Approximate coalitional equilibria in the bipolar world*

Andrei Golman[1][0000−0002−6288−6543] and Daniil Musatov[1,2][0000−0002−1779−2513]

[1] Moscow Institute of Physics and Technology, Russia
[2] Russian Presidential Academy of National Economy and Public Administration, Moscow, Russia

Abstract. We study a discrete model of jurisdiction formation in the spirit of Alesina and Spolaore [1]. A finite number of agents live along a line. They can be divided into several groups. If a group is formed, then some facility is located at its median and every member x of a group S with a median m pays $\frac{1}{|S|} + |x - m|$.
We consider the notion of coalitional stability: a partition is stable if no coalition wishes to form a new group decreasing the cost of all members. It was shown by Savvateev et al [4] that no stable partition may exist even for 5 agents living at 2 points. We now study approximately stable partitions: no coalition wishes to form a new group decreasing all costs by at least ϵ.
In this work we define a relative measure of partition instability and consider bipolar worlds where all agents live in just 2 points. We prove that the maximum possible value of this measure is approximately 6.2%.

Keywords: facility location, group partition, coalitional stability, approximate equilibrium.

1 Introduction

Partitioning of society into groups affects many economic, political and social processes. Everyone is a member of many groups but sometimes the groups must be disjoint. For instance, most people have only one citizenship and only one employer. Almost everywhere a person may be a member of at most one political party. It is very hard to be a fan of two football clubs. An airline may be a member of only one alliance. A bitcoin miner may be a member of only one mining pool.

Why do people and other economic agents unite themselves into groups (clubs, coalitions, communities, jurisdictions etc.)? The general answer is that a group provides some type of club good that is either infeasible or too expensive for an individual. In this paper we study the case when this good is *horizontally differentiated*. It means that the good is described by a set of characteristics and different agents have different tastes. In this case two opposite forces affect

* Supported with RFBR 16-01-00362 and RaCAF ANR-15-CE40-0016-01 grants.

the size of a group. On the one hand, the larger the group the less fraction of the fixed cost is beared by an individual. On the other hand, a smaller group is typically more homogeneous, hence its members could be better satisfied by the club good characteristics.

The main theoretical question is whether these two forces always balance each other and yield a stable partition. The answer crucially depends on the underlying notion of stability. In some frameworks very general existence theorems are proven. In other cases some tricky examples without stable partitioning are constructed. To the best of our knowledge, this paper is the first one where an *approximate* equlibrium is studied in this context. It is more or less obvious that an approximate equilibrium should always exist for coarse approximation factors and may not exist for fine approximation factors if no exact equilibrium exist. Our main contribution is to derive concrete bounds on these factors.

Our study is narrowed to a specific framework where the main features of the model are highlighted. Firstly, we analyze a *bipolar* world with only two types of public good. Secondly, we analyze coalitional stability: a group partition is stable if no new coalition wishes to break out and thus decrease the cost of all its members. The approximate analogue is the following: no new coalition wishes to break out and thus decrease the cost of all its members *by a considerable amount*.

We illustrate our model by the following story: several sportspersons decide which game to play: football or handball. Some people prefer football, the others prefer handball, but everyone prefers to play than not to play. We ignore all limitations on the number of players: any group may rent a hall and play there. The hall is universal, so they may also mix the games: play football for some time and then switch to handball. The rent price is distributed equally among the players. If a player has to participate in the less preferred game, then he gets some additional disutility. The society may divide themselves into several groups. Each group rents a hall, shares the payment equally and choose the game by majority voting. If equal number of agents vote for both games then the playing time may be divided in any proportion.

1.1 Related literature

Here we present a broad perspective of the studies in local public goods, facility location and group partition.

The study of local public goods was started in the 1950s by a discussion between Samuelson [15] and Tiebout [19]. Samuelson stated that a public good can never be financed by a free choice of all society members. The main idea is that if an agent chooses how much public good to purchase then she does not account for positive externalities and thus underfinances the good. Tiebout responded that the situation is different in the case of local public goods. In this case the choice may be internalized by "foot voting": if an agent does not like the tax level and the amount of public good provided in his jurisdiction then he can migrate to another one. Tiebout's hypothesis was that in this case a social optimum is restored.

Unfortunately, Tiebout did not present a formal mathematical model. A subsequent line of research formulated several approaches that dealt with *vertical differentiation*: all agents value the local good but have different willingness to pay for it. The society becomes stratified by the level of public good provision and, consequently, taxes. This framework was analyzed by Westhoff [20], Bewley [3], Greenberg and Weber [8], among others. The latter paper was the first one to distinguish two notions of stability: migrational and coalitional. Under the former type of stability no single agent wishes to change her jurisdiction. Under the latter one no group of agents would like to form a new jurisdiction.[3] Both notions are based on quotes from the Tiebout paper.

Another approach deals with *horizontal differentiation* of local goods and heterogeneous tastes of agents. Mas-Colell [13] introduced a model of determining a type of the public good within a community. In the seminal paper [1] Alesina and Spolaore elaborated a model of stable partitioning into several jurisdictions (nations). Their model is continuous and one-dimensional with uniform distribution of agents. Each jurisdiction determines the location of its public good (the capital) by the median rule.[4] Every agent has two types of cost: monetary cost for providing the good and transportational cost for reaching its location. The paper also employs migrational and coalitional stability.

A bunch of subsequent papers analyze the stability and efficiency issues of the model. Haimanko et al [9] prove that if all coalitions may redistribute the cost in any manner then a coalitionary stable partition does always exist. Bogomolnaia et al [4], [5] study different notions of stability for models with finite number of agents and provide examples where no coalitionary stable structure exists. Savvateev [16] provides a counterexample for a more general notion of stability. Musatov et al [14] explore a very general model with a continuum of agents and migrational stability. It is proven that if a benefit from the local good is bounded and an agent may refuse to join any group then no stable structure may exist even for a very narrow class of models. On the other hand, if participation is mandatory, then there always exists a partition satisfying the border indifference property: no agent wants to migrate to the neighboring community. Under some mild single-crossing conditions such a partition is also a migrational equilibrium: no agent wants to migrate to a distant community as well. The existence result was expanded by Marakulin [12] to the case of a continuous population density with atoms.

Finally we mention some notable works on multi-dimensional models which are due to Drèze et al [7], Marakulin [11] and Savvateev et al [18]. The first paper also employed some kind of approximate equilibrium notion.

---

[3] This is the type of stability developed by Aumann and Drèze [2].

[4] This rule was also employed in the migrational models due to Bolton and Roland [6] and Jehiel and Scotchmer [10]

2 The model

2.1 General setting

Here we formulate a specific model of splitting a bipolar world with a finite number of agents. This model is similar to the model in [17]. A more general model may be found, for instance, in [14]. In our model there is a set X consisting of $N = L + R$ agents: L agents live at point 0 (representing football) and R agents live at point d (representing handball). Denote by $p(x)$ the point where agent x lives. A community S may consist of l agents from the left and r agents from the right. Every community should locate a facility somewhere between 0 and d (endpoints included). This facility represents the distribution of time between the games. We adopt the *median rule*: the facility is located at a median $m \in \mathrm{med}(S)$, i.e., both $\{x \in S \mid p(x)x \leq m\}$ and $\{x \in S \mid p(x)x \geq m\}$ must contain at least half of members of S.It means that if $l > r$ then $m = 0$, if $l < r$ then $m = d$, and if $l = r$ then m may be any point in $[0, d]$. (In this case we say that S is *indeterminate*). The median rule has two advantages. Firstly, a median is a minimizer for the total disutility. Secondly, it beats any other option by majority voting.

If a community is organized then all its members may use its facility. The facility is uncapacitated and the cost of its maintenance (payment for hall rent) does not depend on the number of users. Let this cost be equal to g. Apart from the maintenance cost every member pays for transportation to the facility (disutility from a non-preferred game). If an agent is located at p and a facility is located at m then the agent pays $t \cdot |p - m|$. Thus, the total cost of agent x within coalition S with facility at m equals

$$C(x, S, m) = \frac{g}{|S|} + t \cdot |p(x) - m|.$$

Suppose that a society is split into jurisdictions $S_1, \dots, S_k$ with fixed medians $m_1, \dots, m_k$. The configuration is *coalitionary stable* if no new community can emerge and decrease the cost of all its members. That is, there is no community S with its median m such that for all i and for all $x \in S \cap S_i$ it holds that

$$C(x, S, m) < C(x, S_i, m_i).$$

2.2 Fixing the parameters

In our model we have three arbitrary parameters: d (the distance between the two options), g (the cost of facility maintenance) and t (the unit cost of transportation). It is clear that d and t do not matter by themselves. Only the product $t \cdot d$ does influence the outcome. This value represents the disutility of football players from playing handball and vice versa. So, let us fix $t = 1$ and keep d variating. Now let us notice that g and d do not matter by themselves either, only g/d does matter. This value is the "exchange rate" between the hall rent and the disutility. So, let us fix $g = 1$ and keep (L, R, d) as the parameters of the model.

One could think that the model is also invariant with respect to multiplying L, R and g by the same factor. Indeed, all costs remain the same after this multiplication. But the notion of stability may differ, because new coalitional threats emerge. For instance, for $L = 30$ and $R = 20$ a community of size 27 may be a potential threat that is absent for $L = 3$ and $R = 2$. Nevertheless, we could suppose w.l.o.g. that $L \geq R$. (Football is not less popular than handball).

2.3 Approximate coalitional stability

Now we define the central notions for our analysis.

Definition 1. *Consider a configuration $\mathcal{P}$, i.e. a partition $X = S_1 \sqcup \cdots \sqcup S_k$ with fixed medians $m_1, \ldots, m_k$. Denote by $S(x)$ the group that contains x and by $m(x)$ the respective median. Denote by $C(x, \mathcal{P})$ the cost of x in $\mathcal{P}$, i.e. $C(x, S(x), m(x))$. Then define the absolute instability of the configuration as the value*

$$\Delta_{abs}(\mathcal{P}) = \max_{T \subset X, T \neq \emptyset, m \in \mathrm{med}(T)} \min_{x \in T} \big(C(x, \mathcal{P}) - C(x, T, m)\big). \tag{1}$$

The intuitive meaning is the following: if x agrees to join T with median m then she bears cost $C(x, T, m)$. The difference $C(x, \mathcal{P}) - C(x, T, m)$ is her advantage from joining T, or willingness to join T. The minimum over $x \in T$ is the minimal willingness to join T. We may also call it the willingness to secede of coalition T. If it is positive, then all agents in T would like to break out and establish this community. If it is true for at least one T, then the configuration is unstable. Thus, we have justified the following fact:

Proposition 1. *For any configuration $\mathcal{P}$ it holds that $\Delta_{abs}(\mathcal{P}) \geq 0$. Configuration $\mathcal{P}$ is stable if and only if $\Delta_{abs}(\mathcal{P}) = 0$.*

Proof. If we take $T = S_i$ then the willingness to join T is zero for all its members, hence the maximum is non-negative. The second part was already proved.

Definition 2. *The relative instability of configuration $\mathcal{P}$ is the value*

$$\Delta_{rel}(\mathcal{P}) = \max_{T \subset X, T \neq \emptyset, m \in \mathrm{med}(T)} \min_{x \in T} \frac{C(x, \mathcal{P}) - C(x, T, m)}{C(x, \mathcal{P})}. \tag{2}$$

The intuitive meaning is the same as before, but now the willingness to join is measured in terms of initial costs. Now we expand our notions from a configuration to the whole world.

Definition 3. *The (absolute, relative) instability of a bipolar world (L, R, d) is the minimal (absolute, relative) instability of all configurations $\mathcal{P}$ of this world. We use the same notation as before: $\Delta_{abs}(L, R, d)$ and $\Delta_{rel}(L, R, d)$.*

Our main question is the following:

Problem 1. Find the least upper bound on $\Delta_{abs}(L, R, d)$ and on $\Delta_{rel}(L, R, d)$. Which worlds are the least stable?

Partial solutions to this problem include computing the values of $\Delta_{abs}(L, R, d)$ and $\Delta_{rel}(L, R, d)$ for particular L, R and d, finding parameters with large instability and establishing some upper bounds.

3 Analysis of approximate equilibria: general considerations

It was proven in [17, Th. 2] that any stable configuration must belong to one of three types:

- "Union". The partition consists of one grand coalition. (One hall is rented and all play football).
- "Federation". The partition consists of two coalitions: all agents at 0 and all agents at d. (Two halls are rented and everyone plays their preferred game).
- "Mixed structure". The partition consists of two coalitions: the first one contains R agents from 0 and R agents from d, the second one contains $L-R$ agents from 0. The median of the first coalition lies somewhere between 0 and d. (Two halls are rented, the first one is used for football, the second one is used partially for football and partially for handball, no person who prefers handball plays in the first hall).

In the case of approximate stability things become more complex. But the following theorem still holds (both for absolute and relative instability):

Theorem 1. *There exists a configuration $\mathcal{P}$ minimizing the instability of a bipolar world (L, R, d), such that:*

- *There are at most 3 coalitions in $\mathcal{P}$. Among them, at most one coalition has the median at* 0*, at most one has the median at d and at most one has the median strictly in between.*
- *Among the coalitions with medians at* 0 *and d at most one contains some agents from the other point.*

The proof employs several lemmas. We start with the following one:

Lemma 1. *Among all configurations that minimize the instability of a bipolar world there must exist a strictly Pareto optimal one, i.e., there could not exist another configuration $\mathcal{P}'$ such that for all x it holds that $C(x, \mathcal{P}') \le C(x, \mathcal{P})$ and for some x it holds that $C(x, \mathcal{P}') < C(x, \mathcal{P})$. Also, every minimizing configuration must be weakly Pareto optimal, i.e., there could not exist another configuration $\mathcal{P}'$ such that for all x it holds that $C(x, \mathcal{P}') < C(x, \mathcal{P})$.*

Proof. Suppose that $\mathcal{P}$ is a configuration with minimal absolute instability and that $\mathcal{P}'$ weakly Pareto improves it. It means that for all x, T and m it holds that $C(x, \mathcal{P}') - C(x, T, m) \le C(x, \mathcal{P}) - C(x, T, m)$. The inequality remains valid after minimizing both parts in $x \in T$ and then after maximizing in (T, m). Thus we obtain $\Delta_{abs}(\mathcal{P}') \le \Delta_{abs}(\mathcal{P})$. If $\mathcal{P}$ is a minimizer then so must be $\mathcal{P}'$. If $\mathcal{P}'$ cannot be further Pareto improved, then it must be a strictly Pareto optimal minimizer. It can be easily shown that the set of possible tuples of utilities that Pareto improves the initial one is compact, and thus the mentioned $\mathcal{P}'$ exists. If the initial inequality were strict, then we would obtain $\Delta_{abs}(\mathcal{P}') < \Delta_{abs}(\mathcal{P})$ that contradicts to the choice of $\mathcal{P}$. Thus an optimal configuration cannot be strictly improved and thus must be weakly Pareto optimal. The proof for relative instability is similar.

This lemma immediately implies the following weak version of theorem 1:

Lemma 2. *There exists a configuration $\mathcal{P}$ that minimizes the instability of a bipolar world (L, R, d), such that for any m there exists at most one coalition with median m, and among the coalitions with medians at 0 or d at most one contains agents from the opposite point.*

Proof. From lemma 1 we may suppose that $\mathcal{P}$ is strictly Pareto efficient. If there are two coalitions with the same median then they may unite and thus reduce their costs without affecting any other agent. If there is a coalition S_1 with median 0 that contains agents from d and a coalition S_2 with median d that contains agents from 0 then the extra agents may switch their coalitions. The costs of these agents will fall and the costs of all others remain the same. If no such improvements could occur then we get the claimed configuration.

In order to get the full statement of theorem 1, we must prove that there could not be two jurisdictions with medians strictly between 0 and d. The argument from [17, Th. 1] is no longer valid since not all threats considered there lead to Pareto improvement. Instead, we employ the fact that the question is non-trivial only if no stable partition exists.

Lemma 3. *If a world (L, R, d) does not admit a stable configuration, then*

$$\frac{L}{R(L+R)} < d < \frac{1}{L}. \tag{3}$$

Proof. If no stable configuration exists, then, in particular, configurations "Union" and "Federation" are unstable. Consider the configuration "Union". Agents from 0 get least possible monetary cost and zero transportation cost. Therefore, a separating group may include only agents from the right.

The greater is the separating group the less monetary cost each of them bears. If some group would like to secede then the group of all right agents would like to secede all the more. This implies that if "Union" is unstable then

$$\frac{1}{R} < \frac{1}{L+R} + d. \tag{4}$$

By transposing the terms we get

$$d > \frac{1}{R} - \frac{1}{L+R} = \frac{L}{R(L+R)}$$

and thus establish the first inequality. In the sequel we suppose that (4) holds.

Now consider the configuration "Federation". Let S be the coalition having maximum willingness to secede. Let S contain l agents from the left and r agents from the right. Consider several cases:

1. S is indeterminate (i.e., $l = r$) and chooses median $m \in [0, d]$. If S wants to break out then it must hold that

$$\begin{cases} \frac{1}{L} > \frac{1}{l+r} + m & = \frac{1}{2r} + m; \\ \frac{1}{R} > \frac{1}{l+r} + d - m & = \frac{1}{2r} + d - m. \end{cases}$$

By summing up the two inequalities, we get

$$\frac{1}{L}+\frac{1}{R}>\frac{1}{r}+d.$$

Since $r \le R$ we get

$$d<\frac{1}{L}, \tag{5}$$

as stated.

2. The median of S is at 0 and $l > r$. Because of (4), agents from d cannot win from joining S. But a group of left agents cannot win from breaking out since after seceding they get $\frac{1}{l}$ instead of $\frac{1}{L}$.
3. The median of S is at d and $l < r$. If $l = 0$, then S cannot win from seceding, like in the previous case. If $l > 0$, then consider the group S' consisting of S and $r - l$ agents from the left with the median still at d. Transportation cost in S' is the same as in S and monetary cost is lower. Thus, if S wins from seceding, then S' does win all the more. But S' is indeterminate and thus it implies $d < \frac{1}{L}$, as before.

To complete the proof of theorem 1, we need only one more ingredient. Specifically, the case of two indeterminate coalitions must be excluded.

Lemma 4. *There exists a configuration $\mathcal{P}$ minimizing the instability of a bipolar world (L, R, d) that satisfies the properties from lemma 2 and, moreover, contains at most one indeterminate coalition.*

Proof. Suppose that $\mathcal{P}$ contains at least two indeterminate coalitions with total population $2a$. We prove that after merging into one group they can locate the median such that all their members are better off. Suppose that, on the contrary, some member x is worse off for any location of the median. Suppose that she belongs to coalition B with population $2b$.

Firstly we prove that B must be a large coalition. Specifically, b is greater than $\frac{a}{2}$. Indeed, the initial cost of x must be at least $\frac{1}{2b}$. The cost beared by x in the merged group can be made $\frac{1}{2a}+\frac{d}{2}$. Since x must be worse off, we obtain

$$\frac{1}{2a}+\frac{d}{2}>\frac{1}{2b}. \tag{6}$$

But from (5) we have $\frac{d}{2}<\frac{1}{2L}$. Since $a \le L$, we have $\frac{d}{2}<\frac{1}{2a}$. Plugging this into (6) we get $\frac{1}{2b}<\frac{1}{2a}+\frac{1}{2a}=\frac{1}{a}$, thus $b>\frac{a}{2}$, as stated.

Secondly, we prove that there could not be small coalitions. Specifically, any coalition C with population $c \le \frac{a}{2}$ is better off after joining B. Indeed, the new cost is at most $\frac{1}{2b}+d$. Since $b>\frac{a}{2}$ and $d<\frac{1}{L}$, it is at most $\frac{1}{a}+\frac{1}{L}$. Since $a<L$, this is at most $\frac{2}{a} \le \frac{1}{c}$. Since the old cost is at least $\frac{1}{c}$, all members of C prefer to join B. All members of B do also win if the median does not change: the transportation cost stays the same and the monetary cost decreases.

Thus there could be only two indeterminate coalitions. Finally we show that they are better off if they merge and locate the median at some point m. Suppose that their populations are $2a\gamma$ and $2a(1-\gamma)$, and the distance between their initial medians is q. Suppose that they merge and locate the new median between the old medians within the distances q_1 and q_2 to them, respectively. If the median moves towards an agent then she must be definitely better off. Consider the other agents. The members of the first coalition win $\frac{1}{2a\gamma} - \frac{1}{2a} - q_1$. The members of the second coalition win $\frac{1}{2a(1-\gamma)} - \frac{1}{2a} - q_2$. Both these value can be made non-negative iff their sum is non-negative. This is equivalent to the following:

$$\frac{1}{2a\gamma} + \frac{1}{2a(1-\gamma)} - \frac{1}{a} \geq q. \tag{7}$$

Note that $\frac{1}{\gamma} + \frac{1}{1-\gamma} \geq 4$ if $\gamma \in (0,1)$. Thus the left part of (7) is at least $\frac{2}{a} - \frac{1}{a} = \frac{1}{a}$. But $\frac{1}{a} \geq \frac{1}{L} > d$ and of course $d \geq q$. Thus (7) is established, the merger can be profitably done and lemma 4 and theorem 1 are proven.

4 Approximation algorithm for computing instabilities

The results of the previous section crucially decrease the number of configurations that could potentially minimize the instability. The remaining possibilities can be looked through in polynomial time. Firstly we show how to calculate the instability of a fixed configuration. Note that after the secession all members of the seceding coalition residing at the same point get the same cost. Thus the minima in (1) and (2) are achieved for x with minimal initial cost. Note also that we may consider only the case when the seceding coalition contains only the agents with maximal cost from the two poles. This justifies the following algorithm for computing $\Delta_{abs}(\mathcal{P})$:

1. Calculate $C(x,\mathcal{P})$ for all x. This step is done straightforwardly by definition.
2. Sort all agents from the left by the costs in nonincreasing order. Do the same with the agents from the right.
3. Assign value 0 to the variable Δ. For all $l \in [0,L]$ and $r \in [0,R]$ consider the lth agent from the left and the rth agent from the right in the obtained orders. Denote them by x and y respectively. Let T be the coalition of l agents from the left and r agents from the right. Do the following:
   - If $l > r > 0$ then $\Delta := \max\{\Delta, \min\{C(x,\mathcal{P}) - C(x,T,0), C(y,\mathcal{P}) - C(y,T,0)\}\}$;
   - If $r = 0$ then $\Delta := \max\{\Delta, C(x,\mathcal{P}) - C(x,T,0)\}$;
   - If $0 < l < r$ then $\Delta := \max\{\Delta, \min\{C(x,\mathcal{P}) - C(x,T,d), C(y,\mathcal{P}) - C(y,T,d)\}\}$;
   - If $l = 0$ then $\Delta := \max\{\Delta, C(y,\mathcal{P}) - C(y,T,d)\}$;
   - If $l = r$ then find m that maximizes $\min\{C(x,\mathcal{P}) - C(x,T,m), C(y,\mathcal{P}) - C(y,T,m)\}$. Usually it is just the root of linear equation $C(x,\mathcal{P}) - C(x,T,m) = C(y,\mathcal{P}) - C(y,T,m)$. Then assign $\Delta := \max\{\Delta, \min\{C(x,\mathcal{P}) - C(x,T,m), C(y,\mathcal{P}) - C(y,T,m)\}\}$.
4. Return Δ.

The algorithm for computing $\varDelta_{rel}(\mathcal{P})$ is similar, but all terms of type $C(z,\mathcal{P}) - C(z,T,m)$ are replaced by $\frac{C(z,\mathcal{P})-C(z,T,m)}{C(z,\mathcal{P})}$. The following bound on the time complexity is straightforward:

Proposition 2. *The algorithm for computing $\varDelta_{abs}(\mathcal{P})$ works for $O(N^2)$ steps for worlds with N agents.*

Proof. The first stage is performed in $O(NM)$ steps where M is the number of groups. Clearly $M = O(N)$, and for minimizing configurations we have even $M = O(1)$ by theorem 1. The sorting on the second stage takes $O(N \log N)$ steps. On the third stage the algorithms looks through all pairs (l, r) and makes a small calculation. This takes $O(LR) = O(N^2)$ steps. The total running time is $O(N^2)$, as stated.

Now we describe an approximate algorithm for computing $\varDelta_{abs}(L, R, d)$ (or $\varDelta_{rel}(L, R, d)$) in a bipolar world with parameters (L, R, d). The algorithm also gets the parameter M — the maximum number of considered medians of the indeterminate coalition. The idea is to search through all configurations satisfying the condition of theorem 1 and take the minimal instability. We denote by (l, r) the group consisting of l agents from the left and r agents from the right. If $l = r$, we attach the median m as the third parameter. The procedure is the following:

1. Check the validity of condition (3). If it does not hold, return 0.
2. Assign $\varDelta := \infty$. For all $k \in [0, R]$ and for all $c \in [0, M]$:
   (a) For all $l \in (L - R, L - k]$:
      – $\mathcal{P} :=$ the configuration consisting of groups $(l, 0)$, $(L-k-l, R-k)$, $(k, k, \frac{c}{M}d)$;
      – $\varDelta := \min\{\varDelta, \varDelta_{abs}(\mathcal{P})\}$.
   (b) For all $r \in [0, R - k]$:
      – $\mathcal{P} :=$ the configuration consisting of groups $(L-k, R-k-r)$, $(0, r)$, $(k, k, \frac{c}{M}d)$;
      – $\varDelta := \min\{\varDelta, \varDelta_{abs}(\mathcal{P})\}$.
3. Return $\varDelta$.

Proposition 3. *The algorithm for computing $\varDelta_{abs}(L, R, d)$ works for $O(N^4M)$ steps for worlds with N agents.*

Proof. The algorithm considers all possible k, c and l (or r). There are less than $R \cdot M \cdot L = O(N^2M)$ variants totally. The computation of $\varDelta_{abs}(\mathcal{P})$ takes $O(N^2)$ steps, thus the total time is $O(N^4M)$.

Now we estimate the precision of the described algorithm.

Proposition 4. *Let θ be $\frac{d}{M}$. If the algorithm for computing the absolute instability returns $\varDelta$, then $\varDelta_{abs}(L, R, d) \geq \varDelta - \frac{\theta}{2}$.*

Proof. Suppose that $\Delta_{abs}(L,R,d) = \Delta_{abs}(\mathcal{P})$, where $\mathcal{P}$ satisfies the condition of theorem 1. If $\mathcal{P}$ does not contain an indeterminate coalition then it will be considered during the algorithm and $\Delta = \Delta_{abs}(L,R,d)$ exactly. Suppose that $\mathcal{P}$ contains an indeterminate coalition Q with median q and one or two other coalitions. During the algorithm we will consider the partition $\mathcal{P}'$ with the same coalitions and median q' of Q such that $|q-q'| \leq \frac{\theta}{2}$. Note that for all x it holds that $C(x,Q,q) \leq C(x,Q,q') + \frac{\theta}{2}$. Hence

$$C(x,\mathcal{P}') \leq C(x,\mathcal{P}) + \frac{\theta}{2} \tag{8}$$

for all x. This inequality still holds after subtracting $C(x,T,m)$, taking minimum and maximum. Thus $\Delta_{abs}(\mathcal{P}') \leq \Delta_{abs}(\mathcal{P}) + \frac{\theta}{2}$. Since $\mathcal{P}'$ is considered by the algorithm, we obtain $\Delta \leq \Delta_{abs}(\mathcal{P}')$. By our assumption, $\Delta_{abs}(\mathcal{P}) = \Delta_{abs}(L,R,d)$. Thus the inequality $\Delta_{abs}(L,R,d) \geq \Delta - \frac{\theta}{2}$ is established.

Proposition 5. *If the algorithm for computing the relative instability returns Δ, then $\Delta_{rel}(L,R,d) \geq \Delta - \theta(L+R)$.*

Proof. The argument proceeds along the same line as the previous one. The difference starts when we make implications about Δ_{rel} instead of Δ_{abs}. The inequality (8) still holds and thus

$$\begin{aligned}\frac{C(x,\mathcal{P}') - C(x,T,m)}{C(x,\mathcal{P}')} &\leq \frac{C(x,\mathcal{P}) - C(x,T,m) + \frac{\theta}{2}}{C(x,\mathcal{P}')} \\ &= \frac{C(x,\mathcal{P}) - C(x,T,m)}{C(x,\mathcal{P})} \cdot \frac{C(x,\mathcal{P})}{C(x,\mathcal{P}')} + \frac{\theta}{2C(x,\mathcal{P}')}. \end{aligned} \tag{9}$$

The inequality (8) does also hold after switching $\mathcal{P}$ and $\mathcal{P}'$, i.e., $C(x,\mathcal{P}) \leq C(x,\mathcal{P}') + \frac{\theta}{2}$. Hence $\frac{C(x,\mathcal{P})}{C(x,\mathcal{P}')} \leq 1 + \frac{\theta}{2C(x,\mathcal{P}')}$. Thus the right part of (9) is at most $\frac{C(x,\mathcal{P})-C(x,T,m)}{C(x,\mathcal{P})} + \frac{\theta}{2C(x,\mathcal{P}')}\left(1 + \frac{C(x,\mathcal{P})-C(x,T,m)}{C(x,\mathcal{P})}\right) \leq \frac{C(x,\mathcal{P})-C(x,T,m)}{C(x,\mathcal{P})} + \frac{\theta}{C(x,\mathcal{P}')} \leq \frac{C(x,\mathcal{P})-C(x,T,m)}{C(x,\mathcal{P})} + \theta(L+R)$. The first inequality holds because $\frac{C(x,\mathcal{P})-C(x,T,m)}{C(x,\mathcal{P})} \leq 1$ and the second one because $C(x,\mathcal{P}') \geq \frac{1}{L+R}$. Putting all together, we get

$$\frac{C(x,\mathcal{P}') - C(x,T,m)}{C(x,\mathcal{P}')} \leq \frac{C(x,\mathcal{P}) - C(x,T,m)}{C(x,\mathcal{P})} + \theta(L+R).$$

After sequential minimizing and maximizing we get $\Delta_{rel}(\mathcal{P}') \leq \Delta_{rel}(\mathcal{P}) + \theta(L+R)$, that implies the statement of the theorem, as before.

5 Analysis of approximate equilibria: absolute instability

Now we proceed in the following way: we narrow the set of possible worlds where the maximum instability can be achieved and search through the remaing set exhaustively.

Theorem 2. *The maximal possible absolute instability is achieved for some* (L, R, d) *with* $L < 100$*,* $R < 100$ *and* $d \le 1$*.*

Proof. If $d > 1$ then “Federation” is stable. Indeed, every agent has the cost at most 1, but in any coalition with agents from different poles some agents has the cost greater than 1.

If $L \ge 100$ or $R \ge 100$ then in “Federation” some agents have cost at most 0.01. Thus the absolute instability is at most 0.01. But the configuration $L = 3$, $R = 2$, $d = \frac{14}{45}$ yields absolute instability $\frac{1}{90} > 0.01$. (The proof of this fact is straightforward but tedious and thus is omitted).

The remaining possibilities were analyzed in a brute-force manner. We found only three pairs (L, R) with $\Delta_{abs} > 0.01$ for some d. In all cases optimal d and Δ can be shown to be rational numbers. In table 1 we summarize our findings for optimal distances:

Table 1. The worlds with maximal absolute instabilities.

L	R	d	Δ
3	2	$\frac{14}{45} \approx 0.311$	$\frac{1}{90} \approx 0.0111$
4	3	$\frac{5}{24} \approx 0.208$	$\frac{1}{48} \approx 0.0208$
5	4	$\frac{7}{40} = 0.175$	$\frac{1}{80} = 0.0125$

These findings justify the following theorem:

Theorem 3. *The maximal possible absolute instability in a bipolar world is* $\frac{1}{48}$*.*

6 Analysis of approximate equilibria: relative instability

The case of relative instability is much more complex. The numerical experiments show that sometimes the optimal configuration is neither “Federation”, nor “Union”, nor “Mixed structure”. Specifically, the optimal configuration could be a “Pseudofederation”: there are two communities, one of which contains $L-k$ agents from the left and the other contains k agents from the left and all agents from the right. Thus, an analogue of theorem 2 from [17] does not hold in our setting. On figure 1 we show which configuration is the most stable one for which parameters.

Theorem 4. *There exists a bipolar world with* $\Delta_{rel}(L, R, d) > 0.0615$*.*

Proof. By applying our algorithm for computing Δ_{rel} to various parameters we found the following example: $L = 73$, $R = 56$, $d = 0.0114$. The relative instability returned by the algorithm 0.0622. By subtracting the discrepancy we get the claimed lower bound. In table 2 the relative instabilities for various configurations are shown.

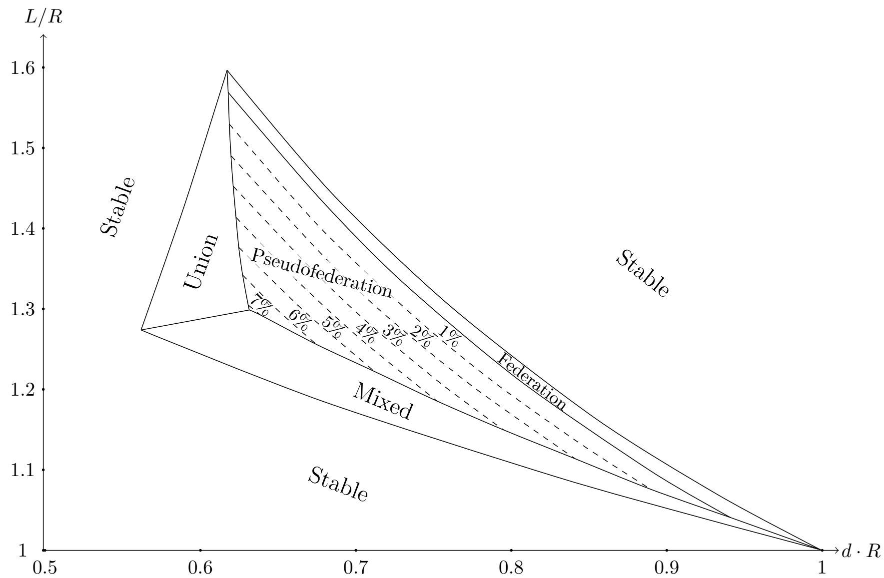

Fig. 1. The figure describes which configurations are the most stable ones for which parameters. Pseudofederation with $a\%$ means that the sizes of the groups are approximately $((1-0.01a)L, 0)$ and $(0.01aL, R)$.

Finally, we present an upper bound on Δ_{rel}.

Theorem 5. *For any bipolar world $\Delta_{rel}(L, R, d) \leq 0.063$.*

Proof (Sketch). The complete proof is rather technical and tedious. As before, it combines analytical and numerical methods. Here we present only the general idea. We consider the three main classes of configurations: "Union", "Federation" and "Mixed". We prove that our bound holds even for these configurations. All the more it does hold for an arbitrary one.

For example, consider the case of "Federation". Here $C(x, \mathcal{P}) = \frac{1}{L}$ for the left agents and $C(x, \mathcal{P}) = \frac{1}{R}$ for the right agents. Let the seceding coalition T be indeterminate. If T is willing to secede, then the group (R, R) is willing stronger,

Table 2. Relative instabilities in the world $L = 73$, $R = 56$, $d = 0.0114$.

Configuration	Groups	Δ
Union	$(73, 56)$	0.066
Federation	$(73, 0) + (0, 56)$	0.062
Mixed	$(56, 56) + (9, 0)$	0.074
Optimal	$(69, 0) + (4, 56)$	0.062

so we may think that $T = (R, R)$. Let t be the median of T. Thus the new costs are $\frac{1}{2R} + t$ for the left agents and $\frac{1}{2R} + d - t$ for the right agents. If the relative instability equals δ then

$$\begin{cases} \delta \geq 1 - \dfrac{\frac{1}{2R} + t}{\frac{1}{R}} = \dfrac{1}{2} + Lt; \\ \delta \geq 1 - \dfrac{\frac{1}{2R} + d - t}{\frac{1}{L}} = 1 - \dfrac{L}{R} \cdot \left(\dfrac{1}{2} + Ld - Lt \right). \end{cases} \tag{10}$$

Note that the right parts of (10) depend not on L, R, d, t themselves, but on the composite values $\rho = \frac{L}{R}$, $\gamma = Ld$, $\eta = Lt$. Thus δ must be not less than

$$\min \left\{ \frac{1}{2} + \eta, 1 - \rho \left(\frac{1}{2} + \gamma - \eta \right) \right\}. \tag{11}$$

This lower bound is maximal for $\eta = \frac{\frac{1}{2} - \rho(\frac{1}{2} + \gamma)}{1+\rho}$. Plugging this into (11) we get a lower bound on δ.

Then we consider other seceding coalitions and obtain other lower bounds on δ. The maximum of these bounds is a lower bound on the relative instability of the configuration. Then we consider other configurations and take the minimal bound. We search for this minimum numerically and find the following values:

$$\frac{L}{R} \approx 1.3065; \quad \frac{d}{R} \approx 0.6328; \quad \delta \approx 0.0623.$$

Adding the discrepancy, we obtain $\Delta_{rel}(L, R, d) \leq 0.063$ for all (L, R, d), as claimed.

7 Conclusion

This work is the first contribution to the literature on jurisdiction partitions that analyzes approximate coalitional equilibria. We have analyzed the bipolar world and found out how far such a world could be from an equilibrium. We considered the cases of absolute and relative instability metrics and established the bounds on them using numerical methods. In the first case the bound is exact. In the second case the upper and the lower bounds almost coincide. We believe that the true value lies somewhere between 0.622 and 0.623. The future work should give the precise value in analytic form and analyze the worlds other than bipolar ones. Our conjecture is that the bipolar worlds are the least stable.

Acknowledgments

We want to thank Alexei Savvateev for his support and advice during the work on this paper.

References

1. Alesina, A., Spolaore, E.: On the number and size of nations. The Quarterly Journal of Economics **112**(4), 1027–1056 (1997)
2. Aumann, R.J., Drèze, J.H.: Cooperative games with coalition structures. International Journal of Game Theory **3**(4), 217–237 (1974)
3. Bewley, T.F.: A critique of Tiebout's theory of local public expenditures. Econometrica: Journal of the Econometric Society pp. 713–740 (1981)
4. Bogomolnaia, A., Le Breton, M., Savvateev, A., Weber, S.: Stability under unanimous consent, free mobility and core. International Journal of Game Theory **35**(2), 185–204 (2007)
5. Bogomolnaia, A., Le Breton, M., Savvateev, A., Weber, S.: Stability of jurisdiction structures under the equal share and median rules. Economic Theory **34**(3), 525–543 (2008)
6. Bolton, P., Roland, G.: The breakup of nations: a political economy analysis. The Quarterly Journal of Economics **112**(4), 1057–1090 (1997)
7. Drèze, J., Le Breton, M., Savvateev, A., Weber, S.: "Almost" subsidy-free spatial pricing in a multi-dimensional setting. Journal of Economic Theory **143**(1), 275–291 (2008)
8. Greenberg, J., Weber, S.: Strong Tiebout equilibrium under restricted preferences domain. Journal of Economic Theory **38**(1), 101–117 (1986)
9. Haimanko, O., Le Breton, M., Weber, S.: Voluntary formation of communities for the provision of public projects. Journal of Economic Theory **115**(1), 1–34 (2004)
10. Jehiel, P., Scotchmer, S.: Constitutional rules of exclusion in jurisdiction formation. The Review of Economic Studies **68**(2), 393–413 (2001)
11. Marakulin, V.M.: On the existence of immigration proof partition into countries in multidimensional space. In: Int Conf DOOR. pp. 494–508 (2016)
12. Marakulin, V.M.: A theory of spatial equilibrium: the existence of migration proof country partition in an uni-dimensional world. Siberian Journal of Pure and Applied Mathematics **17**(4), 64–78 (2017), in Russian
13. Mas-Colell, A.: Efficiency and decentralization in the pure theory of public goods. The Quarterly Journal of Economics **94**(4), 625–641 (1980)
14. Musatov, D.V., Savvateev, A.V., Weber, S.: Gale–Nikaido–Debreu and Milgrom–Shannon: Communal interactions with endogenous community structures. Journal of Economic Theory **166**, 282–303 (2016)
15. Samuelson, P.A.: The pure theory of public expenditure. The Review of Economics and Statistics pp. 387–389 (1954)
16. Savvateev, A.: Uni-dimensional models of coalition formation: non-existence of stable partitions. Moscow Journal of Combinatorics and Number Theory **2**(4), 49–62 (2012)
17. Savvateev, A.: An analysis of coalitional stability in a bipolar world. Journal of New Economic Association **17**, 10–44 (2013), in Russian
18. Savvateev, A., Sorokin, C., Weber, S.: Multidimensional free-mobility equilibrium: Tiebout revisited (2018), https://arxiv.org/abs/1805.11871
19. Tiebout, C.M.: A pure theory of local expenditures. Journal of political economy **64**(5), 416–424 (1956)
20. Westhoff, F.: Existence of equilibria in economies with a local public good. Journal of Economic Theory **14**(1), 84–112 (1977)