Market design and the stability of general equilibrium[☆]

Jacob K. Goeree[a], Luke Lindsay[b,*]

[a] *University of Technology Sydney, Economics Discipline Group, PO Box 123, Broadway NSW 2007 Sydney, Australia; and International Faculty, University of Cologne, Albertus-Magnus-Platz, D-50923 Cologne, Germany*

[b] *University of Exeter Business School, Streatham Court, Rennes Drive, Exeter, EX4 4PU, UK*

Abstract

We employ laboratory methods to study the stability of competitive equilibrium in Scarf's economy (*International Economic Review*, 1960). Tatonnement theory predicts that prices are globally unstable for this economy, i.e. unless prices start at the competitive equilibrium they oscillate without converging. Anderson et al. (*Journal of Economic Theory*, 2004) report that in laboratory double auction markets, prices in the Scarf economy do indeed oscillate with no clear sign of convergence. We replicate their experiments and confirm that tatonnement theory predicts the direction of price changes remarkably well. Prices are globally unstable with adverse effects for the economy's efficiency and the equitable distribution of the gains from trade.

We also introduce a novel market mechanism where participants submit demand schedules and prices are computed using Smale's global Newtonian dynamic (*American Economic Review*, 1976). If the submitted schedules are competitive - sets of quantities that maximize utility taking prices as given - the resulting outcome is the unique competitive equilibrium of Scarf's economy. In experiments using the schedule market, prices converge quickly to the competitive equilibrium. Besides stabilizing prices, the schedule market is more efficient and results in highly egalitarian outcomes.

JEL classification: C92; D50
Keywords: Scarf economy; tatonnement; global Newtonian dynamic; instability; general equilibrium; market design

1. Introduction

This paper investigates how the design of the market influences price dynamics and trading volumes in Scarf's (1960) economy. In Scarf's economy, the tatonnement model

[☆]We would like to thank the Swiss National Science Foundation (SNSF 138162) and the European Research Council (ERC Advanced Investigator Grant, ESEI-249433) for financial support. We are grateful for useful suggestions we received from seminar participants at University of Nottingham and the ESA meeting in New York.

[*]Corresponding author
Email addresses: jacob.goeree@gmail.com (Jacob K. Goeree), luke.lindsay@gmail.com (Luke Lindsay)

April 4, 2016

predicts that prices cycle along a closed orbit around the equilibrium without ever converging. In a series of fascinating experiments, Anderson et al. (2004) implemented a version of Scarf's economy in the laboratory to study how prices evolve in the commonly used double auction market. While the double auction is itself a distinctively non-tatonnement institution, Anderson et al. found strong support for the Walrasian tatonnement hypothesis that price dynamics are largely driven by a market's excess demand.[1] Average trade prices in the experiments cycled along a closed orbit around the unique competitive equilibrium with no clear sign of convergence.[2]

A consequence of out of equilibrium price cycling is that an efficient allocation of resources may never be realized. This motivates our market design question: is there a market mechanism that stabilizes prices in Scarf's economy and leads to higher welfare? The main idea behind our proposed solution is to exploit the price-taking behavior that causes instability in the double auction market, as observed by Anderson et al. (2004). Such price-taking behavior has also been observed in other experimental studies, e.g. Friedman and Ostroy (1995).[3] The proposed mechanism is a call market where agents submit demand schedules, which are aggregated to yield an excess demand function.[4] A Newtonian process suggested by Smale (1976b) is then used to find market clearing prices. Whether this schedule market produces desirable outcomes obviously depends on the types of schedules that get submitted. But if every agent submits a competitive schedule, i.e. a set of quantities that are utility maximizing taking prices as given, then the mechanism produces prices and quantities corresponding to the unique competitive equilibrium of the Scarf economy.

We ran two series of experiments. The first series was devoted to replicating Anderson et al.'s (2004) experiments. One of the major strengths of laboratory experimentation for investigating general equilibrium is control, as the Anderson et al. study exemplifies. By inducing carefully selected demand parameters and initial endowments, the experimenters were able to create a version of Scarf's economy in the laboratory to study its equilibration properties. Another major strength of experimentation is replicability and the importance of Anderson et al's findings motivated our study. Furthermore, we extend the theoretical analysis of Anderson et al. by developing a model of out of equilibrium trading in the double auction, which allows us to model not just prices but also

[1] We use *tatonnement model* and *Walrasian hypothesis* interchangeably to refer to a model that predicts prices adjust in proportion to excess demand. We use *tatonnement institution* to refer to a market institution where prices are centrally adjusted according to excess demand and trade only occurs at equilibrium prices. There have been several other experimental tests of the Walrasian hypothesis. Smith (1962) finds some support for it although the "excess rent" hypothesis he introduces does better. Crockett et al. (2011) find support for the Walrasian hypothesis in an experimental study of Gale's (1963) economy.

[2] While the Scarf economy is an idealized example whose conditions are unlikely to be met in practice, this type of disequilibrium behavior is akin to price cycles observed in some important commodity markets, see for example Cashin and McDermott (2002).

[3] A large market is not a necessary condition for price-taking behavior to be optimal, see e.g. Ostroy (1980).

[4] Submitting demand schedules is a common feature of electricity markets, IPOs, and treasury auctions. Furthermore, this procedure is used prior to the start of the New York Stock Exchange to provide the opening prices for the day. Schedule markets are understudied compared to the double auction market, but an early laboratory test is reported by Smith et al. (1982) who consider a single-commodity market for which stability is not an issue. They find that a schedule market produces efficiency levels similar to those observed in the double auction market.

quantities traded and hence welfare. It is based on Hahn and Negishi's (1962) model of price adjustment with centralized price setting but trading at non-equilibrium prices. A set of prices is called and then trade occurs such that if there is excess demand for a certain good before trading, after trading no one is left holding more of that good than they demand (and vice-versa if there is excess supply). After trading, prices are adjusted according to excess demand.[5] We show that the model predicts behavior in the double auction market experiments remarkably well.

The second series of experiments tested our newly designed schedule market in the same controlled environment. This "engineering" approach, which combines institutional design with laboratory "wind tunnel" testing, has, to the best of our knowledge, not previously been applied to enhance the stability of several interconnected markets. Interestingly, price taking behavior, submitting a competitive schedule, is actually a weakly dominant strategy in the Scarf economy. Furthermore, when others submit competitive schedules, the supply curve that each agent faces is "flat." This non-generic feature follows from the specific parametrization of the Scarf economy.[6]

It is natural to wonder if experimental testing is even necessary, given that submitting a competitive schedule is a weakly dominant strategy. First, the experiments let us test whether the mechanism works when schedules are generated by eliciting quantities with a relatively coarse price grid and are then interpolated to create a continuous function. More importantly, however, there are a number of reasons to be cautious about assuming people will necessarily play weakly dominant strategies. It is not clear that subjects in the experiments will be able to identify the weakly dominant strategy by reasoning, much like they are not able to find the competitive equilibrium by reasoning in the double auction market. Furthermore, previous studies have found that people do not always play weakly dominant strategies. In individual decision making tasks where the weakly dominant strategy may appear obvious, subjects often do not behave optimally (Cason and Plott, 2014). In second-price auctions, subjects typically overbid relative to their values. In call markets, subjects are often insensitive to whether sincere bidding is a weakly dominant strategy and tend to initially bid further from their true value than optimal (Cason and Friedman, 1997). Instead of assuming that weakly dominant strategies are necessarily played, we consider the assumption that subjects use a myopic strategy of submitting a schedule that is a best response to the previously observed price but more elastic or more inelastic than the competitive schedule. We prove that under

[5] There are some other notable models with trading at non-equilibrium prices. Keisler (1995; 1996) introduces a model with decentralized price setting. There is a single market maker who holds an inventory and sets prices. Agents are randomly selected to trade with the market maker. The market maker adjusts prices in such a way that his inventory remains approximately constant over time. Crockett et al. (2008) augment the zero-intelligence trading model with a learning rule that directs convergence to competitive equilibrium. They consider an infinite-horizon model where, in each period, out-of-equilibrium trade yields an allocation in the contract set. The utility gradient at that allocation is then interpreted as a price vector that is used to redistribute wealth to generate a new starting allocation for the next period. The model is a globally stable alternative to Walras' tatonnement. See also Crockett (2008) for an experimental test of the model.

[6] Indeed, there is a large literature on supply function equilibria that studies oligopolistic markets where firms choose supply schedules and do not necessarily face a flat residual demand curve, see for instance, Klemperer and Meyer (1989). Importantly, however, in large economies, the supply curve faced by each agent is approximately flat for arbitrary specifications of preferences and endowments. As a result, submitting competitive schedules is optimal more generally when the economy grows large.

stationary repetition, prices gradually converge to the competitive equilibrium.

1.1. Organization

The next section briefly reviews the tatonnement adjustment process and shows that it is unstable in Scarf's (1960) economy. Section 3 describes the issues in designing a stable market mechanism and puts forth a specific proposal. Section 4 describes the design of an experiment that compares this novel mechanism to the standard double auction market. The results are reported in Section 5. Section 6 concludes. Appendices A and B contain additional theoretical results and proofs. Appendix C provides a detailed comparison with the Anderson et al. (2004) study, which employed only the continuous double auction market.

2. Background

2.1. Walrasian dynamics in Scarf's economy

Scarf (1960) proposes a simple economy with three goods, call them apples (A), bananas (B), and coconuts (C), and three types of agents whose preferences and endowments can be summarized as follows:

	type A	type B	type C
utility	$\min(q_A, q_B)$	$\min(q_B, q_C)$	$\min(q_C, q_A)$
endowment	1 Apple	1 Banana	1 Coconut

Consider a type A agent who is endowed with one apple and has utility function $\min(q_A, q_B)$. For given prices p_A and p_B this agent's demands for apples and bananas are $q_A = q_B = \frac{p_A}{p_A+p_B}$. Notice that there are income effects, i.e. agent A's demands for both apples and bananas rise (fall) when the price of apples (bananas) goes up. The demands for type B and C agents can be derived similarly, and it is readily verified that the equilibrium prices for which demand equals supply satisfy $p_A = p_B = p_C$. Without loss of generality we can single out coconuts to be the numeraire good and fix its price to $p_C = 1$. Then the competitive equilibrium price of each of the goods is one.

How does the economy arrive at competitive equilibrium prices? Consider any set of prices p_A and p_B for apples and bananas respectively expressed in terms of the numeraire $p_C = 1$. In Walras' tatonnement process, the change of price of each good is equal to its excess demand. In vector notation, $dp(t)/dt = z(p)$, or written out in components

$$\begin{aligned} \frac{dp_A}{dt} &= n\left(\frac{1}{1+p_A} - \frac{p_B}{p_A+p_B}\right) \\ \frac{dp_B}{dt} &= n\left(\frac{p_A}{p_A+p_B} - \frac{1}{1+p_B}\right) \end{aligned}$$

where $n \geq 1$ is the number of replicas of each type of agent in the economy. In the first line, the first term between parentheses on the right-hand side is the demand for apples by type C and the second term is the net supply of apples by type A. Likewise, in the second line, the first term between parentheses represents the demand for bananas by agent A and the second term is the net supply of bananas by type B. The price of coconuts is fixed at 1 so there is no price adjustment equation for p_C.

Proposition 1. *In a Scarf economy with $n \geq 1$ agents of each type, the tatonnement process is globally unstable.*

Proof. Consider the Lyapunov function

$$L(p_A, p_B) \;=\; 1 - p_A p_B \exp\bigl(1 - \tfrac{1}{2}p_A^2 - \tfrac{1}{2}p_B^2\bigr)$$

It is readily verified that $0 \leq L \leq 1$ with $L = 0$ if and only if $p_A = p_B = 1$. Moreover, using the tatonnement equations of motion we have

$$\frac{d\log(1-L)}{dt} \;=\; \frac{dp_A}{dt}\Bigl(\frac{1}{p_A} - p_A\Bigr) + \frac{dp_B}{dt}\Bigl(\frac{1}{p_B} - p_B\Bigr) \;=\; 0$$

In other words, the Lyapunov function is constant over time. The combination of prices that yield the same Lyapunov value form closed orbits in (p_A, p_B)-space, see the left panel of Figure 1. So if the process starts with a function value $L \neq 0$, then the prices cannot converge to the competitive equilibrium where the Lyapunov function takes the value 0. Instead, prices cycle in a counter-clockwise manner along the orbit indexed by the value of the Lyapunov function at time zero. □

2.2. Newtonian dynamics in Scarf's economy

Smale (1976a, 1976b) proposes to replace the Walrasian tatonnement process, $dp(t)/dt = z(p)$, by the Newtonian dynamic:[7,8]

$$\frac{dp}{dt} \;=\; -\bigl(\nabla z(p)\bigr)^{-1} z(p)$$

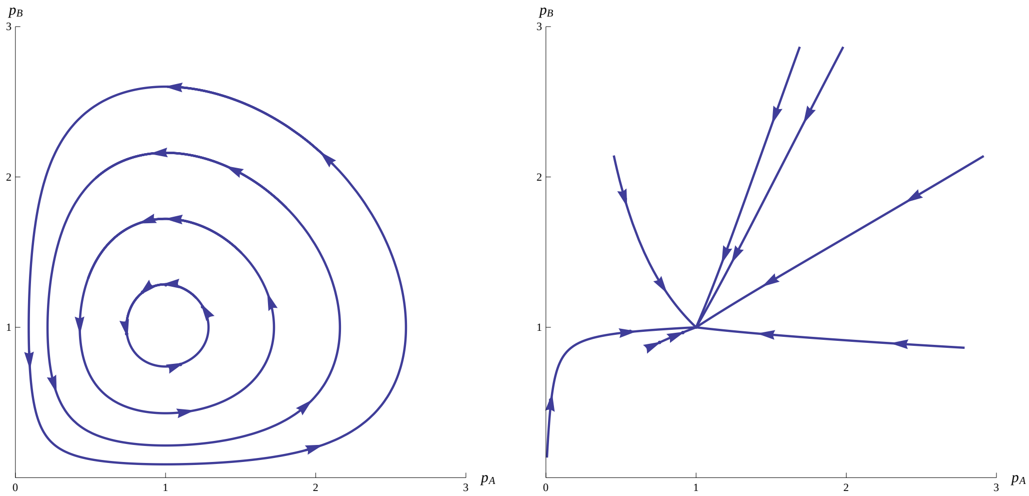
with $\nabla z(p)$ the matrix of partial derivatives of $z(p)$ with respect to p.[9]

Proposition 2. *In a Scarf economy with $n \geq 1$ agents of each type, the Newtonian dynamic is globally stable.*

[7] Newton's method for solving $f(x) = 0$ for some $f : \mathbb{R} \to \mathbb{R}$ can be recovered by taking a discrete approximation: $x_{n+1} - x_n = -f(x_n)/f'(x_n)$.

[8] In writing down the global Newton dynamic we assumed that $\nabla z(p)$ is everywhere non-singular, which is true for the Scarf example. Smale (1976b) discusses a more general form of the Newtonian dynamic that applies also when $\nabla z(p)$ is singular.

[9] Written out in components, $(\nabla z(p))_{ij} = \partial z_i(p)/\partial p_j$.



(a) Tatonnement (b) Newton

Figure 1. Predicted price patterns under the tatonnement dynamic (a) and the global Newtonian dynamic (b) in the Scarf economy. For the tatonnement model, prices cycle in a counter-clockwise manner without converging. In contrast, for the Newtonian dynamic, prices converge exponentially fast to the unique equilibrium $(p_A, p_B) = (1, 1)$.

Proof. Consider the Lyapunov function

$$L(p_A, p_B) = \Big(\frac{1}{1+p_A} - \frac{1}{1+p_A/p_B}\Big)^2 + \Big(\frac{1}{1+p_B} - \frac{1}{1+p_B/p_A}\Big)^2$$

Note that $0 \leq L \leq 1$ and $L = 0$ if and only if $p_A = p_B = 1$. The Newtonian laws of motion for the Scarf economy

$$\begin{aligned} \frac{dp_A}{dt} &= \frac{p_A(1-p_A^2)(1+p_B^2)}{(p_A+p_B)(1-p_A)(1-p_B)+4p_Ap_B} \\ \frac{dp_B}{dt} &= \frac{p_B(1-p_B^2)(1+p_A^2)}{(p_A+p_B)(1-p_A)(1-p_B)+4p_Ap_B} \end{aligned}$$

can be used to verify that

$$\frac{d\log(L)}{dt} = -2$$

Hence, the Lyapunov function decreases exponentially over time to its limit value of zero, corresponding to the competitive equilibrium (see the right panel of Figure 1). $\square$

Remark 1. While Proposition 2 is limited to the Scarf economy, a similar argument applies to more general economies. Define $L = ||z(p)||^2$ then under the Newtonian dynamic $dL/dt = -2L$, i.e. the Lyapunov function is exponentially decreasing. As noted by Smale (1976b,a) this observation can be used to prove existence of a competitive equilibrium for general environments without having to resort to methods of algebraic topology.

3. Designing a new market mechanism

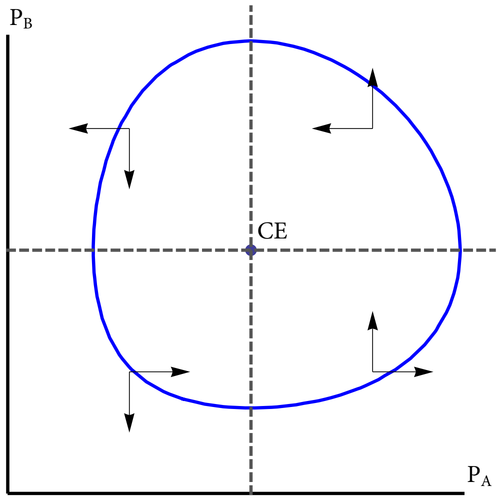
Before we turn to the question of how the Newtonian dynamic can be implemented to stabilize Scarf's economy, it is worth briefly discussing why some alternative mechanisms do not work. The continuous double auction market has typically been used as the standard against which other mechanisms are compared. This is partly because of its practical relevance, i.e. most contemporary financial and commodity markets are run this way, and partly because of its ability to generate competitive equilibrium outcomes in single-commodity markets.[10] The experiments of Anderson et al. (2004), however, demonstrate that for the multi-market Scarf economy the double auction market does not lead to convergence.

The double auction market is a non-tatonnement institution where trade can occur at prices that do not clear the market. Other mechanisms often use some kind of iterative procedure to find prices such that demand equals supply. For example, many valuable public assets (e.g. spectrum that can be used for telecommunication services) are nowadays sold in some type of ascending English auction. This is a tatonnement-like mechanism in that prices increment upwards until there is a unique winner for each item (demand equals supply) at which point the items are assigned (trade occurs). Of course, prices could also start high and decrement downwards as in the multi-unit Dutch auction used to sell flowers in the Netherlands. One could imagine a combination of ascending and descending prices. As we explain next, however, none of these tatonnement-like mechanisms can be expected to stabilize Scarf's economy.

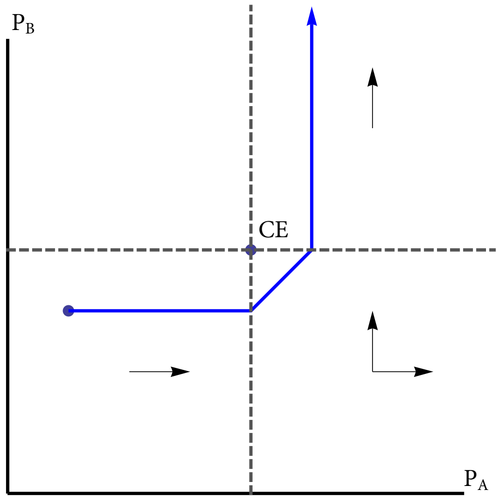
First, consider the tatonnement institution where a Walrasian auctioneer announces a set of prices and participants truthfully announce their demands at these prices. If there is no excess demand, the participants trade and the process terminates. If there is excess demand, the auctioneer adjusts the price of each good in proportion to the excess demand for the good and the process continues. For the reasons described in Section 2.1, prices will not converge to the unique competitive equilibrium of the Scarf economy but instead will cycle as shown in Figure 2a.[11] Now consider the price adjustment rule where the price of a good is increased by a fixed increment if and only if excess demand for the good is strictly positive. In such an institution, prices will not converge to the competitive equilibrium as illustrated in Figure 2b. Similarly, when the price of a good is decreased by a fixed amount if and only if excess demand for the good is strictly negative, prices will not converge to the competitive equilibrium as illustrated in Figure 2c. Finally, for the price adjustment rule where the price of a good is increased by a fixed increment if excess demand for the good is strictly positive and decreased by a fixed

[10] Convergence in single-commodity markets occurs under a wide variety of conditions, see e.g. Smith (1962), Friedman (1984), and Smith (2010).

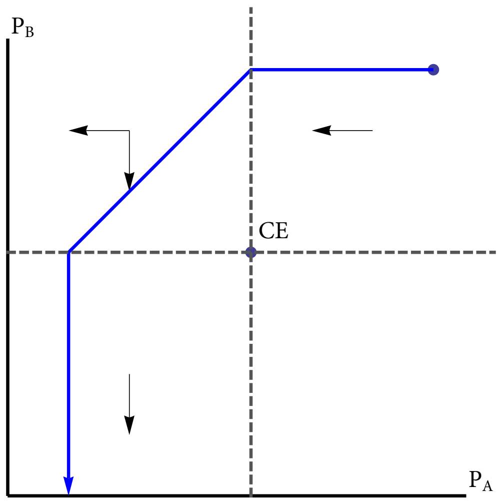
[11] See Smith et al. (1982) and Plott and George (1992) for experimental evidence on the Walrasian auctioneer mechanism.



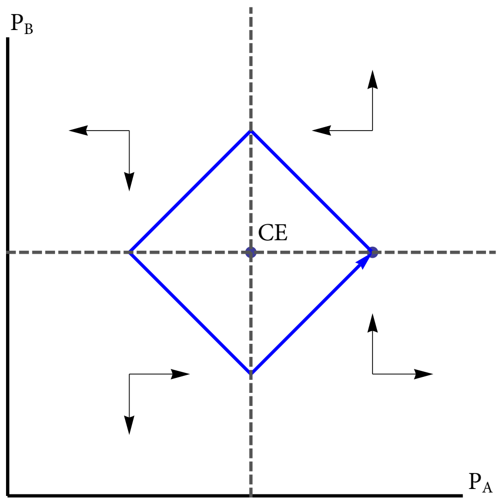
(a) Tatonnement



(b) Ascending price



(c) Descending price



(d) Ascending and descending price

Figure 2. Predicted price dynamics in the Scarf economy under various tatonnement-like institutions: (a) the Walrasian auctioneer, (b) the ascending English auction, (c) the descending Dutch auction, and (d) a combination of the Dutch and English auction. The prices of A and B are plotted in terms of the numeraire $P_C = 1$. The competitive equilibrium is labelled 'CE'. Each plot is divided into four quadrants. The indicator arrows show the effect of excess demand on price in each quadrant.

amount if excess demand for the good is strictly negative, prices will also not converge as illustrated in Figure 2d.

To summarize, commonly used market institutions do not guarantee convergence in Scarf's economy. Our novel design is motivated by the work of the mathematician Stephen Smale on "global Newton" methods. Smale (1976a; 1976b) proposes an alternative to the tatonnement dynamic that is convergent under general conditions, including those defined by the Scarf economy. The main question is how to design a market where prices converge.[12] One could introduce a "Newtonian auctioneer" who (i) announces prices, (ii) elicits agents' demands, and (iii) adjusts prices given reported excess demands according to the global Newton method.[13] One complication is that the Newton method requires information not only about excess demands but also about their derivatives. In addition, an iterative procedure is potentially time consuming and strategically complex. We solve both issues by letting agents submit demand schedules, i.e. a list of quantities demanded at various prices, and then determine the terms of trade by running an automated version of the iterative process (i)–(iii). There are a range of optimization procedures, including the Newtonian dynamic, that could be used to find market clearing prices providing they have the required stability properties.[14] Because submitting entire schedules is more complex and more time consuming than submitting single orders, we consider a call market that is cleared at prespecified times rather than continuously.[15]

3.1. The schedule market

In general terms, the schedule market works as follows. Each participant i reports their excess demand as a function of price $z_i(p)$. Reported demands are aggregated to produce an excess demand function $z(p) = \Sigma_i z_i(p)$. Prices p^* are computed such that supply equals demand for all goods.[16]

To ensure uniqueness of the market clearing prices, some restrictions are placed on the admissible demand schedules. Recall that in Scarf's economy each type of agent derives utility from two of the three goods and is endowed with one of the goods they like. As shown in Table 1, if a type needs good X and has good Y, they submit a schedule specifying the quantity of X demanded at various prices of X relative to Y. Let the demand schedule be denoted by $D_{XY} : \mathbb{R}^+ \mapsto \mathbb{R}^+$, which is a mapping from strictly

[12]The global instability observed in the experiments conducted by Anderson et al. (2004) indicates that the Newtonian dynamic is not at play in the double auction market institution. In a related set of double auction market experiments, Hirota et al. 2005 report that they find no support for the Newtonian model.

[13]Newton's classical method of iteration corresponds to a discrete approximation to Smale's adjustment process.

[14]Notice that a Walraisan Clearing House (Friedman and Rust, 1993, p. 9) where traders submit demand schedules but a tatonnement dynamic is applied to the resultant aggregate demand function would not help. Eliciting demand using schedules rather than iteratively does not solve the instability problem.

[15]McCabe et al. (1990) found that call market based institutions can be highly efficient. Friedman (1993) compares the continuous double auction to a call market institution and finds they produce similar efficiency levels.

[16]In the experiment, participants only report demand at a finite set of prices and interpolation is used to estimate demand at intermediate prices. The interpolation method is described in Section 4.2.

positive prices to strictly positive quantities. The following restrictions are applied to the submitted demand schedules.[17]

(D) If $p_1 < p_2$ then $D_{XY}(p_1) \geq D_{XY}(p_2)$.

(I) If $p_1 < p_2$ then $p_1 \, D_{XY}(p_1) < p_2 \, D_{XY}(p_2)$.

The first restriction states that demand schedules are non-increasing. The second restriction states that demand is inelastic, i.e. the amount spent on a good rises with its price. We say a schedule is *admissible* if it satisfies properties (D) and (I). It is readily verified that admissibility is preserved under aggregation. In particular, let $\mathcal{A}$ be any set of admissible demand schedules and let $\mathcal{D}_{\mathcal{A}} : \mathbb{R}^+ \mapsto \mathbb{R}^+$ denote the aggregate demand schedule $\mathcal{D}_{\mathcal{A}}(p) = \sum_{D \in \mathcal{A}} D(p)$. Then $\mathcal{D}_{\mathcal{A}}$ satisfies (D) and (I).

Define $S_{YX}(q) = q D_{XY}^{-1}(q)$ where D_{XY} is admissible. One can think of $S_{YX}(q)$ as the amount of Y being supplied when q number of units of X are demanded. Since D_{XY} is admissible, $S_{YX}(q)$ is a decreasing function.[18] Consider the aggregate demand schedules of the three types of agents, $\mathcal{D}_{BA}$, $\mathcal{D}_{CB}$, and $\mathcal{D}_{AC}$, with associated supply functions, $\mathcal{S}_{AB}$, $\mathcal{S}_{BC}$, and $\mathcal{S}_{CA}$.

Proposition 3. *If the schedules submitted to the schedule market are admissible, then the amount of each good traded is uniquely determined. If trade occurs, then the price of each good is also uniquely determined.*

Proof. Define $\mathcal{S}_A(q) = \mathcal{S}_{AB}(\mathcal{S}_{BC}(\mathcal{S}_{CA}(q)))$. Admissibility implies each of the supply functions is decreasing, so $\mathcal{S}_A(x)$ is decreasing. Hence, if $\mathcal{S}_A$ has a fixed point, $\mathcal{S}(q_A) = q_A$, it is unique. This fixed point corresponds to the quantity of A traded. (If $\mathcal{S}_A$ has no fixed point, no trade occurs.) The amount of C being traded equals $q_C = \mathcal{S}_{CA}(q_A)$ since $q_C = \mathcal{S}_{CA}(q_A) = \mathcal{S}_{CA}(\mathcal{S}_A(q_A)) = \mathcal{S}_{CA}(\mathcal{S}_{AB}(\mathcal{S}_{BC}(\mathcal{S}_{CA}(q_A)))) = \mathcal{S}_C(\mathcal{S}_{CA}(q_A)) = \mathcal{S}_C(q_C)$, i.e. q_C is the unique fixed point of $\mathcal{S}_C$. A similar logic shows that the amount of B traded equals $q_B = \mathcal{S}_{BC}(q_C)$. Finally, if positive amounts of the goods are traded, prices are $p_A = q_C/q_A$ and $p_B = q_C/q_B$. $\square$

What constitutes an optimal admissible schedule given the schedules submitted by others? If an agent who demands X and supplies Y takes the relative price $p = p_X/p_Y$ as given, the optimal schedule is[19]

$$D_{XY}(p) \;=\; \frac{1}{1+p}$$

[17]One might think a unique equilibrium could be achieved by restricting demand to be a decreasing function and supply to be an increasing function. Unfortunately, this would not allow traders to express their true preferences, because income effects in the Scarf economy result in downward sloping supply curves. The (D) and (I) admissiblity restrictions allow traders to express their true preferences while ruling out multiple equilibria.

[18]Evaluating $S'_{YX}(q)$ at $q = D_{XY}(p)$ where $p = p_X/p_Y$ yields $p + D_{XY}(p)/D'_{XY}(p)$, which is negative by the assumption of inelastic downward sloping demand.

[19]Maximizing $\min(q_X, q_Y)$ over the budget set $pq_X + q_Y \leq 1$ yields $q_X^* = q_Y^* = 1/(1+p)$.

Type	utility	supplies	demands	submits schedule
A	$\min(q_A, q_B)$	A	B	$D_{BA}(p_B/p_A)$
B	$\min(q_B, q_C)$	B	C	$D_{CB}(1/p_B)$
C	$\min(q_C, q_A)$	C	A	$D_{AC}(p_A)$

Table 1. The three types of agents in Scarf's economy and the types of schedules they can submit.

which is admissible. Call this the *competitive* schedule. The associated supply function is $S_{YX}(q) = 1 - q$, which is intuitive since if $1 - q$ units of Y are supplied for q units of X then the agent ends up with equal amounts of X and Y thus maximizing $\min(q_X, q_Y)$.

Given the market-clearing price generally depends on the schedules submitted, is submitting a competitive schedule a best response? We find, surprisingly, it is a weakly dominant strategy. To explore this further consider an example of the Scarf economy with only one agent of each type.

Example 1. Suppose the type-B and type-C agents submit schedules $D_{CB}(p) = D_{AC}(p) = p^{-\alpha}$, which are non-competitive but admissible if $0 < \alpha < 1$. It is readily verified that $S_{BC}(q) = S_{CA}(q) = q^{(\alpha-1)/\alpha}$ so that the supply function that the type-A agent faces is given by

$$S_{BA}(q) \;=\; q^{(\frac{\alpha-1}{\alpha})^2}$$

which is increasing. If the type-A agent submits a competitive schedule she will end up with equal quantities $q_A = q_B = q^*$ where q^* is the unique equilibrium quantity that solves $1 - q = S_{BA}(q)$. If she submits a schedule that gives her $q_B < q^*$ in equilibrium then she is obviously worse off. Suppose she submits a schedule that gives her $q_B > q^*$ in equilibrium. Then she is better off only if also $q_A > q^*$, i.e. if she has to supply less than $1 - q^*$ units of A. But since $S_{BA}(q)$ is increasing, others supply more B only if they get more A. Hence, for a type-A agent to get more than q^* units of B she would have to supply more than $1 - q^*$ units of A. It is thus optimal for the type-A agent to submit a competitive schedule even though others submit non-competitive schedules. □

Remark 2. In the example, the supply function $S_{BA}(q)$ represents the amount of B others are willing to give if they get q units of A. Writing the supply of B as a function of the amount of A taken simplifies the argument for why a competitive schedule is optimal. But the supply function can also easily be expressed in terms of the relative price, $p = p_B/p_A$. From $p_B S_{BA}(q_A) = p_A q_A$ it follows that $q_A = p^{\alpha^2/(2\alpha-1)}$ so

$$S_{BA}(p) \;=\; p^{\frac{(\alpha-1)^2}{2\alpha-1}}$$

which is increasing for $\alpha > \frac{1}{2}$, flat for $\alpha = \frac{1}{2}$, and decreasing when $\alpha < \frac{1}{2}$. The argument that submitting a competitive schedule is optimal now follows from the fact that the elasticity of supply is less than -1 when it is decreasing. This implies that when the type-A agent submits a schedule that gives her more B, she will have to pay more units of A for it. □

In Example 1, the argument that a competitive schedule is optimal for the type-A agent does not depend on the particular functional form of others' schedules but only on the fact that others' supply is increasing, which is more generally true if we impose admissibility.

Proposition 4. *In a Scarf economy with $n \geq 1$ agents of each type, submitting a competitive schedule is a weakly dominant strategy when schedules are restricted to be admissible.*

Proof. Label the type-A agents by $i = 1, \ldots, n$. Suppose there is a schedule $D'_{BA}(p) \neq (1+p)^{-1}$ that gives some type-A agent, denoted i, a higher utility than submitting a competitive schedule. Admissibility implies that the supply function $\mathcal{S}_{BA}(q) = \mathcal{S}_{BC}(\mathcal{S}_{CA}(q))$ is increasing. When agent i submits a competitive schedule she will end up with equal quantities of A and B. By assumption, when agent i submits D'_{BA} she has a higher utility, so agent i must end up with more A and more B. This means that agent i must have given less A and taken more B, so the price of A in terms of B was higher. Since other type-A agents all face the same price, this implies that they must also have taken more B and given less A. But this contradicts the fact that $\mathcal{S}_{BA}$ is increasing. Analogous arguments apply to agents of other types. $\square$

Without the admissibility restriction, the result does not hold; however all agents submitting competitive schedules is a Nash equilibrium (see Appendix A).

The fact that each agent faces a flat supply curve is due to the specific parametrization of the Scarf economy. However, in large economies this would be the case for arbitrary specifications of preferences and endowments. In this sense, Proposition 4 applies more generally when the economy grows large.

To explore the robustness of the mechanism, we consider a repeated market setting where agents submit schedules that are best responses to the previously observed prices but not necessarily best responses to all possible prices. Denote the previous observed prices as p_{t-1}. Consider the following schedule where $\alpha \in [0, \infty]$:

$$\hat{D}_{XY}(p_{t-1}, p) = \frac{p_{t-1} + \alpha}{p + \alpha} \times \frac{1}{1 + p_{t-1}}$$

Notice that $\hat{D}_{XY}(p_{t-1}, p_{t-1}) = \frac{1}{1+p_{t-1}}$ for all values of α. Hence, the schedule is a best response to the previously observed price p_{t-1}. When $\alpha = 1$, $\hat{D}_{XY}(p_{t-1}, p)$ corresponds to the competitive schedule. When $\alpha = 0$, $\hat{D}_{XY}(p_{t-1}, p)$ is the "flattest" admissible schedule that is a best response to p_{t-1} while $\hat{D}_{XY}(p_{t-1}, p)$ limits to the vertical schedule that is a best response to p_{t-1} when $\alpha \to \infty$. The next proposition shows that when agents submit such best response schedules, prices converge to the unique competitive equilibrium for any $\alpha \geq 0$, see Appendix B for a proof.

Proposition 5. *In a repeated Scarf economy with any value of α and starting from any prices p_0, if agents of type A, B, and C submit schedules $\hat{D}_{BA}(p_{t-1}, p)$, $\hat{D}_{CB}(p_{t-1}, p)$, and $\hat{D}_{AC}(p_{t-1}, p)$ respectively, then prices will converge to the unique competitive equilibrium of the economy over successive periods.*

	Original Scarf economy		Experimental economy	
Type	Utility	Endowment	Utility	Endowment
A	$\min(q_A, q_B)$	$(1, 0, 0)$	$40 \min(q_A/10, q_B/20)$	$(10, 0, 0)$
B	$\min(q_B, q_C)$	$(0, 1, 0)$	$40 \min(q_B/20, q_C/400)$	$(0, 20, 0)$
C	$\min(q_C, q_A)$	$(0, 0, 1)$	$40 \min(q_C/400, q_A/10)$	$(0, 0, 400)$

Table 2. Adaptation of Scarf's original economy for the experiment.

This result shows that the schedule market is somewhat forgiving to deviations from submitting competitive schedules and that convergence to the unique competitive equilibrium of the Scarf economy over successive periods can occur even with myopic behavior.

4. Experimental design and procedures

Since one of the goals of the experiment is to replicate Anderson et al.'s (2004) results, we use their (treatment I) parametrization for the Scarf economy. The utility functions and endowments are adapted from those originally used by Scarf as shown in Table 2 below. In Scarf's economy, each agent is endowed with a single unit. In the experiment, this single unit is replaced with multiple units and the utility functions are scaled accordingly. The most numerous good, C, was used as a numeraire and was called "cash" in the experiment. After scaling, the competitive equilibrium prices in terms of cash are 40 for good A and 20 for good B.

There were two treatments: the continuous double auction and the schedule market. A total of 180 subjects participated in the experiment. There were 12 sessions with one group of 15 subjects per session and six sessions per treatment. In a group, five subjects were assigned to each of the three types. Subjects were given the endowments and utility functions shown in Table 2 and were told that they would be paid the value of their holdings after trading, where the value was calculated using their utility functions. There were three unpaid practice periods and 15 paid periods.[20] At the start of each period, endowments were refreshed, no goods were carried over from one period to the next.

At the beginning of the experiment, the instructions were presented using PowerPoint and a paper handout. The instructions included worked examples with Leontief preferences to help subjects understand the induced preferences, and were followed up with a short comprehension test. During the experiment, payoff calculations were performed by the software so that subjects could focus on trading.[21] During the three unpaid practice periods subjects were encouraged to ask questions. An exchange rate of 0.15 Swiss Francs per util was used. The mean payment was 48.74 Francs including a 10 Franc showup fee. The experiment took around 1.5 hours to complete.

[20] The double auction market Session 2 lasted for only ten periods because of a computer crash.

[21] The user interface for both market mechanisms was tailored to Leontief preferences, but it could easily be adjusted to accommodate more general preferences over a pair of goods. For instance, if constant elasticity of substitution utility functions were used, the 'Unused' columns could be replaced by columns showing the marginal utility of each good.

4.1. Continuous double auction implementation

A screenshot of the continuous double auction interface is shown in Figure 3. The left-hand side of the screen shows the subject's utility function, current holdings, and current pay-off. It is also used to construct orders. The right-hand side shows a list of submitted orders, some of which have already transacted. Subjects could submit limit orders to buy and sell the commodities A and B with cash used as the medium of exchange. The price of the last transaction was displayed but subjects could submit orders with any price. As subjects entered figures specifying terms of the order, the payoff consequences of the order were displayed. Transactions occurred as soon as a set of compatible orders had been submitted. Partial filling of orders was allowed. Subjects could cancel or amend orders that had been submitted but had not yet transacted. There was no constraint on the number of orders submitted or the number of transactions. However, subjects could not offer to trade more than they had available. After four minutes had elapsed, the period ended. Subjects were shown a 'results screen' with their earnings for the period, a list of the trades they made, and their total earnings from all completed periods.

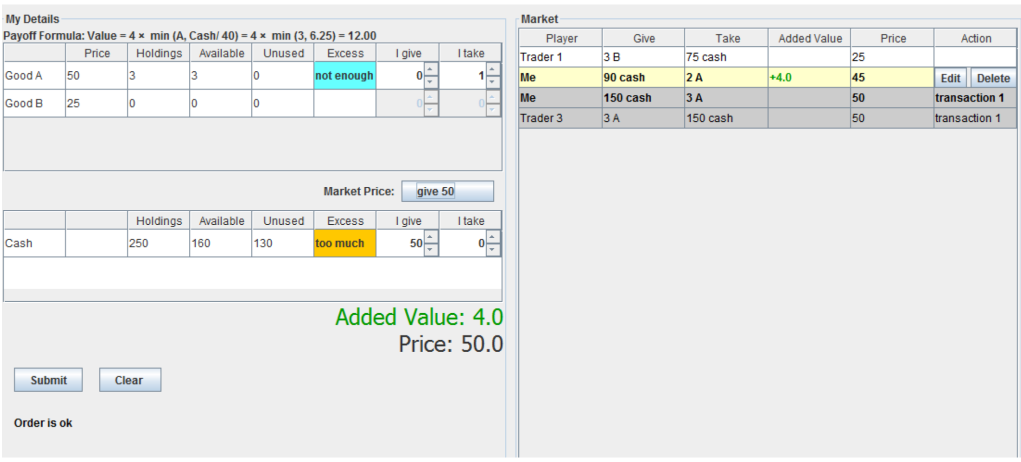


Figure 3. User interface for the continuous double auction market. The screen is from the point of view of a type-C agent who was endowed with cash and needs cash and good A. On the top left of the screen, the text beginning 'Payoff Formula' shows the subject's utility function and the value of the current holdings. Below this is a table with labeled rows for each of the goods. The column headed 'Price' shows the last trade price, 'Holdings' are the current holdings, 'Available' are current holdings which the subject is not currently offering to trade, 'Unused' are the current holdings that are not contributing towards earnings, and 'Excess' indicates similar information in words (in the screen shot, the subject has too much cash and not enough of good A, so can increase earnings by trading cash for good A). The columns 'I give' and 'I take' are used to construct orders. This is done by entering numbers in the columns. As numbers are entered, the 'Added Value' number automatically updates to show how earnings will change if the order transacts. The table on the right hand side shows the orders that have been submitted. There are currently two active orders. Trader 1 is offering to sell good B at a price of 25. Trader 2 (labeled 'Me') is offering to buy good A at a price of 45. There has been one transaction, the current subject bought three units of good A.

4.2. Schedule market implementation

A screenshot of the schedule market interface is shown in Figure 4a. Subjects constructed a schedule by specifying how much they wanted to trade at each of a range of prices. The range of prices was pre-determined and constant throughout the experiment. The admissability restrictions were enforced automatically. When the subject changed their demand at one price, if the restrictions were not satisfied, the computer adjusted demands at other prices to satisfy the restrictions. For example, if demand was initially zero at every price and the subject set demand at the highest price to one, the computer would automatically set the demand at all lower prices to one to satisfy admissibility.

Interpolation was used to produce a continuous demand function from the quantities specified by subjects at each of the pre-specified prices. Suppose a subject is endowed with good A and demands good B. They specify their demand for B at pre-specified prices, giving a list of (p, q_B) pairs. This is converted to a list of (p, q_A) pairs where $q_A = q_B p$. Then the following interpolation is performed. Given two price-supply points (p_1, q_{A1}) and (p_2, q_{A2}), intermediate points are $(p_1\alpha + p_2(1-\alpha), q_{A1}\alpha + q_{A2}(1-\alpha))$ where $\alpha \in [0, 1]$. Interpolation is carried out using price-supply pairs rather than price demand pairs so that intermediate points satisfy the admissibility conditions that ensures a unique equilibrium. The interpolation was performed as subjects were entering the schedule and subjects could see the interpolated points on the graphical representation of the schedule before they submitted it. Although subjects could not directly specify demand at every possible price, they could do so indirectly. For example, suppose a subject wanted to demand a certain quantity at price 7, but the nearest pre-specified prices were 5 and 10. By suitable adjustment of the quantities demanded at 5 and 10, they could specify the desired quantity at price 7.

A period ended if all schedules had been submitted or if four minutes had elapsed. The submitted schedules were summed to produce an aggregate demand function. A numerical optimization procedure was applied to the aggregate demand function to find the market clearing prices. Trade occurred at these prices with each agent trading the quantities specified by their submitted schedule. Figure 4b shows the 'results screen' from the schedule market shown after the period ended. Subjects were shown the demand schedule they had submitted and the residual supply that they faced (the combinations of price and quantity taken by the subject that would equalize supply and demand in all markets).

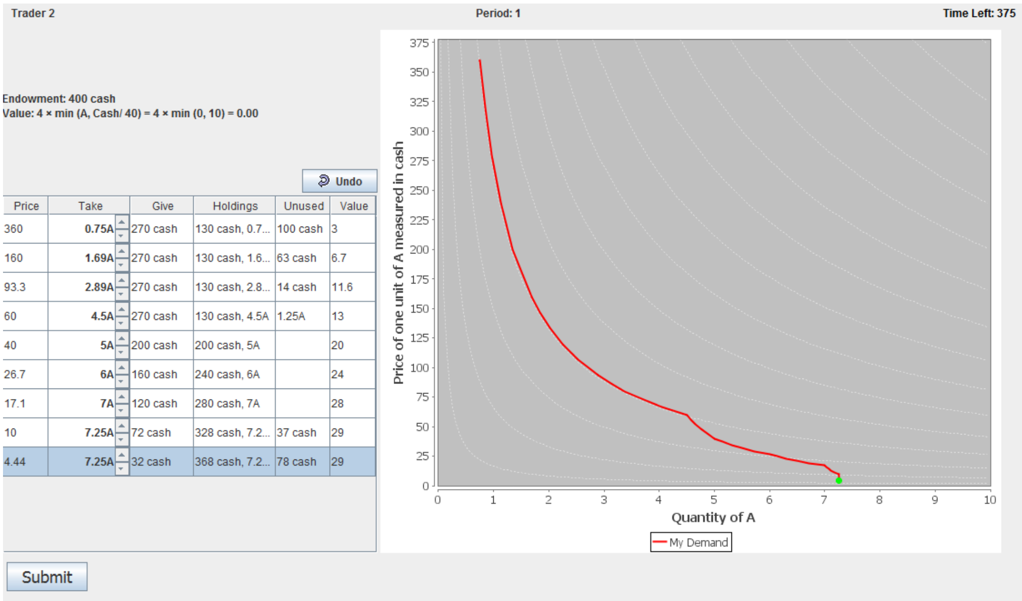


Figure 4a. User interface for the schedule market. The left-hand side of the screen shows the subject's endowment, utility function, and a table that can be used to create demand schedules. The first column of the table is a fixed list of possible prices of the commodity that the subject needs in terms of the commodity with which the subject is endowed. In this case, the subject needs commodity A and has cash, so the price column shows the price of A measured in cash. The subject's task is to fill in numbers in the 'Take' column, representing their demand at each of the prices. As subjects enter numbers, the displayed graph and the relevant numbers in the table (the columns labeled 'Give', 'Holdings', 'Unused', and 'Value') are automatically updated to reflect their choices. The 'Give' column indicates how much subjects will give up for what they want to take at the specified price. The 'Holdings' column shows holdings after trading at each price. The 'Unused' column indicates whether any of the holdings will be unused, i.e. not contribute towards the subject's payoff. The 'Value' column shows the utility of the holdings at each price. The 'Undo' button lets subjects revert to previous states of the schedule, making it easy to correct mistakes and to experiment with different configurations. Once the subject had finished editing the schedule, they pressed 'Submit'. Once submitted, schedules could no longer be altered.

The right-hand side of the screen shows a graphical representation of the schedule being constructed. It shows the quantities chosen at each of the listed prices and the interpolated quantities at intermediate prices. The green dot indicates the point being edited on the left-hand side (currently $p = 4.44$ and $q = 7.25$). The dotted curved grid lines represent the admissibility constraint. For a schedule to be admissible, as q increases, a schedule can only go to a lower grid line.

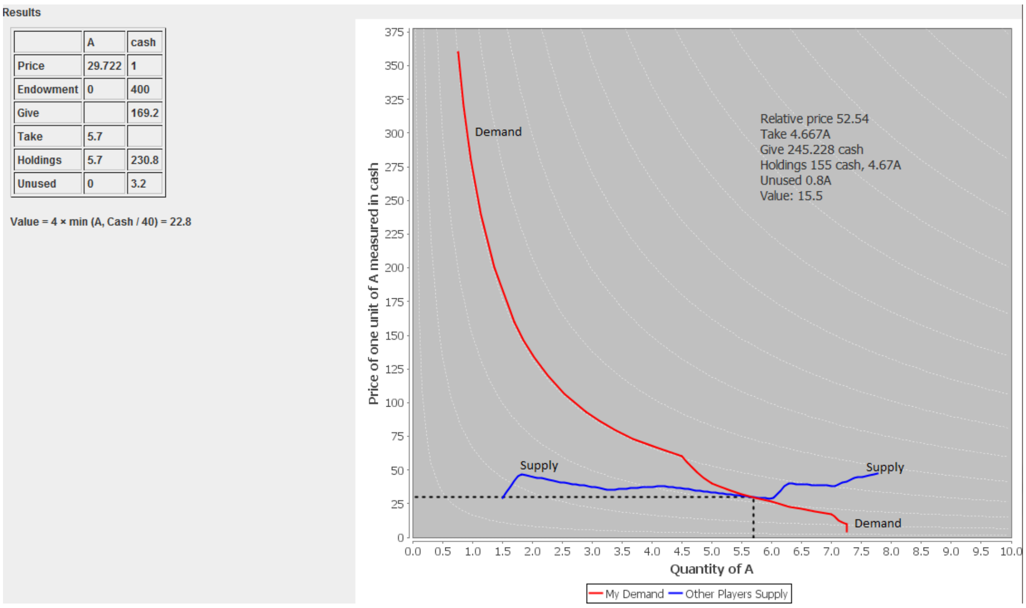


Figure 4b. Results screen for the schedule market. The curve labeled 'Demand' is the demand schedule submitted by the subject. The curve labeled 'Supply' is the residual supply the subject faces. At each point on the residual supply curve, excess demand is zero in all markets. The intersection of the demand and supply curves yields the market equilibrium for the submitted schedules. The text on the left shows prices, quantities traded, holdings after trading, and payoffs. Subjects could also see what would have happened if they had submitted a different schedule. They could do this by moving the mouse pointer to any location in the graph. The text on the top right would then be automatically updated to show the potential payoffs at the position of the mouse pointer.

5. Results

We compare the two market institutions in terms of price stability (Section 5.1), as well as in terms of efficiency (Section 5.2) and equality (Section 5.3). A detailed comparison of the price dynamics observed in our double auction experiments to those of the Anderson et al. (2004) study can be found in the Appendix.

5.1. Price dynamics in the two market institutions

The between-period prices observed in our double auction market experiments are shown in the top six panels of Figure 5, where each panel corresponds to a different session. In each session, prices start in the lower-left corner in the first period of the experiment and then cycle in a counter-clockwise pattern without any obvious tendency for convergence. We thus replicate this main finding of the Anderson et al. (2004) paper, see their Figure 4. Besides counter-clockwise cycling, the observed price paths confirm other features of the tatonnement predictions in Figure 1a. For instance, the further the price of good B falls below its equilibrium level the further the price for good A shoots

	α_1	α_2	α_3	α_4	α_5	α_6	β
D.A	0.68^{**}	0.84^{***}	0.88^{**}	-0.80^{***}	0.81^{***}	0.31	0.041
	(0.23)	(0.16)	(0.20)	(0.18)	(0.18)	(0.19)	(0.021)
Schedule	-2.13^{***}	-2.30^{***}	-2.79^{***}	-2.70^{***}	-2.56^{***}	-2.59^{***}	-0.034^{**}
	(0.08)	(0.15)	(0.13)	(0.22)	(0.21)	(0.13)	(0.011)

Table 3. Estimating the time dependence of the Lyapunov function. Values of the Lyapunov function were calculated with prices normalized so that the competitive equilibrium is $(1, 1)$. A negative/zero/positive β corresponds to prices converging/cycling/diverging. The estimated β is not significantly different from 0 for the double auction market, which indicates that prices cycle. For the schedule market the estimated β is negative, which indicates convergence. Standard errors (robust and without clustering) are in parentheses; * indicates $p < 0.05$, ** indicates $p < 0.01$, *** indicates $p < 0.001$,

out and the less the price path looks like a circle. Compare, for instance, the first and fourth double auction sessions in Figure 5.

> **Result 1.** Prices in the double auction market do not converge to their competitive equilibrium levels.

Support. To test whether prices are converging to equilibrium we evaluate the Lyapunov function of Section 2.1 along the observed price paths. Consider a simple regression model of the form

$$L = \Phi(\alpha_s + \beta(\text{period} - 1) + \varepsilon) \qquad (1)$$

where L is the Lyapunov function and $\Phi(\cdot)$ is the standard normal cumulative distribution.[22] The session-specific constant α_s corresponds to the path's starting point (i.e. the trade price in the first period). The parameter β measures stability: $\beta < 0$ implies convergence, $\beta > 0$ implies divergence, and $\beta = 0$ means that the prices are cycling along a closed path. The estimation results for the double auction market are shown in the top row of Table 3. The estimated β coefficient is not significantly different from 0, indicating that there is no convergence to equilibrium but that, on average, prices cycle along a closed path as the tatonnement model predicts for the Scarf economy. □

The price patterns observed in the schedule market sessions are shown in the bottom six panels of Figure 5. The most striking difference is the absence of any cycling tendencies. The introduction of schedules has stabilized prices, which are close to competitive equilibrium levels.

> **Result 2.** Prices in the schedule market converge to their competitive equilibrium levels.

Support. We apply the regression in (Equation 1) to the Lyapunov function in Section 2.2. The estimated β coefficient is significantly negative for the schedule market, see

[22] We introduce this transformation because the Lyapunov function is bounded between 0 and 1 so a simple linear regression would not be appropriate.

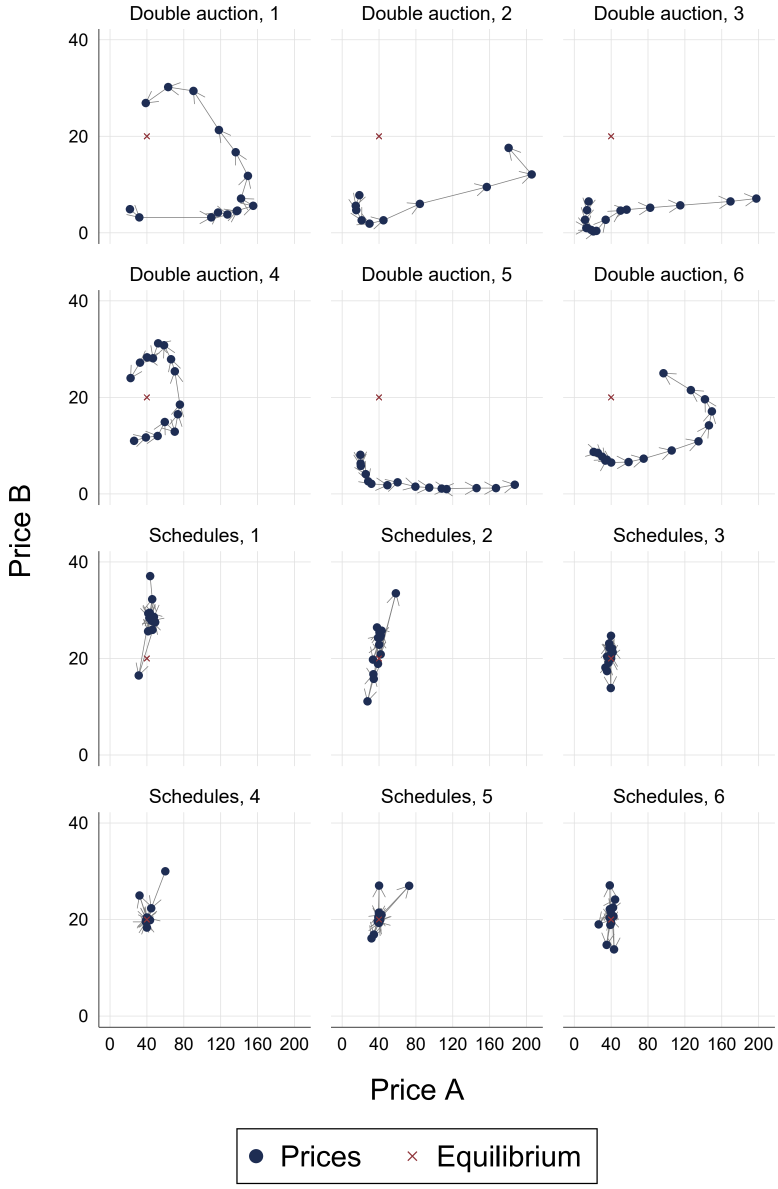


Figure 5. Between-period prices in the double-auction market sessions (top six panels) and the schedule market sessions (bottom six panels). Prices cycle in a counter-clockwise manner in the double auction markets and converge straight to equilibrium in the schedule markets.

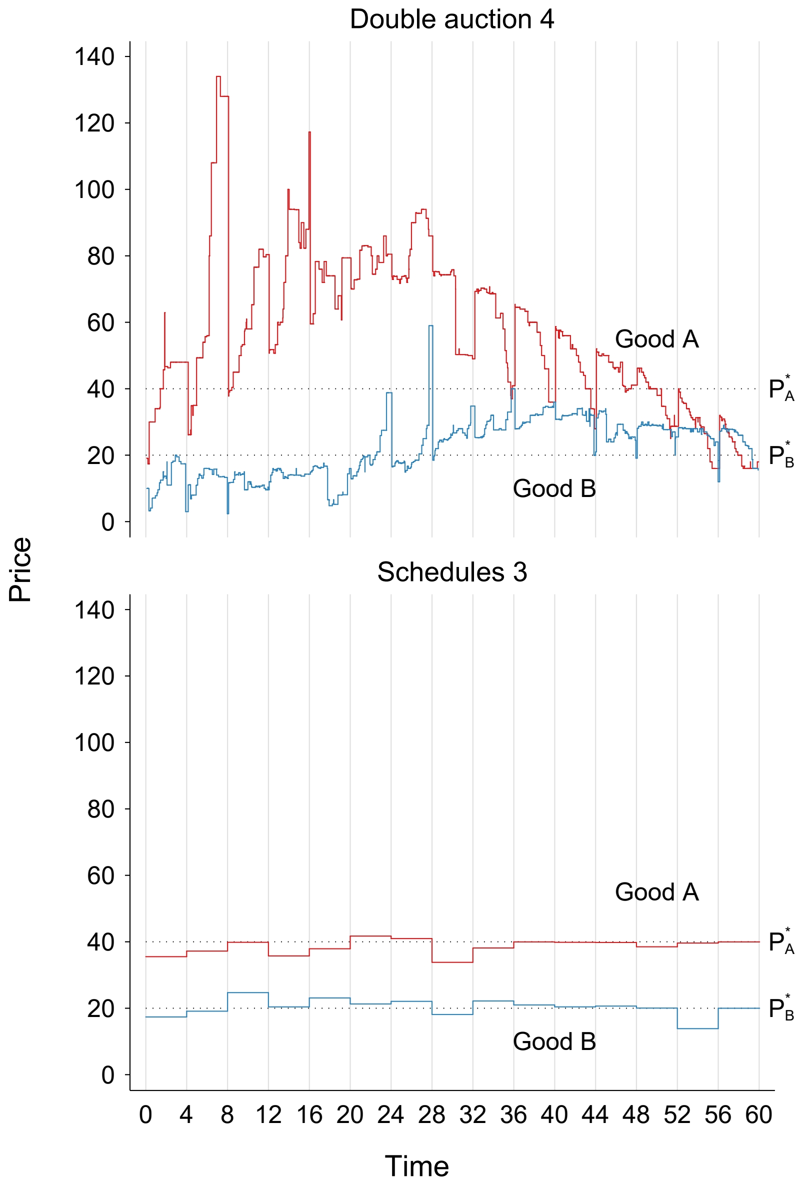


Figure 6. Price evolution in double auction market Session 4 (top) and schedule market Session 3 (bottom). The dotted lines show the competitive equilibrium price for each good. All 15 periods of the experiment are shown, each lasting 4 minutes as indicated by the thin vertical lines. Prices in the double auction market oscillate with substantial intra-period variation. In contrast, prices in the schedule market are steady and close to their competitive equilibrium values.

the bottom row of Table 3. Note, however, that the coefficient is rather small. The reason is that even in the first period of the experiment, prices are in the vicinity of the equilibrium so there is limited scope for further convergence. Notice that the values for α are lower in the schedules treatments indicating different initial conditions across treatments. Another measure of convergence is provided by Anderson et al. (2004) who define prices to be close to equilibrium if they fall in the range $(p_A, p_B) \in [36.5, 43.5] \times [16.5, 23.5]$. Taking averages over all 15 periods in each of the six sessions that used the schedule market, the observed average prices for good A are: 43.6(1.1), 39.5(1.7), 38.6(0.6), 41.2(1.5), 41.3(2.3), 39.2(1.1), with the standard error in parentheses, and for good B they are: 28.2(1.1), 22.4 (1.4), 20.3(0.7), 21.0(0.8), 20.6(0.8), 20.5(0.9). Except for the first session, observed average prices are all close to the equilibrium. The overall price averages over all six sessions are $p_A = 40.6(0.6)$ and $p_B = 22.2(0.5)$. □

The stark difference in price evolution under the two different market mechanisms is illustrated in Figure 6. The top panel shows the observed trade prices in one of the sessions that used the double auction market. Trade prices for all 15 periods are shown, where each period lasted four minutes as indicated by the thin vertical lines, for a total time of 60 minutes. The trade prices oscillate with substantial intra-period variation. The pattern of the oscillations suggests that when one price crosses its equilibrium value the other price is furthest from its equilibrium value, as the tatonnement model predicts (see Figure 1a). The bottom panel shows observed trade prices in the schedule market, using the same scale for the axes to emphasize the lack of price variability under this mechanism. Trade prices start close to their equilibrium values and remain close for the entire duration of the experiment.

The tatonnement model makes more specific predictions than the non-convergence Result 1. For instance, it predicts that prices move along a closed orbit in a counter-clockwise manner. More specifically, for any pair of prices, the tatonnement model makes a precise prediction for the direction of price changes. An alternative prediction is that prices converge along a straight path to the competitive equilibrium. Figure 7 shows how well these two alternatives predict the actual direction of price changes for the double auction (top two panels) and schedule market (bottom two panels). The histograms are based on the difference between the predicted and observed angles of price changes.[23] There is a big spike at 0 degrees for the tatonnement model in the double auction market, and a similar spike at 0 degrees for the convergence model in the schedule market. Applying the convergence model to the double auction market, or the tatonnement model to the schedule market, results in spikes at ±90 degrees in line with the counter-clockwise cycling behavior predicted by the tatonnement model and observed in the double auction market.

> **Result 3.** The direction of price changes in the double auction market is well predicted by the tatonnement model. In the schedule market, prices converge to the competitive equilibrium along a straight path.

[23] Angles were calculated using prices normalized so that the competitive equilibrium is $(1, 1)$. Errors are measured in the clockwise direction. If the predicted direction is north and the observed direction is north east, the error is +45 degrees. If the predicted direction is north and the observed direction is west, the error is -90 degrees.

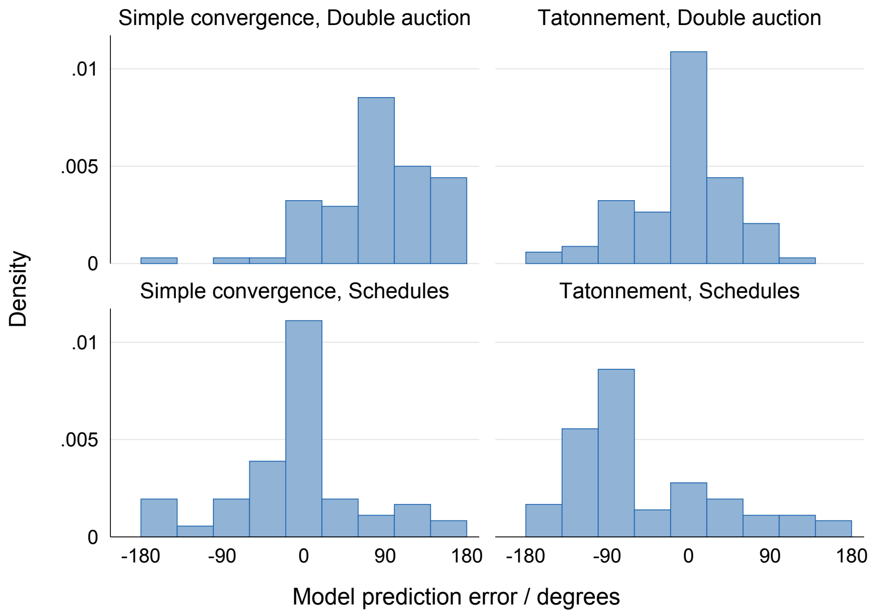


Figure 7. Prediction errors for the simple convergence model (left panels) and the tatonnement model (right panels) in the double auction market (top panels) and the schedule market (bottom panels). For the double auction market, the tatonnement model produces a big spike at zero degrees (no error) while the simple convergence model does the same for the schedule market.

Support. Consider the following simple regression, which explains price movements in the A and B markets in terms of excess demand, as the tatonnement model predicts, and straight convergence:

$$\begin{aligned} p_A(t+1) - p_A(t) &= \beta_{tat}\, z_A(p_A(t), p_B(t)) + \beta_{conv}(p_A^* - p_A(t)) + \varepsilon_A \\ p_B(t+1) - p_B(t) &= \beta_{tat}\, z_B(p_A(t), p_B(t)) + \beta_{conv}(p_B^* - p_B(t)) + \varepsilon_B \end{aligned}$$

The results are shown in Table 4. Note that only β_{tat} is different from 0 in the double auction market while only β_{conv} is different from 0 in the schedule market. □

For additional analysis concerning price dynamics in the double auction market we refer the reader to the Appendix, which compares our results to those of the Anderson et al. (2004) study who used only the double auction market. As discussed in the Appendix, we replicate all the findings that pertain to their "cycling " treatment I that we used for our double auction market experiments. Our main interest is in comparing the two market institutions, in particular, how price (in)stability affects outcomes in terms of efficiency and equality.

	β_{tat}	β_{conv}
Double Auction	0.564^{***}	0.021
	(0.067)	(0.013)
Schedule	-0.183	0.681^{***}
	(0.205)	(0.079)

Table 4. Explaining the direction of price changes using the tatonnement model and a model of straight convergence. In the double auction market, price changes are driven only by excess demands as the tatonnement model predicts. Prices in the schedule market are not affected by excess demands but instead converge along a straight line to the equilibrium. Prices are normalized so that the competitive equilibrium is $(1, 1)$ and excess demand for a good is normalized by dividing by the total quantity of that good in the economy. Standard errors are in parentheses; *** indicates $p < 0.001$.

5.2. The effects of price (in)stability on market performance

This section models the allocations and hence welfare that results from trading at non-equilibrium prices. In a tatonnement institution, an auctioneer adjusts prices in response to reported demands and no trade takes place until market-clearing prices are found. For the Scarf economy this implies that no trade ever takes place. Our double auction market experiments, however, show that there is substantial trade at non-equilibrium prices. A variant of the model proposed by Hahn and Negishi (1962) can be used to model out-of-equilibrium trade. Assume that traders are price-takers and exchange goods in fixed ratios determined by the prices until they hold equal proportions of the goods they want. Of course, if prices are out of equilibrium, not all traders are able to achieve a balanced portfolio of the goods they want. The market does not clear and some traders are left with "unused" goods. However, traders of at least one type will have goods in the desired proportions, making the outcome Pareto optimal and hence no further trade possible.

To derive predictions for the price-taking model, consider the original Scarf economy of Section 2. Let $(p_A, p_B, p_C = 1)$ denote the price vector and let (q_A, q_B, q_C) denote the quantities traded at these prices.

Proposition 6. *For the original Scarf economy, the price-taking model predicts that the quantities traded are*

$$\begin{aligned} q_A &= \min\Big(\overbrace{\frac{p_B}{p_A+p_B}}^{supply}, \overbrace{\frac{1}{1+p_A}}^{demand}\Big) \\ q_B &= \min\Big(\overbrace{\frac{p_A}{p_A+p_B}}^{demand}, \overbrace{\frac{p_A}{p_B+p_Ap_B}}^{supply}, \overbrace{\frac{1}{1+p_B}}^{cash}\Big) \end{aligned} \tag{2}$$

and the resulting welfare is

$$W = q_B + \min(1 - q_B, q_A p_A) + q_A$$

Welfare is maximized at the competitive equilibrium prices, $p_A = p_B = 1$.

In the market for A, type A sells good A to type C. Type C pays using cash from their endowment. When prices are taken as given, there are two constraints on the quantity of A traded: type A's demand and type C's supply. In the market for B, type B sells good B to type A. Type A has no cash in their endowment so must use cash from sales in market A to fund purchases in market B. When prices are taken as given, there are three constraints on the quantity of B traded: type A's demand, type B's supply, and type A's cash receipts from market A. The quantity of cash traded in each market can be simply calculated using the prices and quantities. [24]

Welfare is taken to be the sum of each type's utility. The motivation for this formulation, given that in the Scarf economy preferences are not quasi linear, is as follows. First, given the symmetry between agent's utility functions it seems natural to give the utilities of different types equal weight. Second, it corresponds with subject's earnings in the experiment. The utilities are calculated as follows. Type A's utility is simply the quantity of good B traded. Type B's utility is the minimum of B's holdings of good B after trading and the amount of cash type B can receive from selling good B. Finally, type C's utility is the quantity of good A traded.

We define market efficiency as the fraction of the total gains from trade that are realized:

$$\text{efficiency} = \frac{W_{observed}}{W_{max}}$$

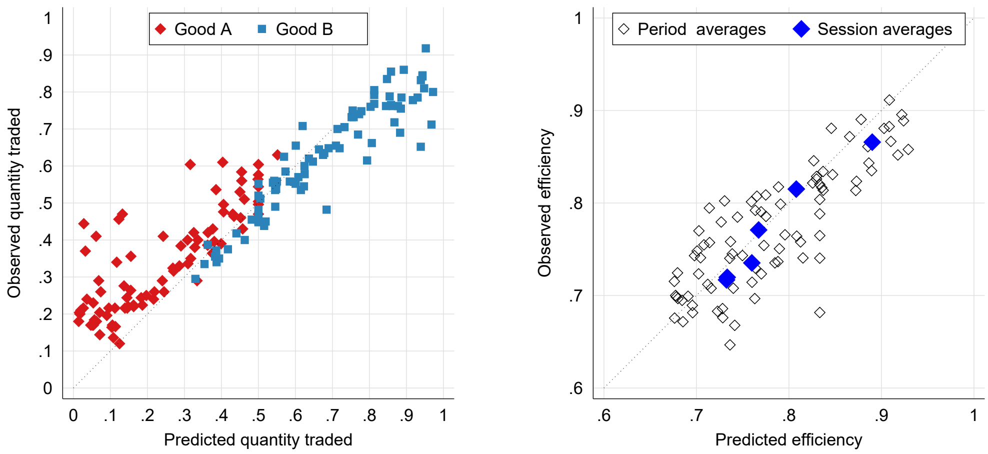
The two panels of Figure 8 show (a) the observed quantities traded and (b) the observed efficiency levels for the double auction market sessions against the predictions of the price taking model. The predictions are based on Proposition 6 using the opening prices for the period: in the first period the opening prices are $(p_A = 20, p_B = 10)$, i.e. half the equilibrium prices, and in later periods the opening prices are equal to the average trading prices in the prior period. For the efficiency levels, the small diamonds correspond to period averages and the large diamonds to session averages. The price-taking model predicts efficiency levels remarkably well.[25] This is especially true for the six session averages, which are all very close to the 45-degree line.[26]

Observed efficiency levels in the double auction market range from 65% to 91% with more observations towards the lower end. The lower bound should come as no surprise since even if prices are completely off the predicted welfare is $W_{min} = 1$, see Proposition 6, corresponding to an efficiency level of 67%.[27]

[24] The predictions of the model can be thought of as the rest point of a dynamic system where each agent attempts the locally optimal trade taking prices as given. This is similar to the Local Marshallian Equilibrium theory proposed by Asparouhova et al. (2011) but with prices fixed. That is, agent i with utility function u^i, and a vector of holdings $\boldsymbol{\omega^i}$ attempts to trade in the direction defined by the vector $\boldsymbol{d^i}$ that solves $\arg\max_{\boldsymbol{d^i}} \nabla u^i(\boldsymbol{\omega^i}) \cdot \boldsymbol{d^i}$ subject to $\boldsymbol{d^i} \cdot \boldsymbol{p} = 0$ and $\boldsymbol{d^i} + \boldsymbol{\omega^i}$ being a feasible allocation. When starting from the Scarf economy endowments, there is one agent supplying and demanding each of the goods and trade can continue until one agent achieves goods in the desired proportions.

[25] The coefficient for correlation between the predicted and observed period averages is 0.770 ($p < 0.0001$).

[26] Two sessions resulted in almost identical and observed efficiency levels, which is why it appears as



(a) Quantities

(b) Efficiency

Figure 8. Observed and predicted quantities traded and efficiency levels in the double auction market. The predictions are based on a model of optimizing behavior taking prices as given, with the prices being equal to the opening prices for the period. The opening prices are set to half the equilibrium prices in the first period and to the average trade prices of the previous period in later periods.

Result 4. In the double auction market, observed efficiency is 77%. In the schedule market, observed efficiency is significantly higher: 95%.

Support. Efficiency levels for the six schedule market sessions are 89.6%, 92.8%, 96.4%, 97.6%, 97.7%, 96.6%, and for the double auction market sessions they are 77.1%, 73.5%, 72.0%, 86.6%, 71.7%, 81.5%. Note that all six efficiency levels for the schedule market are higher, so the null hypothesis that efficiency levels are the same can be rejected ($p = 0.0022$, Median test). □

We next determine how the total gains from trade are divided among the different types of agents.

5.3. *The effects of price (in)stability on equality*

Using the price-taking model of Proposition 6 together with the opening prices for the period we can predict the gains from trade by agent type in the double auction market.

---

if there are only five large filled diamonds.

[27]If all goods are randomly assigned to one type of trader, predicted efficiency is $W_{min} = 1$. A stricter definition of efficiency that corrects for this baseline level would be

$$\text{normalized efficiency} = \frac{W_{observed} - W_{min}}{W_{max} - W_{min}}$$

The normalized efficiency of the double auction market is 31% and that of the schedule market 85%.

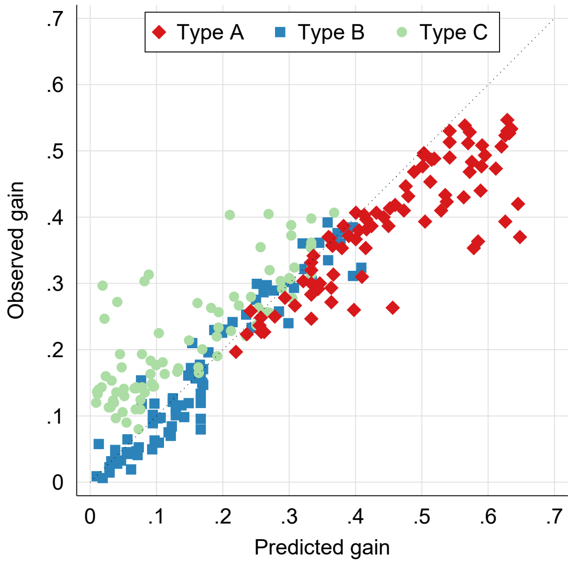


Figure 9. Observed and predicted gains for each agent type in the double auction market. The predictions are based on a model of optimizing behavior taking prices as given and equal to the opening prices for the period. The opening prices are half the equilibrium prices in the first period and the average trade prices of the previous period in later periods.

Figure 9 compares these predictions (horizontal axis) with the observed gains (vertical axis). The price-taking model also does a good job at predicting gains by agent type, although the observed gains for the type-A agent are somewhat lower than predicted while the gains for the type-C agent are somewhat higher than predicted.

To explain this discrepancy, note that the outcomes in Proposition 6 are derived without taking into account the precise details of the trading institution. It is as if the type-A and type-C agents can exchange good A, type-A and type-B agents exchange good B, and type-B and type-C agents exchange good C. As a result, the type-A agent, for instance, would never be left with any excess cash. But in the double auction market, the type-A agent first has to trade with the type-C agent to acquire cash with which she can then buy good B from the type-B agent. And if the type-A agent overestimates how much of good B will be supplied she may be left with excess cash at the end of the period. Note that the other two agent types do not face this problem since they only trade one type of good. The type-B and type-C agents can simply keep trading until (i) they have as much of the good they demand as the good they are endowed with, in which case they have no unused goods or (ii) they hit the limit of the supply of the good they demand, in which case they are left with unused units of the good they are endowed with.

This intuition is confirmed by Figure 10, which shows for each agent type the amount of unused goods (normalized by the total quantity of the good in the economy) averaged

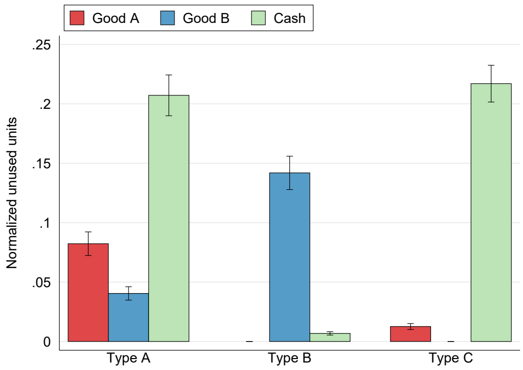


Figure 10. Unused units of each of the three goods by agent type (normalized by the total amount of each good in the economy). The bars show the mean and its standard error. The type-B and type-C agents have unused units only of the goods they demand while the type-A agent has unused units of all three goods.

over all periods and sessions that used the double auction market. The type-B and type-C agents have unused units only of the good they demand while the type-A agent has unused units of all three goods. Because the type-A agent must first trade with the type-C agent to get cash not knowing how much of good B she will be able to buy with that cash, she faces the most uncertainty and, hence, the most difficult task in achieving a balanced portfolio, especially since prices fluctuate within a period. This explains why the type-A agent's gains are somewhat lower then predicted (as indicated by the circles in Figure 9). The tendency of type-A agents to stock up on too much cash benefits the type-C agent, which explains why their observed gains are somewhat higher than predicted (as indicated by the squares in Figure 9).

Notwithstanding these small discrepancies, Figure 9 does a remarkable job at explaining the division of the total gains from trade. The most striking feature of Figure 9, however, is the large variation of gains across agent types. The shares that the type-B agents get (indicated by the squares) are all in the lower-left corner while the shares for the type-A agents (diamonds) are in the upper-right corner and the shares of the type-C agents (circles) are somewhere in the middle. The degree of inequality in the double auction markets is even more clear from Figure 11, which shows the shares by agent type over time in the double auction market sessions (top six panels) and the schedule market sessions (bottom six panels).[28] The white space at the top of each panel indicates the degree to which there was a loss in efficiency.

> **Result 5.** The double auction market results in large inequalities. In contrast, the schedule market results in approximately equal payoffs.

Support. The division of the total gains from trade among the three types of agent

[28]Recall that in the double auction market Session 2 there were only 10 periods due to a computer crash.

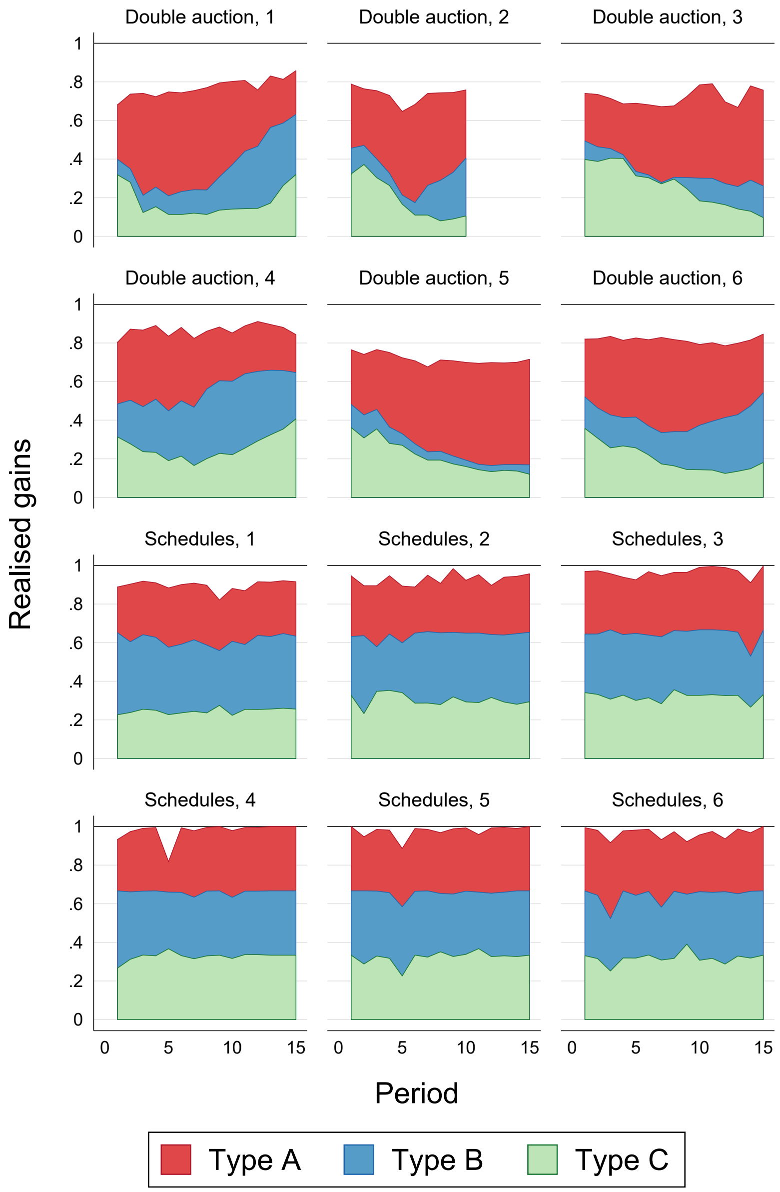


Figure 11. Earnings by agent type in the double auction market sessions (top six panels) and the schedule market sessions (bottom six panels). The white space at the top of each panel indicates the degree to which there was a loss in efficiency. (Recall that double auction market Session 2 lasted for only 10 periods.) There are large inequalities in the double auction markets where shares fluctuate over time. In contrast, the schedule markets result in more equal and stable outcomes.

is roughly even (32.3%, 35.5%, 31.2%) in the schedule markets while it is very uneven (51.1%,20.5%,28.4%) in the double auction markets.[29] Moreover, Figure 11 shows that the division of surplus by agent type is constant over time in the schedule market. In contrast, the shares earned by the different types fluctuate over time in the double auction market and in some instances the outcome is extreme inequality. See, for instance, double auction market Sessions 3 and 5. □

6. Conclusion

General equilibrium theory is one of the triumphs of modern economic analysis. It provides a complete account of entire economies, predicting the exchanges required to arrive at Pareto efficient allocations as well as the prices that define the terms of exchange. The assumptions underlying the theory are that agents maximize their utility at given prices, i.e. *price-taking* behavior, and that prices are such that no good is in excess demand or supply, i.e. prices are *market clearing.*

Despite its powerful mathematical structure and broad applicability, general equilibrium theory is a static theory that does not address how market clearing prices come about. Walras posited a centralized price adjustment process, where a fictitious auctioneer adjusts prices in response to reported demands until market clearing prices are found after which trade occurs. While this "tatonnement" process converges for economies satisfying gross substitutability, Scarf's (1960) simple example demonstrates that without this strong assumption, prices may cycle forever thus precluding trade from occurring. In other words, prices are globally unstable in Scarf's economy, which is perpetually out of equilibrium.

This prediction suggests that the Scarf economy forms an ideal test for general equilibrium theory. And laboratory experiments are the perfect tool to perform such a test. For all the quibbles about representativeness, selection effects, external validity, and lab-field generalizability, one might almost forget about the enormous potential for controlled laboratory experimentation to address questions of basic science. In a pioneering study, Anderson et al. (2004) capitalize on this potential by testing the Scarf economy in a series of double auction market experiments. Their results are fascinating. While the double auction is a non-tatonnement institution, average trade prices in the experiments cycle along a closed orbit around the unique competitive equilibrium with no clear sign of convergence just as the tatonnement model predicts. This is a profound finding that reveals the limits with which general equilibrium models can be applied to predict economic outcomes. Moreover, it has repercussions for the performance of naturally occurring markets, since most contemporary financial and commodity markets employ the double auction institution.

In this paper we replicate the findings of the Anderson et al. (2004) study. Our double auction market experiments confirm that the tatonnement model predicts the

[29] The 95 percent confidence intervals for the division of gains among the three types are (31.7-33.0%, 33.5-37.5%, 30.5-33.8%) for the schedule market and (44.2-58.0%, 12.8-28.1%, 24.8-32.1%) in the double auction. The confidence intervals were calculated using bootstrapping with clustering on groups. The Gini coefficients for the double auction and schedule markets are 0.28 and 0.05 respectively.

direction of price changes remarkably well and that prices are globally unstable as a result. We then ask two important questions that go beyond replication of the Anderson et al. study. First, since competitive equilibrium is never reached in the double auction market, the tatonnement model predicts no trade. But out-of-equilibrium trades occur all the time in the experiment, which raises the question "what model explains out-of-equilibrium trading?" Second, we demonstrate the negative impact of price instability for the economy's performance, in terms of efficiency and equality (see Figure 11), and ask the market design question "how can the economy be fixed, i.e. what institution stabilizes prices and delivers efficient and equitable outcomes?"

With regards to the first question, we provide clear evidence of price-taking behavior in the absence of market clearing. Within a period, traders exchange goods in fixed ratios determined by the prices until they hold equal proportions of the goods. Because prices are out of equilibrium, not all traders can achieve a balanced portfolio resulting in some unused units. These imbalances put upward or downward pressures on prices, which then adjust according to the tatonnement model. The simple price-taking model predicts the allocations of the goods and the division of the total gains from trade extraordinarily well (see Figures 8 and 9).

With regards to the second question, we provide clear evidence that a call market where traders submit demand schedules fixes the Scarf economy. Our proposal is inspired by Smale's (1976a; 1976b) work on Newtonian methods and his desire to "Extend the mathematical model of general equilibrium theory to include price adjustments," which he deemed to be one of the great problems for the 21st century (Smale, 1998). Specifically, our proposed solution is to let agents submit demand schedules, i.e. a list of quantities demanded at various prices, and then an automated market mechanism based on the global Newton method is run to determine the terms of trade.

In the schedule market, price-taking behavior takes the form of submitting a "competitive schedule," i.e. a set of quantities that are utility maximizing taking prices as given. We prove that price-taking behavior is a weakly dominant strategy (see Proposition 4). While the proof relies on the specific parametrization of the Scarf economy, the results are more generally true in large economies where agents face supply curves that are approximately flat. As a consequence, submitting a competitive schedule is optimal for arbitrary specifications of preferences and endowments when the economy grows large.

We also test the schedule market in the laboratory and find that it performs extremely well. In most periods, prices are close to the competitive equilibrium values and between periods converge quickly to the competitive equilibrium (see Figures 5 and 6). Importantly, the schedule market is able to translate price stability into improved performance: observed efficiency is 95% (compared to 77% in the double auction market) and outcomes are highly egalitarian (see Figure 11). Besides the desirable theoretical properties and the excellent performance in our empirical tests, schedule markets are also practical. Electricity markets and treasury auctions allow for schedules, which are also used to determine opening prices for the day on the New York Stock Exchange.

Our results thus have implications for market design that extend beyond the Scarf economy. Nowadays, variants of the tatonnement institution are frequently used in auctions to privatize major public assets. For instance, in the FCC's simultaneous ascending auction, the price of items for which demand exceeds supply is increased until there is no more excess demand. Submitting demand schedules could be an alternative to the iterative adjustment of prices. Univariate functions can be used for schedules in cases (like our experiment) where each trader desires to exchange one type of good for another. Vector functions can be used in cases where traders demand a number of different goods. Our experimental results demonstrate that in settings with complementarities and income effects, an institution whose price adjustments approximate tatonnement dynamics does not necessarily result in competitive equilibrium outcomes while a call market that admits schedules does.

Appendix A. The schedule market without the admissibility restriction

Without the admissibility restriction, submitting a competitive schedule is optimal when others behave competitively. In other words, all agents submitting competitive schedules constitutes a Nash equilibrium even when we drop the admissibility restriction.

Proposition A1. *In a Scarf economy with $n \geq 1$ agents of each type, it is a Nash equilibrium for all agents to submit a competitive schedule.*

Proof. Consider the type-A agents. When the type-B and type-C agents submit competitive schedules the supply of B in terms of A will be one-to-one, i.e. for every q units of B supplied q units of A are demanded. This implies that the relative price of the two goods has to be one, i.e. when others submit competitive schedules then the supply curve the type-A agents face is "flat" at a relative price of 1. A similar logic holds for the other agent types. Hence, no agent can do better by submitting a non-competitive schedule. □

Appendix B. Proof of Proposition 5

Proposition 5. *In a repeated Scarf economy with any value of α and starting from any prices p_0, if agents of type A, B, and C submit schedules $\hat{D}_{BA}(p_{t-1}, p)$, $\hat{D}_{CB}(p_{t-1}, p)$, and $\hat{D}_{AC}(p_{t-1}, p)$ respectively, then prices will converge to the unique competitive equilibrium of the economy over successive periods.*

Proof. The supply schedule $\hat{S}_{YX}(p_{t-1}, q) = q\hat{D}^{-1}_{XY}(p_{t-1}, q)$ is readily calculated as

$$\hat{S}_{YX}(p_{t-1}, q) \;=\; \frac{\alpha + p_{t-1}}{1 + p_{t-1}} - \alpha q$$

The amount of A traded, q^A, as a function of past prices p^A_{t-1} and p^B_{t-1}, follows by solving $q^A = S_{AB}(p^B_{t-1}/p^A_{t-1}, S_{BC}(1/p^B_{t-1}, S_{CA}(p^A_{t-1}, q^A)))$. Analogous expressions apply to q^B and q^C. Period t prices are then given by $p^A_t = q^C(p^A_{t-1}, p^B_{t-1})/q^A(p^A_{t-1}, p^B_{t-1})$ and

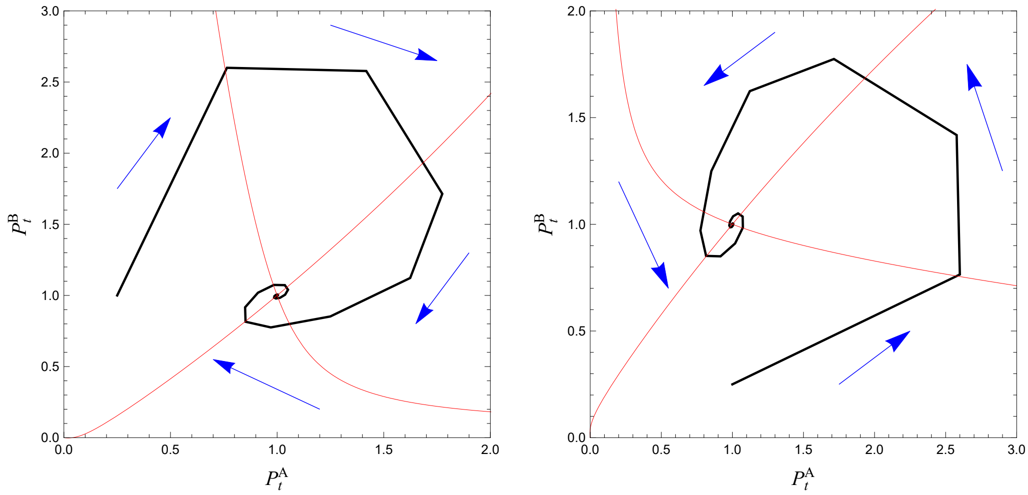


Figure B1. The left panel shows the phase portrait for the dynamical system (Equation B.1) when $\alpha = 5$ together with a typical price path. The right panel shows the same for the reciprocal value $\alpha = \frac{1}{5}$.

$p_t^B = q^C(p_{t-1}^A, p_{t-1}^B)/q^B(p_{t-1}^A, p_{t-1}^B)$. This yields the following discrete dynamical system

$$\begin{pmatrix} p_t^A \\ p_t^B \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} g(p_{t-1}^A, p_{t-1}^B, \alpha) \\ g(p_{t-1}^B, p_{t-1}^A, \frac{1}{\alpha}) \end{pmatrix} \tag{B.1}$$

where

$$g(p,q,\alpha) = (1-\alpha^2)\frac{(p^2-q)(1+q)+\alpha(p+q)(1-pq)}{\alpha^2 p(\alpha+p)+q^2(1+p-\alpha^2(1-\alpha))+(1+p)q(1-\alpha(1-p)+\alpha^3)}$$

It is readily verified that $p^A = p^B = 1$ is the unique rest point, or equilibrium, of the system (Equation B.1). Note that for $\alpha = 1$, prices converge to equilibrium in one iteration and that the dynamical system for $\alpha > 1$ is identical to that for $1/\alpha < 1$ with p_t^A and p_t^B interchanged. In other words, the price paths that occur for $1/\alpha$ are obtained by mirroring the price paths for α in the 45-degree line. Hence, we can focus on the case $\alpha > 1$.

The left panel of Figure B1 shows a typical phase portrait for the system (Equation B.1) when $\alpha > 1$. The upward sloping thin curve corresponds to price pairs where p^A does not change and the downward sloping thin curve corresponds to price pairs where p^B does not change. Their unique intersection corresponds to the equilibrium point $(1,1)$. The arrows indicate the direction of price changes in each of the four regions. The type of price dynamics that is consistent with this phase portrait is (i) convergence to the unique equilibrium, (ii) a limit cycle around the unique equilibrium, or (iii) divergence of prices.

In the latter two cases, the unique equilibrium has to be locally unstable. But a direct computation shows that the eigenvalues of the linearized system around $(p^A, p^B) = (1, 1)$ have squared norm equal to

$$\frac{3}{4}\frac{(1-\alpha)^2}{\alpha + (1-\alpha)^2} < 1$$

So the unique equilibrium is locally stable, and, hence, the phase portrait implies that only (i) can occur, i.e. the system is globally stable. The black piecewise linear curve in the left panel shows an example of a price path that starts at $(p^A, p^B) = (\frac{1}{4}, 1)$ and cycles inwards to the unique equilibrium in a clockwise fashion.

The right panel of Figure A1 shows the phase portrait, the curves that define no change, and a price path starting at $(p^A, p^B) = (1, \frac{1}{4})$ for the reciprocal value of $\alpha = \frac{1}{5}$. Note that the right panel can be simply obtained by mirroring the left panel in the 45-degree line. Also in this case, prices converge to the unique equilibrium but now in a counter-clockwise fashion. $\square$

Appendix C. Replication of Anderson et al. (2004)

In this appendix we discuss in detail Results 1, 2, and 4 from the Anderson et al. (2004) study. These three results pertain to their "cycling" treatment I, which we used for our experiments. Anderson et al. also consider other treatments, which we did not replicate. In this appendix we only consider data from our double auction market sessions.

The first result concerns changes in average prices between periods.

> **Anderson et al. (2004) Result 1.** Changes in price have the same sign as own-market excess demand.

We replicate this result. Following Anderson et al, the excess demand in each market was computed using the average prices in period t. The sign of the excess demand was compared to the sign of the change in average prices in each market between period t and period $t+1$. Data from the markets for good A and good B were pooled. Of the 170 price changes, 131 (77.1%) had the same sign as excess demand. The p-value of a one-tailed sign test of the null hypothesis that the direction of price changes is independent of excess demand is less than 10^{-6}. $\square$

The second result concerns whether prices converge to the competitive equilibrium prices.

> **Anderson et al. (2004) Result 2.** In treatments in which the Scarf model predicts orbits: (i) average prices near the end of the experimental sessions are not close to the equilibrium prices; (ii) prices exhibit no movement toward the equilibrium; (iii) average prices not close to the equilibrium prices are observed moving in the direction predicted by the orbiting model.

We replicate these results. (i) Average prices are said to be close to equilibrium if the prices are in the range $(p_A, p_B) \in [36.5, 43.5] \times [16.5, 23.5]$. Attention is restricted to the

Group	1	2	3	4	5	6	Pooled
β_A^A	0.54	1.64	0.73**	1.77**	1.08**	1.26*	0.76**
	(0.29)	(0.81)	(0.24)	(0.25)	(0.24)	(0.51)	(0.17)
β_A^B	-1.58**	2.03	2.80**	-0.02	2.00*	-0.16	0.28
	(0.28)	(1.45)	(0.81)	(0.24)	(0.84)	(0.53)	(0.32)
# Obs	15	10	15	15	15	15	85
β_B^A	-0.09	0.21*	0.05	0.70*	-0.04*	0.07	0.06
	(0.18)	(0.06)	(0.03)	(0.28)	(0.02)	(0.05)	(0.03)
β_B^B	0.60**	0.83**	0.50**	0.89**	0.42**	0.51**	0.59**
	(0.18)	(0.12)	(0.12)	(0.27)	(0.08)	(0.05)	(0.06)
# Obs	15	10	15	15	15	15	85

Table C1. Estimating a more general price adjustment model using between-period price changes. Prices are normalized so that the competitive equilibrium is $(1, 1)$ and excess demand for a good is normalized by dividing by the total quantity of that good in the economy. Standard errors are in parentheses; * indicates $p < 0.05$, ** indicates $p < 0.01$.

last seven periods of each experimental session. In none of these 42 final seven periods are average prices close to the Walrasian equilibrium. (ii) This is our Result 1, and support for this result is discussed in Section 5.1. Anderson et al. (2004) consider a slightly more general price adjustment model that allows for different speeds of adjustment in the markets for good A and good B. To analyze whether this makes a difference we estimate

$$\begin{aligned} p_A^{t+1} - p_A^t &= \beta_A^A E_A(p_A^t, p_B^t) + \beta_A^B E_B(p_A^t, p_B^t) + \epsilon \\ p_B^{t+1} - p_B^t &= \beta_B^A E_A(p_A^t, p_B^t) + \beta_B^B E_B(p_A^t, p_B^t) + \epsilon \end{aligned}$$

The results are shown in Table C1. First, the speed of adjustments β_A^A and β_B^B are not significantly different. Second, in both markets, price adjustments depend only on the excess demand in that market, i.e. β_A^B and β_B^A are not significantly different from 0. (iii) The clock-hand test and the quadrant test can be used to test whether the observed price paths are consistent with orbiting. Pooling all periods, in 79 of the 85 periods, the clock-hand direction of the price change was in the direction predicted by the orbiting model ($p < 10^{-6}$ under the null hypothesis that clockwise and counter-clockwise movements are equally likely). Similarly, using the quadrant test, in 48 of the 85 periods, the price change was in the quadrant predicted by the orbiting model ($p < 10^{-6}$ under the null hypothesis that movements into each of the four quadrants are equally likely). □

The next result pertaining to the cycling treatment concerns price changes within periods rather than between periods and compares three models of price adjustment.

> **Anderson et al. (2004) Result 4.** In the orbiting treatments, trade-to-trade price movements (i) are not more consistent with simple convergence than with the Scarf model; (ii) are not more consistent with simple convergence than responding proportional to instantaneous excess demand; and (iii) are not more consistent with responding to instantaneous excess demand than with excess demand calculated at initial endowments (the Scarf model).

	Market A	Market B
(I) Price Changes		
#price changes/#transactions	1121/1679 (66.8%)	1892/2863 (66.1%)
(II) Correctly predicted sign of price change		
Convergent model	525/1121* (46.7%)	884 /1892** (46.7%)
Scarf model	748/1121*** (66.7%)	1096 /1892*** (57.9%)
Instantaneous excess demand model	766/1121*** (68.3%)	1114/1892*** (58.9%)
(III) Comparing Models		
Scarf vs. Convergent	429/649*** (66.1%)	698/1227*** (56.9%)
Instantaneous excess demand vs. Convergent	419/609*** (68.9%)	647/1107*** (58.4%)
Instantaneous excess demand vs. Scarf	31/45* (68.9%)	100/182 (54.9%)

Table C2. Trade-to-trade sign test results based on within-period prices. For parts (II) and (III), the null hypothesis that the sign of the price change is correctly predicted with probability 0.5 is tested against a two-sided alternative. * indicates $p < 0.05$, ** indicates $p < 0.01$, *** indicates $p < 0.001$.

We replicate a slightly stronger versions of these results. We find that the simple convergence model predicts the sign of price changes considerably worse than the other two models and that the instantaneous excess demand model does slightly better than the Scarf model. The results of a series of sign tests are reported in Table C2. Part (I) of the table shows that the price changed after approximately two thirds of the transactions in both markets. Part (II) considers the transactions where the price did change and reports the number of times each of the three models correctly predicts the sign of the price change. Part (III) compares pairs of models. Attention is restricted to transactions where the price changed and two models predict different signs for the price change. The number of times the first named model of the pair correctly predicts the price change is reported. In both markets, the simple convergence model does worst. The Scarf and instantaneous excess demand models perform similarly, with the instantaneous excess demand model doing slightly better. □

References

Anderson, C.M., Plott, C.R., Shimomura, K.I.I., Granat, S., 2004. Global instability in experimental general equilibrium: the scarf example. J. Econ. Theory 115, 209 – 249.

Asparouhova, E., Bossaerts, P., Ledyard, J.O., 2011. Price formation in continuous double auctions; with implications for finance. Technical Report. Working paper, University of Utah.

Cashin, P., McDermott, C.J., 2002. The long-run behavior of commodity prices: small trends and big variability. IMF Staff Pap. , 175–199.

Cason, T.N., Friedman, D., 1997. Price formation in single call markets. Econometrica 65, pp. 311–345.

Cason, T.N., Plott, C.R., 2014. Misconceptions and game form recognition: Challenges to theories of revealed preference and framing. J. Polit. Economy 122, pp. 1235–1270.

Crockett, S., 2008. Learning competitive equilibrium in laboratory exchange economies. Econ. Theory 34, 157–180.

Crockett, S., Oprea, R., Plott, C., 2011. Extreme walrasian dynamics: The gale example in the lab. Amer. Econ. Rev. 101, 3196–3220.

Crockett, S., Spear, S., Sunder, S., 2008. Learning competitive equilibrium. J. Math. Econ. 44, 651–671.

Friedman, D., 1984. On the efficiency of experimental double auction markets. Amer. Econ. Rev. 74, pp. 60–72.

Friedman, D., 1993. How trading institutions affect financial market performance: some laboratory evidence. Econ. Inquiry 31, 410–435.

Friedman, D., Ostroy, J., 1995. Competitivity in auction markets: An experimental and theoretical investigation. Econ. J. 105, pp. 22–53.

Friedman, D., Rust, J. (Eds.), 1993. The Double Auction Market: Institutions, Theories, and Evidence. volume 14. Westview Press.

Gale, D., 1963. A note on global instability of competitive equilibrium. Naval Research Logistics Quarterly 10, 8187.

Hahn, F.H., Negishi, T., 1962. A theorem on non-tatonnement stability. Econometrica 30, pp. 463–469.

Hirota, M., Hsu, M., Plott, C.R., Rogers, B.W., 2005. Divergence, closed cycles and convergence in scarf environments: Experiments in the dynamics of general equilibrium systems. Working Papers 1239. California Institute of Technology, Division of the Humanities and Social Sciences.

Keisler, H.J., 1995. Approximate tatonnement processes. Econ. Theory 5, 127–173.

Keisler, H.J., 1996. Getting to a competitive equilibrium. Econometrica 64, pp. 29–49.

Klemperer, P.D., Meyer, M.A., 1989. Supply function equilibria in oligopoly under uncertainty. Econometrica 57, pp. 1243–1277.

McCabe, K.A., Rassenti, S.J., Smith, V.L., 1990. Auction institutional design: Theory and behavior of simultaneous multiple-unit generalizations of the dutch and english auctions. Amer. Econ. Rev. 80, pp. 1276–1283.

Ostroy, J.M., 1980. The no-surplus condition as a characterization of perfectly competitive equilibrium. J. Econ. Theory 22, 183 – 207.

Plott, C.R., George, G., 1992. Marshallian vs. walrasian stability in an experimental market. Econ. J. 102, pp. 437–460.

Scarf, H., 1960. Some examples of global instability of the competitive equilibrium. Int. Econ. Rev. 1, pp. 157–172.

Smale, S., 1976a. A convergent process of price adjustment and global newton methods. J. Math. Econ. 3, 107 – 120.

Smale, S., 1976b. Dynamics in general equilibrium theory. Amer. Econ. Rev. 66, pp. 288–294.

Smale, S., 1998. Mathematical problems for the next century. The Mathematical Intelligencer 20, 7–15. 10.1007/BF03025291.

Smith, V.L., 1962. An experimental study of competitive market behavior. J. Polit. Economy 70, pp. 111–137.

Smith, V.L., 2010. Theory and experiment: What are the questions? J. Econ. Behav. Organ. 73, 3 – 15. On the Methodology of Experimental Economics.

Smith, V.L., Williams, A.W., Bratton, W.K., Vannoni, M.G., 1982. Competitive market institutions: Double auctions vs. sealed bid-offer auctions. Amer. Econ. Rev. 72, pp. 58–77.