Social Choice and Welfare
https://doi.org/10.1007/s00355-025-01606-w

ORIGINAL PAPER

Equilibria and group welfare in vote trading systems

Matthew I. Jones[1,2,3,4]

Received: 11 November 2024 / Accepted: 25 April 2025

Abstract

We introduce a new framework to study the group dynamics and game-theoretic considerations when voters are allowed to trade votes. This model advances prior work by considering vote-for-vote trades in a low-information environment where voters do not know the preferences of their trading partners and do not abstain from voting. All voters draw their preference intensities on two issues from a common probability distribution and then consider offering to trade with an anonymous partner. The result is a strategic game between two voters that can be studied analytically. We compute the Nash equilibria for this game and derive several interesting results involving symmetry, group heterogeneity, and more. This framework allows us to determine that trades are typically detrimental to the welfare of the group as a whole, but there are exceptions. We also expand our model to allow all voters to trade votes and derive approximate results for this more general scenario. Finally, we emulate vote trading in real groups by forming simulated committees using real voter preference intensity data and computing the resulting equilibria and associated welfare gains or losses.

1 Introduction

For good reason, majority rule is the most widely-used decision rule for groups to consolidate members' preferences into a single selection, particularly for binary choices. This process is anonymous, decisive, and neutral (May 1952), but it does not take into account voters' preference intensities; two voters that care deeply may be outvoted by three relatively ambivalent voters. While some may argue that this is by design (Bouton et al. 2021), others conclude that sufficiently motivated minorities should be allowed to exert an oversized influence on certain issues (Jacobs et al. 2009; Casella 2011).

✉ Matthew I. Jones
matt.jones@yale.edu

1 Yale Institute for Network Science, Yale University, 17 Hillhouse Ave, New Haven, CT 06511, USA

2 Department of Sociology, Yale University, 493 College St, New Haven, CT 06511, USA

3 Sunwater Institute, 12358 Parklawn Dr, North Bethesda, MD 20852, USA

4 Department of Mathematics and Computer Science, College of the Holy Cross, 1 College St, Worcester, MA 01610, USA

Published online: 21 May 2025

Unfortunately, extracting preference intensities from voters is not straightforward, since they are incentivized to inflate the intensity of all their beliefs and claim that all issues are of paramount importance. When the group is deciding on multiple issues, one potential remedy for this shortcoming of majority rule is to allow voters to trade votes across issues, accumulating extra votes on their most valued issues and giving up their autonomy on issues they view as unimportant.

The study of vote trading has a long and complicated history that examines many different forms of voting and exchanging of votes to account for preference intensity (see Casella and Macé 2021 for a recent review). The central question, approached from many angles but not completely resolved, is "Does the trading of votes improve outcomes for the entire group?" The trading of votes-for-votes in a majority rule system introduces several complications that make votes unlike traditional goods in a market and require new analyses to make meaningful statements about the value of any particular trade.

Early theorists intuited that vote trading could address two issues in social choice: majority cycles and failure to respond to preference intensity (Buchanan and Tullock 1962; Coleman 1966; Mueller 1967), but both of these claims proved false, in general. The tendency to remove majority cycles was quickly refuted by Park (1967) and Miller (1977). It took only slightly longer to show that although it is straightforward to create situations where the vote trading adds value to the group, it is equally simple to create scenarios where vote trading reduces value or even leads to Pareto inferior outcomes (Riker and Brams 1973; Ferejohn 1974).

After an initial flurry of activity in the 1970s, the study of vote trading diversified, studying various properties of vote trading and different mechanisms by which votes can be traded. One branch looks at vote trading as a dynamical system, studying stability and the convergence of trading to Condorcet winners (Casella and Palfrey 2019, 2021). There are also other ways in which votes can be exchanged besides the classic votes-for-votes framework. One line of research has examined "implicit" vote trading by combining multiple issues into a single bundle (Jones et al. 2023; Câmara and Eguia 2017). Another way to trade votes is to exchange them for a numeraire, an alternative currency, that has value for voters and can be used to buy and sell votes (Casella et al. 2012; Xefteris and Ziros 2017; Lalley and Weyl 2018). Finally, a voter can trade votes with themselves by shifting votes from one issue to another (Casella 2011; Jackson and Sonnenschein 2007). Each trading framework differs from vote-for-vote trading in important ways, and each requires its own analysis.

This paper returns to the classic vote-for-vote model and attempts to study its effect on group welfare in a probabilistic framework. Samsonov, Solé-Ollé, and Xefteris also recently looked at trading votes for votes, but in a system where vote count determines intensity of the reform, removing the payoff discontinuity at the pivotal vote and transforming the voting system back into a traditional market (Samsonov et al. 2023).

In addressing the question of vote trading's impact on group welfare, one must first define group welfare. Early work, driven by economists, tended to focus on Pareto efficiency, with Ferejohn (1974) going so far as to say "Of course Mueller may be using some other notion of welfare [other than Pareto improvements] in which case it is not possible to decide the question of whether vote trading can produce welfare gains." Like many other papers to study vote trading, this paper uses a utility model where each voter gains utility when issues are accepted or rejected by the group (Riker and Brams 1973; Casella and Palfrey 2019) and we define the welfare of the group to be the sum of the voters' final utilities. Because traditional Pareto efficiency approaches avoid interpersonal comparisons, they are ill-suited to measure the value of a group decision. The simple sum of individual utilities, on the other hand, has the benefit that the value for a hypothetical passionate minority is quantified and directly measured against the indifferent majority and the value of the group decision on each issue can be examined in isolation, unlike traditional Pareto efficiency and Condorcet approaches.

Our model incorporates aspects from other papers as well. Like Xefteris and Ziros (2017), our voters have incomplete information, forcing them to make decisions in an ambiguous environment. Much of the literature around vote trading assumes that voters have complete information about the preferences (and preference intensities) of the entire voting population. This severely restricts the space of rational trades to only pivotal votes, where the trade actually changes the outcome of the vote, and this is an assumption worth examining for two reasons. First, even in the most public and high-profile spaces, voters' preferences can be kept secret (Ramzy 2017; McPherson 2019). Second, in a system where voters are required to state their preferences, they will be incentivized to lie about their preferences to make themselves more appealing as trading partners. Since preferences are often unknown (and when they are known, they may be falsified), we assume that voters are unsure of the preferences of others. All voters operate with the same knowledge about the distribution from which voter utilities are sampled but have incomplete information about the preferences of their partners. This model of vote trading represents a critical divergence from the majority of vote trading work. Without complete information, voters must be concerned not just with the value of the issues they are trading on, but also if their trade will make any difference at all!

In the main model, our voters are short-sighted, making myopic decisions as if no other voters will make trades (Casella and Palfrey 2018). We loosen this assumption in Sect. 4 and approximate the effect of randomly pairing voters and allowing them all to trade. Myopic trading also means that voters make sincere trades, working to earn votes that they think are valuable, instead of acquiring votes only to trade them away later (Iaryczower and Oliveros 2016).

As the toy model presented in this paper shows, a voter's decision to offer a potential trade is dependent on their assessment of what their trading partner's preferences will be. Since partner behavior is critical in determining one's own behavior, we study this system via Nash equilibria in which no participants can improve their payoff by changing strategy (Holt and Roth 2004). In this paper, we show how to find these equilibria and then prove several interesting properties of these vote trading systems. We also compute the value of vote trading for the group, specifically we find the probability that a trade has positive expected value for the net utility of the group, which can range from 0 to 1 depending on the underlying utility distribution.

Finally, we use this new tool on real voter preference intensity data (Studies 2021). Gathering empirical vote trading data is nearly impossible and most empirical vote trading studies are conducted in the laboratory with artificially-assigned utilities (Casella and Palfrey 2021; Tsakas et al. 2020; Goeree and Zhang 2017). This analysis illustrates how this analysis of vote trading equilibria can be utilized alongside empirical data. In this example, we study the effect of vote trading in a hypothetical committee made up of real voters deciding on real issues, but this can be easily repeated on data for real decision-makers, assuming accurate information about such preferences could be collected.

This work contributes to the study of vote trading in several ways. First, it brings new mathematical techniques to bear on an old problem. This new perspective showcases the importance of the utility distribution of voters in determining the value of vote trading to both the voters and the entire group. Second, by determining which trades should be offered, we can directly compute the probability that vote trading adds (or subtracts) value and the expected value of a random trade. While distributions exist where trading is extremely valuable, in most scenarios, vote trading removes value from the group by overriding simple majority rule. And third, it demonstrates a new way to study vote trading, by taking real voter preference data and predicting what trades would be offered by rational voters.

2 A model of vote trading

We begin with a group of n voters v_i, where n is odd to avoid ties. Unless otherwise noted, we use $n = 11$ in our examples when computing actual values. Since we consider one-for-one trades of votes, only the two issues are ever relevant when deciding to offer a trade and additional issues do not appear in our analysis, so for all intents and purposes, we only need to work with two issues, t_1 and t_2.

The vote trade game takes place in three stages. First, voters are independently assigned utilities on both issues which can take any value between -1 and 1. Many, if not most, issues in real life are related, and there can be strong correlations between issues. We take this into account by drawing utilities u_{t_1}, u_{t_2} jointly according to the probability distribution $f : [-1, 1]^2 \to \mathbb{R}$ which can take any form desired. This function is known to all voters and is the factor that will determine the equilibria we find for these systems. We assume throughout that f is continuous and positive almost everywhere. We can integrate this function over the four half-squares to get the probabilities that a random voter has positive or negative value on issues t_1 and t_2, denoted Q_1^+, Q_1^-, Q_2^+, and Q_2^-.

In the second stage, two voters can offer to trade away their vote on t_1 for an additional vote on t_2, trade away t_2 for t_1, both, or neither. In the rare case that both voters offer to trade for either issue, the direction of the trade is chosen randomly.

Finally, all voters cast their votes on both issues. The final utility for each voter is the sum of utilities for all issues that pass with a majority of votes, minus the sum of all utilities on issues that do not pass. The welfare of the entire group is simply the sum of individual utilities.

Initially, we consider a game with two players. One player, referred to as the t_1 voter, will be trading away their vote on t_2 for an additional vote of t_1, and the t_2 voter is the opposite. A strategy for a player is a determination, for each utility pair (u_{t_1}, u_{t_2}), of if the player is willing to take on the role of the t_1 voter, the t_2 voter, neither, or both. We will see in Theorem 1 that the rational strategies can be characterized by eight parameters, θ_i, shown in Fig. 1. Each θ_i is dependent on the ratio of utilities, and determines which trades a player will make. If a player's utility pair falls in one of the colored R_i regions, they are willing to trade votes in at least one direction.

The payoff for each player is their expected change in utility when all voters have utilities distributed according to f and the two players trade their votes if their utilities are compatible, meaning one player's utility pair falls in the union of R_1 through R_4 and the other's is in R_5 through R_8. We focus on symmetric equilibria, sets of eight θ_i values such that if both players are playing this strategy, neither player can improve their expected payoff by changing any of the θ_i values.

The calculation varies slightly depending on the sign of u_{t_1} and u_{t_2}, so we break into the eight regions R_1 through R_8 which are determined by the angles θ_1 through θ_8. Larger θ_i values indicate more profitable trades and an increased probability of trade. To get probabilities, we integrate f over all these regions, so $I_i = \iint_{R_i} f(x, y)dxdy$.

We show in "Appendix A" that the sign of the expected value of a trade with utilities u_1 and u_2 is a function of the ratio $\frac{u_2}{u_1}$. Therefore, while the space of all strategies is vast, only strategies of the form shown in Fig. 1, straight lines through the origin defined by the θ_i parameters, can possibly accept all trades with positive expected value and reject all trades with negative expected value. Because of this, we restrict our analysis to the rational strategies of this type.

2.1 Nash equilibria in voting systems

In any vote trading system, there is a *trivial Nash equilibrium* where no voters offer any trades ($\theta_i = 0$). There is no value in being the sole voter who is offering to trade because no one will ever accept your trade. This trivial equilibrium is not a strict Nash equilibrium, is present for all underlying joint utility distributions f, and is not particularly interesting from a voting perspective, so we restrict the rest of our discussion to non-trivial equilibria where it is possible that a trade occurs that changes the outcome of the vote.

By finding the ratio of utilities where the value of the trade is zero, we can separate the trades that have positive and negative value. We simultaneously solve this for all types of trades for v_1 and v_2 and the result is a Nash equilibrium.

Theorem 1 *Suppose n voters are voting on issues t_1 and t_2 where utilities for the two issues are jointly distributed according to f. A strategy (defined by the eight θ_i coefficients) is a non-trivial Nash equilibrium if and only if the θ_i satisfy the following equations:*

$$\theta_1 = \arctan\left(\frac{I_6(\theta_6) + I_7(\theta_7)}{I_7(\theta_7) + I_8(\theta_8)} \, \frac{(Q_1^-)^{\frac{n-1}{2}} (Q_1^+)^{\frac{n-3}{2}}}{(Q_2^-)^{\frac{n-3}{2}} (Q_2^+)^{\frac{n-1}{2}}}\right) \tag{1}$$

$$\theta_2 = \arctan\left(\frac{I_5(\theta_5) + I_8(\theta_8)}{I_7(\theta_7) + I_8(\theta_8)} \, \frac{(Q_1^-)^{\frac{n-3}{2}} (Q_1^+)^{\frac{n-1}{2}}}{(Q_2^-)^{\frac{n-3}{2}} (Q_2^+)^{\frac{n-1}{2}}}\right) \tag{2}$$

$$\theta_3 = \arctan\left(\frac{I_5(\theta_5) + I_8(\theta_8)}{I_5(\theta_5) + I_6(\theta_6)} \, \frac{(Q_1^-)^{\frac{n-3}{2}} (Q_1^+)^{\frac{n-1}{2}}}{(Q_2^-)^{\frac{n-1}{2}} (Q_2^+)^{\frac{n-3}{2}}}\right) \tag{3}$$

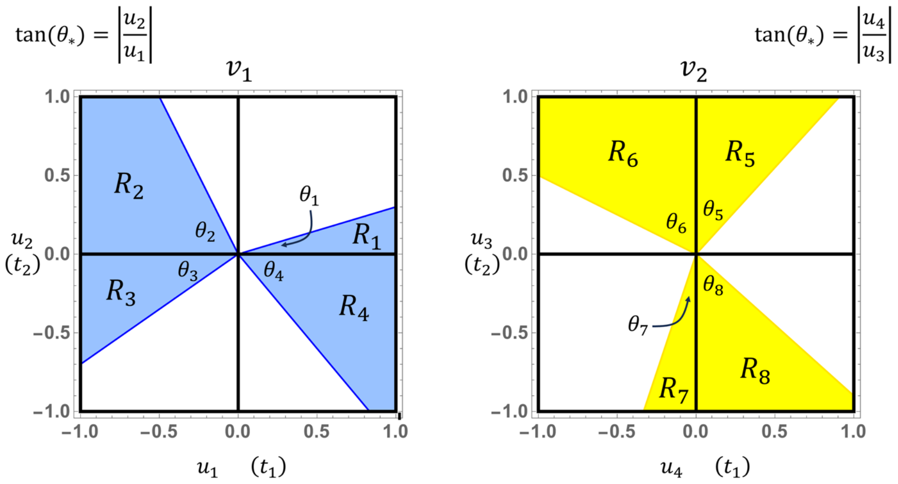

Fig. 1 Trading regions for the t_1 voter (left) and the t_2 voter (right) on two issues. For both plots, the utility on t_1 (u_1 and u_4) is represented by the x-axis, and the utility on t_2 (u_2 and u_3) is represented on the y-axis. The trades that the t_1 voter offers are highlighted in blue, while the t_2 voter's trades are in yellow. Each region R_i is defined by the angle θ_i (alternatively, by the slope of the line). These angles will be adjusted to ensure that all trades in the blue or yellow regions have positive expected value and all trades in white have negative expected value

$$\theta_4 = \arctan\left(\frac{I_6(\theta_6) + I_7(\theta_7)}{I_5(\theta_5) + I_6(\theta_6)}\ \frac{(Q_1^-)^{\frac{n-1}{2}}(Q_1^+)^{\frac{n-3}{2}}}{(Q_2^-)^{\frac{n-1}{2}}(Q_2^+)^{\frac{n-3}{2}}}\right) \tag{4}$$

$$\theta_5 = \arctan\left(\frac{I_3(\theta_3) + I_4(\theta_4)}{I_2(\theta_2) + I_3(\theta_3)}\ \frac{(Q_2^-)^{\frac{n-1}{2}}(Q_2^+)^{\frac{n-3}{2}}}{(Q_1^-)^{\frac{n-3}{2}}(Q_1^+)^{\frac{n-1}{2}}}\right) \tag{5}$$

$$\theta_6 = \arctan\left(\frac{I_3(\theta_3) + I_4(\theta_4)}{I_1(\theta_1) + I_4(\theta_4)}\ \frac{(Q_2^-)^{\frac{n-1}{2}}(Q_2^+)^{\frac{n-3}{2}}}{(Q_1^-)^{\frac{n-1}{2}}(Q_1^+)^{\frac{n-3}{2}}}\right) \tag{6}$$

$$\theta_7 = \arctan\left(\frac{I_1(\theta_1) + I_2(\theta_2)}{I_1(\theta_1) + I_4(\theta_4)}\ \frac{(Q_2^-)^{\frac{n-3}{2}}(Q_2^+)^{\frac{n-1}{2}}}{(Q_1^-)^{\frac{n-1}{2}}(Q_1^+)^{\frac{n-3}{2}}}\right) \tag{7}$$

$$\theta_8 = \arctan\left(\frac{I_1(\theta_1) + I_2(\theta_2)}{I_2(\theta_2) + I_3(\theta_3)}\ \frac{(Q_2^-)^{\frac{n-3}{2}}(Q_2^+)^{\frac{n-1}{2}}}{(Q_1^-)^{\frac{n-3}{2}}(Q_1^+)^{\frac{n-1}{2}}}\right) \tag{8}$$

Proof See "Appendix A". □

We describe a Nash equilibrium that satisfies Eqs. (1)–(8) as non-trivial, in contrast to the trivial Nash equilibrium where no one offers any trades. Equations (1)–(8) form a powerful tool to study equilibria in a vote trading system. For example, we get the existence of a stable equilibrium almost immediately.

Corollary 1 *A non-trivial Nash equilibrium exists.*

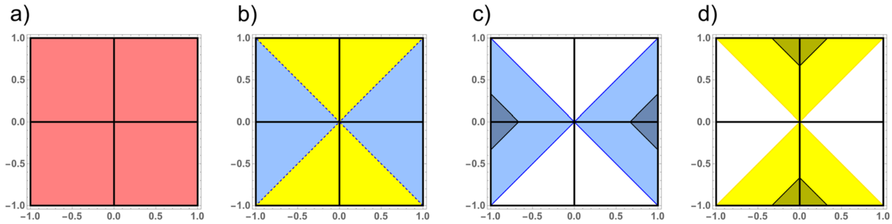

Fig. 2 The results of our vote trading analysis on the uniform distribution. **a** A heatmap of the underlying distribution, in this case a constant $f(x, y) = 1/4$. **b** An equilibrium where the θ_i satisfy Eqs. (1)–(8). We call this particular equilibrium, where all $\theta_i = \frac{\pi}{4}$, the naive Nash equilibrium. **c, d** The trades that improve group welfare (darkened regions) alongside the trades that are offered by the t_1 and t_2 voters, respectively

Proof See "Appendix B", in which we mimic the standard proof for the existence of a mixed Nash equilibrium (Nash 1951) using Brouwer's Fixed Point Theorem (Smart 1980) while bounding away from the origin and avoiding the trivial Nash equilibrium. □

Theorem 1 can also be used to study specific distributions of utilities. For example, in Fig. 2, we consider the uniform distribution $f(x, y) = 1/4$. Plugging in f and solving all eight equations yields the solution $\theta_i = \frac{\pi}{4}$, illustrated in Fig. 2b. We refer to this equilibrium as the "naive strategy" in which a player should offer to give up their vote on an issue that matters less in exchange for another vote on an issue that matters more. This leads us to a powerful symmetry result.

Corollary 2 *Suppose f is point symmetric around the origin, meaning $f(x, y) = f(-x, -y)$. Then the naive strategy is, in fact, a Nash equilibrium.*

Proof Suppose $\theta_i = \frac{\pi}{4}$ for all i. Since f is point symmetric around the origin, $I_1(\frac{\pi}{4}) = I_3(\frac{\pi}{4})$, $I_2(\frac{\pi}{4}) = I_4(\frac{\pi}{4})$, $I_5(\frac{\pi}{4}) = I_7(\frac{\pi}{4})$, and $I_6(\frac{\pi}{4}) = I_8(\frac{\pi}{4})$. These equations make it straightforward to show that $Q_1^+ = Q_1^- = Q_2^+ = Q_2^-$ and then Eqs. (1)–(8) follow easily. □

Even if f is not point symmetric, we can still compute the Nash equilibrium, numerically if necessary. Suppose the two issues have utilities distributed according to the following joint distribution function:

$$f(x, y) = \begin{cases} 1/10 & x, y > 0 \\ 2/10 & x > 0, y < 0 \\ 3/10 & x, y < 0 \\ 4/10 & x < 0, y > 0 \end{cases} \tag{9}$$

The equilibrium for this distribution is shown in Fig. 3. This case differs from the naive equilibrium because of the white and green regions in Fig. 3b. In the white regions, it is not worth trading either issue for the other, even though t_1 is more valuable than t_2. In the green regions, it is worth trading either issue for the other. We provide some intuition for this strange result after another revealing example in Fig. 5.

We present one final claim regarding equilibria of vote trading systems:

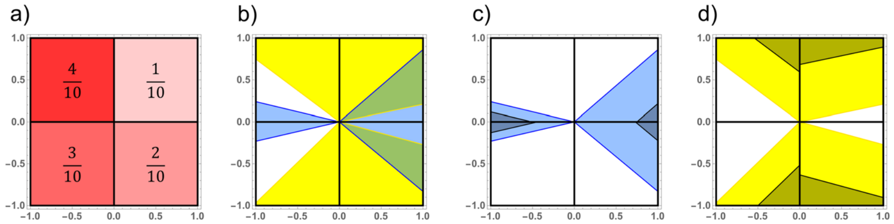

Fig. 3 The results of our vote trading analysis when issues are dependent. **a** A heatmap of the underlying distribution, Eq. (9). **b** The equilibrium. Trades in the green regions are profitable for v_1 and v_2, while positions in the white regions trades are not worth trading for either player. **c, d** The trades that improve group welfare (darkened regions) alongside the trades that are offered by v_1 and v_2, respectively

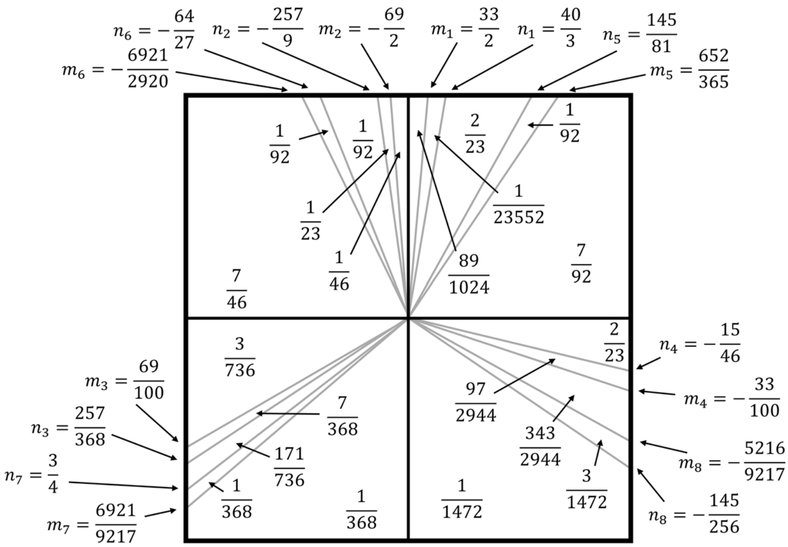

Fig. 4 An integral representation of a function with two Nash equilibria. Each region is defined by one or two lines with slopes indicated by the m_i and n_i. The integral of each region is indicated by the values inside the square. Note that for clarity, these regions are not necessarily to scale. Several of these regions are too small be represented accurately and be readable, so the drawn lines do not represent the true slopes

Corollary 3 *There exist probability distributions f with multiple solutions to Eqs.* (1)–(8) *and therefore have multiple Nash equilibria.*

Proof We prove this with a family of examples. Let f be any joint probability distribution that is positive almost everywhere whose integrals in each of the regions shown is given by the values shown in Fig. 4.

Straightforward computations show that any such function will have at least two equilibria defined by the m_i and n_i coefficients. Therefore, our Nash equilibria are not necessarily unique. Notice that in Fig. 4, $Q_1^+ = Q_1^- = Q_2^+ = Q_2^- = \frac{1}{2}$, so this example holds for all n. □

The distribution shown in Fig. 4 has extremely high density variance. Some tiny slivers of the square have a large fraction of the total utility pairs (like the triangle defined by n_2 and m_2) while other large regions have essentially no mass (like the trapezoid bordered by n_8). All the counterexample families that we have generated seem to have this property, and we conjecture that all "reasonable" utility distributions have a unique non-trivial equilibria, although this seems difficult to state rigorously, let alone prove.

2.2 Group welfare

Theorem 1 has given us a way to determine which strategies are a symmetric Nash equilibrium under any particular density function, so now we can consider the effect such trades will have on the group as a whole. Recall that we define the payoff from any decision for a group to be the sum of payoffs for the individuals in the group. This allows us to compute the expected value of a trade *for the whole group* by extending our proof of Theorem 1 ("Appendix A") to account for the utilities of all members of the group. The derivations and resulting equations are uncomplicated but tedious, so we have placed them in "Appendix C" and simply demonstrate visually with some examples in the main text by plotting the regions where trades have positive expected value for the group.

Again, we first turn to the uniform distribution, where we know that the naive equilibrium holds. Every pair of utilities is offered to trade, but only a fraction of utility pairs have a difference between utilities that is great enough to be worth overriding the majority decision. The darkened regions shown in Fig. 2c, d indicate utility pairs that are beneficial for the entire group. By integrating these regions (or rather, the intersection of these regions with the R_i regions), we compute the probability that a trade improves group welfare. For the uniform distribution, this probability is exactly $\frac{1}{9}$. Furthermore, the expected value (for the whole group) of a trade in the naive strategy equilibrium is ≈ -0.082 when $n = 11$. While the magnitude of this value is dependent on n (because the probability of being pivotal shrinks as n grows), the sign is not, and a trade will have negative expected value for any group size.

For the utility distribution given in Eq. (9), trades are actually slightly more valuable. By integrating the darkened regions of Fig. 3, we see that about 18.5% of trades have positive expected value for the entire group, and the expected value of a random trade is ≈ -0.054.

However, these examples do not imply that vote trading will always have a low probability of improving welfare. Consider the equation

$$g_\alpha(z) = \begin{cases} -1^\alpha \frac{\alpha+1}{2} z^\alpha & z < 0 \\ -1^\alpha \frac{\alpha+1}{2} (z-1)^\alpha & z > 0 \end{cases} \tag{10}$$

and let $f(x, y) = g_\alpha(x) g_\alpha(y)$. An example with $\alpha = 4$ is shown in Fig. 5a. As α approaches ∞, the distribution becomes more and more skewed and the probability of a trade being beneficial approaches 1 since most types of trades become beneficial for the group and the few that do not become increasingly unlikely. In Fig. 5, the probability of a beneficial trade is around 95%, and the expected value of a trade is 1.023.

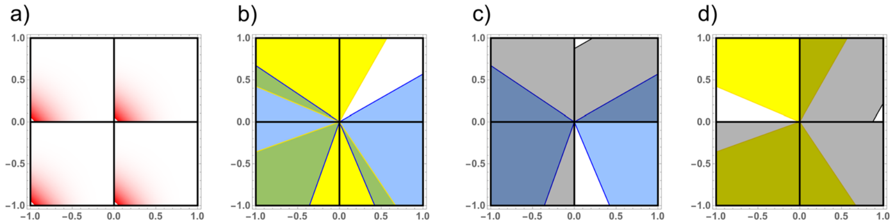

Fig. 5 The results of our vote trading analysis on a highly skewed independent distribution. **a** A heatmap of the underlying distribution from Eq. (10) with $\alpha = 4$. **b** The Nash equilibrium, in which many trades are offered. **c, d** The trades that improve group welfare (darkened regions) alongside the trades that are offered by v_1 and v_2, respectively. The regions of beneficial trades completely cover many of the regions in which trades are being offered, so the vast majority of trades are beneficial for the group

This Nash equilibrium is a revealing example that shows why some utility pairs are worth trading in either direction and others are never worth trading. In this Nash equilibrium, the vast majority of traders are aiming to vote "no" on their preferred issue, so a voter with two negative utilities risks almost nothing by giving away a vote, but there are many traders giving away votes on issues with positive utility, so there is still much to be gained by trading in either direction.

We can also use this technique on simple distributions to draw generalized conclusions about the value of vote trading. For instance, consider two distributions, both symmetric with independent utilities, but where one has mainly uninterested voters and one with predominantly passionate voters (Fig. 6a, d, respectively). Being point symmetric, the naive equilibrium (Fig. 2b) holds for both cases, but trades have a very different impact on group welfare; the former has beneficial trades approximately 24.5% of the time and the expected value of a trade is only ≈ -0.033, while the latter never has beneficial trades, and the expected value is ≈ -0.197.

This result is surprising in the context of previous research which suggests that vote trading is more advantageous *when preferences are heterogeneous*. The key distinction is that these previous works have assumed trading with a numeraire (Casella et al. 2012) or in a power-sharing system (Tsakas et al. 2020), where voters who give up a vote and lose can still gain a small degree of value or representation. Intuitively, the majoritarian system in this model offers no such consolation prize, so trading improves welfare only when the stakes are low enough that most voters are largely indifferent to the outcome of the vote.

Finally, notice that the border of each darkened region, representing beneficial trades, has the same slope as the underlying colored region, representing the trades offered in the Nash equilibrium. In "Appendix C", we compute this vertical shift for each region.

In point symmetric distributions, this vertical shift is the same in all regions, which opens up the possibility of promoting group welfare by imposing a cost on vote trading. Consider the constant distribution in Fig. 2, where each region of group welfare improving trades is shifted up (or down) by $\frac{2}{3}$. In this system, many trades that do not improve welfare could be made. But if every vote trader had to pay a cost of $\frac{2}{3}$ (perhaps

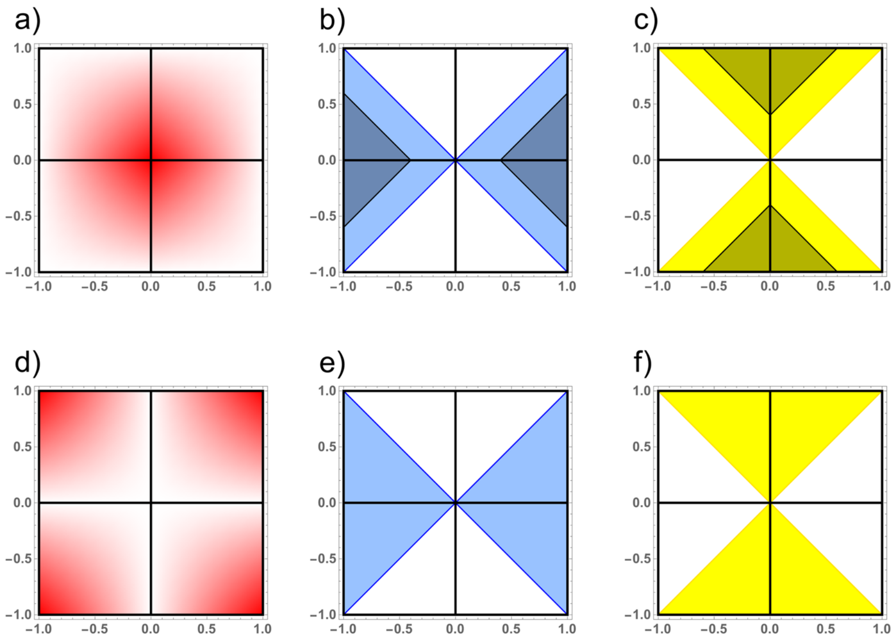

Fig. 6 The difference in benefit of trades in two symmetric distributions. Both distributions have the form $f(x, y) = g(x)g(y)$ where g is a symmetric 1D distribution. **a–c** Correspond to $g(z) = \begin{cases} 1+z & z<0 \\ 1-z & z \geq 0 \end{cases}$ while (**d–f**) correspond to $g(z) = \begin{cases} -z & z<0 \\ z & z \geq 0 \end{cases}$. **a, c** The distribution heatmaps. A group with many low utilities like (**a**) benefits from vote trading much more than a group with many extreme utilities like (**d**)

from some external numeraire), then only trades in the darkened region will be offered and all trades will be welfare improving.

Of course, if the utility distribution is not point symmetric (see Figs. 3 and 5), then different regions have different vertical shifts and a constant cost to trading would not have the same effect. In extreme cases, it would even be necessary to pay voters with specific utilities to trade votes.

3 Vote trading in real populations

These toy examples are good for generalized principles, but still a step away from determining if vote trading improves group welfare in the real world. It is very difficult to study vote trading empirically. In the political sphere, most lawmakers are unwilling to admit to voting against their preferences on an issue but the common suspicion is that vote trading is prevalent throughout legislatures (Stratmann 1995; Aksoy 2012; Cohen and Malloy 2014). In this final section, we form a hypothetical committee made up of real voters with real preferences on real issues and examine the effect vote trading would have on this group.

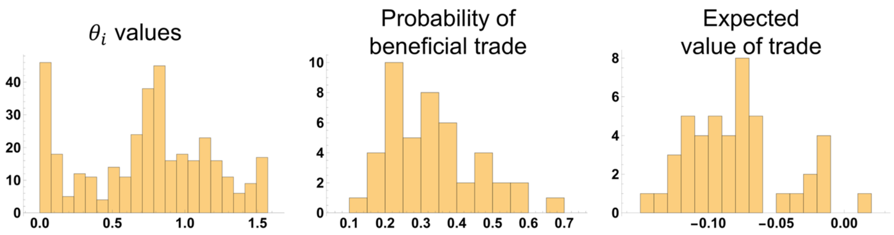

Fig. 7 Examining vote trading equilibria on real joint utility distributions. We examined the 45 joint distributions from all possible pairings of the 10 issues on which we have real voter preferences. On the left, we see a histogram showing the distribution of θ_i values. In the middle, we see the distribution for the probability that a trade is beneficial for the group, and on the right, we see the distribution for the expected values of a trade on each issue pair

The American National Election Survey (Studies 2021) regularly polls voters on a wide range of issues. On some issues, they ask for voters' preferences and the intensity of preferences, so this data allows us to estimate the joint probability distribution of American voter utilities on pairs of issues. The 10 issues where preference intensity was measured involved overall government regulation, government action on income inequality, vaccine requirements in schools, regulation on greenhouse gas emissions, background checks for gun purchases, banning assault rifles, government action on the opioid epidemic, free trade agreements, a universal basic income, and government spending on healthcare. These issues are highly correlated. On all issues, voters expressed their preferences on a scale from 1 (strongly oppose) to 7 (strongly support). Using this discrete data, we apply Gaussian Kernel Density Estimation (Węglarczyk 2018) to get a continuous approximation for the joint probability distribution. Many issues are strongly skewed in one direction or the other, and almost all of the issue pairs are highly correlated.

With this density function, we can find the Nash equilibrium and determine which trades are valuable. In Fig. 7, we can see the results of this analysis across all pairs of issues. First, notice that many θ_i values are close to $\frac{\pi}{4} \approx 0.785$. This suggests that the naive strategy may be a good heuristic for determining rational behavior in realistic scenarios.

We also see that a trade is beneficial to the group about one third of the time, and that the expected value of a random trade is -0.08. Thus, two voters trading votes on real issues is probably harming group welfare by subverting majority rule, rather than benefiting the group by more accurately expressing voter preferences.

4 Group-wide trading

Until now, voters assume that their trade is the only one that will be made, and that the other $n-2$ voters will vote sincerely. Of course, in most scenarios, all voters will have equal opportunity to trade votes, so in this section we extend our model to allow all voters the chance to trade.

Like the single-trade model, we first assign all voters utilities on t_1 and t_2. Then, voters are randomly paired together and allowed to trade like before. Because n is odd, one voter will be left out but the rest of the voters must take into account the other $\frac{n-3}{2}$ pairs of voters that may be trading. Unfortunately, this makes finding an exact closed-form solution difficult, because now voting behavior is dependent across individuals *and* issues, so we introduce an approximation of these dependent variables as independent variables. While imperfect, this allows us to see how vote trading changes when all voters have the opportunity to trade with a random partner.

In “Appendix A”, we computed the probability of the trade being the swing vote as $\binom{n-2}{\frac{n-3}{2}}(Q_1^-)^{\frac{n-1}{2}}(Q_1^+)^{\frac{n-3}{2}}$. When allowing all pairs of voters to trade, we replace these Q terms with new quantities that take into account the probability that the vote was traded away. To see the details of this approximation, see “Appendix C”.

“Appendix C” also includes recreations of all previous results under this new model. Surprisingly, we do not see a large change in voter behavior when all voters are allowed to trade simultaneously. Many of the results are changed only slightly, since it is fruitless to try to take other trades into account unless you can predict the direction of those trades. However, we do see a change in the group welfare implications of trades for the distribution in Fig. 5, an example that was engineered specifically to encourage trades in one direction. In the original model, vote trading was very valuable because it made it more likely that the group would choose the high preference intensity options. Once all voters are allowed to trade, however, the probability of any one trade being influential drops, and the welfare gains of trading also decrease.

5 Discussion

This model of vote trading makes several assumptions of varying strength. Voters have incomplete information about the preferences of their potential trading partners, but complete information about the preferences of the entire group. For simplicity, we restrict voters to one-for-one vote trades, ignoring the possibilities of trading multiple unimportant issues for one supercritical issue or more than two voters gathering to swap a multitude of issues between them all. The primary model also assumes myopic trades where the group votes immediately after a single trade is made, removing the need to consider the impact of other trades on the distribution of utilities. When we loosen this assumption, our approximation reveals changes in the trades being offered when a distribution is only conducive to trades in one direction. All these simplifications and approximations are avenues for future investigation. For example, equipping voters with partial information about the preferences of their partners could illuminate the middle ground between previous complete information models and the incomplete information model presented here.

A crucial component of this analysis is the probability a trade swings the pivotal vote. On balanced issues, where $Q^+ = Q^- = \frac{1}{2}$, the probability of pivotality is about 6% when $n = 11$. Of course, as n grows, this probability approaches zero, but even if the expected value of a trade goes to zero, it is always either positive or negative, so the set of rational trades is unchanged. On unbalanced issues, however, the probability of being pivotal can rapidly go to zero, many trades can become unprofitable, and almost no trades are offered or made. This confirms the natural intuition that if an issue is going to be approved or rejected overwhelmingly, there is no sense trading on that issue, because one additional vote could not possibly overcome the will of the majority.

Despite these limitations, this paper makes several new contributions to our understanding of vote trading. First, it highlights that the welfare implications of vote trading are dependent on the underlying utility distributions. The question of if vote trading improves group welfare does not have a simple yes/no answer. However, for many reasonable utility distributions, including distributions drawn from real data, vote trading adds value for the entire group infrequently, at best. One ray of hope here is that the vote that are beneficial for the whole group are also the most beneficial for the voters themselves, so if voters restrain themselves to only trading on issues where they stand to gain the most utility, the value for the group could also improve.

Second, the toy model presented here gives real insights into the dynamics of vote trading. We see the role of symmetry in the Nash equilibria and we get the counterintuitive result that some utility pairs can be traded in either direction while others cannot be traded for positive value at all. Although these equilibria are not unique in general, our investigation suggests that in most realistic scenarios, rational vote traders should converge to a single Nash equilibrium.

Finally, this paper outlines a probabilistic method for thinking about vote trading, and we believe this could open up new avenues in the study of vote trading.

Acknowledgements We wish to thank Nicholas Christakis, Feng Fu, Larry Samuelson, Dimitrios Xefteris, and Alessandra Casella for their helpful comments. This work is supported by a grant from the Sunwater Institute.

Data/code availability The data and code used for this project are publicly available at https://github.com/MattJonesMath/Vote_Trading_Equilibria.

Declarations

Conflict of interest The author has no Conflict of interest to declare that are relevant to the content of this article. This study was funded by the Sunwater Institute.

Appendix A Proof of Theorem 1

Suppose we have two issues, t_1 and t_2, and two players. The t_1 voter will be giving away their vote on t_2 in exchange for an additional vote on t_1 and the t_2 voter will be giving away their vote on t_1 in exchange for an additional vote on t_2.

Utilities on the two issues are assigned according to a joint probability distribution $f(x, y) : [-1, 1]^2 \to \mathbb{R}$ where x is the utility on t_1 and y is the utility on t_2. By integrating this function, we can determine the probability a random voter supports or opposes t_1 and t_2:

$$\text{Prob}(t_1 \text{ utility} > 0) = Q_1^+ = \int_0^1 \int_{-1}^1 f(x, y)dydx$$

$$\text{Prob}(t_1 \text{ utility} < 0) = Q_1^- = \int_{-1}^0 \int_{-1}^1 f(x, y)dydx$$

$$\text{Prob}(t_2 \text{ utility} > 0) = Q_2^+ = \int_0^1 \int_{-1}^1 f(x, y)dxdy$$

$$\text{Prob}(t_2 \text{ utility} < 0) = Q_2^- = \int_{-1}^0 \int_{-1}^1 f(x, y)dxdy$$

We assume all four quantities are nonzero; otherwise, trading fails as certain individuals have no incentive to trade.To compute the probability that a vote trader supports or opposes an issue, we integrate over these regions and normalize as needed.

$$I_* = \int \int_{R_*} f(x, y)dxdy$$

$$I_{S1} = I_1 + I_2 + I_3 + I_4$$

$$I_{S2} = I_5 + I_6 + I_7 + I_8$$

With this, we can write down all the necessary probabilities:

$$P(u_1 > 0) = \frac{I_1 + I_4}{I_{S1}}$$

$$P(u_1 < 0) = \frac{I_2 + I_3}{I_{S1}}$$

$$P(u_2 > 0) = \frac{I_1 + I_2}{I_{S1}}$$

$$P(u_2 < 0) = \frac{I_3 + I_4}{I_{S1}}$$

$$P(u_3 > 0) = \frac{I_5 + I_6}{I_{S2}}$$

$$P(u_3 < 0) = \frac{I_7 + I_8}{I_{S2}}$$

$$P(u_4 > 0) = \frac{I_5 + I_8}{I_{S2}}$$

$$P(u_4 < 0) = \frac{I_6 + I_7}{I_{S2}}$$

Suppose a trade is about to occur. Let u_1 and u_2 be the utilities for the t_1 voter on t_1 and t_2, respectively. Let u_4 and u_3 be the respective utilities of the t_2 voter. There are four events that are relevant to the expected value for the t_1 voter:

A: The t_1 and t_2 voters have opposite preferences on t_1
B: The t_1 and t_2 voters have opposite preferences on t_2
C: The t_2 voter is the swing vote on t_1
D: The t_1 voter is the swing vote on t_2

Either A and C need to happen, in which case value $2|u_1|$ is gained, or B and D need to happen, in which case $2|u_2|$ value is lost. Therefore, the expected value can be expressed as follows:

$$E(u_1, u_2) = 2|u_1|P(A)P(C|A) - 2|u_2|P(B)P(D|B) \tag{A1}$$

As a note, expectation is a linear operator, so Eq. (A1) holds even if voter utilities are highly correlated. We go through all the details here when $u_1 > 0$ and $u_2 > 0$, which determines the optimal value of θ_1. The other seven cases are similar.

First, consider what happens to the vote on t_1. Nothing happens unless the t_2 voter has negative utility on t_1, i.e. $u_4 < 0$. By above, this occurs with probability $\frac{I_6+I_7}{I_{S2}}$. Conditional on this being true, the vote only changes if the t_2 voter was the swing vote, which means that $\frac{n-1}{2}$ of the non-trading voters have negative utility on t_1 and the remaining $\frac{n-3}{2}$ voters have positive utility. This happens with probability $\binom{n-2}{\frac{n-3}{2}}(Q_1^-)^{\frac{n-1}{2}}(Q_1^+)^{\frac{n-3}{2}}$. We multiply these quantities by $2u_1$ to get the expected value for the t_1 voter of gaining an extra vote on t_1. We repeat the process to compute the expected value of giving up a vote on t_2 to get

$$\begin{aligned}
&E(\text{value of trade})\\
&\quad = 2u_1\frac{I_6 + I_7}{I_{S2}}\binom{n-2}{\frac{n-3}{2}}(Q_1^-)^{\frac{n-1}{2}}(Q_1^+)^{\frac{n-3}{2}}\\
&\qquad - 2u_2\frac{I_7 + I_8}{I_{S2}}\binom{n-2}{\frac{n-3}{2}}(Q_2^-)^{\frac{n-3}{2}}(Q_2^+)^{\frac{n-1}{2}}\\
&\quad = \frac{2}{I_{S2}}\binom{n-2}{\frac{n-3}{2}}\left(u_1(I_6 + I_7)(Q_1^-)^{\frac{n-1}{2}}(Q_1^+)^{\frac{n-3}{2}} - u_2(I_7 + I_8)(Q_2^-)^{\frac{n-3}{2}}(Q_2^+)^{\frac{n-1}{2}}\right)
\end{aligned}$$

By setting the last line of this equality to zero, we find the utility pairs (u_1, u_2) for which the t_1 voter is indifferent between trading and not trading. We see that this defines a line with slope $\frac{u_2}{u_1}$. Any utility pair with a higher u_1 or a lower u_2 will have positive utility, so a rational voting strategy will indeed be determined by regions like those shown in Fig. 1. Solving for $\theta_1 = \arctan\left(\left|\frac{u_2}{u_1}\right|\right)$, we find the value for θ_1 that controls the shape of R_1:

$$\theta_1 = \arctan\left(\frac{I_6 + I_7}{I_7 + I_8} \frac{(Q_1^-)^{\frac{n-1}{2}} (Q_1^+)^{\frac{n-3}{2}}}{(Q_2^-)^{\frac{n-3}{2}} (Q_2^+)^{\frac{n-1}{2}}}\right). \tag{A2}$$

To ensure that this is well-defined, if $I_7+I_8 = 0$, then $\theta_1 = \frac{\pi}{2} = \lim_{x\to\infty} \arctan(x)$. We leave this function undefined when the numerator and denominator are both zero; in this case, no trade will happen and so there is no change in value, positive or negative.

Repeat this process for the other seven θ_i values. Solutions to this set of equations represent a Nash equilibrium, since they are a best response to themselves. In fact, it is a strict Nash equilibrium. Any deviation from this strategy, when played against this strategy, either offers trades with negative expected value or refuses trades with positive expected value; in both cases, the deviant strategy has a lower expected payoff.

Appendix B Proof of Corollary 1

Let $(\theta_1, \ldots, \theta_8)$ be a solution to Eqs. (1)–(8). First, we show that if this point is not exactly the origin, then it must not be located near the origin.

Lemma 1 *Let* $Q_{min} = min\left\{\frac{(Q_1^-)^{\frac{n-1}{2}} (Q_1^+)^{\frac{n-3}{2}}}{(Q_2^-)^{\frac{n-3}{2}} (Q_2^+)^{\frac{n-1}{2}}}, \ldots, \frac{(Q_2^-)^{\frac{n-3}{2}} (Q_2^+)^{\frac{n-1}{2}}}{(Q_1^-)^{\frac{n-3}{2}} (Q_1^+)^{\frac{n-1}{2}}}\right\}$ *and* $\theta_{min} =$ $\arctan(Q_{min})$.

If $\theta_i > 0$ *for any* i*, then all* $\theta_i > 0$ *for all* i.

Furthermore, if any $\theta_i > 0$*, then* $\theta_1 + \theta_3 > \theta_{min}$*,* $\theta_2 + \theta_4 > \theta_{min}$*,* $\theta_5 + \theta_7 > \theta_{min}$*, and* $\theta_6 + \theta_8 > \theta_{min}$*.*

Proof Suppose without loss of generality that $\theta_1 \geq 0$.

Using Eqs. (1)–(8), $\theta_1 > 0 \implies \theta_7, \theta_8 > 0 \implies \theta_2, \theta_3, \theta_4 > 0 \implies \theta_5, \theta_6 > 0$.

By Eq. (7), either $\frac{I_1+I_2}{I_1+I_4} \geq 1$ and therefore $\theta_7 \geq \theta_{min}$, or $I_4 > I_2$, in which case $\theta_5 > \theta_{min}$.

Likewise, by Eq.(8), $\theta_8 \geq \theta_{min}$ or $I_3 > I_1$ and therefore $\theta_6 > \theta_{min}$. With the same method, we get that $\theta_1 \geq \theta_{min}$ or $\theta_3 > \theta_{min}$ and that $\theta_2 \geq \theta_{min}$ or $\theta_4 > \theta_{min}$. □

Note that this implies that if any $\theta_i > 0$, then Eqs. (1)–(8) are all well-defined, since we never have $\frac{0}{0}$.

For each $(\theta_1, \ldots, \theta_8)$, there is a best response, i.e. a set of trades that have positive value and a set of trades that have negative value. Equations (1)–(8) tell us how to define this function $BR : [0, \frac{\pi}{2}]^8 \to [0, \frac{\pi}{2}]^8$.

$$(\theta_1, \ldots, \theta_8) \mapsto \left(\arctan\left(\frac{I_6(\theta_6) + I_7(\theta_7)}{I_7(\theta_7) + I_8(\theta_8)} \frac{(Q_1^-)^{\frac{n-1}{2}} (Q_1^+)^{\frac{n-3}{2}}}{(Q_2^-)^{\frac{n-3}{2}} (Q_2^+)^{\frac{n-1}{2}}}\right), \ldots, \arctan\left(\frac{I_1(\theta_1) + I_2(\theta_2)}{I_2(\theta_2) + I_3(\theta_3)} \frac{(Q_2^-)^{\frac{n-3}{2}} (Q_2^+)^{\frac{n-1}{2}}}{(Q_1^-)^{\frac{n-3}{2}} (Q_1^+)^{\frac{n-1}{2}}}\right)\right) \tag{B3}$$

Let R be the subset of $[0, \frac{\pi}{2}]^8$ where $\theta_1+\theta_3 \geq \theta_{min}$, $\theta_2+\theta_4 \geq \theta_{min}$, $\theta_5+\theta_7 \geq \theta_{min}$, and $\theta_6 + \theta_8 \geq \theta_{min}$. R is clearly convex and compact. By the above lemma, BR maps R to R, is well-defined, and is continuous. Therefore, we can apply Brouwer's Fixed Point Theorem and are done. We have bounded away from the origin, so we know that the fixed point, which is a Nash equilibrium, is not the trivial Nash equilibrium.

Appendix C Group welfare derivations

Here, we derive bounds on the trades that provide benefits (in expectation) for the entire group, not just the vote traders. We use a modification of Eq. (A1) that includes the utility of all members of the group.

$$E(u_1, u_2) = 2E(\text{value for all voters})P(A)P(C|A) - 2E(\text{value for all voters})P(B)P(D|B) \tag{C4}$$

We can write down these terms explicitly for $u_1 > 0$ and $u_2 > 0$. The first term in the parentheses is the vote trader, the second term is the trading partner, the third term is the $\frac{n-1}{2}$ voters that don't agree with the trader, and the fourth term is the $\frac{n-3}{2}$ voters that do agree.

$$\begin{aligned}
E(u_1, u_2) = {} & \frac{I_6 + I_7}{I_{S2}} \binom{n-2}{\frac{n-3}{2}} (Q_1^-)^{\frac{n-1}{2}} (Q_1^+)^{\frac{n-3}{2}} \\
& \Bigg(2u_1 + 2\frac{1}{I_6 + I_7} \iint_{R_6 \cup R_7} x f(x, y) dx dy \\
& \quad + 2\frac{n-1}{2} \frac{1}{Q_1^-} \int_{-1}^{0} \int_{-1}^{1} x f(x, y) dy dx + 2\frac{n-3}{2} \frac{1}{Q_1^+} \int_{0}^{1} \int_{-1}^{1} x f(x, y) dy dx\Bigg) \\
& \quad + \frac{I_7 + I_8}{I_{S2}} \binom{n-2}{\frac{n-3}{2}} (Q_2^-)^{\frac{n-3}{2}} (Q_2^+)^{\frac{n-1}{2}} \cdot (-1) \cdot \\
& \Bigg(2u_2 + 2\frac{1}{I_7 + I_8} \iint_{R_7 \cup R_8} y f(x, y) dx dy \\
& \quad + 2\frac{n-3}{2} \frac{1}{Q_2^-} \int_{-1}^{0} \int_{-1}^{1} y f(x, y) dx dy + 2\frac{n-1}{2} \frac{1}{Q_2^+} \int_{0}^{1} \int_{-1}^{1} y f(x, y) dx dy\Bigg)
\end{aligned}$$

Like before, we can set this expression equal to zero and solve to find the trades with expected value zero from the trade. After removing some common terms, we get

$$\text{Quadrant 1:} \quad \frac{a_1}{c_1} u_1 + \frac{a_1 b_1}{c_1} - d_1 \geq u_2 \tag{C5}$$

where

$$a_1 = (I_6 + I_7)(Q_1^-)^{\frac{n-1}{2}} (Q_1^+)^{\frac{n-3}{2}}$$

$$b_1 = \frac{1}{I_6 + I_7} \iint_{R_6 \cup R_7} x f(x, y) dx dy + \frac{n-1}{2} \frac{1}{Q_1^-} \int_{-1}^{0} \int_{-1}^{1} x f(x, y) dy dx + \frac{n-3}{2} \frac{1}{Q_1^+} \int_{0}^{1} \int_{-1}^{1} x f(x, y) dy dx$$

$$c_1 = (I_7 + I_8)(Q_2^-)^{\frac{n-3}{2}} (Q_2^+)^{\frac{n-1}{2}}$$

$$d_1 = \frac{1}{I_7 + I_8} \iint_{R_7 \cup R_8} y f(x, y) + \frac{n-3}{2} \frac{1}{Q_2^-} \int_{-1}^{0} \int_{-1}^{1} y f(x, y) dx dy + \frac{n-1}{2} \frac{1}{Q_2^+} \int_{0}^{1} \int_{-1}^{1} y f(x, y) dx dy$$

Notice that this line is parallel to the boundary of R_1, just shifted by a factor of $\frac{a_1 b_1}{c_1} - d_1$. If this term is sufficiently negative, there may be no trades with $u_1 > 0$ and $u_2 > 0$ that have positive value for the group.

The exact same process can be repeated for all eight types of trades. We give the final expressions now.

$$\text{Quadrant 2:} \quad -\frac{a_2}{c_2}u_1 - \frac{a_2 b_2}{c_2} - d_2 \geq u_2 \tag{C6}$$

$$\text{Quadrant 3:} \quad \frac{a_3}{c_3}u_1 + \frac{a_3 b_3}{c_3} - d_3 \leq u_2 \tag{C7}$$

$$\text{Quadrant 4:} \quad -\frac{a_4}{c_4}u_1 - \frac{a_4 b_4}{c_4} - d_4 \leq u_2 \tag{C8}$$

$$\text{Quadrant 5:} \quad u_3 \geq \frac{c_5}{a_5}u_4 + \frac{c_5 d_5}{a_5} - b_5 \tag{C9}$$

$$\text{Quadrant 6:} \quad u_3 \geq -\frac{c_6}{a_6}u_4 - \frac{c_6 d_6}{a_6} - b_6 \tag{C10}$$

$$\text{Quadrant 7:} \quad u_3 \leq \frac{c_7}{a_7}u_4 + \frac{c_7 d_7}{a_7} - b_7 \tag{C11}$$

$$\text{Quadrant 8:} \quad u_3 \leq -\frac{c_8}{a_8}u_4 - \frac{c_8 d_8}{a_8} - b_8 \tag{C12}$$

The coefficients can be calculated as

$$\begin{aligned}
a_2 &= (I_5 + I_8)(Q_1^-)^{\frac{n-3}{2}}(Q_1^+)^{\frac{n-1}{2}} \\
b_2 &= \frac{1}{I_5 + I_8}\iint_{R_5 \cup R_8} xf(x, y)dxdy + \frac{n-3}{2}\frac{1}{Q_1^-}\int_{-1}^{0}\int_{-1}^{1} xf(x, y)dydx \\
&\quad + \frac{n-1}{2}\frac{1}{Q_1^+}\int_{0}^{1}\int_{-1}^{1} xf(x, y)dydx \\
c_2 &= c_1 \\
d_2 &= d_1 \\
a_3 &= a_2 \\
b_3 &= b_2 \\
c_3 &= (I_5 + I_6)(Q_2^-)^{\frac{n-1}{2}}(Q_2^+)^{\frac{n-3}{2}} \\
d_3 &= \frac{1}{I_5 + I_6}\iint_{R_5 \cup R_6} yf(x, y)dxdy + \frac{n-1}{2}\frac{1}{Q_2^-}\int_{-1}^{0}\int_{-1}^{1} yf(x, y)dxdy \\
&\quad + \frac{n-3}{2}\frac{1}{Q_2^+}\int_{0}^{1}\int_{-1}^{1} yf(x, y)dxdy \\
a_4 &= a_1 \\
b_4 &= b_1 \\
c_4 &= c_3
\end{aligned}$$

$$d_4 = c_3$$
$$a_5 = (I_3 + I_4)(Q_2^-)^{\frac{n-1}{2}}(Q_2^+)^{\frac{n-3}{2}}$$
$$b_5 = \frac{1}{I_3 + I_4}\iint_{R_3 \cup R_4} yf(x, y)dxdy + \frac{n-1}{2}\frac{1}{Q_2^-}\int_{-1}^{0}\int_{-1}^{1} yf(x, y)dxdy$$
$$+\frac{n-3}{2}\frac{1}{Q_2^+}\int_{0}^{1}\int_{-1}^{1} yf(x, y)dxdy$$
$$c_5 = (I_2 + I_3)(Q_1^-)^{\frac{n-3}{2}}(Q_1^+)^{\frac{n-1}{2}}$$
$$d_5 = \frac{1}{I_2 + I_3}\iint_{R_2 \cup R_3} xf(x, y)dxdy + \frac{n-3}{2}\frac{1}{Q_1^-}\int_{-1}^{0}\int_{-1}^{1} xf(x, y)dydx$$
$$+\frac{n-1}{2}\frac{1}{Q_1^+}\int_{0}^{1}\int_{-1}^{1} xf(x, y)dydx$$
$$a_6 = a_5$$
$$b_6 = b_5$$
$$c_6 = (I_1 + I_4)(Q_1^-)^{\frac{n-1}{2}}(q_1^+)^{\frac{n-3}{2}}$$
$$d_6 = \frac{1}{I_1 + I_4}\iint_{R_1 \cup R_4} xf(x, y)dxdy + \frac{n-1}{2}\frac{1}{Q_1^-}\int_{-1}^{0}\int_{-1}^{1} xfx, y)dydx$$
$$+\frac{n-3}{2}\frac{1}{Q_1^+}\int_{0}^{1}\int_{-1}^{1} xf(x, y)dydx$$
$$a_7 = (I_1 + I_2)(Q_2^-)^{\frac{n-3}{2}}(Q_2^+)^{\frac{n-1}{2}}$$
$$b_7 = \frac{1}{I_1 + I_2}\iint_{R_1 \cup R_2} yf(x, y)dxdy + \frac{n-3}{2}\frac{1}{Q_2^-}\int_{-1}^{0}\int_{-1}^{1} yf(x, y)dxdy$$
$$+\frac{n-1}{2}\frac{1}{Q_2^+}\int_{0}^{1}\int_{-1}^{1} yf(x, y)dxdy$$
$$c_7 = c_6$$
$$d_7 = d_6$$
$$a_8 = a_7$$
$$b_8 = b_7$$
$$c_8 = c_5$$
$$d_8 = d_5$$

This can be used to compute the probability that a random trade is beneficial for the entire group. We can use a similar approach to compute the expected value of a random trade in equilibrium for the group. There are four possible outcomes that could change the value for the group: issue t_1 passes because of the trade, t_1 fails because of the trade, t_2 passes because of the trade, and t_2 fails because of the trade. Consider the first event.

This occurs if t_1 fails by one vote among the other $n-2$ voters and the t_1 voter has positive utility on t_1 and the t_2 voter has negative utility on t_1. The product of the

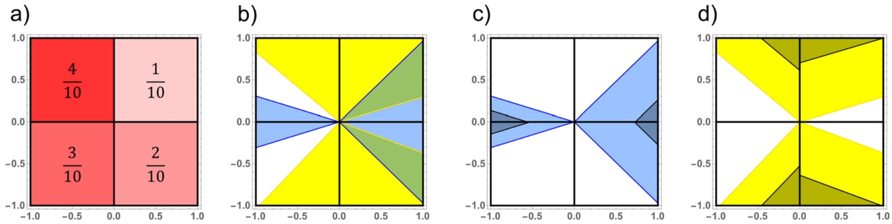

Fig. 8 The Nash equilibrium and welfare-improving trades when all voters are paired up and allowed to trade for the distribution in (**a**). This figure can be compared to Fig. 3 to see the small changes that happen when multiple trades are permitted

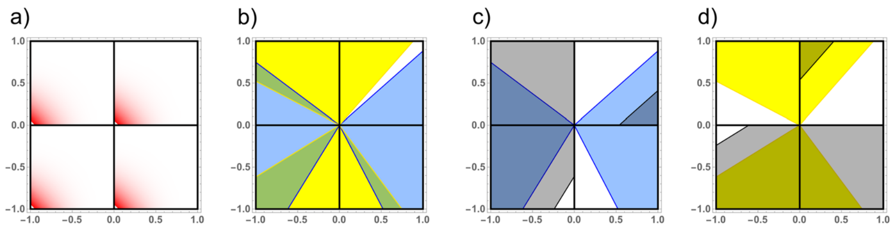

Fig. 9 The Nash equilibrium and welfare-improving trades when all voters are paired up and allowed to trade for the distribution in (**a**). This figure can be compared to Fig. 5 to see that when multiple trades are permitted, some trades located in the first quadrant become detrimental to group welfare

probabilities that each of these occurs is

$$\binom{n-2}{\frac{n-3}{2}}(Q_1^-)^{\frac{n-1}{2}}(Q_1^+)^{\frac{n-3}{2}}\cdot\frac{I_1+I_4}{I_{S1}}\cdot\frac{I_6+I_7}{I_{S2}} \tag{C13}$$

If this happens, the expected value for the group is

$$\begin{aligned}2\Bigg(&\frac{1}{I_1+I_4}\iint_{R_1\cup R_2} xf(x,y)dydx\\&+\frac{1}{I_6+I_7}\iint_{R_6\cup R_7} xf(x,y)dydx\\&+(\frac{n-1}{2})\frac{1}{Q_1^-}\int_{-1}^{0}\int_{-1}^{1} xf(x,y)dydx\\&+(\frac{n-3}{2})\frac{1}{Q_1^+}\int_{0}^{1}\int_{-1}^{1} xf(x,y)dydx\Bigg)\end{aligned} \tag{C14}$$

where the four terms in parentheses are the expected change in utility for the t_1 voter, the t_2 voter, the $\frac{n-1}{2}$ other voters who oppose t_1, and the $\frac{n-3}{2}$ other voters who support t_1, respectively.

Finally, we multiply Eqs. (C13) and (C14), repeat for the other three possible events that could change group welfare, and add them all together to get the final expected value for a random trade:

$$
\begin{aligned}
&E(\text{Value of trade for entire group})\\
&= 2\left(\binom{n-2}{\frac{n-3}{2}}(Q_1^-)^{\frac{n-1}{2}}(Q_1^+)^{\frac{n-3}{2}}\cdot\frac{I_1+I_4}{I_{S1}}\cdot\frac{I_6+I_7}{I_{S2}}\right)\\
&\quad\cdot\left(\frac{1}{I_1+I_4}\iint_{R_1\cup R_2} xf(x,y)dydx+\frac{1}{I_6+I_7}\iint_{R_6\cup R_7} xf(x,y)dydx\right.\\
&\quad\left.+\left(\frac{n-1}{2}\right)\frac{1}{Q_1^-}\int_{-1}^{0}\int_{-1}^{1} xf(x,y)dydx+\left(\frac{n-3}{2}\right)\frac{1}{Q_1^+}\int_{0}^{1}\int_{-1}^{1} xf(x,y)dydx\right)\\
&\quad-2\left(\binom{n-2}{\frac{n-3}{2}}(Q_1^-)^{\frac{n-3}{2}}(Q_1^+)^{\frac{n-1}{2}}\cdot\frac{I_2+I_3}{I_{S1}}\cdot\frac{I_5+I_8}{I_{S2}}\right)\\
&\quad\cdot\left(\frac{1}{I_2+I_3}\iint_{R_2\cup R_3} xf(x,y)dydx+\frac{1}{I_5+I_8}\iint_{R_5\cup R_8} xf(x,y)dydx\right.\\
&\quad\left.+\left(\frac{n-3}{2}\right)\frac{1}{Q_1^-}\int_{-1}^{0}\int_{-1}^{1} xf(x,y)dydx+\left(\frac{n-1}{2}\right)\frac{1}{Q_1^+}\int_{0}^{1}\int_{-1}^{1} xf(x,y)dydx\right)\\
&\quad+2\left(\binom{n-2}{\frac{n-3}{2}}(Q_2^-)^{\frac{n-1}{2}}(Q_2^+)^{\frac{n-3}{2}}\cdot\frac{I_3+I_4}{I_{S1}}\cdot\frac{I_5+I_6}{I_{S2}}\right)\\
&\quad\cdot\left(\frac{1}{I_3+I_4}\iint_{R_3\cup R_4} yf(x,y)dxdy+\frac{1}{I_5+I_6}\iint_{R_5\cup R_6} yf(x,y)dxdy\right.\\
&\quad\left.+\left(\frac{n-1}{2}\right)\frac{1}{Q_2^-}\int_{-1}^{0}\int_{-1}^{1} yf(x,y)dxdy+\left(\frac{n-3}{2}\right)\frac{1}{Q_2^+}\int_{0}^{1}\int_{-1}^{1} yf(x,y)dxdy\right)\\
&\quad-2\left(\binom{n-2}{\frac{n-3}{2}}(Q_2^-)^{\frac{n-3}{2}}(Q_2^+)^{\frac{n-1}{2}}\cdot\frac{I_1+I_2}{I_{S1}}\cdot\frac{I_7+I_8}{I_{S2}}\right)\\
&\quad\cdot\left(\frac{1}{I_1+I_2}\iint_{R_1\cup R_2} yf(x,y)dxdy+\frac{1}{I_7+I_8}\iint_{R_7\cup R_8} yf(x,y)dxdy\right.\\
&\quad\left.+\left(\frac{n-3}{2}\right)\frac{1}{Q_2^-}\int_{-1}^{0}\int_{-1}^{1} yf(x,y)dxdy+\left(\frac{n-1}{2}\right)\frac{1}{Q_2^+}\int_{0}^{1}\int_{-1}^{1} yf(x,y)dxdy\right)
\end{aligned}
\tag{C15}
$$

Appendix D Group-wide trading details

We will need to know the probability that a voter is willing to trade in either direction, so we define the following quantities:

$$J_1 = \iint_{R_1\cap R_5} f(x,y)dxdy \tag{D16}$$

$$J_2 = \iint_{R_2\cap R_6} f(x,y)dxdy \tag{D17}$$

$$J_3 = \iint_{R_3 \cap R_7} f(x, y)dxdy \tag{D18}$$

$$J_4 = \iint_{R_4 \cap R_8} f(x, y)dxdy \tag{D19}$$

A voter v who is deciding what trade to offer needs to know the probability that the trade will change the outcome of the vote. Consider a random other voter w in the population that is not paired with v. First, let us determine the probability that w ultimately casts a vote in support of t_1.

w initially supports t_1 with probability Q_1^+. With probability $\frac{n-3}{n-2}$, w has the opportunity to trade.

There is also a chance that w gives away their vote to someone that opposes t_1. This happens if w's utilities fall in R_5 or R_8 and w's trading partner is in R_2 or R_3, which happens with probability $(I_5 + I_8)(I_2 + I_3)$ which we subtract from the initial probability. However, if w and their partner are both willing to trade in either direction, then w only gives up their vote half the time, so the probability that w has a positive utility on t_1 but trades it away is $(I_5 + I_8)(I_2 + I_3) - \frac{1}{2}(J_1 + J_4)(J_2 + J_3)$.

Similarly, w could also vote for t_1 if they initially oppose the issue but give away their vote to someone who supports it. Therefore, we must add $(I_6 + I_7)(I_1 + I_4) - \frac{1}{2}(J_2 + J_3)(J_1 + J_4)$.

When we add all three of these terms together, we have the probability that voter w votes in favor of t_1:

$$\mathcal{Q}_1^+ = Q_1^+ + \frac{n-3}{n-2}\left((I_6 + I_7)(I_1 + I_4) - \frac{1}{2}(J_2 + J_3)(J_1 + J_4) - (I_5 + I_8)(I_2 + I_3) + \frac{1}{2}(J_1 + J_4)(J_2 + J_3)\right)$$

The J terms cancel out, and we are left with

$$\mathcal{Q}_1^+ = Q_1^+ + \frac{n-3}{n-2}\Big((I_6 + I_7)(I_1 + I_4) - (I_5 + I_8)(I_2 + I_3)\Big). \tag{D20}$$

A similar process gives the other necessary probabilities.

$$\mathcal{Q}_1^- = Q_1^- + \frac{n-3}{n-2}\Big((I_5 + I_8)(I_2 + I_3) - (I_6 + I_7)(I_1 + I_4)\Big) \tag{D21}$$

$$\mathcal{Q}_2^+ = Q_2^+ + \frac{n-3}{n-2}\Big((I_3 + I_4)(I_5 + I_6) - (I_1 + I_2)(I_7 + I_8)\Big) \tag{D22}$$

$$\mathcal{Q}_2^- = Q_2^- + \frac{n-3}{n-2}\Big((I_1 + I_2)(I_7 + I_8) - (I_3 + I_4)(I_5 + I_6)\Big) \tag{D23}$$

Now, when computing the Nash equilibrium using Eqs. (1)–(8), we replace the Q terms with $\mathcal{Q}$ terms from Eqs. (D20)–(D23) to approximate the probability of being the pivotal vote after all other voter pairs have traded. We can make this same replacement in the a_i and c_i terms when computing group welfare implications. Note that we do not replace the Q terms in the b_i and d_i equations, since those are computing welfare, not the probability of being a pivotal vote.

When f is point-symmetric around the origin, it still has the naive Nash equilibrium. All I_is are equal, and almost all terms in Eqs. (D20)–(D23) cancel out. There is no change to the equilibrium between the myopic model and the model that allows all voters to trade votes.

We end by recreating Figs. 3 and 5 with the new equilibria found when all voters are paired up and allowed to trade. The changes are modest, and most pronounced when the distribution f is designed to promote trades in one direction.

References

Aksoy D (2012) Institutional arrangements and logrolling: evidence from the European union. Am J Polit Sci 56(3):538–552. https://doi.org/10.1111/j.1540-5907.2012.00574.x

Bouton L, Conconi P, Pino F, Zanardi M (2021) The tyranny of the single-minded: guns, environment, and abortion. Rev Econ Stat 103(1):48–59. https://doi.org/10.1162/rest_a_00897

Buchanan JM, Tullock G (1962) The calculus of consent: logical fundations of contitutional democracy. The University of Michigan Press

Câmara O, Eguia JX (2017) Slicing and bundling. J Polit 79(4):1460–1464. https://doi.org/10.1086/692807

Casella A (2011) Storable votes: protecting the minority voice. Oxford University Press

Casella A, Macé A (2021) Does vote trading improve welfare? Ann Rev Econ 13(1):57–86. https://doi.org/10.1146/annurev-economics-081720-114422

Casella A, Palfrey T (2019) Trading votes for votes: a dynamic theory. Econometrica 87(2):631–652. https://doi.org/10.3982/ecta15940

Casella A, Palfrey TR (2021) Trading votes for votes: a laboratory study. Games Econ Behav 125:1–26. https://doi.org/10.1016/j.geb.2020.10.004

Casella A, Llorente-Saguer A, Palfrey TR (2012) Competitive equilibrium in markets for votes. J Polit Econ 120(4):593–658. https://doi.org/10.1086/667988

Casella A, Palfrey T (2018) Trading votes for votes. A dynamic theory. Department of Economics discussion papers. https://doi.org/10.7916/d8-jmtm-a154

Cohen L, Malloy C (2014) Friends in high places. Am Econ J Econ Pol 6(3):63–91. https://doi.org/10.1257/pol.6.3.63

Coleman JS (1966) The possibility of a social welfare function. Am Econ Rev 56(5):1105–1122

Ferejohn JA (1974) Sour notes on the theory of vote trading. Working papers 41, California Institute of Technology, Division of the Humanities and Social Sciences. https://EconPapers.repec.org/RePEc:clt:sswopa:41

Goeree JK, Zhang J (2017) One man, one bid. Games Econ Behav 101:151–171. https://doi.org/10.1016/j.geb.2016.10.003

Holt CA, Roth AE (2004) The Nash equilibrium: a perspective. Proc Natl Acad Sci 101(12):3999–4002. https://doi.org/10.1073/pnas.0308738101

Iaryczower M, Oliveros S (2016) Power brokers: middlemen in legislative bargaining. J Econ Theory 162:209–236. https://doi.org/10.1016/j.jet.2015.12.011

Jackson MO, Sonnenschein HF (2007) Overcoming incentive constraints by linking decisions. Econometrica 75(1):241–257. https://doi.org/10.1111/j.1468-0262.2007.00737.x

Jacobs E, Christensen PN, Prislin R (2009) Of practicalities and perspective: what is fair in group decision making? J Soc Issues 65(2):383–407. https://doi.org/10.1111/j.1540-4560.2009.01605.x

Jones MI, Chervenak M, Christakis NA (2023) It is easy for multi-issue bundles to advance anti-democratic agendas

Lalley SP, Weyl EG (2018) Quadratic voting: How mechanism design can radicalize democracy. AEA Pap Proc 108:33–37. https://doi.org/10.1257/pandp.20181002

May KO (1952) A set of independent necessary and sufficient conditions for simple majority decision. Econometrica 20(4):680. https://doi.org/10.2307/1907651

McPherson L (2019) Ocasio-Cortez joins most Republicans in voting against House Democratic bills to reopen government. https://rollcall.com/2019/01/23/ocasio-cortez-joins-most-republicans-in-voting-against-house-democratic-bills-to-reopen-government/

Miller NR (1977) Logrolling, vote trading, and the paradox of voting: a game-theoretical overview. Public Choice 30(1):51–75. https://doi.org/10.1007/bf01718818

Mueller DC (1967) The possibility of a social welfare function: comment. Am Econ Rev 57(5):1304–1311

Nash J (1951) Non-cooperative games. Ann Math 54(2):286. https://doi.org/10.2307/1969529

Park RE (1967) The possibility of a social welfare function: comment. Am Econ Rev 57(5):1300–1304

Ramzy A (2017) McCain's vote provides dramatic moment in 7-year battle over Obamacare. The New York Times. https://www.nytimes.com/2017/07/28/us/politics/john-mccain-vote-trump-obamacare.html

Riker WH, Brams SJ (1973) The paradox of vote trading. Am Polit Sci Rev 67(4):1235–1247. https://doi.org/10.2307/1956545

Samsonov A, Solé-Ollé A, Xefteris D (2023) Logrolling and the power of single-mindedness. SSRN 4549869

Smart DR (1980) Fixed point theorems. Cambridge University Press

Stratmann T (1995) Logrolling in the U.S. congress. Econ Inq 33(3):441–456. https://doi.org/10.1111/j.1465-7295.1995.tb01873.x

Studies ANE (2021) ANES 2020 time series study full release [dataset and documentation]. Technical report, ANES. www.electionstudies.org

Tsakas N, Xefteris D, Ziros N (2020) Vote trading in power-sharing systems: a laboratory investigation. Econ J 131(636):1849–1882. https://doi.org/10.1093/ej/ueaa105

Węglarczyk S (2018) Kernel density estimation and its application. ITM Web Conf 23:00037. https://doi.org/10.1051/itmconf/20182300037

Xefteris D, Ziros N (2017) Strategic vote trading in power sharing systems. Am Econ J Microecon 9(2):76–94. https://doi.org/10.1257/mic.20150254