

STABILIZING THE ECONOMY: MARKET DESIGN AND GENERAL EQUILIBRIUM

Jacob K. Goeree and Luke Lindsay*

September 4, 2012

Abstract

We employ laboratory methods to study 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). We show that for the Scarf economy, submitting a competitive schedule, i.e. a set of quantities that maximize utility taking prices as given, is a weakly dominant strategy. The resulting outcome corresponds to the unique competitive equilibrium of the Scarf economy. In experiments that employ the schedule market, prices do not oscillate but instead converge quickly to the competitive equilibrium. Besides stabilizing prices, the schedule market is more efficient and results in highly egalitarian outcomes.

JEL Codes: C92, D50

Keywords: *Scarf economy, tatonnement, global Newtonian dynamic, instability, general equilibrium, market design*

*Chair for Organizational Design, Department of Economics, University of Zürich, Blümlisalpstrasse 10, CH-8006, Zürich, Switzerland. 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.

1. Introduction

General equilibrium theory embodies our most complete account of the nature of competitive markets. This theory, which originated with the work of Marie-Esprit-Léon Walras in the late nineteenth century and was further developed by Kenneth J. Arrow, Gérard Debreu, and Lionel W. McKenzie in the 1950s, attempts no less than to predict prices and outputs for the entire economy. General equilibrium theory has become an essential ingredient in both theoretical and applied economic modeling.¹ Nowadays, governments and international financial institutions such as the World Bank regularly employ computational general equilibrium models to evaluate the impact of alternative policies.

Despite its central role in economic theory and its widespread use in practice, general equilibrium theory concerns only the prices and quantities that prevail “after the dust has settled,” i.e. when the economy has come to rest. It is a static theory that does not address how the economy gets to a particular rest point nor does it predict what happens if the economy is subject to a temporary shock. Without a better understanding of the dynamic process by which the economy adjusts and the conditions under which it stabilizes, it is hard to assess the usefulness of general equilibrium theory for predicting behavior in competitive markets.

The issue of dynamic stability already features in the work by Walras (1874) who introduced a centralized price adjustment process, called “tatonnement.” In the tatonnement model, a Walrasian auctioneer adjusts prices in response to reported excess demands until equilibrium prices are reached after which trade occurs. Positive results regarding the stability of the tatonnement process have been obtained under the assumption of gross substitutability (e.g. Kenneth J. Arrow and Frank H. Hahn, 1971). However, Herbert Scarf’s (1960) example demonstrates that without this strong assumption, global stability of the tatonnement process cannot be guaranteed. In Scarf’s economy, tatonnement theory predicts that prices cycle along a closed orbit around the equilibrium without ever converging.

Besides this negative theoretical result, the tatonnement model has had limited empirical appeal, ostensibly because the assumption of centralized price adjustment and no trade at non-equilibrium prices is at odds with how real markets operate (see, e.g., Darrell Duffie and Hugo F. Sonnenschein, 1989). Recent laboratory evidence, however, cautions against an early dismissal of tatonnement theory. In a series of fascinating experiments, Christopher M. Anderson et al. (2004) implement a version of Scarf’s economy in the laboratory to study how prices evolve

¹For instance, general equilibrium theory has become the workhorse model in modern finance and macroeconomics. See, e.g., Charles Noussair et al. (1995) for an experimental test of general equilibrium models as applied to international trade.

in the commonly used double auction market. While the double auction trading institution is itself a distinctively non-tatonnement mechanism, Anderson et al. (2004) find strong support for the Walrasian hypothesis that price dynamics are largely driven by a market's excess demand. Average trade prices in the experiments cycle along a closed orbit around the unique competitive equilibrium with no clear sign of convergence.

The experimental findings by Anderson et al. (2004) have far reaching implications. First, they provide empirical proof that price adjustments in markets are governed by "laws of motion." Remarkably, the tatonnement model can be used to derive these laws even for market environments that do not fit its assumptions.² Second, the experiments confirm the main prediction of the Scarf model, i.e. that economic systems can be perpetually out of equilibrium. Obviously, such global instability may have dramatic consequences for the performance of the system, both in terms of overall welfare and the equitable distribution of the gains from trade.

The Anderson et al. (2004) study exemplifies one of the major strengths of laboratory experimentation, namely control. 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 laboratory experimentation is replicability, which seems especially appropriate for the Anderson et al. (2004) study given the importance of their findings. A third major strength of laboratory experimentation is that it offers the possibility to test newly designed institutions 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.

The motivation for our study is thus twofold. First, to replicate the experiments of Anderson et al. (2004). Second, to design a market institution that stabilizes the Scarf economy and forces prices to converge to competitive equilibrium levels. Creating market institutions that

²Hahn and Takashi Negishi (1962) introduce a model of price adjustment with centralized price setting but trading at non-equilibrium prices. A set of prices is called. 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.

Jerome Keisler (1995; 1996) introduces a model with both decentralized price setting and trading at non-equilibrium prices. 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.

Sean Crockett, Stephen Spear, and Shyam Sunder (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.

produce correct and stable prices is important since prices determine traders' opportunity sets as well as how the gains from trade are divided. The problem of stability is non-trivial though, and existing candidate market institutions cannot be expected to solve it. A major challenge is that for the Scarf economy, adjusting prices based on excess demands does not result in convergence even when the economy is "scaled up" by introducing many replicas of each agent type. In other words, even when the impact of every agent in the economy is "small," the truthful revelation of excess demands at current prices does not lead to convergence.

Our novel design finds its origin in 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.³ The main question is how to design a market that implements Smale's Newtonian dynamic.⁴ One could introduce a "Newtonian auctioneer" who adjusts prices given reported excess demands according to the global Newton method. One complication is that the Newton method requires information not only about excess demands but also about their derivatives. In addition, as noted above, the notion of a fictitious auctioneer (i) announcing prices, (ii) agents reporting demand, and (iii) the auctioneer adjusting prices until demand equals supply, all before any trade takes place, lacks realism. 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) using the global Newton method.

Our proposed solution is therefore a call market where agents submit schedules. Submitting demand schedules is a common feature of electricity markets 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 Vernon 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.⁵

Whether our proposed design produces desirable outcomes obviously depends on the types of schedules that get submitted. We show that for the Scarf economy it is a weakly dominant strategy for all agents to submit a competitive schedule, i.e. a set of quantities that are utility

³Newton's classical method of iteration corresponds to a discrete approximation to Smale's adjustment process.

⁴The 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. For a related set of double auction market experiments, Masayoshi Hirota et al. 2005 indeed report that they find no support for the Newtonian model.

⁵For a theoretical treatment of schedule markets, see, for instance, Paul Klemperer and Margaret Meyer (1989) who provide a general analysis of oligopoly markets where firms submit (supply) schedules.

maximizing taking prices as given. A fortiori, all agents submitting competitive schedules constitutes an equilibrium and the resulting prices and quantities correspond to the unique competitive equilibrium of the Scarf economy. Furthermore, when others submit competitive schedules, the supply curve that each agent faces is “flat.” This is a non-generic feature that follows from the specific parametrization of the Scarf economy. 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. To summarize, while scaling up the economy by replicating agents has no bite under the tatonnement dynamic it makes submitting competitive schedules approximately optimal in our design, resulting in competitive equilibrium outcomes.

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. The Appendix 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 proportional to its excess demand. In vector notation, $dp(t)/dt = z(p)$, or written out in components

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

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 side is the supply of apples by type A and the second term is the demand for apples by type C . Likewise, in the second line, the first term between parentheses represents the supply of bananas by type B and the second term is the demand for bananas by agent A . 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

$$L(p_A, p_B) = 1 - p_A p_B \exp\left(1 - \frac{1}{2}p_A^2 - \frac{1}{2}p_B^2\right)$$

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

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

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. ■

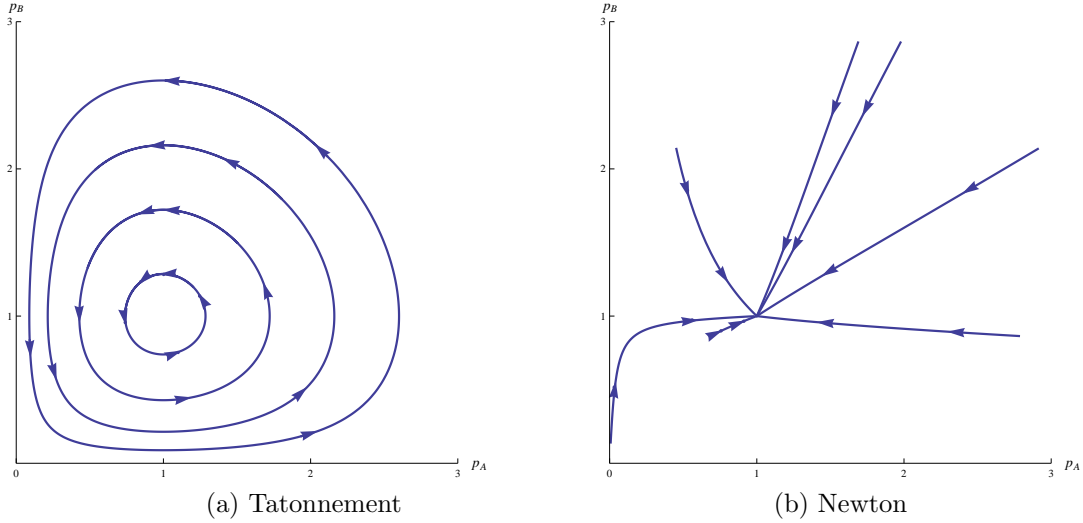


Figure 1. Predicted price patterns under the tatonnement dynamic (a) and the global Newtonian dynamic (b) in the Scarf economy. For the tatonnement model, prices cycle in a counter-clockwise manner without converging. In contrast, for the Newtonian dynamic, prices converge exponentially fast to the unique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

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:^{6,7}

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

with $\nabla z(p)$ the matrix of partial derivatives of $z(p)$ with respect to p .⁸

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

Proof. Consider the Lyapunov function

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

⁶Newton's method for solving $f(x) = 0$ for some $f : \mathbb{R} \rightarrow \mathbb{R}$ can be recovered by taking a discrete approximation: $x_{n+1} - x_n = -f(x_n)/f'(x_n)$.

⁷In 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.

⁸Written out in components, $(\nabla z(p))_{ij} = \partial z_i(p)/\partial p_j$.

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). ■

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 (1976a,b) 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 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.⁹ 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 procedure in that prices increment

⁹Convergence in single-commodity markets occurs under a wide variety of conditions, see e.g. Smith (1962), Dan Friedman (1984), and Smith (2010).

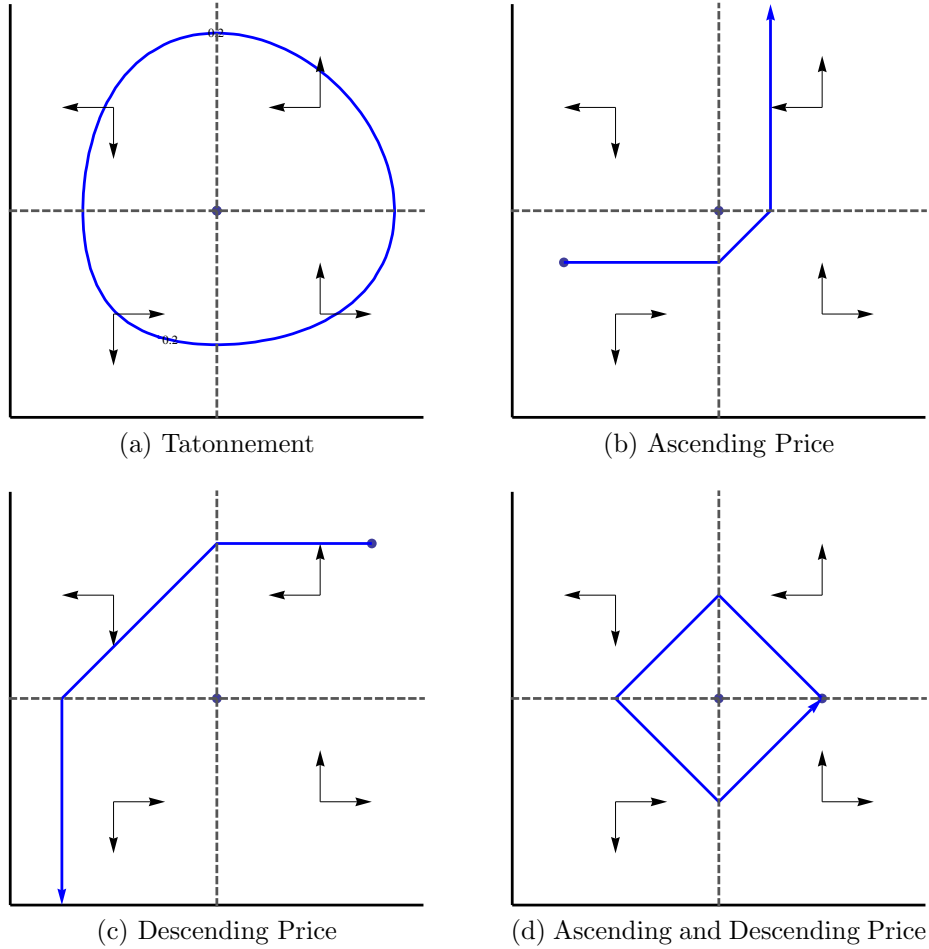


Figure 2. Predicted price patterns under various tatonnement-like procedures: (a) the Walrasian auctioneer (upper-left), (b) the ascending English auction (upper-right), (c) the descending Dutch auction (lower-left), and (d) a combination of the Dutch and English auction (lower-right).

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 increment 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 a Walrasian auctioneer who announces a set of prices and participants 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.¹⁰ 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 a mechanism, 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 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 solution is to leverage the convergence properties of the global Newton method in a market where participants submit demand schedules. 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.¹¹ The clearing mechanism determines the terms of trade by applying the global Newtonian method to the submitted schedules to find equilibrium prices.

3.1. A Schedule Market

In a schedule market, participants report their demand at a series of prices, then equilibrium prices are computed such that supply equals demand for all goods. Since participants only report demand at a finite set of prices, interpolation is used to estimate demand at intermediate prices. To make solving for general equilibrium tractable, 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 positive prices to strictly positive quantities. To ensure uniqueness, the following restrictions are applied to the submitted demand schedules.

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

¹⁰See Smith et al. (1982) and Plott and George (1992) for experimental evidence on the Walrasian auctioneer model.

¹¹McCabe 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. Cason and Friedman (1997) conduct an experimental investigation of price formation in call markets.

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.

(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 $S_{YX}(q) = qD_{XY}^{-1}(q)$ 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.¹² 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 submitted schedules are admissible, the amount of each good being traded is uniquely determined. If trade occurs, 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 $p_B = q_C/q_B$. ■

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 is¹³

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

¹²Evaluating $S'_{YX}(q)$ 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.

¹³Maximizing $\min(q_X, q_Y)$ over the budget set $pq_X + q_Y \leq 1$ yields $q_X^* = q_Y^* = 1/(1+p)$.

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 $\min(q_X, q_Y)$.

Since the market-clearing price generally depends on the schedules submitted, competitive schedules might not be optimal. 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, $p = p_B/p_A$, which is how it was presented to subjects on their result screens after the period ended (see, e.g., the blue line in the lower panel of Figure 4 below). 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 $i = 1, \dots, n$. 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. ■

Submitting a competitive schedule is not necessarily weakly dominant without the admissibility restriction, but doing so remains 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 5. *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. ■

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, Propositions 4 and 5 apply more generally when the economy grows large.

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

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

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 used originally 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 are 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.¹⁴ 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. Subjects completed a short comprehension test and then played three practice periods. During the practice periods, subjects were encouraged to ask questions. An exchange rate of 0.15 Swiss Francs per util was used. The experiment took around 1.5 hours to complete. The mean payment was 48.74 Swiss Francs including a ten Franc showup fee.

4.1. Continuous Double Auction

A screenshot of the continuous double auction interface is shown in Figure 3. The left side of the screen shows the subject’s utility function, current holdings, and is used to construct orders. The right side shows a list of submitted orders, some of which have already transacted.

¹⁴The double auction market Session 2 lasted for only ten periods because of a computer crash.

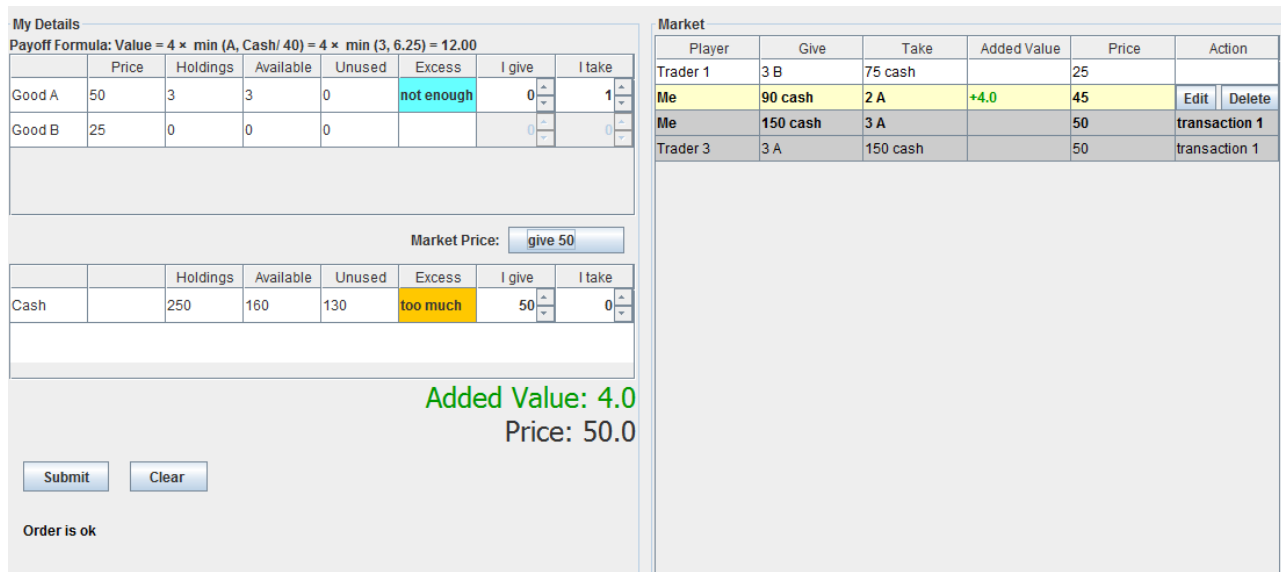


Figure 3. User interface for the continuous double auction market. The screen is from the point of view of a type-*C* agent who was endowed with cash and needs cash and good *A*. On the top left of the screen, the text beginning ‘Payoff Formula’ shows the subject’s utility function and the value of the 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*.

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. Each period lasted for four minutes. 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 on screen their earnings for the period, a list of the trades they made, and their total earnings from all completed periods.

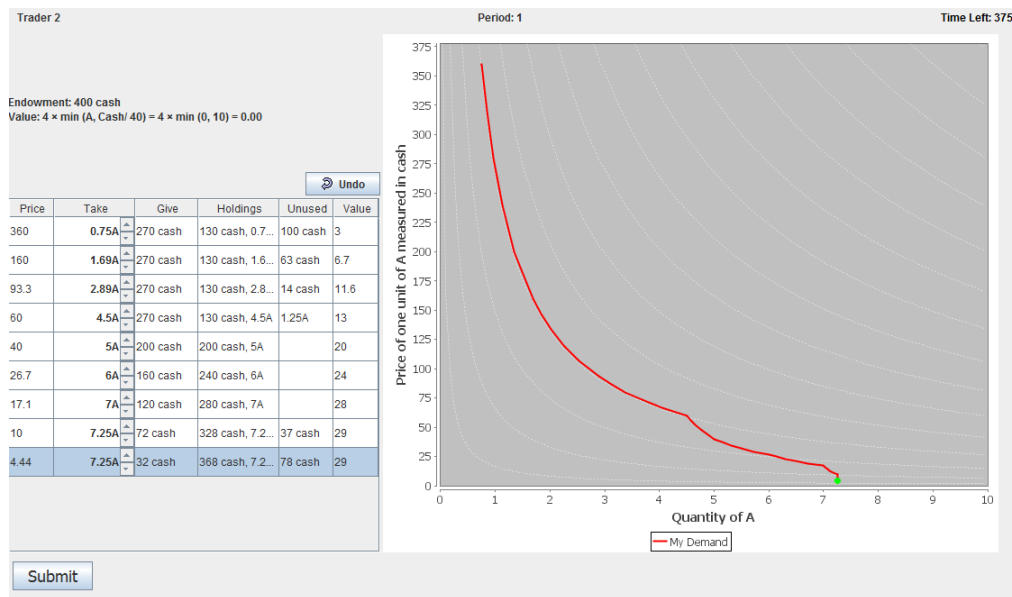


Figure 4a. User interface for the schedules market. The left side of the screen shows the subject’s endowment, utility function, and a table that can be used to create demand schedules. The first column of the table is a fixed list of possible prices of the commodity that the subject needs in terms of 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.

4.2. Schedules Market

A screen shot of the schedules 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 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 admissability. A period ended if all schedules had been submitted or if four minutes had elapsed. Figure 4b shows the results of the schedules

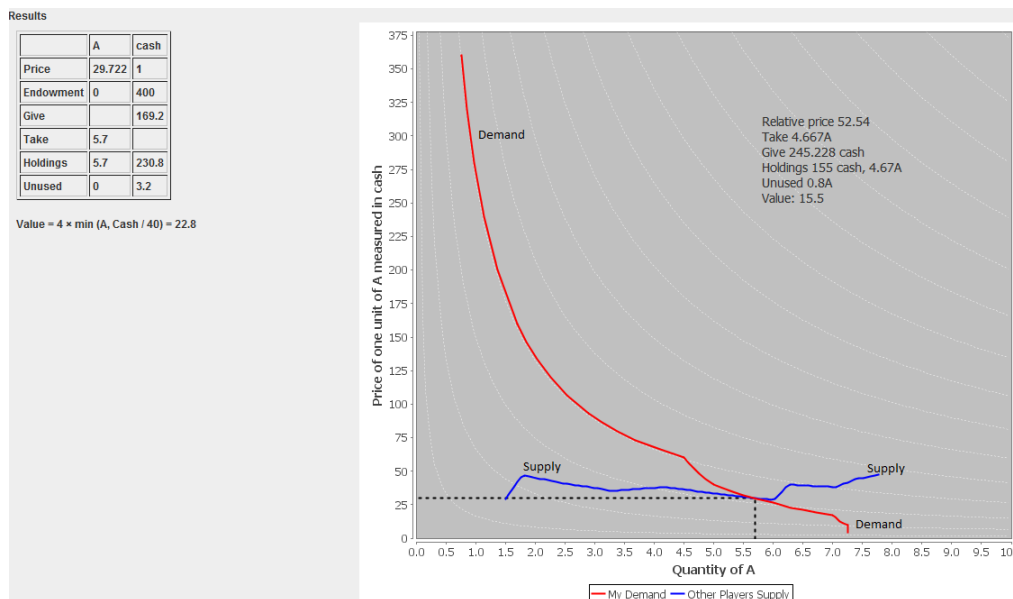


Figure 4b. Results screen for the schedules market. The curve labeled ‘Demand’ is the demand schedule submitted by the subject. The curve labeled ‘Supply’ is the residual supply the subject faces. At each point on the residual supply curve, excess demand is zero in all markets. The intersection of the 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.

market shown to the subjects after the period ended. They 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).

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

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

* indicates $p < 0.05$, ** indicates $p < 0.01$, *** indicates $p < 0.001$, standard errors in parentheses

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.

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 out and the less the price path looks like a circle. Compare, for instance, the first and fourth double auction session 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

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

where L is the Lyapunov function, $\Phi(\cdot)$ is the standard normal cumulative distribution,¹⁵ α_s a session-specific constant that corresponds to the path's starting point (i.e. the trade price in the first period), and β is the parameter that 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. \square

¹⁵We introduce this transformation because the Lyapunov function is bounded between 0 and 1 so a simple linear regression would not be appropriate.

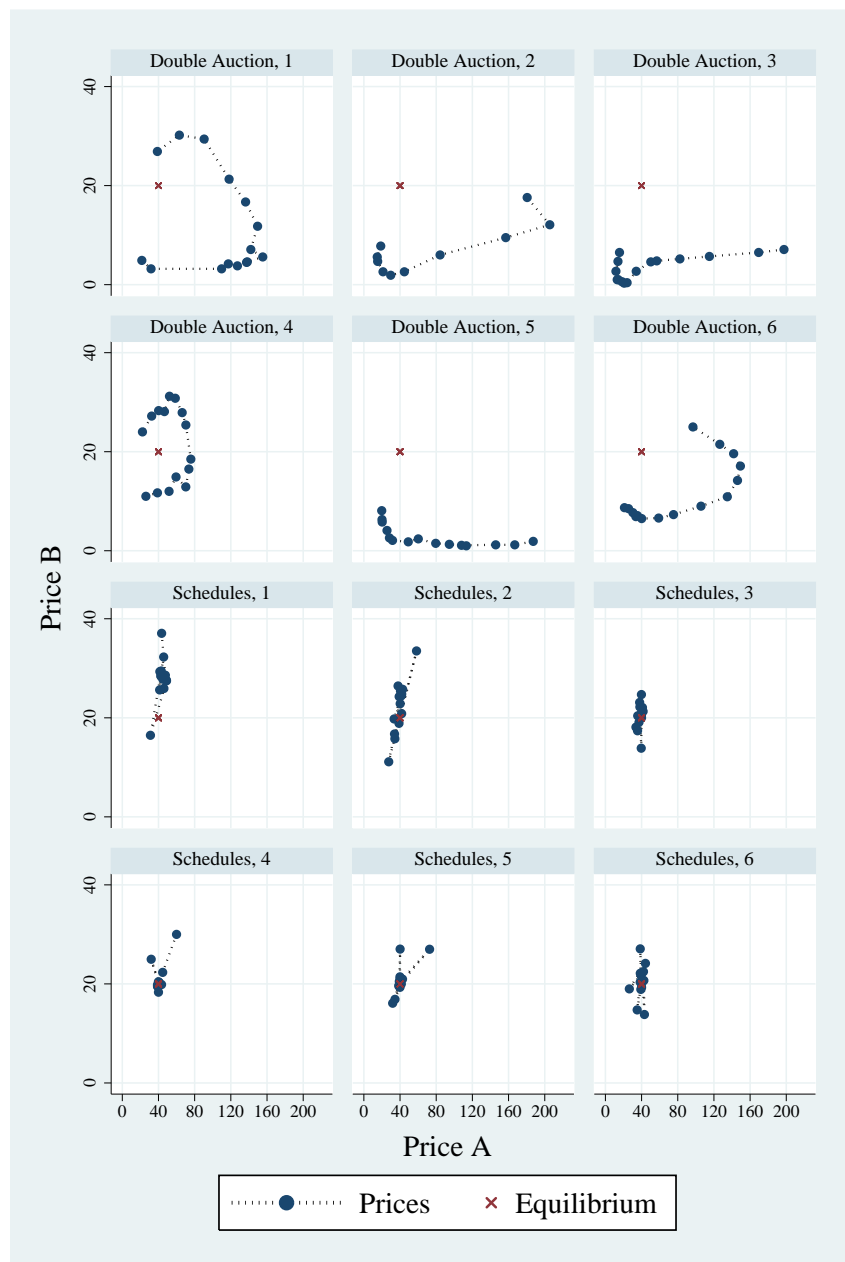


Figure 5. Between-period prices in the double-auction market sessions (top six panels) and the schedule market sessions (bottom six panels). Prices cycle in a counter-clockwise manner in the double auction markets and converge straight to equilibrium in the schedule markets.

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 (1) to the Lyapunov function in Section 2.2. The estimated β coefficient is significantly negative for the schedule market, see 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. 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 $p_B = 22.2(0.5)$. \square

The stark difference in price evolution under the two different market mechanism 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 show a “sine-like” pattern with substantial intra-period variation added. Focusing on the underlying sine-like pattern reveals 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

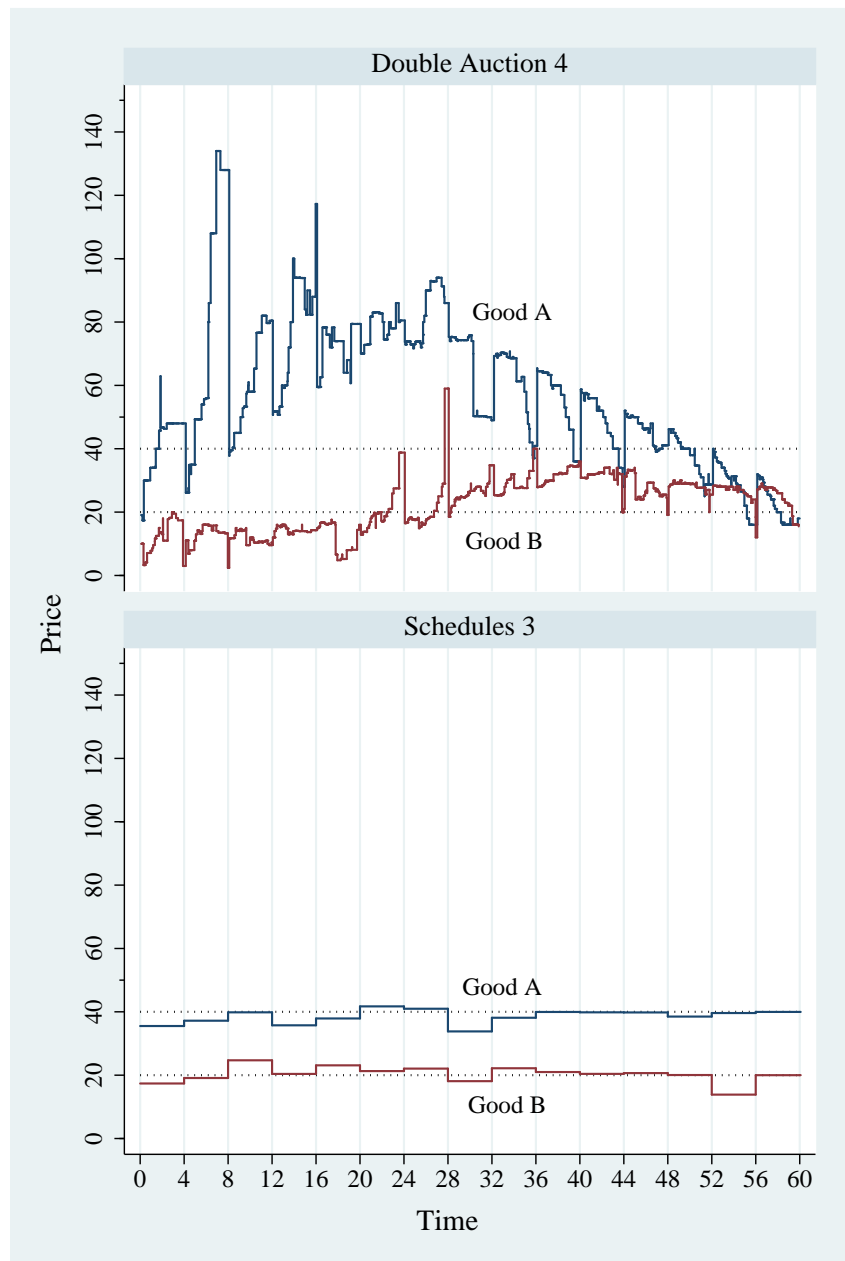


Figure 6. Price evolution in double auction market Session 4 (top) and schedules market Session 3 (bottom). All 15 periods of the experiment are shown, each lasting 4 minutes as indicated by the thin vertical lines. Prices in the double auction market show a sine-like pattern with substantial intra-period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

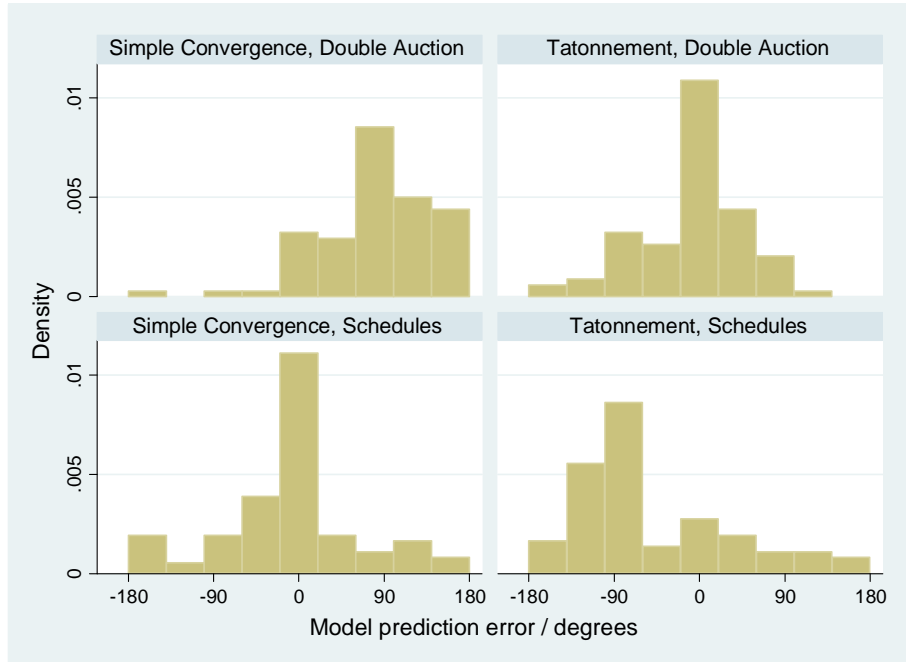


Figure 7. Prediction errors for the simple convergence model (left panels) and the tatonnement model (right panels) in the double auction market (top panels) and the schedule market (bottom panels). For the double auction market, the tatonnement model produces a big spike at zero degrees (no error) while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

(bottom two panels). The histograms are based on the difference between the predicted and observed angles of price changes.¹⁶ 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.

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

¹⁶Angles 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.

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

*** indicates $p < 0.001$, standard errors in parentheses

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.

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. \square

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.

5.2. The Effects of Price (In)stability on Market Performance

In Walras’ tatonnement model, a fictitious 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 amounts traded at these prices.

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

$$\begin{aligned} q_A &= \min\left(\frac{p_B}{p_A + p_B}, \frac{p_B}{p_A + p_A p_B}, \frac{1}{1 + p_A}\right) \\ q_B &= \min\left(\frac{p_A}{p_A + p_B}, \frac{p_A}{p_B + p_A p_B}, \frac{1}{1 + p_B}\right) \\ q_C &= \min\left(\frac{p_A p_B}{p_A + p_B}, \frac{p_B}{1 + p_B}, \frac{p_A}{1 + p_A}\right) \end{aligned}$$

and the resulting welfare is

$$W = 1 + \min\left(\frac{p_A p_B}{p_A + p_B}, \frac{p_B}{p_A + p_A p_B}, \frac{p_A}{p_B + p_A p_B}\right)$$

Welfare is maximized at the competitive equilibrium prices, $p_A = p_B = 1$, and minimized when either price is 0 or ∞ .

Proof. Consider the constrained maximization problem

$$\begin{aligned} \max_{0 \leq q_A, q_B, q_C \leq 1} & \quad (\min(q_B, 1 - q_A) + \min(q_C, 1 - q_B) + \min(q_A, 1 - q_C)) \\ & p_A q_A = q_C \\ & p_B q_B = q_C \\ & 1 - q_A \geq q_B \\ & 1 - q_B \geq q_C \\ & 1 - q_C \geq q_A \end{aligned}$$

The objective is total welfare when (q_A, q_B, q_C) units of goods (A, B, C) are exchanged. The top inequality constraints are the basic resource constraints in the original Scarf economy. The first equality constraint implies that the type- C trader pays the value of how much A she buys in cash and the second equality implies that the amount of cash acquired by the type- B trader equals the value of the amount of B she sells. (The third equality $p_A q_A = p_B q_B$, i.e. the value of how much of good B the type- A trader buys equals the value of how much of good A she sells, is already implied by the other two equality constraints.) The bottom three inequality constraints

reflect that traders do not acquire more of the good they demand than how much they keep of the good they are endowed with. If prices are flexible, the solution to the maximization problem is $q_A = q_B = q_C = \frac{1}{2}$ at prices $p_A = p_B = p_C = 1$. This outcome corresponds to the competitive equilibrium for the Scarf economy and maximizes welfare: $W_{max} = \frac{3}{2}$. For fixed prices, the solutions are as shown in the proposition. Notice that the lowest possible welfare arises when either p_A or p_B is 0 or ∞ , in which case $W_{min} = 1$. ■

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

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

Figure 8 shows observed efficiencies (vertical axis) versus the efficiency levels predicted by the price-taking model (horizontal axis) for the double auction market sessions. The small diamonds correspond to period averages and the large diamonds to session averages. 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. Figure 8 shows that the price-taking model predicts efficiency levels remarkably well.¹⁷ This is especially true for the six session averages, which are all very close to the 45-degree line.¹⁸

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%.¹⁹

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%,

¹⁷The coefficient for correlation between the predicted and observed period averages is 0.770 ($p < 0.0001$).

¹⁸Two sessions resulted in almost identical and observed efficiency levels, which is why it appears as if there are only five large filled diamonds.

¹⁹If all goods are randomly assigned to one type of trader, predicted efficiency is $W_{min} = 1$. A more strict 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%.

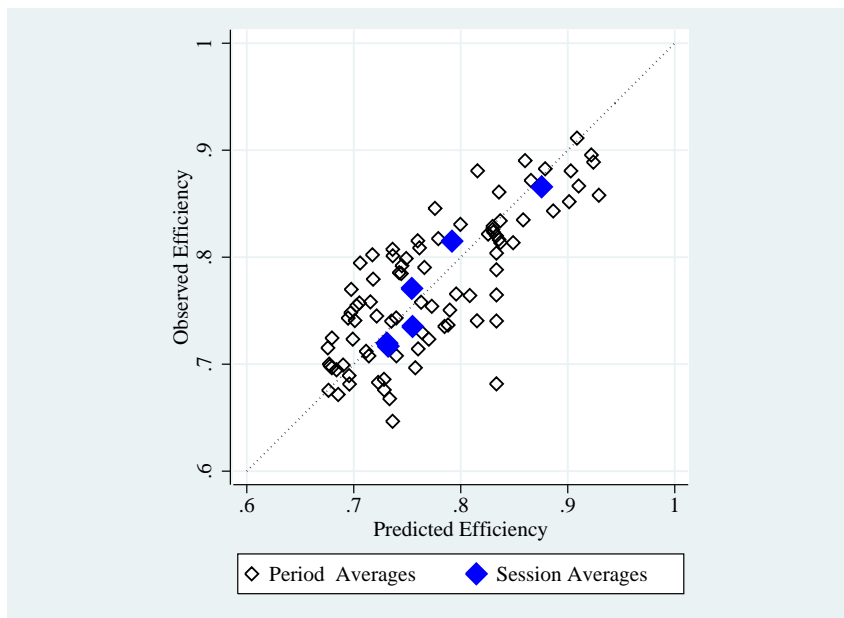


Figure 8. Observed and predicted efficiency levels in the double auction market. The predictions are based on a model of optimizing behavior taking prices as given, with the prices being equal to the opening prices for the period. The opening prices are set to half the equilibrium prices in the first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

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 ($p = 0.0022$). \square

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. 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

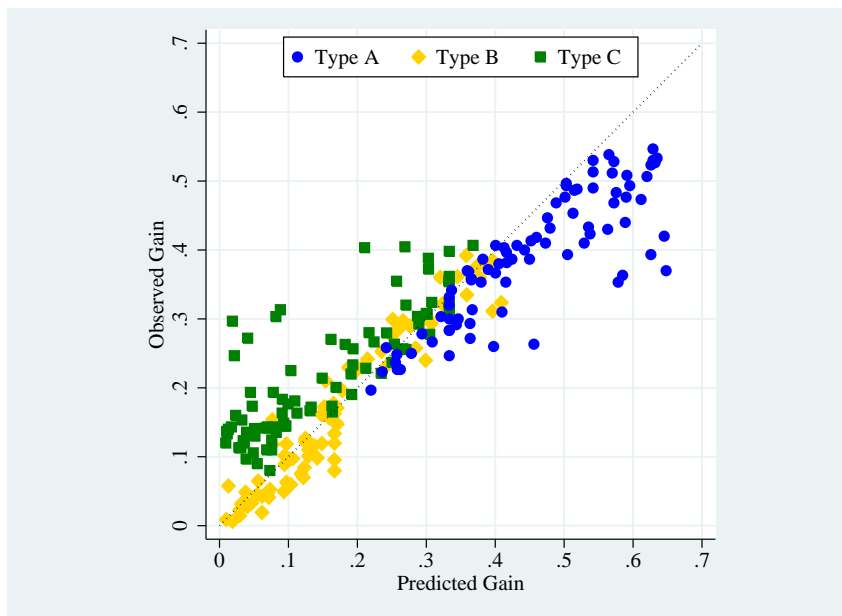


Figure 9. Observed and predicted gains for each agent type in the double auction market. The predictions are based on a model of optimizing behavior taking prices as given and equal to the opening prices for the period. The opening prices are half the equilibrium prices in the first period and the average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

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 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 than predicted (as indicated by the circles in Figure 9). The tendency of type-*A* agents

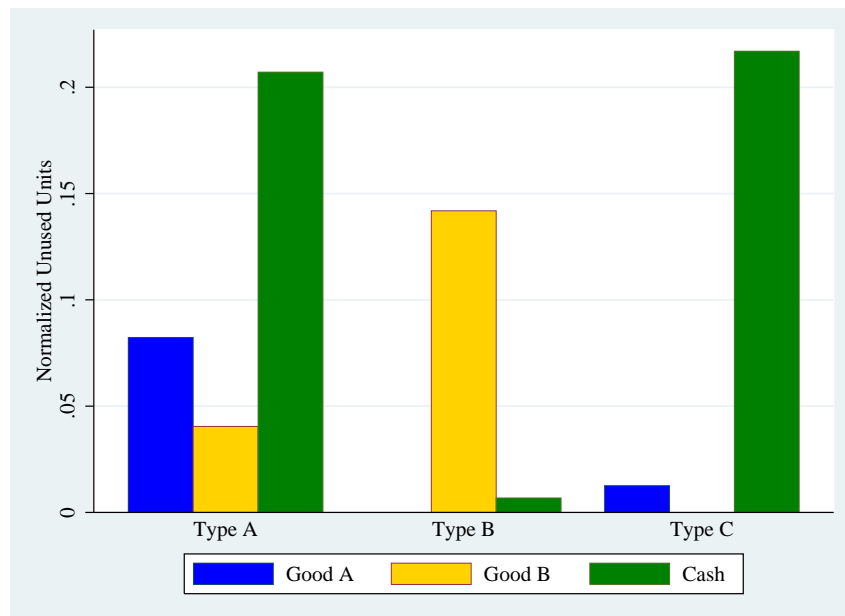


Figure 10. Unused units of each of the three goods by agent type (normalized by the total amount of each good in the economy). The type-*B* and type-*C* agents have unused units only of the goods they demand while the type-*A* agent has unused units of all three goods.

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 diamonds) are all in the lower-left corner while the shares for the type-*A* agents (circles) are in the upper-right corner and the shares of the type-*C* agents (squares) 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).²⁰ 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 fair outcomes.

Support. The division of the total gains from trade among the three types of agent is roughly even (32.3%, 35.5%, 31.2%) in the schedule markets while it is very uneven (51.1%, 20.5%, 28.4%)

²⁰Recall that in the double auction market Session 2 there were only 10 periods due to a computer crash.

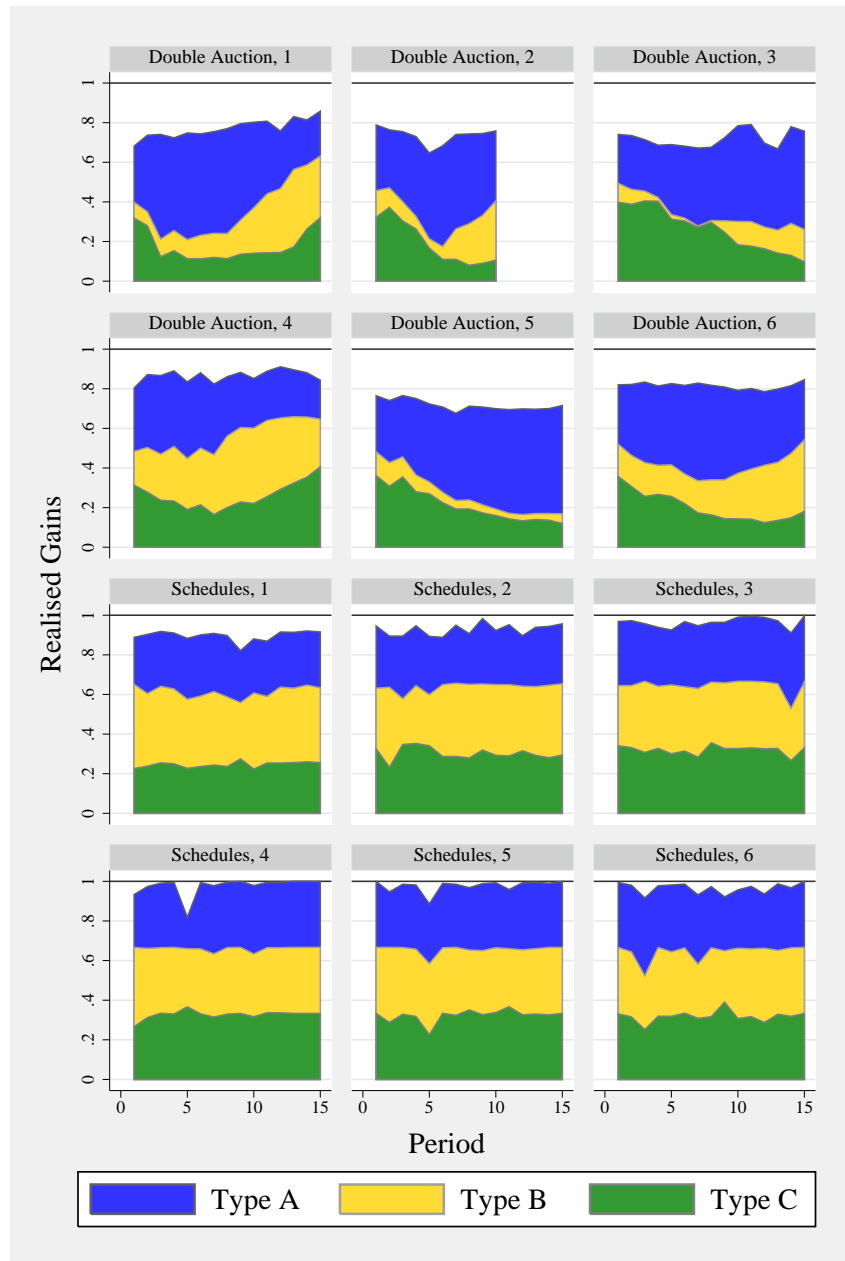


Figure 11. Earnings by agent type in the double auction market sessions (top six panels) and the schedule market sessions (bottom six panels). The white space at the top of each panel indicates the degree to which there was a loss in efficiency. (Recall that double auction market Session 2 lasted for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

in the double auction markets.²¹ 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 outcomes are extremely unfair. See, for instance, double auction market Sessions 3 and 5. \square

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 optimal 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 natu-

²¹The 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 schedules 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.

rally 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 direction of price changes remarkably well and that prices are global unstable as a result. We then ask two important questions that go beyond replication of the Anderson et al. (2004) 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 begs 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 Propositions 4 and 5). 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. Prices are globally stable and 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 process 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. Indeed, our experimental results demonstrate that in settings with complementarities and income effects, institutions based on tatonnement adjustments do not necessarily result in competitive equilibrium outcomes while a call market that admits schedules does.

A. Appendix: 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. (2004) 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 prices 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 results 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 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

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

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

* indicates $p < 0.05$, ** indicates $p < 0.01$, standard errors in parentheses

Table A1. 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.

The results are shown in Table A1. 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). \square

The next result concerns prices 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).

We replicate 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

	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%)

* indicates $p < 0.05$, ** indicates $p < 0.01$, *** indicates $p < 0.001$

Table A2. 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.

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 A2. 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. \square

References

- Anderson, C. M., C. R. Plott, K.-I.-I. Shimomura, and S. Granat: 2004, ‘Global instability in experimental general equilibrium: the Scarf example’. *Journal of Economic Theory* **115**(2), 209 – 249.
- Arrow, K. and F. Hahn: 1971, *General competitive analysis*. San Francisco: Holden-Day.
- Cason, T. N. and D. Friedman: 1997, ‘Price Formation in Single Call Markets’. *Econometrica* **65**(2), pp. 311–345.
- Crockett, S.: 2008, ‘Learning Competitive Equilibrium in Laboratory Exchange Economies’. *Economic Theory* **34**(1), 157–180.
- Crockett, S., S. Spear, and S. Sunder: 2008, ‘Learning Competitive Equilibrium’. *Journal of Mathematical Economics* **44**(7), 651–671.
- Duffie, D. and H. Sonnenschein: 1989, ‘Arrow and General Equilibrium Theory’. *Journal of Economic Literature* **27**(2), pp. 565–598.
- Friedman, D.: 1984, ‘On the Efficiency of Experimental Double Auction Markets’. *The American Economic Review* **74**(1), pp. 60–72.
- Friedman, D.: 1993, ‘How trading institutions affect financial market performance: some laboratory evidence’. *Economic Inquiry* **31**(3), 410–435.
- Hahn, F. H. and T. Negishi: 1962, ‘A Theorem on Non-Tatonnement Stability’. *Econometrica* **30**(3), pp. 463–469.
- Hirota, M., M. Hsu, C. R. Plott, and B. W. Rogers: 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’. *Economic Theory* **5**, 127–173.
- Keisler, H. J.: 1996, ‘Getting to a Competitive Equilibrium’. *Econometrica* **64**(1), pp. 29–49.
- Klemperer, P. D. and M. A. Meyer: 1989, ‘Supply Function Equilibria in Oligopoly under Uncertainty’. *Econometrica* **57**(6), pp. 1243–1277.
- McCabe, K. A., S. J. Rassenti, and V. L. Smith: 1990, ‘Auction Institutional Design: Theory and Behavior of Simultaneous Multiple-Unit Generalizations of the Dutch and English Auctions’. *The American Economic Review* **80**(5), pp. 1276–1283.
- Noussair, C., C. Plott, and R. Riezman: 1995, ‘An Experimental Investigation of the Patterns of International Trade’. *The American Economic Review* **85**(2), pp. 462–491.
- Plott, C. R. and G. George: 1992, ‘Marshallian Vs. Walrasian Stability in an Experimental Market’. *The Economic Journal* **102**(412), pp. 437–460.

- Scarf, H.: 1960, ‘Some Examples of Global Instability of the Competitive Equilibrium’. *International Economic Review* **1**(3), pp. 157–172.
- Smale, S.: 1976a, ‘Dynamics in General Equilibrium Theory’. *The American Economic Review* **66**(2), pp. 288–294.
- Smale, S.: 1976b, ‘Exchange processes with price adjustment’. *Journal of Mathematical Economics* **3**(3), 211–226.
- 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’. *Journal of Political Economy* **70**(2), pp. 111–137.
- Smith, V. L.: 2010, ‘Theory and experiment: What are the questions?’. *Journal of Economic Behavior & Organization* **73**(1), 3 – 15. On the Methodology of Experimental Economics.
- Smith, V. L., A. W. Williams, W. K. Bratton, and M. G. Vannoni: 1982, ‘Competitive Market Institutions: Double Auctions vs. Sealed Bid-Offer Auctions’. *The American Economic Review* **72**(1), pp. 58–77.
- Walras, L.: 1954, ‘Elements of Pure Economics, translated by W. Jaffe (first published in 1874)’.