CORVINUS ECONOMICS WORKING PAPERS

CORVINUS
UNIVERSITY of BUDAPEST

Faculty of Economics

CEWP 19/2014

The Multiple Hierarchical Legislatures in a Representative Democracy: Districting for Policy Implementation

by

Katsuya Kobayashi,

Atilla Tasnádi

http://unipub.lib.uni-corvinus.hu/1774

The Multiple Hierarchical Legislatures in a Representative Democracy: Districting for Policy Implementation

Katsuya Kobayashi* Attila Tasnádi†

December 17, 2014

Abstract

We build a multiple hierarchical model of a representative democracy in which, for instance, voters elect county representatives, county representatives elect district representatives, district representatives elect state representatives, and state representatives elect a prime minister. We use our model to show that the policy determined by the final representative can become more extreme as the number of hierarchical levels increases because of increased opportunities for gerrymandering. Thus, a sufficiently large number of voters gives a district maker an advantage, enabling her to implement her favorite policy. We also show that the range of implementable policies increases with the depth of the hierarchical system. Consequently, districting by a candidate in a hierarchical legislative system can be viewed as a type of policy implementation device.

Keywords: Electoral Systems, Median Voter, Gerrymandering, Council Democracies.

***JEL* Classification Number:** D72.

*Faculty of Economics, Hosei University, 4342 Aihara-machi, Machida-shi, Tokyo 194-0298, Japan. E-mail katsuyak@hosei.ac.jp

†MTA-BCE "Lendület" Strategic Interactions Research Group, Department of Mathematics, Faculty of Economics, Corvinus University of Budapest, Fővám tér 8, Budapest H-1093, Hungary. E-mail attila.tasnadi@uni-corvinus.hu

1 Introduction

In many parliamentary democracies, each representative is elected within a single-member district. However, although having many citizens elect one representative may mean a government is elected efficiently, the relationship between a representative and his/her voters could be weak. This can result in a misrepresentation of citizens' diverse interests, many wasted votes, and lower participation rates. Shugart and Carey (1992) point out that there are trade-offs between the efficiency of citizens' choosing one government and the representativeness of their diverse interests. To reduce this misrepresentation, an electoral system in which a small number of citizens elect one representative can be better than a system in which many citizens elect one representative. While the former system is effective in avoiding misrepresentation, if there are many citizens, they need to be grouped into districts. Then, the system needs to form legislatures, councils, or assemblies comprising the elected representatives as efficiently as possible. However, a legislature may still have too many representatives, making it difficult to organize a single government (e.g., by electing a prime minister). In this case, both efficiency and representativeness may be improved by forming an upper legislature comprising a few representatives from those in the legislature. Depending on the number of citizens, it may then be necessary to add further levels, thus creating a multiple hierarchical legislature. Here, citizens in the same single-member district elect an intermediate representative. Clearly, we can extend this system further by increasing the number of intermediate levels, thus obtaining a multiple hierarchical representative democracy. The purpose of this study is to determine the possible policy outcomes in such a system.

According to their constitution, China is an example of a representative democracy with a multiple hierarchical legislative system, as described above. The legislative system has four-levels, namely the nation, provinces, counties/cities/municipal districts, and townships/towns, with a government at each level. Representatives of legislatures at the national and provincial levels are elected by the legislatures at the next lower level (Article 97 in the Constitution of the People's Republic of China).[1] However, since the central and local governments, including each

[1]Initially, representatives at the county level were also elected by the legislatures at the next lower level. In

legislature, are "under the leadership of the Communist Party" (Preamble of the Constitution) in China, the party controls all elections (Institute of Chinese Affairs 2011).

The proposed model can also be found in related literature and attempts have been made to implement it during the past one and a half centuries, although not in its pure form. In expressing his views on democracy, Jefferson (1816) outlined the so-called "ward republics," in which he distinguished between the national, the state, the county, and the ward level. He characterized this system as follows:

> "It is by dividing and subdividing these republics from the great national one down through all its subordination, until it ends in the administration of every man's farm by himself; by placing under every one what his own eye may superintend, that all will be done for the best."

The so-called council system is based on a similar idea (for its history, we refer to Olson (1997)).

However, even if the problems of misrepresentation or wasted votes are solved by districting citizens prior to an election, there is another potential problem: who groups citizens into districts? Should it be the incumbent representatives, citizens, or perhaps an outsider? If the district maker has a particular political position, she may construct the districts in such a way as to give an advantage to those candidates who have a similar political position to hers. This is referred to as "gerrymandering," and can cause misrepresentations. Gilligan and Matsusaka (2006) measure how misrepresentative a policy determined in the single level legislature is in the case of most extreme gerrymandering. On the other hand, Coate and Knight (2007) show the conditions necessary for socially optimal districting when the independents' policy positions are determined stochastically. From the standpoint of democracy, gerrymandering can be viewed normatively as one of the worst results. However, it can also be viewed as a type of implementation that provides district makers or social planners with a method of implementing their favorite policy.

---

the reform of the electoral law in 1979 (this change was also reflected in the Constitution in 1982), citizens' direct election was expanded (Institute of Chinese Affairs 1980 and 1983). Since then, representatives at the county level have been elected by citizens directly, in addition to the township/town levels.

In this paper, we extend the districting model of Gilligan and Matsusaka and show that, by extreme gerrymandering, a multiple hierarchical model of a representative democracy can serve the interests of a minority. In particular, we show that as the number of voters increases, the district maker can construct more intermediate levels. Accordingly, the district maker's policy implementability becomes stronger and, even in the case of an extreme political position, gains dictatorial power.[2] In addition, we explicitly show the policy range that can be implemented by the hierarchical gerrymandering of every level. From these results, when each voter has the right to become the district maker with an equal probability, the most probable policy is the same as the case of gerrymandering by the extremists. We conclude that gerrymandering districts can be regarded as a method available to a district maker to implement her favorite policy. In addition, her policy implementability is stronger when the legislative system is hierarchical, and is far stronger when there are many citizens in the society. Furthermore, our results may indicate why, despite the advantages pointed out by Jefferson, we find so few historical examples of a multiple hierarchical legislative system such as the ward republic, other than that of China, since the system gives stronger policy implementability to the district maker.

It is worthwhile mentioning that we abstract away from inherent geographical constraints, which pose a problem in the political districting problem. For works on measuring district compactness, we refer the reader to Chambers and Miller (2010, 2013) and Fryer and Holden (2011). In a normative framework, geographical constraints are considered by Puppe and Tasnádi (2014), as well as the references therein. Details on redistricting in practice can be found, for instance, in Altman and McDonald (2010).

The remainder of the paper is organized as follows. Section 2 presents our extended model of a multiple hierarchical representation and a motivating example for the case of an extreme district maker. Section 3 considers the extremists' policy implementability in the case of optimal gerrymandering. Then, Section 4 investigates the policy implementability by politically moderate district makers' gerrymandering. Finally, Section 5 concludes the paper and applications of our model are provided in the Appendix.

[2]On this point, Galam and Wonczak (2000) use a numerical simulation to show that a dictatorship can appear under a hierarchical election. In contrast, we derive more general analytical results along a real line.

2 The model

We use a multiple application of the median voter theorem by nesting Gilligan and Matsusaka's model into itself a finite number of times. According to the well-known median voter model introduced by Black (1958), the policy preferred by the median voter prevails over any other policies in the case of a uni-dimensional policy space and voters' single-peaked preferences.

Gilligan and Matsusaka define and calculate the bias between the median policy and the policy decided in the legislature, which is composed of representatives elected in single-member districts, where each voter is allocated to exactly one district. They find that there is a possibility for gerrymandering in an indirect democracy that divides voters into districts and gives one group an advantage in the election. As a result, the final policy chosen by the representatives in the legislature may not be the policy the median voter prefers.

As we pointed out in the introduction, some societies can adopt a political system in which policies are decided hierarchically. In this paper, by extending the model of Gilligan and Matsusaka, we show how the final policy chosen and implemented by the representative elected from the "multiple-level legislatures" is decided in gerrymandered districts. In addition, we show how the final policy can deviate from that of the median voter theorem and how much power the district maker has to implement her favorite policy through gerrymandering.

The settings and notation we use in our model are essentially the same as those of Gilligan and Matsusaka. The population of citizens[3] consists of N people (hereafter, referred to as voters), all of whom vote. We assume that N is an odd number. The set of voters is defined as $\mathcal{N} = \{1, 2, 3, \ldots, N\}$. We assume that each voter $j \in \mathcal{N}$ has an ideal policy at own position $x_j \in \mathbb{R}$, and that each voter's utility strictly decreases monotonically as an implemented policy moves further from her own ideal position. For convenience, we also assume that voters with smaller numbers (the left wing from the median) are more liberal and that those with larger numbers (the right wing from the median) are more conservative. Thus, we label the voters

[3] In our model, each citizen is a voter who can cast a ballot for her favorite candidate and is a candidate for the final representative who implements her own preferred policy, as in the citizen-candidate models of Besley and Coate (1997).

such that $x_1 < x_2 < \ldots < x_N$. Then, the index of the median of all voters is equal to $\frac{N+1}{2}$. Let $F(x)$ be the cumulative distribution function of voter ideal points, such that $F(x_1) = 1/N$, $F(x_2) = 2/N$, …, $F(x_N) = 1$, which shows the relative position of each voter's ideal policy. We assume that there is a unique median voter in the population with ideal point x^*_{POP} such that $F(x^*_{POP}) = \frac{N+1}{2N}$.[4] The distance of a policy x from the median's ideal policy x^*_{POP}, measured by F, is called a bias, which we define as follows.

Definition 1. *The measure of a policy bias is equal to* $B(x) = |F(x) - F(x^*_{POP})|$.

Observe that policy choice x is unbiased or has minimal bias when $B = 0$, and x has maximal bias when $B = 1 - \frac{N+1}{2N} = \frac{N+1}{2N} - \frac{1}{N} = \frac{N-1}{2N}$.

We consider $t+1$ decision levels, starting from $t = 0$. At decision level $i \in \{1, 2, \ldots, t-1, t\}$, voters grouped into equally sized districts send a representative to a legislature at the $i+1$-th decision level. Let K_i be the number of districts at the i-th decision level. In order to finally elect a single representative who decides on a policy, only one district is formed at the $t+1$-th decision level. Thus, we let $K_{t+1} = 1$. For convenience, let $K_0 = N$.

Assumption 1. *N is odd and the district size K_{i-1}/K_i at level i is an odd integer for any $i \in \{1, \ldots, t+1\}$.*

If not stated otherwise, we assume that Assumption 1 holds.[5]

When $t = 0$, there are no legislatures between voters and the final representative. Only one representative, who implements a policy, is elected directly from among the voters. We view this as a direct democracy. When $t \geq 1$, our model becomes an indirect democratic system and, especially when $t \geq 2$, has hierarchical legislatures with multiple levels. We assume that every voter and representative at each i-th decision level casts a ballot sincerely and that only one among them is elected as a single-member district representative by a majority rule.

[4] In Gilligan and Matsusaka, the cumulative distribution function $F(x^*_{POP}) = 1/2$ at the median's ideal policy is defined. However, this definition is not appropriate because the authors assume discrete voters and N is odd. For example, $N = 27$, $F(x^*_{pop}) = 14/27 \neq 1/2$.

[5] Nevertheless, in the statements of our lemmas, propositions, and corollaries, we explicitly require Assumption 1 whenever needed.

Accordingly, the median voter of each district is elected as the representative, and every district at the $i+1$-th decision level is composed of the median voters from each district at next lower level. Thus, K_i is also the number of representatives at the $i+1$-th decision level. From the assumption of equally sized districts at the same level, the population of each district at the $i+1$-th decision level is K_i/K_{i+1}. Naturally, we must have $K_{i+1} < K_i$, since we elect fewer and fewer representatives as we move upwards in the hierarchy.

In summary, the basic structure of our model is as follows. First, N voters are divided into K_1 equal-sized districts. Then, $N/K_1 = K_0/K_1$ voters per district at the first decision level elect the representatives of legislative level 1 at the second decision level. Consequently, legislative level 1 comprises K_1 representatives. Second, the representatives of legislative level 1 are divided into K_2 equal-sized districts and K_1/K_2 members per district elect the representatives of legislative level 2 at the third decision level. Consequently, legislative level 2 comprises K_2 representatives. Third, the representatives of legislative level 2 are divided into K_3 equal-sized districts and K_2/K_3 members per district elect the representatives of legislative level 3 at the fourth decision level. Consequently, legislative level 3 comprises K_3 representatives, and so on. Finally, since the $t+1$-th decision level is the final level and $K_{t+1} = 1$, the K_t representatives of legislative level t at the $t+1$-th decision level elect only one representative, who is the final representative. Thus, t denotes the number of hierarchical legislative levels inserted between the voters and the final representative. These levels are, for example, ward representatives, county representatives, state representatives, and national representatives. The final representative decides and implements only one policy, which is a number on the real line, that applies to all voters. Table 1 depicts the above structure. Here, we assume that neither voters nor representatives can commit to policies. Thus, the final representative implements his/her own ideal policy. Note that Gilligan and Matsusaka's model is a special case of our model, where $t = 1$. Additionally, there is no vote-value disparity because of the equal-size districts at each level.

From the above structure, we can immediately obtain the following lemma:

Lemma 1. *Under Assumption 1, the number of hierarchical levels is at most the number of*

Table 1: The basic structure of the model

Decision level	Voter or legislative level	Number of district		Total population	Population per district
first	voters	K_1		$K_0 = N$	N/K_1
			$\searrow$		
second	legislative level 1	K_2		K_1	K_1/K_2
			$\searrow$		
third	legislative level 2	K_3		K_2	K_2/K_3
.	.	.	$\searrow$	.	.
.	.	.		.	.
.	.	.	$\searrow$	.	.
$t-1$-th	legislative level $t-2$	K_{t-1}		K_{t-2}	K_{t-2}/K_{t-1}
			$\searrow$		
t-th	legislative level $t-1$	K_t		K_{t-1}	K_{t-1}/K_t
			$\searrow$		
$t+1$-th	legislative level t	$K_{t+1} = 1$		K_t	$K_t/K_{t+1} = K_t$
	$\parallel$				
	t **legislative levels**		$\searrow$		
	final representative			1	

prime factors of N.

Proof. We prove this by contradiction. Let $N = a_1 \cdot a_2 \cdot a_3 \cdots a_t \cdot a_{t+1}$, where each a_i, $i \in \{1, 2, 3, \ldots, t, t+1\}$, is a prime factor of N. We assume that for given N we can make $j+1$ decision levels, noting $K_i > K_{i+1}$ and $K_{j+1} = 1$, where $j > t$. Then, the populations of each district at each level $1, 2, 3, \ldots, j, j+1$ are $\frac{N}{K_1}, \frac{K_1}{K_2}, \frac{K_2}{K_3}, \ldots, \frac{K_{t-1}}{K_t}, \ldots, \frac{K_j}{K_{j+1}}$, respectively. Thus,

$$N = \frac{N}{K_1} \cdot \frac{K_1}{K_2} \cdot \frac{K_2}{K_3} \cdots \frac{K_{t-1}}{K_t} \cdots \frac{K_j}{K_{j+1}},$$

and the number of the factors of N is $j+1$. However, N cannot be the product of more than t integers larger than 1 from its prime factorization, which is a contradiction. $\square$

Hereafter, for convenience, $t^* + 1$ is regarded as the maximum number of decision levels (under Assumption 1) and the number of the prime factors of N. Note that we can set up hierarchical decision levels between 0 and t^*+1 by multiplying together any prime factors of N. For instance, suppose $N = \frac{N}{K_1} \cdot \frac{K_1}{K_2} \cdot \frac{K_2}{K_3} \dots \frac{K_{t^*}}{K_{t^*+1}}$. We can also set up t^*-1 decision levels in addition to t^*+1 levels if we multiply the first three prime factors, $\frac{N}{K_1} \cdot \frac{K_1}{K_2} \cdot \frac{K_2}{K_3}$, with populations per district of $\frac{N}{K_3}, \frac{K_3}{K_4}, \dots, \frac{K_{t^*}}{K_{t^*+1}}$ at each decision level, respectively. See also Example 1.[6]

Let $x^*_{i,k}$ be the ideal point and $j^*_{i,k}$ be the index (i.e., $x_{j^*_{i,k}} = x^*_{i,k}$) of the median representative or voter in district k at the i-th decision level (i.e., legislative level $i-1$). When there are $t+1$ decision levels, at the final decision level, where $K_{t+1} = 1$, the policy outcome decided in the final legislative level becomes $x^*_{t+1,1}$.

How do liberal voters gerrymander to maximize their own political payoff? We shall assume that a liberal extremist, called voter 1, can arrange all districts.[7] Then, this person will attempt to put a representative who is as left as possible on the median position in each district at each decision level. We can say that the more to the left the final representative falls, the stronger is her policy implementability. In this paper, we use the well-known "cracking and packing" algorithm at each decision level, as formulated by Gilligan and Matsusaka. They describe the algorithm as follows.

> "First, citizens with high-value ideal points are "cracked" into districts where they are the minority, maximizing the influence of citizens with low-value ideal points, and second, the remaining high-value citizens are "packed" into districts containing a preponderance of like-minded citizens in order to waste their votes through overkill."

In the next example, we specifically show how far away the final policy in the multiple hierarchical electoral system under the extremist's gerrymandering can be from the policy

[6] It might look as though Lemma 1 imposes a severe restriction on the number of possible decision levels. For instance, if N is a prime number, we can construct only one district at the first decision level because of the assumption of equal-sized districts. We can overcome this problem by changing "equal-sized districts" slightly to "almost equal-sized districts" by rounding K_i/K_{i+1} either up or down. Then, we consistently pick one of the two electors from "the middle" in the case of even district sizes. However, this extension would just complicate our analysis and would produce the same results. See also Example 2.

[7] Of course, in the same way, we can also consider a conservative extremist instead of a liberal extremist.

decided in a direct democracy (i.e., the median voter's ideal policy). We compare the final policy in our model with the policy in the model of Gilligan and Matsusaka in the next example.

Example 1. We consider an example of three decision levels, where $t+1=3$, $N=27$, voters set $\mathcal{N}=\{1,2,3,\ldots,25,26,27\}$, and voters have ideal points at $x_j=j$. We compare the indirect democracies of two decision levels ($t=1$) and three decision levels ($t=2$) with the direct democracy ($t=0$).[8] In this case, the median voter is $\frac{N+1}{2}=14$, as shown in Table 2, so that the final policy is decided at $x^*_{1,1}=x_{14}$ in the direct democracy. However, as shown

Table 2: The case of $N=27$, $K_1=1$

$\{1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27\}$
$\downarrow$
14

by Gilligan and Matsusaka, the final policy is not always decided at the median of all voters in an indirect democracy with gerrymandered districts. Here, we show that the selected policy becomes further from the median as the levels increase. Table 3 and Table 4 illustrate the most extreme gerrymandering to give an advantage to liberal (left) voters in the case of two decision levels ($t=1$). These can be regarded as regular indirect democracies, where the final representative is elected from the legislature composed of the representatives elected in single-member districts.

For $t=1$ and $N=27$, we can consider two cases where $(K_1,K_2)=(3,1)$ and $(9,1)$. Then, we can easily find the final representative who is as close as possible to voter 1, who is a liberal extremist, as in the example with nine voters in Gilligan and Matsusaka. In both Tables 3 and 4, the policies are decided by voter 10, whose ideal position is 4 positions away from the

[8] Here, 27 is equal to $3\cdot3\cdot3$ by the prime factor decomposition. From this fact, in the case of two decision levels, each district at the first decision level and at the second level are composed of 9 voters and 3 representatives, respectively, or of 3 and 9, respectively. In the case of three decision levels, each district is composed of three voters or representatives at every decision level. As a result, we can obtain combinations of $K_1=1$, $(K_1,K_2)=(3,1),(9,1)$, and $(K_1,K_2,K_3)=(9,3,1)$ for the cases of one decision level, two decision levels, and three decision levels, respectively.

Table 3: The case of $N = 27,\ K_1 = 3, K_2 = 1$

$\{1, 2, 3, 4, 5, 24, 25, 26, 27\}$	$\{6, 7, 8, 9, 10, 20, 21, 22, 23\}$	$\{11, 12, 13, 14, 15, 16, 17, 18, 19\}$
$\downarrow$	$\downarrow$	$\downarrow$
$\{5$	10	$15\}$
	$\downarrow$	
	10	

Table 4: The case of $N = 27,\ K_1 = 9, K_2 = 1$

$\{1, 2, 27\}$	$\{3, 4, 26\}$	$\{5, 6, 25\}$	$\{7, 8, 24\}$	$\{9, 10, 23\}$	$\{11, 12, 22\}$	$\{13, 14, 21\}$	$\{15, 16, 20\}$	$\{17, 18, 19\}$
$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$
$\{2$	4	6	8	10	12	14	16	$18\}$
				$\downarrow$				
				10				

median voter. Even if there are several district patterns in a combination of N and t, it can be shown that the maximum distances between the final representative and the median are identical in the case of optimal gerrymandering for either the liberal or conservative extremists who are "partisan."

Finally, what is the policy decided by the upper representatives if an additional level of representation is incorporated in the above decision process in the case of gerrymandering? In the same way as in our previous examples with two decision levels, the districts at the second decision level are arranged by one liberal extremist in order to implement a final policy as close as possible to her own ideal point. In the case of three decision levels, the districts pattern is only $(K_1, K_2, K_3) = (9, 3, 1)$. Table 5 illustrates this case. The policy is decided by voter 8, who is 6 positions away from the median. This is even further from the median than in the case of two decision levels. This example suggests that the final policy decided in the gerrymandered districts becomes further from the median as we add more hierarchical decision levels to the democratic representative system. We revisit this fact in Proposition 1 in the next section. Additionally, from the result of Lemma 1, many voters are required to construct a higher hierarchical representative system. □

Table 5: The case of $N = 27$, $K_1 = 9, K_2 = 3, K_3 = 1$

$\{1, 2, 27\}$	$\{3, 4, 26\}$	$\{5, 6, 25\}$	$\{7, 8, 24\}$	$\{9, 10, 23\}$	$\{11, 12, 22\}$	$\{13, 14, 21\}$	$\{15, 16, 20\}$	$\{17, 18, 19\}$
$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$
2	4	6	8	10	12	14	16	18
				$\downarrow$				
$\{2$	4	$18\}$	$\{6$	8	$16\}$	$\{10$	12	$14\}$
	$\downarrow$			$\downarrow$			$\downarrow$	
	$\{4$			8			$12\}$	
				$\downarrow$				
				8				

Actually, in Example 1, where $N = 27$, the liberal extremist can only construct at most three decision levels if we require strictly equal-sized districts and odd district sizes. We already mentioned in footnote 6 that Lemma 1 is not as restrictive as it might first appear. To see this we provide another example with $N = 27$ voters.

Example 2. Let $N = 27$ and the target district size be 2, which means that districts contain either 1 or 2 voters. Clearly, if we put just one voter in a district, she will be the winner of her district. If we put two voters in a district, the median voter is not uniquely defined. In this example, we assume that the left of the two voters in the middle will be elected. [9] Table 6 illustrates this case. $\square$

We make two observations based on Example 2. First, for a given N, we can have more biased outcomes in the case of indivisibilities and even-numbered district sizes. This statement will become clearer after the discussion of Lemma 2. Second, once almost-equal-sized districts and even-numbered district sizes are allowed, while neither is allowed in Lemma 1, the maximum number of levels monotonically increases as the number of voters increases. In particular, if at the i-th decision level, K_{i-1} is not divisible by K_i, then there exists a unique positive integer s_i and integer $0 \leq r_i < K_i$ such that $K_{i-1} = s_i K_i - r_i$. Then, we have $s_i = \lceil K_{i-1}/K_i \rceil$, and we can form $K_i - r_i$ districts of size s_i and r_i districts of size $s_i - 1$. Clearly, we can obtain a maximum number of decision levels by taking sequence $K_i = N - i$ as the number of districts

[9] Clearly, one could choose the right-hand voter or determine the winner randomly. In the latter case, the most extreme outcome would be if, at each level and each district, the left or right voter was always chosen.

Table 6: The case of $N = 27,\ K_1 = 14, K_2 = 7, K_3 = 4, K_4 = 2, K_5 = 1$

{①, 27}	{②, 26}	{③, 25}	{④, 24}	{⑤, 23}	{⑥, 22}	{⑦, 21}	{⑧, 20}	{⑨, 19}	{⑩, 18}	{⑪, 17}	{⑫, 16}	{⑬, 15}	{⑭}
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
1	2	3	4	5	6	7	8	9	10	11	12	13	14
						↓							
{①	14}	{②	13}	{③	12}	{④	11}	{⑤	10}	{⑥	9}	{⑦	8}
↓		↓		↓		↓		↓		↓		↓	
1		2		3		4		5		6		7	
{①		7}		{②		6}		{③		5}		{④}	
↓				↓				↓				↓	
1				2				3				4	
{①				4}				{②				3}	
↓								↓					
1								2					
{①								2}					
↓													
1													

at level i. In this case, the number of decision levels is equal to the number of voters.

A more natural way to generate a sequence $K_0, K_1, \ldots, K_t, K_{t+1}$ would be to fix a sequence $s_1, s_2, \ldots, s_t$ of target district sizes. This means we have either s_i or $s_i - 1$ representatives in each district at decision level i, the number of districts K_i at decision level i is determined by $K_{i-1} = s_i K_i - r_i$, where $0 \leq r_i < K_i$ and $s_i = \lceil K_{i-1}/K_i \rceil$, and we have $K_i - r_i$ districts of size s_i and r_i districts of size $s_i - 1$. Of course, not every sequence $s_1, s_2, \ldots, s_t$ of positive integers is admissible. It would make sense to look for sequences that remain constant as long as possible (i.e., $s = s_1 = s_2 = \ldots = s_q$ for a $q \leq t$ as large as possible), since then the number of representatives decreases at the same rate as we move upwards in the hierarchy. If choosing s as the target size for each decision level is possible, then the number of hierarchical decision levels is given by $\lceil \log_s N \rceil$ in the function of N. Hence, as the number of voters tends to infinity, the number of decision levels also tends to infinity.

Finally, we investigate whether s_i is a valid target value for the number of representatives at decision level i. Since r_i has to satisfy $0 \le r_i < K_i$, we obtain $0 \le s_i K_i - K_{i-1} < K_i$, which in turn implies

$$\frac{K_{i-1}}{s_i} \le K_i < \frac{K_{i-1}}{s_i - 1}. \tag{1}$$

A sufficient condition for the existence of an integer K_i satisfying (1) is

$$\frac{K_{i-1}}{s_i - 1} - \frac{K_{i-1}}{s_i} \ge 1 \iff s_i(s_i - 1) \le K_{i-1}. \tag{2}$$

Since $s_i(s_i - 1) < s_i^2$, a sufficient condition for the existence of an integer K_i satisfying (1) is given by

$$s_i \le \sqrt{K_{i-1}}. \tag{3}$$

Assuming that we are striving for a constant sequence $s = s_1 = s_2 = \ldots = s_t$, then the violation of condition (3) means that adding two more levels with target district sizes s would be impossible. Therefore, for the top two levels, the target district size s has to be reduced appropriately and, at the same time, we can still achieve at least $\lceil \log_s N \rceil$ decision levels.

3 Extremists' policy implementability

From Example 1, we can conjecture that the final elected voter is getting further from the median of all voters as we add more hierarchical levels of legislatures, when all districts at each level are arranged by a liberal extremist's gerrymandering. The following lemma determines the final elected voter's position in the $t+1$ decision levels, composed of both voters and t hierarchical legislative levels in the indirect democracy. Note that, by Lemma 1, there are at most $t^* + 1$ decision levels in the case of N voters and that the number of decision levels lies between 0 and $t^* + 1$. Then, the final elected voter's position is shown in the next lemma.

Lemma 2. (Generalization of Gilligan and Matsusaka (2006)) *If Assumption 1 is satisfied, then in the case of liberal gerrymandering, the policy outcome is decided by voter*

$$j^*_{t+1,1} = \frac{1}{2}\left(\frac{1}{2}\right)^t \frac{1}{K_t K_{t-1} \cdots K_2 K_1}(K_t + 1)(K_t + K_{t-1})(K_{t-1} + K_{t-2}) \cdots (K_2 + K_1)(K_1 + N) \tag{4}$$

in the political decision system composed of $t+1$ ($t \in \{0, 1, 2, \ldots, t^*\}$) hierarchical decision levels.

Proof. We prove formula (4) by backward induction. First, the final representative (she) is elected at the $t+1$-th decision level. Given that $K_{t+1} = 1$, for her final win, she needs at least a minimum majority: that is, $\frac{1}{2}\left(\frac{K_t}{K_{t+1}} + 1\right) = \frac{1}{2}(K_t + 1)$ representatives with smaller numbers than hers and herself at the $t+1$-th decision level.

Second, since each representative at the $t+1$-th decision level is elected in each district at the t-th decision level, to win a majority at $t+1$-th decision level, she needs at least $\frac{1}{2}(K_t+1)$ districts, including $\frac{1}{2}\left(\frac{K_{t-1}}{K_t} + 1\right)$ representatives with smaller numbers than hers per district and herself at the t-th decision level. Consequently, she needs $\frac{1}{2}\left(\frac{K_{t-1}}{K_t} + 1\right)\frac{1}{2}(K_t+1)$ representatives with smaller numbers than hers and herself at the t-th decision level.

Third, given that each representative at the t-th decision level is elected in each district at the $t-1$-th level, for the final representative to win with $\frac{1}{2}\left(\frac{K_{t-1}}{K_t} + 1\right)\frac{1}{2}(K_t+1)$ representatives supporting her at the t-th decision level, she needs at least $\frac{1}{2}\left(\frac{K_{t-1}}{K_t} + 1\right)\frac{1}{2}(K_t+1)$ districts, including $\frac{1}{2}(\frac{K_{t-2}}{K_{t-1}}+1)$ representatives with smaller numbers than hers and herself per district at the $t-1$-th level. Consequently, she needs $\frac{1}{2}(\frac{K_{t-2}}{K_{t-1}}+1)\frac{1}{2}(\frac{K_{t-1}}{K_t}+1)\frac{1}{2}(K_t+1)$ representatives with smaller numbers than hers and herself at the $t-1$-th decision level. The same logic continues until we reach the voters level.

Finally, at voters level (i.e, the first level), she needs at least

$$\frac{1}{2}\left(\frac{K_0}{K_1}+1\right)\frac{1}{2}\left(\frac{K_1}{K_2}+1\right)\cdots\frac{1}{2}\left(\frac{K_{t-2}}{K_{t-1}}+1\right)\frac{1}{2}\left(\frac{K_{t-1}}{K_t}+1\right)\frac{1}{2}(K_t+1) \tag{5}$$

representatives with smaller numbers than hers and herself. This formula is equal to (4). Hence the position of the final representative with the smallest number is (4). □

Lemma 2 includes the case of the most extremely liberal gerrymandering with $t+1$ decision levels. By symmetry of voters, we can easily obtain the position $N - j^*_{t+1,1} + 1$ by the most extremely conservative gerrymandering. Additionally, (4) also indicates the representative position in the direct democracy at $t = 0$. Thus, the direct democracy can be viewed as a special case of this hierarchical legislature.

Focusing on (5), in the most extremely liberal gerrymandering with $t+1$ decision levels, the final elected voter's position is determined by constructing districts as nested boxes, with minimum majorities composed of her and the voters with lower values than hers so that only they have the decision power. In other words, voters with higher values than hers have no influence in this gerrymandering. Thus, the result does not depend on the order of packing voters into districts.

Example 2 already explained how we can get rid of our assumptions of equal- and odd-sized districts (i.e., Assumption 1). Now, we explain how to obtain an upper bound on the most leftward voter deciding the policy outcome, based on Lemma 2, if we also allow for almost equal and even-sized districts, as described in Example 2.

Lemma 3. *In the case of liberal gerrymandering, we have the following upper bound on the voter deciding the policy outcome:*

$$j^*_{t+1,1} \leq \frac{1}{2}\left(\frac{1}{2}\right)^t \frac{1}{K_t K_{t-1} \cdots K_2 K_1}(K_t+1)(K_t+K_{t-1})(K_{t-1}+K_{t-2})\cdots(K_2+K_1)(K_1+N) \quad (6)$$

in the political decision system composed of $t+1$ *hierarchical decision levels.*[10]

Proof. In order to obtain an upper bound for $j^*_{t+1,1}$ for the more general case, we just have to follow the steps of the proof of Lemma 2. First, if K_t is even, the final representative needs at least a minimum majority: that is, $\frac{1}{2}K_t$ representatives with smaller numbers than hers and herself at the $t+1$-th decision level, which is less than the $\frac{1}{2}(K_t+1)$ obtained for the odd-size case. Second, since each representative at the $t+1$-th decision level is elected in each district at the t-th decision level, to win a majority at the $t+1$-th decision level, she needs at least either $\frac{1}{2}K_t$ or $\frac{1}{2}(K_t+1)$ districts, including either $\frac{1}{2}\frac{K_{t-1}}{K_t}$ representatives from each of the even-sized districts, or $\frac{1}{2}\left(\frac{K_{t-1}}{K_t}+1\right)$ representatives from each of the odd-sized districts with smaller numbers than hers per district and herself at the t-th decision level. Consequently, she does not need more than $\frac{1}{2}\left(\frac{K_{t-1}}{K_t}+1\right)\frac{1}{2}(K_t+1)$ representatives with smaller numbers than hers and herself at the t-th decision level. By analogous modifications in the remaining steps

[10]Note that the indexes $j^*_{i,k}$ can be extended in an obvious way to the case in which Assumption 1 does not hold. More precisely, if an even district size emerges at level i in district k in our electoral system, then the actual value of $j^*_{i,k}$ depends on how we choose between the two voters in the "middle." Here, and in what follows, we assume in the extension of $j^*_{i,k}$ that the left of the two representatives in the middle will be elected to the next higher level, since this results in the most biased outcome. For more on this, we refer to the discussion after Example 2.

of the proof of Lemma 2, we find that the right-hand side of (4) provides an upper bound on $j^*_{t+1,1}$ for the general case. $\square$

Formula (4) is obtained as the representative with the policy that is implementable and nearest to the liberal extremist's ideal policy for given $t+1$ decision levels and N. As a consequence of the formula, the next proposition states the monotonic decrease of the representative's position as t increases for given N.

Proposition 1. *Under Assumption 1, for given N, the final representative position is moving further from the median of voters monotonically as the number of hierarchical stages increases.*

Proof. Let $N = a_1 \cdot a_2 \cdot a_3 \ldots \cdot a_{t^*} \cdot a_{t^*+1}$, where all of a_i, $(i = 1, \ldots, t^* + 1)$ are prime factors of N. From Lemma 1, $N = \frac{N}{K_1} \cdot \frac{K_1}{K_2} \cdot \frac{K_2}{K_3} \cdot \ldots \cdot \frac{K_{t^*-1}}{K_{t^*}} \cdot \frac{K_{t^*}}{K_{t^*+1}}$, noting that $K_{t^*+1} = 1$ in the case of the maximum number of hierarchical levels, $t^* + 1$. Without loss of generality, we let $\frac{N}{K_1}, \frac{K_1}{K_2}, \frac{K_2}{K_3}, \ldots, \frac{K_{t^*-1}}{K_{t^*}}, \frac{K_{t^*}}{K_{t^*+1}}$ equal $a_1, a_2, a_3, \ldots, a_{t^*}, a_{t^*+1}$, respectively. Then, we obtain $K_1 = \frac{N}{a_1}$, $K_2 = \frac{N}{a_1 a_2}$, $K_3 = \frac{N}{a_1 a_2 a_3}$, …, $K_{t^*} = \frac{N}{a_1 a_2 a_3 \ldots a_{t^*}}$, $K_{t^*+1} = \frac{N}{a_1 a_2 a_3 \ldots a_{t^*} a_{t^*+1}} = 1$.

In the case in which there are $t+1$ $(t = 1, \ldots, t^*)$ decision levels, $N = b_1 b_2 \ldots b_{t+1}$, where each b_i $(i = 1, \ldots, t)$ is a product of some prime factors of N. Using the same method, we obtain $K_1 = \frac{N}{b_1}$, $K_2 = \frac{N}{b_1 b_2}$, $K_3 = \frac{N}{b_1 b_2 b_3}$, …, $K_t = \frac{N}{b_1 b_2 b_3 \ldots b_t}$, $K_{t+1} = \frac{N}{b_1 b_2 b_3 \ldots b_t b_{t+1}} = 1$. Here, when the number of total levels is t, we choose any two factors in $\{b_1, b_2, \ldots, b_{t+1}\}$ and find the product. For simplicity, we choose the last two factors, $b_t b_{t+1}$. Then, we get $K_1 = \frac{N}{b_1}$, $K_2 = \frac{N}{b_1 b_2}$, $K_3 = \frac{N}{b_1 b_2 b_3}$, …, $K_{t-1} = \frac{N}{b_1 b_2 b_3 \ldots b_{t-1}}$, $K'_t = \frac{N}{b_1 b_2 b_3 \ldots b_t b_{t+1}} = 1$. Comparing the case of t levels with the case of $t+1$ levels, each of K_1 through K_{t-1} is identical.

We calculate the ratio of $j^*_{t,1}$ and $j^*_{t+1,1}$:

$$\begin{aligned}
\frac{j^*_{t,1}}{j^*_{t+1,1}} &= \frac{\frac{1}{2}\left(\frac{1}{2}\right)^{t-1} \frac{1}{K_{t-1}K_{t-2}\cdots K_2 K_1}(K_{t-1}+1)(K_{t-1}+K_{t-2})(K_{t-2}+K_{t-3})\cdots(K_2+K_1)(K_1+N)}{\frac{1}{2}\left(\frac{1}{2}\right)^{t} \frac{1}{K_t K_{t-1}\cdots K_2 K_1}(K_t+1)(K_t+K_{t-1})(K_{t-1}+K_{t-2})\cdots(K_2+K_1)(K_1+N)} \\
&= \frac{K_{t-1}+1}{\frac{1}{2}\frac{1}{K_t}(K_t+1)(K_t+K_{t-1})} \\
&= \frac{2K_t(K_{t-1}+1)}{(K_t+1)(K_t+K_{t-1})}.
\end{aligned}$$

Here, since $\forall t \in \{1, 2, 3, \ldots, t^*\}$, $K_{t-1} > K_t > 1$, and $2K_t(K_{t-1}+1) - (K_t+1)(K_t+K_{t-1}) = (K_t - 1)(K_{t-1} - K_t) > 0$, the numerator is larger than the denominator. Thus, $j^*_{t,1} > j^*_{t+1,1}$, and therefore, $x^*_{t,1} > x^*_{t+1,1}$. $\square$

By this proposition, we can say that a voter with a policy position nearer to the extremist's ideal policy becomes electable as the decision levels get higher for given N. In other words, the hierarchical legislative system with more levels increases the policy implementability of the extremist.

For given numbers of voters N and levels t, we can determine the number of districts $K_1, \ldots, K_t$ that results in the most biased outcome that is the furthest representative position from the median. Noting that $K_0 = N$, the next proposition can be obtained.

Proposition 2. *If Assumption 1 is satisfied, then in the case of N voters and $t+1$ decision levels for which $\sqrt[t+1]{N}$ is an integer, the multi-level districting given by*

$$\forall i = 1, \ldots, t, t+1 : K_i = N^{\frac{t-(i-1)}{t+1}} \tag{7}$$

admits the largest possible bias.

Proof. Determining the maximum bias is equivalent to minimizing $j^*_{t+1,1}/N$ with respect to $K_1, \ldots, K_t$. Noting that $N = K_0$,

$$\frac{j^*_{t+1,1}}{N} = \left(\frac{1}{2}\right)^{t+1} \frac{1}{K_t K_{t-1} \cdots K_2 K_1 K_0} (K_t + 1)(K_t + K_{t-1})(K_{t-1} + K_{t-2}) \cdots (K_2 + K_1)(K_1 + K_0),$$

we multiply $1/K_0, 1/K_1, \ldots, 1/K_t$ by the terms in the previous formula in reverse order. Hence, we have to minimize the expression

$$\left(\frac{1}{2}\right)^{t+1} \left(1 + \frac{1}{K_t}\right) \left(1 + \frac{K_t}{K_{t-1}}\right) \cdots \left(1 + \frac{K_2}{K_1}\right) \left(1 + \frac{K_1}{K_0}\right), \tag{8}$$

yielding first-order conditions equivalent to

$$0 = -\frac{K_{i+1}}{K_i^2} \left(1 + \frac{K_i}{K_{i-1}}\right) + \left(1 + \frac{K_{i+1}}{K_i}\right) \frac{1}{K_{i-1}},$$

for all $i = 1, \ldots, t$. Thus, by simple rearrangements, we obtain

$$K_{i+1} K_{i-1} = K_i^2, \tag{9}$$

for all $i = 1, \ldots, t$, where $K_{t+1} = 1$. It can be verified that the first-order conditions determine the minimum value for expression (8).

We claim that

$$K_{t-(i-1)} = (K_{t-i})^{\frac{i}{i+1}} \,, \tag{10}$$

for all $i = 1, \ldots, t$, which we prove by induction. Clearly, our claim holds true for $i = 1$, by equation (9). Assume that (10) is valid for i. Therefore, we show that it also holds true for $i + 1$. From (9), we have

$$K_{t-i+1} K_{t-i-1} = (K_{t-i})^2 \,,$$

and by employing our induction hypothesis, we get

$$(K_{t-i})^{\frac{i}{i+1}} K_{t-i-1} = (K_{t-i})^2 \quad \Leftrightarrow \quad K_{t-i} = (K_{t-i-1})^{\frac{i+1}{i+2}} \,,$$

which is what we wanted to show.

Finally, by employing (10) recursively, we obtain the statement of our proposition. □

Assuming in Proposition 2 that $\sqrt[t+1]{N}$ is an integer may appear to be too restrictive.[11] Clearly, for arbitrary combinations of N and t, the sequence $(K_i)_{i=1}^{t+1}$ given by equation (7) is typically a non-integer sequence and does not determine legitimate numbers of districts for all levels. However, we only use the result of Proposition 2 to get an idea of how to find a lower bound for the largest possible bias while keeping t fixed.

Proposition 3. *If the number of levels t is fixed, then as the number of voters N tends to infinity, we have the following lower bound on the bias of the most biased case:*

$$\lim_{N \to \infty} \max \left| F(x^*_{t+1,1}) - F(x^*_{POP}) \right| \geq \frac{1}{2} - \left(\frac{1}{2}\right)^{t+1} .$$

Proof. Noting that $\lim\limits_{N \to +\infty} F(x^*_{POP}) = 1/2$, from Lemma 3, we get

$$\begin{aligned} \lim_{N \to \infty} \max B(x) \;&=\; \frac{1}{2} - \frac{j^*_{t+1,1}}{N} \geq \frac{1}{2} - \frac{\frac{1}{2}\left(\frac{K_0}{K_1} + 1\right) \frac{1}{2}\left(\frac{K_1}{K_2} + 1\right) \cdots \frac{1}{2}\left(\frac{K_{t-1}}{K_t} + 1\right) \frac{1}{2}\left(K_t + 1\right)}{\frac{K_0}{K_1} \frac{K_1}{K_2} \cdots \frac{K_{t-1}}{K_t} \frac{K_t}{K_{t+1}}} \\ &=\; \frac{1}{2} - \frac{1}{2}\left(1 + \frac{K_1}{K_0}\right) \frac{1}{2}\left(1 + \frac{K_2}{K_1}\right) \cdots \frac{1}{2}\left(1 + \frac{K_t}{K_{t-1}}\right) \frac{1}{2}\left(1 + K_t\right). \end{aligned} \tag{11}$$

Now, in the hierarchical construction described in Example 2, we substitute $\lfloor \sqrt[t+1]{N} \rfloor$ for the target district size s. Since we are only interested in the case of many voters, we can assume,

[11] Not to mention, if all prime factors of N are identical, for example, $N = a^{t+1}$, where a is a prime, $\sqrt[t+1]{N}$ is always an integer.

without loss of generality, that $\lfloor \log_s N \rfloor - 1 \geq t$. It can be verified that for fixed t, for the determined value of s, and for sufficiently large N, a sequence $K_1, \ldots, K_t$ of numbers of districts can be chosen in the way described at the end of Example 2. Hence, substituting the target district sizes into (11), we obtain

$$\lim_{N\to\infty} \max B(x) \geq \lim_{N\to\infty} \frac{1}{2} - \left(\frac{1}{2}\right)^{t+1} \left(1+\frac{1}{s}\right)\left(1+\frac{1}{s}\right)\cdots\left(1+\frac{1}{s}\right) = \frac{1}{2} - \left(\frac{1}{2}\right)^{t+1},$$

since s tends to infinity as N tends to infinity for a given t. □

Noting that (8) is not $j^*_{t+1,1}$, but rather $j^*_{t+1,1}/N$, which is the relative position of the final elected voter, this proposition says that the relative position of the voter moves left as N increases. Now, from Proposition 3, it follows that as the maximum number of levels t increases, the maximum bias approaches its highest possible level, which we state in the next corollary.

Corollary 1. *As t tends to infinity, the maximum bias in the case of liberal gerrymandering tends to* $1/2$.

Proposition 3 and Corollary 1 mean that the more voters there are, the more extreme the relative position of the policy that political extremists can realize becomes. As a result, many voters provides the extreme district maker with an expedient way to implement her favorite policy. In particular, if party A is at the left end and party B is at the right end of the unit interval, then, party A's most preferred outcome will be the national outcome, independent of voters' preferences.

Our model is applicable to other democratic issues. In the Appendix, we apply our model to solve and explain the issues of random districting and the partisan bias.

4 Moderates' policy implementability

So far, we have only considered the cracking and packing method in the case in which a political extremist is a district maker who wants to implement a policy as low or as high as possible. However, someone with a moderate political position around the median can also be a district maker, and she would consider how to organize districts so as to implement her favorite policy. If there is at least one districting method that makes it possible to implement the district maker's ideal policy, this districting method is one of her implementations. We have already had the electable and most extreme representative $j^*_{t+1,1}$. Thus, we can say that the districting method is a powerful implementation method in elections for a district maker with any political position if all voters from $j^*_{t+1,1}$ to $\frac{N+1}{2}$ (i.e., from $j^*_{t+1,1}$ to $N - j^*_{t+1,1} + 1$, by symmetry), are also electable as the final representative. In addition, from Proposition 3 and Corollary 1, we can say that almost all voters' favorite policies relative to all voters' become implementable as N increases.

Is any moderate's favorite policy actually implementable? The answer is yes, and she can implement the policy by generalizing the cracking and packing method described in the previous sections, hereafter referred to as the generalized cracking and packing method. Note that the interval we have to consider is only $\{1, 2, \ldots, \frac{N+1}{2}\}$, by symmetry. We define the district number to which $j^*_{t+1,1}$ belongs at each decision level in the cracking and packing method as

$$m_1 \equiv \frac{j^*_{t+1,1}}{\frac{1}{2}(\frac{N}{K_1}+1)} = \frac{1}{2}(K_t+1)\frac{1}{2}\left(\frac{K_{t-1}}{K_t}+1\right)\frac{1}{2}\left(\frac{K_{t-2}}{K_{t-1}}+1\right)\cdots\frac{1}{2}\left(\frac{K_1}{K_2}+1\right),$$

$m_2 \equiv \frac{m_1}{\frac{1}{2}(\frac{K_1}{K_2}+1)}$, $m_3 \equiv \frac{m_2}{\frac{1}{2}(\frac{K_2}{K_3}+1)}$, $\ldots$, $m_t \equiv \frac{m_{t-1}}{\frac{1}{2}(\frac{K_{t-1}}{K_t}+1)}$, and $m_{t+1} \equiv \frac{m_t}{\frac{1}{2}(K_t+1)} = 1$, since there are minimal majority voters or representatives with ideal points left of $j^*_{t+1,1}$. In particular, there are $\frac{1}{2}(\frac{K_{i-1}}{K_i}+1)$ voters of such type for any $i \in \{1, 2, 3, \ldots, t, t+1\}$, including $j^*_{t+1,1}$ at each decision level. For convenience, we also define $m_0 = j^*_{t+1,1}$. Then, in the cracking and packing method, noting that $j^*_{t+1,1}$ is included in the m_1-th district at the first decision level, from the definition of m_i, the cardinality of the maximum minority voters with a higher index is equal to $\frac{1}{2}(\frac{K_{i-1}}{K_i}-1)$, which we insert from the first district through the K_i-th district. This is shown in Table 7.

Table 7: Each district by the cracking and packing method at the i-th decision level

District number	The position of each voter: 1	2	…	$\frac{1}{2}(\frac{K_{i-1}}{K_i}+1)$ = median	$\frac{1}{2}(\frac{K_{i-1}}{K_i}+1)+1$	$\frac{1}{2}(\frac{K_{i-1}}{K_i}+1)+2$	…	$\frac{1}{2}(\frac{K_{i-1}}{K_i}+1)+\frac{1}{2}(\frac{K_{i-1}}{K_i}-1)=\frac{K_{i-1}}{K_i}$
1	1	2	…	A_i	K_{i-1}	$K_{i-1}-1$	…	$K_{i-1}-B_i+1$
2	A_i+1	A_i+2	…	$2A_i$	$K_{i-1}-B_i$	$K_{i-1}-B_i-1$	…	$K_{i-1}-2B_i+1$
$\vdots$		$\vdots$				$\vdots$		
m_i-1	$(m_i-2)A_i+1$	$(m_i-2)A_i+2$	…	$(m_i-1)A_i$	$K_{i-1}-(m_i-2)B_i$	$K_{i-1}-(m_i-2)B_i-1$	…	$K_{i-1}-(m_i-1)B_i+1$
m_i	$(m_i-1)A_i+1$	$(m_i-1)A_i+2$	…	m_iA_i	$K_{i-1}-(m_i-1)B_i$	$K_{i-1}-(m_i-1)B_i-1$	…	$K_{i-1}-m_iB_i+1$
m_i+1	m_iA_i+1	m_iA_i+2	…	$(m_i+1)A_i$	$K_{i-1}-m_iB_i$	$K_{i-1}-m_iB_i-1$	…	$K_{i-1}-(m_i+1)B_i+1$
$\vdots$		$\vdots$				$\vdots$		
K_i-1	$(K_i-2)A_i+1$	$(K_i-2)A_i+2$	…	$(K_i-1)A_i$	$K_{i-1}-(K_i-2)B_i$	$K_{i-1}-(K_i-2)B_i-1$	…	$K_{i-1}-(K_i-1)B_i+1$
K_i	$(K_i-1)A_i+1$	$(K_i-1)A_i+2$	…	K_iA_i	$K_{i-1}-(K_i-1)B_i$	$K_{i-1}-(K_i-1)B_i-1$	…	$K_{i-1}-K_iB_i+1$
	Left wing				Right wing			

$A_i \equiv \frac{1}{2}(\frac{K_{i-1}}{K_i}+1)$, $B_i \equiv \frac{1}{2}(\frac{K_{i-1}}{K_i}-1)$

The numbers in the left wing are shown in ascending order and those in the right wing are shown in descending order. Thus, the number K_iA_i that appears in the K_i-th row and the $\frac{1}{2}(\frac{K_{i-1}}{K_i}+1)$-th column is equal to $K_{i-1}-K_iB_i$, which is one less than $K_{i-1}-K_iB_i+1$ at the K_i-th row and the $\frac{K_{i-1}}{K_i}$-th column. That is, $K_iA_i = \frac{1}{2}(K_{i-1}+K_i) = K_{i-1} - \frac{1}{2}(K_{i-1}-K_i) = K_{i-1}-K_iB_i$. Thus, both numbers are consecutive.

In the cracking and packing method, the numbers of district medians appear at every $\frac{1}{2}(\frac{K_{i-1}}{K_i}+1)$ positions because of the cracking of all voters into minimum majorities with lower values and maximum minorities with higher values, and then packing each into one district. Here, we refer to the minimum majority voters with lower values in each district as "Left wing" and the maximum minority voters with higher values as "Right wing." Now, we focus on the median of each district, especially the m_i-th through the $\frac{K_i+1}{2}$-th district, after the following manipulations. First, in Table 7, we remove the last-position voter in the Right wing of the K_i-th district and slide all voters forward one position in the Right wings of all districts. Then, since voter K_{i-1} also moves forward one position, we can create a vacant position at the first position in the Right wing of the first district. Second, we insert the first-position voter A_i+1 in the Left wing of the second district into the vacant position instead of using voter K_{i-1}. Third, we slide back all voters in the Left wings from the second through the K_i-th districts by one position, and insert the removed voter from the last position in the Right wing of the K_i-th district into the last position in the Left wing of the K_i district, which is now vacant. We refer to this series of manipulations as a cycle. By repeating the cycle, the positions of the voters in the first district become those shown in Table 8.

Table 8: The positions of voters in the first district at the i-th decision level

	Voters' positions						
Sliding by	…	median	median +1	median +2	median+3	…	median+$B_i = \frac{K_{i-1}}{K_i}$
0		A_i	K_{i-1}	$K_{i-1}-1$	$K_{i-1}-2$		$K_{i-1}-B_i+1$
1		A_i	A_i+1	K_{i-1}	$K_{i-1}-1$		$K_{i-1}-B_i+2$
2	…	A_i	A_i+1	A_i+2	K_{i-1}	…	$K_{i-1}-B_i+3$
⋮			⋮			⋮	
$\frac{1}{2}(\frac{K_{i-1}}{K_i}-1)$		A_i	A_i+1	A_i+2	A_i+3		$A_i+B_i=\frac{K_{i-1}}{K_i}$.
		Left wing			Right wing		

By virtue of the cycle, the medians in the second district through the K_i-th district slide and are replaced by each voter with one higher value. Noting that there are $B_i=\frac{1}{2}(\frac{K_{i-1}}{K_i}-1)$ voters in the Right wing in each district, and that there are $(\frac{K_i+1}{2}-1)B_i$ voters in the Right wings of the first through the $\frac{K_i+1}{2}-1$-th districts, we can also apply the cycle in the second through $\frac{K_i+1}{2}-1$-th districts, repeatedly. Then, we can slide $(\frac{K_i+1}{2}-1)B_i$ voters, and voters in the first through the $\frac{K_i+1}{2}-1$-th district in the Right wing are lined up in ascending order. As a result, voters 1 through the median of all voters are lined up in ascending order in the first through $\frac{K_i+1}{2}$-th districts.

Now, we check that any voter between $j^*_{t+1,1}$ and the median of all voters is electable as the final representative. Focusing on the medians of the m_i-th district through the $\frac{K_i+1}{2}$-th district (i.e., the median district), we extract the medians of those districts. Then, we obtain Table 9. Note that, by sliding $(m_i-1)B_i$ positions, since all voters in the Right wing of the first

Table 9: Each district median of the m_i-th through the $\frac{K_i+1}{2}$-th at the first decision level

Sliding by	Median voter of the m_i-th dist.	m_i+1-th dist.	m_i+2-th dist.	…	$\frac{K_i+1}{2}$-th dist.	…
0	m_iA_i	$(m_i+1)A_i$	$(m_i+2)A_i$	…	$\frac{K_i+1}{2}A_i$	…
1	m_iA_i+1	$(m_i+1)A_i+1$	$(m_i+2)A_i+1$		$\frac{K_i+1}{2}A_i+1$	
2	m_iA_i+2	$(m_i+1)A_i+2$	$(m_i+2)A_i+2$		$\frac{K_i+1}{2}A_i+2$	…
$\vdots$		$\vdots$			$\vdots$	
$(m_i-1)B_i$	m_iA_i $+(m_i-1)B_i$	$(m_i+1)A_i$ $+(m_i-1)B_i$	$(m_i+2)A_i$ $+(m_i-1)B_i$	…	$\frac{K_i+1}{2}A_i$ $+(m_i-1)B_i$	…
$(m_i-1)B_i+1$	m_iA_i $+(m_i-1)B_i$	$(m_i+1)A_i$ $+(m_i-1)B_i+1$	$(m_i+2)A_i$ $+(m_i-1)B_i+1$	…	$\frac{K_i+1}{2}A_i$ $+(m_i-1)B_i+1$	…
$(m_i-1)B_i+2$	m_iA_i $+(m_i-1)B_i$	$(m_i+1)A_i$ $+(m_i-1)B_i+2$	$(m_i+2)A_i$ $+(m_i-1)B_i+2$	…	$\frac{K_i+1}{2}A_i$ $+(m_i-1)B_i+2$	…
$\vdots$	$\vdots$	$\vdots$			$\vdots$	
m_iB_i	m_iA_i $+(m_i-1)B_i$	$(m_i+1)A_i$ $+m_iB_i$	$(m_i+2)A_i$ $+m_iB_i$	…	$\frac{K_i+1}{2}A_i$ $+m_iB_i$	…
$\vdots$		$\vdots$			$\vdots$	
$(\frac{K_i+1}{2}-1)B_i$	m_iA_i $+(m_i-1)B_i$	$(m_i+1)A_i$ $+m_iB_i+1$	$(m_i+2)A_i$ $+(m_i+1)B_i$	…	$\frac{K_i+1}{2}A_i$ $+(\frac{K_i+1}{2}-1)B_i$	…
Group number	m_i	m_i+1	m_i+2	…	$\frac{K_i+1}{2}$	…

district through the m_i-1-th district have already been replaced by voter A_i+1 through voter

$A_i + (m_i - 1)B_i$, there is no voter to be replaced before voter $m_iA_i + (m_i - 1)B_i$ in the m_i-th district. Thus, the median of the m_i-th district cannot slide further, so voter $m_iA_i+(m_i-1)B_i$ is unchanged after sliding by $(m_i - 1)B_i$ in Table 9. Similarly, the medians of the $m_i + 1$-th through the $\frac{K_i+1}{2}$-th districts are also unchanged after replacing the voters in the Right wing and lining up the consecutive numbers in each district. This is true each time we slide by B_i positions after sliding by $(m_i - 1)B_i$ positions.

Focusing on the last column in the last row, sliding by $(\frac{K_i+1}{2} - 1)B_i$ positions, in Table 9, voter $\frac{K_i+1}{2}A_i + \left(\frac{K_i+1}{2} - 1\right) B_i$ is equal to $\frac{K_{i-1}+1}{2}$. When sliding by $(\frac{K_i+1}{2} - 1)B_i$ positions, all representatives in the first through $\frac{K_i+1}{2}$-th districts are lined up in ascending order at the i-th decision level. In other words, we can refer to the case shown in the $(\frac{K_i+1}{2} - 1)B_i$-th row of the table as the "full sliding." When $i = 1$ (i.e., the first decision level), voter m_iA_i in the first row and the first column in Table 9 is equal to $j^*_{t+1,1} = m_0$, and is the most extreme voter electable as the final representative, by the definition of m_i. In addition, voter $\frac{K_i+1}{2}A_i + (\frac{K_i+1}{2} - 1)B_i$ in the last row and the last column is equal to $\frac{N+1}{2}$, and is the median of all voters.

Note that the same numbers may appear in Table 9, [12] as shown in Example 3 in the Appendix. Thus, we have all numbers between $j^*_{t+1,1}$ and $\frac{N+1}{2}$, which means those voters are electable as representatives of the $i + 1$-th decision level.

In Table 9, we refer to the medians of the m_i-th district as Group m_i, those of the m_i+1-th as Group $m_i + 1$, and so on. Lastly, those of $\frac{K_i+1}{2}$-th district are Group $\frac{K_i+1}{2}$. To elect a voter as a district representative to the $i + 1$-th decision level, we need to choose a row l including the voter, $l \in \{0, 1, \ldots, (\frac{K_i+1}{2} - 1)B_i\}$, from Table 9. Then, if the voter is in Group m_i, we district all representatives by the cracking and packing method at each decision level from the $i + 1$-th to the t-th levels without sliding any positions. This is because the voter is already at the same position as $j^*_{t+1,1}$'s in the case of the cracking and packing method. If the voter is

[12] In Table 9, the condition of repeated or consecutive numbers lined up between the m_i-th district and the $\frac{K_i+1}{2}$-th district is $m_iA_i + (m_i - 1)B_i + 1 \geq (m_i + 1)A_i$ (i.e., $\frac{K_{i-1}}{K_i} \geq \frac{m_i-1}{m_i-2}$), since the m_i-th column has the least number of sliding positions. Here, the minimum $\frac{K_{i-1}}{K_i}$ is 3, because it is the population per district. Thus, $\frac{K_{i-1}}{K_i} \geq 3 \geq \frac{m_i-1}{m_i-2}$. From this inequality, $m_i \geq \frac{5}{2}$ is obtained. Consequently, noting that m_i is an integer, $m_i \geq 3$ is needed.

in Group $m_i + 1$, she is out by one position from the $j^*_{t+1,1}$'s position. If the voter is in Group $m_i + 2$, she is out by two positions, if in Group $m_i + 3$, she is out by three positions, and so on. Thus, we have to slide the voter by the number of positions that her group is away from Group m_i at the $i+1$-th decision level to elect her as a district representative to the $i+2$-th decision level.

For simplicity, we renumber representatives at the $i+2$-th level as $\{1, 2, 3, \ldots, m_i, \ldots, \frac{K_i+1}{2}, \ldots, K_i\}$ at the $i+1$-th level. When we slide by 0 positions at the $i+1$-th decision level, each representative of Group m_i can become the district median of the m_{i+1}-th district. When we slide by one position at the $i+1$-th decision level, each representative of Group $m_i + 1$ can become the district median of the m_{i+1}-th district, and so on. Sliding each district representative individually, and noting that $m_{i+1} = \dfrac{m_i}{\frac{1}{2}(\frac{K_i}{K_{i+1}} + 1)}$, we have the same table at the $i+1$-th decision level as shown in Table 9, where i is replaced by $i+1$.

Noting that each decision level has the same structure without populations, all elected representatives are renumbered from one to K_i, $i = \{1, 2, 3, \ldots, t\}$ at each level. Then, at each decision level above the first, if representatives elected at the decision level below (i.e., the $i-1$-th level) are between Group m_{i-1} and Group $\frac{K_{i-1}+1}{2}$, we can apply the above manipulation to the i-th level. Thus, we have the following lemma.

Lemma 4. *Under Assumption 1, if a renumbered representative elected at the $i-1$-th decision level is a representative between m_{i-1} and $\frac{K_{i-1}+1}{2}$ at the i-th decision level, where there are K_{i-1} representatives, $i \in \{1, 2, 3, \ldots, t\}$, then she is electable as a district representative between the m_i-th district and the $\frac{K_i+1}{2}$-th to the $i+1$-th decision levels.*

Proof. Note that $m_{i-1} = m_i A_i$ by the definition of m_i, since, at the i-th decision level, there are K_{i-1} representatives elected at the $i-1$-th decision level. Thus, all representatives between m_{i-1} and $\frac{K_{i-1}+1}{2}$ are lined up as district medians of the m_i-th through the $\frac{K_i+1}{2}$-th districts in Table 9. Therefore, at the i-th decision level, all representatives between m_{i-1} and $\frac{K_{i-1}+1}{2}$ are electable as a district representative to the $i+1$-th decision level. $\square$

Applying Lemma 4 repeatedly, we have the following proposition.

Proposition 4. *Under Assumption 1, any voter between $j^*_{t+1,1}$ and the median of all voters $\frac{N+1}{2}$ is electable as the final representative.*

Proof. Since $K_1 > K_2 > K_3 > \ldots > K_t > K_{t+1} = 1$ and $\frac{1}{2}(\frac{K_i}{K_{i+1}} + 1) \geq 1$, we have

$$j^*_{t+1,1} = m_0 > m_1 > m_2 > m_3 > \ldots > m_t > m_{t+1} = 1,$$

from the definition of m_i, and we have

$$\frac{N+1}{2} > \frac{K_1+1}{2} > \frac{K_2+1}{2} > \ldots > \frac{K_t+1}{2} > \frac{K_{t+1}+1}{2} = 1.$$

Thus, by applying the proof of Lemma 4, both m_i and $\frac{K_i+1}{2}$ in Table 9 shrink to 1 as $i \to t$. Lastly, we can say that the voter we want to elect as the final representative between $j^*_{t+1,1}$ and the voters median $\frac{N+1}{2}$ is elected at the final decision level by the sandwich theorem. □

By symmetry, any voters between $j^*_{t+1,1}$ and $N - j^*_{t+1,1} + 1$ are electable as the final representative. Example 3 in the Appendix illustrates the generalized cracking and packing method described here. According to our results, a district maker can implement any policies between $x^*_{t+1,1}$ and $x_{N-j^*_{t+1,1}+1}$ by gerrymandering districts. In addition, from Proposition 3 and Corollary 1, which state that voter $j^*_{t+1,1}$'s relative position to all voters becomes more extreme as the number of voters increases, the more voters there are, the wider is the implementable policy range by gerrymandering.

With regard to the democracy, if we want to avoid the policy bias of a district maker with an extreme political position, we may think that the district maker should be randomly elected from among all voters. However, this may not be effective. We obtain the next corollary from Proposition 4.

Corollary 2. *Under Assumption 1, when each voter becomes the district maker with an equal probability, the voters' positions with the highest probability are those elected by the left and right extremists' gerrymandering districts, $j^*_{t+1,1}$ and $N - j^*_{t+1,1} + 1$, respectively.*

Proof. Since there are N voters, the probability of each voter becoming the district maker is $\frac{1}{N}$. From Proposition 4, any district maker in $\{j^*_{t+1,1}, \ldots, \frac{N+1}{2}\}$ can elect herself as the final representative, or can elect anyone in $\{\frac{N+1}{2}+1,\dots,N-j^*_{t+1,1}+1\}$, by symmetry. Thus, the policies of each $\{j^*_{t+1,1},\dots,\frac{N+1}{2},\dots,N-j^*_{t+1,1}+1\}$ are implemented with probability $\frac{1}{N}$.

On the other hand, voters in $\{1,\dots,j^*_{t+1,1}-1\}$ and $\{N-j^*_{t+1,1}+2,\dots,N\}$ cannot elect themselves as the final representative by gerrymandering districts. Thus, their favorite and implementable policies by gerrymandering districts are the same as those implemented by $j^*_{t+1,1}$ and $N-j^*_{t+1,1}+1$, respectively. Therefore, $j^*_{t+1,1}$ and $N-j^*_{t+1,1}+1$ are elected as the final representative with probability $\frac{j^*_{t+1,1}}{N}$. □

This corollary means that even when the district maker is randomly elected from among all voters, the most likely policies are the same as those implemented by extremists by gerrymandering. As a result, randomly electing a district maker will not reduce the policy bias from in a democracy. In other words, this corollary states that extremists' policy implementability is stronger than that of moderates', even when district makers are chosen randomly.

5 Concluding remarks

In this paper, we introduced a multiple hierarchical legislative system in a representative democracy. Here, a district maker groups voters and intermediate representatives into smaller single-member districts at each legislative level. Then the district maker can gerrymander districts to enable a candidate with a policy position close to hers to win and implement the policy. We showed the range of implementable policies available to the final representative elected in the gerrymandered districts. As the number of legislative levels increases, the range expands. In particular, when the district maker has an extreme policy position, a large number of voters makes it easier for her to implement a policy relatively close to hers because she can construct a legislative system with more levels. In other words, the more voters there are, the stronger is the district maker's policy implementability in the hierarchical representative system. Thus, a hierarchically legislative democracy can be regarded as a method a district maker can use to implement her favorite policy by constructing a multiple hierarchical legislature and gerrymandering districts. In addition, even when the district maker is elected from all voters randomly, extremists have stronger policy implementability than moderate voters. Viewed from the standpoint of a democracy, our results may explain why it is not easy to find examples of multiple hierarchical legislative systems like Jefferson's ward republic.

In our model, we assume that the district maker has all information, and in particular, knows voters' policy preferences. However, in the real world, the district maker has only partial information, although more precise information on the distribution of voters policy preferences is available today than in the past. In particular, the electoral process could be described by the following game to reflect the asymmetric information: 1) each voter sends (intentionally or unintentionally) a message of his/her ideal point; 2) the district maker groups voters into districts hierarchically; 3) each voter casts a ballot sincerely; and 4) the final representative is elected and implements her ideal policy. One important issue in the game is whether voters send messages truthfully. If the district maker announces and commits to districting with t levels and the voters have to commit to their messages (for instance, because they can be monitored by the other members of the same legislature), voters have no incentive to hide or to pretend to have another ideal point in the first stage since a moderate district maker in our hierarchical model selects a voting rule electing the voter with the k-th ranked ideal point from the left. This, according to Moulin (1980) is the strategy proof and the group strategy proof voting rule. As a result, the same outcomes as those in our model are obtained. However, in the literature on asymmetric information, there are several other political situations in which voters cannot commit to anything, do not vote sincerely, or there is no previous announcement. In those cases, various outcomes could appear. These issues need to be addressed by future research. In addition, random districting is another possible application and we could introduce informational structures into our model, such as the verifiability of representatives behavior and the privacy of individual votes.

Appendix

Here, we apply the results of our model to solve or explain two political issues: random districting and the partisan bias.

Application 1. Random districting

In indirect democratic systems, one may hope to prevent bias, and especially the results shown in Corollary 3 or 1, in which some political groups win by intentional gerrymandering. To solve the problem, random districting (i.e., voters are randomly grouped into districts) could be useful. We show that the results on random districting obtained by Gilligan and Matsusaka remain valid for our multiple hierarchical model of a representative democracy.

To compare our model to that of Gilligan and Matsusaka, we first assume that N is odd, all voters support either policy 1 or 2, $x_j \in \{1, 2\}$, and $\lceil N/2 \rceil$ voters have their ideal points at 1 and $\lfloor N/2 \rfloor$ voters have their ideal points at 2.[13] For simplicity, we assume that the number of voters N takes a value such that, for given $K_1, \ldots, K_t$, t-level districting can be carried out in integers.[14] The median voter's ideal point is equal to 1. However, one can group the voters into equal-sized first-level districts such that the ideal point of the first-level median representative is equal to 2, which means that we have more first-level districts with a median representative at 2 than at 1. Now, at the second-level, we can also construct districts with more representatives with ideal points at 2 than at 1. We proceed in the same way until we arrive at the top level, which then has a median representative with an ideal point at 2. Since these types of districts emerge with a positive probability, the expected ideal point will be greater than the voters' median point.

In this example, voters' ideal points are a bit skewed. The next proposition, which is analogous to Proposition 2 of Gilligan and Matsusaka, investigates the cases of symmetric and upwards skewed distributions of voters' ideal points.

Proposition 5. *Let Assumption 1 hold. Assuming that each districting is equally probable, the*

[13]This example is an extension of an example by Gilligan and Matsusaka (2006, p. 387).

[14]Hence, each district has a uniquely determined median voter.

expected bias of random districting

1. *is zero if the voters' ideal points are symmetrically distributed around their median, and*

2. *is biased upwards if the voters' ideal points are skewed upwards.*

Proof. Assume that the voters' ideal points are in ascending order (i.e., $x_1^* < x_2^* < \ldots < x_N^*$). Let $M = (N+1)/2$. Denote by $p(x_i)$ the probability that voter i becomes the top-level representative, that is, the policymaker or legislator.

We start by proving point 1. As a result of the symmetric setting, we must have $p(x_{M-i}) = p(x_{M+i})$ and $x_M - x_{M-i} = x_{M+i} - x_M$, for all $i = 0, \ldots, M-1$. Hence,

$$E\left(x_{LEG}^*\right) = \sum_{i=1}^{N} p(x_i) = \sum_{i=1}^{M-1} p(x_{M-i})x_{M-i} + p(x_M)x_M + \sum_{i=1}^{M-1} p(x_{M+i})x_{M+i} = x_M. \qquad (12)$$

To establish point 2, we just have to replace $x_M - x_{M-i} = x_{M+i} - x_M$ with $x_M - x_{M-i} \leq x_{M+i} - x_M$ for all $i = 0, \ldots, M-1$, which holds since the distribution of ideal points is skewed upwards in equation (12). □

Application 2. The partisan bias

We can apply the result of Lemma 2 to measure the "partisan bias," which is the deviation between the proportion of seats held by a party in the final legislature at the $t+1$-th decision level and that of the votes its members received at the polls. For the measurement, we rearrange voters with each favorite policy to those who are supporting parties of $x_j \in \{0, 1\}$, as in the application in Gilligan and Matsusaka. Here, the population of voters is still N and the number of hierarchical levels is $t+1$.

Let voters who prefer party 0 to party 1 be called partisan. The fraction of voters favoring party 1 is $V = \frac{1}{N}\sum_{i=1}^{N} x_i$ and the fraction favoring party 0 is $1 - V$. The party affiliation of representative k at the $t+1$-th decision level is $x_{t+1,1}^k \in \{0, 1\}$, corresponding to the median representative in the k-th district at the t-th decision level. The fraction of seats held by party 1 and party 0 in the final legislature is $L_t = \frac{1}{K_t}\sum_{k=1}^{K_t} x_{t+1,1}^k$ and $1 - L_t$, respectively. Then, the partisan bias can be defined as $\beta_t = |L_t - V|$. If there is no bias, then the fraction of seats held by party 1 is the same as the fraction of supporters in voters, and $L_t = V$. Since a representative at the $t+1$-th decision level needs at least

$$s \equiv \frac{1}{2}\left(\frac{K_{t-1}}{K_t}+1\right)\frac{1}{2}\left(\frac{K_{t-2}}{K_{t-1}}+1\right)\dots\frac{1}{2}\left(\frac{K_1}{K_2}+1\right)\frac{1}{2}\left(\frac{K_0}{K_1}+1\right)$$

supporters at the voters level, as shown in the proof of Lemma 2, we obtain the next proposition of the partisan bias.

Proposition 6. *Let Assumption 1 hold. Let V be the fraction of voters who support party 1, and suppose there are enough party 1 voters to elect at least one representative of party 1 to the $t+1$-th decision level, but not to elect all representatives:* $\frac{1}{2}(\frac{K_{t-1}}{K_t}+1)\frac{1}{2}(\frac{K_{t-2}}{K_{t-1}}+1)\dots\frac{1}{2}(\frac{K_1}{K_2}+1)\frac{1}{2}(\frac{N}{K_1}+1) < VN < K_t\frac{1}{2}(\frac{K_{t-1}}{K_t}+1)\frac{1}{2}(\frac{K_{t-2}}{K_{t-1}}+1)\dots\frac{1}{2}(\frac{K_1}{K_2}+1)\frac{1}{2}(\frac{N}{K_1}+1)$. *Then*

$$\max\beta_t = \left(\frac{\frac{1}{2K_t}(K_t+1)}{\frac{j^*_{t+1,1}}{N}}-1\right)V.$$

Proof. Since one representative at the $t+1$-th decision level needs at least s voters, when there are $\sum_{i=1}^{K_t} x^i_{t+1,1}$ representatives at the $t+1$-th decision level, the minimum number of supporters at the voters level is $s\sum_{i=1}^{K_t} x^i_{t+1,1}$, which is equal to or less than the actual number of voters supporting party 1: $s\sum_{i=1}^{K_t} x^i_{t+1,1} \le \sum_{i=1}^{N} x_i$. Dividing this by N and multiplying by $\frac{1}{2}(K_t+1)$, we obtain the following inequality:

$$\frac{L_tK_t}{N}\frac{1}{2}(K_t+1)\frac{1}{2}\left(\frac{K_{t-1}}{K_t}+1\right)\frac{1}{2}\left(\frac{K_{t-2}}{K_{t-1}}+1\right)\dots\frac{1}{2}\left(\frac{K_1}{K_2}+1\right)\frac{1}{2}\left(\frac{K_0}{K_1}+1\right) \le \frac{1}{2}(K_t+1)V.$$

Using formula (4) and subtracting V from both sides, we get

$$\beta_t = L_t - V \le \left(\frac{\frac{1}{2K_t}(K_t+1)}{\frac{j^*_{t+1,1}}{N}}-1\right)V.$$

□

The maximum partisan bias β_t is the ratio between the relative position of the median, who is the final representative at the $t+1$-th decision level, and that at the voters level. In fact, this proposition is a generalization of the partisan bias β calculated in Proposition 4 of Gilligan and Matsusaka, and corresponds to the case of $t=1$ in our model. Gilligan and Matsusaka point out that increasing the number of seats K_1 in the legislature decreases the partisan bias, holding the number of voters constant at N. However, one fact needs to be added to their finding. For instance, applying the case of $t = 1$ in our example 1, while both $\{K_1 = 3, K_2 = 1\}$ and $\{K_1 = 9, K_2 = 1\}$ reach the same result (i.e., the 10-th voter becomes the final representative), the maximum partisan biases are $(\frac{2/3}{10/27} - 1)V = 0.8V$ and $(\frac{5/9}{10/27} - 1)V = 0.5V$, respectively. Thus, the legislature with fewer seats has a larger bias than that with more seats, even when the same representative is elected as the final representative in both, with gerrymandering.

Example

Example 3. Let us again consider the example of $N = 27$, $t = 2$, and the voters set $\mathcal{N} = \{1, 2, 3, \ldots, 25, 26, 27\}$. In this case, the median of all voters is 14 and the most extremely liberal voter who is electable as the final representative is $j^*_{3,1} = \frac{1}{2}(3+1) \cdot \frac{1}{2}(\frac{9}{3}+1) \cdot \frac{1}{2}(\frac{27}{9}+1) = 8$. Lemma 4 and Proposition 4 state that voters between 8 and 14, who belong to either the fourth or the fifth district at the first decision level, from the definition of m_i, are electable as the final representative in this example.

Sliding voters: We have to slide voters by zero, one, two, three, and four positions to place voters $8, 9, 10, 11, 12, 13, 14$ at the fourth and fifth district median positions, since $\frac{m_1-1}{2}(\frac{N}{K_1}-1) = \frac{4-1}{2}(\frac{27}{9}-1) = 3$, following Table 8. Then, we have Table 10.

Table 10: Placing voters between 8 and 14 on district median at the first decision level

Sliding	District number								
by	1	2	3	4	5	6	7	8	9
0	$\{1,2,27\}$	$\{3,4,26\}$	$\{5,6,25\}$	$\{7,\boxed{8},24\}$	$\{9,\underline{10},23\}$	$\{11,12,22\}$	$\{13,14,21\}$	$\{15,16,20\}$	$\{17,18,19\}$
				↓	↓				
1	$\{1,2,3\}$	$\{4,5,27\}$	$\{6,7,26\}$	$\{8,\underline{9},25\}$	$\{10,\underline{11},24\}$	$\{12,13,23\}$	$\{14,15,22\}$	$\{16,17,21\}$	$\{18,19,20\}$
				↓	↓				
2	$\{1,2,3\}$	$\{4,5,6\}$	$\{7,8,27\}$	$\{9,\underline{10},26\}$	$\{11,\underline{12},25\}$	$\{13,14,24\}$	$\{15,16,23\}$	$\{17,18,22\}$	$\{19,20,21\}$
				↓	↓				
3	$\{1,2,3\}$	$\{4,5,6\}$	$\{7,8,9\}$	$\{10,\underline{11},27\}$	$\{12,\underline{13},26\}$	$\{14,15,25\}$	$\{16,17,24\}$	$\{18,19,23\}$	$\{20,21,22\}$
				↓	↓				
4	$\{1,2,3\}$	$\{4,5,6\}$	$\{7,8,9\}$	$\{10,\underline{11},12\}$	$\{13,\boxed{14},27\}$	$\{15,16,26\}$	$\{17,18,25\}$	$\{19,20,24\}$	$\{21,22,23\}$

For instance, if we want voter 12 to be elected as a district representative in the second through the final decision levels, we need to slide two positions. Then, we have the representatives $\{2, 5, 8, 10, 12, 14, 16, 18, 20\}$. At the second decision level, we need to slide representatives by one position:

Sliding by	1	2	3
1	$\{2, 5, 8\}$	$\{10, 12, 20\}$	$\{14, 16, 18\}$

Now we have representatives $\{5, 12, 16\}$ in the final decision level. Lastly, 12 is elected as the final representative.

Next, if we want voter 14 to be elected as a district representative in the second through the final decision levels, we need to slide four positions, which is full sliding at the first decision level. Then, we have the representatives $\{2, 5, 8, 11, 14, 16, 18, 20, 22\}$. At the second level, we need to slide representatives by one position, which is also full sliding, for representative 14 to belong to the median district:

Sliding by	1	2	3
1	$\{2, 5, 8\}$	$\{11, 14, 22\}$	$\{16, 18, 20\}$

At the final level, we have $\{5, 14, 18\}$, and voter 14 is elected as the final representative. In this case, voters slide fully at each decision level so that all voters before voter 14 are lined up consecutively in ascending order.

Lastly, if we want voter 11 to be elected as a district representative in the second through the final decision levels, we need to slide voters either one, three, or four positions. If we choose to slide by one position, we have the representatives $\{2, 5, 7, 9, 11, 13, 15, 17, 19\}$, and if we choose three positions, we have $\{2, 5, 8, 11, 13, 15, 17, 19, 21\}$. Sliding by four positions is similar to three positions because 11 is in the fourth district. When sliding by one position, at the second decision level, representatives need to slide by one position. When sliding by three positions, they need to slide by zero positions:

Sliding by first level	Sliding by second level	1	2	3
1	1	$\{2, 5, 7\}$	$\{9, 11, 19\}$	$\{13, 15, 17\}$
3	0	$\{2, 5, 21\}$	$\{8, 11, 19\}$	$\{13, 15, 17\}$.

Since 11 appears as the district median of the fourth and fifth districts when sliding by one, three, and four positions, they achieve the same result at the first decision level. Thus, we can obtain the same result at the final decision level. $\square$

Acknowledgments

This project started when Kobayashi of Hosei University visited the Corvinus University of Budapest as part of an exchange program from February 20 to March 30, 2012. Kobayashi thanks the Corvinus University of Budapest for their financial support and their hospitality. We thank Paolo Balduzzi, Péter Biró, Yoich Hizen, Hiroyuki Ozaki and the participants of seminars and conferences for their helpful comments and useful conversations. Kobayashi gratefully acknowledges the financial support from JSPS KAKENHI (Grant Number 25870793) and Tasnádi gratefully acknowledges the financial support from the Hungarian Scientific Research Fund (OTKA K-112975).

References

[1] Altman, M. and M. McDonald (2010), "The Promise and Perils of Computers in Redistricting," *Duke Journal of Constitutional Law & Public Policy* 5, pp. 69 - 159.

[2] Besley, T. and S. Coate (1997), "An Economic Model of Representative Democracy," *Quarterly Journal of Economics* 112, pp. 85 - 114.

[3] Black, D. (1958), *The Theory of Committees and Elections.* Cambridge: Cambridge University Press.

[4] Chambers, C.P. and A.D. Miller (2010), "A Measure of Bizarreness," *Quarterly Journal of Political Science* 5, pp. 27 - 44.

[5] Chambers, C.P. and A.D. Miller (2013), "Measuring Legislative Boundaries," *Mathematical Social Sciences* 66, pp. 268 - 275.

[6] *Constitution of the People's Republic of China*, the website of the National People's Congress of the People's Republic of China, http://www.npc.gov.cn/englishnpc/Constitution/node_2824.htm, accessed: October 31, 2014.

[7] Coate, S. and B. Knight (2007), "Socially Optimal Districting: A Theoretical and Empirical Exploration," *Quarterly Journal of Economics* 122-4, pp. 1409 - 1471.

[8] Fryer, R.G. Jr. and R. Holden (2011), "Measuring the Compactness of Political Districting Plans," *Journal of Law and Economics* 54, pp. 493 - 535.

[9] Galam, S. and S. Wonczak (2000), "Dictatorship from majority rule voting," *European Physical Journal B* 18, pp. 183 - 186.

[10] Gilligan, T.W. and J.G. Matsusaka (2006), "Public Choice Principles of Redistricting," *Public Choice* 129, pp. 381 - 398.

[11] Institute of Chinese Affairs (1980, 1983, and 2011), China Year Book 1980, 1983, and 2014 (in Japanese), Institute of Chinese Affairs in Japan.

[12] Jefferson, T. (1816), Letter to Joseph C. Cabell. In: Kurland, P.B., Lerner, R. (eds.) *The Founders' Constitution*, Volume 1, Chapter 4, Document 34, http://press-pubs.uchicago.edu/founders/documents/v1ch4s34.html, Chicago: The University of Chicago Press (2000), accessed: March 23, 2012.

[13] Moulin, H. (1980), "On Strategy-proofness and Single Peakedness," *Public Choice* 35, pp. 437 - 455.

[14] Olson, J., (1997), "The Revolutionary Spirit: Hannah Arendt and the Anarchists of the Spanish Civil War," *Polity* 29 - 4, pp. 461- 488.

[15] Puppe, C. and A. Tasnádi (2014), "Axiomatic Districting," *Social Choice and Welfare*, forthcoming.

[16] Shugart, M.S. and J.M. Carey (1992), *Presidents and Assemblies: Constitutional Design and Electoral Dynamics*, Cambridge: Cambridge University Press.