Trading Votes for Votes: A Laboratory Study*

Alessandra Casella[†] Thomas R. Palfrey[‡]

October 28, 2020

Abstract

Vote trading is ubiquitous in committees and legislatures, and yet we know very little about its properties. We explore this subject with a laboratory experiment. We propose a model of vote trading in which pairs of voters exchange votes whenever doing so is mutually advantageous. The resulting trading dynamics always converge to stable vote allocations–allocations where no further improving trades exist. The data show that stability has predictive power: vote allocations in the lab converge towards stable allocations, and individual vote holdings at the end of trading are in line with theoretical predictions. There is less support for the finer details of the trade-by-trade dynamics.

JEL Classification: D70, D72, P16

Keywords: Voting, Majority Voting, Vote Trading, Experiments.

*We thank Kirill Pogorelskiy, Manuel Puente, Enrico Zanardo, and Krzysztof Zaremba for research assistance. We thank Micael Castanheira, Parkash Chander, Andrew Gelman, Michel LeBreton, Cesar Martinelli, Debraj Ray, Richard Van Weelden, Rajiv Vohra, and Alistair Wilson for detailed comments and suggestions. We are grateful to the editor, the advisory editor, and three referees for helpful comments and suggestions. The National Science Foundation (SES-1426560, SES-0617934) and the Gordon and Betty Moore Foundation (SES-1158) provided financial support. Part of the research was conducted while Casella was a Straus Fellow at NYU Law School and Palfrey was a Visiting Scholar at the Russell Sage Foundation. The hospitality and financial support of both institutions are gratefully acknowledged. An earlier version of this paper was part of a working paper entitled "Trading Votes for Votes: A Decentralized Matching Algorithm".

†Departments of Economics and Political Science, Columbia University, 420 West 118th Street, New York, NY 10027. NBER and CEPR, ac186@columbia.edu

‡Corresponding author. Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125 and NBER, trp@hss.caltech.edu

1 Introduction

Considering the very rich literature on voting and committee decision-making, the scarcity of systematic studies on vote trading is remarkable. We use "vote trading" to indicate the exchange of votes on some issues for votes on other issues–lending support to somebody else's preferred position in exchange for that person's support of one's own preferred position on a different issue. Common sense, personal experience, and anecdotal evidence from legislative bodies all point to its extent and importance.

Vote trading is linked to fundamental questions in collective choice: Will trades lead to vote allocations where no further trade is desirable? If so, what efficiency properties will such allocations possess? In *Calculus of Consent* (1962), Buchanan and Tullock advocated vote trading as route to Pareto superior outcomes and conjectured that trading resolves the indeterminacy of voting outcomes in the absence of a Condorcet winner, i.e., an outcome that defeats all other outcomes under majority rule. Is such a conjecture correct? The 60's and 70's saw a flowering of theoretical studies, but the literature was hampered by the lack of a common framework. In the absence of broadly accepted results, analyses that modelled vote trading directly fizzled and eventually stopped.[1].

Not only is the study of vote trading of interest theoretically, but it has long been recognized as being especially significant for our understanding of political institutions. More than a century ago, Arthur F. Bentley argued that logrolling is vital to the practical business of legislatures, which would essentially cease to function if members of legislatures were unable or unwilling to trade votes:

"Log-rolling is a term of opprobrium... Log-rolling is, however, in fact, the most characteristic legislative process... It is compromise, not in the abstract moral form, which philosophers can sagely discuss, but in the practical form with which every legislator who gets results through government is acquainted. It is trading. It is the adjustment of interests. Where interests must seek adjustment without legislative forms, ...they have no recourse but to take matters in their own hands and proceed to open violence or war. When they have compromised and ...process can be carried forward in a legislature, they proceed to war on each other, with the killing and maiming omitted. It is a battle of strength, along

[1]See, among others, Coleman (1966, 1967), Park (1967), Wilson (1969), Tullock (1970), Haefele (1971), Kadane (1972), Bernholz (1974), Riker and Brams (1973), Mueller (1967, 1973), Koehler (1975), Miller (1977), Schwartz (1975, 1977). Systematic empirical evidence is scarce. Focusing on the US Congress, Mayhew (1966) studies agricultural bills in the House; Stratmann (1992, 1995) identifies roll call votes where a legislator votes against his constituency's interest and attribute a substantial fraction of such votes to vote trading. More recently, Guerrero and Matter (2016) measure the extent of vote trading by identifying reciprocity networks in roll call voting and bill cosponsorship.

lines of barter. The process is a similar process, but with changes in the technique. There never was a time in the history of the American Congress when legislation was conducted in any other way."

-from The Process of Government, 1908 (pp.370-371)

One reason for our imperfect understanding of vote trading is that the problem is difficult. Consider the simplest framework, the natural first step proposed by Riker and Brams (1973). A committee with an odd number of members considers several binary proposals, each of which may pass or fail; voters can trade votes with each other without enforcement or credibility problems; after trades are concluded, voting occurs by majority rule, proposal by proposal.[2] Every voter has separable preferences across proposals, with different cardinal intensities. Even in this restricted domain, vote trading is a difficult problem: trades take place without the equilibrating forces of a price mechanism, impose externalities on non-trading voters, change the overall distribution of votes, and with it other voters' incentives and power to affect outcomes and to induce further trades.

Addressing the basic questions raised earlier requires a rigorous definition of stability and a formal model of dynamic adjustment. In this paper we implement in an economics laboratory the general theoretical framework developed in Casella and Palfrey (2019). The theory is based on the concept of a *stable* vote allocation, with the property that no coalition of voters can reach a new allocation that all coalition members strictly prefer by trading votes among themselves. A feasible trade that leads the coalition to a preferred vote allocation is called a *blocking trade*. The framework implies a dynamic trading process with a sequence of blocking trades that continues until a stable (unblocked) allocation is reached. The key theoretical result is that there always exists a path of trades that leads to a stable vote allocation. The result holds regardless of whether there is a Condorcet winner.

The theoretical approach generates sharp predictions about final vote allocations, proposals' outcomes, and even exact sequences of trades. It is a natural framework for a laboratory experiment, where the environment can be implemented and manipulated with strict control and the detailed workings of the dynamic mechanism can be observed without confounding factors. In the simplified laboratory environment trade is restricted to be pairwise: a trade involves the exchange of one vote on one proposal for one vote on another proposal. We impose such constraint in part because pairwise trading is typically considered more empirically relevant,[3] in part to reduce the complexity of the trading

[2]In practice, vote trades often take the form of agreements and promises that are enforced by norms or reputational considerations. Even ignoring such factors, the problem remains complex.

[3]Riker and Brams (1973) for example, argue that the difficulty of organizing a coalition makes non-

mechanism used in the laboratory. An important advantage of this simplified setting is that the theory predicts convergence to a stable vote allocation for *any* sequence of blocking trades.

The experimental design employs three treatments, corresponding to three different preference profiles. All treatments have five member committees, and either two or three proposals. In each case, the stable outcome reachable through the theoretical trading dynamics is unique.

We evaluate the experimental results by distinguishing between hypotheses that concern the final state of the system and those concerning more detailed aspects of the dynamic process of trade. The first set of hypotheses find solid support in the data. First, the data show that stability is a useful predictive tool. In all treatments, two-thirds or more of the final vote allocations after trading are stable. Second, individual vote allocations qualitatively track the theoretical model closely. Across all treatments, across all voters, across all proposals, in every case in which the stable allocation is predicted to reflect a net purchase of votes, or a net sale, we observe it in the data. Third, when looking at the final outcomes reached in the lab, we find that in all treatments the outcome predicted by the theory is either the most frequently observed or the second most frequently observed.

The theoretical hypotheses that relate to the finer details of trading dynamics are less well supported by the data. The main discrepancy, and this is our fourth result, is that while we observe many payoff-improving trades, as posited by the theory, we also observe many trades that shift vote allocations without immediately affecting outcomes. We present the results of a statistical classification model that estimates the relative frequency of different kinds of trades. In addition to payoff-improving trades, the model includes trades that increase the number of votes held on high-value issues, as well as random trades. We find that both of these types of trades are frequent and significant.

Shifting votes towards higher-value proposals suggests some form of prudential behavior. The dynamic process posited by the theory is instead myopic: trades are considered profitable if the resulting vote allocation strictly improves the payoff of the traders, relative to the current vote allocation. Borrowing concepts of farsightedness from the theoretical literature (Chwe, 1994, Dutta and Vohra, 2015, Ray and Vohra, 2017), we extend the model to farsighted vote trading. The definition of farsightedness leads directly to some simple predictions, but in our experiment farsighted behavior of this kind is soundly

pairwise trading unlikely. The restriction to pairwise trades is also consistent with Guerrero and Matter's (2016) empirical strategy.

rejected.

The use of laboratory methods is particularly appropriate for the study of vote trading, given both the difficulty of collecting historical data and the ability of experiments to control the basic features of the environment, such as preferences and the set of issues being voted on, which allows a direct analysis of the theoretical hypotheses. And yet, if empirical and theoretical studies of vote trading are not numerous, experimental studies are even fewer. The study closest to ours is McKelvey and Ordeshook (1980), but the stark differences in procedures and objectives (a comparison of alternative cooperative solution concepts in McKelvey and Ordeshook in contrast to this paper's study of vote trading dynamics) make a direct comparison impossible.[4]

Methodologically related to our trading protocols are some recent experiments on decentralized matching, in particular Echenique and Yariv (2012).[5] In those experiments, as in ours, a central question is the extent to which the experimental subjects succeed in reaching stable outcomes. The details of those environments, however, differ substantially from ours, and the substantive questions we ask are specific to vote trading. There is a more distant relationship between the present paper and experimental studies of network formation. In network models, an outcome is a collection of bilateral links between agents, represented by either a directed or undirected graph, and the structure of payoffs is very different from vote trading games. Some classic theoretical analyses of network formation, however, exploit a pairwise stability concept, as we do (Jackson and Wolinsky 1996). Most experimental papers rely on a different protocol–a simultaneous move game where agents form links unilaterally–but some recent papers are closer to our approach: Carrillo and Gaduh (2016) and Kirchsteiger et al. (2016) examine dynamic sequential link formation with mutual consent[6]; Berninghaus et al. (2006) and Choi et al. (2019, 2020) examine asynchronous unilateral link formation in continuous time.

Finally, as relates to more standard market experiments, a novel feature of our trading environment is the absence of divisible side payments (prices) denominated in a commonly-

[4]Fischbacher and Schudy (2014) conduct a voting experiment to examine the possible behavioral role of reciprocity when a sequence of proposals come up for vote. There is no explicit vote trading, but voters can voluntarily vote against their short term interest on an early proposal in hopes that such favors will be reciprocated by other voters in later votes.

[5]Other related works on matching are Nalbantian and Schotter (1995), Niederle and Roth (2009) and Pais et al. (2011). These papers have incomplete information and study the effects of different offer protocols and other frictions. Kagel and Roth (2000) study forces leading to the unraveling of decentralized matching.

[6]Both papers use the random link arrival protocol of Jackson and Watts (2002): in each period one link is randomly added to the network, and the two newly connected players simultaneously decide to accept or reject the link.

valued currency. That is, these are *barter markets*. To our knowledge, experimental studies of barter markets are rare. Ledyard, Porter and Rangel (1994) is an example that demonstrates the challenges to both design and data analysis.

The paper proceeds as follows. The next section briefly summarizes the theoretical model and results on which our experiment is based; section 3 describes the experimental design; section 4 reports the experimental results, and section 5 concludes.

2 The Model

2.1 The Voting Environment

The environments studied in the laboratory are simplified versions of the trading environments analyzed theoretically in Casella and Palfrey (2019) (CP). A committee of N (odd) voters must approve or reject each of K independent binary proposals, a set denoted by $\mathbf{P}$. Committee members have additively separable preferences represented by a profile of values, Z, where z_i^k is the value attached by member i to the approval of proposal k, or the utility i experiences if k passes. The utility from a proposal failing is normalized to 0, and value z_i^k is positive if i is in favor of k and negative if i is opposed. Proposals are voted upon one-by-one, and each proposal k is decided through simple majority voting.

Before voting takes place, committee members can trade votes. Vote trades can be reversed if the parties to the trade decide to do so, but the agreements suffer no credibility or enforcement problems: one may think of votes as physical ballots, each one tagged by proposal, and of a trade as an exchange of ballots.

After trading, a voter may own zero votes on some proposals and several votes on others, but cannot hold negative votes on any proposal. We denote by v_i^k the votes held by voter i over proposal k, and by $v_i = (v_i^1, ..., v_i^K)$ the set of votes held by i over all proposals. We call $v = (v_1, ..., v_N)$ a *vote allocation*, i.e. the profile of vote holdings over all voters and proposals. The initial vote allocation v_0 equals $(1, 1, , , 1)$, where each 1 denotes a $1 \times K$ unit vector. The set $\mathcal{V}$ contains all feasible vote allocations: $v \in \mathcal{V} \Longleftrightarrow \sum_i v_i^k = N$ for all k and $v_i^k \geq 0$ for all i, k.

CP allows for general trades–trades among coalitions of voters of arbitrary size, where each voter may exchange as many votes as desired over one or more proposals for a possibly different number of votes on other proposals. Here we specialize the model to the design of the laboratory experiment. We restrict all trades to be *elementary* trades. i.e. trades that concern two voters only, with one vote on one proposal exchanged for one vote on another proposal. More precisely:

Definition 1. *An **elementary trade** between voters i and i' is an ordered pair of vote allocations (v, v') such that $v, v' \in \mathcal{V}$ and there exists a pair of proposals k, l such that: (i) $(v_i^k - v_i'^k) = 1$, $(v_{i'}^k - v_{i'}'^k) = -1$, $(v_i^l - v_i'^l) = -1$, $(v_{i'}^l - v_{i'}'^l) = 1$; (ii) $v_j^m = v_j'^m$ for all $m \neq k, l$ and for all j; and (iii) $v_j' = v_j$ for all $j \neq i, i'$.*

Given a vote allocation v, when voting occurs, each voter's dominant strategy is to cast all votes in his possession over each proposal in the direction the voter sincerely prefers–in favor of proposal k if $z_i^k > 0$, and against k if $z_i^k < 0$. We call $P(v) \in \mathbf{P}$ the *outcome* of the vote if voting occurs at allocation v: the set of proposals that receive at least $(N+1)/2$ favorable votes, and therefore pass. The utility of voter i if voting occurs at v is denoted by $u_i(v)$: $u_i(v) = \sum_{k \in P(v)} z_i^k$. Preferences over outcomes are assumed to be strict, that is, $u_i(v) = u_i(v')$ if and only if $P(v) = P(v')$.

Note that with K independent binary proposals, there are 2^K possible outcomes (all possible combinations of passing and failing for each proposal). Although it is convenient to represent preferences in terms of cardinals values Z, our model relies exclusively on the voters' ordinal rankings over the 2^K possible outcomes. All results are unaffected by changes in Z that do not affect individual ordinal rankings.

The focus is on stable vote allocations that hold no incentives for further trading. Define:

Definition 2. *An elementary trade, (v, v'), between voters i and i' is **payoff improving** if $u_i(v') > u_i(v)$ and $u_{i'}(v') > u_{i'}(v)$.*

Definition 3. *An allocation $v \in \mathcal{V}$ is **stable** if, for every pair of voters i and i', there exists no payoff improving elementary trade.*

Note that a stable vote allocation always exists: any feasible allocation of votes where a single voter i holds more than half the votes on every proposal is trivially stable: no exchange of votes involving i can make i strictly better-off; and no exchange of votes that does not involve i can change the outcome. Hence, the interesting question is not whether a stable allocation exists, but whether and under what conditions sequential decentralized trading from the initial vote allocation leads to stable vote allocations.

2.2 Trading Dynamics

To answer the question, the theory needs to specify a dynamic process through which trades take place. Although the literature does not make explicit reference to an algorithm, the sequential myopic trades envisioned by Riker and Brams (1973) and Ferejohn

(1974) lend themselves naturally to such a formalization. *Pivot Algorithms* are defined as sequences of trades yielding myopic strict gains to both traders. When we specialize trades to elementary trades, we can define:

Definition 4. *An* ***Elementary Pivot Algorithm*** *is any mechanism generating a sequence of trades as follows: Start from the initial vote allocation v_0. If there is no payoff improving elementary trade, stop. If there is one such trade, execute it. If there are multiple such trades, choose one according to some tie-breaking rule R. Continue in this fashion until no payoff improving elementary trades exist.*

Rule R specifies how the algorithm selects among multiple possible trades; for example, R may select each potential trade with equal probability; or give priority to trades with higher total gains; or to trades involving specific voters. The definition describes a family of Pivot algorithms, spanning all possible R rules.

Trades are required to be strictly payoff improving for both traders. That means that trades concern pivotal votes: trades of non-pivotal votes cannot affect outcomes and thus cannot induce changes in utility.[7] More than that: since trades are restricted to be elementary, *only* pivotal votes can be traded. It is this property, anticipated by Riker and Brams, that gives the name to the algorithms.

2.3 Pivot-Stable Allocations

An obvious question is whether trading under Pivot algorithms ever stops; in principle there is nothing to rule out trading cycles. If trading does stop, we call the resulting vote allocation a Pivot-stable vote allocation.

Definition 5. *An allocation of votes v is* ***Pivot-stable*** *if it is stable and reachable from v_0 through an elementary Pivot algorithm in a finite number of steps.*

CP's main result in a more general setting is an existence theorem, proving that a finite path of vote trades ending at a stable vote allocation always exists, without requiring restrictions on trades. When trading is limited to elementary trades, the result is stronger:

Theorem. *For any K, N, and z, all elementary Pivot algorithms converge to a stable vote allocation in a finite number of trades.*

[7] Note that strict preferences imply that it is impossible for a trade to cause a strict payoff gain for one side and not change payoffs for the other. Changing payoffs means that pivotal votes have been traded, and thus both sides of the trade must be affected.

The modifier "all" refers to the generality of the result in terms of the choice rule R: convergence is guaranteed for any R. Thus elementary Pivot algorithms *always* reach a stable vote allocation, regardless of the order in which different possible trades are chosen, for any number of voters and proposals, and for all (separable) preferences.

Proof. Consider voter i and a vote allocation v. Let $x_i^k = |z_i^k|$ be voter i's *intensity* for proposal k, and define as voter i's *score* at v the function:[8]

$$\sigma_i(v) = \sum_{k=1}^{K} x_i^k v_i^k.$$

Suppose a payoff improving elementary trade occurs, resulting in a new vote allocation, v'. If i was not a party to the trade, then $\sigma_i(v') = \sigma_i(v)$, by construction. If i was a party to the trade trade, by definition of elementary Pivot algorithm, i must trade away one vote on a proposal k^- that i wins pre-trade and loses post-trade, and aquire one vote on a proposal k^+ that i loses pre-trade and wins post-trade, or $(v_i'^{k^+} - v_i^{k+}) = 1 = -(v_i'^{k^-} - v_i^{k-})$. For the trade to be payoff improving, it must be that i values winning k^+ more than winning k^-, or $x_i^{k^+} > x_i^{k^-}$. Thus:

$$\sigma_i(v') - \sigma_i(v) = x_i^{k^+} - x_i^{k^-} > 0$$

The score of voter i has increased. Hence if i trades, $\sigma_i(v') > \sigma_i(v)$. At any future step, either there is no trade and a stable allocation has been reached, or there is trade, and thus there are two voters i and i' whose score increases. But score functions are bounded and the number of voters is finite. Hence trading must end after a finite number of steps. Note that the argument makes no restriction on R, and thus the result holds for all R. $\square$

Vote trading environments are unusually complex: the implicit value of a vote depends on its pivotality, and thus changes with other voters' allocations; any trade affects the possibility of further Pivot trades and can generate a whole chain of new exchanges; others' trades change outcomes, and generate externalities on all voters. Elementary Pivot algorithms are simple, intuitive rules, describing plausible trades in such a complicated

[8] A voter's score is the intensity-weighted sum of his votes, and can be loosely interpreted as a measure of the potential worth of all the votes currently in the possession of that voter. The score changes as a voter's vote holdings change during trade; for example, it goes up if the voter holds one additional vote on a proposal with a higher value and one less vote on a proposal with a lower value. It will play a role later in our analysis of vote trading in the experiment.

environment. Their simplicity allows some conceptual progress, as in the stability result we just described. But we focus on them for a second reason too: we conjecture that they may have predictive power.

We next turn to a description of the experiment. In what follows, it should be understood that the term "elementary" always applies to trades and to Pivot algorithms, even if not stated explicitly.

3 The Experiment

The experiment was conducted at the Columbia Experimental Laboratory for the Social Sciences (CELSS), with registered Columbia students recruited from the whole campus through the laboratory's ORSEE[9] site. No subject participated in more than one session. After entering the computer laboratory, the students were seated randomly in booths separated by partitions; the experimenter then read aloud the instructions, projected views of the computer screens to be seen during the experiment, and answered all questions publicly. Because the design of the trading platform presents some challenges, we describe it here in some detail. Sample instructions and screenshots are reproduced in the online appendix.[10]

Each subject's computer screen displayed a table with all subjects' values per proposal (in experimental points), and vote holdings. We refer to this matrix as the *vote table.* The vote table conveys to all voters complete information about voters' preferences and the current vote allocation. The interface and the instructions associated the two alternatives for each proposal, Pass or Fail, with two colors, Orange (Pass) and Blue (Fail). Thus a subject's value for a proposal indicated earnings from the subject's preferred alternative winning, relative to zero earnings if it lost. All individual's values were positive and displayed in the color of the individual's preferred alternative.[11] The vote table also showed the vote totals on each issue and the points the subject would win if voting were held immediately. Each subject started with one vote on each proposal.

After observing the vote table, any subject could post a bid: a request for a vote on one of the issues, in exchange for the offer of his vote on a different issue. The bid

[9]Greiner (2015).

[10]The computerized trading platform was implemented using the Multistage software program, an open source software developed at Caltech's Social Science Experimental Laboratory (SSEL) by Chris Crabbe. The software is available for public download at http://multistage.ssel.caltech.edu:8000/multistage/.

[11]Thus, for example, $z_i^1 = -300$ in the notation of the model would appear on the screen as voter i having a value of 300 for proposal 1 highlighted in Blue. A sample subject computer screen shows this in the online appendix.

appeared on all committee members' monitors, together with the ID of the subject who had posted it. A different subject could then accept the bid by clicking on it, or post another bid.

A central feature of vote trading is that the preferences and vote holdings of the specific individuals making a trade determine the effect of the trade. Contrary to standard market experiments, then, subjects must not only post potentially profitable bids, but also consider the identity of their trading partner. In adapting the bidding platform used in market experiments, we added a confirmation step. After a bid was accepted, a window appeared on the bidder's screen detailing the effects of that specific trade–what the outcome would be upon immediate voting–and asking the bidder to confirm or reject the trade. If the trade was rejected, a message appeared on the screen of the rejected trade partner, informing him of the rejection; trading then continued as if the bid had never been accepted (thus the bid remained posted and available for others to accept). If the bidder confirmed the trade, a popup window with the updated vote table appeared on all screens for 10 seconds and trading activity was paused during that 10 second interval, to give traders time to study the new vote allocation that resulted from the trade. The window also reported the post-trade voting outcome that would result if voting were to occur immediately. The vote table that was always visible on the main screen was also updated immediately.

The vote-trading market was open for three minutes.[12] However, in a market where each concluded trade can trigger a new chain of desired trades, it is important to allow adequate time for any subsequent desired trades to be executed. For this reason the time limit was automatically extended by 10 seconds whenever a new trade was concluded.

No bid could be posted if a subject did not have the vote to execute it if accepted; thus a voter could post multiple bids only as long as he possessed the votes to execute them all, had all been accepted. Posted bids could be canceled at any time, an important feature in a market where somebody else's executed trade can make an existing posted bid suddenly unprofitable.

Once the market closed, voting took place automatically, with all votes on each issue cast by the computer in the direction preferred by each subject.[13] Then a new round started, with the initial vote allocation of one vote per voter per proposal.

[12]The market was open for only two minutes in the two-proposal treatment, AB, because the extent of possible trading was more limited.

[13]With binary alternatives for each proposal, voting in the preferred direction is a dominant strategy. We chose to implement the casting of votes automatically to simplify an already complex experiment, save time, and focus all attention on the trades.

The experiment consisted of three treatments, AB, $ABC1$, and $ABC2$. In all three treatments, the size of the voting committee was five ($N = 5$), while the number of proposals depended on the treatment: $K = 2$ in treatment AB, and $K = 3$ in treatments $ABC1$, and $ABC2$. Treatments $ABC1$ and $ABC2$ had different preference profiles. In each committee, subjects were identified by ID's randomly assigned by the computer, and proposals were denoted by A and B (in treatment AB), and A, B, and C (in treatments $ABC1$ and $ABC2$). Each session started with two practice AB rounds; then three rounds of treatment AB, and then five rounds each of $ABC1$ and $ABC2$, alternating the order.[14] We did not alternate the order of treatment AB because its smaller size ($K = 2$) made trading substantially less complicated for the subjects, and thus we implemented it before the more complex treatments. This is also the reason for the fewer number of rounds (three for AB, versus five for $ABC1$ and $ABC2$).

Committees were randomly formed, and ID's randomly assigned at the start of each new treatment, but the composition of each committee and subject ID's were fixed for all rounds of the same treatment. All sessions except for one consisted of 15 subjects, divided into three committees of five subjects each.[15] At the end of each session, subjects were paid their cumulative earnings from all rounds of all treatments, converting experimental points into dollars via a pre-announced exchange rate, plus a fixed show-up fee. Each session lasted about 90 minutes, and average earnings were approximately \$36, including the \$10 show-up fee.

We designed the three treatments according to the following criteria. First, we wanted a $K = 2$ treatment, as a simpler initial task for the subjects as they gained experience with the trading protocol. Second, we chose profiles of values for which the stable vote allocation reachable via Pivot trades is unique but requires multiple trades. In AB, the path to stability is itself unique, while in both $ABC1$ and $ABC2$ the unique stable vote allocation can be reached via multiple trading paths.

Third, we designed preference profiles for which the Condorcet winner exists but need not be the Pivot stable outcome: it is Pivot stable in AB and in $ABC2$, but not in $ABC1$. The two matrices $ABC1$ and $ABC2$ are otherwise similar and in particular have Pivot trading paths of equal multiplicity and length. Note that we do not specify R, the selection rule when multiple trades are possible, but let the experimental subjects select which trades to conclude. For each of the experimental matrices, the Pivot-stable allocation is unique and invariant to R.

[14]Two of the sessions had only two treatments because a programming error made the last five rounds of data unusable: AB and $ABC1$ in one case,and AB and $ABC2$ in the other.

[15]One session had only ten subjects, divided into two committees.

The three preference profiles used in the experiment are given in Table 1 (with the Pass/Fail notation of the theoretical model). In each matrix, rows correspond to proposals, and columns to voters.

AB

	1	2	3	4	5
A	49	−29	−29	12	−12
B	12	−12	−49	29	49

$ABC1$

	1	2	3	4	5
A	23	−23	10	−10	23
B	−10	−10	23	−23	10
C	18	−18	−18	18	−18

$ABC2$

	1	2	3	4	5
A	−21	15	−9	21	9
B	15	9	15	−15	−15
C	−9	−21	21	9	21

Table 1: Preference profiles used in the experiment.

To illustrate the dynamics of the Pivot algorithm, consider the sequence of Pivot trades with value matrix AB. At v_0, A fails and B passes, or $P(v_0) = B$. The outcome is the Condorcet winner. Voters 2, 4 and 5 are all on the winning side of the proposal each of them values most, and have no payoff improving trade. But voters 1 and 3 can gain from a trade reversing the decision on both A and B: voter 1 gives a B vote to voter 3, in exchange for 3's A vote. With no further trade, the outcome would be $P(v_1) = A$, which both 1 and 3 prefer to $P(v_0) = B$. At v_1, however, 2 and 4 have a payoff improving trade: 2 gives a B vote to 4, in exchange for an A vote. At v_2, $P(v_2) = B$. Vote allocation v_2 is stable: all pivotal votes are held by voters 2, 4 and 5, none of whom can gain from trading. No other trading sequence is consistent with a Pivot algorithm; thus trading follows a unique path, of length two (i.e. consists of a sequence of two trades). Indicating first the ID's of the trading partners, and then, in lower-case letters, the proposal on which an extra vote is acquired by the voter listed first, the path is $\{13ab, 42ba\}$. The unique Pivot-stable outcome is $P(v_2) = B$, which is also the Condorcet winner, and thus the two coincide in the case of matrix AB.

With matrix $ABC1$, the Condorcet winner exists and corresponds to $P(v_0) = A$, but the unique Pivot-stable outcome is ABC–all proposals passing. The Pivot algorithm can follow three alternative paths, two of them of length four, and one of length three: $\{13cb, 45bc, 23ab, 45ca\}$, $\{23ab, 45ca, 45bc, 13cb\}$, and $\{23ab, 45ba, 13cb\}$. In matrix $ABC2$, the Condorcet winner again exists. It is ABC, which is also the unique Pivot stable outcome. Again, the Pivot algorithm can follow three alternative paths, two of them of length four, and one of length three. They are: $\{15ab, 34ba, 24cb, 15bc\}$, $\{24cb, 15bc, 15ab, 34ba\}$, and $\{24cb, 15ac, 34ba\}$. Although $ABC1$ and $ABC2$ admit multiple possible trading paths, for all three matrices the Pivot stable vote allocation is unique.

Table 2 reports the experimental design.

Session	Treatments	# Subjects	# Groups	# Rounds
s1	$AB, ABC1, ABC2$	10	2	3,5,5
s2	$AB, ABC2, ABC1$	15	3	3,5,5
s3	$AB, ABC1, ABC2$	15	3	3,5,5
s4	$AB, ABC2, ABC1$	15	3	3,5,5
s5	$AB, ABC2$	15	3	3,5
s6	$AB, ABC1$	15	3	3,5

Table 2: Experimental Design

3.1 Hypotheses

The theoretical model yields two distinct sets of predictions: predictions on the final state of the system once all trading is concluded, and predictions on the trade-by-trade dynamics. We analyze the experimental data by confronting them to the two sets in turn.

Three theoretical hypotheses concern the state of the system when trade has ended: the stability of the final vote allocation; the precise final vote allocations across the five voters on all the proposals; and the final outcome in terms of which proposals pass and which proposals fail. The first two of these hypotheses specifically address *vote allocations*, while the last one addresses *outcomes*. The first hypothesis about vote allocations is that final vote allocations will be stable, i.e., at the final vote allocation there exists no mutually payoff improving elementary trade (H1). The second hypothesis about vote allocations, i.e., the precise predictions of the Pivot stable model about final vote holdings for each voter in each treatment (H2), is presented in Table 3. The hypothesis about which proposals pass or fail, i.e., the Pivot stable outcome in each treatment (H3), is presented in Table 4. That table also shows the Condorcet winning outcome in each treatment for contrast.

Clearly the three hypotheses are linked, and vary in how strongly they restrict the data. H2 places the strongest restrictions, and directly implies H1 and H3: if trading leads to the vote allocations predicted by the model (Table 3), then (a) the outcomes must correspond to the Pivot stable outcomes, and (b) there can be no payoff improving

Voter ID	AB		ABC1			ABC2		
	A	B	A	B	C	A	B	C
1	2	0	1	0	2	2	1	0
2	2	0	2	0	1	1	0	2
3	0	2	0	3	0	0	2	1
4	0	2	0	2	1	2	1	0
5	1	1	2	0	1	0	1	2

Table 3: Theoretical Final Allocations

Theory	AB		ABC1			ABC1		
	A	B	A	B	C	A	B	C
Pivot Stable	Fail	Pass	Pass	Pass	Pass	Pass	Pass	Pass
Condorcet Winner	Fail	Pass	Pass	Fail	Fail	Pass	Pass	Pass

Table 4: Theoretical Final Outcomes

elementary trades. On the other hand, H1 is neither strictly stronger nor strictly weaker than H3. That is, trading may end at the Pivot stable outcome while the final vote allocation is not stable; and, on the flip side, the final vote allocation may be stable while the outcome differs from the Pivot stable outcome.

A distinct hypothesis concerns the exact details of trade-by-trade dynamics. The model predicts Pivot trades (H4), i.e., trades that lead to myopic payoff gains for both voters involved in the trade.

3.1.1 Random trading benchmark

Experimental data are inherently noisy, so a strict test of the model - which is deterministic - will always fail because a single violation is sufficient to falsify the model. In order to obtain a more informative measure of the extent to which the data in the experiment support or contradict the theoretical hypotheses we need as a benchmark comparison an alternative null model of the trading process. Ideally, the alternative should be well defined and self-consistent theoretically, as well as being formulated ex ante, before taking cues from the data itself. A natural null alternative and one that we use because it admits the greatest possible range of different types of trades is that sequences of trades are randomly generated, with all feasible trades at a point in time being equally likely to be executed. In what follows, we begin by comparing the theoretical model to the alternative of random trading to measure the extent to which the Pivot model provides a useful way of organizing the data. In the next section, when studying the detailed trades observed in the lab, we discuss and evaluate plausible ex-post rationalizations of the data.

If the simulation of random trades is the initial yardstick of comparison for our data, it is worth describing the simulation methodology in some detail. We began by calculating, for each treatment, the average length of a round in the experimental data. We then divided that time interval into a grid of 100 time-cells; the program enacts a random trade in each time-cell with probability p, such that $100p$ equals the mean number of trades per round per group in the treatment. Thus the trades arrive according to a Poisson process.

At any point in time during the randomly simulated trading round when the program is called to enact a trade, one random elementary bilateral trade is constructed as follows. First, the program selects randomly a pair of traders, a pair of proposals, and a direction of trade, all with equal probability. If both traders possess the required votes for the trade to be feasible, the trade is implemented. If the trade is infeasible, it is not implemented and the program draws a new possible random trade. The feasibility of that new randomly selected trade is checked, and so forth. Once a feasible random trade has been found and implemented, the clock continues, with further trades being enacted with probability p in each subsequent time-cell until the 100th time cell has been reached, at which point one simulated trading round has been generated. We simulated 5,000 random trading rounds for each treatment.

4 Experimental Results.

The experimental results are presented in two parts. The first part analyzes the properties of the final state of the system after all trading has concluded - final allocations and final outcomes - and addresses hypotheses H1, H2, and H3. The second part of the results section analyzes the trading dynamics and addresses hypothesis H4.

4.1 Part 1: Final Vote Allocations and Final Outcomes

4.1.1 Vote Allocations

Stability of final vote allocations Hypothesis H1 is that final vote allocations will be stable. Are the vote allocations at which the trading process ends such that no further payoff-improving trades are possible?

Table 5 reports, for each treatment and for the full data set, the fraction of final

	Data	Random Trading
	%	%
AB	76.5	68.0
ABC1	64.3	61.6
ABC2	64.3	56.8
All	67.5	62.1

Table 5: Percentage of final vote allocations that are stable.

vote allocations that are stable, and the corresponding fraction under the random simulations.[16] More than three quarters of all final allocations are stable in AB, just short of two thirds in $ABC1$ and $ABC2$, and just above two thirds over the full data set. The proportions of stable final allocations are high, providing solid support for H1. Stability is indeed a useful criterion for thinking about the end points of trading.

Separating the notion of stability as resting point of the system from the dynamic model of Pivot trading, the evidence in favor of the latter is not as strong. The fraction of stable allocations is higher than under simulated random trading in every single case, but the difference is only marginally significant for AB ($p = 0.10$) and for the whole data set ($p = 0.06$).[17] When applied to the stability of final vote allocations, random trading is a demanding comparison because a large fraction of feasible trades take the vote allocation on at least one proposal away from minimal majority–i.e. away from an allocation where the two opposite sides differ by exactly one vote. Once away from minimal majority, elementary payoff-improving trades are impossible, and the vote allocation over the non-minimal majority proposals is necessarily stable.[18]

While Table 5 shows that when trading stops, the final vote allocation is usually stable (H1), it is also informative to ask the converse question: If a stable allocation is reached during the trading process, does trading stop? Table 6 reports the frequency of stable allocations that are not followed by further trade in the experimental data and in simulated random trading. In each of the treatments, the frequency in the experimental data is between 40 and 50%. It is always higher than under random trading, and the difference is significant under a one-sided test for $ABC1$ and pooling across treatments

[16]Note that H1 should hold for all treatments, and there is no strong reason to test it separately on each.

[17]The p-value is calculated from a simple z-test of proportions.

[18]For example, in treatment AB, where breaking minimal majority on a single issue is sufficient to induce stability, a single random vote trade from any unstable allocation has never less than a 30 percent chance of inducing stability.

	Data	Random Trading
	No further trade (%)	No further trade (%)
AB	46.5	40.5
ABC1	43.3	33.7
ABC2	42.1	39.2
All	43.7	37.6

Table 6: Percentage of stable vote allocations followed by no further trade.

($p = 0.02$ in both cases).

Final Vote Holdings For each of the three preference profiles used in our experiment, Pivot algorithms lead to a unique stable vote allocation, summarized in Table 3. Are Pivot-stable vote allocations predictive of the final vote allocations observed in the experimental data?

Figure 1 reports, for each treatment, the number of votes held on each proposal by each voter type in the Pivot-stable vote allocation (dashed black), in the data (blue), and at the end of the random simulations (grey), averaging over all rounds and groups.[19] For clarity, proposal and voter IDs on the horizontal axis are ordered so that theoretical Pivot-stable vote holdings are everywhere weakly decreasing.[20]

Even if vote allocations within a group were generated randomly, they could not be independent, both because of the reciprocal nature of trades and because of the adding-up constraints, forcing total vote holdings across all voters to sum to 5 for each proposal, and total vote holdings across proposals to sum for to 2 in AB or 3, in $ABC1$ and $ABC2$, for each voter. Because of the lack of independence, the confidence intervals reported in the figure (the light blue bands) were obtained by bootstrapping. In the bootstrapping procedure, we populated each sample by drawing with replacement from the data the full vector of vote holdings for a group over all proposals, where a group is identified by its ID, session, and treatment. The confidence intervals correspond to 95% of realizations over 5,000 samples, centered on the experimental data.

The vote distribution in the data is less sharply variable than theory predicts, as expected in the presence of noise, but the qualitative predictions are strongly supported. There are five voters in each treatment, holding votes over two (in AB) or three issues

[19]Thus, for example, the number of votes at A1 corresponds to the number of votes on proposal A held by the subject in role 1 at the end of the round, averaging over all groups, rounds, and sessions of the relevant treatment.

[20]At equal Pivot-stable vote holdings, voters are ordered according to their ID, from 1 to 5.

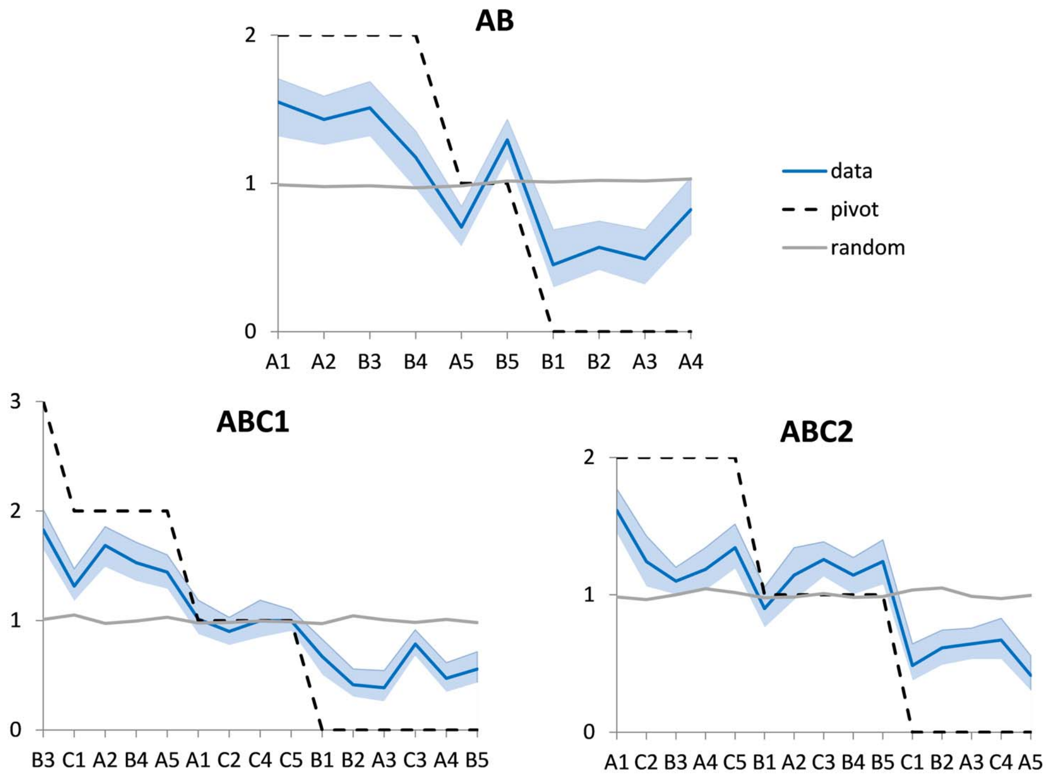

Figure 1: Final vote allocations.

(in $ABC1$ and $ABC2$)–a total of forty points. Of these forty, the theory predicts that 14 should be above 1–the voter should be a net buyer over that issue– and 15 below 1–the voter should be a net seller. The prediction is satisfied in every single case, across all treatments. Note that the pattern cannot have been generated by random trading, the almost fully horizontal line that corresponds to the 5,000 random simulations. As expected, random trading generates individual vote allocations that on average are indistinguishable from the initial vote allocations, or 1 vote to each voter over each proposal.

Returning to the fit between the experimental data and the theoretical model, more informative than the individual points is the pattern across vote allocations. We can capture such a pattern through simple correlation coefficients between the data and the Pivot predictions, as reported, together with bootstrapped confidence intervals, in Table 7.[21]

[21]Note that the correlation coefficients are invariant to the adding-up constraints that tie down each subject's vote holdings across proposals, as well as each group's aggregate vote holdings on each of the proposals–they are unchanged if we drop one proposal and one voter.

Averaging over all groups, per voter ID					
AB		$ABC1$		$ABC2$	
0.90	[0.77, 1.0]	0.96	[0.92, 0.99]	0.87	[0.73, 0.96]
Without averaging over all groups per voter ID					
AB		$ABC1$		$ABC2$	
0.64	[0.42, 0.77]	0.76	[0.50, 0.84]	0.59	[0.45, 0.69]

Table 7: Correlation coefficients between data and Pivot predictions with 95% confidence intervals.

The upper panel in the table corresponds to the figure, and the correlation coefficients, if squared, would correspond to the R^2 of a linear regression of final vote allocations in the experiment on the Pivot predictions. Averaging over all groups and rounds for given treatment, as done in this case, reduces noise, but a more disaggregated analysis delivers a similar message. The lower panel in Table 7 again reports the correlation coefficients between the data and the Pivot predictions, but now without averaging across groups. The presence of inter-group variation reduces the correlation, but the coefficients remain large.

Final Outcomes Figure 2 compares the frequency distribution of different outcomes observed in the experimental data (left bars, in blue) with the simulations with random trading (right bars, in grey). Outcomes are ordered according to the frequency with which they are observed in the data. A star indicates the Condorcet winner, and the Pivot stable outcome is circled.

The figure reveals two regularities. First, in all three treatments, the Condorcet winner is the most frequently observed outcome. Second, the Pivot stable outcome is either the most frequently observed outcome (when it coincides with the Condorcet winner) or the second-most frequently observed outcome (when it differs from the Condorcet winner, in $ABC1$). Random trading shares the first feature with the experimental data, but the frequency of Condorcet outcomes is much higher and the Pivot stable outcome is not the second (or even the third) most frequent outcome in $ABC1$. For each of the three treatments, chi-square tests reject the hypothesis that the frequencies of outcomes are equal in the data and under random trading ($p = 0.010$ for AB, $p = 0.031$ for $ABC1$, and $p = 0.046$ for $ABC2$).

We conjecture that the reason why the Condorcet winner is so often observed in the

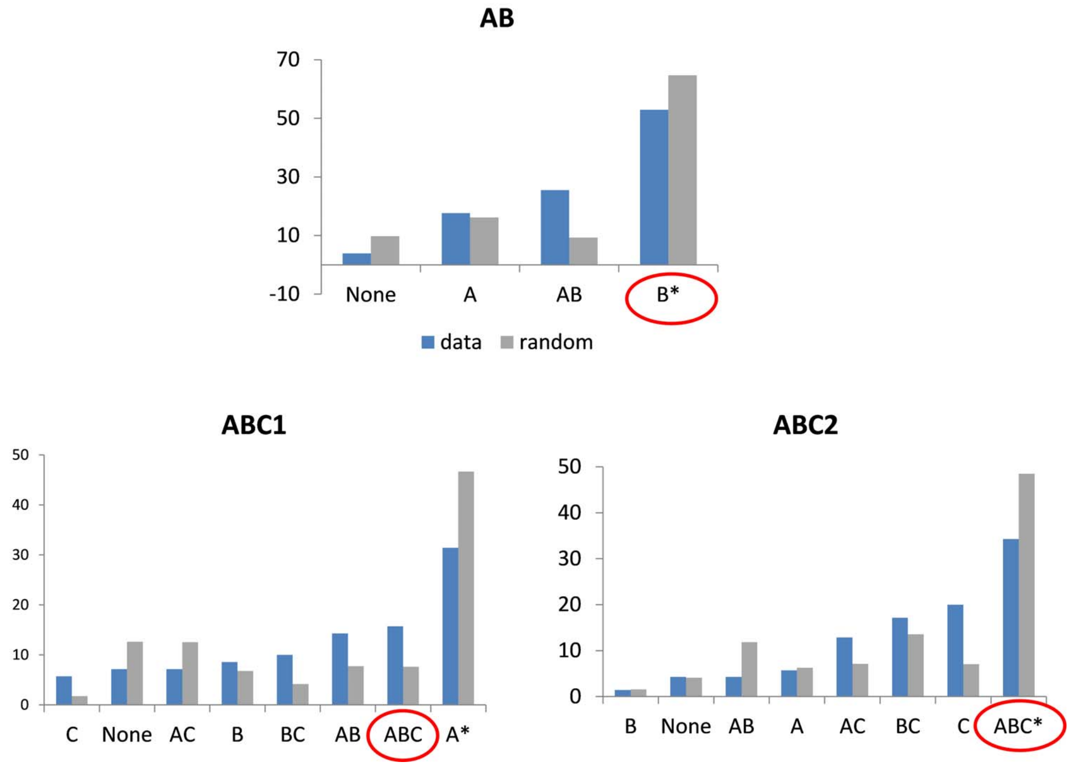

Figure 2: Frequency of outcomes. Data (left) and simulated random trading (right).

data is related to the reason Condorcet outcomes have such high frequency under random trading. It is well-known (Park, 1967; Kadane, 1972) that when the Condorcet winner exists, as in our three treatments, it must coincide with the no-trade outcome. Thus the frequency with which the Condorcet winner is reached goes hand in hand with the persistence of pre-trade outcomes. Both reflect the inertia resulting from noisy trades, and more precisely by trades that do not change outcomes but result in non-minimal majority vote allocations. Such allocations are stable, and moves away from the status quo are in general more difficult.[22]

[22]We can compute the frequency of the different outcomes focusing only on those rounds in which the voting outcome, if voting were held, changes during trade. The relative frequency of the Condorcet winner falls in all three treatments, supporting the conjecture.

4.2 Part 2: Trading behavior and dynamics

4.2.1 The Dynamics of the Trading Process

If the final state of the system–vote allocations and outcomes–is well captured by a snapshot at the system's resting point, the sequence of trades is a dynamic process tracing the system's evolution. Before analyzing the trade data in detail, we display the trade-by-trade dynamic process in Figure 3.

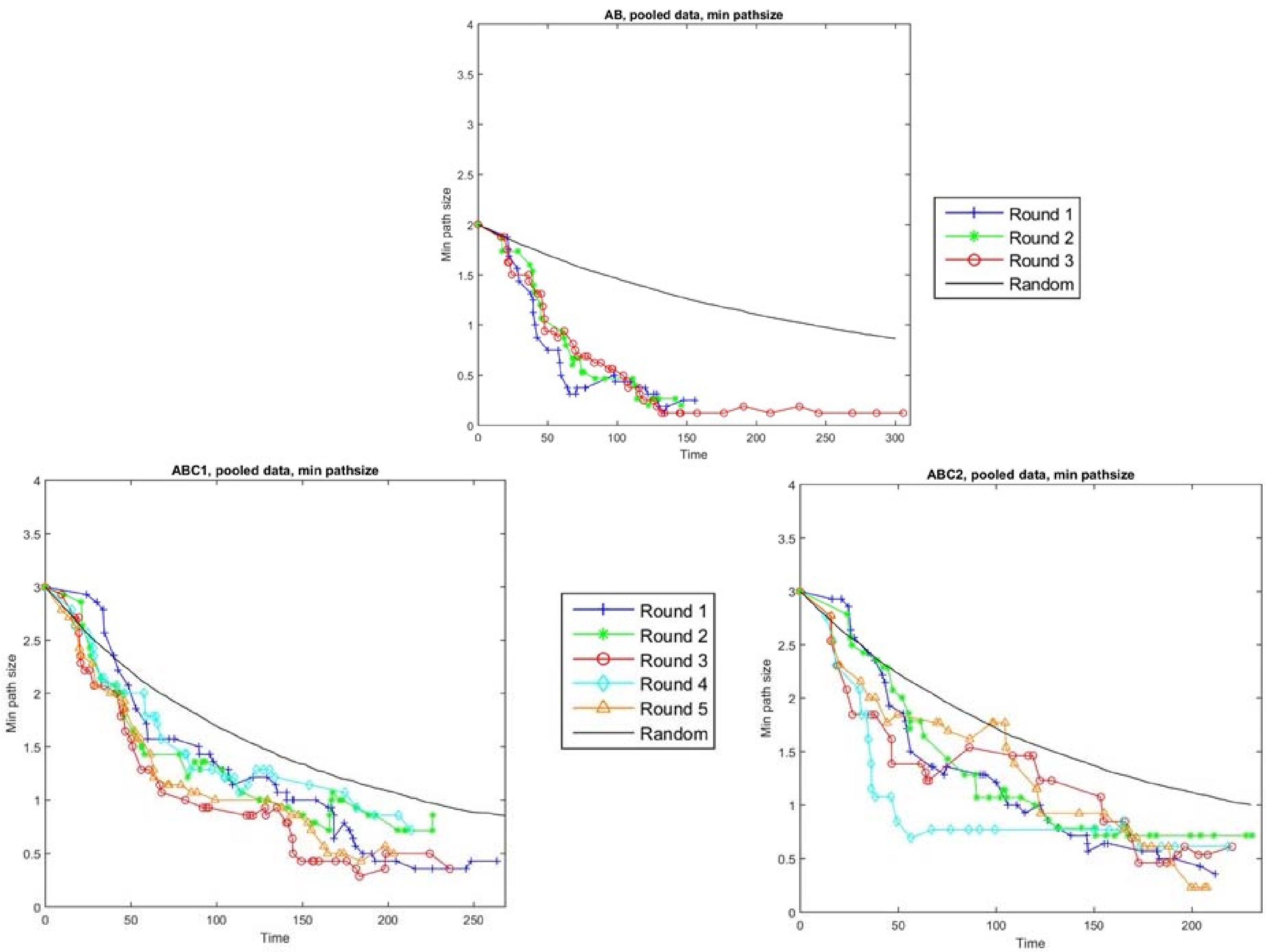

Figure 3: Dynamic convergence to Pivot stable outcomes.

The figure shows, for each round of each treatment, the dynamic path of the vote allocations, as it is modified by the succession of trades. The horizontal axis measures time, in seconds. For each round and treatment, a new marker represents a single trade in one of the sessions at that point in time in that round/treatment. The vertical height of each marker in the graph measures the distance from stability of the vote allocations after the new trade, averaged across all groups for that round and treatment. In line with our notion of stability, the distance measure is the minimum number of payoff-improving trades leading to a stable vote allocation.[23] Thus, a new marker lower than the previous marker indicates a trade that moves the vote allocation closer to stability, and a new marker higher than the previous marker indicates a trade that moved the vote allocation further from stability. Because the vertical axis measures distance from stability, the AB curves begin at 2, since two payoff-improving trades are required from the initial vote allocation in order to reach a stable vote allocation. For the same reason, the $ABC1$ and $ABC2$ curves begin at 3.

The curves representing the trading sequences decline in all rounds of all treatments, almost perfectly monotonically, clearly showing the steady dynamic convergence towards stability: almost every trade moves the system closer (at least weakly) to stability. To evaluate such convergence against a benchmark, the black curve in each panel reports the distance from stability over time with random trading, averaged over 5000 simulated trading periods. With the exception of two trades in round 5 in $ABC2$, after the first 40 seconds of the trading period, the random trading curve is always higher than the curve corresponding to the experimental data, in all rounds and in all three treatments, indicating that the actual trading in the experimental vote markets consistently converged to stability faster than random trading. In the AB treatment, actual trading converges to an averaged distance from stability that is virtually 0 in all rounds, while random trading only reduces the distance to stability by half, from 2 to 1.[24]

Figure 3 is consistent with subjects' intentional search for gains from trade. Is this hypothesis robust to more detailed analysis of the specific trades we see in the lab? In the remainder of the paper, we address hypothesis H4: are experimental trades well-described by the Pivot algorithms? We begin with a broad summary of the properties of the experimental bids and trades. We then test whether the observed trades are compatible with Pivot trading, explore possible rationales for the deviations we observe, and construct a statistical model estimating the relative frequency of different types of

[23]Since the measure of distance is averaged over all groups in all sessions for that round/treatment, the vertical shift resulting from a single trade in a single group is small and always less than one, reflecting the unchanged allocations of the other groups.

[24]It is noteworthy that, except for first 40 seconds of the trading period, convergence to stability is faster than random trading even in the very first round, before subjects have experience with the market. Notice the lack of learning in the data–the figure shows little if any systematic difference between earlier and later rounds.

Treatment	Tot trades	groups $\times$ rounds	Mean trades	Median	s.d	Max	Model
AB	115	51	2.25	2	1.92	13	2
$ABC1$	211	70	3.0	3	1.67	9	3,3,4
$ABC2$	175	70	2.5	2	1.36	7	3,3,4

Table 8: Number of trades.

trades. Finally, we investigate the possibility of farsighted trading behavior.

4.2.2 Summary of bids and trades

Table 8 reports basic summary statistics on the number of observed trades. The last column, labeled "Model" refers to the predicted number of trades under the Pivot algorithm. The unit of analysis is the group per round.

A histogram of the number of trades per treatment (per group per round) (Figure 4) shows the higher frequency of shorter trade paths in the AB treatment. Between the two $K = 3$ treatments, $ABC2$ has higher fractions of shorter paths, but the difference is not large–56% of rounds end with two or fewer trades in $ABC2$, as opposed to 41% in $ABC1$, and 80% end with three or fewer trades in $ABC2$, as opposed to 76% in $ABC1$. In all treatments, few rounds include five or more trades.

Table 9 summarizes bidding and trading behavior. Recall that bidding for votes is more complex than bidding for goods in a typical double auction market experiment. First, once a trade has occurred, previously posted bids may have become unprofitable, and subjects can cancel them. In our data, canceled bids reflect both mistaken bids, and bids that were correctly posted but were later rendered counterproductive.[25] Second, the preferences of the subject accepting a bid determine whether or not the trade is profitable for the bidder. Hence once a posted bid is accepted, the bidder is asked to confirm or reject the trade.

The first two rows of Table 9 report the total number of bids, and how many of these bids were canceled by the bidder. Row 3 shows the number of bids that were accepted. i.e that found a taker. Rows 4-6 break down these total accepted bids into three categories, depending on whether the voter accepting the bid would have a payoff gain, a payoff loss, or no change in payoff if the acceptance were confirmed by the bidder. Row 7 displays

[25]Some of the bid cancelations were required by the trading rules which did not allow a bidder to be exposed to a possible negative vote holding. Thus, for example, if voter 1 posted a bid that offered an A vote in exchange for a C vote, and voter 2 was offering a B vote in exchange for an A vote, then voter 1 would have to first cancel his bid in order to accept 2's bid, if he had a single A vote.

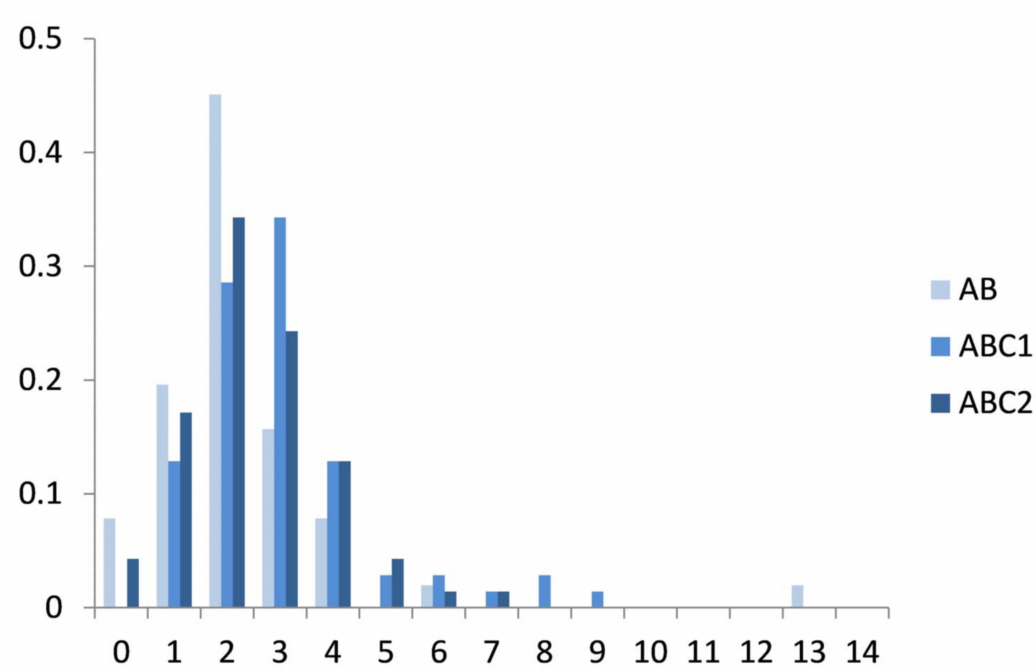

Figure 4: Number of trades. Frequencies by group per round.

the number of accepted bids that were confirmed by the bidder, with rows 8-10 breaking down this total into three categories depending on whether confirmation of the trade would give the bidder a payoff gain, a payoff loss, or no change. Row 11 displays the number of accepted bids that were rejected by the bidder, with rows 12-14 again breaking down this total into three categories depending on whether confirmation of the trade would have given the bidder a payoff gain, a payoff loss, or no change.

There are several observations about these summary bid statistics. First, in all treatments, a large fraction of un-canceled bids were accepted, and a majority of these un-canceled bids ultimately led to a transaction. The percentage of un-canceled bids that ultimately resulted in a transaction ranged from a minimum of 51% in $ABC2$ to more than 65% in AB.

Second, the option for a bidder to cancel bids was regularly exercised. The percentage of bids that were canceled was 27% in AB, 34% in $ABC1$, and 27% in AB.

Third, the bidder's option of rejecting trades, and thus discriminating over who accepted the original bid, played an important role. It was exercised frequently and in the expected direction. Over the three treatments, nearly a third (231 out of 732, or 32%) of accepted bids were rejected by the bidder. Of these 231, only 32 (14%) were bids that would have lead to a payoff gain for the bidder. The very large majority of rejected

	AB	ABC1	ABC2
TOTAL BIDS	243	544	519
Bids Cancelled	66	186	174
BIDS ACCEPTED	169	296	267
With Gain	44	109	109
With Loss	26	48	44
With No Change	99	139	114
TRADES CONFIRMED	115	211	175
With Gain	42	94	72
With Loss	13	23	34
With No Change	60	94	69
TRADES REJECTED	54	85	92
With Gain	6	15	11
With Loss	9	25	36
With No Change	39	45	45

Table 9: Summary of bids and trades

acceptances (199 out of 231) were for trades that would have led to either a loss or no gain for the bidder.

A large percentage of accepted bids and consummated trades did not change any vote outcome and hence led to no payoff gain or loss for either of the trading parties. Overall, 48% (352 out of 732) of accepted bids involved (potential) trades with no change for either party, with similar frequency across the three treatments. We see a similar pattern in the consummated trades, where 45% of those trades resulted in no change in outcomes.

The data can be sorted under several dimensions–round and order, voter id, individual experimental subject. Such break-downs do not add to the substance of what we report in the text, so these finer details are relegated to the appendix. In particular, the appendix shows that there is no evidence of learning or of order effects–behavior appears very consistent across rounds, and regardless of whether $ABC1$ or $ABC2$ was played first.

Pivot trades According to hypothesis H4, trading activity should be dominated by Pivot trades, i.e. by payoff-increasing trades. In line with the theory, we test the hypothesis by considering the fraction of trades associated with myopic strict increases in payoff *for both traders*. In Figure 5, the left bars in the graph correspond to the experimental data, the right bars to the simulations with random trading. The error bars indicate 95%

confidence intervals.[26]

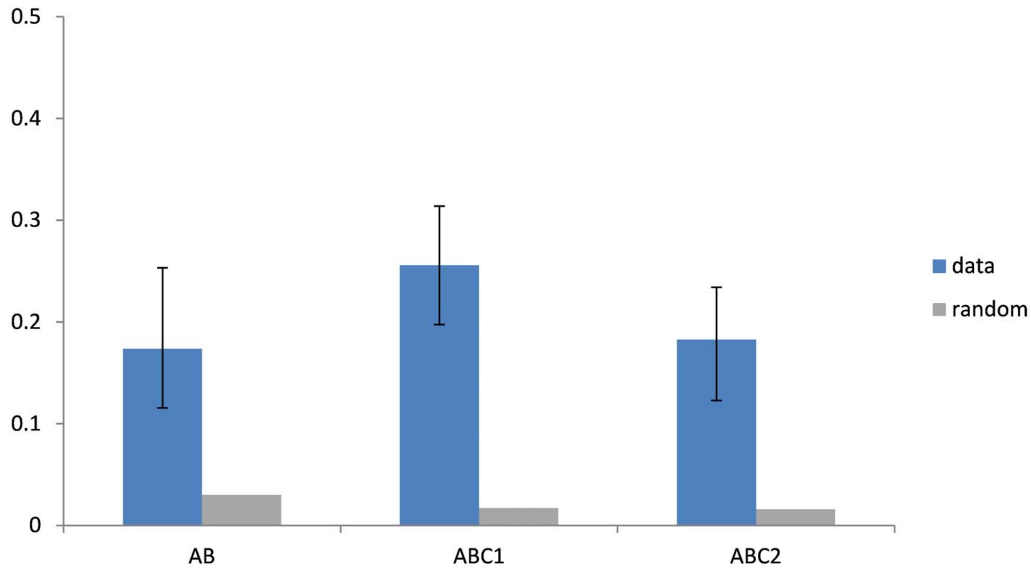

Figure 5: Fraction of Pivot trades.

The figure shows clearly the subjects' search for gains. With random trading, payoff gains for both traders rarely occur (3% in AB and 1% in $ABC1$ and $ABC2$) – less than one fifth of what we observe in AB, and less than one tenth in $ABC1$ and $ABC2$. In all cases, the probability that the data are generated by random trades is negligible. While this is further evidence that the trading behavior of the experimental subjects is not random, it is also true that the fraction of trades consistent with the Pivot algorithm is far from 100% (17% in AB, 26% in $ABC1$ and 18% in $ABC2$). What other kinds of trades occur?

Other trades Besides Pivot trades, observed trades fall mainly in two not mutually exclusive categories. First, rather than requiring strict payoff gains to the traders, a weaker rationality condition is that no trader loses from the trade. *Weak payoff-increasing trades* include zero-gain trades, or trades that involve the exchange of non-pivotal votes and hence leave the current outcome unchanged. Admitting such trades contradicts the focus on stability, since trading need never stop. But zero-gain trades are not irrational because they do not impact the traders' (myopic) payoffs. The fraction of weak payoff-

[26] The confidence intervals are calculated assuming independence, under the null of random trading. Note however that even under random trading the assumption of independence cannot be strictly correct because trades are linked dynamically.

increasing trades (i.e., the sum of Pivot trades and zero-gain trades) is 70% in AB and $ABC1$ and 58% in $ABC2$.[27]

Second, every Pivot trade requires increasing the number of votes held on high-value proposals while reducing the number of votes held on low-value proposals. However, not all such trades are Pivot trades: a trade that induces strict payoff gains must also change the resolution of the proposals concerned. Recall the definition of a voter's score as the product of the subject's number of votes and intensities, summed over all proposals: $\sigma_i(v) = \sum_{k=1}^{K} x_i^k v_i^k$. The score is a shadow value of the total votes held by a voter, reflecting the voter's intensity of preferences and the number of votes held, and remains unchanged whether the voter wins or loses any proposal. *Score-improving trades* are trades that increase both traders' score. All Pivot trades are score-improving, but zero-gain trades can also be score-improving while not improving payoffs (and thus without being Pivot trades).[28] The fraction of observed trades that are score-improving is 61% in AB, 64% in $ABC1$, and 63% in $ABC2$.

Figure 6 shows, for each treatment, the fraction of trades consistent with Pivot trades (left bar), zero payoff-change trades (middle bar), and score increasing trades that are not Pivot trades (right bar).

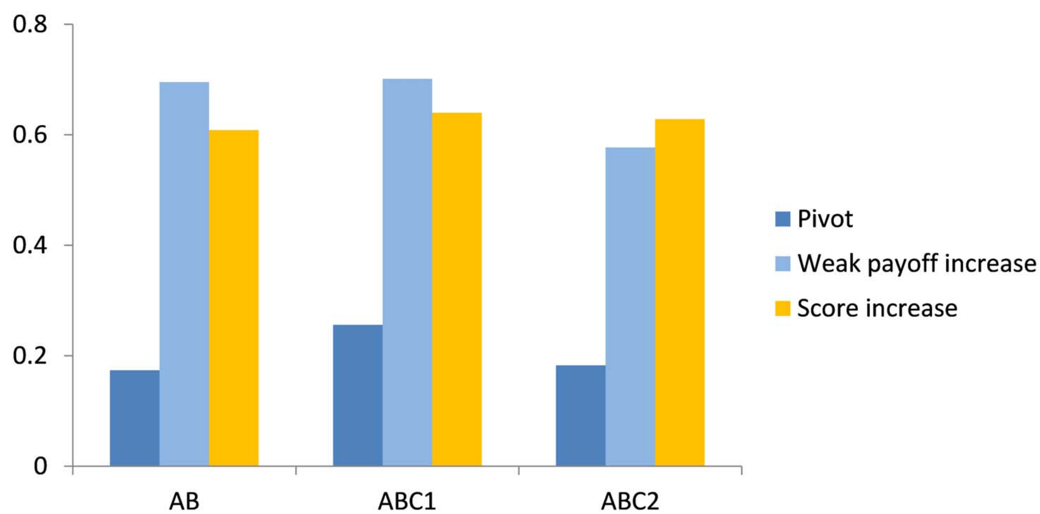

Figure 6: Types of trades.

[27]Note that it is impossible for payoffs to remain unchanged for one of the two traders only: either at least one pivotal vote is exchanged, and the change in outcome affects both traders, or none is, and no outcome changes.

[28]Score-improving trades can also occur with only one of the traded votes being pivotal, in which case one of the traders makes a loss and other trader makes a gain.

Figure 5, above, showed that Pivot trades occur much more frequently than can be accounted for by random trading. Is the same true of weak payoff-improving and score-improving trades? Figure 7 plots, for each treatment, the observed fractions of Pivot trades, zero-payoff change trades, and score-increasing-not-Pivot trades (left bars), together with the corresponding fractions under random trading (right bars), and the 95% confidence interval under the null hypothesis of random trading.

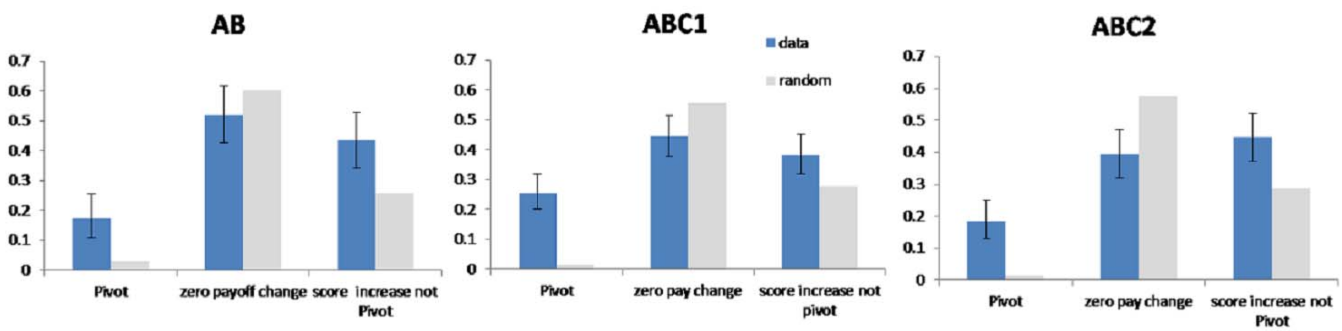

Figure 7: Different types of trades, relative to random trading.

The figure makes clear that although the fraction of zero-payoff changing trades is large, we cannot statistically rule out that it is the result of noise trading: because all non-pivotal trades have zero effect on payoffs, for any given vote distribution a large share of feasible trades belong to this class and thus are realized with high probability under random trading. In fact, in both $ABC1$ and $ABC2$, the evidence suggests that zero-payoff trades are if anything less than random trading generates. This is not true for non-Pivot score-increasing trades: the fraction observed in the data is significantly higher than under random trading ($p < 0.001$ in all treatments).

The figure is suggestive but, because trade categories are not mutually exclusive, it provides a less than complete description of the trade types: many, though not all, of not-Pivot score-improving trades are zero-change trades, and many, though not all, of zero-change trades, are not-Pivot score-improving trades. A more rigorous approach to testing together both the frequency and the intentionality of the different types of trades requires developing and estimating a statistical model of trade classification.

4.2.3 A statistical model of trade types

The statistical model of trade classification we estimate assumes that executed trades fall into four possible categories, not mutually exclusive: (1) Pivot trades; (2) zero-payoff changing trades; (3) score-improving trades; (4) and a residual category of random trades. The model estimates by maximum likelihood four parameters, i.e., the propensities to consummate a trade in each category: π (Pivot); ζ (zero-payoff); ξ (score-improving); and ϵ (random).[29] Each observation is a consummated trade. Given a vector $(\pi, \zeta, \xi, \epsilon)$, the probability the model assigns to an observed trade depends on the categories the trade could theoretically belong to, and on the number of feasible trades (at the current vote allocation) in each category. Because multiple trades are possible from any vote allocation, and the set of feasible trades available at any point in time is largely outside the control of any individual voter, the model assumes the trades are independent observations.[30] The likelihood of observing the data set is then simply the product of the probabilities of all observed trades, and the vector $(\pi, \zeta, \xi, \epsilon)$ can be estimated by maximum likelihood.

An example will help. Suppose, for purposes of explanation, that there were only two observed trades in our entire experiment, one Pivot trade when the vote allocation was v, denoted observation x_1, and one score-improving trade when the vote allocation was v', denoted observation x_2, such that one of the traders suffered a strict myopic payoff loss. Further suppose that at v the number of available Pivot trades is $T^P(v)$, the number of available zero-payoff trades is $T^0(v)$, the number of available score-improving trades is $T^S(v)$, and the total number of available feasible trades is $T(v)$, Then, for a given vector of parameters, $(\pi, \zeta, \xi, \epsilon)$, the model assigns the probability of observation x_1, the Pivot trade at v, to be equal to:

$$p(x_1|\pi, \zeta, \xi, \epsilon) = \frac{\pi}{T^P(v)} + \frac{\xi}{T^S(v)} + \frac{\epsilon}{T(v)}.$$

Similarly, the model assigns the probability of observation x_2, the score-improving trade with myopic loss at v', to be equal to:

$$p(x_2|\pi, \zeta, \xi, \epsilon) = \frac{\xi}{T^S(v')} + \frac{\epsilon}{T(v')}.$$

The likelihood function to be maximized in this simple illustration would then be:

$$L(\pi, \zeta, \xi, \epsilon) = p(x_1|\pi, \zeta, \xi, \epsilon) \cdot p(x_2|\pi, \zeta, \xi, \epsilon)$$

The parameters $(\pi, \zeta, \xi, \epsilon)$ can be interpreted as propensities for the vote market to

[29]The classification model nests both the Pivot trade model ($\pi = 1$) and the simulated random trading model ($\epsilon = 1$), but also allows for the two additional trade types identified in the previous section, zero-payoff and score-improving.

[30]As discussed earlier, the assumption of independence is not strictly correct because of the dynamic linkages across trades.

	AB		$ABC1$		$ABC2$	
π	0.06	[0, 0.14]	0.19	[0.13, 0.25]	0.11	[0.05, 0.17]
ζ	0.11	[0, 0.23]	0.07	[0, 0.16]	0	[0, 0.10]
ξ	0.41	[0.28, 0.55]	0.34	[0.25, 0.43]	0.39	[0.29, 0.49]

Table 10: Model parameter estimates with 95% confidence intervals.

generate trades in the corresponding category. Thus the probability of observing a specific trade is given by the sum of all propensities to choose categories to which the trade belongs, with each propensity divided by the number of feasible trades in that category at the time the trade is executed.

The model implicitly assumes that all four criteria are always feasible, or equivalently that the sets $T(v)$, $T^P(v)$, $T^0(v)$, and $T^S(v)$ are non-empty at all v. Given the three matrices AB, $ABC1$, and $ABC2$ and the one-to-one constraint on trading, $T(v)$ and $T^0(v)$ indeed can never be empty.[31] But this is not true in principle for $T^S(v)$ and for $T^P(v)$. In practice, the problem can be ignored for $T^S(v)$, which is empty only very rarely[32], but not for $T^P(v)$. In AB, 39% of trades (45/115) occur from vote allocations for which $T^P(v)$ is empty; in $ABC1$, the frequency of such trades is 28% (59/211), in $ABC2$, 35% (62/175). Whenever a Pivot trade is not feasible, and thus not observed, the statistical model records the realized trade as setting a zero weight on π. Hence the estimates of π should be considered lower bounds.

We report our estimates of π, ζ, and ξ in Table 10, together with the 95% confidence intervals.[33]

In practice, the model provides a simple way of organizing the trading data, yielding estimates that mirror the observed frequencies of the different trades. There is one im-

[31] One-to-one trades imply that no subject can ever hold more than two (in AB) or three (in $ABC1$ and $ABC2$) total votes. Hence it is impossible for the same subject to hold all five votes on each proposal, the condition required for $T(v)$ to be empty. Similarly, 0-payoff trades are always possible. In each treatment there are at least two subjects with identical directions of preferences. In $ABC1$ and $ABC2$, each holds three votes and there are five votes in total dedicated to each proposal. Thus it is impossible for them to each hold all votes on the same proposal. Hence they can always make a 0-payoff trade. In AB, if the subjects who agree hold all their votes on the same proposal–which in this case, with two total votes per subject, is feasible–then both proposals must be decided by non-minimal majority, and again trades that do not affect payoffs are always possible.

[32] The frequency of trades from vote allocations such that $T^S(v)$ is empty is 5% percent for AB and less than 1% for both $ABC1$ and $ABC2$.

[33] We constructed the confidence intervals by bootstrapping the data. We estimated the probabilities of each type of trade from the data; we then constructed confidence intervals by running 1000 simulations using those estimates and re-estimating the probabilities. Confidence intervals do not account for possible correlations in the data.

portant exception: once the overlap across categories is recognized, there is essentially no evidence of intentional zero-profit trades in any of the three treatments (in all treatments the 95% confidence interval for ζ includes 0). There is however a significant probability of Pivot trades in treatments $ABC1$ and $ABC2$, and of score-improving trades in all three treatments. Keeping in mind that the π estimates are lower bounds, the estimation confirms the message of the figures. Again as implied by the figures, π and ξ are not collinear and can be estimated separately.[34]

4.2.4 Score Improving Trades

The high estimated fraction of score-improving trades across all three treatments invites the question of why subjects engage in such trades. One possibility is that score improvement is the ultimate objective, either for motives we do not understand or because of some confusion regarding payoffs or pivotality. However, there is evidence to suggest that this is not the case.

A first reason to be skeptical is the frequency of rejected trades reported in Table 9 above.Across all treatments more than two thirds of the rejected trades would have led to a score improvement for the bidder.[35] In contrast to payoff-improving trades, score improving trades do not depend on the identity of the trading partner, and thus all proposed trades that find a taker should be confirmed by the bidder.

In addition, we find that opportunities for score improvement almost always remain open when subjects stop trading. We have defined stability as the absence of any feasible strictly payoff-improving pairwise vote trade. One can define the alternative concept of *score stability* as the absence of any feasible score-improving pairwise vote trade.[36] But score stability is not a useful characterization of final vote allocations. As shown in Table 11, the fraction of score-stable final vote allocations is 34% in AB, 14% in $ABC1$, and 6% in $ABC2$; the corresponding frequencies of Pivot stable vote allocations are 76%, 64%, and 64%, respectively.[37] Final vote allocations are far more likely to be payoff stable than

[34] To check the conjecture of a downward bias in our estimates, we re-estimated the model focusing only on the subsample for which Pivot trades were available. The resulting estimates of the Pivot propensity π increase by a large factor in all three treatments: from 0.06 to 0.15 in AB; from 0.19 to 0.29 in ABC1, and from 0.11 to 0.20 in ABC2, with all confidence intervals bounded away from 0.

[35] Recall that about a third of all accepted bids are rejected by the bidder.

[36] A score-stable vote allocation always exists with pairwise trading.

[37] As a finer analysis, one can define distance from score stability, analogously to the earlier definition of distance from payoff stability. The distance of final vote allocations from score stability is significantly greater than the distance from stability. Specifically, the distribution of steps to score stability stochastically dominates the distribution of steps to payoff stability, and the difference is large and highly significant.

	AB	ABC1	ABC2
Payoff Stable	76.5	64.3	64.3
Score Stable	34.0	14.3	6.0

Table 11: Percentage of final vote allocations that are score stable, compared to payoff stable.

score stable.

These observations suggest that score-improvement in itself does not produce a strong incentive for trading in our experiment. Although speculative and post-hoc, an alternative rationalization of these trades is to attribute them to precautionary behavior in the face of uncertainty about future trades. In the complex environment of our experiment, subjects may want to accumulate votes on high value proposals, because of a concern that further trades are likely to take place before voting actually occurs. A forward-looking voter may then choose to buy non-pivotal votes for a high-value proposal her favorite side is already winning in order to increase the margin of victory and lock-in the outcome. Is there evidence then that trading is far-sighted?

4.2.5 Is trading behavior farsighted?

Forward-looking behavior can be modeled rigorously. Vote trading is a form of dynamic barter in which others' trades affect both the feasibility and the desirability of one's own trades, making a fully strategic analysis particularly complex.[38] Still, it is possible to make some progress by exploiting concepts that have been developed in cooperative game theory.[39]

The formalization requires three standard definitions, adapted to elementary trades:

Definition 6. *Given two vote allocations v and v', a pair of voters $D = \{i, j\}$ is said to be* effective *for (v, v') if (v, v') is an elementary trade between i and j.*

Definition 7. *A* chain *from v to v' is an ordered sequence of vote allocations $v^1, v^2, ..v^m$, with $v^1 = v$ and $v^m = v'$, and a corresponding sequence of effective pairs $D^2, .., D^m$ such*

[38]The difficulty is shared by other games with similar structure, for example matching and network formation games. And indeed such games are typically analyzed under myopia or other strongly restrictive conditions. See for example, Roth and Vande Vate (1990), Diamantoudi et al (2004), Watts (2001), Jackson and Watts (2002).

[39]See for example Harsanyi (1974), Chwe (1994), Mauleon et al. (2011), Ray and Vohra (2015), Dutta and Vohra (2017), and the references therein. Note however that because of the externalities involved and because the opportunities for trade depend on the current vote allocation, vote trading cannot be represented under any of the existing cooperative models of farsightedness.

that for all $t = 1, ..m-1$, D^{t+1} *is effective for* (v^t, v^{t+1}).

Finally:

Definition 8. *A collection of vote allocations* $v^1, v^2, ..v^m$, *with* $v^1 = v$ *and* $v^m = v'$ *is a* farsighted chain *if it is a chain, and, in addition,* $u_j(v^t) < u_j(v')$ *for all* $t = 1, ..., m-1$ *and all* $j \in D^{t+1}$. *If there exists a farsighted chain from* v *to* v', *then* v' *is said to* farsightedly dominate *(F-dominate)* v.

Using these basic concepts, there are several possible ways to define farsightedly stable vote allocations. The most intuitive is the pairwise parallel of the farsighted core: it states that an allocation v is (pairwise) farsightedly stable if there exists no v' that F-dominates v.[40] Other definitions are possible, and in general problems of existence are not trivial.[41] Developing a full analysis is beyond the scope of this paper, but our goal is much more limited: farsightedness builds on Harsanyi's (1974) notion of indirect dominance, as defined above. If subjects in our experiment are farsighted, then their trades should be such that the final vote allocation reached at the end of the round should be associated with a payoff gain for each trader, relative to the vote allocation at which the subject traded. Was this the case?

Table 12 reports, for each treatment, the fraction of trades associated with farsighted gains for both traders (row 2), with farsighted losses for both traders (row 3), and, for comparison, the fraction of trades that were Pivot trades (that is, trades associated with myopic gains, row 4).

In all treatments, the fraction of trades with farsighted gains is less than 10%, and less than about a third of the fraction of Pivot trades (half of that for $ABC1$); in the two three-proposal treatment, it is less than half of the fraction of farsighted losses. On the basis of the these numbers alone, it is hard to put much weight on farsighted domination as engine of trade.

The evidence of score-improving trades suggests that subjects give some thought to future trades, but standard notions of farsightedness adapted from cooperative game theory cannot explain the experimental data.

[40] Defining farsighted stability in terms of the farsighted core is the direction followed by Casella and Palfrey (2018). Note that farsighted stability is much more demanding that myopic stability: v' can F-dominate v even if trades generate temporary myopic losses, as long as the final allocation v' is preferred to the allocation at which each voter trades. What matters is the utility comparison between the end point of the chain and the vote allocation at which trading occurs.

[41] A vote allocation where all votes are held by a single voter is in the farsighted core and thus is pairwise farsighted stable, according to this definition. Other plausible definitions, however, do not guarantee existence in our setting. In addition, none addresses the more interesting question of whether stability can be reached from the starting vote allocation.

	AB	ABC1	ABC2
Farsighted gains	5.2	3.8	6.3
Farsighted losses	2.6	11.8	14.3
Pivot trades	17.4	25.6	18.3

Table 12: Percentage of trades yielding farsighted gains and losses, compared with percent Pivot trades.

5 Conclusions

This paper presents the results of a laboratory experiment designed to explore the theoretical implications of a dynamic model of vote trading. The theoretical approach has two essential features: (1) an equilibrium concept based on stability; and (2) a rational vote trading process, which we call Pivot trading. Stable vote allocations are those for which there are no strictly myopic payoff-improving vote trades for any pair of voters. The trading process defines the possible sequences of payoff-improving trades that converge to a stable vote allocation.

The experiment delivers four main findings. First, two-thirds of all final vote allocations in the experiment are stable, supporting the hypothesis of stability as the equilibrium criterion. Second, individual final vote holdings line up closely with the theory: in all treatments, each trader type's vote allocation at the end of the round, averaged over rounds and groups, always changes in the direction predicted by the theory: increasing, relative to the initial allocation, when the theory predicts that the subject will be a net buyer of votes, and decreasing when the theory predicts net selling. Third, final proposal outcomes most frequently correspond to Condorcet winners. In two of the treatments this coincides with the Pivot-stable outcome; in the third treatment it does not, but the Pivot-stable outcome is the second-most frequently observed outcome. Because the Condorcet winner is the pre-trade outcome, this is suggestive of a bias in favor of the pre-trade outcome, analogous to a status quo bias. Fourth, we find weaker support for the dynamic process of trade posited by the model, as many of the observed trades do not yield mutual strict myopic payoff increases.

To identify the relative frequency of different types of trades, we develop and estimate a statistical model to classify trades, and the estimation indicates that, when noise trading is filtered out, a large fraction of trades are score-improving, but not necessarily payoff-improving. A cautionary note about this finding is that the estimation procedure biases downward the estimated fraction of payoff improving trades, because such trading opportunities often fail to exist. Score-improving trades are vote exchanges in which each voter trades a vote on a less salient issue in exchange for a vote on a more salient issue, but does not necessarily immediately benefit from the trade in terms of changed outcomes. We conjecture that score-improving trades may be pursued for their precautionary value, suggesting the possibility of some farsighted thinking. However, rational farsighted trading is unambiguously rejected by our data: on average, a trade is twice as likely to leave the two traders eventually worse off in the final outcome as it is to make them better off.

This study only scratches the surface of possibilities for laboratory analyses of vote trading and logrolling. There are many interesting environments that are not represented by the three studied in the paper. First, a Condorcet winner exists in all three environments in this study, but we know that more generally a Condorcet winner may not exist. It would be interesting to explore such preference configurations and study whether the inertia towards pre-trade outcomes we observe in our data remains true in the absence of a Condorcet winner. Second, the experiment studies pairwise trading, but it would also be interesting to explore more complex coalitional trades. The pairwise vote trading model extends quite naturally to coalitional vote trading, although designing a user-friendly trading interface would be a major challenge. Related to this point, there are alternative ways one could organize the market. For example, one could allow communication among voters, either concurrently with or prior to the actual trading protocol. Communication might help voters identify beneficial trading partners. Other possible extensions of the trading process include allowing package trades, or allowing voters to target their offers to specific other members.

The experimental findings are also suggestive of useful extensions of the theoretical framework. The evidence we find for score-improving trades suggests precautionary incentives. Understanding such precautionary motives requires allowing for risk aversion and modeling the strategic uncertainty faced by vote traders - uncertainty about trades that future voters might engage in. As presently formulated, the model of vote trading operates only on the ordinal preferences of voters over the profile of final outcomes. With uncertainty, preferences would be defined on the space of lotteries over outcomes and would require a somewhat different theoretical approach.

References

[1] Bentley, A. F. 1908, *The Process of Government*, University of Chicago Press: Chicago.

[2] Bernholz, P. 1973, "Logrolling, Arrow Paradox and Cyclical Majorities", *Public Choice*, 15, 87-96.

[3] Bernholz, P., 1974, "Logrolling, Arrow Paradox and Decision Rules: A Generalization", *Kyklos,* 27, 49–62.

[4] Berninghaus, S. K., K.-M. Ehrhart, and M. Ott, 2006, "A Network Experiment in Continuous Time: The Influence of Link Costs," *Experimental Economics*, 9, 237-251.

[5] Buchanan, J.M. and G. Tullock, 1962, *The Calculus of Consent*. Ann Arbor: University of Michigan Press.

[6] Carrillo, J. and A. Gaduh, 2016, "Are Stable Networks Stable? Experimental Evidence." *Working Paper*. University of Southern California.

[7] Casella, A. and T. Palfrey, 2018, "Trading Votes for Votes. A Dynamic Theory", *Social Science Working Paper No. 1444*, California Institute of Technology.

[8] Casella, A. and T. Palfrey, 2019, "Trading Votes for Votes. A Dynamic Theory", *Econometrica*, 87, 631-652.

[9] Choi, S., S. Goyal, and F. Moisan, 2019, "Brokerage", *Cambridge Working Papers in Economics* # 2005, University of Cambridge.

[10] Choi, S., S. Goyal, and F. Moisan, 2020, "Connectors and Influencers", *Working Paper*, University of Cambridge.

[11] Chwe, M., 1994, "Farsighted Coalitional Stability," *Journal of Economic Theory*, 63, 299-325.

[12] Coleman, J., 1966, "The possibility of a social welfare function", *American Economic Review*, 56, 1105–1122.

[13] Coleman, J., 1967, "Reply", *American Economic Review*, 57, 1311–1317.

[14] Dutta, B. and R. Vohra, 2017, "Rational Expectations and Farsighted Stability," *Theoretical Economics*, 12, 1191-1227.

[15] Echenique F. and L. Yariv, 2012, "An Experimental Study of Decentralized Matching", Working Paper, Caltech.

[16] Fischbacher U. and S. Schudy, 2014, "Reciprocity and Resistance to Comprehensive Reform", *Public Choice* 160:411-428.

[17] Greiner, Ben, 2015, "Subject pool recruitment procedures: organizing experiments with ORSEE", *Journal of the Economic Science Association*,1: 114. https://doi.org/10.1007/s40881-015-0004-4

[18] Guerrero, O. and U. Matter, 2016, "Revealing the Anatomy of Vote Trading", Saïd Business School Research Paper 2016-26, University of Oxford.

[19] Haefele, E., 1971, "A Utility Theory of Representative Government", *American Economic Review*, 61, 350-365.

[20] Harsanyi, J., 1974, "An Equilibrium-Point Interpretation of Stable Sets and a Proposed Alternative Definition," *Management Science*, 20, 1472-495.

[21] Jackson, M. and A. Watts, 2002, "The Evolution of Social And Economic Networks," *Journal of Economic Theory*, 106(2): 265-295.

[22] Jackson, M. and A. Wolinsky, 1996, "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory*, 71(1): 44-74.

[23] Kadane, J., 1972, "On the Division of the Question", *Public Choice*, 13, 47-54.

[24] Kagel, J. and A. Roth, 2000, "The Dynamics of Reorganization in Matching Markets: A Laboratory Experiment Motivated by a Natural Experiment", *Quarterly Journal of Economics,* 115, 201-235.

[25] Kirchsteiger, G., M. Mantovanid, A. Mauleone, and V. Vannetelbosche. 2016. "Limited farsightedness in network formation". *Journal of Economic Behavior and Organization* 128, 97-120.

[26] Koehler, 1975, "Vote Trading and the Voting Paradox: A Proof of Logical Equivalence", *American Political Science Review*, 69, 954-960.

[27] Ledyard, J.O., D. Porter, and A. Rangel, 1994, "Using Computerized Exchange Systems to Solve an Allocation Problem in Project Management", *Journal of Organizational Computing,*4, 271-296.

[28] McKelvey, R. and P. Ordeshook, 1980, "Vote trading: An experimental study", *Public Choice*, 35, 151-184.

[29] Mauleon, A., V. Van Netelbosch and W. Vergote, 2011, "von Neumann-Morgenstern farsighted stable sets in two-sided matching," *Theoretical Economics*, 6, 499-521.

[30] Mayhew, D. 1966, *Party Loyalty Among Congressmen*, Harvard University Press: Cambridge.

[31] Miller, N., 1977, "Logrolling, Vote Trading and the Paradox of Voting: A Game Theoretical Overview", *Public Choice*, 30, 51-75.

[32] Mueller, D., 1967, "The Possibility of a Social Welfare Function. Comment", *American Economic Review*, 57, 1304-1311.

[33] Mueller, D.C., 1973, "Constitutional Democracy and Social Welfare", *Quarterly Journal of Economics*, 87, 61–79.

[34] Nalbantian, H. and A. Schotter, 1995, "Matching and Efficiency in the Baseball Free-Agent System: An Experimental Examination", *Journal of Labor Economics*, 13, 1-31.

[35] Niederle, M. and A. Roth, 2009, "Market Culture: How Rules Governing Exploding Offers Affect Market Performance,". *American Economic Journal: Microeconomics* 1: 199-219.

[36] Pais,J., A. Pinter and R Veszteg, 2011, "Decentralized Matching Markets: A Laboratory Experiment", mimeo, ISEG/Technical University of Lisbon.

[37] Park, R.E., 1967, "The Possibility of a Social Welfare Function. Comment", *American Economic Review*, 57, 1300-1304.

[38] Ray, D. and R. Vohra, 2015, "The Farsighted Stable Set," *Econometrica*, 83, 977-1011.

[39] Riker, W. and S. Brams, 1973, "The Paradox of Vote Trading", *American Political Science Review*, 67, 1235-1247.

[40] Schwartz, T., 1975, "Vote Trading and Pareto Efficiency", *Public Choice*, 24, 101-110.

[41] Schwartz, T., 1977, "Collective Choice, Separation of Issues and Vote Trading", *American Political Science Review*, 71, 999–1010.

[42] Stratmann, T. , 1992, "The Effects of Logrolling on Congressional Voting", *American Economic Review*, 82, 1162-1176.

[43] Stratmann, T. , 1995, "Logrolling in the U.S. Congress", *Economic Inquiry*, 33, 441-56.

[44] Tullock, G., 1970, "A Simple Algebraic Logrolling Model", *American Economic Review*, 60, 419-426.

[45] Wilson, R., 1969, "An Axiomatic Model of Logrolling", *American Economic Review*, 59, 331-341.

	Final outcome is stable	Stable outcome is final
Round	0.003	0.021
	(0.028)	(0.021)
Order	0.034	-0.111
	(0.174)	(0.151)
Treatment ABC1	-0.189	0.006
	(0.144)	(0.13)
Treatment ABC2	-0.186	-0.011
	(0.158)	(0.148)
Constant	0.722***	0.511***
	(0.125)	(0.125)
Observations	191	295
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 13: Probability of stable outcomes

Appendix

Supplementary Data Analysis

Learning and order effects

We state in the text that we do not observe either learning or order effects in the data. We report here the evidence on which the statement is based.

Table 13 reports the results of a linear probability model, estimating the probability that the final vote allocations is stable (column 2) and that a stable allocation leads to no further trade (column 3) as functions of the round number, the order of the treatments, the treatments themselves and a constant.[42] The estimation is by OLS with standard errors clustered at the (group × treatment × session) level, thus ensuring that the group composition is constant. With the exception of the constant term, none of the parameters is significant.

Table 14 reports the results of a similar linear regression for the probability that a trade is Pivot. As above, the estimation is via OLS and errors are clustered at the (group × treatment × session) level. None of the estimated parameters is significant, here including

[42]Note that the variable Order is always set to 1 for treatment AB, which is always ordered first. Hence for AB Order is simply an addition to the constant. But the order changes across sessions for $ABC1$ and for $ABC2$.

	Dependent variable: Pivot trade
Round	-0.01
	(0.017)
Order	0.104
	(0.117)
Treatment ABC1	-0.008
	(0.09)
Treatment ABC2	-0.07
	(0.102)
Constant	0.103
	(0.087)
Observations	501
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 14: Probability that a trade is pivotal

the constant term, reflecting the low frequency of Pivot trades.

The two tables presented above evaluate learning in terms of the predictions of the model. This is an unusually complex experiment, and we can ask a more basic question too: do subjects learn to avoid irrational actions, in the sense of agreeing to a trade that induces a myopic payoff loss?[43] Such actions can take three possible forms: a subject may accept a posted bid that would induce a loss for the subject; a bidder may confirm a trade that causes the bidder a loss, or reject a trade that would result in a gain.

Table 15 reports, for each treatment, the results of a linear regression where the dependent variable is the frequency of myopic loss actions and the explanatory variables are the round number, the order, and the voter type, i.e. the subject ID, from 1 to 5. The estimation is by OLS and the standard errors are clustered at the (group $\times$ session) level. The subject ID matters because assigned preferences vary by ID, and thus so do opportunities to trade. For example, note that in treatment AB the probability of a myopic loss action is highest for subject 5 (as well as being the most significant of that regression's coefficients). According to the theoretical model, subject 5 should not trade at all; every trade by 5 is classified as a myopic loss action.

In treatment AB, all voter types make significantly more myopic loss actions than ID 1, but there is no significant difference across voter types for the other treatments.

[43]In principle forward-looking behavior could justify myopic losses; in practice we suspect they mostly reflect confusion.

	AB	ABC1	ABC2
Round	-0.019	-0.009	-0.015
	(0.025)	(0.096)	(0.012)
Order		0.078	0.03
		(0.052)	(0.067)
ID 2	0.145**	-0.021	0.031
	(0.062)	(0.052)	(0.056)
ID 3	0.058*	-0.053	0.019
	(0.033)	(0.041)	(0.059)
ID 4	0.102**	0.005	-0.023
	(0.042)	(0.045)	(0.056)
ID 5	0.280***	0.035	0.1
	(0.092)	(0.063)	(0.071)
Constant	0.094	0.042	0.195**
	(0.082)	(0.080)	(0.097)
Observations	338	592	592
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table 15: Probability of a myopic loss action

Treatment	Voter type (ID)	Bids	Accept	Concluded as seller	Concluded as buyer	Concluded total	Myopic Loss
AB	1	56	31	32	20	52	3
	2	52	22	25	19	44	10
	3	55	37	22	26	48	6
	4	40	61	19	36	55	12
	5	40	18	17	14	31	14
ABC1	1	109	76	38	54	92	20
	2	113	67	45	46	91	17
	3	101	61	37	44	81	11
	4	133	55	59	44	103	22
	5	88	37	32	23	55	16
ABC2	1	105	62	40	36	76	15
	2	116	51	35	41	76	19
	3	78	47	20	29	49	14
	4	104	30	41	23	64	10
	5	116	77	39	46	85	31

Table 16: Summary of bids, acceptances, and concluded trades.

The constant term is relatively high and significant for treatment $ABC2$ but no other coefficient is statistically significant. In particular, there is no evidence of learning.

Bids and trades by voter type

Besides the context of learning, trading activity by voter type is interesting for its own sake. Table 16 reports the relevant data. For each voter type, in each treatment, we report the number of bids posted (column 3), bids accepted (column 4); trades concluded as bidder (i.e. bids that were posted by the corresponding ID, accepted by another trader, and confirmed by the bidder with the given ID–column 5); trades concluded as acceptor (i.e. bids posted by others that the voter with the corresponding ID accepted and saw confirmed by the bidder–column 6); and finally the total number of trades in which the corresponding ID took part, either as bidder or as acceptor (column 7, the sum of columns 5 and 6).

The table shows that all voter types were active and no gross asymmetry appears in the data. The most extreme is the disparity in the trades concluded by IDs 4 and 5 in $ABC1$, which differ almost by a factor of 2. But one problem in interpreting the table is the possibility that it reflects unusual behavior by individual subjects who happen to be given specific voter types in the lab.

Individual trading behavior

The focus of analysis in this paper is the trade–the pairwise exchange between two subjects. Because trading requires the joint action of two different voter types, and the opportunities and gains from trade depend on the current vote allocation, as well as on the voters' preferences, each trade has an idiosyncratic element. It is not informative to classify experimental subjects in terms of the type of trades they conclude–Pivot trades, or zero-payoff changing trades, or score-improving trades. Although still dependent on assigned preferences and on others' behavior, the frequency of myopic loss actions is the one category of trading activity that more than others reflects individual choices, and it on its basis that we can evaluate heterogeneity in our population of subjects. The frequency of myopic loss actions must be considered in relation to the total number of trading actions the subject makes–the number of bids the subject accepts, and, when bidder, the number of accepted bids the subject either accepts or rejects.

Figure 8 reports, for each experimental subject, the number of myopic loss actions on the vertical axis, versus the subject's total number of trading actions on the horizontal axis. Because the focus is on the individual subject, and subjects take actions in all treatments, the data are not separated by treatment.[44] In the figure, bubbles come in four different sizes, from 1 to 4, corresponding to the number of points that are superimposed.

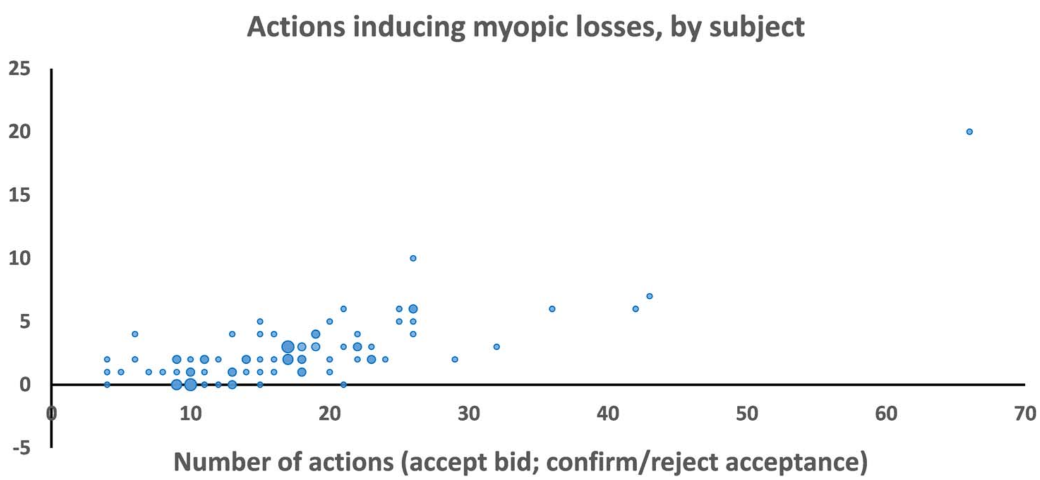

Figure 8: Number of myopic loss actions by subject.

The minimum total number of trading actions any subject took is 4; i.e., every subject engages in some trading activity. More than 75% of subjects (65/85) made between 10 and 26 actions in total; more than 75% (64/85) made between 0 and 3 irrational actions (more than 60% (53/85) made 0 to 2 irrational actions). If we classify as outliers subjects making more than 30 total trading actions and 10 or more irrational ones, then we have

[44]Because two sessions had only two treatments, not all subjects played the same number of rounds, another reason why accounting for the total number of trading actions is important.

6 outliers, in a population of 85. It is clear for example that the subject on the upper far right corner of the figure is an outlier, with 66 total trading actions, of which 20 were irrational.

Experimental Instructions

VOTE TRADING INSTRUCTIONS

Make yourself comfortable, and then please turn off phones and don't talk or use the computer. Thank you for agreeing to participate in this decision making experiment. You will be paid for your participation in cash, at the end of the experiment. Different participants may earn different amounts. What you earn depends partly on your decisions and partly on the decisions of others. If you have any questions during the instructions, raise your hand and your question will be answered. If you have any questions after the experiment has begun, raise your hand and an experimenter will come and assist you.

The experiment today is a committee voting experiment, where you will have an opportunity to trade votes before voting on an outcome. The experiment will be in three parts. At the end of the experiment you will be paid the sum of what you have earned in all three parts of the experiment, plus your promised show-up fee of 10 dollars. Everyone will be paid in private and you are under no obligation to tell others how much you earned. Your earnings during the experiment are denominated in POINTS. For this experiment every 100 POINTS earns you 6 DOLLARS.

Here are the instructions for Part 1.

You will be randomly assigned to one of 3 committees, each composed of 5 members. Each committee is completely independent of the others, and the decision taken in one committee has no effect on the others. The committee will vote using majority rule to decide on 2 different motions, denoted A and B. Each motion can either pass or fail depending on how the committee votes. There will be a separate vote on each motion. The computer will assign you a committee member number (1, 2, 3, 4, or 5). Part 1 consists of 3 rounds.

You will be told, for each motion, whether you prefer it to pass or to fail. The computer will assign you (and each other member) a value for each motion which will be a number between 1 and 100. You will earn your value for a motion if you prefer that motion to pass and it passes, or if you prefer it to fail and it fails. This is your only source of earnings. Your earnings for the round are equal to the sum of your earnings over the two motions.

Each committee member starts a round with 1 vote to cast on each motion. Then there will be a 2 minute trading period, during which you and the other members of your committee will have an opportunity to trade votes with each other. For example, you may wish to trade your A vote in exchange for some other member's B vote. We will describe exactly how to do this shortly.

After the trading period ends, you will proceed to the voting stage. Once everyone has voted, you will be told what the final votes were in your committee and how much you earned in that round. This will complete the first round. The remaining 2 rounds in Part

1 follow the same rules. Each committee member starts the round with a single vote on each motion. Your committee member number, preferences for each motion (pass or fail), your value for each motion, and the preferences and values of the other four members of your committee all stay the same for all 3 rounds of part 1 of the experiment.

Your earnings for part 1 are the sum of your earnings in all 3 rounds. After round 3 ends, I will read you instructions for part 2 of the experiment.

We now describe in detail how you and the other members of your committee can trade votes. When we begin a round, you will see a screen like this, although the exact numbers may be different. [Display Screen 1] On the right of the screen is an Information Table that contains a lot of information, so please listen carefully. It displays each member's preference for each motion (pass or fail), value, and number of votes. If the member prefers the motion to fail, then the value is written in a blue color. If the member prefers the motion to pass, then the value is written in an orange color. You can simply think of there being two sides - the orange side and the blue side - on each motion. The number of votes held by each member on each motion is in parentheses. Because no trading has occurred yet, each member holds exactly one vote on each motion.

Your own row is specifically labeled and the label is highlighted in gray. The last row in the table is labeled "outcome". This row tells you, for each motion, what the total vote would be if voting took place now, by showing the column sum of votes on each motion. The number of votes for is given first, in orange, and the number of votes against is given second, in blue. If the votes in favor of a motion exceed the votes against, then all voters who prefer the motion to pass will earn their value for that motion, and all voters who prefer the motion to fail will earn zero for that motion. Similarly, if the votes in favor of a motion failing exceed the votes in favor of it passing, then all voters who prefer the motion to fail will earn their value for that motion, and all voters who prefer the motion to pass will earn zero for that motion. There is a check mark next to your value if the outcome of that motion is the outcome you prefer. This means that you earn your value for that motion. In this example, if there were no votes traded at all, then on motion A, there are 2 votes held by members who prefer A to pass and 3 held by members who prefer A to fail, so motion A fails. On motion B, there are 3 votes held by members who prefer B to pass and 2 held by members who prefer B to fail, so motion B passes. Since ID 1 (You) prefers both motions to pass, he earns his value for motion B but earns 0 for motion A.

To the left of the table, in grey, is the trading window. At any time during the trading period, any committee member may post a trade offer by requesting 1 vote on one motion in exchange for 1 vote on some other motion. Suppose the participant on the slide in front of the room wanted to post a trade requesting one A vote in exchange for one B vote. This is done by entering a 1 in the A box under "Requests" and a 1 in the B box under "Offers". [Screen 2]. You can only trade 1 vote for 1 vote; you can neither request nor offer multiple votes.

After you have entered this trade request and clicked the "submit trade offer" button, the trade is posted in the trading panel for everyone in your committee to see. [SCREEN 3] If another committee member wants to accept your trade request, they may click on it to highlight it, and then click on the "accept selected offer" button.[SCREEN 4] You now have 10 seconds to either confirm or reject the accepted trade. A message will pop-up on your screen. [SCREEN 5]. The message tells you what the outcome of the vote would be if you either accept or reject the trade and voting took place without any further trade. If you reject the trade or do nothing for 10 seconds, the trade does not occur. The committee member who had accepted your offer is informed that you declined to confirm the trade.[SCREEN 6]. Your offer is re-posted in the trading window, and some other voter can accept it. If you confirm the trade, then the voter who accepted the offer now holds 0 A votes and 2 B votes, you now hold 2 A votes and 0 B votes, and the Information Table is updated accordingly. The new Information Table is displayed for 10 seconds on a popup screen for everyone in your group to see. [SCREEN 7]

If you have a standing offer listed in the trading window, you may cancel it by first clicking on it and then clicking the "cancel selected offer" button.[SCREEN 8]

The trading period continues for 2 minutes. The timer at the top tells you how much time remains in the trading period. The clock is frozen when the Information Table is shown after a trade, with the new vote holdings. If a trade occurs within 10 seconds of the end of the trading period, the trading period is automatically lengthened by 10 more seconds.

You are free to post trade requests at any time, but you are not allowed to offer to trade away a vote on a motion if you currently hold 0 votes for that motion or already have an offer posted on the trading window that would result in holding 0 votes if accepted. In that case you would first have to cancel your existing posted offer. Also remember that you can only trade one vote for one motion in exchange for one vote for another motion. If you try to do a trade that is not allowed, you will either receive an error message, or the action buttons will become gray and be deactivated, preventing you from proceeding with that trade.

When the trading period for the round is over, we proceed to the voting stage. Your screen will now look something like this: [SCREEN 9]. In this stage you do not really have any choice. You are simply asked to click a button to cast all the votes you hold at the end of trading. The computer will automatically cast your votes on each motion according to the preferences you were assigned. For example, if you prefer motion B to fail and you hold two B votes after the trading period, those two votes will be cast automatically against motion B. Please cast all your votes without delay by clicking on the vote button.

After you and the other members of the committee have voted, the results are displayed and summarized. [SCREEN 10]

As the experiment proceeds, at the bottom of each screen you will see a history table, summarizing the results of the previous rounds [SCREEN 11. Go over the different columns] If you switch to tab view, each round will be shown separately].

We then proceed to the next round, where you again start out with one vote on each motion and the rules are the same as in the first round. Remember that your assigned committee number, preferences for motions, values for motions, and those of the other members of your committee all stay the same for all 3 rounds of part 1 of the experiment. After the first 3 rounds are completed, we will read instructions for the second part of the experiment.

To give you some experience with the trading screen, we will conduct two practice rounds. The rules will be the same as they will be in the paid rounds, but the values and preference assignments, for or against a motion, are not the same as they will be in the paid rounds. You are not paid for the practice rounds, so they have no effect on your final earnings. The only purpose of the practice rounds is to help you become familiar with the computer interface and the trading rules.

This summary slide [SCREEN 12: Summary slide] will remain up during the experiment to remind you of the rules on trading and on time.

Are there any questions before we proceed to the first practice round? [START SERVER]

Please click on the icon marked Multistage Client on your desktop. Then enter the number of your carrel (on the right side of the carrel), click enter, and then wait. Remember that you are not allowed to use the computer for any other purposes while waiting during the experiment (email, browsing, etc.).

[CONNECT EVERYONE AND START]

Please complete the practice rounds on your own. Feel free to raise your hand if you have a question.

[WAIT FOR SUBJECTS TO COMPLETE PRACTICE ROUNDS]

The practice rounds are now over. Remember, you will not be paid the earnings from the practice rounds.

If you have any questions from now on, raise your hand, and an experimenter will come and assist you. We will now begin the paid rounds.

(Play 3 real rounds for Part 1) [After last ROUND, read:]

We will now proceed to Part 2. The rules for part 2 are the same as for part 1, but there are now 3 motions for your group to vote on. You can only trade one vote on one motion for one vote on another motion. The trading period will last 3 minutes. As before, 10 seconds will be added to the clock if a trade takes place within 10 seconds of the time limit.

The values and pass/fail preferences will be different from part 1, and your committee number as well as the composition of your committee may change. However, both the preferences and the composition of the committee will remain the same for all of Part 2. Part 2 will last for 5 rounds. At the end of the 5 rounds, we will stop and read the instructions for Part III.

Are there any questions before we begin?

(Play 5 real rounds for part 2) [After last ROUND, read:]

We will now proceed to Part 3. Part 3 is identical to Part 2, but the values and pass/fail preferences may be different. Your committee number as well as the composition of your committee may also change. Part 3 will again last for 5 rounds and again the trading period is 3 minutes (plus 10 seconds if a trade is concluded within 10 seconds of the time limit).

This is the end of the experiment. You should now see a popup window, which displays your total earnings in the experiment. Please record this and your Computer ID on your payment receipt sheet, rounding up to the nearest dollar. After you are done, please, click ok to close the popup window. Do not close any other windows on your computer and do not use your computer for anything else. Also enter 10 dollars on the show-up fee row. Add the two numbers and enter the sum as the total.

[Write output]

We will pay each of you in private in the next room in the order of your computer numbers. Remember you are under no obligation to reveal your earnings to the other players. Please do not use the computer; be patient, and remain seated until we call you to be paid. Do not converse with the other participants or use your cell phone. Thank you for your cooperation.

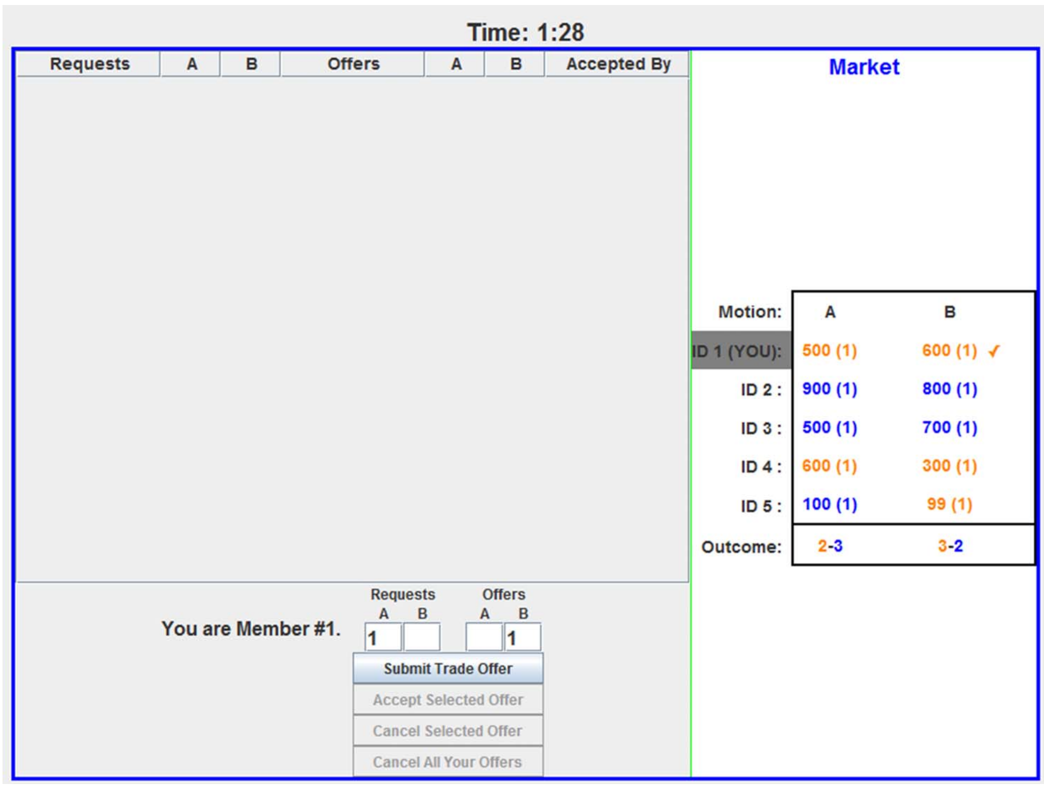

Figure 9: Screenshot for a subject posting a bid.

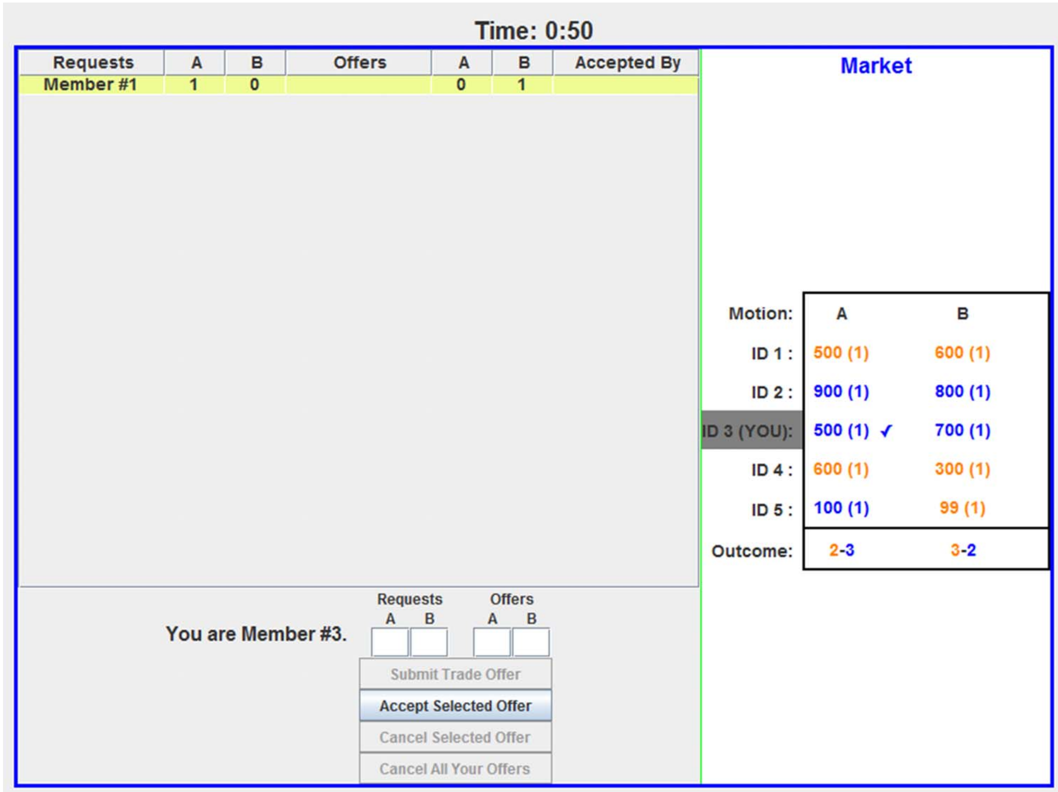

Figure 10: Screenshot for a subject accepting a posted bid.

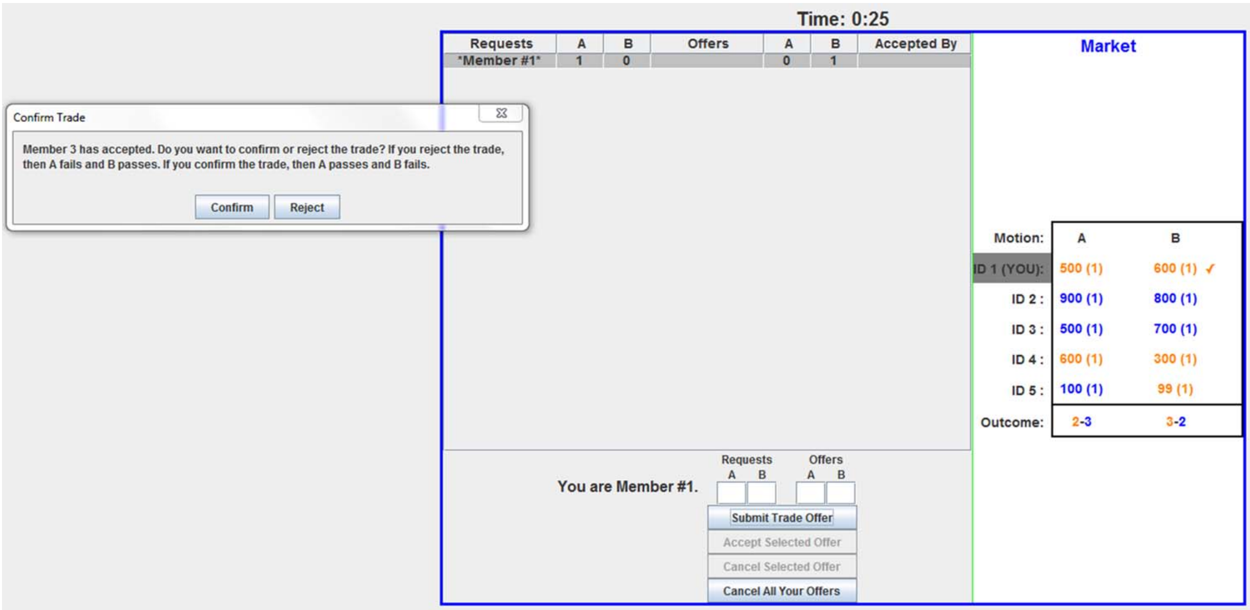

Figure 11: Confirmation request for the bidder.