Trading Votes for Votes.*

A Decentralized Matching Algorithm.

Preliminary and Incomplete

Alessandra Casella† Thomas Palfrey‡

March 10, 2015

Abstract

Vote-trading is a central component of decision-making in groups. Yet we know very little about its properties. Inspired by the similarity between the logic of sequential rounds of vote-trading and matching problems, we borrow three central questions from the matching literature: (1) Does a stable allocation of votes always exists? (2) Is it reachable through a decentralized algorithm? (3) What welfare properties does it possess? We find that if vote-tradng is pair-wise, a stable allocation is always reached in a finite number of trades, for any number of voters and issues, for any separable preferences, and for any rule on how trade are prioritized. Its welfare properties however are guaranteed to be positive only in special cases. A laboratory experiment confirms that stability has predictive power on the vote allocation achieved via trades.

JEL Classification: C62, C72, D70, D72, P16

Keywords: Voting, Majority Voting, Vote Trading, Condorcet Winner, Matching, Matching Algorithm, Assignment Problems

*We thank Enrico Zanardo, Kirill Pogorelskiy and Manuel Puente for their assistance, and participants to the workshop on "Institutions, Individual Behavior and Economic Outcomes" held in Alghero, Italy, on June 23-July 1, 2013, for their comments. We thank the National Science Foundation for its financial support through grants SES-0617820, SES-0617934, and SES-0962802. Part of the research was conducted while Casella was a Straus Fellow at NYU Law School. The hospitality and financial support of the Straus Institute are acknowledged with gratitude.

†Columbia University, NBER and CEPR, ac186@columbia.edu

‡Division of Humanities and Social Sciences, Caltech, trp@hss.caltech.edu

1 Introduction

Trading support for one proposal in exchange for someone else's support of a different proposal is a common aspect of voting. Whether as exchanges of favors in small informal committees or as more elaborate deals in legislatures, common sense, anecdotes, and systematic evidence, all suggest that the practice is a central component of decision-making in groups.[1] Yet we know very little about the properties of vote trading. Efforts at a theory were numerous and enthusiastic in the 1960's and 70's but have fizzled and disappeared in the last 40 years. John Ferejohn's words in 1974: "[W]e really know very little theoretically about vote trading. We cannot be sure about when it will occur, or how often, or what sort of bargains will be made. We don't know if it has any desirable normative or efficiency properties" (Ferejohn, 1974, p. 25) remain true today.

One reason for the lack of progress is that the problem is difficult: not only does vote bartering occur without the equilibrating properties of a fully flexible price, not only does it cause externalities to allies and opponents of the trading parties, but each exchange triggers new profitable exchanges. If we think of the trades sequentially, as a subset of voters trade votes on some proposals, the default outcomes of these proposals change, generating incentives for a new round of vote trades, which will again change outcomes and trigger new trades. What is the most productive approach to a question of this type? This paper is inspired by the similarity between the logic of sequential rounds of vote trading and the problem of achieving stability in sequential rounds of matching among different agents, as originally proposed by Gale and Shapley (1962).[2] We borrow from the matching literature the idea of an algorithm through which a sequence of decentralized myopic pairings are realized, and in particular we borrow the notion of stability: an allocation of votes such that no welfare-improving trade exists. We ask whether a stable allocation of votes exists, whether the specific algorithm we construct converges to a stable allocation, and whether we can say anything about individuals' preferences over the outcomes induced by stable vote allocations.

We should note at the outset that the parallel to matching problems is imperfect. The closest analogue is to a one-sided matching problem with exter-

[1] Stratmann (1992) provides evidence consistent with vote trading in agricultural bills in the US Congress. (Add references)

[2] Roth and Sotomayor (1990) and Gusfield and Irving (1989) survey some of the main results, from the perspective of economic theory and computer science, respectively. The research has remained very active. See for example, Kojima and Pathak (2009), Kojima (2011),....

nalities: a single group of individuals who match in pairs but such that everyone has preferences not only over his own partner, but over the composition of all matches. Here too there is only one group–voters–and in principle everyone can match with everyone else,[3] and preferences are defined over the full set of matches. But in addition in our voting problem preferences evolve endogenously in response to executed trades. Trades by others can reverse a voter's status as winner or loser and affect his desire to trade, as well as his attractiveness as trading partner.

The committee we study is formed by an odd number of voters and faces several proposals, each of which may pass or fail and, after trade, is voted upon separately through majority voting. Every committee member can be in favor or opposed to any proposal and attaches some cardinal value to his preferred direction prevailing. Members' preferences are separable across proposals. A vote trade is a physical exchange of ballots. For most of our analysis and in the experiment, we allow pair-wise trading only: two voters engage in a trade if one delivers his vote to the other on one proposal, in exchange for the other's vote on a different proposal. As in the matching literature, a trading pair is said to *block* a given allocation of votes if it can be better-off under an alternative allocation that is in the pair's power to achieve, keeping fixed the votes held by the other committee members. A stable allocation is then an allocation of votes that cannot be blocked, i.e. such that no pair-wise improving trade exists.

As remarked in Riker and Brams (1973), the requirement that a trade be welfare improving for the two traders implies that the traders must be pivotal, and we call our algorithm the *Pivot Algorithm.* Our first result is that the Pivot algorithm *always* generates a stable vote allocation in a finite number of steps, for any number of voters, any number of proposals, and any configuration of (separable) preferences.

This is an interesting result, not only for its generality but also because stability–convergence to a vote allocation such that no further vote trade is profitable–was one of the two central questions of the early literature on vote trading. The literature addressed the ambitious conjecture that vote trading may offer the solution to majority cycles in the absence of a Condorcet winner. The original analysis (Park (1967)), studied non-binding agreements when voters vote on the full package of proposals (as opposed to voting separately on each proposal). Park considered only majority coalitions and concluded that

[3] In contrast to two-sided matching problems where matches only occur between members of two separate groups–men and women, students and schools, workers and firms.

the process can converge only if a Condorcet winner exists.[4] Riker and Brams (1974) and Ferejohn (1974) simplified the problem by considering binding agreements and proposal-by-proposal voting, as in our model. Their conclusions are ambiguous: Riker and Brams conjecture that even if stability held for pair-wise trading it would be compromised by allowing trades among larger coalitions of voters; Ferejohn suggests that a stable vote allocation may not hold even for pair-wise trading if voters are forward-looking, but does not fully specify the game structure.

The Pivot algorithm through which we model vote trades implies that voters are myopic, and if pair-wise trades only are allowed, a stable vote allocation always exists. In the second part of the paper, we allow vote-trading among coalitions of voters of any sizes. Our results confirm Riker and Brams' conjecture. With coalition-trading stability cannot be guaranteed: we construct an example with well-behaved preferences where a cycle develops and trading need never end. In addition, in our model, with proposal-by-proposal voting and binding trades, there is no logical connection between coalitional stability in vote-trading and the existence of a Condorcet winner. A coalition-stable vote allocation may exist in the absence of a Condorcet winner, and may differ from the Condorcet winner when the Condorcet winner exists.

The second conjecture at the core of the interest in vote trading in the 60's and 70's concerned not the existence of stable vote allocations but their welfare properties. It held that vote trading leads to Pareto superior outcomes because it allows the expression of the intensity of preferences. The conjecture stemmed from an early debate between Gordon Tullock and Anthony Downs[5] and was stated explicitly in Buchanan and Tullock (1965).[6] As a general result, the claim was rejected by Riker and Brams' (1973) influential "paradox of vote trading": Riker and Brams showed that if vote trading is pair-wise and binding, there are non-pathological preferences such that each pair of voters individually gains from vote trading and yet everyone strictly prefers the no-trade outcome. Opposite examples where vote trading is Pareto superior to no-trade can easily be constructed too[7], and the literature eventually ran dry with the tentative conclusion that no general statement on the desirability of vote trading can be made.

[4]The result was later echoed by other studies, for example Berholz (1973), Koehler (1975), Schwartz (1981).

[5]Tullock, 1959 and 1961, Downs, 1957, 1961.

[6]See also Coleman (1966), Haefele (1970), Tullock (1970), and Wilson (1969).

[7]For example, Schwartz (1975).

Our algorithm leads us to the same conclusion, although we reach some unexpected results in special cases. In particular, when the committee is faced with only two proposals (and thus, since each proposal can either pass or fail, four possible outcomes), then for any number of voters and any (separable) preferences, the outcome associated with all stable vote allocations must be unique, is always Pareto optimal, is the Condorcet winner, if a Condorcet winner exists, and must be preferred by the majority to the no-trade outcome if it differs from it. These results hold whether trade is restricted to pairs of voters or coalitions are allowed. They are surprising because it has always been understood that vote trades' ambiguous welfare properties are due to the externalities inherent in the exchanges. But externalities are clearly present in the two-proposal case, and yet the algorithm delivers an outcome with desirable welfare properties.

Approaching vote trading through a mechanical algorithm allowed us to make some progress by avoiding the difficulties of a strategic model. We have chosen it, however, for a second reason too: we conjectured that it may have predictive power. In the second part of the paper, we test the Pivot algorithm in the laboratory. The barter nature of the task, and thus the lack of a common unit of exchange, the changing profitability of trades in response to others' trades, the role of pivotality, all make the experiment unusually complex.[8] For this reason, we limit trades in the laboratory to pair-wise trades.

We study three treatments, corresponding to three sets of cardinal values for each voter over each proposal. All treatments have five voters, but differ in the number of proposals (two in treatment A, and three in treatments $M1$ and $M2$), and in the prediction of the Pivot algorithm. The Condorcet winner exists in all three cases; it coincides with the unique stable outcome reachable through the Pivot algorithm in treatments A and $M2$; it differs in treatment $M1$.

Our experiment reaches several results. First, we find that stability is a useful predictive tool. In all treatments, two thirds or more of the vote allocations reached by experimental subjects are stable. Second, the dynamic data showing the evolution of the votes allocations during the trading periods indicate reasonably fast convergence towards allocations that if not fully stable are in close proximity of stability, as measured by the number of further trades necessary to reach stability. The graphs seem consistent with the willful search for a stable vote allocation. Indeed, and this is our third result, final vote allocations are

[8] To our knowledge, barter experiments are rare. Ledyard, Porter and Rangel (1994) is a very interesting example and makes clear the challenges of the design and the data analysis.

well-predicted by the Pivot algorithm, at least qualitatively. Across all treatments, across all voters, across all proposals, in every single case in which the stable allocation reachable via the Pivot algorithm reflects a net purchase of votes, or a net sale, we observe it in the data. Fourth, very few trades are associated with myopic losses. However, a large fraction are associated with no strict gains–purchases of votes from weak allies, or purchases of losing votes. There is no logical reason why these trades should be ruled out, and to explain the data the algorithm must include them. Including zero-gain trades, however, expands the set of stable outcomes. We find some weak evidence that different frequencies of outcomes between the $M1$ and the $M2$ treatments are consistent with the different theoretical predictions. However, observed outcomes suggest that for each proposal the resolution preferred by the majority in the status quo pre-trade is a strong attraction point. [Note: Our analysis of the experimental results is not completed].

To our knowledge, the only experimental paper studying vote trading is McKelvey and Ordeshook (1980). The paper reports results from a large series of experiments, all done face-to-face and under different creative protocols designed to allow either pair-wise only or coalitional trades, and either binding or non-binding agreements. The methodologies are different enough to make a direct comparison of results difficult, and McKelvey and Ordeshook's focus on alternative cooperative solution concepts has no counterpart in our experiment. Closer to our computerized experimental protocols are recent experiments on decentralized matching, in particular Echenique and Yariv (2012).[9] In that work, as in ours, a central finding is the extent to which the experimental subjects succeed in reaching a far-from-apparent set of stable matches. The set-up however differs substantially, even within the perspective of matching theory–a two-sided matching problem with no externalities and fixed preferences, in Echenique and Yariv; a one-sided matching problem with externalities and evolving preferences in our case. In addition, of course, the substantive questions we ask are specific to vote-trading.

The paper proceeds as follows. The next section presents our model and derives the model's theoretical predictions; section 3 discusses the experimental design; section 4 reports the experimental results, and section 5 concludes.

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

Longer proofs are collected in the Appendix 1, and the instructions from a representative experimental session are in Appendix 2.

2 The Model

Consider a committee composed of N odd voters who must approve or reject each of K independent binary proposals. Committee members have separable preferences summarized by a set of cardinal values Z, where z_i^k is the value attached by member i to the approval of proposal P^k, or the utility i experiences if P^k passes. Value z_i^k is positive if i is in favor of P^k and negative if i is opposed. Proposals are voted upon one-by-one, and each proposal P^k is decided through simple majority voting.

Before voting takes place, committee members can trade votes. The initial allocation of votes assigns one vote to each committee member over each proposal. We think of votes as physical ballots, specialized by proposal–for example, imagine ballots of different colors for different proposals. A vote *trade* is an actual exchange of ballots, with no enforcement or credibility problem, where by exchange we mean that each trader must give away and receive at least one vote. After trading, a voter may own zero votes over some proposals and several over others, but cannot hold negative votes. We call v_i^k the votes held by voter i over proposal P^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 profile (or *allocation*) of vote holdings over all voters and proposals. $\mathcal{V}$ denotes the set of 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 $v_i^k \in V$.[10] The initial allocation of votes equal $v_{i0}^k = 1$ for all k, or $V_0 = \{\mathbf{1}, \mathbf{1}, .., \mathbf{1}\}$, where $\mathbf{1}$ indicates a unitary vector of dimension K.

Given a feasible vote allocation V, we assume that at the time of voting, voters who attach positive value to a proposal cast all votes they own over that proposal in its favor, and voters who oppose it cast all available votes against it. We indicate by $\mathbf{P}(V)$ 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

[10] 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.

occurs at V: $u_i(V) = \sum_{k \in \mathbf{P}(V)} z_i^k$.

Our focus is on the existence and properties of vote allocations that hold no incentives for further trading. In line with the matching literature, a pair of voters i, i' is said to block V if there exists a feasible vote allocation $\widehat{V} \in V$ such that $\widehat{V}_j = V_j$ for all $j \neq i, i'$, and $u_i(\widehat{V}) > u_i(V)$, $u_{i'}(\widehat{V}) > u_{i'}(V)$. We can then define:

Definition. *We say that the allocation $V \in \mathcal{V}$ is* pair-wise stable *if there exists no pair of voters i, i' who can block V.*

We can remark immediately that a feasible allocation of votes that yields dictator power to a single voter is trivially pair-wise 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.[11] Hence:

Remark. *A pair-wise stable vote allocation V exists for all Z, N, and K.*

2.1 Dynamic adjustment.

Pair-wise stable allocations exist, but are they reachable through sequential decentralized trades? To answer the question, we need to specify the dynamic process through which bilateral trades take place. Our focus is on simple myopic algorithms.

We begin with the following definition:

Definition. *A pair-wise welfare improving trade is* minimal *if eliminating any vote from the exchange causes at least one trader a strict myopic loss.*

Recall that a trade involves the exchange of at least one vote on each side.[12] Concentrating on minimal trades allows to "unbundle" complex trades into elementary trades. For any individual voter, 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

[11] Other examples are easy to construct. For example, any allocation such that for all P^k, $v_{i_k}^k \geq \frac{n+1}{2}$, where i_k is the voter who, among all, attributes highest value to winning issue k, is pair-wise stable.

[12] Thus surrendering a single vote with no myopic utility change is not a trade. All trades involving a single vote exchange on each side are minimal.

earlier analyses, we define the *Pivot Algorithms* as sequences of pair-wise trades yielding myopic strict gains to both traders:

The Pivot Algorithms: *Start from any vote allocation* V_0*. If there is no minimal pair-wise strictly welfare-improving trade, stop. If there is one such trade, execute it. If there are multiple, choose one according to 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 (fully random); or give priority to trades with higher total gains; or to trades involving specific voters. The class of Pivot algorithms corresponds to the class of possible R rules, and individual algorithms differ in the specification of rule R. At this stage, it is not necessary to be more specific about R.

Pivot trades are not restricted to two proposals only: a voter can trade his vote, or votes, on one issue in exchange for other voters' vote(s) on more than one issue. The only constraint is the requirement that trades be minimal: zero-utility trades cannot be bundled with welfare improving trades. If a trade is welfare improving and minimal, it is a legitimate trade under Pivot.

A crucial property was anticipated by Riker and Brams and gives the name to our algorithm:

Lemma 1. (Riker and Brams)**.** *Under the Pivot algorithms, all votes transferred must be pivotal.*

Proof. Immediate from the requirement of minimal trades and the definition of Pivot algorithms.

2.2 Existence of stable vote allocations

The question we want to ask is whether a stable vote allocation is reachable through the Pivot algorithms. From here onward, we maintain $V_0 = \{\mathbf{1}, \mathbf{1}, ..\mathbf{1}\}$. We define:

Definition. *An allocation of votes* V *is* Pivot-stable *and is denoted by* $V_T(R)$ *if it is stable and reachable through a Pivot algorithm in a finite number of steps, following rule* R*.*

Does a Pivot-stable allocation always exist? Surprisingly, the answer is clear-cut and positive. Pivot-stable vote allocations always exist, for the entire class of

Pivot algorithms, independently of the rules R through which competing claims to trade are resolved. We can state:

Theorem. *For all K, N, Z, a Pivot-stable allocation of votes exists for all R.*

Proof. Consider trades dictated by the Pivot algorithm. If trade occurs at V_0 it can only concern proposals that at V_0 are decided by minimal majority. But by minimality of trade, it then follows that the same proposals must still be decided by minimal majority in any subsequent votes allocation V_t, with $t > 0$. But since $V_0 = \{\mathbf{1}, \mathbf{1}, \mathbf{1}, ..\}$, no more than one vote is ever traded on any given proposal (although trades could involve bundles of proposals). Now consider voter i with values Z_i and absolute values $|Z_i| \equiv X_i$. We call i's *score* at step t the function $\sigma_{it}(X_i, V_{it})$ defined by:

$$\sigma_{it} = \sum_{k=1}^{K} x_i^k v_{it}^k$$

where x_i^k is the (absolute) value i attaches to each proposal k, and v_{it}^k is the number of votes i holds on that proposal at t. If i does not trade at t, then $\sigma_{it+1} = \sigma_{it}$. If i does trade, then, by the argument above, i's vote allocation must fall by one vote on some proposals $\{k, k', ..\}$ that i was winning and increase by one vote on some other proposals $\{\widetilde{k}, \widetilde{k}', ..\}$ that i was losing. Call the first set of proposals $P_{i,t}$ and the second $P_{-i,t}$. Note that although the two sets may have different numerosity, by definition of pair-wise improving trade, $\sum_{k \in P_{i,t}} x_i^k < \sum_{k \in P_{-i,t}} x_i^k$ and, since a single vote is traded on each proposal, $\sum_{k \in P_{i,t}} x_i^k v_{it}^k < \sum_{k \in P_{-i,t}} x_i^k v_{it+1}^k$. Hence if i trades at t, $\sigma_{it+1} > \sigma_{it}$: for all i, $\sigma_{it}(X_i, V_{it})$ must be non-decreasing in t. At any t, either there is no trade and the Pivot-stable allocation has been reached, or there is trade, and thus there are two voters i and i' for which $\sigma_{it+1} > \sigma_{it}$ and $\sigma_{i't+1} > \sigma_{i't}$. But $\sigma_{it}(X_i, V_{it})$ is bounded above and the number of voters is finite. Hence trade must stop in finite steps: a Pivot-stable allocation of votes always exists.[13] Note that we have made no assumptions on R, the rule through which trades are selected when multiple are possible. A Pivot-stable allocation of votes exists for any R.$\square$

The generality of the result is surprising: a Pivot-stable allocation always

[13] It is not difficult to find the upper boundary on the number of trades needed to reach a Pivot-stable allocation. It equals the maximum number of trades that could shift all individuals' votes to their respective highest-value proposal, or $\left[\frac{K(K-1)}{2}\right]\left[\frac{(N-1)}{2}\right]$.

exists, regardless of the number of voters and proposals, for all (separable) preferences, and regardless of the order in which different possible trades are chosen. As we said, the parallel to the matching literature is imperfect, and indeed no such result can be found there. In one-sided matching problems, it is well-known that a stable match may not exist.[14] When it does exist, it is not the case that any sequence of decentralized myopic matchings will converge to a stable matching. Cycles are possible. If preferences are strict, one converging sequence of matchings always exists, but if matchings are decentralized, guaranteeing convergence requires some randomness in the selection of blocking pairs: a random rule assigning a positive probability of selection to any blocking pair.[15] The difficulty of achieving stability is increased by the presence of externalities. We are not aware of comparable results for one-sided matching problems with externalities. In two-sided matching problems, guaranteeing the existence of a stable match in the presence of externalities requires a very stringent definition of blocking.[16] In our problem, the score function we have defined above is not subject to cycles. Because it is always non-decreasing in t, convergence to a stable allocation of votes is guaranteed for any selection rule among blocking pairs.

2.3 Preferences over stable outcomes

Definition. *An outcome* $\mathbf{P}(V)$ *is Pivot-stable if it is achieved from a Pivot-stable allocation of votes.*

We denote by $\mathcal{P}(V_T(R))$ the set of all stable outcomes reachable with positive probability through a Pivot algorithm with rule R. What are the welfare properties of $\mathcal{P}(V_T(R))$? We have modeled vote trading through an algorithm, and our institution-free approach demands a welfare evaluation that is equally institution-free. We ask whether outcomes in $\mathcal{P}(V_T(R)$ must belong to the Pareto set; whether they must include the Condorcet winner, if one exists, and more generally whether they can be ranked, in terms of majority preferences, relative to the no-trade outcome.

Our first set of answers is unexpectedly positive. Because we characterize

[14] Gale and Shapley (1962).

[15] Diamantoudi et al. (2004). The result that randomness in the selection of the blocking pair induces convergence builds on Roth and Vande Vate (1990), who established it in the case of two-sided matching.

[16] Sasaki and Toda (1996). Two individuals block an existing match if they strictly gain from matching with one another under any possible rematching by all others.

results that hold for all R, we can use the simpler notation $\mathcal{P}(V_T)$ with element $\mathbf{P}(V_T)$. We can show:

Proposition 1. *If $K = 2$, then, for all N, Z, and R: (1) $\mathbf{P}(V_T)$ is unique.*[17] *(2) $\mathbf{P}(V_T)$ is Pareto optimal. (3) If a Condorcet winner exists, then $\mathbf{P}(V_T)$ is the Condorcet winner. (4) $\mathbf{P}(V_T)$ can never be the Condorcet loser. (5) If $\mathbf{P}(V_T) \neq \mathbf{P}(V_0)$, then a majority prefers $\mathbf{P}(V_T)$ to $\mathbf{P}(V_0)$.*

Proof in the Appendix.

Proposition 1 is interesting because it highlights that the lack of Pareto optimality in vote trading examples, in particular Riker and Brams' paradox of vote trading, is not an immediate result of voting externalities. Externalities are not eliminated when $K = 2$, and yet the outcome of the Pivot algorithm (the same myopic vote-trading rule studied by Riker and Brams) is always Pareto optimal. Similarly, when $K = 2$ vote trading performs well in terms of majority preferences.[18]

In fact, there is another scenario in which, for all values Z, the Pivot-stable outcome is related to majority preferences:

Proposition 2. *If $N = 3$, then for all K, Z, and R: (1) if a Condorcet winner exists, $\mathbf{P}(V_T)$ is unique and is the Condorcet winner.(2) $\mathbf{P}(V_T)$ can never be the Condorcet loser.*

Proof in the Appendix.

The intuition behind Proposition 2 is straightforward: with three voters, any Pivot trade between a pair reflects the majority's preferences. We know from Park (1967) and Kadane's (1972) that if there is a Condorcet winner, it can only be the no-trade outcome[19], and thus with three voters there cannot be any trade under the Pivot algorithm. If the Condorcet loser exists, all proposals must be decided in the direction favored by the minority, but since any trade reflects the majority's preferences, such an outcome is impossible to reach.

The results from Pivot trading are less predictable in the general case:

Proposition 3. *If $K > 2$ and $N > 3$, then: (1) there exist Z such that for any R no outcome in the Pareto set is Pivot-stable. (2) There exist Z such that the Condorcet winner exists, but for any R it is not Pivot-stable.*

[17] Note that uniquess of $\mathbf{P}(V_T)$ does not imply uniqueness of V_T.

[18] Possibly, but not necessarily also in terms of total utilitarian welfare. In a finite electorate, results on utilitarian welfare depend on the distributions from which values are drawn.

[19] Because the majority must prefer the no-trade outcome to any outcome that differs from no-trade in the resolution of a single issue.

Proof. We prove the two statements by example, and because the examples are simple and instructive, we report them here in detail. (1) Consider the following example, with $K = 5$ and $N = 5$:

	1	2	3	4	5
A	10	-1	-1	-1	-1
B	-1	10	-1	-1	-1
C	-1	-1	10	-1	-1
D	-1	-1	-1	10	-1
E	-1	-1	-1	-1	10

Each row is a proposal (A, B, C, D, and E) and each column a voter ($1, 2, 3, 4$, and 5). Each cell $\{k, i\}$ reports z_i^k, the value attached by voter i to proposal k passing. No voter is pivotal, and thus V_0 cannot be blocked. The unique stable outcome is $\mathbf{P}(V_T) = \mathbf{P}(V_0) = \{\varnothing\}$, all proposals fail. Yet, all proposals failing is not Pareto optimal: it is Pareto-dominated by all proposals passing. The example suggests the importance of allowing for trades among coalitions of more than two voters, a point to which we return below.

(2) Consider the following example, with $K = 3$ and $N = 5$:

	1	2	3	4	5
A	4	-7	1	-1	4
B	1	1	-4	4	-1
C	-3	4	2	-2	2

$\mathbf{P}(V_0) = \{P^1, P^2, P^3\}$ is the Condorcet winner but this example has a unique Pivot-stable outcome $\mathbf{P}(V_T) = \{P^1\}$. It is not difficult to verify that there are three possible trade chains but all stop at $\mathbf{P}(V_T) = \{P^1\}$. Indicating in order the voters engaged in the trade, the proposals on which they trade votes, and, in parenthesis, the outcome corresponding to that allocation of votes, we can describe the three chains as $\{\{13BC(P^1)$, $45BC(P^1P^2P^3)$, $23AB(P^3)$, $45AC(P^1)\}$, $\{23AB(P^3)$, $45AB(P^1P^2P^3)$, $13BC(P^1)\}$, and $\{23AB(P^3)$, $45AC(P^1)$, $45BC(P^1P^2P^3)$, $13BC(P^1)\}\}$. □

Note that result (2) in Proposition 3 immediately implies:

Corollary to Proposition 3. *If $K > 2$ and $N > 3$, then there exist Z such that for all R, $\mathbf{P}(V_T) \neq \mathbf{P}(V_0)$, but $\mathbf{P}(V_0)$ is majority preferred to $\mathbf{P}(V_T)$.*

The first example is an immediate implication of Lemma 1: under the Pivot algorithm, trade can occur only between pivotal voters. If the vote allocation does not correspond to minimal majority, no pivotal voters exist. Thus the status quo is Pivot-stable, and delivers the unique Pivot-stable outcome; if such an outcome is Pareto-inferior, then the stable outcome does not belong to the Pareto set.

The second example is more unexpected. Why does the positive result with $K = 2$ not extend to a larger number of issues? Intuitively, the problem is that previous trades over some issues k and k' can make it impossible for a pair of voters to execute a different, desired trade over k' and k''. Thus, contrary to the $K = 2$ case, the Pivot algorithm does not allow voters to exploit all opportunities for mutual agreements.

2.4 Coalitional Trades

A natural aspect of vote trading is the possibility of forming coalitions, indeed the incentive to do so. The experiment we describe below focuses on pair-wise trades, but our approach can be extended to the study of coalitions and sheds some light on the debates in the early literature on vote-trading. In this subsection, we derive two main results. First, the stability highlighted by our theorem on pair-wise trades does not generalize to vote-trading within larger coalitions of voters. Second, in our model, there is no logical connection between stability under coalitional trade and existence of the Condorcet winner.

We begin by redefining stability in the presence of coalitions. A coalition of voters $C = \{i, i', i'', ..\}$ is said to block V if there exists a feasible vote allocation $\widehat{V} \in \mathcal{V}$ such that $\widehat{V}_j = V_j$ for all $j \notin C$, and $u_i(\widehat{V}) > u_i(V)$ for all $i \in C$. The allocation $V \in \mathcal{V}$ is *coalition-stable* if there exists no coalition of voters C who can block V. Up to now, we have restricted C to be of size 2; here we allow C to have any size between 2 and N.

As in the case of pair-wise trades, the first observation is that a coalition-stable allocation always exists: a feasible allocation of votes that gives decision power to a single voter over all proposals remains trivially stable because no coalition that excludes the dictator can change the outcome, and the dictator cannot strictly gain from participating in any coalition. The interesting question is not whether a stable allocation exists, but rather whether it can be reached through the relevant extensions of our algorithm to coalitional trades.

To extend the algorithm, the previous definitions need to be amended. Keeping in mind that a vote trade must, by definition, include at least two voters and at least two issues, we define:

Definition. *A coalition-improving trade is* minimal *if it concerns: (1) the minimal package of votes such that all members of the coalition weakly gain from the trade; and (2) the minimal number of members such that the outcome corresponding to* $\widehat{V}$ *can be achieved.*

The C-Pivot Algorithms: *Start from the initial vote allocation* $V_0 = \{\mathbf{1}, \mathbf{1}, .., \mathbf{1}\}$*. If there is no minimal strictly coalition-improving trade, stop. If there is one such trade, execute it. If there are multiple, choose one according to rule* R_C*. Continue in this fashion until no further pair-wise improving trade exists.*

The *C-Pivot Algorithm* is the natural generalization of the Pivot Algorithm to coalitions. Note again that coalitions can be of any size and we have imposed no rule selecting among them in the order of trades, when several coalition-improving trades are possible. We call *C-Pivot stable* an allocation of votes that cannot be blocked by any coalition and is reachable via the C-Pivot algorithm in a finite number of steps, and define a *C-Pivot stable* outcome as an outcome $\mathbf{P}^C(V_T)$ that corresponds to a C-Pivot stable allocation of votes.

In the presence of coalitions, stability becomes a more elusive goal:

Proposition 4. *There exist* K*,* N*,* Z*, and* R_C *such the C-Pivot algorithm never reaches a stable vote allocation.*

Proof. The proposition is proved by example in the Appendix.

In the presence of coalitions, stability may fail because trades can be profitable for the coalition even when the pair-wise trades that are part of the overall exchange are not: coalition members benefits from the positive externalities that originate from the trades of other members. As a result, the score function defined in the proof of the Theorem in section 3 is no longer monotonically increasing in the number of trades, and the logic of that proof does not extend to coalition trades.

But if a C-Pivot stable allocation exists, does it have desirable welfare properties? A first, positive remark is immediate:

Remark. *if a C-Pivot stable outcome exists, then it cannot be Pareto dominated.*

Regardless of the history of previous trades, if an outcome is Pareto dominated, then the coalition of the whole can always reach the Pareto superior outcome. But then the allocation corresponding to the Pareto dominated outcomes cannot be C-Pivot stable.[20]

But further results are more ambiguous:

Proposition 5. *Consider Z such that the Condorcet winner exists. (1) If either $K = 2$ or $N = 3$, then the C-Pivot stable outcome always exists, is unique, and coincides with the Condorcet winner. (2) If $K > 2$ and $N > 3$, if a C-Pivot stable outcome exists, it need not coincide with the Condorcet winner.*

Proof. (1) See the Appendix. (2) Consider Example 2, in the proof of Proposition 3. There are two C-stable outcomes: $P(V_T) = \{P^A\}$, reached through pair-wise trades as described earlier, and $\mathbf{P}(V_T) = \{P^A, P^B\}$, if we allow $|C| > 2$. The second stable outcome is reached thought the following trades: after voters 2 and 3 have traded votes on A and B, a coalition of voters 1, 4, and 5 is formed; 4 gives an A vote to 1; 5 gives a B vote to 4, and a C vote to 1. Note that the trade is minimal, and the resulting vote allocation cannot be blocked. Hence $\mathbf{P}(V_T) = \{P^A, P^B\}$. Neither outcome is the Condorcet winner.□

Proposition 5 is interesting because it clarifies that the existence of the Condorcet winner and the existence and properties of C-stable outcomes are logically independent. In some cases, ($K = 2$ or $N = 3$), the two must coincide; in others ($K > 2$ and $N > 3$), the existence of the Condorcet winner gives no information about the existence and welfare properties of C-stable outcomes. The result is driven by two central assumptions of our model: (1) vote trades are binding, and (2) voting occurs proposal-by-proposal.

[The public posting of retractable bids during the experiment amounts to some form of public communication. Whether this could be enough to coordinate on coalitions is a question that deserves more thought. We will address it in the next version of the paper.]

3 The Experiment

The experiment was run at the Columbia Experimental Laboratory for the Social Sciences (CELSS) in November 2014, with Columbia University students recruited from the whole campus through the laboratory's ORSEE site. No subject participated in more than one session. After entering the computer

[20] Note that if the coalition trade is not minimal, it can be made so by eliminating redundant trades or traders.

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.

Because the design of the trading platform presents some challenges, we describe it here is some detail. A copy of the instructions for a representative session is included in Appendix 2.

At the start of each treatment, a subject saw on his computer screen the matrix of values, denominated in experimental points, and the vote allocation. To help intuition, the two alternatives for each issue–Pass or Fail–were identified with two colors–Orange or Blue–, and each individual's values were written in the color of the individual's preferred alternative. 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 bid by clicking the offer and highlighting it. Figure 1 reproduces two of the screenshots showed during instructions: the screen of the subject making a bid (ID 1), and the screen of a subject accepting the bid (ID 3).

Figure 1. Screenshots for a subject posting a bid (on the left), and for a subject accepting a posted bid (on the right).

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. In the example shown in the figure, both 1 and 3 would be trading pivotal votes, and thus the vote balance would change on both A and B: if voting took place just after the trade, A would be won by Orange and B by Blue, with the result that 1's payoff would fall by 100 points, while 3's would increase by 200. Note that if 1's bid had been accepted by 4, no change in outcomes would result from the trade (because 1 and 4 have identical preferences), and, if voting occurred, neither trader would experience any change in payoff. If instead the bid had been accepted by 5, then the majority would prefer Orange in both issues, and thus, if voting took place just after the trade, 1 would gain 500 points, and 5 lose 100.

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 is accepted, a window appears 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 (Figure 2). If the trade is rejected, a message appears on the screen of the rejected trade partner, informing him of the rejection.

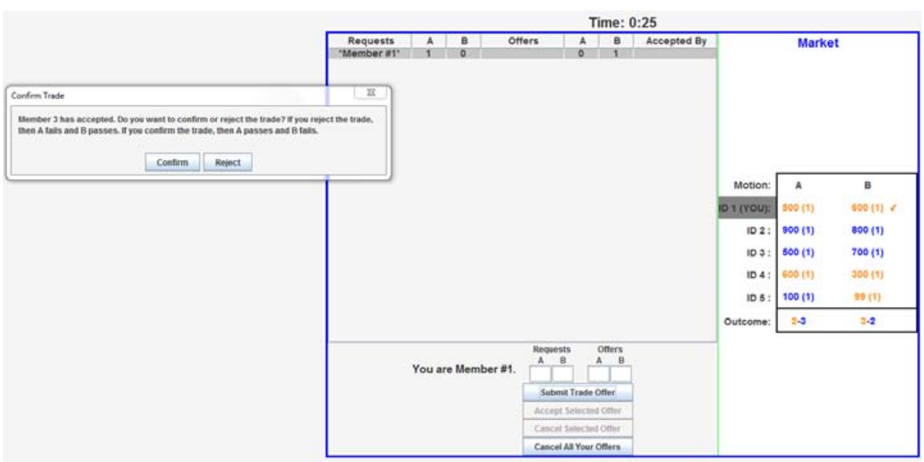

Figure 2. Confirmation request for the bidder.

After a trade was concluded, the vote tally on each issue was updated and conveyed to all subjects via a specific message on all screens. The message also reported the post-trade voting outcome if voting were to occur immediately. Note that the value matrix and the vote holdings were always present on the screen.

The market was open for three minutes.[21] However, in a market where each concluded trade can trigger a new chain of desired trades, it is particularly important to ensure that all desired trades have the time to be executed. Thus the time limit was automatically extended by 10 seconds whenever a new trade occurred within 10 seconds of the expected closure.

Only trades of a single vote on one issue against a single vote on a different issue were allowed, again to limit the complexity of the task. 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, A, $M1$, and $M2$, 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 A, and $K = 3$ in treatments $M1$, and $M2$. In each committee, subjects were identified by ID's randomly assigned by the computer, and issues were denoted by A and B (in treatment A), and A, B and C (in treatments $M1$ and $M2$). Each session started with two practice rounds; then three rounds of treatment A, and then five rounds each of $M1$ and $M2$, alternating the order.[22] We did not alternate the order of treatment A 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 A, versus five for $M1$ and $M2$).

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.

[21] Two minues in treatment A, with two issues only.

[22] Two of the sessions had only two treatments: A and $M1$ in one case, and A and $M2$ in the other.

All but one sessions consisted of 15 subjects, divided into three committees of five subjects.[23] At the end of each session, subjects were paid their cumulative earnings from all rounds, 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 $34.[24]

We designed the three treatments according to the following criteria. First, we wanted a $K = 2$ treatment, both as further training for the subjects and because of the sharp theoretical predictions of the Pivot algorithm in this case. Second, we chose value matrices for which the stable outcome reachable via Pivot trades is unique but requires multiple trades. In A, the path to stability is unique, while in both $M1$ and $M2$ the Pivot stable outcome can be reached via multiple paths, with no path being clearly focal. Third, we chose matrices such that not only is the Pivot stable outcome unique, but the stable vote allocation reached via Pivot trades is unique, even with multiple possible trading paths. Fourth, we designed matrices for which the Condorcet winner exists, but need not correspond to the Pivot stable outcome: it does in A (by necessity–see Proposition 1), and in $M2$ (by construction), but not in $M1$. Finally, we wanted $M1$ and $M2$ to be superficially very similar and to have predicted trading paths of similar multiplicity and length, allowing us to test whether the different force of attraction of the Condorcet winner predicted by the theory is reflected in the data. 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. Our theoretical results hold for all R.

The three value matrices are the following:

A

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

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

[24] We used the Multistage Game software package developed jointly between the SSEL and CASSEL labs. This open-source software can be downloaded from http://software.ssel.caltech.edu/

M1

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

M2

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

In all three cases, the initial vote allocation V_0 is unstable. In the case of matrix A, $\mathbf{P} = \{P_B\}$ is the Condorcet winner and the unique Pivot-stable outcome. The Pivot algorithm follows a unique path, of length two (i.e. consists of a sequence of two trades). Matrix $M1$ has identical properties to the matrix of values discussed in the proof of Proposition 3. The Condorcet winner exists and corresponds to $\mathbf{P} = \{P_A\}$, but the unique Pivot-stable outcome is $\mathbf{P} = \{P_A, P_B, P_C\}$. In matrix $M2$, the Condorcet winner is $\mathbf{P} = \{P_A, P_B, P_C\}$, and corresponds to the unique Pivot stable outcome. With both matrices $M1$ and $M2$, the Pivot algorithm can follow three different paths, and for both matrices two of these paths have length four, and one has length three.[25]

Table 1 reports the experimental design.

Session	Treatments	# Subjects	# Groups	# Rounds
s1	$A, M1, M2$	10	2	3,5,5
s2	$A, M2, M1$	15	3	3,5,5
s3	$A, M1, M2$	15	3	3,5,5
s4	$A, M2, M1$	15	3	3,5,5
s5	$A, M2$	15	3	3,5
s6	$A, M1$	15	3	3,5

Table 1: Experimental Design[26]

4 Experimental Results.

4.1 Trading

How much trading did we see? Table 2 reports basic statistics on observed trades. The unit of analysis is the group per round.

[25] The possible paths are detailed in the Appendix.

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

Treatment	Tot trades	groups × rounds	Mean trades	Median	s.d	Max	Pivot
A	115	51	2.25	2	1.92	13	2
$M1$	211	70	3.0	3	1.67	9	3,3,4
$M2$	175	70	2.5	2	1.36	7	3,3,4

Table 2. Number of trades.

A histogram of the number of trades per treatment (Figure 3) shows clearly the higher frequency of few trades in the A treatment, with $K = 2$. Between the two $K = 3$ treatments, $M2$ has consistently higher fractions of low trades, but the differences are not striking–56 percent of rounds end with two or fewer trades in $M2$, as opposed to 41 percent in $M1$, and 80 percent end with three or fewer trades in $M2$, as opposed to 76 percent in $M1$. In all treatments, few rounds include five or more trades.

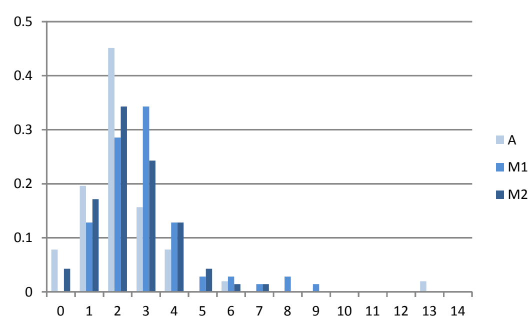

Figure 3. Number of trades. Frequencies

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 $M1$ or $M2$ 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 4 shows the CDF of steps to stability for the three treatments, in blue, as well as in 1,000 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.

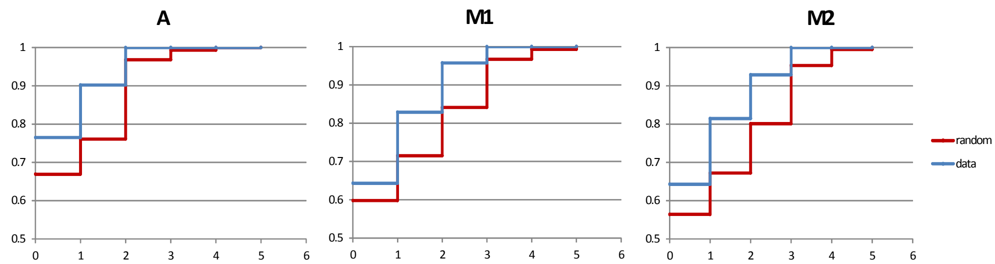

Figure 4. Steps to stability. Cumulative distribution functions.

The fraction of stable vote allocations in the experimental data was 76 percent in A, and 64 percent in both treatments $M1$ and $M2$. 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 A 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 provide 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, a partner, and a direction of trade, all with equal probability, and enacting the trade as long as both traders' budget constraints are 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 length of rounds and the average number of trades in the treatment. For each treatment, we repeated the procedure 1,000 times, each time focusing on a group.

Figure 4 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 5 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 dot corresponds to a trade. Thus, for any given dot, the horizontal axis indicates when the trade took place, within the maximal interval of 250 seconds for each round.[27] The vertical axis measures distance from stability, defined, as in Figure 4, by the minimal number of Pivot trades necessary to reach a stable allocation. Such number is calculated first for the vote allocation characterizing 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 colored curve reports data from the same round (1-3 for A and 1-5 for $M1$ and $M2$). The jumps between dots are relatively small because a trade concerns a single group, while the others' vote allocations remain unchanged.

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 1000 simulations with random trading. In each simulation, 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 period in the treatment. If a trade occurs, as described above, it is selected randomly among those not violating the budget constraint. After the trade has taken place, the average steps to stability for each vote allocation are calculated and averaged over all groups. The curve reports the result when the process is repeated 1,000 times.

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.[28] Notice also the lack of learning in the data–there is no systematic difference between earlier and later rounds.

[27] The trading period lasts 180 seconds, but there is a 10 second delay with each trade, to give time to subjects to study the new vote allocations. In addition, 10 more seconds are added at the end of the period if any trade takes place in the last 10 seconds. Trading never lasted more than 250 seconds.

[28] With the exception of two trades in round 5 in $M2$.

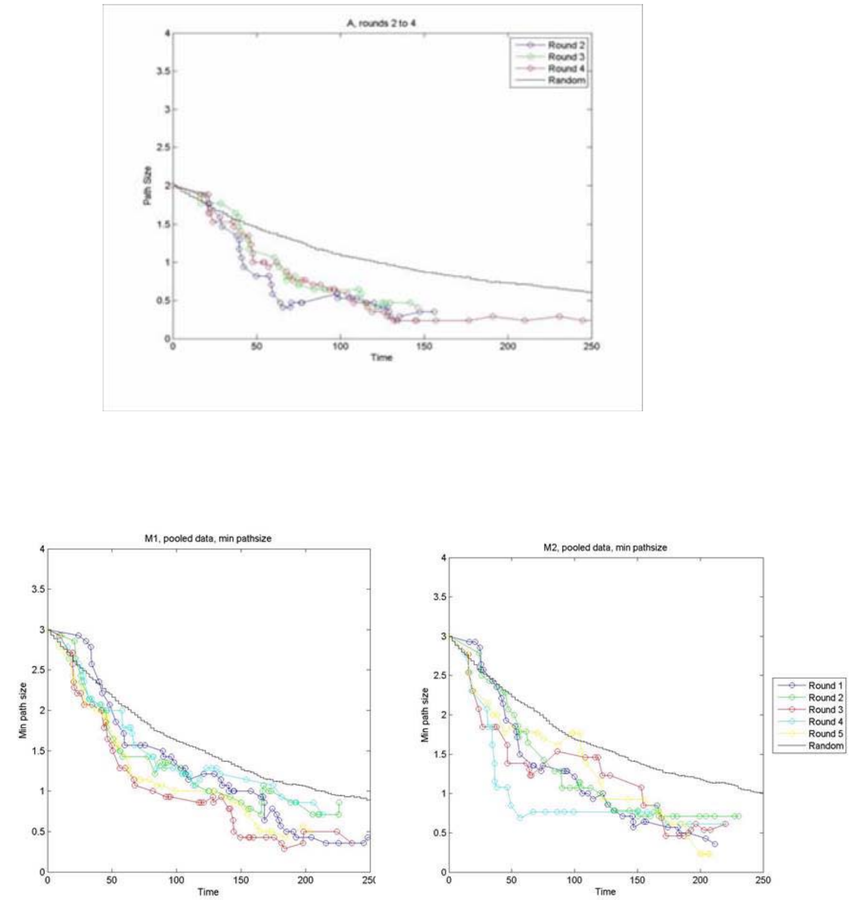

Figure 5. Path to stability

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 6 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 vote allocations only.

It is clear from the figure that the vote distribution in the data is less sharply variable across issues than theory predicts, as we would expect in the presence of both zero-gain trades and noise. Yet, the qualitative predictions are strongly supported. There are five voters in each treatment, holding votes over two (in A) or three issues (in $M1$ and $M2$)–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.[29]

[29] The theory predicts that voter 3 in treatment $M1$ should hold three votes.

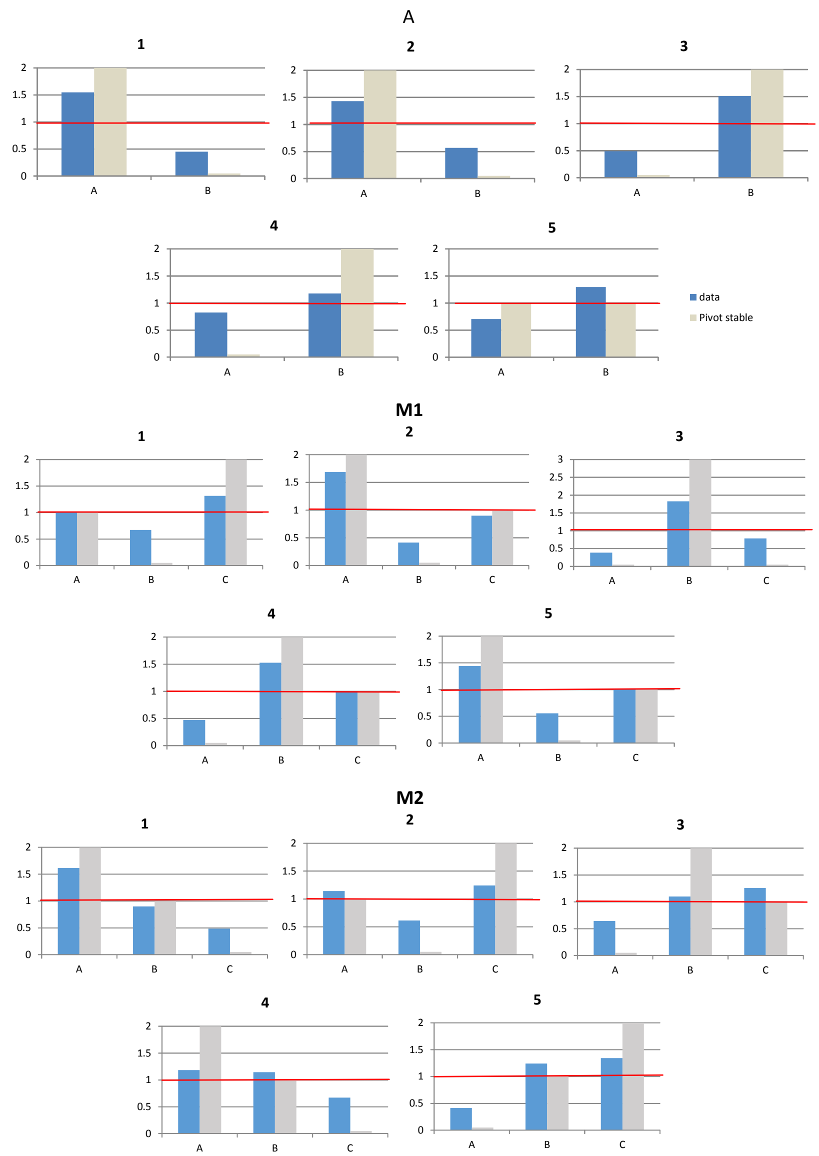

Figure 6. Average vote allocations at the end of each round by voter id.

4.4 Trades

Does the Pivot algorithm explain the data? Figure 7 shows, for each treatment, the fraction of trades associated with myopic strict increases in payoff, considering first individual trades (in panel a), and then binary trades (in panel b). Panel (b) thus corresponds to Pivot trades: both partners strictly benefit from the trade (myopically). The blue columns correspond to the experimental data, the red columns to the simulations with random trading, and the error bars indicate 95 percent confidence intervals (under the null of random trading).[30]

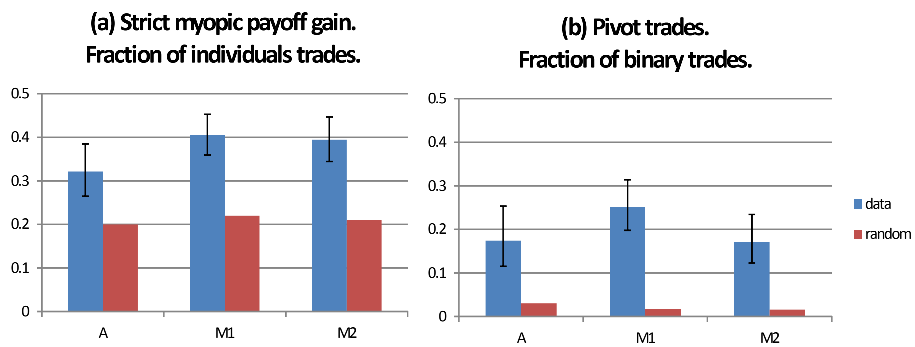

Figure 7. Myopic payoff gains in individual (panel a) and binary (panel b) trades, as fraction of all trades.

Relative to random trading, individual trades resulting in myopic strict gains are twice as frequent in the data in $M1$ and $M2$, and more than 50 percent more frequent in treatment A (panel (a)). In all three treatments, the difference is highly significant ($p < 0.0001$).

The Pivot algorithm is formulated for pair-wise trades, and it is on pair-wise trades that it should be tested. In panel (b), the subjects' search for gains is even clearer. With random trading, the frequency of strict payoff gains for both traders is 3 percent in A and 1 percent in $M1$ and $M2$, or less than one fifth of what we observe in A, and less than one tenth in $M1$ and $M2$. 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 A and $M2$, and 25 percent in $M1$. Which other trades

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

are subjects concluding? Figure 8 reports, for each treatment, the fraction of individual trades that leave the traders' payoff unchanged or lower.

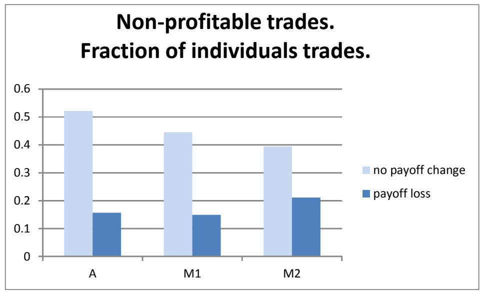

Figure 8. Non-profitable individual trades.

The figure makes clear that a large number of trades has no immediate impact on payoffs. In every treatment, more than 40 percent of all individual trades result in no change in payoff, and in every treatment this is more than the fraction of strictly-improving trades. 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. Note that no (myopic) rationality requirement is violated by trades that are only weakly-improving, either for one or both traders. Empirically, they are an important share of observed trades; theoretically, we see no reason to rule them out. Thus we need to take a short theoretical detour to evaluate how the theory changes if we expand the definition of the Pivot algorithms to include trades with zero payoff changes.

4.5 Weak Pivot

We can define the class of *weak-Pivot algorithms*, executing at each stage a minimal pair-wise weakly improving trade. We adopt the term *weak-Pivot* for obvious analogy to the Pivot algorithms. But note that when weak improvements are allowed, as mentioned above, traded votes need not be pivotal. The only requirement is that the trade be minimal.

The first question is whether the theorem in section 2.2 extends to weak trades: is a stable vote allocations always reached in a finite number of steps via weak-Pivot? Trades dictated by the Pivot algorithm are a strict subset of weak-Pivot trades. Thus if Pivot-stable allocations always exist, such allocations are also reachable through weak-Pivot trades. However, they are not reachable for all rules R: for example a rule that always selects back-and-forth exchanges between two traders with identical preferences generates a trivial but not-ending cycle.[31] On the other hand, consider a rule that first prioritizes strictly pairwise improving trades and then selects among them according to some R. Call such a rule R'. We know that weak-Pivot trades with rule R' must converge to a stable allocation. Hence we can state a weaker result:

Remark. *For all K, N, Z, a weak-Pivot-stable allocation of votes exists for some R.*

A specific type of R rule is full randomization: at any vote allocation V_t, if multiple feasible trades exist, the algorithm selects any one of them with equal probability. Because Pivot-stable vote allocations always exist, full randomization guarantees that weak-Pivot trades converge to a stable allocation with probability one.[32] Thus full randomization belongs to the set of rules to which the Remark above applies.

If a weak-Pivot stable outcome exists, what can we say about its welfare properties?

Under R', weak-Pivot trades not only lead to stability, but yield stable outcomes identical to those reached under strict Pivot and rule R. In particular, under such rule all the results in Propositions 1-3 extend to weak-Pivot stable outcomes. However, there also exist rules R^w such that stable outcomes reachable under weak-Pivot are not reachable for any R under strict Pivot. The effect on the welfare evaluation of vote trading is ambiguous:

Proposition 6. *For all K, N, Z such that the Condorcet winner exists: (1) there always exists $R^w(Z)$ such that the Condorcet winner is weak-Pivot stable. (2) If V_0 is not stable, there always exists $R^{w'}(Z)$ and a weak-Pivot stable outcome $\mathbf{P}^w(V_T(R^{w'}))$ that differs from the Condorcet winner.*

Proof. In the Appendix.

Allowing for weak-Pivot trades (and arbitrary selection rule R) expands the set of reachable stable outcomes. Relative to our experiment, the Condorcet winner becomes reachable with matrix $M1$, and other stable outcomes besides

[31] Note that such trades are minimal and legitimate under the weak-Pivot algorithm. The possibility of such a cycle always exists for any $N \geq 5$ since with five or move voters there must always be at least two who agree on any pair of issues.

[32] The reason is that any cycle is broken with probability one, and any strictly-improving trade takes the algorith up one step towards stability. The logic is identical to Roth and Vande Vate (1990).

the Condorcet winner become reachable with A and $M2$. Note also that weak-Pivot trades have weaker predictions on the number of trades necessary to reach stability. As shown in the proof of Proposition 6, a stable outcome can always be reached with a single weak-Pivot trade. On the other hand, if Z is such that at least two voters agree on the desired resolution of two issues, as indeed is the case in all three of our matrices, then the weak Pivot algorithm sets no limit to the number of zero-gain trades between them, and thus no limit on the number of trades occurring before a stable outcome is reached, if it ever is.

4.6 Outcomes

Which outcomes did the experimental subjects reach? Figure 9 plots the frequency of different outcomes observed over the full data, or restricting attention to stable outcomes only.

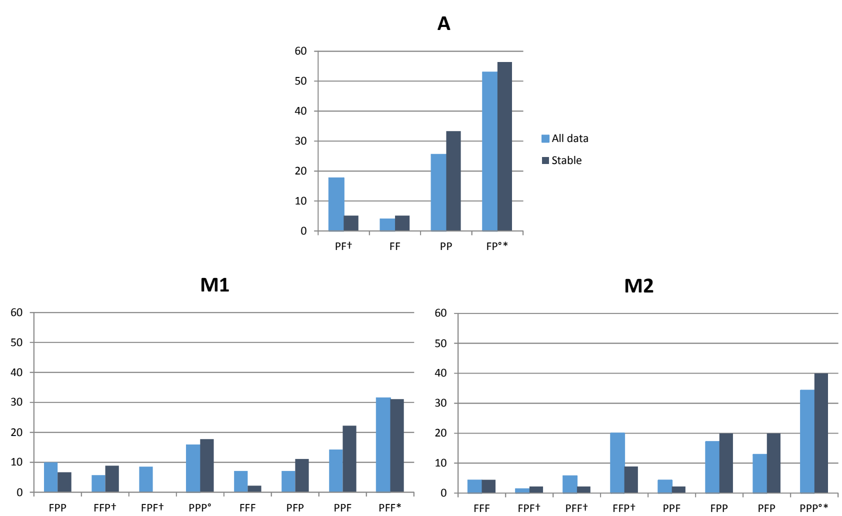

Figure 9. Frequency of outcomes.

Outcomes are ordered from lowest to highest aggregate payoff (thus the order is different in $M1$ and $M2$). A star indicates the Condorcet winner, and a dot the Pivot stable outcome. Both such outcomes are also weak-Pivot stable; other possible weak-Pivot stable outcomes are indicated by a dagger.

The figure shows two immediate regularities. First, in all treatments, the Condorcet winner is the most frequent outcome, whether we consider all outcomes, or stable outcomes only. Second, in all treatments, the frequency of outcomes correlates positively and significantly with aggregate payoffs. However, because aggregate payoffs also correlate perfectly with persistence of pre-trade outcomes–the fewer the changes in the resolution of the different issues, the higher the aggregate payoff– 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 A and $M2$, where the Condorcet winner is Pivot-stable. And among stable outcomes we see a higher frequency of $\{P^A, P^B, P^C\}$ in treatment $M1$, where it is Pivot-stable. On the whole, however, the clean predictions on outcomes derived from the Pivot algorithm, either in its strict or in its weak form, are not evident in the data. One possible reason is that, contrary to a goods market, the outcomes of vote-trading are sensitive to individual mistakes–one subject's missed trading opportunity affects the final result of voting for all.

To be continued.

5 Conclusions

Appendix 1

Proposition 1. *If $K = 2$, then, for all N, Z, and R: (1)* $\mathbf{P}(V_T)$ is unique. *(2) $\mathbf{P}(V_T)$ is Pareto optimal. (3) If a Condorcet winner exists, then $\mathbf{P}(V_T)$ is the Condorcet winner. (4) $\mathbf{P}(V_T)$ can never be the Condorcet loser. (5) If $\mathbf{P}(V_T) \neq \mathbf{P}(V_0)$, then a majority prefers $\mathbf{P}(V_T)$* to *$\mathbf{P}(V_0)$.*

Proof. (1). At any step t, if there is a unique blocking pair, then V_{t+1} is determined uniquely and to any unique vote allocation corresponds a unique outcome. Multiplicity can arise only if at any step there are multiple blocking pairs (possibly with some voters belonging to more than one pair). But if $K = 2$ any trade by any of the blocking pairs must induce the identical change in outcome. Thus at any step the outcome of the vote is unique. And since a Pivot-stable outcome exists, it must be unique. Note that the Pivot-stable vote allocation need not be unique.

(2) With $K = 2$, there are four possible outcomes: $\{\{\varnothing\}, \{P^1\}, \{P^2\}, \{P^1, P^2\}\}$. Suppose, with no loss of generality, that $\mathbf{P}(V_0) = \{P^1, P^2\}$. With $K = 2$, $\mathbf{P}(V_T) \in \{\{P^1, P^2\}, \{\varnothing\}\}$. Majority voting at V_0 implies that at least one trader

must rank $\mathbf{P}(V_0)$ strictly above all other outcomes. But if trade occurs and V_0 is blocked, there must be at least two voters who strictly prefer $\{\varnothing\}$ to $\mathbf{P}(V_0)$. One of those voters must have preferences: $\{P^2\} \succ \{\varnothing\} \succ \{P^1, P^2\} \succ \{P^1\}$, and the other $\{P^1\} \succ \{\varnothing\} \succ \{P^1, P^2\} \succ \{P^2\}$. Thus, if V_0 is blocked, no outcome is Pareto-dominated and all are Pareto-optimal. If V_0 is not blocked, $\mathbf{P}(V_0) = \mathbf{P}(V_T)$ by definition, and again it is Pareto-optimal.

(3) Again suppose, with no loss, that $\mathbf{P}(V_0) = \{P^1, P^2\}$. There are four possible preference types, which we call $T_{1,2}, T_\varnothing, T_1$, and T_2, with rankings as in Table 1 below. Preferences are transitive and the outcome in any cell is preferred to the outcome(s) below it; two outcomes are in the same cell when either could be ranked above the other.

$T_{1,2}$	$T_\varnothing$	T_1	T_2
P^1, P^2	$\varnothing$	P^1	P^2
P^2 P^1	P^2 P^1	$\varnothing$ P^1, P^2	$\varnothing$ P^1, P^2
$\varnothing$	P^1, P^2	P^2	P^1

Table 1. $K = 2$, possible preference types.

Call $n_{1,2}$ the number of voters of type $T_{1,2}$, and similarly for the other types. With $\mathbf{P}(V_0) = \{P^1, P^2\}$, both proposals are supported by a majority of voters. Call M_1 (M_2) the number of voters who prefer P^1 (P^2) to pass rather than fail. Then $M_1 = n_{1,2}+n_1 \geq (N+1)/2$, and $M_2 = n_{1,2}+n_2 \geq (N+1)/2$. We begin by asking whether any of the four possible outcomes could be the Condorcet winner. With $\mathbf{P}(V_0) = \{P^1, P^2\}$, it follows immediately that $\{P^1, P^2\}$ is preferred by the majority to $\{P^1\}$ and to $\{P^2\}$–more generally the no-trade outcome must be majority preferred to any outcome that differs in the direction in which a single proposal is decided. This is Park (1967) and Kadane's (1972) result: if there is a Condorcet winner, it can only be the no-trade outcome. Here that is $\mathbf{P}(V_0) = \{P^1, P^2\}$. Suppose then that $\mathbf{P}(V_0)$ is the Condorcet winner.

If either $M_1 > (N+1)/2$ or $M_2 > (N+1)/2$, no voter is pivotal, V_0 cannot be blocked, and $\mathbf{P}(V_T)$ trivially equals $\mathbf{P}(V_0)$, the Condorcet winner. Hence the proposition holds. The more interesting case is when $M_1 = M_2 = (N+1)/2$, V_0 is blocked, and trade takes place. Notice that in such a case $(N+1)/2 = n_{1,2}+n_1 = n_{1,2}+n_2$. Hence $n_1 = n_2$. Call such a number m. $\{P^1, P^2\}$ ($\mathbf{P}(V_0)$) is the Condorcet winner if it beats $\{\varnothing\}$ but the ranking of the two outcomes by types T_1 and T_2 is ambiguous. Call m_1^{PP} ($m_1^\varnothing$) the number of voters of type T_1 who rank $\{P^1, P^2\}$ above (below) $\{\varnothing\}$, and similarly for voters of type T_2.

$\mathbf{P}(V_0) = \{P^1, P^2\}$ beats $\{\varnothing\}$ if and only if $(n+1)/2 + m_1^{PP} + m_2^{PP} > (n-1)/2 + m_1^{\varnothing} + m_2^{\varnothing}$, or:

$$1 > \left(m_1^{\varnothing} + m_2^{\varnothing}\right) - \left(m_1^{PP} + m_2^{PP}\right) \tag{1}$$

Note that $\left(m_1^{\varnothing} + m_2^{\varnothing}\right) + \left(m_1^{PP} + m_2^{PP}\right) = 2m$, an even number. But if the sum of two numbers is even, the difference of those two numbers is also even. Hence $\{P^1, P^2\}$ is the Condorcet winner if and only if $\left(m_1^{\varnothing} + m_2^{\varnothing}\right) = \left(m_1^{PP} + m_2^{PP}\right) - 2R$, where R is an integer strictly larger than 0.

We now show that if $\{P^1, P^2\}$ is the Condorcet winner, than it must also be the Pivot-stable outcome. Any pair of traders blocking V_0 must be such that one of them is counted in $m_1^{\varnothing}$ and one is counted in $m_2^{\varnothing}$. With $K = 2$, voters can only trade votes once. Hence V_0 is blocked once if $\min(m_1^{\varnothing}, m_2^{\varnothing}) = 1$ and any V_t such that $\mathbf{P}_t = \{P^1, P^2\}$ can potentially be blocked s times if $\min(m_1^{\varnothing}, m_2^{\varnothing}) = s$. Similarly, if V_0 is blocked and V_1 is such that $\mathbf{P}_1 = \{\varnothing\}$, V_1 will be blocked if $\min(m_1^{PP}, m_2^{PP}) = 1$. As above, any V_t such that $\mathbf{P}_t = \{\varnothing\}$ can potentially be blocked s' times if $\min(m_1^{PP}, m_2^{PP}) = s'$. Thus $\{P^1, P^2\}$ is the Pivot-stable outcome if and only if $\min(m_1^{PP}, m_2^{PP}) \geq \min(m_1^{\varnothing}, m_2^{\varnothing})$. Now recall that $(m_1^{\varnothing} + m_1^{PP}) = (m_2^{\varnothing} + m_2^{PP})$ since both sums must equal m. Thus if $\{P^1, P^2\}$ is the Condorcet winner and $\left(m_1^{\varnothing} + m_2^{\varnothing}\right) = \left(m_1^{PP} + m_2^{PP}\right) - 2R$, it must be that $m_1^{\varnothing} = m_2^{PP} - R$ and $m_2^{\varnothing} = m_1^{PP} - R$. Hence $\min(m_1^{PP}, m_2^{PP}) = \min(m_1^{\varnothing}, m_2^{\varnothing}) + R$, or $\min(m_1^{PP}, m_2^{PP}) \geq \min(m_1^{\varnothing}, m_2^{\varnothing})$, and $\{P^1, P^2\}$ is the Pivot-stable outcome. Note that identifying $\mathbf{P}(V_0)$ with $\{P^1, P^2\}$ is with no loss of generality. The proof can be restated as follows: the only candidate for Condorcet winner is $\mathbf{P}(V_0)$, and when $\mathbf{P}(V_0)$ is the Condorcet winner, then it must also be the Pivot-stable outcome.

(4) By Kadane's "improvement algorithm", if $\mathbf{P}(V_0) = \{P^1, P^2\}$, then not only is $\{P^1, P^2\}$ majority preferred to $\{P^1\}$ and to $\{P^2\}$, but $\{P^1\}$ and $\{P^2\}$ are majority preferred to $\{\varnothing\}$. Hence if there is a Condorcet loser, it can only be $\{\varnothing\}$. But if $\{\varnothing\}$ is the Condorcet loser, it means that $\{P^1, P^2\}$ is majority preferred to $\{\varnothing\}$. Hence $\mathbf{P}(V_0) = \{P^1, P^2\}$ is the Condorcet winner, and by (2) above $\mathbf{P}(V_T) = \mathbf{P}(V_0)$.

(5). From (2) above, if $\mathbf{P}(V_0)$ is the Condorcet winner,then $\mathbf{P}(V_T) = \mathbf{P}(V_0)$. Thus if $\mathbf{P}(V_T) \neq \mathbf{P}(V_0)$, $\mathbf{P}(V_0)$is not the Condorcet winner. If, with no loss of generality, $\mathbf{P}(V_0) = \{P^1, P^2\}$, then if $\mathbf{P}(V_T) \neq \mathbf{P}(V_0)$, $\mathbf{P}(V_T) = \{\varnothing\}$. By Kadane's argument, $\mathbf{P}(V_0) = \{P^1, P^2\}$ is majority preferred to $\{P^1\}$ and to $\{P^2\}$. Hence if $\mathbf{P}(V_0)$is not the Condorcet winner,$\{\varnothing\}$ must be majority preferred to $\{P^1, P^2\} = \mathbf{P}(V_0)$. But $\{\varnothing\} = \mathbf{P}(V_T)$. Hence if $\mathbf{P}(V_T) \neq \mathbf{P}(V_0)$,

$\mathbf{P}(V_T)$ must be majority preferred to $\mathbf{P}(V_0)$.□

Proposition 2. *If $N = 3$, then for all K, Z, and R: (1) if a Condorcet winner exists, $\mathbf{P}(V_T)$ is unique and is the Condorcet winner.(2) $\mathbf{P}(V_T)$ can never be the Condorcet loser.*[33]

Proof. (1). Select any k proposals, with $k = 2, .., K-1$. Call $\mathbf{P}(V_0, k^-)$ the outcome that would follow if the k proposals were decided against the majority preference at V_0, and the remaining $K-k$ according to the majority preference at V_0, i.e. as if one vote trade was executed on the k proposals and none on the remainder. If the Condorcet winner exists, by Park and Kadane, it can only be $\mathbf{P}(V_0)$. Thus for any $\mathbf{P}(V_0, k^-)$ at least two of the three voters prefer $\mathbf{P}(V_0)$ to $\mathbf{P}(V_0, k^-)$. But then no trade can take place. If the Condorcet winner exists, V_0 cannot be blocked. Thus $\mathbf{P}(V_T)$ equals $\mathbf{P}(V_0)$ and is the Condorcet winner.

(2) We begin by reiterating, and generalizing, an argument we used above in the proof of Proposition 1.

Lemma 2. *Call $\mathbf{P}_-(V_0)$ the outcome obtained by choosing the minority's preferred direction for each proposal at V_0. If a Condorcet loser exists, it can only be $\mathbf{P}_-(V_0)$.*

Proof of Lemma 2. The Lemma follows from Kadane's improvement algorithm. Select k from the K proposals ($k = 1, .., K-1$). Consider the outcome $\mathbf{P}(V_0, k^-)$ obtained by deciding those k proposals in the direction favored by the minority at V_0, and the remainder $K-k$ in the direction favored by the majority. Now consider the outcome obtained by switching one additional proposal from the majority to the minority's preferred direction at V_0: $\mathbf{P}(V_0, k^-, (k+1)^-)$, where the argument k^- is maintained to make clear that the selection of the original k proposals has not changed. Then, by construction, $\mathbf{P}(V_0, k^-)$ is majority preferred to $\mathbf{P}(V_0, k^-, (k+1)^-)$. It follows that for any $\mathbf{P} \neq \mathbf{P}_-(V_0)$ there always exists an outcome $\mathbf{P}'$ that differs only by switching the direction of one proposal from the majority's to the minority's at V_0 such that $\mathbf{P}$ is majority preferred to $\mathbf{P}'$. Hence if a Condorcet loser exists, it can only be $\mathbf{P}_-(V_0)$. □

We can now prove (2). By Lemma 2, if a Condorcet loser exists it can only be $\mathbf{P}_-(V_0)$. The result follows if we can show that if $\mathbf{P}_-(V_0)$ is Pivot-stable, it cannot be the Condorcet loser. Suppose $\mathbf{P}_-(V_0)$ is Pivot-stable, and call V_0^- any allocation of votes such that $\mathbf{P}(V_0^-) = \mathbf{P}_-(V_0)$. Given preferences Z, consider any allocation of votes $V(Z)$ such that V_0^- can be reached from

[33] Result (1) replicates the result in Koehler (1975), under slightly different assumptions. We find however that the result is limited to $N = 3$.

$V(Z)$ through a strictly pair-wise improving trade. Since $\mathbf{P}_-(V_0) \neq \mathbf{P}(V_0)$ and $\mathbf{P}_-(V_0)$ is Pivot-stable, $V(Z)$ must exist. With $N = 3$, the existence of a strictly pair-wise improving trade implies that $\mathbf{P}(V_0^-) = \mathbf{P}_-(V_0)$ is majority preferred to $\mathbf{P}(V(Z))$. Hence if $\mathbf{P}_-(V_0)$ is Pivot-stable, it cannot be the Condorcet loser, and the result is proven. □

Proposition 4. *There exist K, N, Z, and R_C such the C-Pivot algorithm never reaches a stable vote allocation.*

Proof. Consider the following example, where as usual rows represent proposals, columns represent voters and the entry in each cell is z_i^k, the value attached by voter i to proposal k passing.

	1	2	3	4	5	6	7
A	3	-2	-2	-2	1	1	1
B	-2	3	-2	-2	1	1	1
C	-2	-2	3	-2	1	1	1
D	-2	-2	-2	3	1	1	1

At V_0, all proposals pass, and $u_i(V_0) = -3$ for $i = \{1, 2, 3, 4\}$. Consider a coalition composed of such voters, and the following coalition trade: voter 1 gives his A vote to voter 2, in exchange for his B vote; voter 3 gives his C vote to voter 4, in exchange for his D vote. At V_1, all proposals fail and $u_i(V_1) = 0$ for all $i \in C$. The trade is strictly improving for all members of the coalition. In addition, it is a minimal trade, since V_1 cannot be reached by the coalition by trading fewer votes, nor can it be reached by a smaller coalition (each of the two pair-wise trades alone is welfare decreasing for the pair involved). But note that V_1 is not C-Pivot stable: voters 1 and 2 can block V_1 by trading back their respective votes on A and B, reaching outcome $P(V_2) = \{P^A, P^B\}$, and enjoying a strictly positive increase in payoffs: $u_j(V_2) = 1$ for $j = \{1, 2\}$. The same argument applies to voters 3 and 4. At V_2, $u_s(V_2) = -4$ for $s = \{3, 4\}$, but 3 and 4 can block V_2, trade back their votes on C and D, and obtain a strict improvement in their payoff: $P(V_3) = \{P^A, P^B, P^C, P^D\}$, and $u_s(V_3) = -3$ for $s = \{3, 4\}$. Note that the sequence of trades has generated a cycle: $V_3 = V_0$, an allocation that is blocked by coalition $C = \{1, 2, 3, 4\}$, etc.. Hence for R_C that selects the blocking coalitions in the order described, no C-Pivot stable allocation of votes can be reached.□

The logic we exploited in the proof of the Theorem in section 3 does not extend to coalition trades. Because of the externalities present in coalitional trades, the score function we defined earlier is no longer monotonically increasing in the number of trades. In the example, voters $1, 2, 3$, and 4 have a score of 9 before the coalition trade and a score of 8 after the trade.[34] Cycles become possible, and stability cannot be guaranteed.

Proposition 5. *Consider Z such that the Condorcet winner exists. (1) If either $K = 2$ or $N = 3$, then the C-Pivot stable outcome always exists, is unique, and coincides with the Condorcet winner. (2) If $K > 2$ and $N > 3$, if a C-Pivot stable outcome exists, it need not coincide with the Condorcet winner.*

Proof. (1) (i) Consider first the case $K = 2$. The following Lemma establishes that the existence of a C-Pivot stable outcome:

Lemma 3. *If $K = 2$, then for all N, Z, and R a Pivot stable allocation V always exists.*

Proof of Lemma 3. We need two further Lemmas:

Lemma 4. *If at V_0 both proposals pass by minimal majority, then if $K = 2$ all C-Pivot trades must be pair-wise trades.*

Proof of Lemma 4. By minimality of the coalition trades, if at V_0 both proposals pass by minimal majority then at each step of the algorithm both proposals must be decided by minimal majority. At each step t any coalition must include a voter with preferences $\{P^1\} \succ P(V_t)$ and a voter with preferences $\{P^2\} \succ P(V_t)$, and each must have at least one vote to trade away, or they would not be part of the minimal coalition. But if the proposals are decided by minimal majority, the two voters could just trade among themselves. Hence the minimal coalition must be a pair.□

Lemma 5. *If $K = 2$, there can be at most one trade that is not pair-wise, and it can only be the first.*

Proof of Lemma 5. By Lemma 4, if non pair-wise coalition trades occur, it must be that at V_0 at least one proposal is not decided by minimal majority. Any blocking coalition at V_0 must then involve more than two voters. But by minimality, after the coalition trade both proposals must be decided by minimal majority. But then from $t = 1$ onward, all trades can only be pair-wise trades.□

We can now prove Lemma 3. If V_0 cannot be blocked, then it is trivially stable. If it can be blocked,then Lemma 4 and 5 imply that any voter can at

[34] In the example, the coalition trade can be divided into two separate pair-wise trades, but this feaure plays no important role. It is easy to generate examples where coalition trades are linked in a chain. What matters is that scores can fall after a welfare-improving coalition trade.

most trade once. Hence the trades must converge to a stable votes allocation. We can say more: a stable votes allocation must be reached in at most $(N-1)/2$ steps, the maximal number of possible pair-wise trades.$\square$

We can now prove that the C-Pivot stable outcome must coincide with the Condorcet winner, when the Condorcet winner exists, and thus be unique.[35] If at V_0 both proposals pass by minimal majority then by Lemma 4 Proposition 1 applies and the result follows. Suppose then that at V_0 at least one proposal is not decided by minimal majority. If there is no blocking coalition, V_0 is trivially stable, and since the Condorcet winner can only be $P(V_0)$, then the stable outcome coincides with the Condorcet winner, if it exists. Suppose then that at least one blocking coalition exists (and note that it must include more than two voters). If several exist, select one by rule R_C. After the first coalitional trade, $P(V_1) = \{\varnothing\}$. With $K = 2$, all members of a minimal coalition can only trade once, regardless of the coalition's size. We can construct a fictional vote allocation $\widetilde{V}_0$ and preferences $\widetilde{Z}$ such that $\widetilde{v}^k_{0i} = v^k_{1i}$ and $\widetilde{z}^k_i = z^k_i$ for all $i \notin C$, and $\widetilde{v}^k_{0i} = 1$ and $\widetilde{z}^k_i = -1$ for all $i \in C$, $k = 1, 2$. Note that $P(\widetilde{V}_0) = P(V_1) = \varnothing$, and $\widetilde{Z}$ respects all individuals' ranking between the only possible outcomes, $\{P^1, P^2\}$ and $\{\varnothing\}$, but $\widetilde{V}_0$ and $\widetilde{Z}$ guarantee that, as required, all $i \in C$ will not trade any further. At $\widetilde{V}_0$, both proposals are decided by minimal majority, and starting from $\widetilde{V}_0$ all minimal coalitions will be pair-wise. As a result, starting from $\widetilde{V}_0$, Proposition 1 applies. Note that, if $P(V_0) = \{P^1, P^2\}$ is the Condorcet winner, then a majority prefers $P(V_0)$ to $P(\widetilde{V}_0)$. Hence, by Proposition 1, $P(\widetilde{V}_0)$ cannot be Pivot stable. The C-Pivot stable outcome is then $P(V_0)$. But $P(V_0)$ is the Condorcet winner and the result is proven.

(ii). Consider now the case $N = 3$. Recall that the Condorcet winner can only be $\mathbf{P}(V_0)$. If $N = 3$, and the Condorcet winner exists, then no welfare improving trade exists at V_0. Hence $P(V_0)$ is trivially C-Pivot stable, and is the Condorcet winner.$\square$

Trade paths for the experimental matrices.

(1) Matrix A. There is a unique Pivot path. At V_0, with $\mathbf{P}(V_0) = \{P_B\}$, 1 trades his B vote to 3 in exchange for 3's A's vote, leading to $\mathbf{P}(V_1) = \{P_A\}$; then 2 trades his B vote to 4 in exchange for 4's A's vote. The resulting vote allocation is stable, and $\mathbf{P}(V_2) = \mathbf{P}(V_T) = \{P_B\}$. (2) Matrix $M1$. $\mathbf{P}(V_0) = \{P_A\}$

[35] In fact, with $K = 2$ it is possible to show that the C-Pivot stable outcome is always unique, whether or not the Condorcet winner exsts, and it is always Pareto optimal and always majority preferred to $\mathbf{P}(V_0)$.

is the Condorcet winner, but V_0 is not stable. The unique Pivot-stable outcome is $\mathbf{P}(V_T) = \{P_A, P_B, P_C\}$. Three alternative paths, of length $\{4, 4, 3\}$ lead to it. Indicating first the ID's of the trading partners, and then the issue on which an extra vote is acquired by the voter listed first, the three paths are: $\{13CB, 45BC, 23AB, 45CA\}$, $\{23AB, 45CA, 45BC, 13CB\}$, and $\{23AB, 45BA, 13CB\}$. (3) Matrix $M2$. $\mathbf{P}(V_0) = \{P_A, P_B, P_C\}$ is the Condorcet winner and the unique Pivot-stable outcome. However V_0 is not stable. Three alternative paths, of length $\{4, 4, 3\}$ lead to $\mathbf{P}(V_T) = \{P_A, P_B, P_C\}$. They are: $\{15AB, 34BA, 24CB, 15BC\}$, $\{24CB, 15BC, 15AB, 34BA\}$, and $\{24CB, 15AC, 34BA\}$.

Under weak-Pivot, the trades just described remain legitimate and the Pivot-stable outcomes are weak-Pivot stable (under the appropriate R rule). However different paths and outcomes are also possible. Given our matrices, a necessary constraint is that the number of issues on which the direction preferred by a majority of votes changes as a result of a trade must be even–either 0 or 2 (because if a single issue changes direction the trade must be welfare-decreasing for one the two trading partners). Thus additional weak-Pivot stable outcomes are: $\mathbf{P}(V_T) = \{P_A\}$, with matrix A; and $\mathbf{P}(V_T) \in \{\{P_A\}, \{P_B\}, \{P_C\}\}$ with matrices $M1$ and $M2$.

Proposition 6. *For all K, N, Z such that the Condorcet winner exists: (1) there always exists $R^w(Z)$ such that the Condorcet winner is weak-Pivot stable. (2) If V_0 is not stable, there always exists $R^{w'}(Z)$ and a weak-Pivot stable outcome $\mathbf{P}^w(V_T(R^{w'}))$ that differs from the Condorcet winner.*

Proof. Recall that if an outcome is the Condorcet winner, it can only be $\mathbf{P}(V_0)$.

(1). If at V_0 all but at most one issue are decided by non-minimal majority, then no pair of voters can block V_0, and $\mathbf{P}(V_0)$ is trivially weak-Pivot stable. Suppose that there exists a set of issues $\widehat{K}_0 \subseteq K$, with $|\widehat{K}_0| \geq 2$, such all $k \in \widehat{K}_0$ would be decided by minimal majority at V_0. V_0 can be blocked only if there exist at least two issues k and $k' \in \widehat{K}_0$ such that there exist a voter i who wins on k and loses on k', and a voter j who loses on k and wins on k'. But for any such pair of issues, i can give his k' vote to j in exchange for j's k vote. The trade is legitimate under weak-Pivot: at V_1, the outcome has not changed, but k and k' are not decided by minimal majority any longer. We can define a new subset of issues, $\widehat{K}_1 \subset \widehat{K}_0$ such that all $k \in \widehat{K}_1$ would be decided by minimal majority at V_1, and proceed in the same fashion. The process converges to a vote allocation V_T such that there exist no two issues that would both be decided by minimal majority at V_T and over which at least two voters have opposite preferences. But then V_T is stable, and $\mathbf{P}(V_T)$ is weak-Pivot stable. By construction, $\mathbf{P}(V_T) = \mathbf{P}(V_0)$, and thus $\mathbf{P}(V_T)$ is the Condorcet winner.

(2). Because V_0 is not stable, there must exist at least two voters, i and j, who can block V_0, i.e. such that a pair-wise trade of votes between them (a) results in a strict (myopic) gain for both of them, and (b) changes the outcome of at least two issues k and k'. Execute the trade, which is legitimate under weak-Pivot. Thus $\mathbf{P}(V_1) \neq \mathbf{P}(V_0)$. If V_1 is stable, statement (2) follows immediately. If V_1 is not stable, proceed as under (1) above, until, by construction, $\mathbf{P}(V_T) = \mathbf{P}(V_1)$. But $\mathbf{P}(V_1)$ is not the Condorcet winner, and thus neither is $\mathbf{P}(V_T)$. $\square$

The intuition is straightforward. The Condorcet winner can only be the pre-trade outcome. If V_0 is stable, then (1) follows trivially. If V_0 is not stable, then there always exists a zero-gain trade that does not change the resolution of any issue but moves the allocation closer to stability. On the other hand, if V_0 is not stable, it is always possible to realize a Pivot trade, and reach vote allocation $V_1 \neq V_0$, with $\mathbf{P}(V_1) \neq \mathbf{P}(V_0)$. But then, by the argument above, there exists a further set of trades such that the resolution of any issue does not change, relative to V_1 but but moves the allocation closer to stability. $\mathbf{P}(V_t) = \mathbf{P}(V_1)$. Thus there exists $\widetilde{R}^w(Z)$ such that $\mathbf{P}(V_T) \neq \mathbf{P}(V_0)$.

The difficulty with weak-Pivot trades is that many stable outcomes become reachable. It all depends on the selection rule R, the selection among the many possible paths. Note that in the experiment, we will have no control on how different trades are prioritized.

Trade paths for the experimental matrices: Weak-Pivot.

The trade paths described for the Pivot algorithm remain legitimate under weak-Pivot, and the Pivot-stable outcomes are weak-Pivot stable (under the appropriate R rule). However different paths and outcomes are also possible. Given our matrices, a necessary constraint is that the number of issues on which the direction preferred by a majority of votes changes as a result of a trade must be even–either 0 or 2 (because if a single issue changes direction the trade must be welfare-decreasing for one the two trading partners). Thus additional weak-Pivot stable outcomes are: $\mathbf{P}(V_T) = \{P_A\}$, with matrix A; and $\mathbf{P}(V_T) \in \{\{P_A\}, \{P_B\}, \{P_C\}\}$ with matrices $M1$ and $M2$.

References

[1] Balinski, M. and T. Sonmez, 1999, "A Tale of Two Mechanisms", *Journal of Economic Theory*, 84, 73-94.

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

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

[4] Echenique F. and L. Yariv, 2012, "An Experimental Study of Decentralized Matching", mimeo, Caltech.

[5] Ferejohn, J, 1974, "Sour notes on the theory of vote trading", Caltech, Social Science Working Paper No.41.

[6] Fischbacher U. and S. Shudy, 2010, "Reciprocity and Resistance to Comprehensive Reform", Research Paper No.119864, Turgau Institute of Economics and Department of Economics, University of Konstanz.

[7] Gale D. and L. Shapley, 1962, "College Admissions and the Stability of Marriage, *American Mathematical Monthly*, 69, 9-15.

[8] Gusfield, D. and R. Irving, 1989, *The Stable Marriage Problem*, Cambridge, Ma: MIT Press.

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

[10] 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.

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

[12] 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.

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

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

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

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

[17] 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.

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

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

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

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

[22] Roth, A. and M. Sotomayor, 1990, *Two-Sided Matching: A Study in Game-Theoretic Modelling and Analysis*, Cambridge, UK: Cambridge University Press.

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

[24] Sasaki, H. and M. Toda, 1996, "Two-Sided Matching Problems with Externalities", *Journal of Economic Theory*, 70, 93-108.

[25] Schwartz, 1977, *American Political Science Review*

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

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

VOTE TRADING INSTRUCTIONS
A (2x3), M (5x3), M (5x3). Known Values. No Chat. 11/14/14; 11/21/14 Columbia

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 $10.00. 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. [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 reposted 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.).

[CONNNECT 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 $10.00 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.