# "Reverse Gerrymandering": Manipulation in Multi-Group Decision Making

**Omer Lev**
Ben-Gurion University of the Negev
Beersheba, Israel
omerlev@bgu.ac.il

**Yoad Lewenberg**
The Hebrew University
of Jerusalem, Israel
yoadlew@cs.huji.ac.il

## Abstract

District-based manipulation, or gerrymandering, is usually taken to refer to agents who are in fixed location, and an external division is imposed upon them. However, in many real-world setting, there is an external, fixed division – an organizational chart of a company, or markets for a particular product. In these cases, agents may wish to move around ("reverse gerrymandering"), as each of them tries to maximize their influence across the company's subunits, or resources are "working" to be allocated to areas where they will be most needed.

In this paper we explore an iterative dynamic in this setting, finding that allowing this decentralized system results, in some particular cases, in a stable equilibrium, though in general, the setting may end up in a cycle. We further examine how this decentralized process affects the social welfare of the system.

## Introduction

In October 2016, just before the US presidential elections, the New York Times published an article titled "Go Midwest, Young Hipster" (MacGillis 2016). In it, the author emphasized how crucial elections could be won if supporters of a party would move from where they are concentrated to areas where they are more sparse. While this may be a ridiculous suggestion for national elections and the large number of people involved (and the author did not claim otherwise), in other settings such a proposition is not as preposterous.

An obvious, more realistic, setting of a similar idea in organizations is that of change agents (Rogers 1962). When trying to change a corporate behavior, change agents move around the organization, trying to form coalitions to bring about some form of a change they are advocating for, so that it encompasses the whole company. Many people in large bureaucratic organizations are familiar with the sets of committees in which decision are made. In such organizations, people seeking to push an idea (or a person) try to gain influence in multiple committees.

A wider angle reveals this to be, more generally, a model of decentralized resource allocation. Understanding what would be the outcome if many of our resources were autonomous and could attempt to allocate themselves to their

optimal destination, while their ultimate goal is succeeding in as many destinations as possible. This is a desirable feature in many industrial Internet of Things (IoT) applications, but more concretely, one can think of a medical resource (or robot) deciding which of multiple field operations it should go to without central direction. It can be a batch of advertising resources directed to appropriate product/market without the company headquarters deciding on it; or it can be grassroots political campaigners going to volunteer in the place they think they will be able to challenge rival campaigns, with no need for their candidate telling them where to go.

All these settings have to do with agents realizing whether they are needed in their current location, and if not – looking for a better place they can contribute in, possibly only seeing the places that are reachable by them from their area. In this work, we are less interested in how they might reach their destination, but in the fundamental question of what are the effects of using a decentralized system, instead of a hierarchal, centralized one.

Since this setting is, fundamentally, equivalent to an attempt by supporters of some candidate – which can be a person, a policy, or an idea – to spread themselves so they can encourage other supporters and together form a majority in many sub-units or districts, we will use the terminology of elections and voting. Voting and elections have been widely studied in artificial intelligence as natural tools for aggregating preferences of independent, self-interested agents in multiagent systems. Agents have a preference for some candidate and can see if their involvement in their district is affecting some change (we use "district" as a shorthand for a subunit, we do not employ geographical considerations). If it does not, they may examine other possible districts, and they can move to and be more helpful in (this is a similar setting to that of (Bervoets and Merlin 2012), though they were looking for novel voting techniques). Each agent operates independently of others, raising the question of whether this process can converge; what are its outcomes when agents move about; and what structural issues influence this outcome.

## Related Work

As a motivation, resource allocation lies at the core of much economic/Algorithmic Game Theory (AGT) research, and in particular, a desire to avoid a central, all-knowing, entity

that divides the resources. Indeed, both auctions and pricing, the main components of the AGT toolkit, are ways in which resources can be allocated more efficiently, without requiring a centralized decision-making process (Krishna 2002; Nisan et al. 2007). However, in cases like ours, where the "resources" may be autonomous, these approaches are not practical.

Discussion of gerrymandering and district division and the effects of population distribution on them have been growing significantly in the past several years in the computer-science community (Bachrach et al. 2016; Lewenberg, Lev, and Rosenschein 2017; Cohen-Zemach, Lewenberg, and Rosenschein 2018; Borodin et al. 2018). But our particular model is more similar to hedonic games (Aziz and Savani 2016): agents are choosing groups to participate in to increase their utility. However, while in hedonic games agents care for the success of their coalition, in our model they are concerned with the overall outcomes of every coalition in the game. Probably the closest paper to our setting is (Bervoets and Merlin 2012), which, like our setting, moves agents between different gerrymandered districts. However, it seeks to find mechanism that prevent such movement, while we accept it as part of the model.

Somewhat related to our model, (van Bevern et al. 2015) modeled a setting motivated by gerrymandering, in which agents are moved according to a fixed rule in a graph in order to "reorganize" them. In a certain manner, the work of (van Bevern et al. 2015) is a less general case of our model, as it follows only a particular pattern.

We employ an iterative framework, which bears similarity to an extensive line of research, such as (Meir et al. 2010; Lev and Rosenschein 2016; Grandi et al. 2013; Meir 2016; Rabinovich et al. 2015). While the existence of districts causes some complications beyond regular iterative voting, we are less interested in the change of voters' actual vote, but on their manipulation by changing their district which is a type of iterative manipulation not previously considered.

Calculating the optimal allocation is complicated by the "resources" in this case being the agents with the utility. Finding an optimal allocation using an iterative process (which we use) has been shown to be in most cases NP-hard (Aziz et al. 2016).

## Model

We have a set $C$ of $m$ candidates $\{c_1, \ldots, c_m\}$, and a set $V$ of $n$ voters $\{v_1, \ldots, v_n\}$. Each $v_i \in V$ is associated with a preference order of candidates (i.e., if $\pi(C)$ is the set of linear orders of the elements of $C$, each $v_i$ is associated with its preference order, an element in $\pi(C)$). For any number of voters $1 \le s \le n$ we have a voting rule $f : (\pi(C))^s \to 2^C \setminus \emptyset$, and from here we have two options for tie-breaking:

**deterministic** Using tie-breaking rule $t : 2^C \to C$, so that the outcome of the election is $t \circ f$.

**fractional** Each candidate in the winning set receives an equal fraction of the win (the meaning of this will become apparent shortly).

In this paper, we will focus on plurality – the most commonly used voting rule – in which each vote gives a point

to its top-ranked candidate, and the candidate with the most points is the winner. The model can easily be extended to any other voting rule.

In addition to this basic voting setting, we also have a set $\mathcal{D}$ of $k$ districts $\{D_1, \ldots, D_k\}$. There is an initial division of voters to the different districts, which can be considered as $g^0 : V \to \mathcal{D}$. We will denote the initial voters of district $D_i$ as $V_i^0 = \{v \in V | g^0(v) = D_i\}$. From the voters of $(V_1, \ldots, V_k)$ one can construct a district scoring vector $p \in \mathbb{R}_+^m$, giving each candidate a point (or its fraction, in the fractional case) if they have won a district. An inter-district voting rule takes this vector $p$ and outputs a set of overall winners: $\tilde{f} : \mathbb{R}_+^m \to 2^C$ (with $f$ taking $p \in \mathbb{R}_+^m$ as its input). Once again, a tie-breaking rule can be used, to make the overall election winner $\tilde{t} \circ \tilde{f}$ (here we do not have a fractional variant, as we seek to end with an overall winner). In this paper (as in other gerrymandering papers (Bachrach et al. 2016; Lewenberg, Lev, and Rosenschein 2017; Borodin et al. 2018)), we shall use the plurality rule as our inter-district voting rule, making the winner the candidates which won the most districts.

Each voter $v_i$ has a utility function which gets as an input the district scoring vector $u_i : (\mathbb{R}_+^m) \to \mathbb{R}$. We are going to consider two types of utility functions for our agents:

**global** These agents are only concerned with the ultimate outcome. That is, the final outcome of $\tilde{t} \circ \tilde{f}$ and where the winner is located in their own preference order. This type of utility is more appropriate for presidential systems, where the only thing that matters is the ultimate outcome of the winner.

**lexicographic** These agents consider the input district scoring vector. They re-order it according to their own preference order, constructing $(score(c_1^i), \ldots, score(c_m^i)) \in \mathbb{R}_+^m$ such that $c_1^i$ is voter $v_i$'s top ranked candidate, $c_2^i$ the 2nd favorite and so on. The utility function works lexicographically, such that $(x_1, x_2, \ldots, x_m) \succeq_i (x_1', x_2', \ldots, x_m')$ if there is $1 \le j \le m$ such that for all $j' < j$, $x_{j'} = x_{j'}'$ and $x_j > x_j'$. This type of utility is more appropriate for parliamentary system, where agents seek as many representatives as they can of parties they favor.

## Iterative Process

At time $h$ an agent can decide that they wish to move from their district to another. An agent $v$ will move from $D_i$ to $D_j$ at time $h$ if $u_i(p^h) > u_i(p^{h-1})$ ($p^0$ is the starting point) for the district score vector $p^{h-1}$ created from $f(V_1^{h-1}) \ldots, f(V_i^{h-1}), \ldots, f(V_j^{h-1}), \ldots, f(V_k^{h-1})$ and $p^h$ created from $f(V_1^{h-1}) \ldots, f(V_i^{h-1} \setminus \{v\}), \ldots, f(V_j^{h-1} \cup \{v\}), \ldots, f(V_k^{h-1})$ (this shows for $j > i$, but this is not a condition).

If there are multiple changes available to $v$, the voter will choose the one that maximizes $u_i$ (i.e., a best-response).

That is, an agent will move if they can improve the overall result of the election, regardless of their effect in their previous district. As in other iterative voting papers (from

(Meir et al. 2010) onwards), we will allow only one agent to change their district at each time point.

In various settings (e.g., workplaces), there may be limitations on agents leaving an organizational sub-unit understaffed, or arriving at an already crowded sub-unit. Hence, we will explore the implications of limitations on the size of districts. We will do so by imposing upper and lower thresholds on district sizes, so that for any $i$, $b_- \leq |D_i| \leq b^+$.[1] That is, if some district is at its lower threshold, no agent will be allowed to leave. If a district is at its upper threshold, no agent will be allowed to join.

Beyond the manipulation of agents moving between districts, they could combine such a move with a strategic changing of their announced vote. This applies less to resource allocation settings (a resource can not change its essence), but more to settings that involve agents pushing for some change (such as in an organizational setting). We do not allow agents to change their vote while in the same district (that would basically reduce to the iterative voting setting in any case), but they can change their vote when changing district. Thus, an agent $v \in V$ currently in $D_i$ voting with a preference order $v' \in \pi(C)$ will move to district $D_j$ with a vote $v''$ in time $h$ if $u_i(p^h) > u_i(p^{h-1})$ for the district score vector $p^{h-1}$ created from $f(V_1^{h-1}) \ldots, f(V_i^{h-1}), \ldots, f(V_j^{h-1}), \ldots, f(V_k^{h-1})$ and $p^h$ created from $f(V_1^{h-1}) \ldots, f(V_i^{h-1} \setminus \{v'\}), \ldots, f(V_j^{h-1} \cup \{v''\}), \ldots, f(V_k^{h-1})$.

Ultimately, what we wish to know, is if this iterative process converges – does it end in a state in which no agent wishes to change their location (such a state, when no agent wishes to change their strategy, is called a *Nash equilibrium*).

To recap, we have several variables in flux:

**tie-breaking** Either deterministic (one winner per district) or fractional. This serves to model different district settings. When, suppose, each district selects a delegate, deterministic tie-breaking makes sense. In cases of, for example, money allocation, a fractional tie-breaking makes more sense.

**utility** Either globalists (caring only for the global result) or lexicographic. Globalists makes sense, again, for cases of representation when, ultimately, the district process will select an overall winner. When the districts are making a monetary allocation, for example, someone fighting for a particular cause, first wants their cause to get the most it can, and only secondary to that, to allocate money to more minor concerns.

**strategic actions** Either just moving between districts, or also vote-strategic. This depends on the information structure – if all know a particular agent as a supporter of a particular candidate, they may not be able to switch their allegiance easily.

---

[1]This exists, technically, in many US states' congressional districts, which are forbidden by some states to vary in size beyond a limited amount, such as 1%. Of course, in such cases people are not prevented from leaving, but rather the district will have to be redrawn following the next census.

Table 1: Agent variables and their possible values

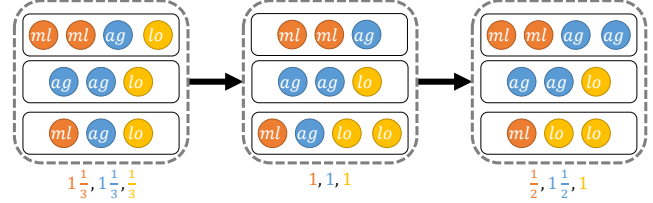| Category | Option I | Option II |
|---|---|---|
| **Tie-breaking** | deterministic | fractional |
| **Utility** | globalist | lexicographi |
| **Strategy** | district move only | vote-strategic |



Figure 1: Example 1 — The $lo$ agent (represented as a yellow circle) in the top unit moves to the bottom unit, winning it over. This makes it possible for the $ag$ agent in the bottom unit to move to the top. The number indicate each groups score

Naturally, in addition to that we have as parameters the environment's variables: number of voters ($n$), candidates ($m$), districts ($k$), and bounds on district sizes ($b_-, b^+$), the initial district allocation, etc.

**Example 1.** *Consider a multi-unit technology company. Each unit is in charge of one product, and each product can be developed using one of the three available technologies: machine learning (ML), algorithmic game theory (AGT), and logic. There are 10 developers in the company and 3 units, and in every unit there must be either 3 or 4 developers. The developers want that as many products as possible will be developed using their favorite technology. In case of a tie in a unit, the product will be developed using hybrid technology (i.e., fractional districts).*

*3 developers in the company are machine learning enthusiasts ($ml$), 3 believe in algorithmic game theory ($ag$) and the other 4 support logic ($lo$) (the full preferences order is not needed for this example). The initial company partition is $V_1^0 = \{ml, ml, ag, lo\}$ $V_2^0 = \{ag, ag, lo\}$ and $V_3^0 = \{ml, ag, lo\}$. In the initial state, the first product is developed using ML techniques, the second product is developed using AGT and third combines all three approaches. As can be seen in Figure 1, in the initial state, only a developer from the first unit can move to other units, due to the size constraints, and indeed the developer that supports "lo" wants to move to the third unit, as this move will result in the third product be developed using logic approaches. That is, $V_1^1 = \{ml, ml, ag\}$ $V_2^1 = \{ag, ag, lo\}$ and $V_3^1 = \{ml, ag, lo, lo\}$. Now, the developer that supports "a" wants to move to the first unit and cause the first product to partly be developed using algorithmic game theory approaches and $V_1^2 = \{ml, ml, ag, ag\}$, $V_2^2 = \{ag, ag, lo\}$ and $V_3^2 = \{ml, lo, lo\}$. Assuming that the machine learning developers from $V_1^2$ prefer $lo \succ ag$, then no developer has an incentive to switch units and therefore the third state is stable.*

# Convergence

The first step in our analysis is to try and see if there are cases in which the iterative process will converge.

Table 2: Summary of the cases our theorems prove for $b^+ - b^- = 1$. For $b^+ - b^- > 1$, see Theorem 4 (1), (2).

|  | globalist utility | lexicographic utility |
|---|---|---|
| **fractional tie-breaking** | Theorem 4 (3) *no convergence* | Theorem 3 *converges* |
| **deterministic tie-breaking** | Theorem 1 *converges* | Theorem 2 *converges* |

**Theorem 1.** *If $b^+ - b^- = 1$ and agents are globalists and not vote-strategic and districts are not fractional, the iterative process will always converge to a Nash equilibrium, in which no agent wishes to change.*

*Proof.* First, we will show that under the theorem's conditions, the score of each district's winner is weakly monotonically increasing with the iterative process. If a district's size is $b^-$, no agent can leave, and if a new agent arrives, it will do so to change the winner outcome, and hence the new winner's score will be the same as the previous one (since all of its voters are still there), or strictly more. If the district size is $b^+$, at each point, the last transfer of a voter to the district (which made its size $b^+$) was needed to make this candidate win (otherwise the voter would not have moved). This means no supporters of the current winner can leave the district, therefore the score of the winner of the district will never decrease.

Assume for contradiction that the process does not converge and results in a cycle. Since each district's maximal score is weakly monotonically increasing, districts' maximal scores during the cycle always stay the same. Of all the voters that move let $C'$ be the set of candidates they vote for, and let $c' \in C'$ be the candidate of that set that is highest ranked in the tie-breaking rule. Let $v$ be a voter that votes for $c'$. Once $v$ moves to a new district and $c'$ wins there, there is no way for any other candidate to dislodge $c'$. So $v$ has no reason to ever continue moving in the cycle, contradicting its existence. □

**Theorem 2.** *If $b^+ - b^- = 1$ with lexicographic, not vote-strategic voters and districts are not fractional the iterative process will always converge to a Nash equilibrium, in which no agent wishes to change.*

*Proof.* Assume for contradiction that there are cycles and let $\mathcal{S} = S^1 \to S^2 \to \cdots \to S^l \to S^1$ be a cycle, where each $S \in \mathcal{S}$ is a state that holds all the relevant information.

First, observe that in every district one voter leaves, one voter joins and so on, because $b^+ - b^- = 1$. Further, voters would join districts only if they could influence the result in their new districts. For every state $S \in \mathcal{S}$, district $D \in \mathcal{D}$ and candidate $c \in C$ let $s(c, D, S)$ be the number of votes in district $D$ at state $S$ for candidate $c$. Let $w(D, S) \in C$ be the winner in district $D$ at state $S$.

Let $C' \subset C$ be the set of candidates that become victorious during the cycle $\mathcal{S}$, and let $z \in C'$ be the lowest candidate in $C'$ according to the tie breaking rule. Let $D \in \mathcal{D}$ be a district in which $z$ becomes victorious during $\mathcal{S}$ and let $(v_1, \ldots, v_{\ell_D}, v_1)$ be the ordered set of voters that joined district $D$ during $\mathcal{S}$ and let $(r_1, \ldots, r_{\ell_D}, r_1)$ be the ordered set voters that left district $D$ during $\mathcal{S}$. Assume without loss of generality, that first voter $v_1$ joined district $D$ and then voter $r_1$ left district $D$. Further, assume without loss of generality that voter $v_1$ supports candidate $z$. Consider the move $S^1 \overset{v_1 \to D}{\to} S^2 \in \mathcal{S}$ when voter $v_1$ joined district $D$.

Now, there are two possible cases:

**Case I: Voter $r_1$ does not support candidate $z$** Since candidate $z$ is the lowest ranked candidate, according to the tie breaking rule, who became victorious in district $D$ during the cycle $\mathcal{S}$, it must hold that $s(z, D, S^2) \geq s(c, D, S^2) + 1$,[2] for every other candidate $c \in C$, $c \neq z$. As voter $r_1$ does not support candidate $z$, before $v_2$ joined district $D$, the winner was candidate $z$, and therefore after $v_2$ joined district $D$ in state $S^h$ there were exactly two candidates (including $z$) with $s(\cdot, D, S^h) \geq s(z, D, S^2)$. Because voters alternately join and leave district $D$ and voters would join a district only if they could influence the result, it holds that in every $S \in \mathcal{S}$ there is a candidate with $s(\cdot, D, S) \geq s(z, D, S^2)$. However, if in state $S^1$ there was a candidate with $s(c, D, S^1) \geq s(z, D, S^2)$ then voter $v_1$ could not make candidate $z$ victorious.

**Case II: Voter $r_1$ does support candidate $z$** When $r_1$ left district $D$ they joined district $D'$. If the next voter to leave district $D'$ does not support candidate $z$ then we are back at Case I. Therefore, there is a cycle of $z$'s supporters that join and leave districts. Consider voter $v$ that left district $D$ and join district $D'$ in the move $S \overset{D \to v \to D'}{\to} S'$. As the last voter to join district $D$ also supports candidate $z$ it must hold that $w(D, S) = z$, $w(D', S) \neq z$, $w(D, S') \neq z$ and $w(D', S') = z$.[3] Let $w(D', S) = c$ and $w(D, S') = c'$, because $v$ moved from $D$ to $D'$ it must hold $c \npreceq_v c'$, i.e., voter $v$ prefers candidate $c'$ over candidate $c$. Consider the directed graph $G_v = (C, E_v)$ where the nodes are the candidates and there is a directed edge $(c, c') \in E_v$ if and only if there is a move $S \overset{D \to v \to D'}{\to} S' \in \mathcal{S}$ such that $w(D', S) = c$ and $w(D, S') = c'$. Because $\mathcal{S}$ is a cycle, after $v$ joins a district they leave the district and therefore the in-degree of a node $c \in G_v$ equals to its out-degree, and thus there is a cycle in $G_v$. However, for every directed edge $(c, c') \in E_v$ it holds that $c \npreceq_v c'$, reaching a contradiction.

□

---

[2] $s(x, y, z)$ denotes score of candidate $x$ in district $y$ for state $z$.
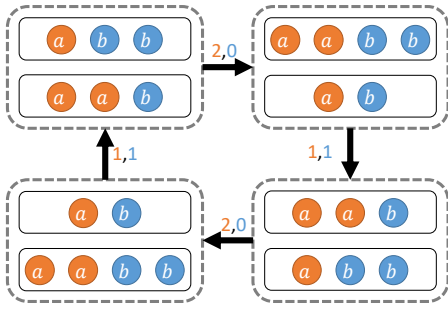[3] $w(x, y)$ denotes the winner in district $x$ in state $y$.

Figure 2: A cycle occurs, assuming districts are deterministic tie-breaking in each district is in $a$'s favor, and either agents are lexicographic, or the global tie-breaking is for $b$. On each edge is the score vector induced by this move.

**Theorem 3.** *If $b^+ - b_- = 1$ with fractional districts and lexicographic, not vote-strategic voters, the iterative process will always converge to a Nash equilibrium, in which no agent wishes to change.*

*Proof.* Again, assume for contradiction that there are cycles and let $\mathcal{S} = S^1 \to S^2 \to \cdots \to S^l \to S^1$ be a cycle.

As before, note that in every district one voter leaves, one voter joins and so on; and voters would join districts only if they could influence the result in their new districts.

Now there are three possible cases:

**Case I:** This is the case if during the cycle at some state $S \in \mathcal{S}$ for some $D \in \mathcal{D}$: $|\arg\max_{c \in C} s(c, D, S)| = 1$. Let $c^* \in C$, $D \in \mathcal{D}$, $S^1 \overset{v \to D}{\to} S^2 \in \mathcal{S}$ such that $\arg\max_{c \in C} s(c, D, S^2) = \{c^*\}$ and in this case I, the next voter to leave district $D$ does not support candidate $c^*$. Let voter $v \in V$ be the voter that joined $D$ in the move $S^1 \overset{v \to D}{\to} S^2$, let voter $v' \in V$ be the next voter to join $D$ in the move $S^h \overset{v' \to D}{\to} S^{h+1} \in \mathcal{S}$, and let voter $u$ be the voter that left $D$ after $v$ has joined and before $v'$ has joined. As voters only join districts if they can influence the results, and voter $u$ did not support candidate $c^*$ it holds that there are two candidates in state $S^{h+1}$ with $s(\cdot, D, S^{h+1}) = s(c^*, D, S^2)$, and therefore for every $S \in \mathcal{S}$ there exists a candidate with $s(\cdot, D, S) = s(c^*, D, S^2)$. It is a contradiction to the fact that before $v$ joined district $D$, $\max_{c \in C} s(c, D, S^1) = s(c^*, D, S^2) - 1$.

**Case II:** This is the case if during the cycle at some state $S \in \mathcal{S}$ for some $D \in \mathcal{D}$: $|\arg\max_{c \in C} s(c, D, S)| = 1$, however (unlike case I), this is the case if at any point that $\arg\max_{c \in C} s(c, D, S) = \{c^*\}$ for $c^* \in C$, $D \in \mathcal{D}$, $S \in \mathcal{S}$ the next voter to leave district $D$ supports candidate $c^*$. In that case, let $c^* \in C$, $D \in \mathcal{D}$, $S^1 \overset{v \to D}{\to} S^2 \in \mathcal{S}$ such that $\arg\max_{c \in C} s(c, D, S^2) = \{c^*\}$ and let $u \in V$ be the next voter to leave district $D$ and joined district $D'$ in the move $S^h \overset{D \to u \to D'}{\to} S^{h+1} \in \mathcal{S}$. As voter $u$ supports candidate $c^*$, and $\arg\max_{c \in C} s(c, D, S^h) = \{c^*\}$ it must hold that $\arg\max_{c \in C} s(c, D', S^{h+1}) = \{c^*\}$ and the next voter to leave district $D'$ also supports candidate
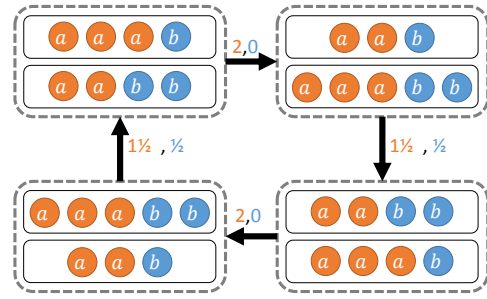


Figure 3: A cycle occurs, assuming districts are fractional and agents are lexicographic. On each edge is the score vector for the candidates induced by this move.

$c^*$. Hence, there is a cycle of $c^*$'s supporters that leave join and districts. Similarly to the proof of Theorem 2 (case II), because voters leave their district only if they strictly prefer the new outcome, this dynamic cannot cycle.

**Case III:** This is the case if during the cycle $\mathcal{S}$, for every $S \in \mathcal{S}$, $D \in \mathcal{D}$: $|\arg\max_{c \in C} s(c, D, S)| > 1$. In that case, when a voter $v \in V$ that supports candidate $c^*$ leaves district $D$ and joins district $D'$ in the move $S^1 \overset{D \to v \to D'}{\to} S^2 \in \mathcal{S}$ it must hold that $c^* \in \arg\max_{c \in C} s(c, D, S^1)$ and $c^* \in \arg\max_{c \in C} s(c, D', S^2)$ and in addition, due to the fractional character of districts,

$$\left| \arg\max_{c \in C} s(c, D', S^2) \right| < \left| \arg\max_{c \in C} s(c, D, S^1) \right| \quad (1)$$

also note that

$$\left| \arg\max_{c \in C} s(c, D, S^1) \right| = \left| \arg\max_{c \in C} s(c, D, S^2) \right| + 1 \quad (2)$$

and

$$\left| \arg\max_{c \in C} s(c, D', S^1) \right| = \left| \arg\max_{c \in C} s(c, D', S^2) \right| - 1. \quad (3)$$

Consider the following potential function[4]

$$\phi(S) = \sum_{D \in \mathcal{D}} \left| \arg\max_{c \in C} s(c, D, S) \right|^2 \quad (4)$$

due to Equations 1, 2, 3 and 4 it holds that $\phi(S^1) > \phi(S^2)$, which cannot happen in a cycle, as games with potential functions are guaranteed to converge to a Nash equilibrium (Monderer and Shapley 1996). $\square$

These convergence proofs are tight, in the sense that almost any deviation from them, results in a setting that does not converge.

---

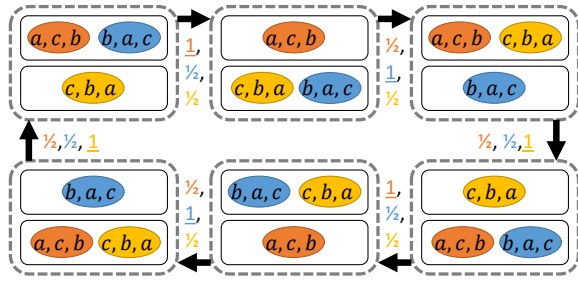[4]Definition of which can be found in (Maschler, Solan, and Zamir 2013).

Figure 4: A cycle occurs, assuming districts are fractional and agents are globalist. On each edge is the score vector for the candidates induced by this move.

**Theorem 4.** *The following cases are not guaranteed to converge, and contain cases with cycles:*

1. *If $b^+ - b_- \geq 2$ when districts are deterministic (not fractional) and agents are vote-strategic or not.* [**bounds of Theorem 1 and Theorem 2**]

2. *If $b^+ - b_- \geq 2$ when districts are fractional, and agents are lexicographic and either vote-strategic or not.* [**bound of Theorem 3**]

3. *If $b^+ - b_- = 1$ when districts are fractional and voters are globalist.* [**bound of Theorem 1**]

*Proof.* For item 1, when districts are deterministic, Figure 2 shows a cycle when tie-breaking $t$ in each district is $a \succ b$, and overall, $\tilde{t}$ is $b \succ a$.

For item 2, when districts are fractional, Figure 3 shows a cycle, when agents are lexicographic.

For item 3, when districts are fractional and voters globalist, Figure 4 shows a cycle. $\square$

## Simulations

While the results on convergence are tight, we are interested to see the effects of the decentralized iterative dynamic on the overall welfare of the system. Moreover, as we have multiple parameters of the type of districts and types of agents, we are interested in exploring how (and if) are these parameters affect the efficiency of the process and how this affects the overall social welfare.

While the examples above demonstrated a relatively small number of candidates, when trying to simulate more of them, the question arises of determining the preference order of the voters. We have chosen to run each simulation both with randomized preferences as well as with single-peaked preferences. Single-peaked preferences are relevant when there is an agreed upon ordering of candidates on some axis (e.g., political right to left; location of public parks along a street; etc.). Every voter has a particular location on the axis which is its most favored location, and the farther away an option is from that location, the lower it is in the voter's preference order. Unlike randomized preferences, which can create quite unrealistic preference orders, there is a case to be made regarding these preferences (see (Sprumont 1991)) and resembling realistic preferences (as (Mattei and Walsh 2017) note,
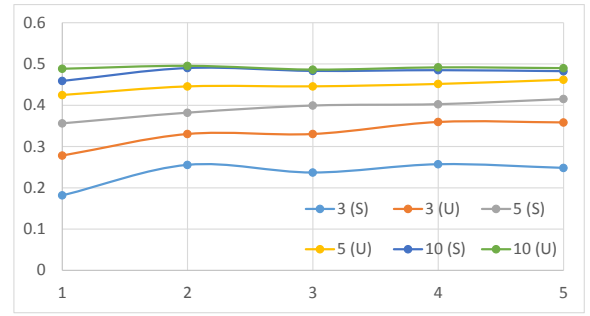


Figure 5: The average proportion of agents that prefer the final position over the opening position for non-strategic agents with deterministic tie-breaking, lexicographic utility, according to the gap constraint (the $x$-axis) and number of districts for elections with 8 candidates. (S) marks single-peaked preferences vs. uniformly random ones (U) Generally, as the gap constraints are loosened, agents can increase their welfare.

elections are rarely purely single peaked, but they are closer than purely random models, as (Kedar 2009) indicates).

In order to calculate social welfare, in each simulation, we measure the percentage of voter that prefer the final state over the initial state.[5]

In order to simplify data analysis and comparison, we ran a few experiments with a large variety of voter numbers, but here we will present the extensive simulations we have done of the iterative dynamic with 53 voters (the number was useful due to the district size). We examined the effects of the number of districts (we ran simulations with 2,3,5 and 10), and the size of the gap between maximal and minimal group size – $b^+ - b_-$ (we ran simulations with 1,2,3,4,5).

A simulation setting included a choice of number of the district, a choice of the number of candidates, a choice of the size of gap, whether voter preferences were uniformly randomly generated or single-peaked, whether districts were fractional or deterministic, whether agents' utilities were global or lexicographic, and whether agents were vote-strategic or not. For each of these 320 potential settings, we ran 1,000 scenarios, each beginning in the truthful state (for vote-strategic agents), and advancing from there.

## Results

Despite the theoretical convergence results, the most striking of the simulation results was that the rate of non-convergence was so small. It was, overall, slightly less than 0.66%. While no setting had a particularly large amount of cycles, the number of districts did slightly increase them, as did using lexicographical utilities and using fractional districts. On average, 57% of initial states were already in equilibrium, but this was highly volatile, and mostly appeared in settings with a small number of districts.

Convergence happened in almost all cases within 13 steps. Single-peaked preferences took markedly less time (presum-

---

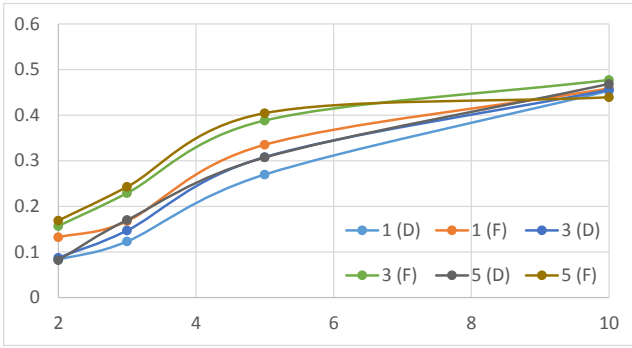[5]In a case where a cycle occurred, we averaged over all the states of the cycle.

Figure 6: The average proportion of agents that prefer the final position over the opening position for non-strategic agents with lexicographic utility and 5 candidates, according to number of districts (the $x$ axis) and various gaps. (F) marks fractional tie-breaking vs. deterministic one (D).
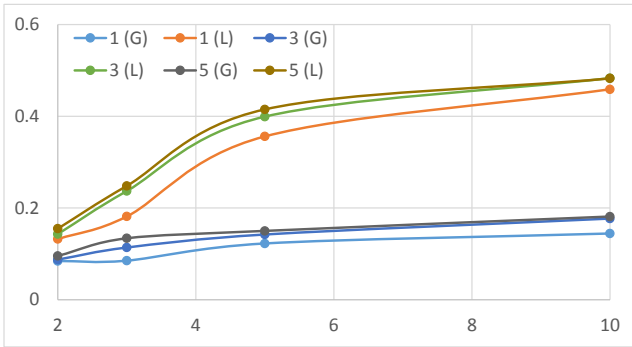


Figure 8: The average proportion of agents that prefer the final position over the opening position for non-strategic agents with deterministic tie-breaking, global utility and 5 candidates, according to number of districts (the $x$ axis) and various gaps. (S) marks single-peaked preferences vs. uniformly random ones (U).



Figure 7: The average proportion of agents that prefer the final position over the opening position for strategic agents with deterministic tie-breaking and 8 candidates, according to number of districts (the $x$ axis) and various gaps. (L) marks lexicographic preferences vs. globalist ones (G).

(though with different magnitudes) when changing number of candidates, tie-breaking system, and whether agents are strategic or not. This has to do with the greater ability of lexicographic agents to be partially satisfied – several agents, with opposing views can be satisfied. However, the growth is quite dramatic. Furthermore, notice that smaller gaps produce less utility for the agents, while larger gaps presumably allow greater flexibility to the manipulating agents.

Beyond lexicographic agents' greater utility, global agents' distribution can effect their utility significantly. As can be seen in Figure 8, single-peaked agents were almost half the utility compared those whose preferences were allocated uniformly at random. Notice that, as before, in general, higher gap agents were more successful.

We should note that one of the most surprising outcomes is that strategic agents did not, ultimately, have a meaningfully better utility than non-strategic agents. In a sense, all agents could save themselves the effort, and just not bother with strategizing.

## Discussion

The "reverse gerrymandering" setting, presented in this paper, may sound slightly unnatural at first blush, since people do not usually get to jump between voting districts (though (Bervoets and Merlin 2012) worked on such a setting). However, we believe that when viewed from the perspective of people participating in workplace committees, with their overlapping organizational influence, they do indeed strategize on where they could be more influential, and they move if they find a better position. In a more futuristic outlook, as autonomous systems become more common, the issue of these agents will need to be finding on their own where they are pivotal to help, and when will it be wrong to move. We presented here both theorems on the properties of this dynamic, and also explored it empirically (including for cases which we showed could not converge, hence an empirical examination is the main tool for analysis), discovering some key issues on the effect of changing the agent preference

ably, thanks to their more structured form), and global utility converged faster than lexicographic utility.

The flexibility that is given to players to manipulate – the maximal/minimal bounds over district sizes – affects the outcome, though it was not necessarily monotonic. As can be seen in Figure 5 (the $x$-axis is $b^+ - b_-$) the effect of the gap on all cases is almost identical – even when the "amplitude" of the agents' social welfare is different (mostly dependent on district size and on distribution), the effect of changing the constraints is fairly consistent across all settings. This is also true when comparing the social welfare as a function of the number of district for fractional vs. deterministic tie-breaking. As can be seen in Figure 6, the utility increase as the number of districts grows (the $x$-axis), and generally speaking under fractional settings, the agents tend to prefer the final state more than under deterministic tie-breaking settings.

The difference in utility between globalists and lexicographic agents is quite significant (almost $250\%$ more), as can be seen in Figure 7, an advantage that is consistent
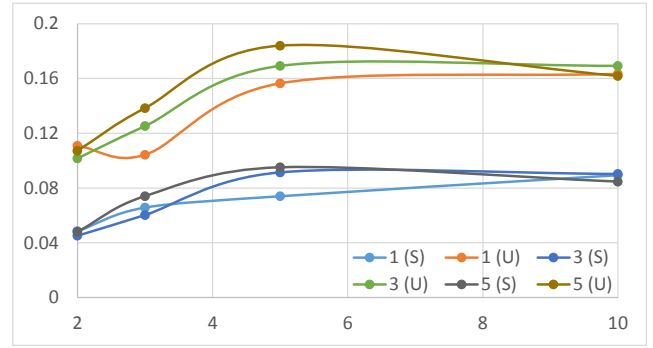
model, and the effect of district size on the agent. There is still much to discover – what other preference models work well with this setting; understanding better the effect of the gap constraints on social welfare; and combining various different types of agents in the same simulation, examining how their differences interact.

## Acknowledgments

## References

Aziz, H., and Savani, R. 2016. *Handbook of Computational Social Choice*. Cambridge University Press. chapter 15: Hedonic Games.

Aziz, H.; Kalinowski, T.; Walsh, T.; and Xia, L. 2016. Welfare of sequential allocation mechanisms for indivisible goods. In *Proceedings of the 22nd European Conference on Artificial Intelligence (ECAI)*, 787–794.

Bachrach, Y.; Lev, O.; Lewenberg, Y.; and Zick, Y. 2016. Misrepresentation in district voting. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*, 81–87.

Bervoets, S., and Merlin, V. 2012. Gerrymander-proof representative democracies. *International Journal of Game Theory* 41:473–488.

Borodin, A.; Lev, O.; Shah, N.; and Strangway, T. 2018. Big city vs. the great outdoors: Voter distribution and how it affects gerrymandering. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*. Stockholm, Sweden: to be published.

Cohen-Zemach, A.; Lewenberg, Y.; and Rosenschein, J. S. 2018. Gerrymandering over graphs. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, 274–282. International Foundation for Autonomous Agents and Multiagent Systems.

Grandi, U.; Loreggia, A.; Rossi, F.; Venable, K. B.; and Walsh, T. 2013. Restricted manipulation in iterative voting: Condorcet efficiency and Borda score. In *Proceedings of 3rd International Conference of Algorithmic Decision Theory (ADT)*, 181–192.

Kedar, O. 2009. *Voting for Policy, Not Parties: How Voters Compensate for Power Sharing*. Cambridge University Press.

Krishna, V. 2002. *Auction Theory*. Academic Press.

Lev, O., and Rosenschein, J. S. 2016. Convergence of iterative scoring rules. *Journal of Artificial Intelligence Research (JAIR)* 57:573–591.

Lewenberg, Y.; Lev, O.; and Rosenschein, J. S. 2017. Divide and conquer: Using geographic manipulation to win district-based elections. In *Proceedings of the 16th International Coference on Autonomous Agents and Multiagent Systems (AAMAS)*, 624–632.

MacGillis, A. 2016. Go midwest, young hipster. New York Times.

Maschler, M.; Solan, E.; and Zamir, S. 2013. *Game Theory*. Cambridge University Press.

Mattei, N., and Walsh, T. 2017. *Trends in Computational Social Choice*. AI Access. chapter 15: A PrefLib.org Retrospective: Lessons Learned and New Directions.

Meir, R.; Polukarov, M.; Rosenschein, J. S.; and Jennings, N. R. 2010. Convergence to equilibria of plurality voting. In *Proceedings of the 24th National Conference on Artificial Intelligence (AAAI)*, 823–828.

Meir, R. 2016. Strong and weak acyclicity in iterative voting. In *Proceedings of the 9th International Symposium on Algorithmic Game Theory (SAGT)*, 182–194.

Monderer, D., and Shapley, L. S. 1996. Potential games. *Games and Economic Behavior* 14(1):124–143.

Nisan, N.; Roughgarden, T.; Tardos, E.; and Vazirani, V. V., eds. 2007. *Algorithmic Game Theory*. Cambridge University Press.

Rabinovich, Z.; Obraztsova, S.; Lev, O.; Markakis, E.; and Rosenschein, J. S. 2015. Analysis of equilibria in iterative voting schemes. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI)*, 1007–1013.

Rogers, E. M. 1962. *Diffusion of Innovations*. Free Press of Glencoe.

Sprumont, Y. 1991. The division problem with single-peaked preferences: A characterization of the uniform allocation rule. *Econometrica* 59(2):509–519.

van Bevern, R.; Bredereck, R.; Chen, J.; Froese, V.; Niedermeier, R.; and Woeginger, G. J. 2015. Network-based vertex dissolution. *SIAM Journal on Discrete Mathematics* 29(2):888–914.