides [2002] proposes that agents adopt the best-response from the set of actions taken by all agents of their own type in the most recent period of play, subject to a small probability of 'trembling' to an action chosen at random. These trembles are interpreted as mistakes or experimentation. He demonstrates that the Nash/Walrasian equilibrium in this game is the only stochastically stable state of the economy. As with Chatterji and Ghosal, the limitation of this result to two commodities is severe. Also, a practical accounting problem is that agents must keep track of the actions and payoffs of an infinite number of agents of their own type, or a central agency must do it for them.

[3] They also show that the converging sequence of allocations generated by myopic retrading can also be supported along some retrade-proof Subgame Perfect Equilibrium path when traders anticipate future rounds of retrading

Finally, Kumar and Shubik [2004] demonstrate a mechanism (the Cournot-Shubik mechanism) that generates convergence to competitive equilibrium in Scarf's example adapted to the market game setting. However, they go on to make the observation via other examples that the convergence properties of a given mechanism in a market game depend on the underlying parameters of the economy, an observation which is consistent with the findings in the *tatonnement* literature.

A third alternative framework for explaining equilibrium adjustment exploits the game theoretic concept of the core. Game theorists acknowledge that the solution of a two-person bargaining game is indeterminate. However, as the number of participants is increased, the set of solutions that survive recontracting among the players shrinks, and in the limit converges to the set of competitive equilibria [Aumann 1964]. When preferences are private information and recontracting is negotiated directly among individuals, the equilibration process is still missing here. Alternatively, Feldman [1974] and Green [1974] construct a market in which potential coalitions and contracts are periodically selected at random. Recontracting takes place upon the approval of all members of a selected coalition, and it is shown that the core and the set of absorbing states of this recontracting process coincide when each coalition member approves of any utility-improving recontracting proposal. Serrano and Volij [2002] extend this work to the case where players tremble in their decision-making with positive probability (that is, players may occasionally support a coalition that gives them less utility than they currently enjoy), and show that competitive equilibria are the only stochastically stable states of the recontracting process. While this work is encouraging, the specification of a market equilibration process is bought at the price of instituting an abstract process of coalition formation. This drawback is particularly acute in economies with many agents, the very ones where recontracting would appear to offer the most promise of an equilibration process, because the number of coalitions from which to select increases exponentially with the number of agents.

Gode and Sunder [1993] take a very different approach to the problem of implementing competitive equilibrium. They analyze a single market with many interacting agents, based on the standard double auction supply and demand experiments pioneered by Vernon Smith and Charles Plott. In the experimental version of this market, one group of agents plays the role of buyers, the other the role of sellers. Buyers can purchase one unit of the good, and this one unit is worth a given reservation price to them. Hence, if they buy the good for a price at or below their reservation value, they earn a profit. Sellers can each sell up to one unit of the good. If they sell their unit, they incur a given production cost. Hence, if they sell at a price at or above their cost, they make a profit. It is well established in the literature on experimental markets that human traders in this environment eventually end up trading the competitive amount of the good at prices that closely approximate the competitive equilibrium (i.e., transactions take place according to the price and aggregate quantity specified by the intersection of the supply and demand schedules for the experiment). It is important to note, however, that subjects in this experiment generally require several periods of trading before they learn what the relevant equilibrium prices are, so the data generated in such experiments exhibit a convergence of prices and quantities to the predicated competitive equilibrium prices and quantities, rather than an abrupt and direct implementation of the equilibrium.

Gode and Sunder ask whether this process of finding the right prices and allocations is one that requires very sophisticated learning, or whether it could be implemented with "zero intelligence" (ZI) search procedures. They proceeded to replicate the basic experimental setup using computerized robots. The robot traders in their model generated simple random bids (if they were buyers) or offers (if they were sellers) with the only restriction on behavior being that no bid or offer, if accepted, should make an agent worse off. Thus, buyers were restricted to bid below their reservation prices, while sellers were restricted to offer above their costs.[4] In simulations of the model, Gode and Sunder found that while prices don't converge to the competitive equilibrium (CE) prices (as they do with human subjects), the inframarginal prices (i.e. the prices of the last observed transactions) always occur at or near the CE price, while the efficiency of the market is about 98% of the maximum (which occurs when the quantity of the good traded is the CE quantity). These results tell us that the double auction mechanism of the classic supply and demand experiment will implement the competitive equilibrium allocation under very mild conditions on agents' behavior. If we interpret the prices at which the inframarginal trades occur as the limit of the pricing sequence generated in the simulation, then the ZI procedure is also capable of finding the CE prices, as well.

The zero intelligence trading result does not, however, answer the question of whether the competitive paradigm can be implemented easily in environments where many agents trade many goods. Follow-on work by Gode, Spear and Sunder [2000] shows that, at least in the context of a two agent, two good pure exchange economy, simple random search easily finds Pareto optimal equilibria. The random search process does not, however, find the competitive equilibrium. The reason for this is self-evident. The random search process generates a set of random trajectories from the initial endowment to the contract curve, so the ending allocations are not typically competitive except by chance. Crockett, Spear and Sunder [2005b] propose a simple learning rule that, coupled with the utility-improvement rule, is sufficient to drive exchange economies to competitive equilibrium. The next section will review the model in some detail, and extend it to a laboratory environment so that market decision rules can be identified.

Finally, several experimental general equilibrium papers have exhibited promise of decentralized market coordination. Gjerstad [2004] finds much stronger evidence of learning competitive equilibrium than in the present paper, but does corroborate evidence of decision rule heterogeneity and various levels of sophistication in early periods of play. However, the identification of decision rules in Gjerstad's research is apparently more difficult than when using a CSS institution, because the strategy and state space for subjects is much larger. Therefore, the analysis is couched predominantly in per capita terms. It is clear these particular markets typically converged, but *why* remains an open question, as does the robustness

[4] Actually, this a description of their Zero Intelligence - Constrained behavior. They also specified an unconstrained zero intelligence behavior where bids and asks were not restricted to be utility-improving. All references to Gode-Sunder zero intelligence behavior in this paper will be to the constrained version.

of the result to alternative parameterizations. In a macroeconomic environment, Noussair, Plott and Reizman [1995] and Crockett, Smith and Wilson [2005a] demonstrate that subjects are generally adept at coordinating productive activity according comparative advantage and that there is convergence *towards* competitive equilibrium. However, in these papers the focus is not on understanding the underlying price and allocation convergence dynamics, nor does convergence *to* competitive equilibrium generally occur.

The Negishi Algorithm

A finite number of agents trades a finite number of goods in a pure exchange economy. Agents are indexed as $i = 1, ..., M$ and goods as $j = 1, ..., \ell$. Each agent is completely characterized by his preferences, endowments, and decision rules. Preferences are represented by a utility function,

$$u_i : \mathbb{R}^{\ell}_{+} \to \mathbb{R},$$

which is strictly increasing, strictly quasi-concave, and at least twice differentiable. An *allocation* $\mathbf{x}$ is defined as an M-tuple of commodity vectors (these vectors will be referred to as *portfolios*), $[x_1, ..., x_M]$, with $x_i \in \mathbb{R}^{\ell}_{+}$ for $i = 1, ..., M$, so that x_{ij} is agent i's inventory of good j. Let $\omega_i \in \mathbb{R}^{\ell}_{+}$ be the initial portfolio possessed by agent i. Feasible allocations satisfy

$$\sum_{i=1}^{M} x_i = \sum_{i=1}^{M} \omega_i.$$

An allocation is Pareto optimal if it is feasible and there exists no other feasible allocation that makes some agent better off while making no agent worse off. An allocation $[\widehat{x}_1, ..., \widehat{x}_M]$ is said to be a competitive equilibrium if it is Pareto optimal, and such that for each agent

$$p \cdot \widehat{x}_i = p \cdot \omega_i,$$

where p is a vector of strictly positive prices.

The market institution

Fix a number $\varepsilon > 0$, and let

$$\mathbf{Z} = \{\mathbf{z} : |z_{ij} - \omega_{ij}| \le \varepsilon, \; i = 1, 2, \ldots, M, \; j = 1, 2, \ldots, \ell\}$$

Now choose an allocation $\mathbf{z}$ uniform randomly from this set. Call $\mathbf{z}$ a *proposed allocation* and z_i a *proposal* for agent i.

The market institution operates according to the following rules. A proposed allocation within an ε-cube (that is, an $M^{\ell-1}$-dimensional cube with side 2ε) of the current endowment, as above, is generated randomly. Each agent is given the opportunity to accept or reject his individual proposal, where this proposal and his decision are private information. The votes are then collected and tabulated by the institution. If acceptance is unanimous, $\mathbf{z}$ becomes the current endowment; otherwise, the current endowment remains unchanged. Regardless of the outcome, a new proposed allocation is generated in an ε-cube about the current endowment, and agents again vote to accept or reject their individual proposals. This process is repeated until gains from trade have been exhausted. In CSS, exhaustion is defined as such a time when the standard deviation between individual normalized utility gradients is less than a certain threshold. In the experiments, a fixed number of minutes, μ, will transpire, which is assumed to be sufficient for subjects to pursue reasonable trading strategies and exhaust gains from trade. In either case, after the end of a period, goods are consumed, and a new period of trade begins, with each agent possessing his initial endowment. The CSS institution will sustain this process *ad infinitum*, although, of course, in the experiments it will only be possible to administer a finite number of periods.

A theory of behavior

CSS provides two decision rules for agents in this economy and proves that together they guarantee convergence to a near Pareto optimum at the end of each period, and the convergence of end-of-period allocations to a near competitive equilibrium (here "near" conventionally implies arbitrarily close). The first rule is that a trade proposal must improve utility to be accepted. If adopted as a necessary and sufficient condition for accepting a trade proposal, this rule alone, hereafter referenced as *zero intelligence*, guarantees convergence to a near Pareto optimum at the end of each period (see Gode et al.).

The second rule "reduces" the extent of wealth redistribution in each period. The second welfare theorem guarantees any Pareto optimum can be supported as a competitive equilibrium with an appropriate redistribution of endowments in these economies. CSS stands this theorem on its head. Rather than redistribute endowments to make the period-ending allocation a competitive equilibrium, agents are instead given property rights to their endowments. At a near Pareto optimum, agents will agree on relative prices and, hence, on the value (at these prices) of their end-of-period portfolios and endowments. If the agent's wealth at his period-ending portfolio is less than the wealth of his initial endowment, he determines that he is *subsidizing* other agents. The second decision rule is that in all future periods, the agent must provide a strictly smaller subsidy, where this subsidization constraint is redefined each time a subsidizing period-ending portfolio is reached.

CSS adopts these two rules jointly as a necessary and sufficient condition for accepting a trade proposal. The operative hypothesis of this paper is that these rules together are a necessary condition for subjects to accept a trade proposal, implying that subjects will be more selective than CSS trading programs.

In the first period, period 0, the subsidization constraint is not binding, so for now let's simply define the first decision rule:

Assumption 1 Weak Utility-improvement: *For all i, if agent i possesses a current portfolio x_i and is considering a proposal z_i, then he will accept the proposal if and only if $u_i(z_i) \geq u_i(x_i)$ and Assumption 2 is satisfied.*

The second assumption does not bind in the initial period, so if all subjects in the experiment follow this decision rule and there is sufficient time in the period (a consideration discussed in the next section), the final allocation $\widehat{\mathbf{x}}^0$ that is adopted must be approximately Pareto optimal, so the normalized utility gradients for all agents at this allocation will be nearly equal.[5] Let p^0 equal the centroid normalized utility gradient this Pareto optimal allocation. Now, define the i^{th} agent's *gain* at this allocation as

$$\lambda_i^1 = p^0 \cdot \left(\widehat{x}_i^0 - \omega_i\right).$$

Note that here the superscript for the gain references the following period, not the current; this is because λ_i^1 will only have significance in future periods. If $\lambda_i^1 < 0$, agent i is said to be *subsidizing* other agents. Note that if no agent in the economy is providing any subsidies, then the economy has reached a near competitive equilibrium, since $\lambda_i^1 \geq 0$ for all i implies

$$\sum_{i=1}^{M} p^0 \left(\widehat{x}_i^0 - \omega_i\right) = p^0 \cdot \sum_{i=1}^{M} \left(\widehat{x}_i^0 - \omega_i\right) \geq 0,$$

where p^0 is the centroid normalized utility gradient. From the assumptions on utility functions, it must be the case that $p^0 \gg 0$, while, from feasibility, $\sum_{i=1}^{M} \left(\widehat{x}_i^0 - \omega_i\right) = 0$. It may then be inferred that $p^0 \cdot \left(\widehat{x}_i^0 - \omega_i\right) \approx 0$, and the economy has reached a near competitive equilibrium.

Suppose some agent is a subsidizer at the end of the first period, and a new period of the trade proposal process is begun with each player possessing his initial endowment. As before, each agent will only accept trades that are utility-improving in the new period. However, for any agent i such that $\lambda_i^1 < 0$, all acceptable proposals must also provide strictly greater wealth, given prices p^0, than that provided by the initial portfolio, $p^0 \cdot \widehat{x}_i^0$. This process generalizes to future periods until all subsidization constraints are within a small neighborhood of zero, and thus the economy has reached a near competitive equilibrium.[6]

Assumption 2 Decreasing subsidization: *Suppose period T has just ended. Define $\lambda_i^0 = -\infty$ and $\lambda_i^{t+1} = p^t \cdot \left(\widehat{x}_i^t - \omega_i\right)$ for all $t \in [0, T]$ and all agents i. Let $\tau \leq T$ be the most recent period such that $\lambda_i^{\tau+1} < 0$ (that is, if $t \in (\tau, T]$, then $\lambda_i^{t+1} \geq 0$ or*

[5] Eventually proposals will be restricted to a discrete grid. Therefore, while an allocation may be Pareto optimal in the sense that no Pareto-improving trades can be found on the grid, it generally will not lie on the continuously defined contract set. Of course, by making the grid arbitrarily fine we can guarantee that such allocations are as close to the contract set as we would like.

[6] An observant reader will notice there is no guarantee an economy must reach a near competitive equilibrium in each future period after a near competitive equilibrium has been implemented. In CSS, the algorithm is simply halted once a near competitive equilibrium is found. If the process continues indefinitely, then it is possible to go in and out of equilibrium from period to period. However, eventually it must be the case that all agents possess subsidization constraints within a certain tolerance δ of being zero. At this time, all future periods of trade will end near CE.

$t = T$ and $\lambda_i^{T+1} < 0$). $\lambda_i^{\tau+1}$ is agent i's subsidization constraint for period $T+1$. Fix $\eta > 0$ and let x_i^{T+1} be any proposal under consideration for agent i in period $T+1$. If $p^\tau \cdot \left(x_i^{T+1} - \omega_i\right) < \lambda_i^{\tau+1} + \eta_i$, then the proposal will be rejected by agent i. Here, p^τ denotes the Pareto optimal "price" (i.e., normalized utility gradient) at the end of period τ.[7]

The intuition for *Assumption 2* is straightforward. If a period ends and agent i, given "prices" defined by the utility gradient at his current portfolio, cannot buy back his initial endowment at those prices, he will infer that he should have gotten a better return for his endowment. Thus, in the future he will demand to remain above the wealth hyperplane defined at the period-ending portfolio in question, until a new subsidization constraint is defined. A subsidization constraint serves as a signal to ratchet up expectations on the portfolio one should receive for one's endowment.

CSS proves that the intersection of the Pareto set and the set of proposals that satisfy the first and second assumptions is non-empty and nested over time, and that its limiting point is a competitive equilibrium. Therefore, as long as subjects do not violate these assumptions, and have sufficient time to exhaust gains from trade in each period, then with a sufficient number of periods they will converge to a near-competitive equilibrium allocation. Fix a small δ and suppose period T has just ended. If $\lambda_i^{T+1} < \delta$ for all i, the economy has converged to a δ-competitive equilibrium. See Figure 1 for a visual depiction of convergence in a two-agent, two-good economy. The point ω is the initial endowment. The random proposal generation process in the first period will reach some point on the contract set, like point A. At point A, agent 1 is the subsidizer, and agent 2 the subsidy receiver. Therefore, in all future periods, agent one will never accept a trade below the budget line through A, ratcheted up by some positive amount η. Suppose in the second period the economy reaches point B. Now agent two is the subsidizer, and defines his new subsidization constraint. In all future periods, trade (in addition to being Pareto-improving) must take place in the shaded gray region. CSS proves this process must converge to a δ-competitive equilibrium.

Note that in passing from one period to the next, each agent must always carry along his most recent subsidization constraint, even if he moves to a new portfolio in which he is currently receiving a subsidy . If an agent "forgets" his past subsidization constraint, then the trading process could go back to an allocation in which this agent was again making losses, possibly larger than in the previous period. Hence, at each stage t, the data required for each agent is $\left[\widehat{x}_i^t, \lambda_i^{\tau+1}, p^\tau\right]$, where $\lambda_i^{\tau+1}$ and p^τ were the price and loss in the last period τ at which agent i incurred a loss.

[7]When there are more than two goods in the economy, CSS strengthens the subsidization constraint a bit to rule out cycles between subsidization constraints, although this strengthening is not necessary to obtain a weaker convergence result. Since only two goods appear in this experiment, cycling is not an issue.

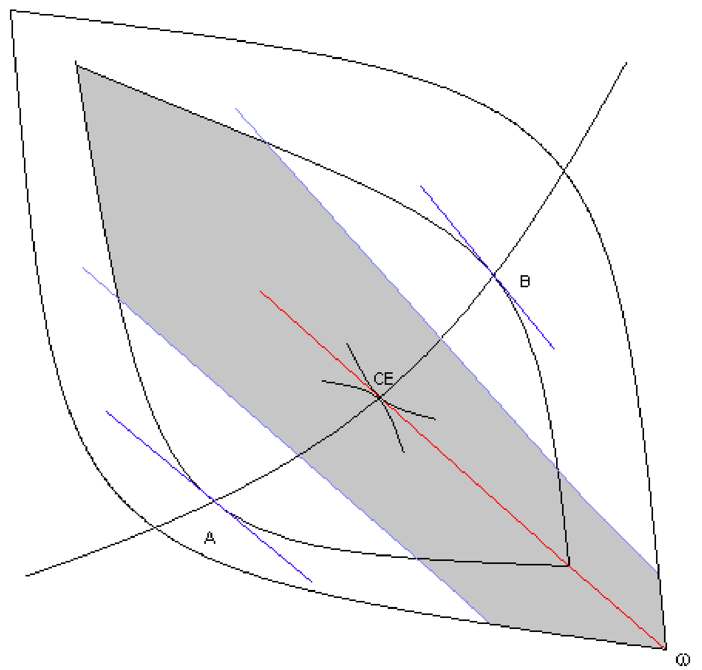


Figure 1: Convergence to Competitive Equilibrium

Numerical and practical constraints in implementation

Restriction of trade to integer quantities

A random generator will only generate numbers to a certain level of precision (typically double precision, which utilizes roughly 17 significant digits), so an unavoidable restriction when implementing the market institution in practice is to impose a grid on the space of feasible allocations. The loss of the off-grid points poses a problem for the CSS convergence proof. Intuitively, consider a two agent, two good economy and an allocation such that the two indifference curves through this point are close to tangent, but not quite 'near Pareto optimal' (recall that we have loosely defined "near" to be as close to tangent as we please). CSS proves it must be the case that the subsidy-constrained set of Pareto-improving trades at this allocation is non-empty. However, this set does not necessarily intersect the grid, and may, in fact, lie between grid points. Of course, in theory we may make the grid as fine as we like, but not in practice. Therefore, there is no guarantee the economy can reach a near Pareto optimum unless the grid is sufficiently fine.

The problem can't be eliminated in general. The only thing to be done is make certain that for the specific parameterization chosen, the closest grid points to the contract set are sufficiently close, so that they satisfy the adopted condition for near Pareto optimality. If not, the proposed decision rules may not be sufficient to generate a δ-competitive equilibrium.

Since proposed allocations must lie on a grid any way, it is without loss of generality to restrict endowments and trade proposals to integers. Of course in doing so, endowing agents with less than $1 * 10^{17}$ of each good loses available "fineness." As stated above, this is not problematic as long as the δ-competitive equilibrium is attainable along any path, given the chosen grid and obeyance of the proposed decision rules. It should also be noted that the experimenter's budget constraint does not permit salient incentives to be induced over too fine a grid, any way. If the grid is too fine, subjects may see proposals that do not impact their earnings (the experimenter can only pay to the nearest penney), and yet the theory will make strong predictions of behavior in these circumstances. Parameterization of the economy is done on a grid as fine as salience will allow, and then tested on a case-by-case basis to determine if induced preferences are such that subjects faithful to the theory will indeed generate allocations sufficiently close to the contract set to permit convergence to a δ-competitive equilibrium.

Periods are of constant length

Another practical constraint in the implementation of this experiment is to allow subjects the opportunity to exhaust gains from trade while not forcing them to do so. In the Negishi algorithm, near Pareto optimality is reached when the normalized utility gradients of all agents are approximately equal (specifically, the standard deviation of their normalized utility gradients must be less than a small, pre-specified number). Since subjects are not forced to exhaust gains from trade in the experiments, a different criterion is necessary. The market institution in the experiments declares the end of a period after μ minutes of trading. Therefore, μ must be large enough to allow subjects the opportunity to be sufficiently selective in the proposals they choose to accept, but short enough to minimize boredom and permit enough periods of trade for learning to be exhibited. In pilot studies, 5 minutes appeared to be the minimum amount of time necessary for the most sophisticated subjects to exhaust gains from trade (in general, these subjects make very few utility-diminishing trades, and hold out for trade proposals that meet certain conditions, so they typically take longer to reach a Pareto optimum than other subjects), so 5 minutes was adopted as the length of each period in the experimental economies. It should be noted that in pilot sessions with longer periods, boredom among many subjects was pronounced.

Restriction to 2×2 economies

The experiment was restricted to two agent, two good economies. Given the need for a sufficiently fine good-space as outlined above, a graphical presentation of preferences (rather than tables) was deemed necessary. Restricting the graphical field to two dimensions was driven by the desire to keep things as simple as possible for the subjects in an already complicated experiment.

The restriction to two agents is motivated less by wanting to "keep it simple" for subjects, and more because larger economies create analytical complications that would be difficult to disentangle. First, if there are more than two subjects in the economy, and if subjects reject proposals at the same rate in 2- vs. M-subject economies (as they would in CSS), it will take larger economies longer to converge. If periods remain restricted to 5 minutes, subjects may have little choice but to accept many more utility-improving trades in order to exhaust gains from exchange. Thus, the time length restriction mentioned above would have important consequences in the comparative static analysis of economies of various sizes. Second, the more agents are added, the more substantive the departure of the CSS institution from most other market institutions. In particular, in the CSS institution, each agent has veto power over all trades, regardless of how many agents populate the economy. This is not the case in double auctions for more than two people, for example, since one's abstention from trade does not affect the ability of others to trade amongst themselves. However, there is apparently no good way to allow for trade amongst subsets of traders in the CSS institution, given the need to maintain feasibility. Thus, while the relative bargaining power of an agent in a double auction decreases as the number of agents gets large, this is not necessarily true in the CSS institution.

The insight that can nevertheless be gained by studying 2×2 economies in the CSS institution is to determine the extent to which subjects can learn to demand better terms of trade over time. In fact, one would expect they should be able to do so more easily in 2-person economies rather than larger ones, since according to cooperative game theory the entire individually rational set is up for grabs. Therefore, by analyzing 2×2 CSS economies, we give learning its greatest chance of success. To anticipate the results, the fact that satisficing is so prevalent even this very explicit bargaining environment is evidence that such behavior may be robust to many environments. However, we do find evidence of sophisticated learning by a smaller subset of subjects, although this learning is typically not consistent with CSS learning. The 2×2 treatment likely gives us the most mileage we can get out of the CSS institution: Satisficing may be a very robust phenomenon, but some people are clearly more sophisticated.

Trade proposals are occasionally "directed"

With a uniform random generator, it generally takes longer to generate a Pareto-improving proposed allocation as Pareto-improving trades are consummated. To refrain from setting μ impractically but necessarily high to present subjects with the opportunity to exhaust gains from trade, I incorporate a "directed" proposal generator to periodically draw only from the set of Pareto-improving proposed allocations. Each proposed allocation is calculated by a uniform random generator with probability $m\alpha$, or a directed generator with probability $1 - m\alpha$, where $\alpha < 1$. On the first proposal, m equals zero, on the second one, on the third two, etc. Once a Pareto-improving proposed allocation is generated (regardless of whether or not it is adopted), m is reset to zero. In all of the treatments in this experiment, α was set to 0.2. It should be emphasized here that the directed generator only restricts proposals to be Pareto-improving, not necessarily subsidy-surviving. The reason for not restricting the directed proposals to the subsidy-surviving set is that if subjects wait (collectively) for Pareto-improving proposals, and these proposals are restricted to be within the subjects' subsidization constraints, then it may appear that subjects are demanding to stay above these constraints, when in fact they may not be doing so at all.

Uniform random proposal generator: Consider trade in two goods, X and Y, and let T_{xi} and T_{yi} be the length of the axes represented on agent i's monitor. In each treatment I have chosen endowments such that $T_{xi} = T_{yi} = 1000$ for all i. Ignoring the subscript i, let $\varepsilon = \frac{1}{10}T_x = 100$. The uniform random generator computes a proposed allocation such that (1) Each individual's proposal is within an ε-cube of the current allocation; (2) Each agent is asked to give up some amount of one good and receive some amount of the other (i.e., no free lunch); (3) No agent can be proposed to assume a zero or negative amount of a good; and (4) The proposed allocation must reside within the Edgeworth box.

The trade generator begins by choosing one agent at random to complete the Edgeworth box. For the other agent, two random integers are generated. The first will be the proposed net change in X, the second number the proposed net change in Y; one is randomly assigned to be negative. The absolute value of each number will be uniform randomly generated between 0 and the minimum of ε and the maximum quantity that will keep the proposal within the Edgeworth box. Minus each of these net proposals will be tentatively assigned to the agent who has been chosen to complete the Edgeworth box. If the proposal would put the economy outside of the Edgeworth box, the uniform random trade generator is restarted. This process continues until an appropriate trade proposal has been generated.

Directed random proposal generator: The directed random proposal generator is similar, except the proposal generated must be Pareto-improving. The net trade vector generated this time is restricted to be an acute deviation from the agent's utility gradient, and the agent who most values X is required to receive X, in order to improve the speed with which a Pareto-improving proposed allocation is found. After a pre-determined number of failed attempts, the directed generator is discarded in favor of the uniform random generator; presumably, the economy is close to the Pareto set.

A criticism of using the directed proposal generator is that the market institution occasionally uses private preference information to generate trade proposals. CSS is subject to a similar complaint because the period termination condition utilizes private information. However, in both cases this information is only used to enhance the opportunity to converge to a near Pareto optimum in a reasonable amount of time, and is not strictly necessary. The directed generator's purpose is to shorten the length of the experiment and thus reduce subject boredom, and potentially mitigate the influence of subjects' own time preferences on the results.

Experimental Design

Subjects were recruited through the University of Pittsburgh Experimental Economics Laboratory (P.E.E.L.) during academic year 2003-4. Most subjects were undergraduates enrolled at the University of Pittsburgh or Carnegie Mellon

University. Each was paid \$5 for completing the experiment and additional performance-based compensation. Specifically, subjects were awarded points at the end of each period and were paid the sum of these period-ending points divided by 100, in dollars, at the end of the session (so each point was worth 1 cent).

Subjects were presented with two minutes of verbal instruction to highlight payment procedures. They were then instructed to familiarize themselves with the trading interface and rules of the game by completing an on-line tutorial, which can be downloaded from `http://www.ist.caltech.edu/~scrockett/`. Subjects were permitted to complete the tutorial at their own pace, and signal the experiment supervisor when ready to begin. They also received printed instructions for convenient reference during the experiment.

Four treatments were developed for this experiment. All imposed preferences were a monotonic transformation of the constant elasticity of substitution function:

$$U(x,y) = \frac{\left[\left(p_1^{\frac{1}{p_3}} x^{\frac{p_3-1}{p_3}} + p_2^{\frac{1}{p_3}} y^{\frac{p_3-1}{p_3}}\right)^{\frac{p_3}{p_3-1}}\right]^{p_4}}{p_5}$$

The transformation involved (i.) Exponentiating the utility function (with $p_4 = 2$ in all treatments) to magnify utility differences between portfolios (increasing the salience of decision-making with respect to payment), and (ii.) Dividing the transformed function by an amount (p_5) designed to make the mean expected performance-based payment equal to \$10. The chief advantage of using the CES utility function is that parameters p_1, p_2 and p_3 can be chosen such that goods are strongly complementary at the initial endowment. This generates a large individually rational set, which permits an economy populated with utility-improving agents the opportunity to exhibit learning between periods (or not). Consider the other extreme, where preferences are nearly linear. In this case, any sequence of Pareto-improving trades to the contract set will end near the competitive equilibrium. Therefore, agents who follow *Assumption 1* and ignore *Assumption 2* will "converge" near the competitive equilibrium in each period without learning. Imposing near-Leontief preferences at the initial endowment provides the best opportunity to reject the motivating theory for paper. In all treatments Player 1 began with an initial endowment of 800 units of X and 200 units of Y, while Player 2's initial endowment was 75 units of X and 925 units of Y.

Treatment 1 was designed to test *Assumption 2*. The role of Player 1 was assigned to a subject, and Player 2 to a CSS computer trading program.[8] Preferences were skewed towards the subject being a subsidizer under zero intelligence play; in particular, $p_1 = 0.4$ and $p_2 = p_3 = 0.6$ for the subject, while these values equalled 0.75, 0.25, and 0.99, respectively, for the CSS robot. This treatment thus presents a uniform and reasonably facilitating trading partner against which to assess subject behavior, unless subjects are very assertive with regards to what proposals they will accept. Figure 2 presents a scatter plot of 1000 final allocations in ZI simulations of

[8] Subjects were informed they could be matched with other subjects or computerized trading programs. The trading programs were designed with a built-in decision lag of 2 to 6.5 seconds to simulate human response. This lag may have been slightly too long, as the number of proposals generated in Treatment 1 was about 15% less than the others, on average.

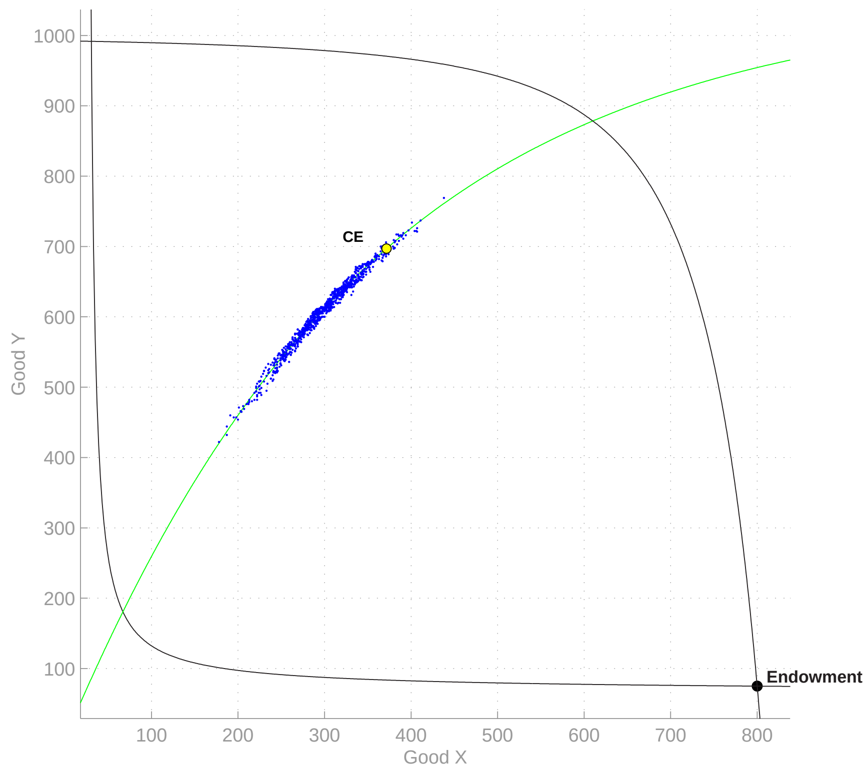


Figure 2: Final allocations for 1000 ZI simulations, Treatment 1

this parametrization. Clearly, if subject i were to adopt a ZI trading rule in period t, it would most likely be the case that $\lambda_i^t < 0$. Only 5.1% of the simulations resulted in Player 1 not being a subsidizer in the first period.

Treatment 2 shares the identical parametrization of Treatment 1, but with both Players 1 and 2 now being assigned to subjects. This treatment was designed to facilitate the analysis of non-CSS learning behavior. If subject 1 in this treatment is either zero intelligence or CSS, then subject 2 should reject relatively few utility-improving trade proposals (particularly in early periods) if he is CSS, since most of these proposals will not violate his subsidization constraint. If, on the other hand, subject 2 rejects utility-improvements that do not violate his subsidization constraint, then this fact would suggest the presence of a non-CSS learning rule that might not be picked up by Treatment 1 because it is correlated with CSS when the subject is a subsidizer. That is, a non-CSS learning process might be observationally similar to CSS for subject 1 but not for subject 2. In Figure 3 is depicted the median period-ending allocations for periods 1-10 of 1000 CSS simulations. This statistic is computed by finding the median utility for agent 1 and choosing the corresponding allocation. As is apparent from the figure, the median period of convergence in these simulations was the sixth, although the fourth and fifth periods are relatively close.

Treatment 3 provides an opportunity to analyze how subjects behave when each is equally likely to be a subsidizer under zero intelligence play. Here $p_1 = p_2 = 0.5$ and $p_3 = 0.6$ for both players. To the extent that learning is found in the first two treatments, this treatment lends itself to evaluating what happens to the decision

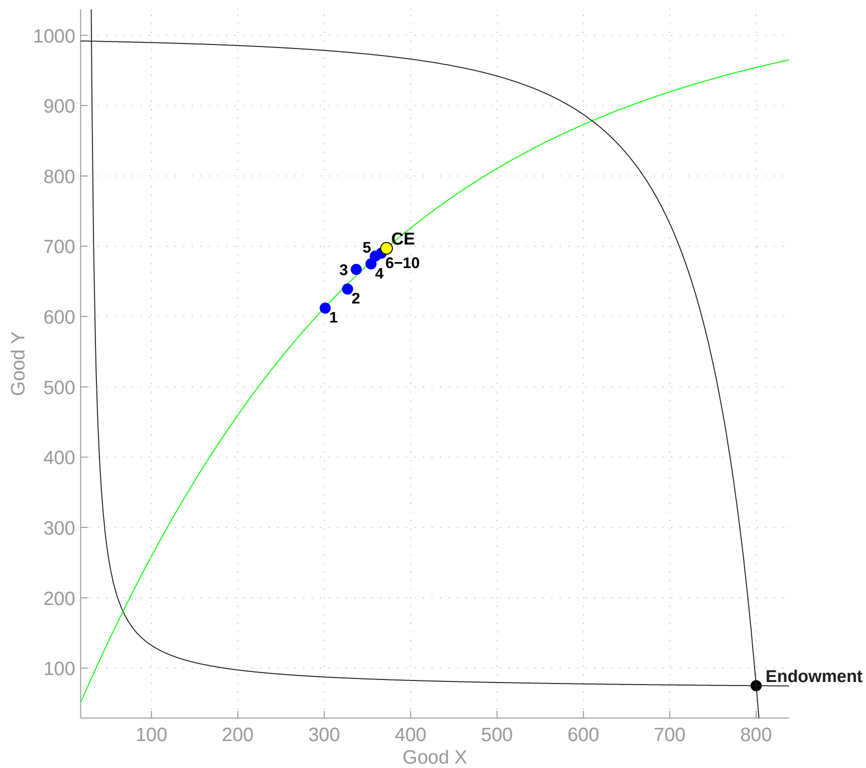


Figure 3: Median final allocations for 1000 CSS simulations, Treatment 1

process when both subjects are potentially frequent subsidizers. Finally, it should be noted that $p_5 = 7,656$ for Player 1 in this treatment, and 12,656 for Player 2. In Treatments 1 and 2, these numbers were 9,000 and 10,000, respectively.

Irrespective of treatment, the presentation of constraint information to subjects was a difficult issue. It was certainly not desirable for them to be coached on the issue of subsidization, but they needed to be given some sort of relevant information in order to test the theory. Define the *value* of good X per unit of Y as the marginal rate of substitution of X for Y at the period-ending portfolio (in the instructions it is done with less jargon and using illustrations). Define *portfolio value* as current wealth given this value. At the end of each period, subjects were presented with *value* and *portfolio value* for their period-ending portfolio and endowment, and this data remained on-screen for the remainder of the session. They were not instructed how to use this information, or if they should use it at all. If subjects were subsidizers, the portfolio value of their endowment will necessarily be larger than the portfolio value of their period-ending portfolio. In addition, subjects were given the "value line" (that is, the budget line) through their period-ending portfolios. Again, they were not instructed how to use this line, and they could toggle it off and on for the reminder of the session as desired. See Figure 4 for a screen-shot of the user interface that contains all of this information. Here, Player 1 is a subsidizer.

After running Treatments 1 through 3, there was little evidence to support CSS learning. This may be in part due to the fact that subjects are not trading in prices, but rather are engaged in barter. The presentation of "price" information at the

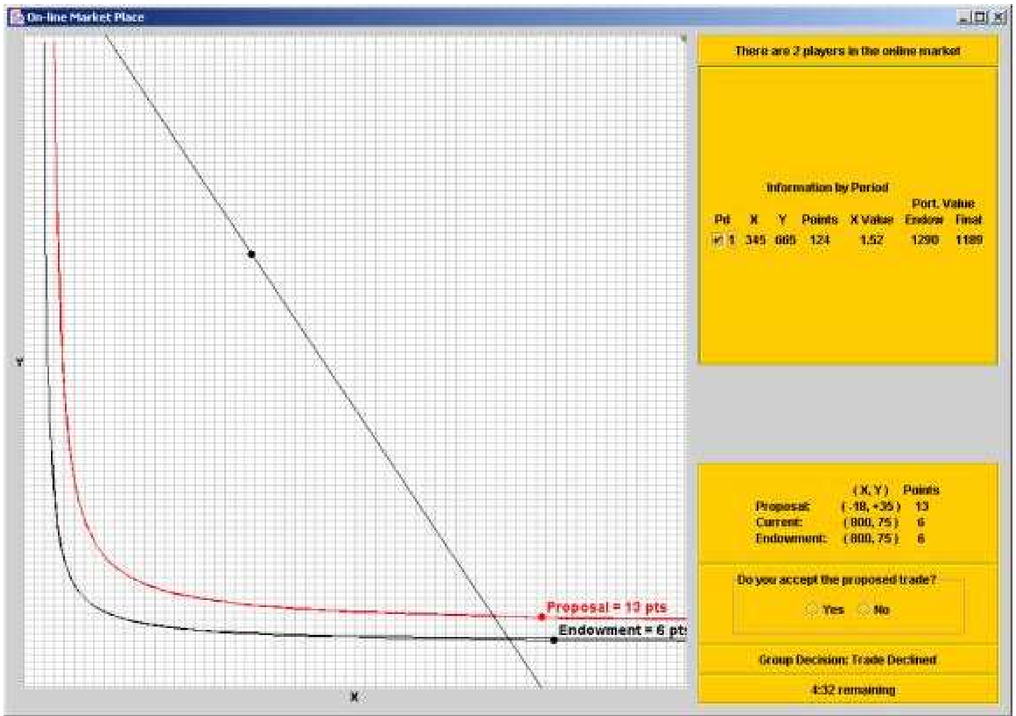


Figure 4: User Interface, Second Period of Play

end of each period is apparently artificial, since prices are not used explicitly during the trading process as in a double auction or posted price mechanism. Therefore, **Treatment 4** was designed to draw more attention to period-ending marginal rates of substitution and the resulting implication for wealth redistribution. This was done by explaining that when the relative valuation of goods across subjects is approximately equal, no more gains from trade are available, in which case the current period will end early. They were then coached on the implications of this common valuation for implied wealth redistribution. They are told that being a subsidizer is a signal that someone has gotten the better of them in the trading process, since they did not get the full termination value of the goods they 'sold.'

Results for Treatments 1-3

Because Treatments 1-3 were designed in tandem and Treatment 4 as an ex post robustness check, the results for the first three treatments are presented separately from the last. It should be noted before proceeding that there does not appear to be a significant difference in behavior between any of the four treatments.

Exhausting gains from trade

Across all treatments, subjects typically reached a final allocation very close to the Pareto set in each period. One measure of convergence is the ratio of the distance

	Mean	Median	Standard Deviation	Observations
Treatment 1	0.019	0.011	0.020	18
Treatment 2	0.058	0.017	0.079	8
Treatment 3	0.058	0.019	0.086	9

Figure 5: Relative distance of final allocations from the Pareto set

	Period									
	1	2	3	4	5	6	7	8	9	10
Treatment 1	0.033	0.014	0.020	0.013	0.025	0.023	0.018	0.016	0.019	0.010
Treatment 2	0.171	0.063	0.084	0.023	0.016	0.017	0.032	0.029	0.063	0.087
Treatment 3	0.139	0.113	0.063	0.084	0.026	0.052	0.019	0.023	0.036	0.021

Figure 6: Mean distance of final allocations from Pareto set

between an allocation and some 'nearby' point on the Pareto set to the distance between this point and the endowment. That is, if x is the final allocation, ω is the initial endowment, and γ is a point on the Pareto set, then the measure is:

$$\frac{\sqrt{(x-\gamma)'(x-\gamma)}}{\sqrt{(\omega-\gamma)'(\omega-\gamma)}}$$

In choosing γ, it makes intuitive sense to take the mean marginal rate of substitution at x and find the intersection of the Pareto set with the line defined by x and (minus) this slope. The point thus found will be a plausible member of the contract set relative to x, which may not be true of the nearest Pareto optimal point to x.

Figure 5 was derived with this formula. The figure indicates that subjects were generally adept at not leaving gains from trade on the table. For example, in Treatment 1 the mean distance of these economies from the Pareto set is less than 2% relative to the endowment, and the median ratio is just over 1%. So across periods the typical individual gets nearly 99% of the way from the initial endowment to the Pareto set, or goes just past it. Because points are rounded up to the nearest penny, it is, in fact, almost never the case that the typical individual at the end of an average period could earn even one more penny by way of Pareto-improvement. This indicates very strong convergence to the Pareto set. In Treatments 2 and 3, the median individual ends up a little further away (roughly 2% rather than 1%), but is still quite close. A couple of sessions that do not generally exhibit convergence bias the mean distance upwards in both treatments. The primary reason that subjects in Treatment 1 get a little closer to the Pareto set is that the CSS robots accept nearly all Pareto-improving trades, since their subsidization constraints are largely untested, whereas some subjects frequently reject very small utility improvements.

In Figure 6 is displayed the mean normalized distance to the Pareto set across individuals by period. Note that the first period distance is substantially larger than the rest in all treatments, but the bulk of subjects adapt quickly to not leaving

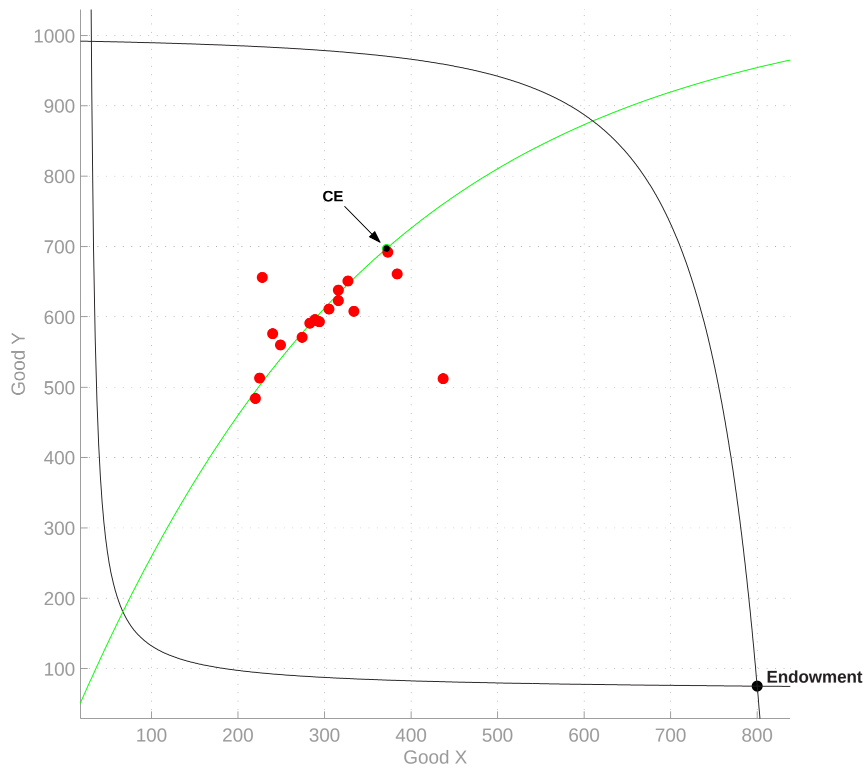


Figure 7: Treatment 1, period-ending allocations in first period

significant gains from trade on the table. The medians by period (not displayed) are typically much smaller than the means in Treatments 2 and 3.

Aggregate evidence of learning

We may adopt a similar measure to determine how near an economy is to the competitive equilibrium. As it turns out, the median distance does not tell a story of convergence. Across treatments, the median distance from competitive equilibrium at the end of the first period is nearly equal to the median distance in ZI simulations (the ZI simulated median distance was 0.15 for Treatments 1 and 2, and 0.076 for Treatment 3; the first period median distances in the experiment were 0.173, 0.160, and 0.086 for Treatments 1-3, respectively). In Treatment 1, this distance generally decreased over time, but remained above 0.10; in the other two treatments, it fluctuated quite a bit.

So we know that at least half of the economies do not typically converge to the competitive equilibrium. Do any? In all treatments, some subjects do take strong bargaining positions and steer final allocations in their economies towards the boundary of the other subject's individually rational set. In Treatment 1, subjects can only push things as far as the competitive equilibrium due to the robot's subsidization constraints. Therefore, it is interesting to consider how many subjects can learn to push their CSS trading partners close to the competitive equilibrium.

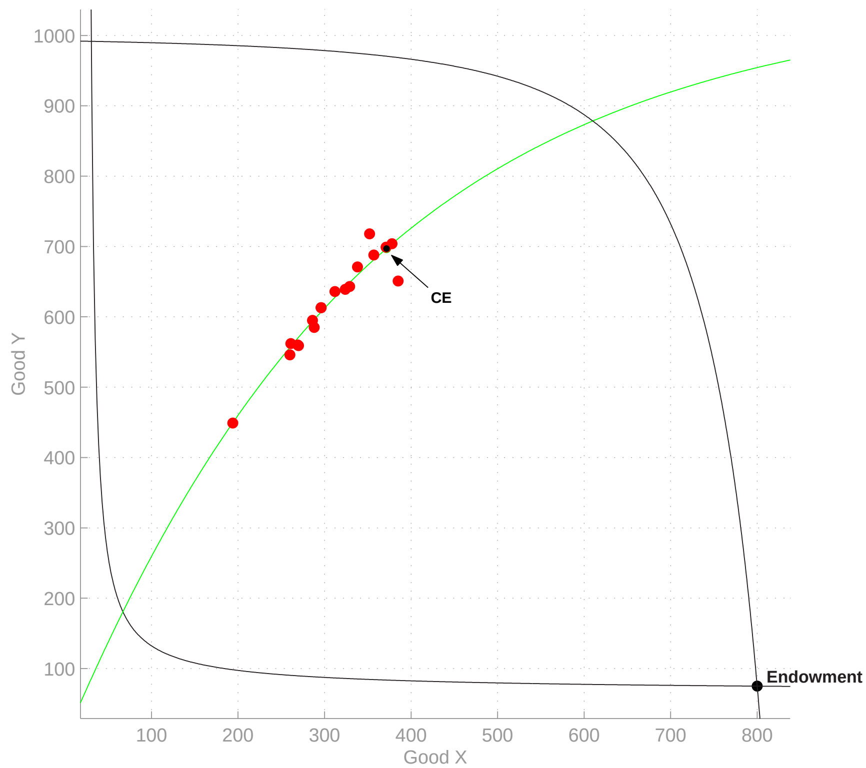


Figure 8: Treatment 1, period-ending allocations in tenth period

The results suggest that some subjects learn to do so while others do not, which anticipates a closer look at behavioral heterogeneity in these experiments.

Figures 7 and 8 present period-ending allocations for each session in Treatment 1 for the first and tenth periods, respectively. The data suggest that some subjects may, in fact, learn competitive equilibrium, because there is a clustering of allocations near the competitive equilibrium in the tenth period that did not exist in the first. Consider Figure 9. In it is presented the proportion of sessions that finish within a certain normalized distance of competitive equilibrium, by period, and corresponding zero intelligence benchmarks. On the horizonal axis is period number. The vertical axis measures the proportion of economies 'near' competitive equilibrium. The red lines represent the proportion of simulated ZI economies within a certain distance of CE, and the blue lines track the proportion of sessions in a particular treatment within such a distance. For example, 5.4% of zero intelligence simulations end within 2.5% of competitive equilibrium. In period 1, 5.6% of the economies in Treatment 1 finish as close; they are nearly equal. By period 10, however, 16.7% of these economies finish within 2.5% of competitive equilibrium, three times more likely than ZI traders. This difference across periods is almost surely due to learning by some individuals rather than chance; the probability of ZI traders generating such results is nearly zero. This trend is mirrored in consideration of other "rings" about competitive equilibrium.

It appears to be the case, at least in Treatment 1, that human subjects actually do a little worse than zero intelligence subjects on average in the first period and

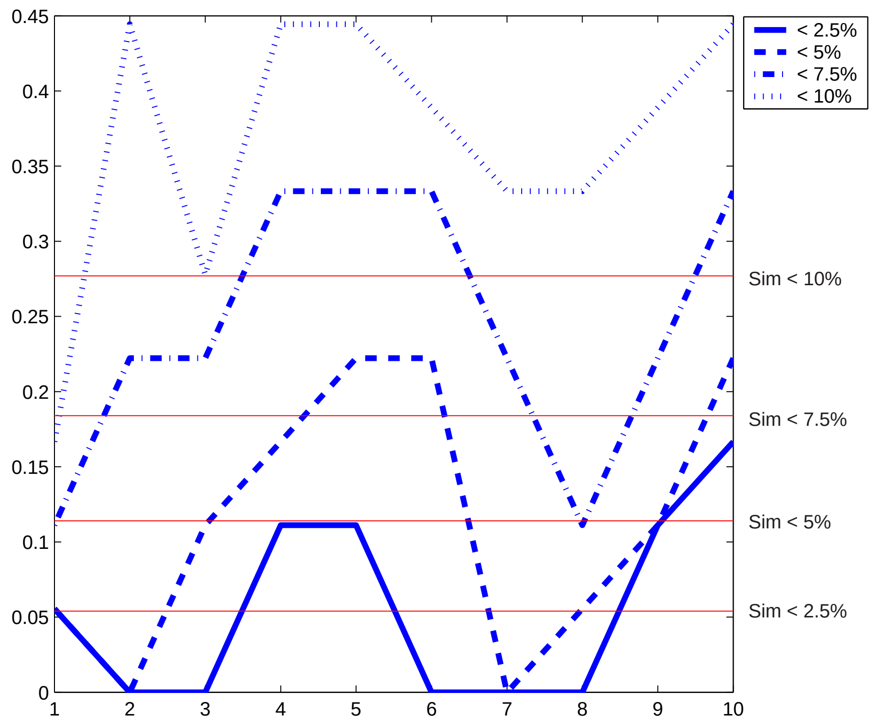


Figure 9: Proportion of economies 'near' competitive equilibrium in Treatment 1

then quickly learn to do at least as well, with some subjects learning to extract the full extent of gains from trade available from their CSS trading partners. To better understand the specific form of this apparent behavioral heterogeneity, and to make use of the data from all treatments, it is necessary to shift the analysis from the aggregate to the subject level.

Individual decision-making

The analysis presented thus far anticipates several potentially interesting statistics to consider at the individual decision-making level. Chief among these, given Assumption 1 and the fact that the period-ending data appear to exhibit some similarities to zero intelligence play, is the proportion of proposed losses rejected and gains accepted. Another is the correlation between the proportion of accepted proposals and the number of periods: CSS learners, and perhaps other types of learners as well, will generally become more selective as time passes, so the rate of accepted trade proposals would be expected to decrease across periods. A third potentially interesting statistic is the correlation between the rate of proposal acceptance and the magnitude of a proposed gain (or loss). It seems reasonable to expect that most learners will be more likely to reject small gains than larger ones, particularly as time passes. Figure 10 displays this information along with payment data, excluding the $5 participation fee.

				Correlation w/ Loss Accept			Correlation w/ Gain Accept	
Subject	**Treat**	**Pay**	**Prob. Loss Accept**	**Period**	**(-) Size of Loss**	**Prob. Gain Accept**	**Period**	**Size of Gain**
1	1	8.74	0.08	0.08	0.15	0.93	-0.24	0.12
2	1	8.75	0.04	-0.15	0.14	0.95	0.21	0.20
3	1	8.28	0.04	-0.28	0.10	0.62	0.03	-0.07
4	1	9.90	0.01	0.10	0.05	0.77	-0.42	0.18
5	1	9.39	0.08	-0.32	-0.05	1.00	0.09	0.06
6	1	9.36	0.04	-0.04	0.22	0.99	-0.07	-0.05
7	1	9.67	0.00	-0.07	-0.09	0.99	0.07	0.04
8	1	8.79	0.06	-0.05	0.21	1.00	-0.05	0.02
9	1	7.26	0.25	-0.15	0.45	0.95	-0.11	0.13
10	1	8.82	0.02	-0.08	0.03	0.97	0.06	0.09
11	1	9.13	0.00	-0.09	0.06	1.00	0.07	0.04
12	1	8.87	0.00	NaN	NaN	0.98	0.20	0.04
13	1	11.29	0.13	-0.14	0.26	0.67	-0.06	0.38
14	1	9.75	0.19	0.00	0.32	0.80	-0.08	0.21
15	1	10.29	0.04	-0.09	-0.07	0.73	0.25	0.26
16	1	9.65	0.09	0.12	0.10	0.44	-0.05	0.43
17	1	9.95	0.01	-0.08	0.09	0.94	-0.02	0.21
18	1	8.52	0.06	0.02	0.24	0.98	-0.02	0.08
19a	2	4.31	0.03	0.04	0.05	0.98	-0.11	0.02
19b	2	17.19	0.11	0.13	0.25	0.66	-0.50	0.03
20a	2	10.08	0.02	-0.11	0.05	0.79	-0.20	0.18
20b	2	10.51	0.01	0.02	0.00	0.98	0.00	0.09
21a	2	13.34	0.25	0.32	0.30	0.75	-0.20	0.34
21b	2	8.26	0.12	-0.38	0.07	0.94	0.18	0.15
22a	2	8.94	0.01	-0.11	0.03	0.97	-0.09	0.08
22b	2	11.26	0.03	-0.07	0.05	0.98	0.09	0.14
23a	2	12.05	0.01	-0.04	-0.04	0.79	0.59	0.02
23b	2	6.82	0.50	0.21	0.07	0.59	-0.11	0.20
24a	2	7.86	0.02	0.04	0.10	0.99	0.13	0.05
24b	2	12.34	0.01	-0.01	-0.02	1.00	0.03	-0.04
25a	2	9.85	0.08	-0.26	0.04	0.49	-0.07	0.43
25b	2	10.39	0.04	-0.08	0.19	0.67	-0.29	0.20
26a	2	10.17	0.02	-0.04	0.00	0.98	-0.06	0.07
26b	2	10.20	0.01	0.03	0.04	0.98	-0.08	0.11
35a	3	6.99	0.19	-0.52	0.00	0.86	0.41	-0.16
35b	3	13.93	0.00	-0.07	0.00	0.99	-0.04	0.00
36a	3	9.72	0.00	NaN	NaN	0.99	0.02	0.02
36b	3	10.31	0.02	0.08	0.10	0.93	0.01	0.05
37a	3	9.56	0.00	NaN	NaN	1.00	0.08	-0.04
37b	3	10.34	0.17	-0.29	0.17	0.75	0.47	0.20
38a	3	12.71	0.04	0.04	0.17	0.74	-0.30	-0.16
38b	3	8.28	0.01	0.07	0.02	0.99	0.06	0.08
39a	3	9.67	0.02	-0.09	0.12	0.99	0.11	0.02
39b	3	10.43	0.03	0.20	0.05	0.96	-0.24	0.12
40a	3	9.53	0.33	-0.39	0.32	0.86	0.01	0.23
40b	3	8.59	0.17	-0.40	0.08	0.56	-0.02	0.18
41a	3	10.62	0.09	0.14	0.25	0.90	-0.29	0.20
41b	3	9.61	0.07	-0.20	0.25	0.99	0.01	0.04
42a	3	9.32	0.01	-0.04	0.08	0.98	0.04	0.05
42b	3	10.88	0.05	-0.09	0.19	0.94	-0.05	0.14
43a	3	9.29	0.02	-0.10	0.11	0.74	-0.31	-0.36
43b	3	10.71	0.00	NaN	NaN	0.96	-0.11	0.20
Mean	all	**9.82**	**0.07**	**-0.07**	**0.11**	**0.87**	**-0.02**	**0.10**
Median	all	**9.67**	**0.04**	**-0.07**	**0.09**	**0.95**	**-0.02**	**0.08**
Mean	1	9.25	0.06	-0.07	0.13	0.87	-0.01	0.13
Median	1	9.25	0.04	-0.08	0.10	0.95	-0.02	0.11
Mean	2	10.22	0.08	-0.02	0.07	0.85	-0.04	0.13
Median	2	10.19	0.02	-0.03	0.05	0.96	-0.08	0.10
Mean	3	10.03	0.07	-0.11	0.13	0.90	-0.01	0.05
Median	3	9.70	0.03	-0.09	0.11	0.95	0.01	0.05

Figure 10: Loss and gain acceptance data for Treatments 1-3

Zero intelligence

A striking aspect of the data are signs of prevalent zero intelligence play. More than half of the subjects (27 out of 52) accepted at least 95% of the utility-improving proposals offered to them and about 62% (32 of 52) accepted at least 93% of such proposals. Only two subjects accepted more than 10% of their utility-diminishing proposals, and only seven accepted more than 5% of such proposals. A reasonable hypothesis is that subjects who accepted at least 93% of utility-improving proposals behave as ZI traders with a bit of trembling. I will assume that there is no difference in the distribution of these subjects across treatments. This assumption is supported by the fact that the median probability of rejecting a utility-improving trade proposal is either 95 or 96% in all three treatments, and the median probability of accepting a utility-diminishing proposal is 2-4% in each treatment.

Let ΔV_i be the difference in expected value between accepting and rejecting the current proposal for subject i (time has been suppressed for notational convenience, since in this model I will assume that any dependency of the value function on time will be differenced out when comparing the expected value of accepting versus rejecting a proposal). If the subject follows a zero intelligence decision rule with a constant probability of making a mistake, this difference can be expressed as:

$$\Delta V_i = \beta_{0i} + \beta_{1i}\delta + \beta_{2i}\nu + \varepsilon_i.$$

Here, δ is an indicator variable equal to one if $u_i(z_{it}) - u_i(x_{it}) > 0$, zero otherwise, and the indicator ν is equal to one if $u_i(z_{it}) - u_i(x_{it}) = 0$, zero otherwise. I shall assume $\varepsilon \sim N(0,1)$; the assumption of unit rather than constant variance is without loss of generality since ε can only be estimated in ratio to the estimated coefficients. Subject i will accept the current proposal if and only if $\Delta V_i > 0$. While this variable is unobserved, it can be estimated using a simple probit model. The null hypotheses are $\beta_0 < 0$ and $\beta_1 > -\beta_0$. If the 'mistakes' are symmetric for gains and losses, we obtain the additional restriction $\beta_1 = -2\beta_0$.

It turns out that this model is too simple for most of the subjects. There are two primary deviations from it. Six of the 32 potential ZI subjects exhibit strong evidence of a regime change in their decision-making across periods. Much more prevalent still, most of the subjects tend to 'tremble' from ZI for small changes in utility more frequently than for large ones. Apparently, at least some of these deviations from ZI behavior are intentionally made.

What might motivate these conscious deviations from ZI? As mentioned previously, a subject might reject a small proposed increase in utility if he expects a better proposal might soon be generated and adopted. Conversely, accepting small utility improvements, particularly early in a period, can lock a subject onto a path to the least desirable region of the contract set. Actually, one of the biggest surprises in the data is that the rejection of small gains is not more pervasive.

The deliberate acceptance of utility-diminishing trade proposals is more difficult to explain. At least 40% of the group are significantly more likely to accept small losses than larger ones, which is reassuring (many others accepted so few losses that the trades could not be characterized as clustering significantly in the 'small loss' region, even though this is where their several losses occurred). One possibility is that subjects conjecture that accepting small losses will signal cooperativeness to the other trader, increasing the likelihood that he will reciprocate by being more cooperative herself in the future. Of course, preferences are private information in this experiment, so no one knows when anyone else is contemplating a gain or a loss, which weakens the signal considerably. However, a subject does signal general cooperativeness to some extent through his overall rate of proposal acceptance. A subject may also attempt to signal cooperativeness in a particular context, as well; for example, by accepting a loss after a string of failed proposals.

Given these concerns, a probit model was estimated for each of the 32 potentially near-ZI subjects that extends the model above by allowing estimated acceptance rates to vary discretely with changes in utility. A variety of specifications were fitted for each of these subjects. The variable $w_i = u_i(z_{it}) - u_i(x_{it})$ was also included in each specification, to test whether the covariation between proposal acceptance and change in utility was adequately controlled in the models by the indicator variables. The best fit for each subject[9] is reported in Figure 2.11, after dropping w_i (in all but one case its coefficient was insignificant at the 15% level, and in the remaining case the associated p-value was 0.075) and re-estimating the model. Change in utility appears on the axis in terms of cents per proposal, and estimated probabilities of proposal acceptance for given changes are reported, along with the 1-tail p-value from the associated t-test. For example, subject 6 (in Treatment 1) is estimated to be 15.8% likely to accept a loss between 1 and 3 cents, 1.1% likely to accept a loss greater than 4 cents, and 98.9% likely to accept any utility gain.

The estimated decision rules for these agents clustered into four types. Six subjects were nearly perfect ZI. They accepted far less than 1% of their utility-diminishing proposals, and rejected 1% or less of their utility-improving proposals. Because there is so little variation in the data for these subjects, no models can be appropriately fit for them. Three more subjects 'learned' ZI or near-ZI behavior after one to three periods of frequently violating this decision rule.

The other three ZI types involve a tendency to deviate from ZI behavior for small utility changes, gains and/or losses. One characteristic shared by all of these subjects is that a 'small' change in utility was no more than 10 cents; none showed evidence of weighting a 41 cent change in payoffs (the largest recorded proposed change in utility) more heavily than an 11 cent change. Some of these subjects exhibited evidence of a more restrictive definition of small; indeed, some treated 1 cent changes in payoffs differently from other changes, and showed no evidence of distinguishing between changes of greater magnitude. For all three types, no subject apparently required the two regions between $|1|$ and $|10|$ to be subdivided into more than three parts, and most could be subdivided into two. It may be the case that some bell-shaped function (decreasing in absolute value from 1) could better fit the data for some of these subjects rather than relying on discretized decision rules. If the function was flexible enough to vary between a step and a line with zero slope depending on the parameterization, a mixed effect probit model might be estimated for this entire subset of the data, to obtain a 'mean' near-ZI decision rule and an estimated distribution of behavior about this mean.

[9] Subjects in the role of Player 1 are denoted a, and b for Player 2.

Figure 2.11: Near-ZI Probit estimates of acceptance probabilities

Near (perfect) ZI

7. Accepted 2 of 364 losses and rejected 3 of 291 gains through ten periods.

11. Accepted 1 of 365 losses and rejected 1 of 388 gains through ten periods.

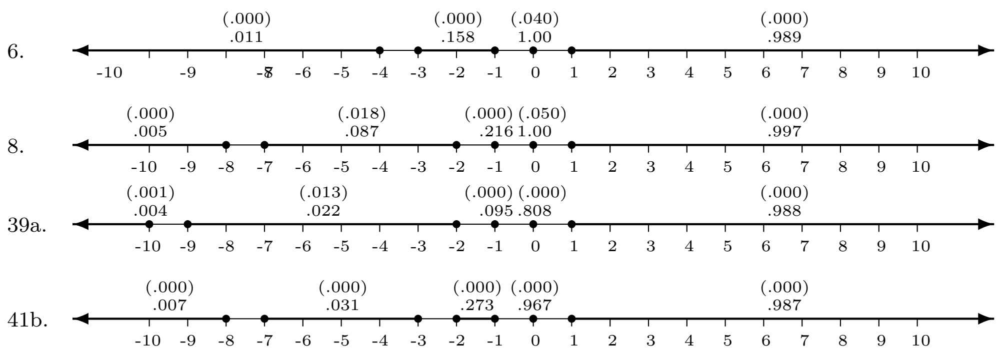
24b. Accepted 3 of 540 losses and rejected 2 of 528 gains through ten periods.

35b. Accepted 1 of 614 losses and rejected 4 of 532 gains through ten periods.

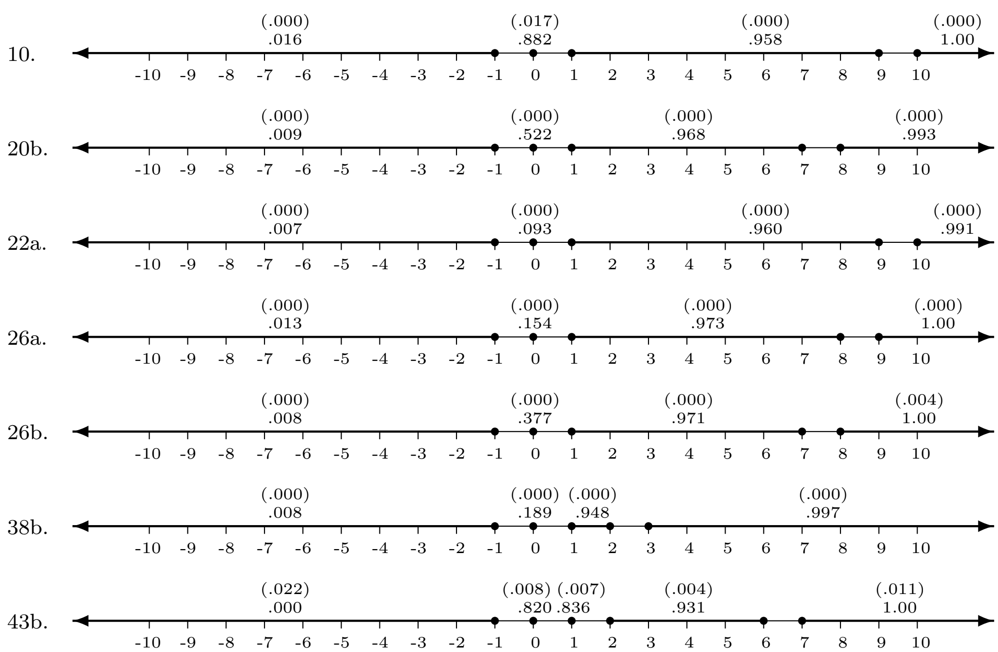
36a. Accepted no losses and rejected 2 of 549 gains through ten periods.

37a. Accepted no losses and rejected 1 of 460 gains through ten periods.

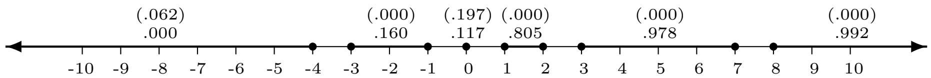
Near-ZI with conscious small loss acceptance

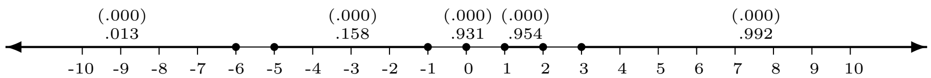


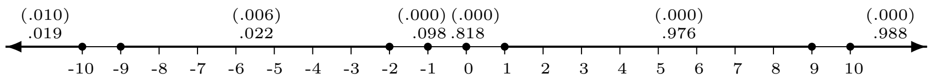
Near-ZI with conscious small gain rejection

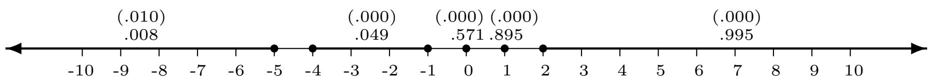


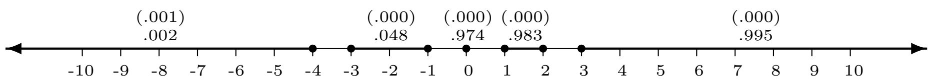
Near-ZI with conscious small utility change deviations in both directions

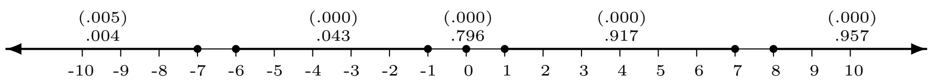
17. 

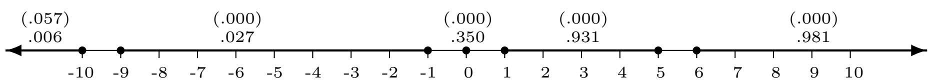
18. 

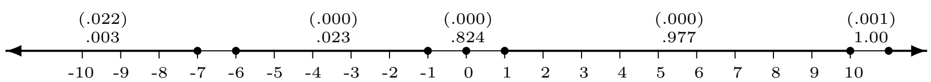
19a. 

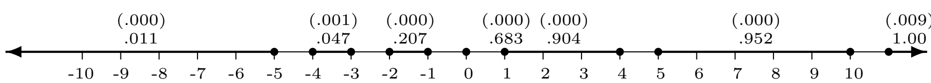
22b. 

24a. 

36b. 

39b. 

42a. 

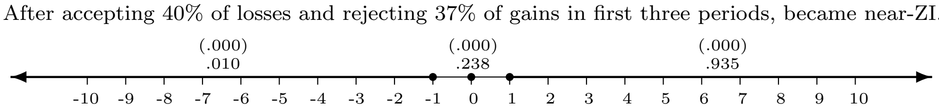
42b. 

Learn near-ZI

5. Perfect ZI in periods 2-10. Accepted 11 of 13 losses and rejected 1 of 20 gain in first period.

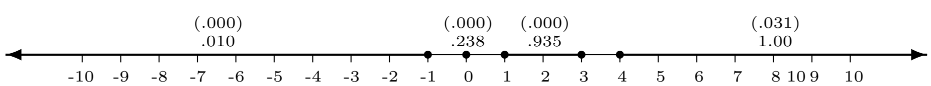
12. Perfect ZI in periods 2-10. Accepted no losses and rejected 6 of 40 gains in first period.

35a. After accepting 40% of losses and rejecting 37% of gains in first three periods, became near-ZI.

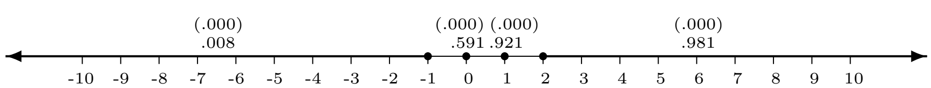


Learn near-ZI with conscious small gain rejection

2. After accepting 16% of losses and rejecting 13% of gains in first two periods, became near-ZI.



21b. After accepting 38% of losses and rejecting 15% of gains in first three periods, became near-ZI.



Four subjects show gradation in their willingness to accept losses, but do not exhibit signs of distinguishing between proposed gains. The models estimate that three of these subjects are quite willing to accept 'small' losses (they variously accept 1-3 cent losses 16-27% of the time), but are unwilling to accept larger ones more than 1% of the time. All four accept at least 99% of their utility-improving proposals.

Seven subjects rarely accept utility-diminishing proposals (fitted probabilities between 0 and .016), but demonstrate a willingness to reject small utility improvements more often than larger ones. The fitted models predict they will accept nearly any 'large' utility improvement (above 3 to 10 cents for the various subjects), but reject smaller ones 3-16% of the time. Importantly, none of these subjects exhibit a strong correlation between period and earnings or between period and the rate at which small utility improvements are rejected, which might indicate they should be considered candidates for a learning decision rule. Two other subjects adopted this decision rule after 2-3 periods of substantial deviations.

Finally, nine subjects are generally ZI but more likely to deviate when utility changes are small, in both directions. Some of these subjects exhibit very sharp distinctions in the predicted probabilities with which they will accept various proposals. For instance, subject 17 has a predicted probability of accepting a 1-3 cent loss of 16% and a predicted probability of accepting a loss greater than 3 cents of 0, while he is predicted to reject 20% of his 1-2 cent payoff increases, but 98% (or more) of the larger ones. For others, the distinctions between various levels of utility change are not so sharp, although statistically significant distinctions exist in all cases.

Of the 32 subjects who accepted at least 93% of their proposed utility improvements, 30 were able to be successfully classified as using one of these four near-ZI decision rules. One additional subject, number 35a, was also so-typed, as he 'learned' near-ZI behavior by the fourth period. The two subjects who were not successfully typed as near-ZI, numbers 1 and 9, will be discussed later. Therefore, 60% of the subjects in Treatments 1-3 can be fairly labeled "zero intelligence," if we extend the definition to include decision-makers who, on occasion, consciously reject small utility improvements or accept small utility losses. A full 25% of the subjects were nearly zero intelligence even without this small utility change buffer. This result was quite surprising.

The fact that these subjects are near-ZI indicates that salience was achieved and they understood the presentation of utility/payment information. That is, they were motivated to be fairly precise in the rules they followed (in general, at least 19 of 20 actions were ZI, and the remaining 1 action typically deviated in a fairly predictable manner). However, it is surprising that more sophisticated rules did not evolve over time for so many subjects. After all, these are 2-person barter economies where both agents have veto power over each proposed exchange; the entire individually rational Pareto set is potentially in-play. Gains from trade were typically exhausted in less than 3 minutes in each period for these subjects, so one might reasonably expect that over time they would have explored the extent of their negotiating power. Nevertheless, time and time again one can observe a subject who follows a near-ZI strategy and a sequence of good draws to a reasonably high payoff in one period(say, $1.10), who in the next period follows the same strategy and a

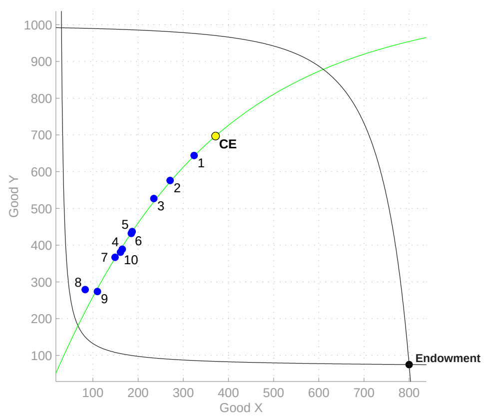


Figure 12: Session 19 (Treatment 2)

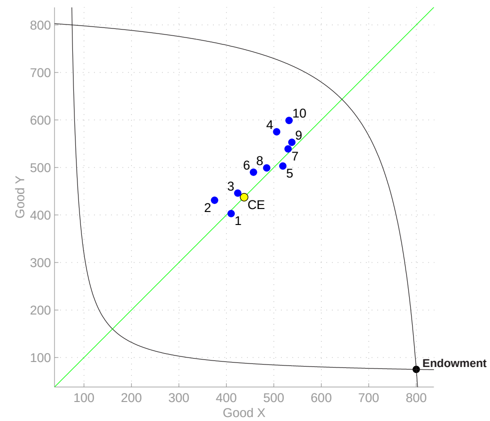


Figure 13: Session 38 (Treatment 3)

sequence of poor draws to \$0.70. If \$1.10 was available in the last period, why not reject an early 3 cent-improving proposal and attempt to adopt a trading path to get at least that much again?

It is interesting to note that near-ZI behavior is adopted in response to a wide range of trading partner strategies. In Treatment 1, the automated CSS trading partner is effectively ZI when the subject is near-ZI, since its subsidization constraints are quite infrequently binding (there is only a 5% chance of a ZI subject forcing the robot to subsidize in a given period). In Treatments 2 and 3, near-ZI behavior is robust to near-ZI trading partners, trading partners who exhibit very erratic behavior, and increasingly aggressive trading partners. For striking examples of the latter case, consider period-ending allocations in Sessions 19 and 38, displayed in Figures 12 and 13. Subjects 19a and 38b are near-ZI and are matched with trading partners who begin the first period following a near-ZI trading rule (in fact, neither rejects a utility-improving trade in the first period). However, subjects 19b and 38a become increasingly more demanding over time, while their trading partners maintain their near-ZI decision rules throughout. The result is the more aggressive subjects push the near-ZI subjects towards the boundary of their respective individually rational sets! It was expected that most subjects would exhibit behavior similar to 19b and 38a, and that subjects like 19a and 38b would 'fight back,' at least to some extent.

A 'rational' explanation for the prevalence of near-ZI behavior is that subjects believe their trading partners are very sensitive to signals of cooperation. Perhaps one rejected trade might push a partner into becoming uncooperative, possibly (in the most extreme case) not ever trading again. Of course, when faced with such a partner it may even make sense to accept losses now and again, but is such a strong prior sensible, particularly given that no one in the subject pool exhibits anything close to such behavior? This explanation might be more plausible if preferences were known, in which case the rejection of utility-improving proposals, even if marginal, might trigger uncooperativeness analogous to the well-documented fact in ultimatum games that small proposals are often refused at an explicit loss to the subject. It seems more likely to be the case that near-ZI subjects think they are performing

				Period Earnings (in cents)									
Subject	Treatment	Median ZI	CE	1	2	3	4	5	6	7	8	9	10
4	T1	88	124	91	89	68	109	88	94	110	93	130	123
13	T1	88	124	82	101	111	102	101	97	121	124	171	125
14	T1	88	124	69	103	116	82	116	113	54	106	103	119
15	T1	88	124	98	80	88	125	112	91	147	101	96	97
20a	T2P1	88	124	100	76	51	83	104	108	178	91	121	96
21a	T2P1	88	124	62	143	299	124	119	126	134	96	94	137
23a	T2P1	88	124	6	104	82	94	147	115	163	133	179	182
19b	T2P2	113	84	103	125	142	184	170	171	193	230	220	186
38a	T3P1	95	100	87	85	99	152	137	117	150	127	156	167
41a	T3P1	95	100	103	100	111	77	93	97	136	129	116	106
43a	T3P1	95	100	94	60	77	89	85	104	72	121	116	111

Figure 14: Payments for potential learners, Treatments 1-3

'well enough' and don't won't to jeopardize the distribution of their current payment stream by trying something more aggressive. They may alternatively (or in addition) think it is 'unfair' to reject utility-improving trade proposals if they think their partners are reciprocally fair, or they may simply not realize the opportunity to potentially increase their earnings by rejecting some proposed gains.

Subjects who learn

In an attempt to assess which subjects might be learning, it is instructive to consider payments in simulated ZI play. In Treatments 1 and 2, the median (mean) simulated ZI payment per period of Player 1 was \$0.88 (\$0.90), with a standard deviation of 19.3 cents. For Player 2, the median (mean) ZI payment was \$1.13, with a standard deviation of 17.4 cents. In Treatment 3, the median (mean) simulated ZI was \$0.95 (\$0.95) for Player 1 and \$1.04 (\$1.05) for Player 2, with standard deviations of 18.1 and 15.0 cents, respectively.

If a subject consistently secures greater than ZI payments, or does so with increasing frequency over time, it may be a signal that he is using a strategy more sophisticated than near-ZI (it may also be luck, or the fact that he is matched with a partner who accepts many utility-diminishing trade proposals). This is particularly true in Treatment 1, where the actions of the trading partner are CSS. Figure 14 presents payment data for eleven subjects who received significantly more than ZI payments, at least in several periods. Shaded in the figure are periods in which the player never rejected a utility-improving offer; with one exception, such activity is typically confined to the first couple of periods. Consider first the seven subjects who played the role of player one in Treatment 1 or 2. In the ZI simulations, only 5.1% of the time did player 1 receive at least 124 points, 6.6% of the time player 1 received at least 120 points, and 15.2% of the time player 1 received at least 110 points. Treating these statistics as coming from the true probability distribution can give us a rough estimate of the likelihood of seeing certain payment data from a ZI subject. For example, the probability of observing no payments in ten periods of at least competitive utility is equal to $0.949^{10} = 0.5925$. The probability of observing zero or one at-least-competitive payment is $0.5925 + 10 * 0.949^9 * 0.051 = 0.9109$. The probability of observing two competitive payments in ten periods is

$$(0.949)^8 (0.051)^2 \frac{10!}{8!2!} = 0.077,$$

so the probability of observing more than two competitive payments in ten periods is just over 1%. Therefore, one might strongly suspect more than ZI sophistication if three or more competitive payments are observed, and be quite suspicious in the event two are observed. Similarly, the probability of observing at least three payments greater than 120 points is 2.43%, and the probability of observing four or more payments of 110 points is 5.21%.

Subject number 13 has clearly learned to extract full gains from trade from his CSS counterpart. He earned at least 121 points in the final four periods, and at least 124 points in the final three. Also, he forced the robot onto a constraint just below competitive equilibrium (relative to the robot) in the final period, so he would have been unable to extract more than competitive utility from the robot in any future period. This economy is an example of learned competitive equilibrium. Subject number 4 may also have learned competitive equilibrium, as well, but the evidence is not quite as conclusive, since only the last two periods result in at-least-competitive payments. While the probability of any two periods bearing competitive payment to a ZI agent is about 8%, the probability of it happening in the last two periods is about a quarter of one-percent, so it seems rather likely that this is another example of learning competitive equilibrium.

Subject 23a, matched against a human trader, earned more than competitive payments in 5 of the last 6 periods. However, did so mostly through ZI play because of erratic behavior from his trading partner. In the first period, subject 23a accepted only 16% of the utility-improving proposals offered to him, and none of them resulted in a trade. This may have confounded his trading partner, to the point that the partner accepted about 50% of the utility-diminishing proposals made to him in the period, and continued this remarkable trend throughout the game. Subject 23a gradually began accepting more and more proposals over subsequent periods (and trade did take place), eventually adopting a pure ZI strategy to his great advantage.

Subject 21a reached at-least-competitive utility a remarkable 7 of 10 times. The first two times were in large part due to a great number of utility-diminishing proposals accepted by his trading partner. However, his partner only accepted one such proposal in the remaining seven periods, so it is clear 21a adopted a strategy that allowed him to extract large gains from trade, although they did not uniformly get larger over time.

The remaining three potential learners of the seven player 1s from Treatments 1 and 2 appear to earn significantly more than a ZI player (each earns less than median ZI utility in only two or three periods, and in all but one of these instances does a 'low' payment occur after the fourth period). However, the earnings are not sufficiently high or clustered enough in later periods to suggest obvious learning. These subjects reject utility-improving proposals more frequently than the 32 subjects that have been classified as ZI, and earn more money; beyond these facts, it would be difficult to say more without estimating a particular learning model, which is beyond the scope of this paper.

Subjects 41a and 43a from Treatment 3 each perform better in the last five periods than the first five, and only once does one of these subjects (43a) earn less than competitive utility in the final five periods, so some learning process appears to be present (these subjects turn out to be matched with near-ZI partners).

We have already seen subjects 19b and 38a briefly in Figures 12 and 13. Analogous to subject 13 (and likely 4), these subjects have learned to extract nearly the full gains from trade available given the strategy of their trading partners (in these two cases, both near-ZI). Furthermore, they are not CSS learners; from the figures it is clear that subsidization constraints are rarely, if ever, tested (or, in the case of 19b, never even defined).

Finally, in two sessions not listed in the table (25 and 40), the trading partners demonstrate strong responsiveness to each others' level of cooperation. One subject becomes particularly demanding in one period while the other accepts most proposals, then over the next several periods they will flip positions, with the formerly passive trader being much more aggressive while the formerly more aggressive trader is willing to accept most utility-improving proposals. These subjects often leave significant gains from trade on the table in their unwillingness to budge from a position even late in the period. The process of switching the aggressive/passive roles can cycle several times.

CSS learning is apparently not practiced by any of these "learners," not even in a subset of periods. Various CSS probit models were fit for all of the agents who developed subsidization constraints, and while some of the coefficients on the interaction term between the indicator for passing the subsidization constraint and the indicator for positive utility-improvement were positive and significant, none of them appear to be "driving" the data. So what other learning processes might be present?

Bearing in mind that some subjects try to achieve more than competitive payments (like 19b, 38a, and likely 13), some subjects apparently do "ratchet" expectations, even in circumstances where they are subsidy receivers. One particularly simple reinforcement learning model is to restrict oneself to accepting proposals above the net terms of trade established in the previous period. That is, at the end of each period, calculate $m = -(Y - \omega_y) / (X - \omega_x)$, where X and Y are the period-ending amounts of goods X and Y held by the agent in question. Then, in the next period, accept a trade proposal if and only if it lies above the line defined by this slope and the endowment, and it improves utility. To the extent that gains from trade are exhausted, this rule will procure higher utility in each period. This rule might be complemented by another that allows it to be violated when it becomes apparent that gains from trade might not be exhausted if the rule continues to be followed. That is, perhaps a certain amount of time has passed, or a certain number of consecutive proposals has been rejected.

I attempted to fit numerous probit models for subjects 19b and 38a using decision rules in this class. These subjects were prime candidates, since their payments were almost monotonically increasing over time. In the few periods that payments decreased, it is possible the consecutive rejection rule would explain the deviation from the terms of trade rule. It does not appear to be the case that the model fits the data very well. While the terms of trade line is statistically significant, its effect is not nearly large enough to account for the pronounced ratcheting that takes place.

Upon viewing animated graphics of decision-making in these economies, a stylized fact emerged in both cases: These subjects were much more demanding than the terms of trade rule early in the period. Then, once the terms of trade constraint was crossed, they relaxed their demands to simply stay above this line (unless time was running out, in which case the subjects occasionally go below the constraint). These subjects were thus taking a parabolic track to the Pareto set, and typically finished with net terms of trade better than the previous period. This makes a lot of sense, since their near-ZI trading partners are most receptive to really bad terms of trade early in the period. That is, preferences are such in these experiments that a ZI trader will accept nearly any trade proposal in the appropriate quadrant at the start of a period, but not so near the Pareto set.

Parabolic tracks to the Pareto set were attempted by other subjects, as well, although with less consistent success. In the fifth period, subject 43a (who had only rejected one utility-improving proposal in the first four periods combined) accepted his first utility-improving proposal, then rejected the next 25 until accepting one that netted him a 27-cent improvement, after which he accepted all others until the end of the period. In the remaining 5 periods, he demanded 35 or 36 cent-improvements on the first trade, rejecting all 10 and 24 cent proposed utility improvements before accepting the 35-36 cent gain. This seemed to be a very specific demand. Some of these gains were extremely good terms of trade (giving up hardly any of one good for a bunch of the other, which typically implies that great terms of trade will be in the individually rational set for the next set of proposals, as well), and others were just a little better than average. Some of the rejected proposals were large payment improvements (like 43 cents) and some were extremely good terms of trade, so it seems like both of these values (payoff and terms of trade) were important to the subject, as well as perhaps the desire to see many proposals in order to assess the data generating process. However, until the last two periods, this subject returned to near-ZI behavior once the first trade was secured, so trade tended to take a parabolic track to the Pareto set. In the last two periods, the subject went through a similar process of rejecting many proposals until a good first and second trade was adopted (high utility, good to excellent terms of trade), then returned to near-ZI behavior.

Most of the "learners" characterized in this section appear to develop strong demands with respect to the terms of trade and/or utility improvements for the first trade or two, and then follow-up with less-demanding behavior, so this appears to be a sketch of the learning model with the most promise. Any apparent significance of the CSS rule is likely coincident with such a rule.

Other subjects

Six subjects remain to be catalogued. One, subject 23b, apparently became confused after his trading partner rejected so many trade proposals that no trade took place in the first period. He accepted an average of 50% of the utility-diminishing proposals presented to him throughout the session, which translated into his earning a small amount for the session ($6.82). Two others in this group accepted far too many utility-diminishing proposals, as well (17% and 25%), but not to such an extent.

The other three subjects accepted 4-9% of such proposals, and were very selective in the utility-improving proposals they would accept (44-75%). However, even though they were matched against a CSS trading partner, their payments were no better than typical ZI payments, and showed no trends over time. It is unclear why they were being so selective, but it did not translate into greater payoffs.

Extension of the analytical framework

A richer econometric identification of decision rules could be implemented with the present data. The development of econometric models to identify heterogeneous learning in experiments has grown in recent years. For example, in normal and extensive form games there is adaptive experience-weighted attraction [Camerer, Ho and Chong 2002] and quantal response equilibrium [McKelvey and Palfrey 1995; 1998]. Unfortunately, these techniques are insufficiently flexible to accommodate the diverse range of behavior that has been observed in the present experiment. El-Gamal and Grether [1995] develop a technique which optimally assigns each subject to one of a (potentially large) set of decision rules.

Houser, Keane and McCabe [2004] endogenously identify the number and characteristics of decision rules in a class of stochastic investment experiments, using a Bayesian classification algorithm that is driven by the assumptions that all state variables are known and the decision rules are polynomial in these state variables (which are reasonable for their data). The fact that the state space is potentially vast in this experiment makes the application of Houser et al. daunting, as the exclusion of many potentially relevant state variables would almost necessarily be ad hoc, and the interpretation of estimated rules not at all intuitive. More relevant to this environment would be the application of an econometric model more akin to El-Gamal and Grether, although this would require the imposition of more *a priori* structure on the set of potentially relevant decision rules.

However, the development of such structure may not add value to what is already apparent: In an environment permitting substantial exercise of market power, satisficing is pervasive, although a significant minority appear to be considerably more sophisticated. It may not be all that interesting to know more precisely what subjects are doing, because the relevance of distinctions between them may not transfer to institutions we care about, like the posted price or double auction. The contribution of this paper is that heterogeneity in sophistication, bounded somewhere between zero intelligence and optimizing, can be of first order of importance in the process of market equilibration. It would be difficult to identify such heterogeneity in other institutions with much larger action spaces without knowing to look for it. Now we can anticipate such heterogeneity, which should assist in the identification of decision rules in more standard institutions, and also inspire new ways of thinking about the theory of market equilibration in general. Nearly all of the existing literature on out-of-equilibrium dynamics, game theoretic or general equilibrium, assumes homogeneity in sophistication; this experiment suggests that this may not be the right way to go (as do Houser, Keane and McCabe [2004] and El-Gamal and Grether [1995], for that matter).

Treatment 4

The chief difficulty of implementing this experiment was providing the appropriate tools to use subsidization constraints without coaching the subjects on how to use these tools, or that they should be using these tools at all. Apparently, not too much coaching was given, since none of the 52 subjects made obvious use of the subsidization constraints. After consideration of Treatment 4 data, it is obvious that these constraints are simply not natural in this environment.

Treatment 4 is parameterized identically to Treatment 1, including one subject assigned the role of player 1 and player 2 being played by a CSS robot. Several differences in the treatment are presented in the instructions. First, subjects are instructed that a period may end prior to the expiration of the period clock if "all traders agree on the relative values of X and Y." That is, if their trading partner has an X value (i.e., MRS) approximately equal to their own. It is explained to the subjects that no more significant gains from trade can be obtained in this circumstance, so it makes sense to let time expire.

This change alone should boost average payoffs from this treatment over Treatment 1 a little. A median 23 cents (mean 33 cents) per subject was lost in Treatment 1 *after* convergence to the Pareto set (defined by no 1-cent Pareto-improving increases available). Recall that gains from trade were typically exhausted about half-way through each 5-minute period, although many proposals were generated afterward. Occasionally, utility-diminishing proposals were accepted by the subjects that were acceptable to the CSS robot. Only two of the eighteen subjects lost more than 42 cents this way, although those two lost $1.84 and $1.33.

The second change in the instructions was to "coach" the subjects to use their subsidization constraints. This was done by informing them that "if the value of your current portfolio is less than the value of your endowment [these values are calculated and displayed using the period-ending MRS], it does indicate that your trading partner got a better deal than you in some sense," and that "your trading partner experienced a greater period-ending portfolio value than endowment value."

It was hoped that a much more sensible motivation for the appearance of marginal rates of substitution (that their equivalence would prematurely stop the period) would facilitate understanding the budget line/subsidization constraint concept, and that this understanding, along with the message that falling below this line implied your partner was getting a better deal than you, would motivate more use of CSS constraints. This did not take place.

The mean and median payments in Treatment 24 (23 subjects) were $9.01 and $9.06, respectively, a little lower than the corresponding statistics for Treatment 1 ($9.25 mean and median payment), even with the benefit of no losses after convergence to the Pareto set. More importantly, the five highest-paid subjects in Treatment 1, who were identified as potential learners, received at least $9.75; similarly, six subjects received at least this much money in Treatment 4. These learning candidates in Treatment 4 did earn 20 cents more on average than their Treatment 1 counterparts, but this can be more than accounted for by the new stopping rule (two learners in

Treatment 1 didn't lose any money after reaching the Pareto set, one lost 5 cents, one lost 27 cents, and one lost $1.33).

It is reasonably clear that more wide-scale learning did not take place than in Treatment 1. It is useful to note that the median payoff for a CSS robot after 10 periods of play was $11.46 in the simulation of Treatment 1. Only one subject reached a payment this high in Treatment 1 and in Treatment 4, and a second subject in Treatment 4 came somewhat close with a payment of $11.00.

Missing data from Treatments 1-4

There was a flaw in the software that occasionally caused an economy to crash mid-session. Below is a list of partially recorded data, which has not been included in the preceding analysis.

Treatment 1. One unreported subject received the highest performance-based payment in the treatment, $12.80. He earned $1.06 in the first period, so he may have been somewhat more selective than a zero intelligence player in the first period and quickly learned competitive equilibrium afterwards. This can be inferred from the fact that competitive play for 10 periods would have netted $12.40; his first period was $0.18 below this benchmark, but he still managed to make $12.80. The trading program's subsidization constraints would have made it nearly impossible to make $12.80 while earning significantly less than the competitive payment in more than one or two periods. I also recall observing this subject's screen several times and noting how closely the period-ending portfolios were clustered near the competitive equilibrium. Unfortunately, no decision data was recorded, and period-ending allocation information was lost.

Treatment 3. All decision data for one session was captured through 7 periods when the software crashed. Neither subject rejected any utility-improving proposals. One subject accepted one utility-diminishing proposal, while the other accepted nearly 10% of such proposals. These subjects were clearly near-ZI, although it is possible that one or both switched strategies in the last three periods.

Treatment 4. One economy crashed near the end of the sixth period, and the only data captured was period-ending utility and allocations. The subject's period-ending payments in periods 1-6 were 78, 85, 113, 165, 137, 124 points, respectively. The first 5 periods were Pareto optimal, while the last one was only about 80% of the way to the Pareto set when the server crashed. At this point the subject was hugging the robot's subsidization constraint, leading me to suspect that he would have received payment for the period between 124 and 137 points. He clearly had learned to extract full gains from trade from the CSS robot.

Therefore, one might safely conclude there were two strong learners and two near-ZI subjects in addition to the subjects previously characterized.

Conclusion

This experiment was motivated by a result establishing the existence of a boundedly rational learning strategy which, when shared by all agents in an exchange economy, is sufficient to implement competitive equilibrium. Crockett, Spear, and Sunder characterize this strategy as ε-intelligent, implying only a slight edge over their zero intelligence counterparts. The data suggests ε is, in fact, is quite large.

Subjects participated in a 2-person bargaining game. Neither participant was able to post terms of trade, but both had veto power over each element of a sequence of randomly generated proposed exchanges. Thus, subjects had significant market power, but could only exercise this power by rejecting trade proposals, not making them. This institution therefore severely restricted the space of actions while providing subjects with plenty of scope to exhibit the ability to extract increasing gains from trade over time, thus facilitating decision rule identification. A majority of subjects were apparently not concerned with extracting the best terms of trade (or maximizing, as it were), but instead were content to accept any 'reasonable' terms of trade. There are many possible explanations for this finding. Subjects may have been sensitive to the possibility that aggressive play would adversely affect the level of cooperation exhibited by their trading partner. They may also (or in addition) have considered it 'unfair' to skim only the particularly good trade proposals. Or perhaps they are simply 'happy enough' with ZI payments.

The prevalence of satisficing behavior may have important implications if it is robust to alternative institutional specifications. In institutions like the double auction, where agents have the opportunity to request terms of trade or adopt terms posted by others, an interesting conjecture is that satisficers may tend to focus on the submission of utility-improving market orders rather than formulate limit order strategies. If true, this would provide incentive for more sophisticated agents to focus on the limit order book. Some finance researchers have wondered if there could be a bias in the composition of who submits market versus limit orders; here is a hypothesis that can be tested in the laboratory. Another question that has been asked in experimental double auction markets is why so many subjects tend to act quickly rather than wait for more information to become available. The presence of a substantial number of satisficers may provide an answer: The presence of satisficers should influence the decision of more sophisticated traders to enter the market early, in order to snatch up the satisficers before the other sophisticated traders do.

However, the existence of a more sophisticated learning group, about 30% of the subjects, is evidence that competitive equilibrium can possibly be learned in other institutional settings. While the particular market institution adopted in this experiment permits few signals between participants, in others signals abound. It is possible that in institutions like the double auction, satisficers can coordinate on values of decision variables, like prices, that are driven by competing learners in the direction of competitive equilibrium. If learners on opposite sides of the market can push current prices in the direction of competitive equilibrium (a leap of faith at this point), then the satisficers may reinforce this momentum. In fact, the presence of a substantial subset of satisficers may assist, rather than hinder, price equilibration. This idea will be the topic of further research.

References

Aumann, R. (1964). "Markets with a continuum of traders." *Econometrica*, **32**, 39–50.

Camerer, C., T. Ho and J. Chong (2002). "Sophisticated experience-weighted attraction learning and strategic teaching in repeated games." *Journal of Economic Theory*, **104**, 137–188.

Camerer, C. and K. Weigelt (1993). "Convergence in experimental double auctions for stochastically lived assets." In D. Friedman and J. Rust, eds., *The Double Auction Market: Institutions, Theories, and Evidence*. Addison-Wesley Publishing Company.

Chatterji, S. and S. Ghosal (2004). "Local coordination in a competitive market." *Journal of Economic Theory*, **114**(2), 255–279.

Crockett, S., V. Smith and B. Wilson (2005a). "Spontaneous specialization and exchange." Working Paper.

Crockett, S., S. Spear and S. Sunder (2005b). "Learning competitive equilibrium: the negishi algorithm." Working Paper.

El-Gamal, M. and D. Grether (1995). "Are people bayesian? uncovering behavioral strategies." *Journal of the American Statistical Association*, **90**, 1137–1145.

Feldman, A. (1974). "Recontracting stability." *Econometrica*, **42**, 35–44.

Fudenberg, D. and D. Levine (1998). *The Theory of Learning in Games*. MIT Press.

Ghosal, S. and M. Morelli (2004). "Retrading in market games." *Journal of Economic Theory*, **115**(1), 151–181.

Gjerstad, S. (2004). "Market dynamics in edgeworth exchange." Working paper.

Gode, D., S. Spear and S. Sunder (2000). "Convergence of double auctions to contract curve in an edgeworth box." GSIA Working Paper #2001-E28.

Gode, D. and S. Sunder (1993). "Allocative efficiency of markets with zero intelligence traders: Market as a partial substitute for individual rationality." *Journal of Political Economy*, **101**(1), 119–137.

Green, J. (1974). "The stability of edeworth's recontracting process." *Econometrica*, **42**, 21–34.

Houser, D., M. Keane and K. McCabe (2004). "Behavior in a dynamic decision problem: An analysis of experimental evidence using a bayesian type classification algorithm." *Econometrica*, **72**(3), 781–822.

Kumar, A. and M. Shubik (2004). "Variations on the theme of scarf's counter-example." *Computational Economics*, **24**(1), 1–19.

McKelvey, R. and T. Palfrey (1995). "Quantal response equilibria for normal-form games." *Games and Economic Behavior*, **7**, 6–38.

McKelvey, R. and T. Palfrey (1998). "Quantal response equilibria for extensive-form games." *Experimental Economics*, **1**, 9–41.

Noussair, C., C. Plott and R. Reizman (1995). "An experimental investigation of the patterns of international trade." *American Economic Review*, pp. 462–491.

Saari, D. and C. Simon (1978). "Effective price mechanisms." *Econometrica*, **45**, 1097–1125.

Scarf, H. (1960). "Some examples of global instability of the competitive equilibrium." *International Economic Review*, **1**, 157–172.

Serrano, R. and O. Volij (2002). "Mistakes in cooperation: The stochastic stability of edgeworth's recontracting." Institute for Advanced Studies Working Paper 29.

Shapley, S. and M. Shubik (1977). "Trading using one commodity as a means of payment." *Journal of Political Economy*, **84**, 937–968.

Temzelides, T. (2002). "A note on learning to bid walrasian in 2×2 market games." Working paper.