American Economic Journal: Microeconomics 2017, 9(4): 1–41
https://doi.org/10.1257/mic.20150201

Call Market Experiments: Efficiency and Price Discovery through Multiple Calls and Emergent Newton Adjustments†

By Charles R. Plott and Kirill Pogorelskiy*

We study multiple-unit, laboratory experimental call markets in which orders are cleared by a single price at a scheduled "call." The markets are independent trading "days" with two calls each day preceded by a continuous and public order flow. Markets approach the competitive equilibrium over time. The price formation dynamics operate through the flow of bids and asks configured as the "jaws" of the order book with contract execution featuring elements of an underlying mathematical principle, the Newton-Raphson method for solving systems of equations. Both excess demand and its slope play a systematic role in call market price discovery. (*JEL* C92, D41, D44, G14)

This paper studies the principles governing price discovery and dynamics in call markets. Call markets accumulate orders until a scheduled time at which a "call" takes place, a single market-clearing price is determined, and all exchanges take place at that price. Accumulation of orders over time creates "market depth," which can conceivably lead to reducing the price variability. By contrast, the widely used continuous double auctions are founded on a different architecture in which order flow takes place continuously and the timing of contract executions is endogenous and possibly at different prices.

The three broad research questions are: (i) Do the basic laws of supply and demand operate as they are known to operate in continuous markets? (ii) What are the behavioral principles that guide the price dynamics? (iii) How do the institutions and rules together with behavioral principles operate to guide market performance?

* Plott: Division of the Humanities and Social Sciences, MC 228-77, California Institute of Technology, Pasadena, CA 91125 (email: cplott@hss.caltech.edu); Pogorelskiy: Department of Economics, University of Warwick, Coventry CV4 7AL, UK (email: k.pogorelskiy@warwick.ac.uk). We thank Skylar Cook and Megan Cutrofello, and other students in the Caltech experimental economics class who actively participated in this project in the beginning and contributed to the experimental design. We are grateful to Brian Merlob for helping us monitor the experimental sessions and recording the demonstration video. We express our warm thanks to Pablo Beker, Paul Brewer, Tim Cason, Stefano DellaVigna, David Eil, Dan Friedman, Ben Gillen, Dan Levin, Tibor Neugebauer, Andrew Oswald, Stefan Palan, Tom Palfrey, Charles Sprenger, Rakesh Vohra, Bill Zame, and participants of the 2014 Santa Barbara Conference on Experimental and Behavioral Economics and 2015 Econometric Society World Congress for useful comments. We acknowledge the financial support of the Caltech EEPS laboratory. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper. The laboratory experiments received Institutional Review Board approval from Caltech.

† Go to https://doi.org/10.1257/mic.20150201 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

The call markets studied are organized as an exchange in which agents have multiple units. The analysis is restricted to the "pure" case of price discovery that is not complicated by agent uncertainty about the personal value of the traded item.[1] The underlying flow of incentives to trade arrives in a series of independent periods that include an unannounced structural shift in the market parameters. A period is like a trading day in which two calls occur, prices are determined and exchanges take place.

Our results demonstrate that in the two-call, multiple-unit auctions market prices and volumes converge close but not perfectly to the competitive equilibrium derived from the underlying incentives. Efficiency is relatively high, increases over time and converges to near the competitive equilibrium level. Results are robust to the presence of a structural shift in the market parameters.

The results provide insights about a long-standing mystery of how markets achieve an equilibrium defined as a solution to the equations created by the underlying incentives. An interpretation is that the market "discovers" the solutions to a system of equations that no one in the market knows. The model describes the formation process as working within a period through the multiple market "calls" to create a series of steps of information aggregation and computation leading to price changes and then to ultimate convergence across periods. The model begins with the order flow shaping "market jaws." The price and time priority of the orders placed in the open order book produce a graphical representation (jaws) that approximates the slope of excess demand; and the difference between the number of buy and sell orders arriving shortly before a call price announcement approximates the excess demand at the price. Together, when operating in the multiple calls, the mechanisms exhibit features of the Newton-Raphson method of finding the solutions to a system of equations as will be discussed in detail later. To emphasize the combined model, we will refer to the model of the price discovery process as "Newton-Jaws." We feel compelled to warn the casual reader that this model has nothing to do with the jaws of the outstanding mathematician and physicist Sir Isaac Newton, after whom the respective numerical method is named. We also need to emphasize that relationships are confined to strong similarities and that differences do exist between Newton-Jaws and the Newton method as used in numerical analysis.

Walrasian adjustment, the main alternative to Newton-Jaws, also finds support, but Newton-Jaws performs better by comparison. Simulations with zero-intelligence agents in Appendix B demonstrate that the Newton-Jaws model has a solid foundation: it is a property of the call market institution (together with underlying demand/supply parameters) rather than a consequence of special or idiosyncratic features of traders' strategic behavior.

[1] A long history of research exists on the relationship between information aggregation and institutions. Experimental environments studied range from multiple states and multiple markets (Arrow-Debreu securities), and single markets (winner's curse), to cascades and bubbles. The institutions range from various forms of continuous double auctions, call markets, quote markets, dealer markets, auctions (sealed bid, ascending price, etc.) and special mechanisms designed explicitly for the purpose of information aggregation. Our focus is on call markets with independent values about which there is no uncertainty. Even with the focus so restricted a substantial range of institutions exist.

The remainder of the paper is organized as follows. Section I provides a brief review of the background literature. Section II describes the experimental economic environment. Subsection IIA presents the call market institution we implemented in the lab. Subsections IIB and IIC describe our experimental procedures and the basic parameters, respectively. Sections III and IV describe models and theory behind their application to the actual data. Section V presents our main results. Section VI concludes. Appendices A and B contain additional estimation details, while Appendix C contains experimental instructions.

I. Background and Related Literature

Call markets share institutional features with many other types of markets. The term "call market" or "clearing house" is typically reserved for a complex class of institutions with a designated time for tenders and simultaneous price discovery, operating in environments with multiple buyers and sellers. This class of institutions is large. For example, auctions, including any form of sealed-bid ones, can be viewed as special cases of call markets with a single seller (or buyer).

Major institutional differences aside, the principles that govern the behavior of call markets potentially have broad applications and motivate questions that run through several decades of experimental research. Experimental attention was drawn to similar institutions by the discovery[2] that a posted price process exhibits different efficiency and price performance than the continuous double auction. The subsequent, overarching literature represents a meticulous exploration of blends of call markets and the continuous double auction that has led us to the experiments and models developed here. Cason and Friedman (1997) nicely summarize the issues: "The general question of price formation thus resolves into three research questions. What are the relevant market institutions? What are the equilibrium properties of such institutions? And to what extent do human traders come to approximate the equilibrium outcomes?"

The experimental focus was first drawn to periodic call markets by Smith et al. (1982), who observed that call market prices demonstrated a convergence process in a repetitive, stationary environment with multiple units. Price convergence was slow relative to continuous markets; the ultimate efficiency was below but comparable to the continuous double auction. McCabe, Rassenti, and Smith (1993) studied performance of call markets with differing features, including multiple units, open/closed book, freedom to modify or cancel (at a cost) orders during bid tenders, different call and pricing rules, and different order submission rules.[3] Similarly to the earlier work, they observed convergence fell short of the competitive equilibrium. Why convergence was slow and what changes might make it faster were open questions that emerged from the work.

Guided by the theoretical development of Wilson (1987) and Satterthwaite and Williams (1993), the experimental work of Friedman (1993), Cason and Friedman

[2] See Plott and Smith (1978).

[3] Their research was motivated by the rule used in the Arizona stock exchange (1992–2001) and by an interest in isolating procedures and rules that might enhance the performance of call markets.

(1997), and Kagel (2004) explored principles of call market price formation under very strong conditions that allowed a test of the Bayes Nash Equilibrium model. Presumably, a better understanding of the details of bidder behavior would produce insights about the behavior of the system.[4] The environment included a closed book during bid tenders, a one-unit restriction on individual preferences, randomly changing costs and values, and price determination rules that provided a clear view of individual strategies. Their experiments produced systematic deviations from the Bayes Nash Equilibrium model and the patterns they observed motivated a conjecture that a learning aspect was needed. Models based on exposed decision errors and missed trades seemed promising. Such models were consistent with behavior observed in second-price auctions by Garvin and Kagel (1994) and also by Cason and Friedman (1999a).

Suggestion of a learning process in the randomly changing environment led naturally to a question about whether having multiple calls in a single period would lead to the emergence of price convergence and efficiency. Cason and Friedman (1999b, 2008), investigated this possibility with a mechanism they called the Multiple Call Market. They explored the question in a "thin" market environment, which classical theory suggests is extremely challenging, especially for the study of delicate strategic relationships. They observed substantial inefficiencies that they attribute to the thin markets.

A natural question motivated by the Cason and Friedman experiments is whether or not thicker markets with public (open) book, bid adjustment flexibility, and multiple calls will enhance call market performance. The issue receives some support from the experiments of Cason and Plott (1996), who study call markets in a replicating environment with individual bidder incentives determined at random. When viewed from one call to the next, the replicating environment has coordination and information similarities to those of multiple calls within a period. Cason and Plott observe both efficiency and price convergence to near competitive equilibrium levels. More importantly, they also reported value revelation of extra marginal units, which is directly related to the role that value revelation of marginal and extra marginal units can play in forming a process of convergence when market environments are repeated across periods.

A connection between excess demand and price changes was established early (Smith 1965). A connection between prices and order flow as represented by excess bids (i.e., total buy orders minus total sell orders) was established later (Smith, Suchanek, and Williams 1988), leading to a long-standing challenge to understand the mechanisms at work. That work as well as Selten and Neugebauer (2014) find substantial support for the excess bids model as a predictor of prices. Their analysis leaves open the question of whether or not the excess bids model is more accurate than the classical excess demand model or the Newton-Jaws model developed below, and what might be the source of its accuracy in predicting price changes.

Studying the market adjustments in response to an unstable competitive equilibrium, Plott and George (1992) demonstrated that a special type of call market with

[4] See also Friedman and Ostroy (1995) who investigated several equilibrium models in a quantity-only call market they called CHQ.

price changes responding to bids and asks through an explicit tâtonnement secant mechanism (a "smart" market approximation of the Newton method) converges to the nearest Walrasian stable competitive equilibrium.[5] This discovery re-emphasizes the central role of excess demand, as well as its slopes, in determining price adjustment and discovery process, and thus the importance of excluded bids and asks in approximating those slopes. The insight becomes enhanced with the idea that the excluded bids and asks in the continuous double auction, captured by the order book shape, perform the same function. The jaw-type structure of the order flow recorded in the order book is related to the rules governing the order book for continuous, multiple-unit double auctions.[6] That possibility was formalized by Bossaerts and Plott (2008) as the market "jaws": a Newton adjustment process based on the jaw-shaped order book could contribute to price convergence in the continuous double auction.[7] Whether a Newton-Jaws type adjustment operates in a discrete, multiple-call market environment has not been investigated until now.

The call market exchange we explore in this paper incorporates several features shown to be important in the literature. The exchange consists of multiple (two) calls in each of a series of periods replicated under stationary market demand and supply schedules. The order book is open so all participants can view the order flow and the tentative price, which is continuously computed and displayed. Following the rules of the call market that has become known as the uniform price double action (UPDA) as introduced by McCabe, Rassenti, and Smith (1993), bids and asks can be tendered, adjusted, or cancelled at any time during the order submission period. After a call, each participant sees the volume, own transactions, and the untraded bids and asks, which remain in the book, as is the case with the Cason and Friedman rules. Markets are not thin in the sense of Cason and Friedman since agents have multiple units and there are typically more than ten buyers and ten sellers. Given previous experiments and theory, all of the features suggest that we should observe price discovery and efficiency convergence. The experimental questions are whether convergences indeed occur and if so, what dynamic model can approximate the process.

Our results demonstrate the existence of a price formation process that embodies the logic of the Newton-Raphson method of solving systems of equations, building on and extending the previous results. The information used in price formation exists in the order flow and encompasses both the information contained in the excess demand and the information in the excess bids (i.e., total buy orders minus total sell orders). However, additional information about the separate slopes of the demand and supply functions is supplied by the "jaws": the values of the excluded

[5] Gjerstad (2013) studied the price dynamics in a continuous double auction, and used a Hahn stochastic process to estimate disequilibrium price adjustment within a period, which is an alternative approach to modeling the dynamics.

[6] The definition and rules as first developed by Plott and Gray (1990) use price/time priority for listing in the book. Subsequent computerized markets such as MUDA (Johnson, Lee, and Plott 1989) and the more advanced Marketscape (for a 1997 animation illustrating the jaws dynamics of the order flow using Marketscape visit http://eeps.caltech.edu/mov/jaws.html) made a graphical representation of the data available to traders in real time.

[7] A Newton-based adjustment process was tested and rejected for the continuous double auction when operating in an environment with unstable equilibria. See Hirota et al. (2005).

bids and asks. This information is used by a Newton-like process of price formation and discovery.

There is a growing interest in call markets applications inspired by the possibility that call markets can be useful supplements to other forms of markets. While such possibilities raise basic science challenges far beyond the questions posed here, introducing the respective connections puts the research reported here in a broader context. Budish, Cramton, and Shim (2015) argue that call markets might avoid difficulties caused by high-frequency trading. Brewer, Cvitanic, and Plott (2013) suggest call markets as a tool to deal with flash crashes that might occur in continuous markets. The Euronext and Xetra exchanges use call markets combined with other forms of markets to open and close trading based on the theory that it improves price discovery.[8] A completely different approach is taken by Selten and Neugebauer (2014), who attempt to create phenomena reported in the finance literature as "puzzles" in the laboratory. Notwithstanding the design differences, they also find support for the predictive model of price formation based on excess bids. They argue that path dependence between current and past excess bids, i.e., the adaptive model of price formation, operates at the individual level, while in our experiment this mechanism is eliminated by book clearing at the end of each period.[9]

What form a call market should take to meet these challenges, how they would perform, or what forms the theory might take to unravel the challenges presented by field observations are beyond the scope of this paper.

II. Environment

In this section, we describe the experimental call markets implemented in the laboratory.

A. *Institution, Rules, and Timing*

The call market we study is based on a double auction design in which both bids to buy and asks to sell are tendered. Unlike the continuous time double auction, trades only happen at a call. Before the start of the experiment, the subjects are designated as buyers or as sellers, which they keep for the entire experimental session. Each session consists of several periods, developed as follows.

Before the start of a period, costs and redemption values are induced. Costs are distributed in terms of buy orders from the experimenter to the sellers (who would buy from the experimenter and resell to buyers) and redemption values are distributed in terms of sell orders from the experimenter to the buyers (who would buy from sellers and resell to the experimenter). These incentive-based orders are placed in a private market accessible only by the subject for whom they are intended. No

[8] See van Bommel and Hoffmann (2011).

[9] See also Selten and Neugebauer (2015), who compare call markets and double auctions, and report the call markets as less effective. They do not explore variations of the two institutions or isolate the principles that seem to be operating.

inventories in terms of units or orders are carried over from period to period. All values, costs, and prices were specified in experimental currency called "francs."

A public (trading) market opens at the beginning of each period. In this public market, two calls are performed each period. The first call is 1.5 minutes into the period and the second is 4.5 minutes into the period (3 minutes later), leaving 1.5 minutes to redeem units purchased or return unsold units to the experimenter when no more calls remain in the period.[10]

At any time during the period, sellers can place sell orders and buyers can place buy orders to the public market. Orders are ranked (buy orders from high to low and sell orders from low to high) according to the execution mechanism, should a call take place. The orders are published on a screen so any trader can see everyone's orders in the sequence of potential executions. Orders are also displayed in the graphical form by means of demand and supply curves based on the current order book. Orders can be canceled and resubmitted at any time before the call so the curves and prices can shift around before the call. No constraints are placed on orders except by limiting the number of units to six. Subjects are allowed to tender potentially unprofitable offers. Thus, the technology allows subjects to attempt to manipulate the price. The number of orders a subject can have simultaneously placed on the public market at any given time is limited by the number of units made available by the experimenter.

At each call, all buy and sell orders in the order book are simultaneously considered and a market price is established. It is determined as follows:

- Based on all orders in the book, the system sorts buy orders by their respective prices per unit from high to low. Sell orders are sorted by their respective prices per unit from low to high.
- The system matches two sorted series selecting all pairs for which the purchase price is greater than the sale price, and stops at the last pair for which this is true.
- The market price is calculated midway between the last accepted (the lowest filled) buy order and the last accepted (highest filled) sell order. Except for ties, all buy orders with prices above the market price will trade at the market price. All sell orders with prices below the market price will trade at the market price. All other orders will remain unfilled.

Technically, the call price (the market clearing price announced at a call) must be computed from discrete or integer-valued bids and asks, and is determined from submitted orders as follows. Let z be an index of buy orders (bids) ordered from high to low and sell orders (asks) ordered from low to high. Thus, z is an index of ordered pairs $(b(z), a(z))$, where $b(z)$ is the bid, and $a(z)$ is the ask of the zth pair. Let z^* be the smallest z for which $b(z+1) < a(z+1)$. Thus, z^* is the index of the

[10] The last 1.5 minutes of a period were unnecessary for the call market functioning, but allowed the subjects to learn the outcome of the second call trade and manually convert their units on hand into francs if they wished to do so. As a convenience feature, the software automatically converted all units on hand into francs at the end of the period using traders' true value and cost schedules.

"last trade," the last accepted bid, and the last accepted ask. A market clearing price is any $p^* > 0$ such that:

- for $z \leq z^*$, $b(z) \geq p^*$ and $a(z) \leq p^*$; and
- $p^* \in \left[\max\{b(z^*+1), a(z^*)\}, \min\{b(z^*), a(z^*+1)\}\right]$.

The concept of a market clearing price is related to the concept of a competitive equilibrium in the sense that the competitive equilibrium is a market clearing price but the competitive equilibrium price is based on the concepts of market demand and supply and not just the bids and asks that happened to have been submitted prior to a call. The distinctions will be addressed in the Section III discussion about the dynamics of convergence in an ongoing market system.

Participants have profits continuously updated. A history of all trades up to the current time is always available. Remaining orders in private markets are always displayed. Untraded units are returned to the experimenter at the private market price at the end of each period.

B. *Experimental Procedures*

Subjects were recruited from Caltech and Purdue University. In total, 123 subjects participated. Upon sign-up for the experiment, subjects received an email with the hyperlink to the actual experiment webpage, instructions, and the demo. We also recorded and uploaded a short video describing the details of the experiment using the software interface.[11] The instructions are available in Appendix C. Subjects were paid after the end of the session by checks mailed to the addresses they specified at the sign-up. In all sessions, subjects made decisions via Internet using a web browser. Each session had 18 to 19 periods and lasted about 2 hours. Subjects were not informed about the last period unless it was over.

The first three periods were practice periods using a specially designed set of parameters that allowed low gains and low losses. Subjects were told that the first three periods were designed to help them understand how the software worked. Subjects were instructed that if they failed to make a profit in the first three periods to demonstrate their understanding of the trading system, they would receive a show-up payment but would not be used in the experiment. A frequent mistake was related to thinking that they should exercise all orders found in their private order books, e.g., sell all units they could independently of profitability.

Subjects were randomly assigned as buyers and sellers, and their types were fixed during the session. However, buyer redemption values and seller costs were changed once after the practice and once after a parameter shift, as explained below. Buyers (sellers) could submit buy (sell), multiple-unit orders in a public market and redeem their values (costs) from the experimenter using their private values (costs) markets.

Table 1 presents the summary of the experimental sessions.

[11] The video is available at http://tinyurl.com/kcq6pmb.

Table 1—All Experimental Sessions

Session no.	Date	Practice, periods	Pre-shift, periods	Post-shift, periods	Initial subjects	Paid subjects	Average payoff, $	Exchange rate
1	2012–05–11	3	8	8	17	14	45.21	1f = 3.5¢
2	2012–05–12	3	8	8	13	13	48.85	1f = 3.5¢
3	2012–12–01	4	7	7	15	15	36.13	1f = 3.5¢
4	2012–12–13	3	8	8	17	17	41.06	1f = 3.5¢
5	2013–02–23	3	8	8	21	19	34.79	1f = 2.5¢
6	2013–03–02	3	8	8	21	21	37.38	1f = 2.5¢
7	2013–05–16	3	9	7	19	17	30.59	1f = 2.5¢

Note: After end of practice and after shift, all types were rotated.

Table 2—Experimental Parameters

		Private values/costs per unit, francs						
Sessions	Types	Unit 1	Unit 2	Unit 3	Unit 4	Unit 5	Unit 6+	Shift
1–4	B1	210	195	190	155	150	120	75 added
	B2	220	200	175	165	145	115	to each unit
	B3	215	195	185	160	140	125	
	S1	120	150	155	175	195	205	
	S2	115	145	165	185	200	210	
	S3	125	140	160	180	190	215	
5–7	B1	250	208	203	170	144	120	83 added
	B2	246	213	199	165	140	118	to each unit
	B3	242	218	195	160	133	115	
	B4	238	224	190	155	130	112	
	B5	234	230	186	150	125	110	
	S1	110	150	155	185	214	235	
	S2	115	146	160	190	218	240	
	S3	120	142	165	196	222	245	
	S4	122	137	170	204	226	250	
	S5	126	132	174	209	230	255	

Notes: Types indicators correspond to (B)uyers and (S)ellers. All sessions included three-period practice with different test parameters. After end of practice and after shift, all types were rotated.

Subjects' earnings in francs were exchanged into US dollars at the end of the experiment. Average earnings were $19.57 per hour.

C. *Basic Parameters*

We chose the basic parameters in the experiment to satisfy a wide range of criteria implied by our focus on convergence and market dynamics. In sessions 1–4, we used three types of buyers and three types of sellers, where each subject's type defines her private costs/values. In sessions 5–7, we used five types of buyers and five types of sellers. Table 2 contains the costs and values for all our setups.[12] In all

[12] Multiple-unit demand and supply creates a possibility for strategic price manipulation by "withholding" units to create shortages, which could, in turn, influence the speed and efficiency of market adjustments and induce strategic behavior that game-theoretic models attempt to understand. Our choice of parameters reflects our interest in price discovery and the fundamental principles of convergence and price dynamics across calls in competitive markets. Researchers interested in the challenge posed by withholding strategies could explore the incentives for

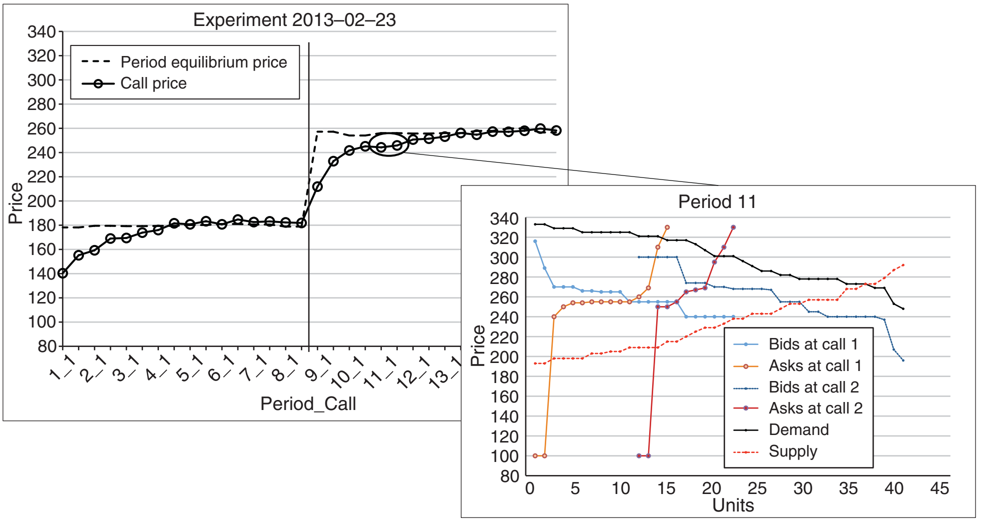

FIGURE 1. CALL PRICE DYNAMICS IN THE EXPERIMENT ON 2013–03–02

Notes: Call prices shown are the actual prices from the experiment. Competitive equilibrium (CE) period prices, which can differ from experiment to experiment due to different numbers of participants, are shown as the dashed line, which is average of the CE price over the corresponding period across all our experiments. After period 8, there is an exogenous shock (parameter shift). The market jaws are formed each call by the excluded bids (lower jaw) and excluded asks (upper jaw) that resemble an open mouth of a fish swimming to the right.

sessions, we also implemented a parameter shift around period 9 after the end of practice. The shift increased all costs and values by the specified amount of francs. Types were rotated after the parameter shift. Subjects were assigned to types uniformly so that the market contained multiple traders of each type.

Figure 1 illustrates the main features of the experimental call market. The left panel of Figure 1 shows the time series of the call prices for each call in the experiment of 2013–03–02. We see how call prices converge to the competitive equilibrium (dashed line) as determined by the classical demand and supply model up to period 9, when the upshift of equilibrium price by about 80 francs takes place, which resets the convergence anew. The right panel of Figure 1 depicts the order book of the experiment (2013–03–02), period 11. The market demand and supply are based on the induced values and costs used in that period, as well as the revealed demand and supply based on the order flow, the buy orders (bids) and sell orders (asks) in the book for the two calls. The patterns of the orders in the book resemble a "hockey stick" with the handle appearing flat and near the market price and the blades at angles reflecting and approximating the relative values of excluded units. The submitted values of the marginal and extra-marginal units along the "blade" of the hockey stick, play an important role in the price dynamics, as we demonstrate below. These values taken together will be called the market "jaws"—an open mouth ready to bite as illustrated in the figure.

withholding by tweaking our parameters near the equilibrium prices in order to reduce the elasticity of demand and supply.

III. Models

The price discovery model explored here is a process of convergence based on insights from three sources, each of which suggests a separate element of the overall model. The first is the classical theory of price adjustment that supplies the most basic of principles that price changes are responsive to excess demand. The second is a more abstract literature that points to the possible role of the slope of the excess demand. Both are introduced in Subsection IIIA. The third, introduced in Subsection IIIB, is the understanding provided by "market jaws" that order flow as shaped by specific market making institutions is a fundamental source of the information and commitments required by an equilibration process. Together, these principles operate with properties similar to the Newton method of finding zeroes of a differentiable function as will be illustrated in Subsection IIIC. The connection with numerical methods provides a framework for understanding how institutions coordinate decisions and interact with decentralized information, and helps focus on where institutional changes might improve market functioning.

A. *The Classical Concepts of Demand, Supply, and Equilibration*

The market demand function, $D(p)$ and the market supply function, $S(p)$, express the quantity that buyers are willing to buy at price p and the quantity that sellers are willing to sell at price p and are derived from respective utility and profit maximization with the assumption that decision makers treat prices as constants. Excess demand at a market price, p, is defined as $ED(p) = D(p) - S(p)$. The classical concept of a competitive equilibrium price is a price p^* such that $ED(p^*) = 0$. Since the demand and supply functions are generally not observable, the concept of "price discovery" has emerged over the decades in response to the view that the equilibrium price is the solution to equations that no one in the market knows.

Classical theory of price discovery works through an abstract adjustment process termed "tâtonnement." It is as if a price is announced (by a fictional agent sometimes called the "Walrasian auctioneer"), the excess demand is observed and, without trading taking place, a new adjusted price is announced based on the revealed excess demand quantity. The price movement motivated by the model is summarized by the classical price adjustment equation:

$$(1) \qquad \frac{dp}{dt} = A\,[D(p) - S(p)], \qquad A > 0.$$

Under appropriate conditions this adjustment will converge to p^*: $ED(p^*) = 0$.[13]

A natural technical generalization of the classical model, that strengthens the conditions under which price discovery can happen, is based on a Newton adjustment derived from the Taylor expansion:

$$(2) \qquad \frac{dp}{dt} = -\frac{D(p) - S(p)}{D'(p) - S'(p)}.$$

[13] For the background and development of this class of models see McKenzie (2002).

This extension postulates that the constant, A, in the tâtonnement model is replaced by a function of the excess demand derivative, $-1/(D'(p) - S'(p))$. Thus, the process could be dependent on both $ED(p)$ and $ED'(p)$ as opposed to $ED(p)$ alone, and, as has been explored by Saari and Simon (1978) and Saari (1985), the additional information has implications for price discovery, i.e., convergence to a competitive equilibrium. However, the Walrasian auctioneer would not have access to the additional information and the range of institutions that might carry the information has remained an open question.

B. *Market Jaws*

Market institutions that have evolved over the decades are populated with additional institutional features that can be a source of information. For example, evidence exists that suggests that such source could be the order flow of the continuous double auction (Asparouhova, Bossaerts, and Plott 2003; Bossaerts and Plott 2008; Barner, Feri, and Plott 2005). Bossaerts and Plott (2008) suggest that the information is in the market order "book" and term the source "market jaws." The question explored here is whether a similar process and source of information might exist for call market institutions.

The structure of the data in the order book is sometimes described as an open mouth with an upper jaw and a lower jaw, together with a tongue, which is a curve tracing the average of the two. Market jaws provide a snapshot of commitments by potential traders that approximate important features of observed market adjustments in continuous markets. Here, the model is adapted for call markets institutions. When a call takes place, these commitments together with market making rules define exchanges that are executed at the call.

The basic intuition is as follows. As order flow develops and bids and asks are submitted, the shape of the order book changes in a self-organizing and specific way that reflects aggregate demand and supply. In order to become "provisional traders" who would trade if a call took place, traders can revise their bids and asks in an attempt to meet or beat the competition in light of the offers tendered by the other side of the market. The tendency of bidders to anticipate the bids of others is an established property of call markets (Cason and Friedman 1997; Kagel 2004). For the marginal and extra marginal units outside the expected set of provisional trades, the possibility that the price changes randomly can create incentives for value revelation. Value revelation is encouraged by the possibility to trade in the case the market "jumps" from expected price.

As a result, the order book is continuously updated as traders update their orders in light of the orders of others. The orders of those traders whose true values are extra marginal at the current price are pushed out. Bidders change their offers with increasing revelation of the marginal values in response. The shape of the book resembles jaws, hence, the name.

To illustrate, consider Figure 2. Consider p_1, the actual price at the end of the first call and q_1, the actual total volume at the end of the first call. We fit a line to the unfilled asks that exist in the order book at the end of the first call. We only use the first τ unfilled asks in order to avoid the extremely high asks that sometimes show

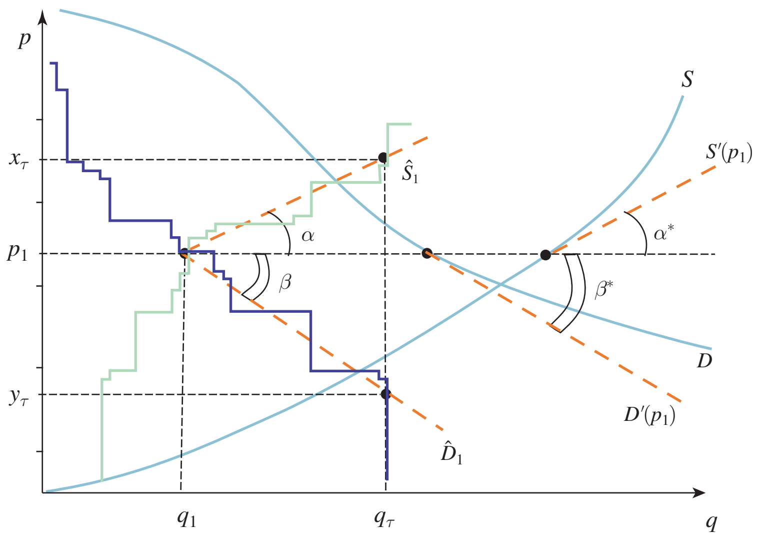

Figure 2. Theory Behind the Jaws, Applied after the First Call

Note: The jaws are formed by the fitted line segments of demand $\hat{D}_1$ and supply $\hat{S}_1$, remaining after the call.

up in the book, and also due to the local nature of Newton approximation. We also restrict the fitted line to go through the point (p_1, q_1), and denote x_τ the price at which the line cuts through $q_\tau = q_1 + \tau$, the corresponding quantity. Similarly, we fit a line to the unfilled bids in the book at the end of the first call, denoting y_τ the price at which the line cuts through q_τ.

When the two fitted lines are imposed at the point (p_1, q_1), as illustrated in Figure 2, we obtain a graphical representation of the market jaws: the "upper jaw" given by the line connecting (p_1, q_1) to (x_τ, q_τ), and the "lower jaw" given by the line connecting (p_1, q_1) to (y_τ, q_τ). We will refer to these lines as $\hat{S}_1$ and $\hat{D}_1$, respectively, and treat them as tangents to a smoothed out model of the revealed supply and demand as represented by the unfilled asks and bids in the order book.[14] That is, we are going to use their slopes as an estimate of the excess demand slope when evaluated at the price of the first call. The excess demand slope is one of the two key features of the Newton adjustment, as explained below.

Now, define α, the angle between the upper jaw, $\hat{S}_1$, and price, p_1, using the slope of $\hat{S}_1$:

$$\tan\alpha = \frac{x_\tau - p_1}{\underbrace{q_\tau - q_1}_{=\tau}},$$

and define β, the angle between price p_1 and the lower jaw, $\hat{D}_1$, via

$$\tan\beta = \frac{y_\tau - p_1}{\underbrace{q_\tau - q_1}_{=\tau}}.$$

[14] We fit $\hat{D}_1$ and $\hat{S}_1$ to the data by ordinary least squares in a non-trivial way as described in detail in Appendix A.

The true demand and supply curves (based on the private values and costs of the traders present at both calls) in Figure 2 are given by $D(\cdot)$ and $S(\cdot)$, respectively. For the sake of clarity, they are drawn as continuous curves, but in the actual experiment $D(\cdot)$ and $S(\cdot)$ are discrete step-wise curves, just like the corresponding actual order book. The market level, induced (true) demand/supply curves have the corresponding angles at price p_1 being α^* and β^*, respectively (see Figure 2).

The jaws model of the order book dynamics postulates two important properties in the call market environment.

First, **excess demand slope revelation**: the excluded traders' orders accumulate according to the ranking of their true values and costs, and therefore the book at the call reflects the true slopes of aggregate demand and supply at the call price. Thus, the first property says: the slopes of both revealed jaws closely approximate the slopes of the true demand and supply at price p_1, or equivalently, $\alpha \approx \alpha^*$ and $\beta \approx \beta^*$.

Second, **excess demand revelation**: the number of bids in excess of the number of asks arriving in a fixed period shortly before a call is proportional to excess demand at the call price. The theoretical intuition behind the second property is that an approaching call motivates the traders to actively submit and adjust their bids and asks. The rate of bids and asks depends on the number of units desired and the number of buyers and sellers, and thus the excess demand at the current price.

As we show below, these properties are crucial for interpreting the price dynamics across calls as one step of the Newton-Raphson method of finding zeroes of a differentiable function. Multiple steps involving multiple periods require additional abstraction. The two properties, slope and excess demand revelation, also rely on the more fundamental feature of Marshallian adjustment path.[15]

C. *Newton and Walrasian Adjustment*

In this section, we operationalize the technical properties of Walras and Newton adjustment, which are useful in testing and estimation. We begin by discretizing the tâtonnement equations (1) and (2) by replacing the derivative on the left-hand side of the expressions by a price difference, $\Delta p = p_{t+1} - p_t$. This represents one step of an iterative process that starts from an initial price p_0 and converges to an equilibrium price p^* that solves $ED(p^*) = 0$. In a Newton-type process each successive root approximation at time $t+1$ depends on both the excess demand and the slope of the excess demand, and can be written formally as

$$p_{t+1} - p_t = -\frac{D(p_t) - S(p_t)}{D'(p_t) - S'(p_t)}, \tag{3}$$

[15] The Marshallian path is an empirical property that buyers with high values and sellers with low costs are those that first find their way to trade, and if they have multiple units, they trade their most profitable ones first. This is the mysterious property predicted by the Wilson model, observed as part of BNE performance by Cason and Friedman (1997), and Kagel (2004), and documented as a feature of the continuous double auction by Plott, Roy, and Tong (2013).

where p_t is the market price at time t, $D(\cdot)$ and $S(\cdot)$ denote the true demand and supply curves, and $D'(\cdot)$ and $S'(\cdot)$ are their respective derivatives.

In the Newton numerical method of finding a root of a differentiable function, iterations in (3) are repeated until the stopping criterion (e.g., the desired tolerance) is reached. Unlike the traditional scheme, we take only one step of this scheme. Furthermore, the Newton numerical process has all information about the functions when it starts, while the market price discovery process requires new information at each step, and the new information is based on the results of the previous step. From a bidder's point of view, new bids at any call are based on the previously announced price, a feature that suggests a need for theory.

The price difference on the left-hand side of (3) can be taken either (i) between prices realized at two subsequent time moments (e.g., two calls within one period, producing the difference $(p_2 - p_1)$), or (ii) between a Walrasian theoretical market-clearing price and the price at a given call. In other words, we could have replaced the second time period of our model with the theoretical perfect-competition Walrasian market. If we did so, we would then be able to test convergence to the Walrasian model. Thus, we can also use the equilibrium price p^* instead of p_{t+1} in (3) for the purpose of an alternative test of equilibration: asymptotically as t grows large, the iterations should converge to the equilibrium point. This results in two additional price differences: $(p^* - p_1)$ and $(p^* - p_2)$. We report these alongside the between-call difference in Results 5, 6, and 7 below. Notice that if the price dynamics followed Newton and we knew the excess demand and its derivative at time t, then we could use the price at time t and (3) to predict price at time $t+1$.

A natural alternative to the Newton dynamics is Walrasian adjustment, where

$$p_{t+1} - p_t = A\big[D(p_t) - S(p_t)\big]. \tag{4}$$

Thus, the difference of Walrasian adjustment from Newton is that the former does not utilize the slopes of excess demand, assuming that the price change is proportional to the excess demand with some positive constant factor A.[16]

The information about the excess demand and its slope can be related to the market jaws. This relationship is based on two hypothesized properties of the jaws we described in Subsection IIIB. First, (slope revelation:) the slopes of revealed jaws will closely approximate the slopes of the true demand and supply at price p_1, i.e., in terms of Figure 2, $\alpha \approx \alpha^*$ and $\beta \approx \beta^*$. Second, (excess demand revelation:) the pattern of jaws changing shortly before the call will reflect excess demand at price p_1, the second key property of the dynamic model, via the relative excess of the number of bids over the number of asks.

[16] In some continuous time environments, in particular, in unstable environments of Scarf (1960), where prices do not converge to the competitive equilibrium, Walrasian adjustment finds more support than Newton (Hirota et al. 2005). Assured global convergence to the competitive equilibrium via an iterative procedure in general environments cannot be guaranteed (Saari 1985). However, the information about excess demand and its slope suffices for local convergence (Saari and Simon 1978), and is particularly relevant in our environment with quasi-linear supply and demand, and no income effects. Hence one would expect that in a particular environment like our experiment, with less stringent information requirements, the knowledge of the first derivative of the excess demand should allow Newton to perform better than Walras.

Denote $\widehat{ED}(p_1)$ the revealed excess demand. Consider Figure 2 again. If the first property of the jaws holds as a perfect equality,

$$\hat{S}' = S'(p_t) = \tan\left(\frac{\pi}{2} - \alpha^*\right) = \cot\alpha^*;$$

$$\hat{D}' = D'(p_t) = \tan\left(\frac{\pi}{2} + \beta^*\right) = -\cot\beta^*.$$

If the second property holds perfectly, $\widehat{ED}(p_t) = D(p_t) - S(p_t)$. Then using (3), the price dynamics across calls follows Newton if and only if

$$p_{t+1} - p_t = \frac{\widehat{ED}(p_t)}{\hat{S}' - \hat{D}'}.$$

These relationships summarize the main property of Newton-Jaws.

IV. Statistical Models

A. *Statistics of Convergence*

We use a simple dynamic model to assess convergence to theoretical predictions, which in our case is the competitive equilibrium. The basic idea is to establish whether the difference between the data and the corresponding equilibrium goes to a common asymptote of zero as periods in an experiment proceed. The model was developed by Noussair, Plott, and Riezman (1995).[17]

The original model is modified to account for a shift in parameters that occurs after the first several periods. The model for price convergence is as follows:

$$(5)\quad p_{it} - p_{it}^* = \left(\frac{\alpha_{\tilde{t}_1}\delta_1}{t - \tilde{t}_1 + 1} + \frac{\beta_1}{t}\right) d_1 + \cdots + \left(\frac{\alpha_{\tilde{t}_i}\delta_i}{t - \tilde{t}_i + 1} + \frac{\beta_i}{t}\right) d_i$$
$$+ \cdots + \left(\frac{\alpha_{\tilde{t}_K}\delta_K}{t - \tilde{t}_K + 1} + \frac{\beta_K}{t}\right) d_K + \gamma\left(1 - \frac{1}{t}\right) + \varepsilon_{it},$$

where i indexes experimental market sessions; t indexes periods in a session starting from 1; p_{it} is the average market price in period t of experimental market session i; p_{it}^* is the equilibrium market price in period t of the same session; K is the total number of experimental sessions (we ran 7); $d_i, i \in \{1, \ldots, K\}$ is a dummy variable corresponding to experimental session i; β_i is the origin of the corresponding time series; $\tilde{t}_i$ indexes the period when the parameter shift[18] occurs in session i; $\alpha_{\tilde{t}_i}$ is a dummy variable that corresponds to the shift, i.e., $\alpha_{\tilde{t}_i} = 0$ for $t < \tilde{t}_i$, and $\alpha_{\tilde{t}_i} = 1$ for $t \geq \tilde{t}_i$; δ_i captures the "new origin" effect, created by the shift; γ is the

[17] Noussair, Plott, and Riezman named it the AElG model after Orley Ashenfelter and Mahmoud El-Gamal whose suggestions led to the development of the model.

[18] In our experiments, most shifts occurred in the ninth period following the practice period.

asymptote of the series common to all experimental sessions; and, finally, ε_{it} is a random error.

The same equation with $p_{it} - p_{it}^*$ replaced by $q_{it} - q_{it}^*$ is used to estimate volume convergence.

The basic idea behind this dynamic model is as follows. In experimental market session i, the difference between the data and the competitive equilibrium starts from some random origin, captured by β_i, and moves closer to the common asymptote γ as time (i.e., period number) increases from 1 to the time of the parameter shift, if there is convergence. At the time of the shift $t = \tilde{t}_i$, the term $\frac{\alpha_{\tilde{t}_i}\delta_i}{t - \tilde{t}_i + 1}$ becomes nonzero if $\delta_i \neq 0$, and so it serves as an updated origin from which the difference on the left hand side of (5) starts to converge anew.

In theory, as time increases toward infinity, the updated origin receives less and less weight (and the initial origin even less), so the difference between the equilibrium of the model and the data converges to the asymptote γ. Thus, if the estimate of γ is not significantly different from zero, we conclude that the data series converges to the equilibrium prediction perfectly.

Since we have two calls per period, there exist alternative ways to define p_{it} and p_{it}^* because the model does not predict the dynamics within a period. We explicitly address this in subsection IIIC, where we describe the application of the Newton method to our data. For our convergence results, we defined the observed market price in a period, p_{it}, as the average realized price across two calls, and the equilibrium market price in a period, p_{it}^*, as the competitive equilibrium price, based on the private values and costs of buyers and sellers who were actively present[19] in at least one of the two calls in the period. Since there was no carry-over cost from call to call, the model predicts that the two calls should create the same price and that the total volume should be distributed to maintain the equal prices. The theoretical equilibrium price as well as volume could change in every period, depending on the number of traders who are present. Thus, we defined the actual volume in a period as the total number of units traded at both calls, and the equilibrium trading volume as the volume that corresponds to the equilibrium price p_{it}^*.

B. *Market Efficiency*

As a measure of efficiency in each period, we used the consumer plus producer surplus expressed as a percentage of the maximum possible (Plott and Smith 1978). We define it as the difference between the total "consumption," i.e., the franc redemption value of the purchased units, and the total franc cost of those units, divided by the maximum possible difference between total of redemption values and costs that

[19] By "actively present" we mean those participants who submit public orders before a call, i.e., reveal their wish to participate in trade. Note that it may happen that their orders do not trade at the call, but such orders form a part of the market supply and demand at a given call, and, hence, are taken into account. In experiments conducted with remotely located subjects, as opposed to all subjects confined to the laboratory, a degree of experimental control is lost. Subjects can become distracted or simply quit without warning. From one point of view, this phenomenon is a lack of control, but a bid or ask reveals presence and parameters can be adjusted accordingly, so from another point of view the appearance or disappearance of subjects illustrates the robustness of a model that works.

can be achieved during a period. The maximum is achieved at the competitive equilibrium allocation, which, if attained, is 100-percent efficient.

Let $R = \{r\}$ be the set of all redemption values to all buyers that participated in one or more of the two calls in period t. Let $C = \{c\}$ be the set of all costs to all sellers that participated in one or more of the two calls in period t. Order the elements of R from highest to lowest, with r_i being the ith element. Order the elements of C from lowest to highest with c_i being the ith element. Let R^* and C^* denote the sets of redemptions and costs that were actually exercised during the period. Maximum Surplus is

$$\mathrm{MS} = \max_z \sum_{i=1}^{z} (r_i - c_i).$$

Realized Surplus is

$$\mathrm{RS} = \sum_{r \in R^*} r - \sum_{c \in C^*} c.$$

We define efficiency in period t as the ratio of the two quantities:

$$\text{Efficiency} = \frac{\text{Realized Surplus}}{\text{Maximum Surplus}}. \tag{6}$$

Notice that subjects can submit multiple-unit orders, and we explicitly account for this possibility in the efficiency score.

V. Results

All results in this section are presented in the form of a "result" statement followed by the "support." The result statement summarizes the authors' qualitative interpretation of the data within the context of the abstract theory, and the support provides the precise relationships and technical details that justify the interpretation. We present several types of results describing the market-level properties (macro-properties) of the call markets (convergence, efficiency, and price dynamics.) Subsection VA addresses the traditional measures of market performance such as convergence of prices, volumes, and efficiency relative to the competitive equilibrium. Subsection VB addresses the more detailed model of the nature of price adjustment as suggested by the Newton dynamic.

A. *Market Performance Relative to the Competitive Equilibrium*

The section contains three results related to broad properties of the call markets. Together the results say that market behavior is captured reasonably well by the competitive equilibrium model. Prices, volumes, and efficiencies all converge to near the quantities predicted by the model. Price and volume behave substantially as predicted when parameters change. These are evident in Figures 3, 4, and 5 showing, respectively, price convergence, volume convergence, and efficiencies in response to a market with a stationary demand and supply punctuated by a parameter shift in period 9, and then returning to a stationary level.

Result 1: In the two-call multiple-unit market, price and trading volume converge to near the equilibrium levels of the competitive equilibrium model. Equilibrium price upward shift affects price but not volume convergence, as predicted.

Support: Using data from seven market experiments, we estimate a simple dynamic model of convergence described in Subsection IVA for price (and volume) by equation (5).

We estimate the model in (5) using ordinary least squares with bootstrapped standard errors.[20] Since we have the order book cleared across the periods, we can treat periods as independent observations.[21] An estimate of the common asymptote, γ, close to zero implies that the actual price/volume converges to the price/volume of the static theoretical model as time proceeds. The results of our estimation are reported in Table 3.

As Table 3 shows, we reject perfect price (volume) convergence: the estimated value of the asymptote γ, 7.643 (-1.785, resp.), is significantly different from zero, with its bootstrap standard error of 2.472 (0.465, resp.) Nevertheless, the estimated asymptotic differences are rather small: the equilibrium price in these experiments ranged from 165 to 280.5 francs (against the error of 7.643 francs, or about 3 percent to 5 percent of the equilibrium), and the equilibrium volume ranged from 17 to 32 units (against the error of 1.785 units, or about 6 percent to 11 percent of the equilibrium).

Figures 3–4 show graphically the price and volume dynamics across periods, averaged over all experiments. The spike at period 9 corresponds to the shift in parameters. Figure 3 also shows that for almost all periods, price at the second call is closer to equilibrium than price at the first one.

Notice also that the parameter shift effects are consistent with model predictions. In our experiments, the shift increases the equilibrium price by a constant, but does not change the equilibrium volume. The estimations in Table 3 display this feature: the updated origins after the shift, $\delta_1 - \delta_7$ (except δ_3 and δ_4), are highly significant

[20] We programmed the ordinary nonparametric bootstrap with bias correction in R. All data and code are available from the authors upon request. Regression coefficient estimates in the tables are bias-corrected, i.e., equal $\tilde{\beta} \equiv \hat{\beta} - \widehat{Bias} = \hat{\beta} - \left(\frac{1}{B}\sum_{b=1}^{B}\hat{\beta}_b^* - \hat{\beta}\right)$, where $\hat{\beta}$ is an OLS estimate from the original data sample, $\hat{\beta}_b^*$ is an OLS estimate from bth bootstrap sample, and B is the total number of bootstrap samples (we use $B = 10{,}000$). Bootstrap standard error is $s(\hat{\beta}^*) = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}\left(\hat{\beta}_b^* - \frac{1}{B}\sum_{b=1}^{B}\hat{\beta}_b^*\right)^2}$. Notice that while bootstrap standard errors can be used to test for significance of regression coefficients in a straightforward way by plugging them into the usual t-statistic, doing so does not fully utilize the advantage of the bootstrap. A better test procedure we implemented (percentile-t bootstrap test) uses bootstrap to compute the critical values from the finite sample distribution of the test statistic. Namely, to test hypothesis $H_0 : \beta = \beta^0$ versus the two-sided alternative, we bootstrapped a symmetrical recentered t-statistic $\hat{t}_b^* = \frac{|\hat{\beta}_b^* - \hat{\beta}|}{s(\hat{\beta}_b^*)}$ to obtain the $1-\alpha$ quantiles of the bootstrap distribution $\{\hat{t}_b^*\}_{b=1}^{B}$, and compared them with the test statistic $\hat{t} = \frac{|\hat{\beta} - \beta^0|}{s(\hat{\beta})}$. Hypothesis is rejected at level α if $\hat{t} > q_{1-\alpha}^*$. See Horowitz (2001) for details.

[21] Allowing for autocorrelation does not noticeably change the results.

TABLE 3—ESTIMATION OF CONVERGENCE: PRICE AND QUANTITY

Regressor	Dependent variable			
	Price difference, $p_{it} - p_{it}^*$		Volume difference, $q_{it} - q_{it}^*$	
Asymptote (γ)	7.643	(2.472)	−1.785	(0.465)
Origin (before shift) in				
session 1 (β_1)	−17.390	(25.901)	−1.814	(5.236)
— 2 (β_2)	−2.432	(12.215)	−1.521	(2.268)
— 3 (β_3)	−22.616	(12.228)	−10.560	(4.473)
— 4 (β_4)	−18.052	(12.681)	−5.760	(4.089)
— 5 (β_5)	−64.243	(12.970)	−1.433	(2.702)
— 6 (β_6)	−64.531	(23.390)	−8.590	(3.369)
— 7 (β_7)	−22.606	(27.945)	−3.940	(2.475)
Origin (after shift) in				
session 1 (δ_1)	−42.382	(13.394)	−2.497	(2.439)
— 2 (δ_2)	−53.154	(8.109)	−5.620	(3.094)
— 3 (δ_3)	1.697	(8.906)	3.734	(1.698)
— 4 (δ_4)	−10.334	(10.608)	1.534	(2.844)
— 5 (δ_5)	−51.431	(11.579)	−1.286	(3.461)
— 6 (δ_6)	−67.122	(15.202)	−1.441	(3.836)
— 7 (δ_7)	−38.202	(14.426)	−0.752	(1.341)
Observations	94		94	
R^2_{adj}	0.732		0.638	
F-stat	20.88		13.82	

Notes: Bootstrap standard errors are in parentheses. Estimates are bootstrap-corrected for bias.

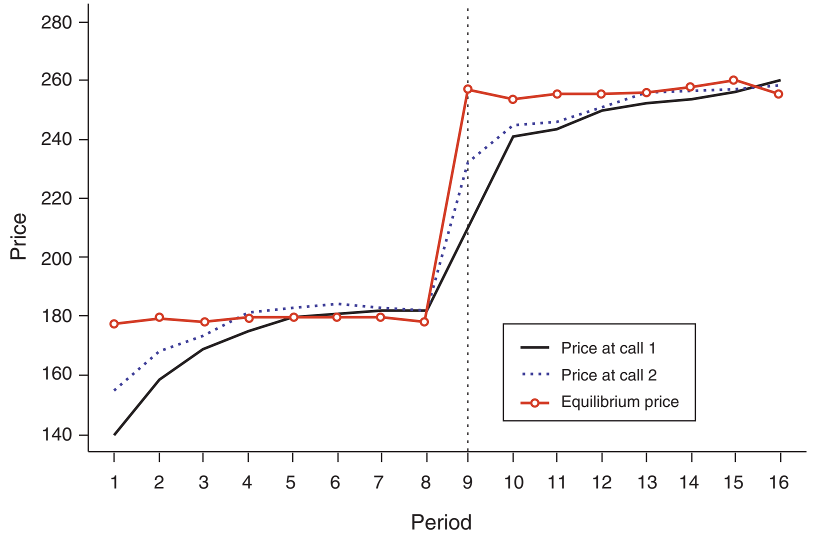

FIGURE 3. AVERAGE PRICE DYNAMICS ACROSS PERIODS

Note: There is a parameter shift at period 9.

and large for the price convergence equation, whereas for the volume convergence, only δ_2 and δ_3 are significant.

Taking together data from Table 3 and Figures 3–4, we argue that the price, and to a lesser extent, volume, converge close to their equilibrium levels.

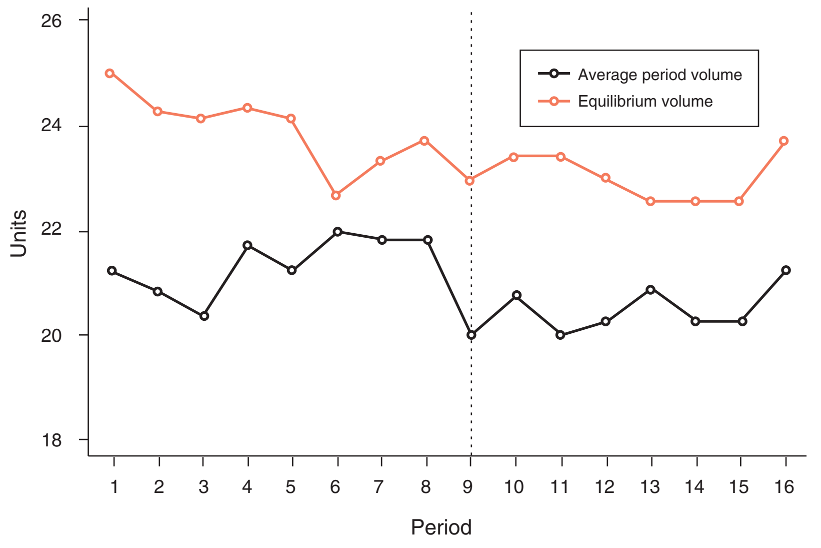

Figure 4. Average Volume Dynamics across Periods

Note: There is a parameter shift at period 9.

Result 2: The average efficiency score increases over time as price and volume converge to their equilibrium levels. The two-call market does not achieve full efficiency, but is about 83 percent efficient on average.

Support: We computed efficiency in each period of all experiments as the normalized total surplus, defined in subsection IVB by equation (6).

The average efficiency score is 82.54 percent. This is a bit less than levels typically reported in single-call market experiments (e.g., Cason and Friedman 1997 report an efficiency score of 87.3 percent). However, the average efficiency score increases over time, as price and volume converge to their equilibrium levels, with a sharp drop after the parameter shift, which corresponds to the market adjustment. Figure 5 illustrates. It seems intuitive that the efficiency score should roughly correspond to how well the total volume in a period matches the equilibrium volume, provided the actual price is close to the equilibrium price.[22] However, there is more to this than a simple comparison of total volumes: since subjects can make multiple-unit orders, it is also matters that all subjects do not over- or under-acquire their inventory.

Result 3: The parameter shift of demand and supply upward by a constant only affects the price, and not the volume of trading, as should be expected. There is no significant effect on the observed efficiency.

[22] When the actual price is far from the equilibrium price, efficiency is low even if the volumes are matched exactly. This is the case, for example in the first period after practice in the experiment of Session 5 in Table 1: the equilibrium price was 188, the actual price was 117; subjects acquired 26 units (with 28 units in equilibrium) and the efficiency score was the lowest: 28.28 percent.

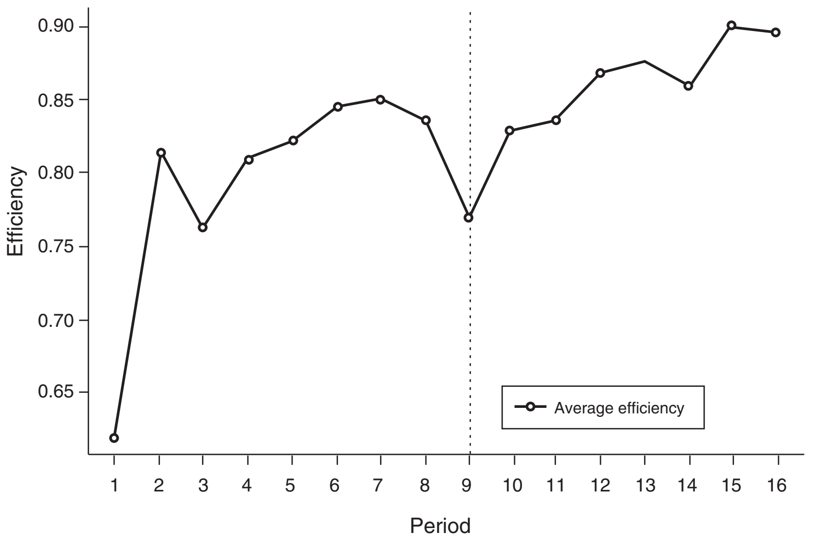

Figure 5. Average Efficiency Score by Period (*average across periods: 82.6 percent, converges to 89.1 percent*)

Support: Table 4 presents the summary statistics of the data across periods, grouped by the parameter shift. As expected, the period volume before and after the parameter shift does not change much, since the shift only affected price, and not volume. The changes in equilibrium volume are due to the varying number of traders. The changes in efficiency are not significantly different (Mann-Whitney $p = 0.172$).

B. *Principles and Models of Price Discovery: Newton-Jaws and Alternatives*

In this subsection, we demonstrate that key elements of market jaws and the Newton method, characterized in subsections IIIB and IIIC, are observed in the call markets. We organize the results into two parts, focusing first on structural and specification tests (i.e., how well the models explain the data conditioned on known parameters) and next on parameter sensitivity and prediction properties (i.e., relative model performance conditioned on estimated parameters).

Structural Approach.—The next three results address three different models that ultimately become combined into the Newton-Jaws model. Each has its own structure that can be tested separately. Result 4 addresses the market jaws model. Result 5 examines the Newton model and Result 6 examines the Walrasian model.

Result 4: The two main properties of the market jaws find limited support in the data:

(i) (Slope revelation). The excess demand slopes based on true parameters of the model are well approximated by the slope estimates obtained from the jaws.

Table 4—Summary Data Statistics across Periods

Variable	Before shift		After shift	
	Mean	Range	Mean	Range
Total period volume, units	21.38	[13.00 .. 32.00]	20.43	[13.00 .. 29.00]
Equilibrium period volume, units	23.93	[17.00 .. 32.00]	23.00	[18.00 .. 32.00]
Average period price, francs	174.50	[121.50 .. 225.00]	248.30	[191.50 .. 288.50]
Equilibrium period price, francs	179.60	[165.00 .. 201.00]	256.70	[240.00 .. 280.50]
Efficiency, percent	79.81	[28.28 .. 99.54]	85.32	[61.80 .. 99.21]

Table 5—Estimation of Supply/Demand Slope Revelation at Each Call by Market Jaws

	Supply at call 1		Demand at call 1		Supply at call 2		Demand at call 2	
Regressor	Estimate	(SE)	Estimate	(SE)	Estimate	(SE)	Estimate	(SE)
Slope(γ_1)	0.693	(0.065)	1.518	(0.203)	0.661	(0.045)	1.005	(0.104)
$H_0 : \gamma_1 = 1$	Rejected at < 0.001 level		Rejected at < 0.05 level		Rejected at < 0.001 level		Not rejected	

Notes: Bootstrap-corrected estimates of regression (A1) terms (for jaws computed using fixed τ) and their standard errors using 10,000 replications. The null of perfect revelation by jaws, $H_0 : \gamma_1 = 1$, is bootstrap-tested.

(ii) (Excess demand revelation). The jaws-based estimates of the excess demand converge to near the actual excess demand, especially at call 2. At call 1, substantial variance precludes tight convergence.

Support: Consider (i), *slope revelation*. The slopes of the jaws imperfectly but robustly reflect the true slopes of aggregate demand and supply at each call. The null hypothesis says both estimated jaw slopes closely approximate the slopes of the true demand and supply at price p_1, or equivalently, in terms of Figure 2, $\alpha \approx \alpha^*$ and $\beta \approx \beta^*$. To test that this holds for demand and supply remaining after the first call, we use the algorithm that has a resemblance to the two-stage least squares, and essentially compares $S'(p_1)$ with $\cot\alpha$ (or $D'(p_1)$ with $-\cot\beta$, respectively). The null hypothesis says that the jaws perfectly reveal each slope (e.g., that in regression $S'(p_1) = \gamma_1 \cot\alpha + \varepsilon, \gamma_1 = 1$). The detailed description is provided in Appendix Section A1. We report the more conservative estimates (with fixed τ, the number of orders used to estimate the slopes from the data) in Table 5, and additional estimates in Table A1 in Appendix A.

After the first call, the slope of the revealed supply can be positively related to the slope of the equilibrium supply, but the null of full supply revelation is rejected at < 0.001 significance level. By contrast, the true demand revelation at call 1 is only marginally rejected. We also checked revelation after the second call. The null hypothesis of perfect revelation by jaws for demand is not rejected.

Overall, our results in Table 5 show that the slopes of the jaws can imperfectly but meaningfully approximate the slope of the excess demand even with substantially infrequent trade opportunities, as in a call market.

Consider (ii), *excess demand revelation*. The second hypothesized property of jaw adjustment is that the arrivals of bids relative to asks shortly before a call is

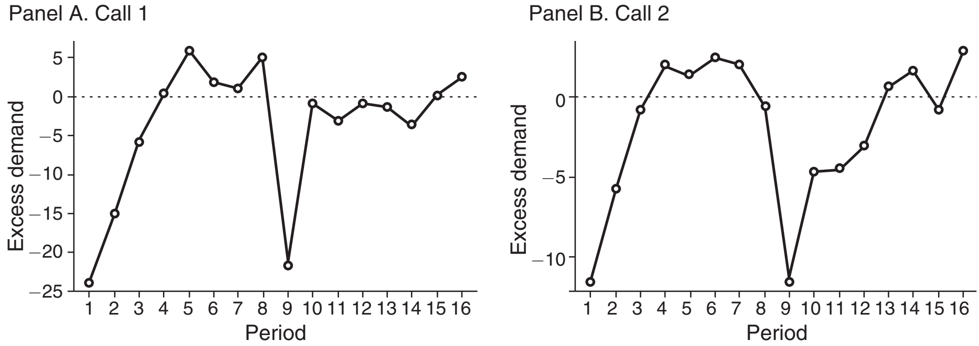

Figure 6. Average Difference between Excess Demand as Revealed by Jaws and the Actual Excess Demand after Each Call

Note: In period 9, there is a parameter shift.

proportional to excess demand at the call price. The intuition is that an approaching call motivates the traders to actively submit and adjust their bids and asks, with adjustment rate depending on excess demand at the current price.

We test this hypothesis by estimating a model of convergence of the excess demand estimated using bid-ask relative difference[23] to the actual excess demand at the call price.

We use the difference between (bid-ask based) excess demand and the actual excess demand at the call as a dependent variable in regression (5) that was previously used to estimate convergence to equilibrium price and volume. The resulting table is similar to Table 3 (see Table A2 in Appendix A.) The null of perfect convergence is rejected at call 1, since the estimate of the asymptote coefficient in regression (5) applied to the difference between the bid-ask arrival excess demand and the actual excess demand at call 1 is 7.620, significantly different from zero. However, we cannot reject convergence at call 2: the asymptote of 3.904 is insignificant. Figure 6 illustrates that revelation of excess demand via bid-ask arrival differences improves over time approaching the actual excess demand, while substantial variance in estimated excess demand at call 1 precludes tight convergence.[24]

Result 5: The price movement toward the equilibrium can be described by the Newton method of solving systems of equations. However, Newton does not capture well the price change across the calls. The relation is significant and particularly strong for predicting the equilibrium price given the actual induced parameters, excess demand and excess demand slope evaluated at the call price.

Support: The theory behind this result is described in subsection IIIC. There are several ways to estimate Newton (3). The simplest one is to use the actual excess

[23] To obtain this, we added up the bid-ask arrival differences over the last 30 seconds before each call. We chose this time interval to capture the most intense period of trading activity shortly before a call.

[24] Spearman rank correlation between excess demand and bid-ask arrival differences is 0.141 at call 1, and 0.139 at call 2.

Table 6—Estimation of Newton at Each Call Using Actual Excess Demand and Its Slope

	Call 1		Call 2
Regressor	Dep. var.: $p^* - p_1$	Dep. var.: $p_2 - p_1$	Dep. var.: $p^* - p_2$
Newton part (δ_1)	−0.991 (0.090)	−0.441 (0.039)	−1.070 (0.049)
Observations	108	108	108
RMSE	6.142	5.904	3.171
Newton hypothesis	Not rejected	Rejected at < 0.001 level	Not rejected

Notes: Bootstrap-corrected estimates of regression (7) terms (using 10,000 replications) with their standard errors in parentheses. The hypothesis $H_0 : \delta_1 = -1$ is bootstrap-tested.

demand and slope as dictated by parameters and the independent variables and directly estimate a linear relation

$$p_{t+1} - p_t = \delta_1 \frac{ED(p_t)}{ED'(p_t)} + \varepsilon, \tag{7}$$

where $t = 1, 2$ is the call number, $ED(p_t)$ is excess demand at price p_t, $ED'(p_t)$ is the slope of excess demand at p_t, and ε is the random error, and then test if $\delta_1 = -1$. Alternatively, one can add the intercept

$$p_{t+1} - p_t = \delta_0 + \delta_1 \frac{ED(p_t)}{ED'(p_t)} + \varepsilon, \tag{8}$$

and test the joint hypothesis $H_0 : \delta_0 = 0, \delta_1 = -1$.

There are also at least three choices for the time difference in (7) and (8). For $t = 1$, we have a choice between setting p_{t+1} equal to the price at the second call and the equilibrium price. For $t = 2$, we set p_{t+1} equal to the equilibrium price in the current period, since books are cleared each period.

Table 6 presents the main estimation results for the Newton regression (7) using the actual excess demand and its slope.

The main insight from Table 6 is that there is a significant relation between the Newton term and the price dynamics. More specifically, when regression (7) is estimated for the difference between the equilibrium price and the price at either call, the null hypothesis of perfect Newton dynamics is not rejected. When we estimate (7) for the difference between the call prices, the null of perfect Newton is rejected, indicating that the price change across calls is too large to approximate the instantaneous rate of change sufficiently well as required by Newton.

Thus, there is strong evidence for a Newton structure. The price movement is to the equilibrium but not to the next call price (unless it is near the equilibrium), which emphasizes that the Newton method is a theory of equilibration, not a theory of price movement independent of its equilibration tendencies. In particular, the price change across calls in a period seems to incorporate factors in addition to the Newton term, like, e.g., strategic considerations.

Table 7—Estimation of Walras at Each Call Using Actual Excess Demand and Its Slope

	Call 1		Call 2
Regressor	Dep. var.: $p^* - p_1$	Dep. var.: $p_2 - p_1$	Dep. var.: $p^* - p_2$
Walras part (γ_1)	1.474	0.558	1.562
	(0.085)	(0.106)	(0.067)
Observations	108	108	108
RMSE	5.162	7.205	3.550
Walras hypothesis	Not rejected	Not rejected	Not rejected

Notes: Bootstrap-corrected estimates of regression (9) terms (using 10,000 replications) with their standard errors in parentheses. The null hypothesis $H_0 : \gamma_1 \geq 0$ versus alternative $H_a : \gamma_1 < 0$ is bootstrap-tested.

Result 6: Walrasian model based on excess demand given by parameters at the call price also explains price change in a period but less accurately than Newton.

Support: The simplest way to estimate Walrasian adjustment (4) is to directly estimate a linear relation using the theoretical excess demand as dictated by parameters as the independent variable,

$$p_{t+1} - p_t = \gamma_1 ED(p_t) + \varepsilon, \tag{9}$$

and then test $H_0 : \gamma_1 \geq 0$ versus one-sided alternative $\gamma_1 < 0$.

Thus, the fundamental difference from Newton adjustment is that (9) does not utilize the excess demand slope.

Table 7 shows the results of the model estimation.

While we observe a significant and positive coefficient on the excess demand, γ_1, as predicted by Walras, the overall regression fit, as measured by the root-mean-square error (RMSE), is worse than that produced by Newton (7) in Table 6 in two out of three cases, estimated from the same data (RMSE is larger). The exception is the estimate of call 1 price in which the RMSE is 5.162 for Walras and 6.142 for the Newton model. Thus, Newton seems to outperform Walras adjustment by utilizing the information about the slopes of excess demand in price adjustment between calls and after the second call.

Parameter Information Sensitivity and Relative Model Performance.—When the structure of the Newton method of solving systems of equations is supplemented by the behavioral features of Jaws a new, Newton-Jaws, model takes shape. The Newton-Jaws model merges two variables known to be associated with price discovery, excess demand and order flow. Results 4 and 5 together with the models from subsections IIIB and IIIC demonstrated that these variables both can stand alone and provide the ingredients for useful models of market movement. Each provides its own view of market adjustment. However, a more powerful model emerges when the two types of variables become integrated into the Newton-Jaws model. Result 4 and Result 5, together with (7) outline a precise way this can be done.[25]

This seems particularly relevant for settings where the underlying supply and demand parameters are not observed by the econometrician, e.g., in the field. In the analysis that follows, we investigate the relative performance of Newton-Jaws and its sensitivity to the information about parameters. The model makes very precise predictions about the price change, and so should be easy to reject in the data. Nevertheless, Result 7 demonstrates that Newton-Jaws is on par with less precise adjustment models, like Walras and excess bids, that only make predictions about the sign of the price change.

Result 7: For predicting price change in a period, the Newton-Jaws model fit is similar to that of Walrasian and excess bids, and better in all three of our price change comparisons when estimated conditioned on known excess demand and jaws-estimated excess demand slope.

Support: The excess bids model is examined through the application of the same methodology as used to test the Newton-Jaws model and the Walrasian excess demand model. Let $XB(p_t)$ be the total number of buy orders (bids) minus the total number of sell orders (asks) existing in the market at time t. With t being the end of second call, we should note that the excess bid measure includes all bids and asks at the call, including those unfilled orders that remained from the first call.

We estimate

$$p_{t+1} - p_t = \beta_1 XB(p_t) + \varepsilon. \tag{10}$$

Table 8 shows the results of the model estimation.

The results in Table 8 demonstrate that the model of price dynamics based on the number of buy orders (bids) minus the number of sell orders (asks) cannot be rejected.

Both Walras and Newton show a better overall fit than excess bids when evaluated using the true parameters, as indicated by uniformly lower RMSE in Tables 6 and 7 compared to Table 8. Since excess bids do not rely on experiment parameters, it is important to check whether the better fit of Walras and Newton continues to hold when the true parameters are unknown, as in the field. Therefore, we also compared the fit when these models are evaluated conditioned on estimated measures of excess demand and its slope, so that all three models are on an equal footing. This comparison is reported below (see Tables A3 and A5 in Appendix A for additional details).

[25] Namely, Result 5 shows support for the Newton method of price adjustment across calls when the right-hand side of equation (7) is evaluated conditioned on known experimental parameters. Jaws provides a way to recover the Newton part in (7) directly from the data, as we established in Result 4, and allow us to estimate these quantities in various combinations. First, we can take both slopes and excess demand estimated from the jaws. Second, we can take the jaws-estimated slopes and use the true excess demand. Third, we can take the true slopes and use the jaws-estimated excess demand. Finally, for jaws-based slope estimates, we can use either the fixed revealed jaws, or the best-fitted revealed jaws. All of these cases are reported in Tables A3 and A4 in Appendix A, which include Table 6 as a special case.

Table 8—Estimation of Price Dynamics Using Excess Bids

	Call 1		Call 2
Regressor	Dep. var.: $p^* - p_1$	Dep. var.: $p_2 - p_1$	Dep. var.: $p^* - p_2$
Excess bid part (β_1)	1.413	0.576	0.612
	(0.340)	(0.145)	(0.247)
Observations	108	108	108
RMSE	19.368	10.042	13.954
Excess bid hypothesis	Not rejected	Not rejected	Not rejected

Notes: Bootstrap-corrected estimates of regression (10) terms (using 10,000 replications) with their standard errors in parentheses. The null hypothesis $H_0 : \beta_1 \geq 0$ versus alternative $H_a : \beta_1 < 0$ is bootstrap-tested.

Table 9—Fit of Newton-Jaws, Walras, and Excess Bids Using Empirical Measures of Excess Demand (ED) and Its Slope

	Call 1		Call 2
Model	Dep. var.: $p^* - p_1$	Dep. var.: $p_2 - p_1$	Dep. var.: $p^* - p_2$
Excess bids	19.638	10.042	13.954
Walras	21.166	10.529	14.071
Newton-Jaws:			
ED_{jaws}, $ED\ slope_{jaws}$	20.670	10.493	14.147
ED_{jaws}, true $ED\ slope$	21.121	10.468	14.288
true ED, $ED\ slope_{jaws}$	11.878	7.920	8.044

Notes: For each model in the first column, RMSE fit is reported in the corresponding cell. Subscript "jaws" indicates the quantity estimated from jaws as opposed to true parameters.

From Table 9, we determine that the Newton-Jaws model is more accurate than either Walras or excess bids when jaws are used to estimate excess demand slopes (Newton RMSE of 11.877, 7.920, and 8.044 versus excess bids RMSE of 19.638, 10.042, and 13.954, respectively, and Walras RMSE of 21.121, 10.468, and 14.288, respectively), and similar but slightly less accurate when jaws are used to estimate excess demand. In the latter case, Newton-Jaws is a bit more accurate than Walras (Newton RMSE is lower except at call 2.) Newton is slightly worse than excess bids when both excess demand and its slope are jaws-estimated, and slightly better than Walras except at call 2. Thus, Newton-Jaws fit is similar to both of these alternatives.

In other words, the empirical variant of Newton performs at least as well as alternative empirical models. Since at the same time, Newton predictions are much more precise than those of the alternatives, Newton overall performance is strictly better.

An explanation of the accuracy differences among the three models when comparably evaluated can be provided by adding an assumption about the subjects' bidding strategies. Given the nature of bidding strategies as postulated by Jaws, the different measures bring different information content to the model as follows.

Excess demand provides no direct information about the distance of the price from the equilibrium price. The excess demand measure as contained in the parameters evaluated at a price, contains only the qualitative information in the sign of the excess demand, which suggests an upward or downward movement. By contrast, excess bids reflect behaviors and contain limited information about excess demand as well as some information about the distance from the equilibrium price. However, the bids and asks order flow can also reflect unrealistic expectations, attempted signals, and other complex phenomena, so the quality of the information when aggregated can be poor. Under an assumption about the nature of tenders, the excess bids can contain information about the sign of the excess demand as well as limited information about the excess demand slope: if traders tend to restrict bids and asks to those for which values reside within a common, fixed interval of the price, then as predicted by the Jaws model, the total number of bids or asks placed will increase with the inverse of the respective slopes. Thus, according to the model, the relative numbers of bids over asks contain more information about limit values than just excess demand.

The difference of information about limit values differentiates the information content supplied to a model by the two variables, excess bids and excess demand, and their integration by the Newton-Jaws. The information in excess bids is indirect since it depends on the consistency with which bidders submit bids and asks given their incentives and the excess bids model does not have the information needed to produce a calculation. By contrast, in the Newton-Jaws, the information about limit values exists directly and separately to be used by the model. Thus, the key information about both excess demand and slopes of the demand and supply is contained in the Newton-Jaws model.[26] Given the microstructure of the price determination in the call markets the information is sufficient to provide a prediction of both direction and magnitude of price movement toward the competitive equilibrium.

VI. Conclusion

This paper initiates an investigation of principles of price adjustment in experimental multiple-call, multiple-unit markets. As such, it extends other research challenged by the possibility that call markets might provide a tool that helps solve problems encountered in markets operating in field environments. The challenge is made complex by wide ranging institutional features that can be assembled in many different configurations to create alternative call market architectures. The strategy is to experimentally probe theories of how selected institutions work together.

We report evidence that multiple calls, the shape of the associated order book, and a natural profit-maximizing behavior of individual traders organize themselves to produce an underlying price discovery process similar to one iterative step of a powerful tool for finding solutions to systems of equations (the Newton-Raphson method).

More specifically, we ask two main general questions:

- Do call market exchanges converge to the classical demand and supply?
- Do major patterns of convergence follow those suggested by market jaws and Newton?

[26] Table 9 shows how the goodness of fit of the Newton regression (7) changes as actual parameter values are replaced with their data-driven estimates. See also Table A3 in Appendix A.

We provide positive answers to both questions. Market behavior is captured by the competitive equilibrium model. The shape of the order book, captured by market jaws, reveals useful information about the slope of excess demand, which becomes part of the price change dynamics across calls. The change in the price toward equilibrium follows a single iteration of the Newton method for solving equations remarkably closely and produces a Newton-Jaws model. While the Walrasian adjustment, which does not include slopes of excess demand, finds support in the data, structural tests of the two models demonstrate that the Newton-Jaws model provides a better description of how the markets operate. At the same time, we conduct nonstructural tests of the models like those that might be possible in field applications, and find that the performance of both models is similar. Interestingly the performance of the two models in that testing environment is also similar to the excess bids model that has price changes predicted by the difference between the total number of bids and asks. However, close examination of the excess bids model suggests that the reason for its predictive power resides in its close proximity to excess demand and the market jaws.

Our results reveal a systemic compatibility between the self-organizing and coordinating features that emerge from individual behavior and the institutional features that guide it. The combination shows that a price discovery process can be related to the Newton method based on the order flow approximated by the market jaws. The question suggested is whether or not other market features can be combined with even more powerful tools to produce better market performance. Are there methods better than Newton when put to this purpose? What institutional modifications might be needed to establish compatibility?

Appendix A. Additional Estimation Details

This Appendix supplies additional estimation details for the results in Section V.

A1. *Algorithm for Testing Slope Revelation by Market Jaws*

For brevity, the description here focuses on supply only. Stages I and II address measurement challenges created by the discrete nature of the data. Stage III tests whether the estimated slopes reveal the true ones.

Stage I: For each data point $i \in \{1, \ldots, N\}$ corresponding to the first call in a period (subindex i is suppressed below):

(i) Estimate the slope of the revealed supply at p_1, $\hat{S}_1$, by best fitting $\tau + 1$ remaining (i.e., unfilled) asks in the book with a line that goes through the actual price-quantity point at the first call, (p_1, q_1). That is, estimate a regression of the form

$$u_k - q_1 = \hat{\theta}(a_k - p_1),$$

where $a_k > p_1$ is the kth remaining ask in the book, $k \in \{0, \ldots, \tau\}$, u_k is the corresponding kth unit, and $\hat{\theta} = \cot\hat{\alpha}$ is the estimate of the slope of $\hat{S}_1$. Keep $\hat{\theta}$.

(ii) Estimate the slope of the true supply at p_1, $S(p_1)$, by best fitting $2\tau'$ private costs around the true supply curve at p_1 (i.e., those $c_k \in [p_1 - \Delta, p_1 + \Delta]$ for some $\Delta > 0$ and each $k \in \{-\tau', \ldots, 0, \ldots, \tau'\}$) with a line that goes through supply at the actual price at the first call, $(p_1, S(p_1))$. That is, estimate a regression of the form

$$u_k - S(p_1) = \hat{\theta}^* (c_k - p_1),$$

where c_k is the kth component in the vector $(c_{-\tau'}, \ldots, c_0, \ldots, c_{\tau'})$ of costs for a fixed τ' around the true supply[27] at p_1, with $c_{-\tau'} \leq \cdots \leq c_0 \leq \cdots \leq c_{\tau'}$; u_k is the corresponding kth unit around $S(p_1)$, and $\hat{\hat{\theta}}_1^* = \cot\hat{\alpha}^*$ is the estimate of the slope of $S(p_1)$. Keep $\hat{\theta}^*$.

(iii) Record the pair $(\cot\hat{\alpha}_i^*, \cot\hat{\alpha}_i)$ as one observation in the new dataset.

Stage II: Using the data constructed at stage I, estimate

$$\cot\hat{\alpha}_i = \gamma_1 \cot\hat{\alpha}_i^* + \varepsilon_i, \tag{A1}$$

where ε_i is the random error. We bootstrapped the regression in (A1) using 10,000 replications.

Stage III: Test the null hypothesis $H_0 : \gamma_1 = 1$ versus the two-sided alternative. If the null is not rejected, then the slope of the remaining supply after the first call perfectly reveals the true supply slope. Alternatively, we can add an intercept γ_0 in (A1) and test the joint null $H_0 : \gamma_0 = 0,\ \gamma_1 = 1$.[28] If the null is rejected but γ_1 is positive and significant, the slope is revealed imperfectly, still providing some useful information about the underlying market parameters.

The coherence of true supply and revealed supply slopes after the second call, as well as coherence of true demand and revealed demand after each call, can be estimated and tested in the same way.

The procedure above is silent about how τ and τ' are specified. For τ' this is not an issue, since the equilibrium demand/supply at any price is recovered from our parameters, and is close to a linear curve in all of our experimental sessions. To have a sufficiently smooth approximation, we chose $\tau' = 7$, thus, estimating the equilibrium slopes on 14 points around the point where the equilibrium curve intersects with the actual price.

[27] For example, if we take three true costs below p_1 and three true costs above p_1, then $\tau' = 3$.
[28] We checked both specifications and found minimal differences.

To determine τ, we used two approaches. In the first one, we exogenously fixed τ at three-fourths of excluded orders for each data point $i \in \{1, \ldots, N\}$. This helps exclude the extreme orders that may have large impact on linear estimates. In the second one, we repeated steps (i) $-$ (ii) of stage I, varying τ from 3 units[29] up to the length of the book (in units) at the call in question, and then picked the value that produced the best match (in terms of minimizing the absolute difference) between the slopes of equilibrium demand/supply and the slopes of the actual data at that price, estimated using τ units of remaining demand/supply. Thus, in this case, we best-fitted the slopes for individual data points of the regression at stage II. Note that this does not automatically imply that regression (A1) is trivial, since different data points may require different values of τ. This approach nests the estimation technique from Asparouhova, Bossaerts, and Plott (2003), who find limited support for the jaws (in the form of correlation between the order book and excess demand) using a small and exogenously set τ.

Table A1—Estimation of Supply/Demand Slope Revelation at Each Call by Market Jaws

	$S(p_1)$		$S(p_2)$	
Variable	$\tau =$ best-fit	$\tau =$ fixed	$\tau =$ best-fit	$\tau =$ fixed
γ_1	0.762 (0.043)	0.693 (0.065)	0.823 (0.030)	0.661 (0.045)
R^2_{adj}	0.786	0.574	0.891	0.710
H_0	Rejected at $<$ 0.001 level	Rejected at $<$ 0.001 level	Rejected at $<$ 0.001 level	Rejected at $<$ 0.001 level
Observations	101	101	101	101
	$D(p_1)$		$D(p_2)$	
Variable	$\tau =$ best-fit	$\tau =$ fixed	$\tau =$ best-fit	$\tau =$ fixed
γ_1	0.966 (0.052)	1.518 (0.203)	1.024 (0.043)	1.005 (0.104)
R^2_{adj}	0.751	0.326	0.809	0.468
H_0	Not rejected	Rejected at $<$ 0.05 level	Not rejected	Not rejected
Observations	101	101	101	101

Notes: Bootstrap-corrected estimates of regression (A1) main terms and their standard errors are in parentheses, using 10,000 replications. The null hypothesis $H_0 : \gamma_1 = 1$ is bootstrap-tested.

[29] We chose three as the minimal number of units that allows a non-singular OLS fit.

Table A2—Estimation of Excess Demand Revelation by Jaws: Convergence of Bid-Ask Arrival Difference to the True Excess Demand at Each Call

Regressor	Dependent variable: Excess demand difference			
	Call 1		Call 2	
Asymptote (γ)	7.555	(2.086)	3.883	(1.492)
Origin (before shift) in				
session 1 (β_1)	−24.013	(21.036)	−10.993	(13.984)
— 2 (β_2)	4.661	(7.066)	3.842	(9.944)
— 3 (β_3)	−36.091	(23.792)	−14.199	(6.332)
— 4 (β_4)	−21.816	(10.822)	−1.666	(7.335)
— 5 (β_5)	−48.882	(10.664)	−37.351	(8.960)
— 6 (β_6)	−59.861	(26.377)	−29.484	(18.563)
— 7 (β_7)	−2.616	(10.467)	5.470	(11.578)
Origin (after shift) in				
session 1 (δ_1)	−35.522	(16.313)	−4.510	(7.214)
— 2 (δ_2)	−40.565	(9.519)	−25.600	(4.650)
— 3 (δ_3)	11.767	(17.212)	−6.184	(3.268)
— 4 (δ_4)	−2.134	(8.397)	−6.213	(7.223)
— 5 (δ_5)	−16.124	(13.886)	−12.023	(8.145)
— 6 (δ_6)	−52.884	(14.898)	−28.367	(8.096)
— 7 (δ_7)	−25.708	(9.739)	−12.985	(6.329)
Observations	94		94	
R^2_{adj}	0.614		0.547	
F-stat	12.570		9.774	

Notes: Bootstrap standard errors are in parentheses. Estimates are bootstrap-corrected for bias.

Table A3—Estimation of Newton (7) at Each Call

Regressor	S, ED	S, ED_j	S_j^{fx}, ED	S_j^{bf}, ED	S_j^{fx}, ED_j	S_j^{bf}, ED_j
Call 1, dependent variable $p^* - p_1$						
δ_1	−0.991	−0.520	−0.784	−0.026	−0.140	−0.231
	(0.090)	(0.157)	(0.096)	(0.244)	(0.079)	(0.056)
R^2_{adj}	0.922	0.073	0.707	0.343	0.008	0.112
RMSE	6.142	21.121	11.878	17.780	21.850	20.670
H_0	Not rejected	Rejected at < 0.01 level	Rejected at < 0.05 level	Rejected at < 0.001 level	Rejected at < 0.001 level	Rejected at < 0.001 level
Call 1, dependent variable $p_2 - p_1$						
δ_1	−0.441	−0.257	−0.316	−0.008	−0.084	−0.090
	(0.039)	(0.090)	(0.065)	(0.105)	(0.041)	(0.030)
R^2_{adj}	0.706	0.075	0.470	0.222	0.016	0.070
RMSE	5.904	10.468	7.920	9.599	10.791	10.493
H_0	Rejected at < 0.001 level	Rejected at < 0.001 level	Rejected at < 0.001 level	Rejected at < 0.001 level	Rejected at < 0.001 level	Rejected at < 0.001 level
Call 2, dependent variable $p^* - p_2$						
δ_1	−1.070	−0.743	−0.630	−0.874	−0.377	−0.763
	(0.049)	(0.319)	(0.105)	(0.054)	(0.224)	(0.308)
R^2_{adj}	0.955	0.079	0.708	0.909	0.029	0.097
RMSE	3.171	14.288	8.044	4.488	14.668	14.147
H_0	Not rejected	Not rejected	Rejected at < 0.01 level	Rejected at < 0.05 level	Rejected at < 0.01 level	Not rejected
Observations	108	108	108	108	108	108

Notes: Bootstrap-corrected estimates of regression (7) terms with their standard errors using 10,000 replications are in parentheses. Notation: S is true slope, ED is true excess demand, j subscript stands for jaws-based, fx is using fixed book share for estimating jaws, and bf is using best-fitted jaws. The null of perfect Newton, $H_0 : \delta_1 = -1$, is bootstrap-tested.

Table A4—Estimation of Newton (8) at Each Call

Regressor	S, ED	S, ED_j	S_j^{fx}, ED	S_j^{bf}, ED	S_j^{fx}, ED_j	S_j^{bf}, ED_j
Call 1, dependent variable $p^* - p_1$						
δ_0	0.738	7.857	2.210	7.897	8.829	8.018
	(0.495)	(1.858)	(0.796)	(2.831)	(1.914)	(1.860)
δ_1	−0.976	−0.316	−0.784	0.000	−0.096	−0.200
	(0.095)	(0.139)	(0.093)	(0.246)	(0.064)	(0.043)
R^2_{adj}	0.907	0.024	0.657	0.298	0.000	0.088
H_0	Not rejected	Rejected at $<$ 0.001 level	Rejected at $<$ 0.05 level	Rejected at $<$ 0.001 level	Rejected at $<$ 0.001 level	Rejected at $<$ 0.001 level
Call 1, dependent variable $p_2 - p_1$						
δ_0	1.027	3.955	1.966	4.155	4.406	4.164
	(0.621)	(0.836)	(0.582)	(1.251)	(0.932)	(0.916)
δ_1	−0.422	−0.155	−0.284	0.002	−0.062	−0.073
	(0.044)	(0.080)	(0.064)	(0.100)	(0.036)	(0.025)
R^2_{adj}	0.652	0.024	0.390	0.172	0.006	0.047
H_0	Rejected at $<$ 0.001 level	Rejected at $<$ 0.001 level	Rejected at $<$ 0.001 level	Rejected at $<$ 0.001 level	Rejected at $<$ 0.001 level	Rejected at $<$ 0.001 level
Call 2, dependent variable $p^* - p_2$						
δ_0	0.113	3.632	1.411	0.148	4.098	3.491
	(0.220)	(1.341)	(0.609)	(0.383)	(1.340)	(1.351)
δ_1	−1.068	−0.579	−0.612	−0.872	−0.265	−0.619
	(0.049)	(0.310)	(0.105)	(0.055)	(0.210)	(0.310)
R^2_{adj}	0.950	0.046	0.688	0.900	0.011	0.062
H_0	Not rejected	Not rejected	Rejected at $<$ 0.01 level	Rejected at $<$ 0.05 level	Rejected at $<$ 0.001 level	Not rejected

Notes: Bootstrap-corrected estimates of regression (8) terms with their standard errors using 10,000 replications are in parentheses. Notation: S is true slope, ED is true excess demand, j subscript stands for jaws-based, fx is using fixed book share for estimating jaws, and bf is using best-fitted jaws. The joint hypothesis of perfect Newton, $H_0: \delta_0 = 0, \delta_1 = -1$, is bootstrap-tested.

Table A5—Estimation of Price Dynamics Using Excess Bids from the Last 30 Seconds Before Call

	Call 1		Call 2
Regressor	Dep. var.: $p^* - p_1$	Dep. var.: $p_2 - p_1$	Dep. var.: $p^* - p_2$
Excess bid part (β_1)	0.864	0.410	1.471
	(0.279)	(0.132)	(0.580)
Observations	108	108	108
RMSE	21.166	10.529	14.071
Excess bid hypothesis	Not rejected	Not rejected	Not rejected

Notes: Bootstrap-corrected estimates of regression (10) terms are based on the bids and asks in the last 30 seconds before a call (using 10,000 replications) with their standard errors in parentheses. The null hypothesis $H_0: \beta_1 \geq 0$ versus alternative $H_a: \beta_1 < 0$ is bootstrap-tested.

Appendix B. Robustness Check: Newton Applied to Zero-Intelligence Robots

In order to check the robustness of our main result, the Newton-like price dynamics across calls, we also simulated artificial data using zero-intelligence robots (see, e.g., Gode and Sunder 1993, Cason 1992). From the analysis of zero-intelligence

Table B1—Estimation of Newton at Each Call, Zero-Intelligence Data

	Call 1		Call 2
Regressor	Dep. var.: $p^* - p_1$	Dep. var.: $p_2 - p_1$	Dep. var.: $p^* - p_2$
Intercept (δ_0)	0.245	6.979	−1.045
	(0.114)	(0.288)	(0.131)
Newton part (δ_1)	−1.068	−1.154	−0.889
	(0.016)	(0.044)	(0.012)
$H_0 : \delta_0 = 0, \delta_1 = -1$	Rejected at 0.001 level	Rejected at < 0.001 level	Rejected at < 0.001 level

Notes: Table lists OLS estimates of regression (8) terms for ZI data ($N = 1{,}024$), shift fixed effects included, with standard errors in parentheses. The joint hypothesis $H_0 : \delta_0 = 0, \delta_1 = -1$ is tested.

trades, described below, we conclude that Newton is a property of institution rather than an outcome of strategic interaction, and is directly related to convergence towards the competitive equilibrium.

The simulations were done as follows. We ran 512 replications of each of pre-shift and post-shift experimental setting. In each replication period, there are 2 calls, and zero-intelligence buyers and sellers make random bids and asks as follows. Buyer bids for each unit are uniform from [110, *unit value*]. Seller asks for each unit are uniform from [*unit cost*, 250]. The unit values and costs are exactly as used in Sessions 5–7 of the experiment (see Table 2). Each of 15 buyers (three buyers for each of 5 buyer types) submits independent bids for all 6 units. Each of 15 sellers (three sellers for each of 5 seller types) submits independent asks for all 6 units.[30] At the first call, price and trade volume are determined based on the orders in the book as in the experiment. Then buyers and sellers make random bids and asks for the remaining units. At the second call, price and trade volume are determined based on the orders in the book, and the replication period ends.

Since bidding is completely random, only the price dynamics across calls are studied. The randomness also precludes the use of the difference between the number of bids and asks as a measure of revealed excess demand. Table B1 contains the main results based on the actual excess demand and its slope. Results using jaws-based slope estimates are reported in Table B2.

Table B1 demonstrates Newton predictions are robust even when applied to agent bids produced by zero intelligence. In all of the estimations the coefficient at the Newton term in (8), δ_1, has correct negative sign. At the same time, we reject the null hypothesis of full consistency with Newton: the constant term δ_0 is significantly different from zero. This seems to be due to the fact that tight convergence to competitive equilibrium (a necessary assumption for Newton adjustment to work) is lacking under random bidding. Nevertheless, note that price at the first call points toward equilibrium price from below ($\delta_0 > 0$), while price at the second call overshoots and points toward equilibrium from above ($\delta_0 < 0$). Importantly, the price movement indication (the Newton part) is toward equilibrium price in both cases.

[30] An equivalent representation is to assume that each buyer or seller only has 1 unit, and there are 90 buyers and 90 sellers.

Table B2—Estimation of Newton at Each Call, ZI Data

Regressor	S, ED	S_j^{fx}, ED	S_j^{bf}, ED
Call 1, dependent variable $p^* - p_1$			
δ_0	0.245	0.185	0.197
	(0.114)	(0.118)	(0.118)
δ_1	−1.068	−2.210	−1.211
	(0.016)	(0.042)	(0.027)
H_0	Rejected at $<$ 0.001 level	Rejected at $<$ 0.001 level	Rejected at $<$ 0.001 level
Call 1, dependent variable $p_2 - p_1$			
δ_0	6.979	6.915	6.928
	(0.288)	(0.288)	(0.288)
δ_1	−1.154	−2.392	−1.312
	(0.044)	(0.094)	(0.054)
H_0	Rejected at $<$ 0.001 level	Rejected at $<$ 0.001 level	Rejected at $<$ 0.001 level
Call 2, dependent variable $p^* - p_2$			
δ_0	−1.045	−1.642	−1.168
	(0.131)	(0.129)	(0.132)
δ_1	−0.889	−1.213	−0.885
	(0.012)	(0.016)	(0.012)
H_0	Rejected at $<$ 0.001 level	Rejected at $<$ 0.001 level	Rejected at $<$ 0.001 level

Notes: Table lists OLS estimates of regression (8) terms for ZI data ($N = 1{,}024$), shift fixed effects included, with standard errors in parentheses. Notation: *S* is true slope, *ED* is true excess demand, *j* subscript stands for jaws-based, *fx* is using fixed book share for estimating jaws, and *bf* is using best-fitted jaws. The joint hypothesis $H_0 : \delta_0 = 0$, $\delta_1 = -1$ is tested.

Appendix C: Instructions Emailed to the Participants

C1. *Experiment Overview*

There are two types of participants on the public market (labeled **'Market *X*'**): public buyers and public sellers. The public buyers place orders to buy and public sellers place orders to sell in the public *X* market. Public buyers have odd IDs and public sellers have even IDs.

The experiment uses a currency called "francs." The exchange rate between francs and real money that you get paid is fixed and will be announced at the start of the experiment.

The experiment consists of several six-minute periods. Periods are independent from each other, and your payment is based on your total earnings in all periods. This means that you should try making profit in each period, but if you make a mistake and lose in some periods, you'll have a chance to recover in future periods.

Before the actual experiment begins, there will be three low-paid practice rounds. If you are consistently losing money during the practice rounds, you will be declared bankrupt and the system will block you from further participation.

You have been guaranteed a minimum, but you will receive it only if you participate for the full duration of the experiment.

C2. *Prices and Calls*

All prices are determined at (and only at) the time of market calls. In each period, there are two market calls. One is at 1.5 minutes after a period begins and one is at 4.5 minutes. During the period, the order book of the public X market accumulates buy and sell orders from public buyers and public sellers, but trade in the public X market can only happen at a market call. If you trade, it will happen at the market price, not the price that you state in your orders.

At each call, all buy and sell orders in the order book are simultaneously considered and a market price is established. It is determined as follows:

- Based on all orders in the book, the system sorts buy orders by their respective prices per unit from high to low. Sell orders are sorted by their respective prices per unit from low to high.
- The system matches the two sorted series selecting all pairs for which the purchase price is greater than the sale price, and stops at the last pair for which this is true.
- The market price is calculated midway between the last accepted (the lowest filled) buy order and the last accepted (highest filled) sell order. Except for ties, all buy orders with prices above the market price will trade at the market price. All sell orders with prices below the market price will trade at the market price. All other orders will remain unfilled.

This means that if your buy order has a price below the market price, it won't trade. Similarly, if your sell order has a price above the market price, it won't trade. However, your order may change the expected market price, if, once inserted in the sorted series, it changes the intersection point (the last pair of matched orders where the purchase price is greater than the sale price).

Ties happen when there are several orders at the same price per unit in the order book and the total quantity demanded at the market price does not match the total quantity supplied at that price. In this case, the orders at the market price are filled in the first come first served manner as long as there is a match.

If all this sounds too technical, just remember that the market price is based on all orders present in the book at the time of the call, and trade only happens at the call.

You can always see the current market price if it exists (based on the orders currently present in the book) in the Best Buy/Best Sell Offer columns. The number before the symbol indicates the total number of units available at the corresponding Best Buy/Best Sell price.

The first period will have an additional five minutes before the first call so you have time to figure out what to do. The end of the experiment will be announced without warning after the last period.

C3. *Information for Buyers*

If you are a PUBLIC BUYER (an odd ID number): you will buy units in the public X market and collect the values in your Value Opportunities List, just like a middleman buying in one and selling in the other.

Profit on unit = Value Opportunity of unit − Price paid for unit in the public X market.

You can lose money if you pay a price in the X market that is higher than your value for the unit, so make sure to look up your Value Opportunities List before you buy. You can also miss an opportunity to make money if you do not buy enough in the public X market when it is profitable to you. Notice that if you submit a multi-unit order, your profit will likely be different for each consecutive unit, so you can lose if the total profit from a multi-unit order is negative.

You cannot buy more than six units in each call, and if you run out of Values in your Value Opportunities List, all units in excess will be redeemed by the system at the end of the period at the worst possible price to you.

Values in your Value Opportunities List can change and do expire (each has a time tag). Refresh the frame to see an updated time tag (no automatic update). You do not need to collect all of the values in your list just because they are there. Some can be bad deals, depending on the public X market so that they can cause you to lose money.

At the end of the period your inventory is worthless. That is, if you simply spend money and accumulate units of X in your inventory at the end of a period, you will lose what you have spent. Neither francs nor inventory will store across periods, only your profits.

C4. *Information for Sellers*

If you are a PUBLIC SELLER (an even ID number): You will short sell units in the public X market that you afterwards procure at a cost from your Cost Opportunities List. You are just like a middleman who short sells in one market, and then, after figuring out how much to deliver, buys back in the other market.

Profit on unit = Price received for unit in the X market − cost of unit from your Cost Opportunities List.

Your optimal strategy may seem a bit tricky because it involves short selling. You should sell units in the public market before you actually have them. Of course, your inventory will go negative until you procure the units needed to cover your sales. Once you trade in the public X market, you should procure the units using your Cost Opportunities List to cover what you have sold. This strategy allows you to avoid the risk from trade, because if you did otherwise, i.e., first paid the cost of the units and then tried selling them in the public X market, you would be likely to lose money, as the market price might happen to be less than your cost and your units won't sell.

You can lose money if you sell in the X market at a price that is lower than your cost, so make sure to look up your Cost Opportunities List before your sell. You can also miss an opportunity to make money if you do not sell enough in the public X market when it is profitable to you. Notice that if you submit a multi-unit order, your profit will likely be different for each consecutive unit, so you can lose if the total profit from a multi-unit order is negative.

You cannot short sell more than six units in each call, and if you run out of Costs in your Cost Opportunities List, all remaining standing units will be covered by the system at the end of the period at the worst possible price to you.

Costs in the Cost Opportunities List can change and do expire (each has a time tag). Refresh the frame to see an updated time tag (no automatic update). You do not need to use all of the costs in your list just because they are there. Some can be bad deals, depending on the public *X* market.

At the end of the period your inventory is worthless. That is, if you accumulate debt (i.e., negative inventory) at the end of a period, your cash on hand will be spent to cover it, so you will lose what you have earned in that period. Neither francs nor inventory will store across periods, only your profits.

C5. *Practice and Demo*

Instructions for the trading technology and practice are available at URL.

It is in your best interest to understand how this program works.

Do not confuse the experiment and the demo. You cannot participate in the experiment from the demo page.

It is possible that your computer will not be able to load the demo. If your computer can load the demo, then it can load the experiment.

If you have any further questions, please e-mail us at EMAIL.

REFERENCES

Asparouhova, Elena, Peter Bossaerts, and Charles Plott. 2003. "Excess demand and equilibration in multi-security financial markets: The empirical evidence." *Journal of Financial Markets* 6 (1): 1–21.

Barner, Martin, Francesco Feri, and Charles R. Plott. 2005. "On the microstructure of price determination and information aggregation with sequential and asymmetric information arrival in an experimental asset market." *Annals of Finance* 1 (1): 73–107.

Bossaerts, Peter, and Charles R. Plott. 2008. "From Market Jaws to the Newton Method: The Geometry of How a Market Can Solve Systems of Equations." In *Handbook of Experimental Economics Results*, Vol. 1, edited by Charles R. Plott and Vernon L. Smith, 22–24. Amsterdam: North-Holland.

Brewer, Paul, Jaksa Cvitanic, and Charles R. Plott. 2013. "Market Microstructure Design and Flash Crashes: A Simulation Approach." *Journal of Applied Economics* 16 (2): 223–50.

Budish, Eric, Peter Cramton, and John Shim. 2015. "The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response." *Quarterly Journal of Economics* 130 (4): 1547–1621.

Cason, Timothy N. 1992. "Call market efficiency with simple adaptive learning." *Economics Letters* 40 (1): 27–32.

Cason, Timothy N., and Daniel Friedman. 1997. "Price Formation in Single Call Markets." *Econometrica* 65 (2): 311–45.

Cason, Timothy N., and Daniel Friedman. 1999a. "Learning in a Laboratory Market with Random Supply and Demand." *Experimental Economics* 2 (1): 77–98.

Cason, Timothy N., and Daniel Friedman. 1999b. "Price Formation and Exchange in Thin Markets: A Laboratory Comparison of Institutions." In *Money, Markets and Method: Essays in Honour of Robert W. Clower*, edited by Peter W. Howitt, Elisabetta de Antoni, and Axel Leijonhufvud, 155–79. Cheltenham, UK: Edward Elgar.

Cason, Timothy N., and Daniel Friedman. 2008. "A Comparison of Market Institutions." In *Handbook of Experimental Economics Results*, Vol. 1, edited by Charles R. Plott and Vernon L. Smith 264–72. Amsterdam: North-Holland.

Cason, Timothy N., and Charles R. Plott. 1996. "EPA's New Emissions Trading Mechanism: A Laboratory Evaluation." *Journal of Environmental Economics and Management* 30 (2): 133–60.

Friedman, Daniel. 1993. "How Trading Institutions Affect Financial Market Performance: Some Laboratory Evidence." *Economic Inquiry* 31 (3): 410–35.

Friedman, Daniel, and Joseph Ostroy. 1995. "Competitivity in Auction Markets: An Experimental and Theoretical Investigation." *Economic Journal* 105 (428): 22–53.

Garvin, Susan, and John H. Kagel. 1994. "Learning in common value auctions: Some initial observations." *Journal of Economic Behavior and Organization* 25 (3): 351–72.

Gjerstad, Steven. 2013. "Price dynamics in an exchange economy." *Economic Theory* 52 (2): 461–500.

Gode, Dhananjay K., and Shyam Sunder. 1993. "Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality." *Journal of Political Economy* 101 (1): 119–37.

Hirota, Masayoshi, Ming Hsu, Charles R. Plott, and Brian W. Rogers. 2005. "Divergence, Closed Cycles and Convergence in Scarf Environments: Experiments in the Dynamics of General Equilibrium Systems." Caltech Humanities and Social Sciences (HSS) Working Paper 1239.

Horowitz, Joel L. 2001. "The Bootstrap." In *Handbook of Econometrics*, Vol. 5, edited by James J. Heckman and Edward Leamer, 143–44. Amsterdam: North-Holland.

Johnson, Alonzo, Hsing Yang Lee, and Charles R. Plott. 1989. "Multiple Unit Double Auction User's Manual." Caltech Humanities and Social Sciences (HSS) Working Paper 676.

Kagel, John H. 2004. "Double Auction Markets with Stochastic Supply and Demand Schedules: Call Markets and Continuous Auction Trading Mechanisms." In *Advances in Understanding Strategic Behaviour: Game Theory, Experiments and Bounded Rationality*, edited by Steffen Huck, 181–208. New York: Palgrave Macmillan.

McCabe, Kevin A., Stephen J. Rassenti, and Vernon L. Smith. 1993. "Designing a uniform price double auction: An experimental evaluation." In *The Double-Auction Market: Institutions, Theories and Evidence*. Sante Fe Institute Studies in the Sciences of Complexity, edited by Daniel Friedman and John Rust, 307–32. Cambridge: Perseus.

McKenzie, Lionel W. 2002. *Classical General Equilibrium Theory*. Cambridge: MIT Press.

Noussair, Charles N., Charles R. Plott, and Raymond G. Riezman. 1995. "An Experimental Investigation of the Patterns of International Trade." *American Economic Review* 85 (3): 462–91.

Plott, Charles R., and Glen George. 1992. "Marshallian vs. Walrasian Stability in an Experimental Market." *Economic Journal* 102 (412): 437–60.

Plott, Charles R., and Peter Gray. 1990. "The multiple unit double auction." *Journal of Economic Behavior and Organization* 13 (2): 245–58.

Plott, Charles R., and Kirill Pogorelskiy. 2017. "Call Market Experiments: Efficiency and Price Discovery through Multiple Calls and Emergent Newton Adjustments: Dataset." *American Economic Journal: Microeconomics*. https://doi.org/10.1257/mic.20150201.

Plott, Charles R., Nilanjan Roy, and Baojia Tong. 2013. "Marshall and Walras, disequilibrium trades and the dynamics of equilibration in the continuous double auction market." *Journal of Economic Behavior and Organization* 94 (1): 190–205.

Plott, Charles R., and Vernon L. Smith. 1978. "An Experimental Examination of Two Exchange Institutions." *Review of Economic Studies* 45 (1): 133–53.

Saari, Donald G. 1985. "Iterative Price Mechanisms." *Econometrica* 53 (5): 1117–31.

Saari, Donald G., and Carl P. Simon. 1978. "Effective Price Mechanisms." *Econometrica* 46 (5): 1097–1125.

Satterthwaite, Mark, and Steven Williams. 1993. "The Bayesian Theory of the *k*-Double Auction." In *The Double-Auction Market: Institutions, Theories, and Evidence*. Sante Fe Institute Studies in the Sciences of Complexity, edited Daniel Friedman and John Rust, 99–124. Cambridge: Perseus.

Scarf, Herbert. 1960. "Some Examples of Global Instability of the Competitive Equilibrium." *International Economic Review* 1 (3): 157–72.

Selten, Reinhard, and Tibor Neugebauer. 2014. "Experimental stock market dynamics: Excess demand, adaptation, and style investing in a call-auction with multiple multi-period lived assets." Unpublished.

Selten, Reinhard, and Tibor Neugebauer. 2015. "Stock market 'puzzles' observed in the experimental call auction and continuous double auction market: A comparison of the two market institutions." Unpublished.

Smith, Vernon L. 1965. "Experimental Auction Markets and the Walrasian Hypothesis." *Journal of Political Economy* 73 (4): 387–93.

Smith, Vernon L., Gerry L. Suchanek, and Arlington W. Williams. 1988. "Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets." *Econometrica* 56 (5): 1119–51.

Smith, Vernon L., Arlington W. Williams, W. Kenneth Bratton, and Michael G. Vannoni. 1982. "Competitive Market Institutions: Double Auctions vs. Sealed Bid-Offer Auctions." *American Economic Review* 72 (1): 58–77.

van Bommel, Jos, and Peter Hoffmann. 2011. "Transparency and Ending Times of Call Auctions: A Comparison of Euronext and Xetra." University of Luxembourg School of Finance (LSF) Research Working Paper 11-09.

Wilson, Robert B. 1987. "On Equilibria of Bid-Ask Markets." In *Arrow and the Ascent of Modern Economic Theory*, edited by George R. Feiwel, 375–414. London: Palgrave Macmillan UK.