Trading Votes for Votes: An Experimental Study*

Alessandra Casella† Thomas Palfrey‡

May 8, 2017

Abstract

Trading votes for votes is believed to be ubiquitous in committees and legislatures, and yet we know very little of its properties. We return to this old question with a laboratory experiment and a simple theoretical framework. We posit a family of minimally rational trading rules such that pairs of voters can exchange votes when mutually advantageous. Such rules always lead to stable vote allocations–allocations where no further improving trades exist. Our experimental 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 our theoretical predictions. However, there is only weak support for the dynamic trading rules themselves, and although trading is frequent final outcomes show significant inertia around pre-trade outcomes.

JEL Classification: D70, D72, P16

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

---

*We thank Enrico Zanardo, Kirill Pogorelskiy and Manuel Puente for research assistance, and participants to numerous seminars and conferences for comments. We especially wish to thank Micael Castanheira, Andrew Gelman, Michel LeBreton, Debraj Ray, Richard Van Veelden, Rajiv Vohra, and Alistair Wilson for detailed comments and suggestions. The National Science Foundation (SES-1426560) 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".

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

‡Corresponding author. HSS Division, Mail Code 228-77, 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. Political scientists have long emphasized the vital role of vote trading and logrolling in collective decision making. Common sense, personal experience, empirical and historical studies all suggest the extent and the importance of such institutions.

To many, such behavior is not only widespread, but marginally unethical. A legislator voting against the interests of the voters who elected him runs counter to basic democratic principles of representation. However, well over a century ago, an early pioneer in political science, Arthur F. Bentley, argued that this view was shortsighted and unrealistic; that logrolling was 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 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)

There is a relatively small literature that attempts to document specific cases of vote trading, mostly in the context of the U.S. Congress. Mayhew's (1966) book is the first comprehensive study, focusing on agricultural bills in the house, and there is much anecdotal evidence in earlier research. Stratmann (1992) and Stratmann (1995) identify roll call votes where a legislator votes against his constituency's interest and exploit econometric techniques to 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 cosponsorhip through big data techniques.

Outside these studies, systematic evidence on vote trading remains scarce, in contrast to the common belief in its prevalence. The disparity between evidence and perceptions can be attributed in part to vote trading's tainted reputation–representatives voting against voters' interests are unlikely to publicize the fact–, in part to institutional features that can effectively serve to "hide" vote trading–for example the committee system in the US Congress, through which logrolling is embedded in the writing of the bills. But more generally, the belief that vote trading is common is due to our anecdotal experience of its ubiquity wherever power is delegated to committees, across institutions, settings and countries, not only in formal but also in relatively informal settings: professional associations, school boards, faculty committees, neighborhoods and buildings' owners associations, cooperatives, cultural and civic institutions boards, and many more, all settings that do not lend themselves easily to the collection and analysis of systematic data.

The scarcity of empirical studies is matched by the scarcity of rigorous theoretical works. Notwithstanding general agreement on the importance of understanding vote trading, after an early, enthusiastic wave of work in the 1960's and 70's,[1] the theoretical literature mostly ran dry. One reason is that the problem is difficult. Consider the simplest framework, the natural first step studied 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. Every committee member can be in favor or opposed to any proposal and has separable preferences across proposals, with different cardinal intensities. Even in this environment, 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' power to affect outcomes, and induce further trades.

In a companion paper (Casella and Palfrey, 2017) we develop the bare bones of a dynamic theory of vote trading that applies to this simple environment. In

[1] See, among others, Buchanan and Tullock (1962), Coleman (1966, 1967), Park (1967), Wilson (1969), Tullock (1970), Haefele (1971), Kadane (1972), Bernholz (1973), Riker and Brams (1973), Mueller (1967, 1973), Koehler (1975), Miller (1977a, 1977b), Schwartz (1975, 1977).

the present paper, we present the results of testing the theory's predictions in a laboratory experiment. The use of laboratory methods to study vote trading seems particualrly apporpriate given both the difficulty of collecting historical data and the ability of controlled experiments to address fundamental, micro-level questions of behavior, the crucial questions of *why, when, and how* vote trades emerge from the chaos of committee wheeling and dealing. And yet, if empirical and theoretical studies of vote trading are not numerous, experimental studies are even fewer. To our knowledge, the study closest to ours is McKelvey and Ordeshook (1980), but the differences in methodologies (face-to-face exchanges in McKelvey and Ordeshook, computer-mediated platforms in this paper) and especially in objectives (a focus on alternative cooperative solutions in McKelvey and Ordeshook, on dynamics in this paper) make a direct comparison impossible.[2] We think of this paper as a first exploratory step towards a full-fledged non-cooperative understanding of the dynamics of vote trading.

In studying their original framework, Riker and Brams conjectured that restrictions on trading may be needed to prevent continuing cycles: one exchange of votes changes the outcomes that would be reached if voting were held, and hence makes other voters consider new trades, which again induce further trading. Evaluating whether Riker and Brams' conjecture is correct requires a rigorous definition of stability and a formal model of dynamic adjustment. Stability is identified with an allocation of votes such that no pair of voters can trade votes and induce an out-

[2] 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.

come they both prefer. Dynamic adjustment occurs via an algorithm that selects, with some arbitrary rule, a pair of voters with strictly improving trades; if the trade induces stability, then trading stops; if not, a new pair of voters is selected. The process continues until a pairwise stable allocation is reached. We show in Casella and Palfrey (2017) that, contrary to Riker and Brams' conjecture, such a process is always guaranteed to converge to a stable vote allocation: continuing cycles of trading will not occur.

This minimal theoretical apparatus - a definition of stability and the specification of a dynamic trading process - is enough to generate strong predictions that can be tested in an experimental setting: specific predictions about final vote allocations, proposals' outcomes, and even exact sequences of trades.

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

We reach three main conclusions. First, we find that stability is a useful predictive tool. In all treatments, two thirds or more of the final vote allocations after trading are stable.

Second, the final vote allocations in the experiment are in line with the theoretical predictions. 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.

And yet the final outcomes deviate significantly from the model and display some inertia around the no-trade outcome.[3] The reason, and this is our third result, is that while we do see the gain-searching trades predicted by the theory, a larger fraction of trades do not lead to strict improvement for the two voters engaging in the trade. Rather, while the trades do indeed increase the number of votes held on high-value issues, this often happens without changing the outcomes associated with the new vote allocation. Interestingly, trades that increase the number of votes on high-value issues are not predictive of final vote allocations: trading stops when no opportunity for payoff gains remains, as in our stability concept, even when it is still possible to increase votes on high-value proposals.

Shifting votes towards higher-value proposals suggests some form of prudential behavior. The theoretical trading dynamic is instead myopic: trades are considered profitable if the vote allocation immediately resulting from the trade strictly improves the payoff of the traders, relative to the current vote allocation. The trading data from the experiment suggest that the myopia assumption should be considered more carefully. We conclude the paper with an exploration of possible extensions of the model to allow for farsighted vote trading, and re-examine our experimental data in this new light. We show that the definition of farsightedness leads directly to some simple predictions that can be confronted with the data. In our experiment, fully farsighted behavior is soundly rejected.

Methodologically related to our trading protocols are some recent experiments on decentralized matching, in particular Echenique and Yariv (2013).[4] In those experi-

[3] This is not saying that there is frequently no trade. On the contrary, 96% of our groups make at least one vote trade. The final outcome often corresponds to the no-trade outcome, but the final vote allocation does not.

[4] Other related works are Nalbantian and Schotter (1995), Niederle and Roth (2011) and Pais,

ments, as in ours, a central finding 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 2000). 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.[5]

Finally, if seen as a trading experiment, in the spirit of good markets experiment, a peculiarity of our design is the lack of a common unit of value. 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 discusses

Pinter and Vesztegz (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.

[5] 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.

the experimental design; section 4 reports the experimental results, and section 5 concludes. A short appendix reports the detailed trading paths predicted by the theory with our experimental parameters. The instructions from a representative experimental session are available in a second appendix online.

2 The Model

2.1 The Voting Environment

A committee of N (odd) voters must approve or reject each of K independent binary proposals, a set denoted by P. Committee members have 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. 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: it is helpful to 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 over some proposals and several votes over others, but cannot hold negative votes on any proposal. We call v_i^k the votes held by voter i over proposal k, $V_i = \{v_i^k,\, k = 1, .., K\}$ the set of votes held by i over all proposals, and $V = \{V_i,\, i = 1, .., N\}$ the *vote allocation*, i.e., the profile of vote holdings over all voters and proposals. The initial vote allocation is denoted by V_0, and we set

$V_0 = \{\mathbf{1}, \mathbf{1}, ..\}$: prior to any trade, each voter has a single vote over each proposal. 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 $v_i^k \in V$. [6]

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 P_k if $z_i^k > 0$, and against P_k if $z_i^k < 0$.[7] We indicate by $\mathbf{P}(V) \in P$ the set of proposals that receive at least $(N+1)/2$ favorable votes, and therefore pass. We call $\mathbf{P}(V)$ the *outcome* of the vote if voting occurs at allocation V. Note that with K independent binary proposals, there are 2^K potential outcomes (all possible combinations of passing and failing for each proposal). Finally, we define $u_i(V)$ as the utility of voter i if voting occurs at V: $u_i(V) = \sum_{k \in \mathbf{P}(V)} z_i^k$.

Although the theory allows for trading within coalitions of arbitrary size, in the experiment trades are restricted to be bilateral. We impose such a constraint in part because pairwise trading is typically considered more empirically relevant[8], and in part to limit complexity in what already is an unusually complicated experimental platform. We thus specialize the model to pairwise trades only.

Our focus is on the properties of vote allocations that hold no incentives for further trading. We define:

Definition 1 *An allocation $V \in \mathcal{V}$ is **stable** if there exists no pair of voters i, i' and*

[6] Note that $\sum_k v_i^k \neq K$ is feasible because we are allowing a voter to trade votes on multiple issues in exchange for one or more votes on a single issue. Of course, the aggregate constraint $\sum_i \sum_k v_i^k = NK$ must hold.

[7] We assume that all preferences are strict, and hence rule out $z_i^k = 0$ for all k and all i.

[8] Riker and Brams (1973) for example, argue that the difficulty of organizing a coalition makes non-pairwise trading unlikely. Guerrero and Matter (2016) build their empirical strategy on the asusmption of pairwise trades.

no $\widehat{V} \in \mathcal{V}$ *such that* $\widehat{V}_j = V_j$ *for all* $j \neq i, i'$, *and* $u_i(\widehat{V}) > u_i(V)$, $u_{i'}(\widehat{V}) > u_{i'}(V)$.

Note that a stable vote allocation always exists: a feasible allocation of votes that yields dictator power to a single voter i is trivially stable: no exchange of votes involving voter i can make i strictly better-off; and no exchange of votes that does not involve voter i can make anyone else strictly better-off. The interesting question is not whether a stable allocation exists, but whether it is reachable through sequential decentralized trades.

2.2 Trading Dynamics

We next specify the dynamic process through which trades take place. We begin with the following definition.

Definition 2 *A trade is* ***minimal*** *if it consists of a minimal package of votes such that both members of the pair strictly gain from the trade.*

Concentrating on minimal trades allows us to "unbundle" complex trades into elementary trades. Multiple welfare-improving trades cannot be bundled, and zero-utility trades cannot be bundled with strictly welfare-improving trades.

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. In line with these earlier analyses, we define the *Pivot Algorithms* as sequences of trades yielding myopic strict gains to both traders:

Definition 3 *A **Pivot Algorithm** is any mechanism generating a sequence of trades in the following way: Start from the initial vote allocation V_0. If there is no minimal strictly improving trade, stop. If there is one such trade, execute it. If there are multiple improving trades, choose one according to some choice rule R. Continue in this fashion until no further improving trade exists.*

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, corresponding to the family of possible R rules.

Pivot trades are not restricted to two proposals only: a voter can trade his vote, or votes, on one issue in exchange for the other voter's vote(s) on more than one issue. The only constraint is that trades be minimal: any reduction in the number of votes traded prevents the trade from being strictly payoff-improving for at least one of the two voters. If a trade is welfare improving and minimal, it is a legitimate trade under Pivot.[9]

Trades are required to be strictly welfare improving for the participating pair. That means that pivotal votes must be traded: trades of non-pivotal votes cannot affect outcomes and thus cannot induce changes in utility. More than that: since we restrict trades to be minimal, *only* pivotal votes can be traded. It is this property,

[9] Ruling out the bundling of multiple payoff improving trades is for simplicity only. Ruling out the bundling of zero-utility trades with welfare improving trades plays instead a substantive role. Zero-utility trades cause no immediate gains or losses, but affect the feasibility of future profitable trades. Allowing them to be bundled would affect the dynamics of the vote allocations, without the discipline provided by the requirement of payoff gains.

anticipated by Riker and Brams, that gives the name to our algorithm.

2.3 Pivot-Stable Allocations

An obvious question to ask is whether trading under Pivot algorithms will ever stop: in principle there is nothing to rule out the possibility of trading cycles. Fortunately, the answer to the question is positive. For all K, N, Z, all Pivot algorithms converge to a stable vote allocation in a finite number of steps. The term "all Pivot algorithms" refers to the arbitrariness of the choice rule R: convergence is guaranteed for any R.[10]

The generality of the result is unexpected: the Pivot algorithms *always* reach a stable vote allocation, regardless of the number of voters and proposals, for all (separable) preferences, and regardless of the order in which different possible trades are chosen. No such general result applies, to our knowledge, to other games in which successive moves occur in the absence of an equilibrating price process–for example in matching, or network formation, or barter trading, all cases in which convergence to stability requires some randomness in rule R.[11] In vote trading, Riker and Brams (1973) conjectured that convergence required limiting the number of allowed trades per proposal; Ferejohn (1974) believed that it may fail.

In fact the intuition is surprisingly simple. When trades occur under a Pivot algorithm, both voters trade away votes on proposals they value less (on which they have a relatively low $|z_i^k|$), in exchange for votes on proposals they value more.

[10] See Casella and Palfrey (2017).

[11] Randomness in R ensures that any cycle will be broken. See Roth and Vande Vate (1990), and Diamantoudi et al. (2004) for matching; Jackson and Watts (2002) for network formation games; Feldman (1973) and Green (1974) for barter trading.

Given the current vote holdings for voter i, we can define the total intensity-weighted value of i's vote holdings, or *score,* as $S_i(v_i) = \sum_k |z_i^k| v_i^k$. When i trades under a Pivot algorithm $S_i(v_i)$ increases, and therefore so does the total group score, $S(v) = \sum_i S_i(v_i)$. Since there are a finite number of issues and votes, $S(v)$ is bounded above, so at most a finite number of Pivot trades are possible.[12]

We call any vote allocation reachable by a Pivot algorithm a *Pivot-stable Vote Allocation*, and any outcome associated with a Pivot-stable vote allocation a *Pivot-stable Outcome.*

Vote trading environments are unusually complex: votes' values depend on their pivotality, and thus change with others' allocations; trades by others affect the desirability of further trades, and thus a single trade can generate a whole chain of new exchanges; externalities ensure that individuals' welfare depends on others' trades; no continuous price exists. Pivot algorithms are simple, intuitive rules, describing plausible trades in such a complicated environment. Their simplicity allows some conceptual progress, as in our stability result. But we have posited them for a second reason too: we conjecture that they may have predictive power. We now turn to testing the Pivot algorithms in the laboratory.

3 The Experiment

The experiment was conducted at the Columbia Experimental Laboratory for the Social Sciences (CELSS) in November 2014, with Columbia University registered students recruited from the whole campus through the laboratory's ORSEE site.

[12] See Casella and Palfrey (2017).

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 during the experiment, and answered all questions publicly.[13]

Because the design of the trading platform presents some challenges, we describe it here is some detail.[14]

At the start of each treatment, each subject's computer screen displayed the matrix of values, denominated in experimental points, and the vote allocation. We refer to this matrix as the *vote table.* The interface and the instructions associated the two alternatives for each issue, Pass or Fail, with two colors, Orange (Pass) and Blue (Fail). Each individual's values were written in the color of the individual's preferred alternative. All experimental values were positive and indicated earnings from one's preferred alternative winning, relative to zero earnings if it lost.[15] The screen also showed the votes totals and the points the subject would win if voting were held immediately. Each subject started with one vote on each issue.

After having observed the matrix of values and the current vote allocation, a subject could post a bid for a vote on one of the issues, in exchange for his vote on a different issue. The bid appeared on all committee members' monitors, together with the ID of the subject posting the bid. A different subject could then accept the

[13] Sample instructions are provided in the online appendix.

[14] 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/).

[15] 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.

bid by clicking the offer and highlighting it.[16]

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 specific 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 market was open for three minutes.[17] However, in a market where each concluded trade can trigger a new chain of desired trades, it is important to ensure adequate time for all desired trades to be executed. For this reason the time limit

[16] Sample screenshots are provided in the online appendix.

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

was automatically extended by 10 seconds whenever a new trade was concluded.

The theory allows for trades of multiple votes and over multiple proposals, but with the matrices of values assigned to subjects during the experiment minimal Pivot trades would amount to trades of a single vote on one issue against a single vote on a different issue. In the experiment then we allowed only such trades, with the goal of limiting the complexity of the task (without affecting our theoretical predictions). No bid could be posted if a subject did not have enough votes to execute it if accepted; thus a voter could post multiple bids only as long as he had enough 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. Then a new round started.

The experiment consisted of three treatments, AB, $ABC1$, and $ABC2$, each corresponding to a different matrix of values. In all three treatments, the size of the voting committee was five ($N = 5$), while the number of issues depended on the treatment: $K = 2$ in treatment AB, and $K = 3$ in treatments $ABC1$, and $ABC2$. In each committee, subjects were identified by ID's randomly assigned by the computer, and issues 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 rounds; then three rounds of treatment AB, and then five rounds each of $ABC1$ and $ABC2$, alternating the order.[18] We did not alternate the order of treatment AB because its smaller size ($K = 2$) made it substantially easier for the subjects, and thus we used it as further practice before the more complex treatments. This is also the reason for the smaller 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 group and subjects' ID's were kept unchanged for all rounds of the same treatment, to help subjects learn. All but one sessions consisted of 15 subjects, divided into three committees of five subjects.[19] At the end of each session, subjects were paid their cumulative earnings from all rounds, converting experimental points into dollars via a preannounced exchange rate, plus a fixed show-up fee. Each session lasted about 90 minutes, and average earnings were $36, including a $10 show-up fee.

We designed the three treatments according to the following criteria. First, we wanted a $K = 2$ treatment, as further training for the subjects. Second, we chose value matrices 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 stable allocation can be reached via multiple paths, with no path being clearly focal. Third, the older literature discussed at length, and with contradictory results, the relationship between stable vote allocations reachable via vote trading and the existence of the Condorcet winner. We designed matrices for which the Condorcet winner exists, but need not correspond to the Pivot stable out-

[18] Two of the sessions had only two treatments: AB and $ABC1$ in one case, and AB and $ABC2$ in the other.

[19] One session had only ten subjects, divided into two groups.

come: it does in AB and in $ABC2$, but not in $ABC1$. The two matrices $ABC1$ and $ABC2$ are superficially very similar and have Pivot trading paths of comparable multiplicity and length, allowing us to test whether the Condorcet winner has stronger attraction. 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.

The three value matrices used in the experiment are given in Table 1.

AB

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

*ABC*1

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

*ABC*2

	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. *Matrices of values used in the experiment.*

In all three cases, the initial vote allocation $V_0 = \{\mathbf{1}, \mathbf{1}, ..\}$ is unstable. Consider for example matrix AB. At V_0, proposal A fails and proposal B passes; 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 $\mathbf{P}(V_1) = \{A\}$, which both 1 and

3 prefer to $\mathbf{P}(V_0) = \{B\}$. 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, and with no further trade the outcome reverts to $\{B\} = \mathbf{P}(V_2)$. Indeed, no further trade can occur: all pivotal votes are held by voters 2, 4 and 5, none of whom can gain from trading. It is straightforward to verify that there are no other trading sequences that are consistent with a Pivot algorithm. The Pivot algorithm 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 issue on which an extra vote is acquired by the voter listed first, the path is $\{13ab, 24ab\}$. The unique Pivot-stable outcome is $\mathbf{P} = \{B\}$, 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 $\mathbf{P} = \{A\}$, but the unique Pivot-stable outcome is $\mathbf{P} = \{A, B, C\}$. The Pivot algorithm can follow three alternative paths, two of them of length four (i.e. consisting of four trades), and one of length three. The three paths are: $\{13cb, 45bc, 23ab, 45ca\}$, $\{23ab, 45ca, 45bc, 13cb\}$, and $\{23ab, 45ba, 13cb\}$. In matrix $ABC2$, the Condorcet winner is $\mathbf{P} = \{A, B, C\}$, and corresponds to 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 $\{24cb, 15ac, 34ba\}$.[20]

Table 2 reports the experimental design.

[20] Notice that for all three matrices, the limitation that trades must be one-for-one was inessential, as the only theoretically possible Pivot trading sequences involved only such trades.

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.*[21]

4 Experimental Results.

4.1 Trading

How much trading did we see? Table 3 reports basic statistics on observed trades. "Pivot" refers to the predicted number of trades under the Pivot algorithm. The unit of analysis is the group per round.

Treatment	Tot trades	groups × rounds	Mean trades	Median	s.d	Max	Pivot
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 3. *Number of trades.*

A histogram of the number of trades per treatment (per group per round) (Figure 1) shows the higher frequency of shorter trade paths in the AB treatment, with

[21] A programming error in sessions s5 and s6 made the last five rounds of data unusable.

$K = 2$. Between the two $K = 3$ treatments, $ABC2$ has higher fractions of shorter trades, but the differences are not striking–56 percent of rounds end with two or fewer trades in $ABC2$, as opposed to 41 percent in $ABC1$, and 80 percent end with three or fewer trades in $ABC2$, as opposed to 76 percent in $ABC1$. In all treatments, few rounds include five or more trades.

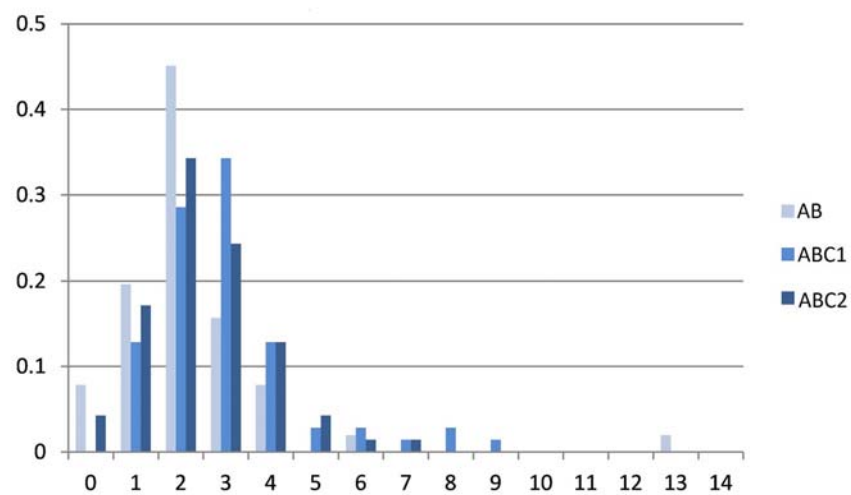


Figure 1. *Number of trades. Frequencies.*

As expected, the bidder's option of rejecting trades, and thus discriminating over who accepted the original bid, was important. In columns 2-4 of Table 4, we report the total number of bids, how many of these bids found a taker in the market, and how many of these acceptances were then rejected by the bidder. A large fraction of all posted bids found a counterpart–from a minimum of 77 percent in $ABC2$ to more than 95 percent in AB–but about a third of these accepted trades were rejected by the bidder– 32 percent in A, 29 percent in $ABC1$, and 34 percent in $ABC2$. As the last column of the table shows, some rejected trades were associated with a strict increase in myopic payoff for the bidder, but the number is small–between 10 and 20 percent of rejections in all treatments.

Treatment	Tot bids	Accepted	Rejected by bidder	Rejected with payoff gain
AB	177	169	54	6
$ABC1$	368	296	85	15
$ABC2$	345	267	92	11

Table 4. *Bids, accepted bids, and rejected trades.*[22]

Whether in terms of number of trades or of any other variable studied below, the data show no evidence of learning or of order effects–behavior appears very consistent across rounds, and regardless of whether $ABC1$ or $ABC2$ was played first. Thus we present the experimental results aggregating over rounds and order.

4.2 Stability

Our point of departure is the definition of stable vote allocations. Is the stability requirement satisfied in the vote allocation to which our subjects converge at the end of each round? Figure 2 shows the CDF of steps to stability for the three treatments, in black, as well as in 5,000 simulations with random trading, in red. The horizontal axis measures the minimal number of Pivot trades necessary to reach stability, and the vertical axis the proportion of final vote allocations not further from stability than the corresponding number of trades.

[22] Tot bids excludes canceled bids.

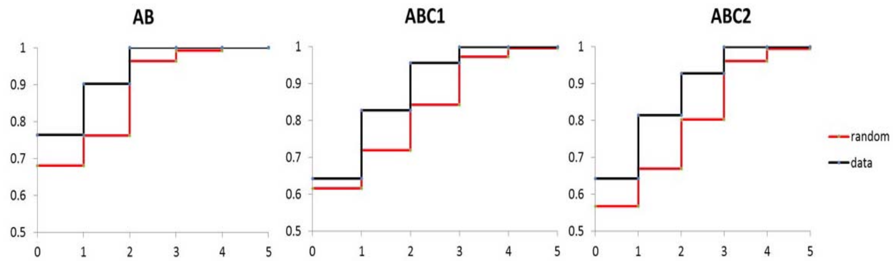


Figure 2. *Steps to stability.* Cumulative distribution functions.

The fraction of stable vote allocations in the experimental data was 76 percent in AB, and 64 percent in both treatments $ABC1$ and $ABC2$. In all treatments, more than 80 percent of all vote allocations were within one step (one trade) of stability, although the figure also shows the predictably easier convergence to stability in the AB treatment, with only two proposals. In all three treatments, the distribution corresponding to random trading FOSD's the distribution for the experimental data.

The simulation of random trades provides the yardstick of comparison for our data. We will use it repeatedly in what follows, and it is worth describing the methodology in some detail. In each treatment, we constructed the random trades by randomly selecting an individual, one or two issues (in the two- and three-issue treatments, respectively), a partner, and a direction of trade, all with equal probability, and enacting the trade as long as both traders' budget constraints were satisfied. If budget constraints are violated, we cancel the proposed trade and restart. In each group, a trade occurs with specified probability over a short time interval, with both parameters calculated to match the observed average length of rounds and the average number of trades in the treatment.[23] For each treatment, we repeated the procedure 5,000 times, each time focusing on a group.

Random trading is a demanding comparison when applied to the stability of vote allocations because a large fraction of feasible trades take the vote allocation away from minimal majority, and hence make pivot trades impossible, and the allocation stable.[24] But Figure 2 is informative beyond the comparison to random trading, and that is because our soft timing constraint de facto allows subjects to choose when to stop trading. A high fraction of stable allocations at the end of the rounds is indicative of either a search for or at least of a recognition of stability, of opportunities for payoff gains having been exploited.

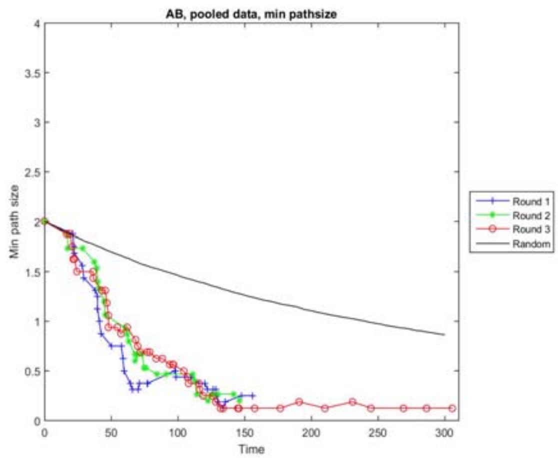
Figure 2 reports information on the stability of the vote allocations reached at the end of trading. But our data also give us information on dynamic convergence. Do successive trades move the vote allocation towards stability?

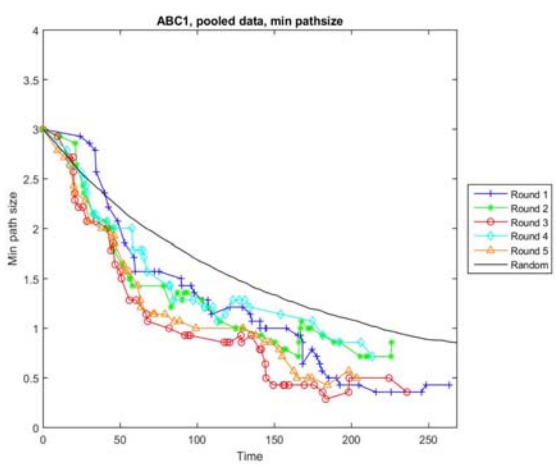
Figure 3 shows, for each treatment, the dynamic path of the vote allocation, as captured by the succession of trades. The horizontal axis measures time, in seconds. A marker corresponds to a trade. Thus, for any given marker, the horizontal axis indicates when the trade took place, within the maximal round length observed in the data for each treatment. The vertical axis measures distance from stability, defined, as in Figure 2, by the minimal number of Pivot trades necessary to reach a stable allocation. Such number is calculated first for the vote allocation characterizing

[23] Given the average length of a round in the treatment, time is divided into a grid of 100 cells, and in each cell a group can trade with probability p, such that $100p$ equals the mean number of trades per round in the treatment.

[24] 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.

each group in the treatment at that moment in that round, and then averaging over the groups. The figure is drawn pooling over all groups and all sessions, for given treatment, and each curve, with its own shade and marker symbol, reports data from the same round (1-3 for AB and 1-5 for $ABC1$ and $ABC2$). The jumps between dots are relatively small because a trade concerns a single group, while the others' vote allocations remain unchanged.





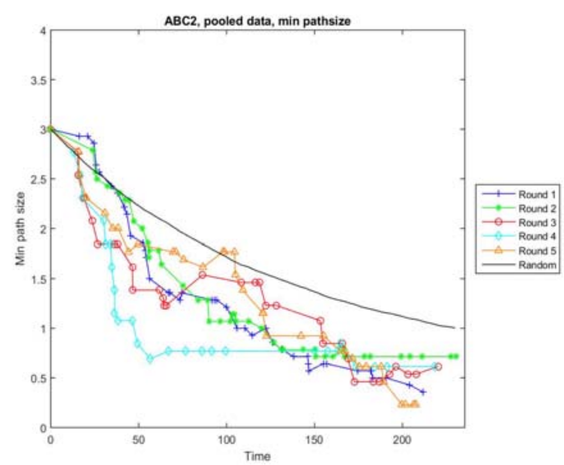


Figure 3. *Dynamic convergence to Pivot stable outcomes.* Data vs. Random.

All curves decline, almost perfectly monotonically, showing the dynamic convergence towards stability. To help us evaluate such convergence, the black curve in each panel reports the steps from stability calculated from the 5,000 simulations with random trading. After the first minute, in all three treatments, the curve corresponding to random trades remains higher than the curve corresponding to any round of experimental data.[25] Notice also the lack of learning in the data–there is no systematic difference between earlier and later rounds.[26]

4.3 Vote Allocations

For all three value matrices used in our experiment, the Pivot algorithms predict a unique stable vote allocation. Is such an allocation reached by the experimental subjects? Figure 4 reports the number of votes held by each voter at the end of a round, averaged over all rounds of the same treatment. Each panel corresponds to a treatment and reports the number of votes by voter ID, i.e. by the vector of values corresponding to each column of the value matrix. The blue columns represent the experimental data, the grey columns the Pivot prediction, and the red line the no-trade status quo (or equivalently, the average vote holding after random trading). The figure reports data from all rounds, but remains effectively identical if we select stable vote allocations only.

[25] With the exception of two trades in round 5 in $ABC2$.

[26] To verify that results were not driven by averaging, we computed CDF's of steps to stability for the data and for the random simulations, as in Figure 2, at all 30-second intervals. In all treatments and at all times, the CDF corresponding to random trading FOSD's the CDF from the data.

The vote distribution in the data is less sharply variable across issues than theory predicts, as we would expect in the presence of noise. Yet, the qualitative predictions are strongly supported. There are five voters in each treatment, holding votes over two (in AB) or three issues (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. When the theory predicts holding a single vote–11 cases for which the voter should exit trade with the same number of votes held at the start–, the data show three cases where the average vote holding is below 1, five where it is above, and three where it is effectively indistinguishable from 1. On average, our subjects hold 0.56 votes when the theory predicts 0; 1.05 when the theory predicts 1, and 1.43 when the theory predicts 2.07.[27]

[27] The theory predicts that voter 3 in treatment $ABC1$ should hold three votes.

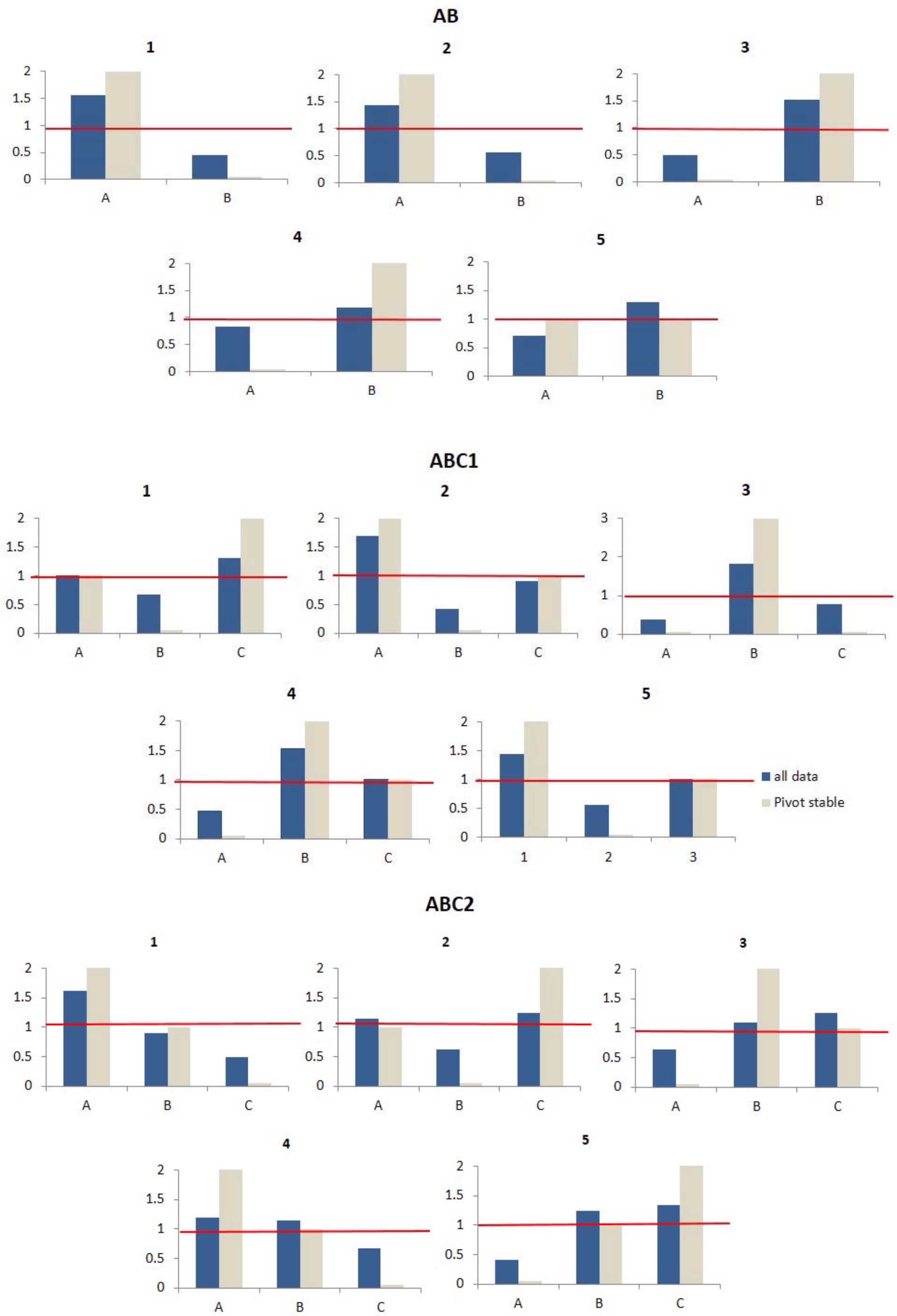


Figure 4. *Average vote allocations at the end of each round, by voter type.*

4.4 Trades

According to our results so far, final vote allocations tend to be stable; dynamic trading moves towards stability, and final individual vote holdings mirror qualitatively the properties of Pivot-stable allocations. But can we say more about the specific trades we see in the lab? In particular, are these trades compatible with the Pivot algorithm?

4.4.1 Pivot trades.

The class of pairwise Pivot algorithms is a class of mechanical selection rules among feasible pairwise trades. Accordingly, we test it on binary trades–i.e. by considering the fraction of all trades associated with myopic strict increases in payoff for both traders.[28] We plot such a fraction in Figure 5. The blue columns correspond to the experimental data, the light grey columns to the simulations with random trading, and the error bars indicate 95 percent confidence intervals (under the null of random trading).[29]

[28] Alternatively, we could consider the fraction of *individual* trades that induce strict (myopic) gains, a weaker test of our model. But Pivot algorithms are not equivalent to optimizing rules of individual behavior–the latter would presumably include a search for maximal gain, competition for specific traders, endogenous surplus division, etc..

[29] Note that under the null all observations are independent. Thus no correction for correlation is required.

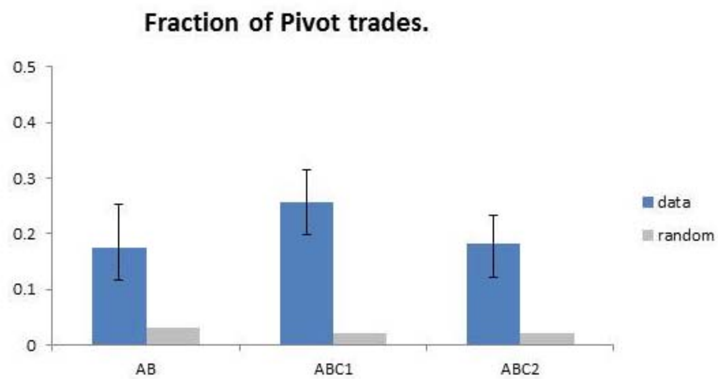


Figure 5. *Fraction of Pivot trades.*

The figure shows clearly the subjects' search for gains. With random trading, the frequency of payoff gains for both traders is 3 percent in AB and 1 percent in $ABC1$ and $ABC2$, or 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.

But if 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 small: 17 percent in AB, 26 percent in $ABC1$ and 18 percent in $ABC2$. Which other trades are subjects concluding?

4.4.2 Other trades.

We find that a much larger share of the data can be explained by extending the Pivot algorithms in one of two directions. First, while the Pivot algorithms select trades with strict gains in payoffs, in every treatment more than 40 percent of all trades result in no change in payoff for either trader. Zero-gain trades are trades involving non-pivotal votes, and thus preserving the status quo outcome; they could be the result of buying votes from allies with weak preferences, for example, or of buying losing votes, to strengthen one's favorite side's margin of victory. Pivot algorithms can be extended to weakly-improving trades without violating any rationality requirement.[30] The fraction of observed trades consistent with the model would then increase to 70 percent in AB and $ABC1$ and 58 percent in $ABC2$.[31] But our goal here is not to find support for the model, but to understand whether the zero-gain trades were intentional, and if so why.

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 our previous definition of a voter's *score* (at time t) as the product of the subject's number of votes and absolute valuation, summed over all proposals:

$$S_{it} = \sum_{k=1}^{K} |z_i^k| v_{it}^k$$

Note that the score reflects the voter's intensity of preferences and the number of

[30] Pivot trades are a subset of weak Pivot trades, and thus a Pivot stable allocation of votes is also reachable via weak Pivot trades. It follows that convergence to stability extends to weak Pivot trades under some constraint on the rules R through which trades are prioritized. For example, a rule R that executes first trades with strict payoff gains will reproduce the Pivot stable allocations reachable via strict Pivot algorithms; a rule R that allows infinite back-and-forth trades between two voters with identical preferences will not lead to convergence.

[31] And if the model is evaluated in terms of the fraction of trades weakly-improving for the indvidual making the trade (as opposed to the pair), then the support from the data is higher still: 84 percent in AB, 85 percent in $ABC1$, and 79 percent in $ABC2$.

votes held but remains unchanged whether the voter wins or loses any proposal. We call score-improving trades all trades that increase a subject's score. Trades may be score-improving but not payoff-improving (and hence Pivot trades) either because the proposals on which votes are traded continue to be lost or because they were already won. Such trades could reflect difficulties understanding pivotality, but could also mirror behavior that is more forward-looking than Pivot algorithms. Myopic gains are evaluated assuming voting occurred immediately. In fact, in the uncertain and complex environment of our experiment, subjects may want to accumulate votes on high value proposals, regardless of their resolution under immediate voting, because they conjecture that further trades are likely to take place before voting actually occurs.

Figure 6 shows, for each treatment, the fraction of binary trades consistent with Pivot trades (in dark blue), weak payoff increases for both traders (light blue), and score increases, again for both traders (in orange).[32]

Figure 6. *Types of trades.*

[32] The experimental matrices do not allow for weak score increases.

By construction, Pivot trades are a subset of both of the other two categories, and thus must explain a smaller fraction of observed trades. What is surprising is how much smaller. The figure shows that Pivot trades are of the order of one third of all weakly-payoff-improving trades in treatments A and $ABC2$, and about two fifths in treatment $ABC1$. Similar numbers apply to score-improving trades.

The frequency of different types of trades is informative, but what we need to understand is the intentionality of such trades. As we remarked about Figure 5, Pivot trades are not very frequent, but they appear intentional: they cannot be explained by random trading. Is that true of other types of trades?

Figure 7 plots, for the representative case of the AB treatment, the observed fractions of Pivot trades, zero-payoff change trades, and score-increasing-not-Pivot trades, together with the corresponding fractions under random trading, and the 95 percent confidence interval under the null hypothesis of random trading.

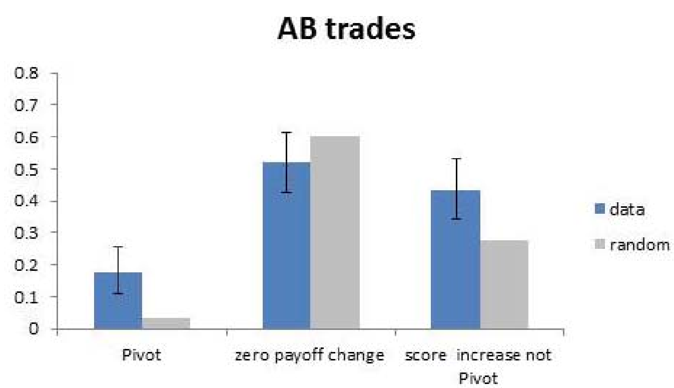


Figure 7. *AB trades by type v/s random.*

The figure makes clear that although the fraction of zero-payoff changing trades is large, we cannot rule out that it is the result of noisy trading: because all non-pivotal trades have zero effect on payoffs, for any given vote distribution a large share of feasible trades belongs to this class and thus is chosen under random trading. The figure does show, however, that this is not true for non-Pivot-score-increasing trades: the fraction observed in the data is significantly higher than under random trading ($p < 0.0001$).

We can make these observations more precise through a simple statistical model.

4.4.3 A simple statistical model

The model we discuss in this subsection is purely statistical, i.e. it aims not at explaining behavior but at classifying the types of trades, lending some rigor to the comments suggested by the figures. In line with the data just reported, we suppose that executed trades are selected according to four myopic criteria, synthetic summaries of the rules followed by the pairs of traders: (1) Pivot trades; (2) zero-payoff changing trades; (3) score-improving trades; (4) some other criterion we ignore, and such that the trade appears to us fully random. When executing a trade, each pair of traders follows one of these rules. Each trade can then be written in terms of the probability of following the four criteria: $probP$ for Pivot trading; $prob0$ for zero-payoff changing trades, $probS$ for score improving trades, and $probR$ for random trades. Call T_t the set of all trades feasible at t, where a trade is defined by a pair of traders, a pair of proposals, and the direction of trade. Similarly, call T_t^P the set of all feasible Pivot trades, T_t^0 the set of all feasible zero-payoff trades, and T_t^S the set of all feasible score-improving trades. Suppose that we observe a Pivot trade. The probability of such a trade equals $probP/|T_t^P| + probS/|T_t^S| + probR/|T_t|$. Similarly, the probability of a score-improving but not Pivot trade is given by $probS/|T_t^S| + probR/|T_t|$. Assuming that different trades are independent, the likelihood of observing the data set is simply the product of the probabilities of each trade. The probabilities $probP$, $prob0$, $probS$, and $probR$ can then be estimated immediately through maximum likelihood. The only challenge is that the sets of feasible trades, T_t, T_t^P, T_t^0, and T_t^S, all evolve over time, as budget constraints become binding, and the changes in vote allocations alter the payoff effects of different vote exchanges.[33]

We report our estimates in Table 5, together with the 95 percent confidence intervals.[34]

	AB		$ABC1$		$ABC2$	
$probP$	0.06	[0, 0.14]	0.19	[0.13, 0.25]	0.11	[0.05, 0.17]
$prob0$	0.11	[0, 0.23]	0.07	[0, 0.16]	0	[0, 0.10]
$probS$	0.41	[0.28, 0.55]	0.34	[0.25, 0.43]	0.39	[0.29, 0.49]
$probR$	0.42	[0.27, 0.57]	0.40	[0.29, 0.52]	0.50	[0.35, 0.59]

Table 5. *Model parameter estimates with 95% confidence intervals.*

[33] The unit of analysis is the trade itself, evaluated with respect to the set of feasible trades at t. It is the constantly changing set of feasible trades that determines the classification of the individual trade. Although the data were collected over multiple rounds, the changing set T_t, outside the control of any individual trader, makes the assumption of independence less problematic than in a standard set-up with individual decision-making and a small set of possible states. It radically simplifies an estimation procedure that is computationally quite demanding.

[34] We constructed the confidence intervals by bootstrapping the data and estimating the model's parameters 1000 times.

According to our statistical model, trade is very noisy and, as Figure 7 lead us to expect, there is no evidence of intentional zero-profit trades in any of the three treatments (in all treatments the 95 percent confidence interval for $prob0$ 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. Again as implied by the figures, $probP$ and $probS$ are not fully collinear and can be estimated separately.

4.4.4 Score-improving trades

Observing that a rising score can explain a substantial fraction of the experimental trades does not imply that the increase in score is the final objective pursued by our subjects. A first reason to be skeptical is the frequency of rejected trades reported in Section 4.1. Recall that about a third of all accepted bids are rejected by the bidder. Contrary to payoff changes, score increases do not depend on the identity of the trading partner: if score increases were the goal of the trades, they could be secured by the bidder and there would be no reason to reject any partner. In all treatments more than two thirds of the trades rejected by the bidder would have caused the bidder an increase in score.

A second cause for doubt comes from investigating whether subjects have indeed exploited all opportunities for score increases when trade comes to an end. We have defined stability as the absence of any feasible pairwise strictly payoff-improving trade. We can construct the similar concept of *score stability*, defined as the absence of any feasible pairwise score-improving trade, and enquire whether score stability is a useful characterization of final vote allocations.[35] Figure 8 plots the CDF's of minimal steps from score stability in the three treatments (in orange), together with the CDF of minimal steps to payoff stability (in blue).

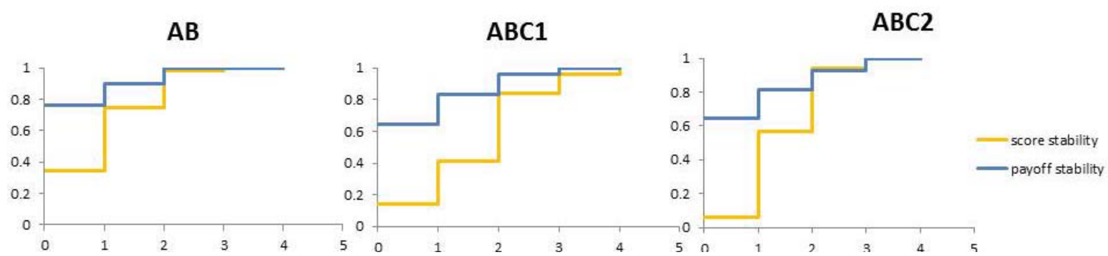


Figure 8. *Score and payoff stability.* Cumulative distribution functions.

Score stability is a much weaker explanation of final vote allocations than payoff stability: the fraction of score-stable final vote allocations is 34 percent in AB, 14 percent in $ABC1$, and 6 percent in $ABC2$; the corresponding numbers for payoff-stability are 76 percent, 64 percent, and again 64 percent. Not only does the orange CDF FOSD's the blue CDF (in AB and $ABC1$), but the gaps are large.

The important message from these numbers is that subjects appear to recognize payoff-stable vote allocations and tend to stop trading at that point, but they stop trading long before achieving maximal score improvements. Our conclusion is that subjects do not pursue score improvements for their own sake. Thus we conjecture that non-Pivot score-improving trades are unlikely to reflect primarily confusion about pivotality or payoffs, and more likely to result from some cautionary behavior

[35] A score-stable allocation always exists in pairwise trading. Indeed it is this property that leads to the convergence of the Pivot algoritms to payoff-stable vote allocations.

in front of uncertainty about future trades.

4.4.5 Are Subjects Farsighted?

Can the conjecture of forward-looking behavior be made more rigorous? The characteristics of votes trading for votes–a dynamic barter model in which others' trades affect both the feasibility and the desirability of one's own trades–make a fully strategic analysis a daunting prospect.[36] It is possible however to make some small progress by borrowing from cooperative games. Again, 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.[37] We can however adopt to our problem, and test on our data, some basic concepts from this literature.

We need three definitions:

Definition 4 *Given two vote allocations V and V', a pair of voters $D = \{i, j\}$ is said to be* effective *for (V, V') if $V' \in \mathcal{V}$ (V' is feasible) and $V'_s = V_s$ for all $s \neq i, j$.*

That is, voters i and j can move the vote allocation from V to V' by reallocating votes among themselves only.

Definition 5 *A* pairwise chain *from V to V' is a collection of vote allocations $V^1, V^2, ..V^m$, with $V^1 = V$ and $V^m = V'$, and a corresponding collection of effective pairs $D^2, .., D^m$ such that for all $t = 1, ..m-1$, D^{t+1} is effective for (V^t, V^{t+1}).*

[36] 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.

[37] See for example Chwe (1994), Mauleon et al. (2011), Ray and Vohra (2015), Dutta and Vohra (2015), and the references therein.

Finally:

Definition 6 *(Harsanyi, 1974) A a collection of vote allocations* $V^1, V^2, ..V^m$*, with* $V^1 = V$ *and* $V^m = V'$ *is a* pairwise farsighted chain *if it is a pairwise chain, and, in addition,* $u_j(V^t) < u_j(V')$ *for all* $j \in D^{t+1}$*. If there exists a pairwise farsighted chain from* V *to* V'*, then* V' *is said to* pairwise farsightedly dominate *(PF-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 PF-dominates V.[38] Other definitions are possible, and in general problems of existence are not trivial.[39] Developing a full analysis goes well beyond the scope of this paper, but our goal is much more limited: farsightedness builds on Harsanyi's notion of indirect dominance, defined above. If subjects in our experiment were 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 6 reports, for each treatment, the fraction of trades associated with farsighted gains for both traders (F-gains, in column 2), with farsighted losses (F-losses,

[38] Note the difference between this definition and the definition we used earlier. Myopic stability holds if there is no alternative vote allocation that a pair of voters can move to such that the pair would gain if voting occurred without further trades. Farsighted stability is much more demanding: V' can PF-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 beween the end point of the chain and the vote allocation at which trading occurs.

[39] A vote allocation that gives dictatorship power to a single voter is in the fairsghted 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. For questions of existence, in environments that differ from ours, see the discussions in the references cited above.

in column 3), and, for comparison, the fraction of Pivot trades (that is, trades associated with myopic gains, in column 4).

	F-gains	F-losses	Pivot
AB	5.2	2.6	17.4
$ABC1$	3.8	11.8	25.6
$ABC2$	6.3	14.3	18.3

Table 6. *Shares of trades yielding farsighted gains and losses, and share of Pivot trades.*

In all treatments, the fraction of trades with farsighted gains is less than 10 percent, and about a third of the fraction of Pivot trades; 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.[40]

The evidence of score-improving trades suggests that subjects gave some thought to the possible path of future trades, but standard notions of farsightedness adapted from recent approaches in cooperative game theory do not help explain the experimental data.

4.5 Outcomes

Which outcomes did the experimental subjects reach? Figure 9 plots the frequency of different outcomes observed over the full data (light blue), or restricting attention

[40] Note that the test here is weak: farsighted domination rests on farsighted chains, whose logic requires that *all* trades on a chain be farsighted.

to stable outcomes only (dark blue).

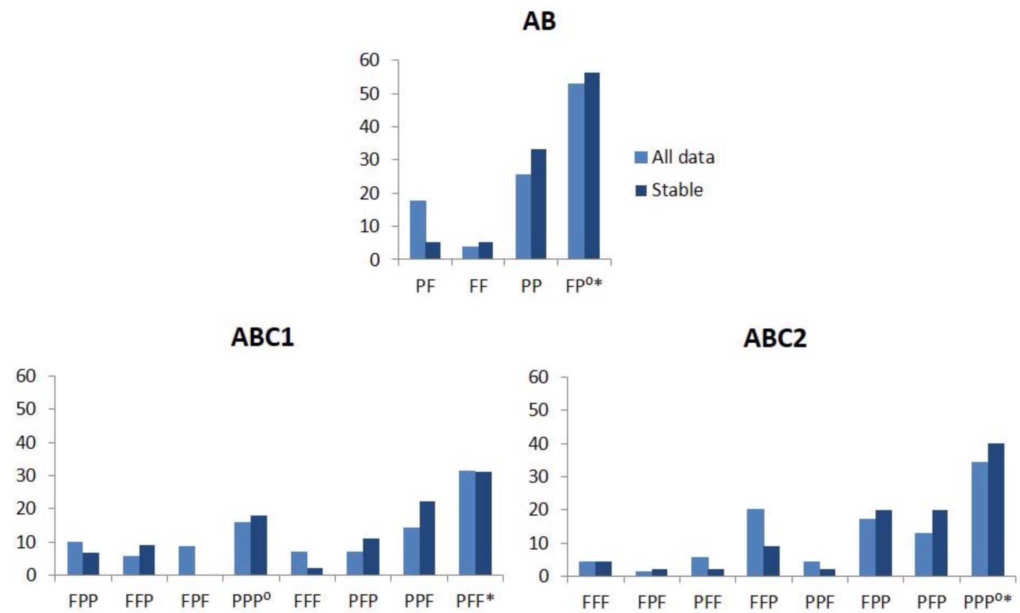


Figure 9. *Frequency of outcomes.* All data and stable allocations only.

Outcomes are ordered from lowest to highest aggregate payoff (thus the order is different in $ABC1$ and $ABC2$). A star indicates the Condorcet winner, and a dot the Pivot stable outcome. The Condorcet winner always corresponds to the no-trade outcome.[41]

The figure shows two immediate regularities. First, in all treatments, the Condorcet winner is the most frequent outcome, whether we consider all outcomes, or

[41] This is a well-known result when the initial allocation of votes grants to each voter one vote per proposal, as in our experiment. See Park (1967) and Kadane (1972).

stable outcomes only. Second, in all treatments, the frequency of outcomes correlates positively and significantly with aggregate payoffs. However, because both the Condorcet winner and aggregate payoffs also correlate perfectly with persistence of pre-trade outcomes[42], both results may reflect the inertia built into the market by the frequent zero-gain trades.

In terms of Pivot predictions, we see a higher frequency of the Condorcet winner, relative to the second most frequent outcome, in treatments AB and $ABC2$, where the Condorcet winner is Pivot-stable. And among stable outcomes we see a small spike in the frequency of outcome $\{A, B, C\}$ in treatment $ABC1$ where it is Pivot-stable, relatively to the outcome's low payoff-rank. On the whole, however, the clean predictions on outcomes derived from the Pivot algorithm are not evident in the data. Since final vote allocations, on the contrary, are in line with theory, the divergence of outcomes from predictions is surprising: outcomes are the automatic result of vote allocations. The divergence highlights the high sensitivity of outcomes to noise–contrary to good markets, one subject's missed trading opportunity (and thus a small deviation of final vote allocations from the theory) affects the final result of voting for all.

As shown by Figure 10 , the outcomes we observe are consistent with the trades' characteristics highlighted by the statistical model. The figure reports the frequency of different outcomes in the data (considering here all final vote allocations, whether stable or unstable) and, in columns denoted by diagonal stripes, in 5,000 trading simulations in which, given the vote allocation, a trade is selected randomly, following

[42] In our matrices, the fewer the changes in the resolution of the different issues, the higher the aggregate payoff.

the estimated probabilities in Table 5. As in all simulations in the paper, at each time interval the probability of a trade occurring is calculated so as to replicate, on average, the observed number of trades in the treatment. The model simulations match the ordinal ranks of the different outcomes' frequencies, although they consistently overestimate the frequency of the Condorcet winner. Such overestimation, however, is mostly mechanical: the result of the relatively high probability of random trades, and the likelihood that such random trades leave outcomes unchanged. Because zero-gain trades result in non-Pivotal vote allocations, they make Pivot-trades impossible, and thus bias the simulations towards pre-trade outcomes and the Condorcet winner.

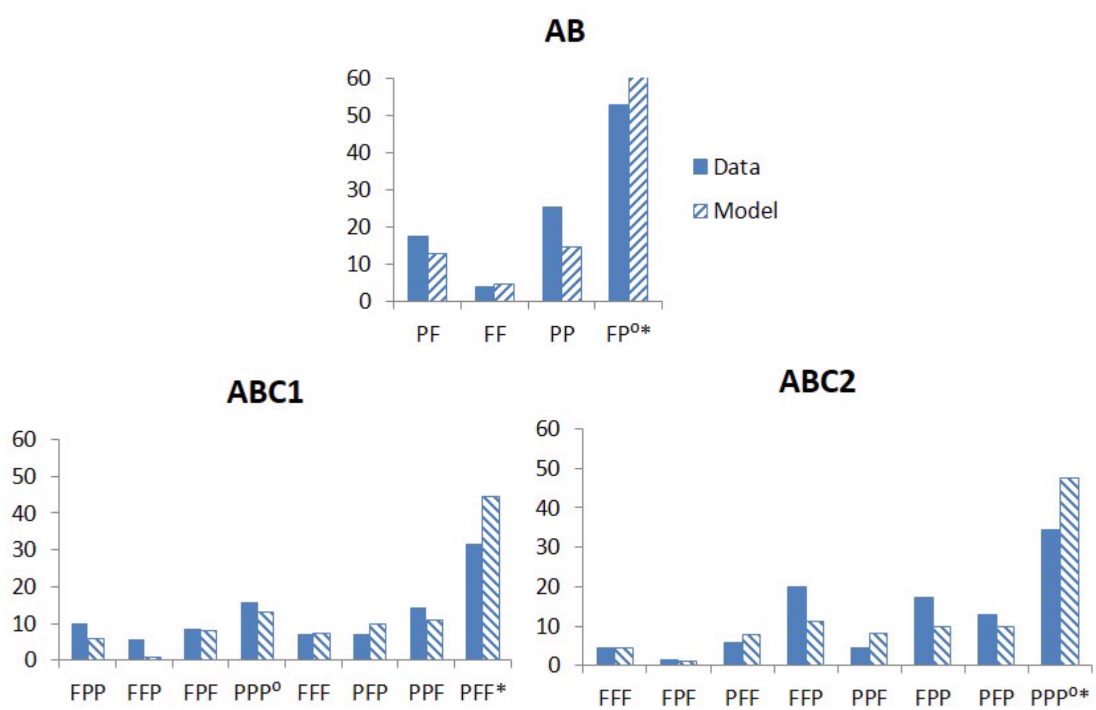


Figure 10: *Simulated outcomes v/s data.*

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) a notion of stability; and (2) a rational vote trading process. Stable vote allocations are those for which there are no strictly 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, the stability concept is useful in organizing the experimental data. Overall, two-thirds of all final vote allocations in the experiment are stable, and more than eighty percent are at most one trade away from stability. Second, final vote allocations are in line with the theory, but final proposal outcomes show a clear bias towards the pre-trade outcome. In vote trading environments, there is great scope for path dependence. A single deviation from predicted trading behavior can have large impacts on proposal outcomes, because subsequent trades are easily triggered (or inhibited) by the current trade. In particular, trades that accumulate votes on the side that is already winning make pivotal trades impossible and consolidate the pre-trade outcome.

Third, an analysis of trade-by-trade data lends weaker support to the theoretical trading process itself. We classify trades though a simple statistical model and find that when noise is accounted for most trades are score-improving, but not necessarily payoff-improving–the trades that are posited by the theoretical trading process. Score-improving trades are vote exchanges in which each voter trades a vote on a less important issue in exchange for a vote on a more important issue, without necessarily benefiting from the trade (either because the outcome does not change, or because the directions of preferences are such that one voter suffers a loss). Because such trades coexist with the subjects' ability to recognize stable vote allocations–i.e. to stop trading in the absence of further opportunities for payoff increases, we conjecture that they may be precautionary more than irrational, suggesting the possibility of some farsighted behavior. However, and this is our fourth result, rational farsighted trading behavior is unambiguously rejected by our data: on average, a trade is twice as likely to leave the traders worse off in the final outcome as it is to make them better off.

This study only scratches the surface of possibilities for laboratory studies of vote trading and logrolling. There are many interesting environments that are not represented by the three that are studied in the paper. First, a Condorcet winner exists for all three environments in this study, but we know that more generally Condorcet winners 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 to organize the market. For example, one could allow communication to take place either concurrently with or prior to the actual trading protocol. This might make it easier for voters to identify beneficial trading partners. In the current trading scheme, voters who offer a trade might have to reject a trading partner, which leads to delays and leaves room for accidental trades. Other extensions of the trading process would 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 vote hoarding, whereby voters acquire extra *non-pivotal* votes on high-salience issues, is indicative of precautionary incentives to trade for votes, so as to guarantee passage or failure of those issues. Understanding such precautionary motives requires modeling the strategic uncertainty faced by vote traders - uncertainty about trades that future voters might engage in and allowing for risk aversion. 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] Buchanan, J.M. and G. Tullock, 1962, *The Calculus of Consent.* Ann Arbor: University of Michigan Press.

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

[6] Casella, A. and T. Palfrey, 2017, "Trading Votes for Votes. A Dynamic Theory." *Working Paper.* Columbia University.

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

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

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

[10] Diamantoudi, E., E. Miyagawa, and L. Xue, 2004, "Random Paths to Stability in the Roommate Problem", *Games and Economic Behavior*, 48, 18-28.

[11] Dutta, B. and R. Vohra, 2015, "Rational Expectations and Farsighted Stability," *Working Paper*, Brown University.

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

[13] Feldman, A., 1973, "Bilateral Trading Processes, Pairwise Optimality, and Pareto Optimality", *Review of Economic Studies* 40: 463-73 .

[14] Ferejohn, J, 1974, "Sour notes on the theory of vote trading", *Social Science Working Paper No.41*, California Institute of Technology.

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

[16] Green, J. 1974, "The Stability of Edgeworth's Recontracting Process," *Econometrica*, 42, 21-34.

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

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

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

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

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

[22] 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.

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

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

[25] 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.

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

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

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

[29] Miller, N., 1977a, "Graph-Theoretical Approaches to the Theory of Voting", *American Journal of Political Science*, 769-803.

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

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

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

[33] 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.

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

[35] Niederle, M. and L.Yariv, 2011, "Decentralized Matching with Aligned Preferences", unpublished.

[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] Roth A. and J. Vande Vate, 1990, "Random Paths to Stability in Two-Sided Matching, *Econometrica*, 58, 1475-1480.

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

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

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

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

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

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