Pork Barrel Politics and Gerrymandering*

Hideo Konishi†    Chen-Yu Pan‡

September 22, 2015

Abstract

In this paper we propose a tractable model of partisan gerrymandering followed by electoral competitions in policy positions and transfer promises in multiple districts. With such pork-barrel considerations, we generally find that gerrymandering results in (i) packing the opponent party’s supporters in losing districts, and (ii) more extreme policies in winning districts. However, depending on the party leaders’ and voters’ preference intensity for policy positions, and depending on how much freedom the leader has in redistricting, we obtain a variety of optimal gerrymandering policies. The well-known pack-and-crack gerrymandering is not necessarily optimal: the party leader may choose to create some extremely polarized districts to avoid making costly pork-barrel promises, even with candidates who have more polarized positions than hers.

Keywords: electoral competition, gerrymandering, policy convergence/divergence, pork-barrel politics

JEL Classification Numbers: C72, D72

*Still preliminary. Not for circulations.
†Hideo Konishi: Department of Economics, Boston College, USA. Email: hideo.konishi@bc.edu
‡Chen-Yu Pan: School of Economics and Management, Wuhan University, PRC. E-mail panwhu@126.com
1 Introduction

It is widely agreed that the US Congress has polarized quite a bit in the last half century. The distribution of the House representatives’ political positions were more concentrated at the center of political spectrum with considerable overlaps between Republican and Democratic representatives’ positions in 1960s, while it became sharply twin-peaked without overlaps in 2000s (see Figure 1: by Fiorina, Abrams, and Pope, 2011). Simultaneously, Fiorina, Abrams, and Pope (2011) argue that the US voters have not polarized so much during the same time period. These conflicting observations generate an obvious puzzle: How could the Congress polarize if voters didn’t?

One possible explanation is that voters sorted out into Republican and Democratic parties by their political positions during the period, and that the parties’ political positions were polarized in party members’ preference aggregation. Levendusky (2009) suggests that party elites’ polarization led voter sorting, although it is controversial how much mass polarization actually occurred by voter sorting. Recently, Krasa and Polborn (2015) propose a two-party electoral competition model with multiple districts, voters care not only about the political positions of their local candidates but also about the positions of the parties (which are determined by the elected party representatives in the legislatures), and show that elected representatives’ positions are pulled towards the position of the party’s median voters of the districts. Their model can explain the House representatives’ polarization at least partially, as long as the party median voters are pulled away from the center by voter sorting.

However, more direct explanation for that is partisan gerrymandering in the US politics. Since the one-person, one-vote decisions by the US Supreme Court in 1960s, the courts closely examine population equality in districts. In practice, however, redistricting is not simply about equaling the population in each district. It is clear that there is a strong incentive for the party in charge of redistricting to change the political identity of each district to favor

---

1 It is now standard to use a one-dimensional scaling score (DW-Nominate procedure on economic liberal-conservative, Poole and Rosenthal, 1997) to measure representatives’ political positions.

2 Levendusky (2009) asserts that party elites’ polarization clearly caused limited scale of mass polarization. Fiorina et al. (2011) say that there is little increase in mass polarization, although they admit that party activists’ political positions have polarized.
the partisan interests.\textsuperscript{3} Although Ferejohn (1977) finds little support for gerrymandering being the cause of declines in competitiveness of congressional districts from mid 1960s to 1980s, Fiorina et al. (2011) state “Many (not all) observers believe that the redistricting that occurred in 2001-2002 had a good bit to do with this more recent decline in competitive seats—the party behaved conservatively, concentrating on protecting their seats rather than attempting to capture those of the opposition.” (see Fiorina et al. pp. 214-215). They further argue that this decrease in competitiveness is a driving force behind the recent political polarization in Congress (see also Gilroux, 2001).\textsuperscript{4}

In partisan gerrymandering, two tactics are often discussed. The first is to concentrate or “pack” those who are supporting opponent party in losing districts. The second is to evenly distribute or “crack” supporters in winning districts. Introducing uncertainty in each district’s median voter’s position, Owen and Grofman (1988) consider the situation where a partisan gerrymanderer redesigns districts in order to maximize the expected number of seats or the probability to win a working majority of seats for her party. They assume that the uncertainty in median voter’s political position is local and is independent across districts when the objective is expected number of seats, and that the uncertainty is global when the goal is the probability of winning the majority. Also, they assume that the sum of the positions of district median voters must stay the same after redistricting (their feasibility constraint). They show that the optimal strategy is “packing” the opponents in losing districts, and “cracking” the rest of voters evenly across the winning districts with substantial margins so that the party can win districts even in the cases of negative shocks under both cases.\textsuperscript{5} Gul and Pesendorfer (2010) extend Owen and Grofman (1988) by introducing a continuum of districts, and voters’ party affiliations, and further allow for bipartisan gerrymander-

\textsuperscript{3}One recent example is the 2003 Republican redistricting in Texas that has been considered contributing to Democrat’s defeat in the subsequent election.

\textsuperscript{4}McCarty et al. (2006, 2009) document that the political polarization of the House of Representatives has increased in recent decades, using data on roll call votes. Although they find only minor relation between polarization and gerrymandering, Krasa and Polborn (2015) argue that their answer may be incomplete if the political positions of district candidates are mutually interdependent.

\textsuperscript{5}The original “cracking” tactics create the maximum number of winning districts with the smallest margins. In the traditional literature, some argue that gerrymandering will increase political competition by this reason. In this paper, we use “cracking” tactics in the sense of Owen and Grofman (1988).
ing. They assume that each party leader maximizes the probability of winning the majority of seats under more sophisticated feasibility constraints by assuming a continuum of districts and generate a “pack-and-crack.”

Friedman and Holden (2008), on the other hand, assume that a partisan gerrymanderer has full freedom in allocating population over a finite number of districts, and that she maximize the expected number of seats when there are only valence uncertainty in median voters’ utilities (thus, there is no uncertainty in median voter’s political position). In this idealized situation, they find that the optimal strategy is “slice-and-mix”: mixing a slightly larger number of the most extreme supporters and a slightly smaller number of the most extreme opponents in the first district, then mixing a slightly larger number of the second most extreme supporters and a slightly smaller number of opponents in the second district, and so on and so forth. Although this is a very different prescription for a partisan gerrymanderer, but it is a nice benchmark to see what can be done under ideal situation.

However, these papers do not model spatial competition in policy positions, and the elected representatives’ positions are implicitly assumed to be the district median voters’ positions (Downsian competition). In contrast, we assume that policy-motivated party leaders compete with their political positions and pork barrel promises, and we find that policy divergence actually occurs even though our setup is deterministic. We analyze the optimal gerrymandering policies building on this political competition model. With pork-barrel politics, the party leader is aware of the cost of pork-barrel policies, and therefore she has strong partisan gerrymandering incentives to collect their supporters in the winning districts in order to avoid large pork-barrel promises. Our simple model is flexible enough to analyze the optimal policies under various feasibility constraints on gerrymandering. It turns out that we can generate “pack-and-crack” in Owen and Grofman (1988), “slice-and-mix” in Friedman and Holden (2008), and even “pack-and-slice” (see the

---

6They consider two feasibility constraints. The first is the constant mean of median voters’ positions which is the same as the one in Owen and Grofman (1988). The second one is that the status quo needs to be a mean-preserved spread of a feasible redistricting plan.

7In two-party electoral competition model a la Wittman (1983), the equilibrium outcome is the Downsian unless there is uncertainty in median voter’s position (see Roemer 2001). In our model, if we place the upperbound on the level of pork-barrel promise, then the political competition equilibrium converges to the Downsian as the upperbound goes to zero.
following paragraph) as the optimal gerrymandering policies depending on feasibility constraints, although our model is a deterministic one.

The rest of the paper is organized as follows. Section 2 discusses some related literature. In Section 3, we start with analyzing political-position and pork-barrel competition and characterizing the party leader’s payoff from each winning district by the district median voter’s position (Proposition 1). In Section 4, we investigate the optimal gerrymandering strategy when the party leader has complete freedom as in Freedman and Holden (2008), and show that their “slice-and-mix” is also an optimal strategy in our model (Proposition 2). In Section 5, we proceed to the cases where gerrymanderer’s freedom is limited by indivisibility of localities. For tractability of our analysis, we assume voters’ and leaders’ cost function in political distance have common constant elasticity $\gamma > 1$. We also assume that each district has normally distributed voters to justify the feasibility constraint imposed by Owen and Grofman (1988). We show that the optimal gerrymandering policy may be “pack-and-crack” (Owen and Grofman, 1988), or “pack-and-slice,” in which the gerrymanderer pack the opponent supporters, and slice her supporters from the strongest to moderate in order, depending on the value of elasticity $\gamma$ of cost functions (Propositions 3 and 4). We also investigate how the results could be affected by an introduction of uncertainty in median voters’ positions. Unlike in Owen and Grofman (1988), we find that “pack-and-slice” becomes more likely than “pack-and-crack” since in our model, party leaders care about policy positions as well as expected number of seats (Proposition 5). Dropping normality of voter distributions, the optimal gerrymandering strategy can be more complicated. Even when indivisible localities can be ordered in first-order stochastic dominance, we can only show “packing” is optimal, but we cannot generally say much about how to allocate supporters over districts as well as the optimal number of winning districts. Section 6 concludes. All proofs are collected in Appendix A.

2 Related Literature

Our paper is related to two branches of literature. The first one is the pork-barrel literature. In this branch, our model is most related to Lindbeck and Weibull (1987) and Dixit and Londregan (1996). The former introduces a two-party competition model in which (extreme) parties use pork-barrel
policies to attract agents with heterogeneous policy preferences. The latter generalizes Lindbeck and Weibull (1987) to allow parties having different abilities in practicing pork-barrel policies, and this difference determines the pork-barrel policy’s target being swing voters or loyal supporters. Our model is different from theirs in that we introduce parties’ platform decisions besides pork-barrel politics, and party leaders choose these two policies simultaneously. Since we try to explain the House representatives’ polarization, having platform decisions is essential in our analysis. Moreover, the political competition result is deterministic in our model which is different from the setup with uncertainty in the literature.

Second, other than the optimal gerrymandering literature we discussed in the previous section, there is gerrymandering literature from normative point of view as well. This literature focuses on the how the gerrymandering affects the relation between seats and the vote shares won by a party, so-called “seat-vote curve”. Coate and Knight (2007) identify the social welfare optimal seat-vote curve and then the conditions under which the optimal curve can be implemented by a districting plan. With fixed and extreme parties’ policy positions, they find that the optimal seat-vote is biased toward the party with larger partisan population. However, Bracco (2013) shows that, when parties strategically choose their policy position, the direction of seat-vote curve bias should be the opposite. Besley and Preston (2007) construct a model similar to Coate and Knight (2007) and show the relation between the bias of seat-vote curve and parties policy choices. They further empirically test the theory and the result show that reducing the electoral bias can make parties strategy more moderate.

3 The Model

We consider a two-party ($L$ and $R$) multi-district model with one party being entitled to redistrict a state. There are many localities in the state, each of which is considered the minimal unit in redistricting (a locality cannot be divided into smaller groups in redistricting). We assume that there are $L$ dis-

---

8Dixit and Londregan (1998) propose a pork-barrel model with strategic ideological policy decision based on their previous work. However, the ideology policy in their paper is the equality-efficiency concern engendered by parties’ pork-barrel strategies. Therefore, the ideology decision in their work is a consequence of pork-barrel politics instead of an independent policy dimension.
crete localities each of which has population 1. The state must have $K$ equally populated districts, and $\mathcal{L}$ is a multiple of $K$. That is, the party in power needs to create $K$ districts by combining $\frac{L}{K} = n$ localities in each district. Locality $\ell = 1, \ldots, \mathcal{L}$ has a voter distribution function $F_\ell : (-\infty, \infty) \to [0, 1]$, where $(-\infty, \infty)$ is a space of ideologies in the society and $F_\ell(\theta)$ is non-decreasing with $F_\ell(-\infty) = 0$ and $F_\ell(\infty) = 1$. Ideology $\theta < 0$ is regarded left, and $\theta > 0$ is right. With a slight abuse of notations, we denote the set of localities also by $\mathcal{L} \equiv \{1, \ldots, \mathcal{L}\}$. A redistricting plan $\pi = \{D^1, \ldots, D^K\}$ with $|D^k| = n$ for all $k = 1, \ldots, K$, is a partition of $\mathcal{L}$. The gerrymandering party’s leader chooses the optimal district partition $\pi$ from the set of all possible partitions $\Pi$.

In each district $k$, the probabilistic distribution function $F^k$ is an average of distribution functions of $n$ localities: $F^k(\theta) = \frac{1}{n} \sum_{\ell \in D^k} F_\ell(\theta)$. District $k$’s median voter is denoted by $x^k \in (-\infty, \infty)$ with $F^k(x^k) = \frac{1}{2}$. We also assume the uniqueness of $x^k$ in each districting plan.

We model pork-barrel elections in a similar manner with Dixit and Londregan (1996). A type $\theta$ voter in district $k$ evaluates party $L$ or $R$ according to the utility function with two arguments: one is the policy position of the candidate representing the corresponding party, $\beta \in \mathbb{R}$, and the other is the party’s pork-barrel transfer $t \in \mathbb{R}_+$. We interpret this pork-barrel transfer as a promise of local public good provision (measured by the amount of monetary spending) in the case that the party’s candidate is elected. Formally, a voter $\theta$ in district $k$’s utility when party $L$ is the winner is

$$U_\theta(L) = t^k_L - c(|\theta - \beta^k_L|)$$

where $\beta^k_L$ is party $L$’s district $k$ candidate’s policy position, $t^k_L \geq 0$ is party $L$’s district-$k$-specific pork-barrel transfer, and $c(d) \geq 0$ is the ideology cost function which is increasing in the distance between a candidate’s position and her own position. We assume that $c(\cdot)$ is continuously differentiable, and satisfies $c(0) = 0, c'(0) = 0$, and $c'(d) > 0$ and $c''(d) > 0$ for all $d > 0$ (strictly increasing and strictly convex). Similarly, we have

$$U_\theta(R) = t^k_R - c(|\theta - \beta^k_R|)$$

ootnote{A partition $\pi$ of $\mathcal{L}$ is a collection of subsets of $\mathcal{L}$, $\{D^1, \ldots, D^K\}$, such that $\bigcup_{k=1}^K D^k = \mathcal{L}$ and $D^k \cap D^{k'} = \emptyset$ for any distinct pair $k$ and $k'$.}

ootnote{In reality, there are many restrictions on what can be done in a redistricting plan. For example, a district is required to be connected geographically. Despite the complication involved, our analysis can still be extended to the case with geographic restrictions by introducing the set of admissible partitions $\Pi^A \subseteq \Pi$ (see Puppe and Tasnadi, 2009).}
where $\beta_R^k$ is party $R$’s district $k$ candidate’s policy position and $t_R^k$ is party $R$’s district-$k$-specific transfer.

Therefore, voter $\theta$ votes for party $L$ if and only if

$$U_{\theta}(L) - U_{\theta}(R) = [c(|\theta - \beta_R^k|) - c(|\theta - \beta_L^k|)] + t_L^k - t_R^k > 0$$  \hspace{1cm} (2)$$

Since the median voter’s type in district $k$ is $x^k$, given $\beta_L^k$, $\beta_R^k$, $t_L^k$ and $t_R^k$, $L$ wins in district $k$ if and only if

$$U_{x^k}(L) - U_{x^k}(R) = [c(|x^k - \beta_R^k|) - c(|x^k - \beta_L^k|)] + t_L^k - t_R^k > 0$$  \hspace{1cm} (3)$$

Each party leader in the state (of these $K$ districts) cares about (i) the influence within her party, (ii) the candidate’s policy position in each district, and (iii) the district-specific pork-barrel spending. We assume that the party leader prefers to win a district with a candidate’s position closer to her own ideal ideological position and a less pork-barrel promise. The former is regarded as the “policy-motivation” in the literature. By formulating the latter, we consider a situation where the leader bears some costs when implementing the promised local public goods provision. For example, the bargaining efforts needed to push for federal funding. To simplify the analysis, we assume that the negative utility by pork-barrel is measured by the amount of money promised. We denote the ideal political positions of the leaders of party $L$ and $R$ by $\theta_L$ and $\theta_R$, respectively, with $\theta_L < \theta_R$. Without loss of generality, we will set $\theta_L = -1$ and $\theta_R = 1$ in the end, but we will stick to notations $\theta_L$ and $\theta_R$ until the gerrymandering analysis starts, which would be more helpful to the readers. Formally, by winning in district $k$, party $L$’s leader gets utility

$$V^k_L = Q - t_L^k - C(|\beta_L^k - \theta_L|),$$

where $Q > 0$ is the fixed payoff (e.g., getting recognition among her party elites) for winning another district, and $C(d)$ is a party leader’s ideology cost function with $C(0) = 0$, $C'(0) = 0$, $C''(d) > 0$ and $C'''(d) > 0$ (strictly increasing and strictly convex). If she loses in district $k$, she gets zero utility from the district. Thus, party $L$’s state leader’s utility from $K$ districts in the state is

$$V_L = \sum_{k=1}^K I_L(k)V^k_L = \sum_{k=1}^K I_L(k)[Q - t_L^k - C(|\beta_L^k - \theta_L|)],$$  \hspace{1cm} (4)$$
where $I_L(k) = 1$ if $L$ wins, $I_L(k) = 0$ if $L$ loses, and $I_L(k) = \frac{1}{2}$ if $L$ ties with $R$ in district $k$. Similarly, the utility for $R$ is defined as

$$V_R = \sum_{k=1}^{K} I_R(k) V_R^k = \sum_{k=1}^{K} I_R(k)[Q - t_R^k - C(|\beta_R^k - \theta_R|)].$$

(5)

We introduce a *tie-breaking rule* in each district based on the relative levels of the state party leaders’ utilities $V_L^k$ and $V_R^k$. We assume that if two parties’ offers are tied for the median voter $x^k$ ($U_{x^k}(L) = U_{x^k}(R)$) while one party’s leader gets a higher (indirect) utility than the other’s, the median voter will vote for the party. That is,

**Assumption 1. (Tie-Breaking)** Given two parties’ offers are such that $U_{x^k}(L) = U_{x^k}(R)$, $L$ wins if $V_L^k > V_R^k$ ($V_L^k < V_R^k$).

This assumption is justified by the fact that the higher utility is equivalent to the higher ability to provide a better offer to the median voter. Especially, consider the case in which two parties are tied and, say, $V_L^k > V_R^k = 0$, party $L$ has the ability to provide $\epsilon > 0$ more pork-barrel promise. Therefore, we break the tie by assuming the median voter prefers $L$, which is the standard assumption.

Our second assumption is a simple sufficient condition that assures interior solutions for both parties.

**Assumption 2. (Relatively Strong Office Motivation)** For all feasible $x^k$, $Q \geq \min_{\beta} \{C(|\theta_j - \beta|) + c(|\beta - x^k|)\}$ holds for $j = L, R$.

Notice that if the party leader gets 0 utility, he must offer pork-barrel promise equal to $Q - C(|\theta_j - \beta|)$. Therefore, the median voter get utility $U_{x^k} = Q - C(|\theta_j - \beta|) - c(|\beta - x^k|)$. Then, this assumption means that the payoff from winning a district, $Q$, is large enough so that for any $x^k$, both party can offer the median voter positive utility, which is a sufficient condition for the candidate selection problem has interior solution. Also notice that since the model only allows a finite median voters positions, there must exist a $Q$ to satisfy this assumption. Moreover, the implication of this assumption is that it guarantees that in equilibrium both parties promise positive pork-barrel. We will see this more clearly in the next section.\(^\text{11}\)

---

\(^{11}\)This assumption can be weakened significantly. In the Appendix B, we present a weaker condition (Assumption 2') for interior solutions, and discuss the case of corner solutions. However, our argument still extend to the cases with corner solutions to some extent.
The timing of the game is as follows:\footnote{We can separate stage 2 into two: policy position choices followed by pork-barrel promises. If we do so, the loser of a district $k$ will get zero payoff in every subgame, so it becomes indifferent among policy positions. Thus, we need equilibrium refinement to predict the same allocation. By assuming that the loser party chooses the policy position that minimizes the opponent party leader’s payoff, we can obtain the exactly the same allocation in SPNE.}

1. One party, say $L$, chooses a redistricting plan $\pi = \{D^1, \ldots, D^K\}$ of $L$.

2. Given the districting plan in stage 1, party leaders $L$ and $R$ simultaneously choose local policy positions and pork-barrel promises $(\beta^k_L, t^k_L)_{k=1}^K$ and $(\beta^k_R, t^k_R)_{k=1}^K$, respectively.

3. All voters vote sincerely (with our tie-breaking rule). The winning party is committed to its policy position and its pork-barrel promise in each district $k = 1, \ldots, K$. All payoffs are realized.

We will employ weakly undominated subgame perfect Nash equilibrium as the solution concept. We require that in stage 2, party leaders play weakly undominated strategies so that the losing party leader does not make cheap promises to the district median voters.\footnote{The losing party does not suffer from doing this, since she gets zero utility in losing districts anyway. The winning party needs to match the offer as long as she can get a positive payoff by doing so.Demanding players to play weakly undominated strategies, we can eliminate these unreasonable equilibria.} We will call a weakly undominated subgame perfect Nash equilibrium simply an equilibrium.

### 3.1 Electoral Competition with Pork-Barrel Politics

We solve the equilibria of the game by backward induction. We start with stage 2, knowing that voters vote sincerely in stage 3. Notice that the key player is the median voter in the voting stage. Thus, when the leader of party $L$ makes her policy decisions in district $k$, she at least needs to match $R$’s offer in terms of median voter’s utility in order to win. First, we consider the case that party $L$ wins (the party $R$’s leader wins only by providing a strictly better offer to the median voter). In this case, the leader of party $L$ tries to offer the same utility to the median voter $x^k$ and wins in the district $k$ with
the tie-breaking rule. Formally, the party leader’s problem is described by

\[ \max_{t^k_L, t^k_R} \{ Q - t^k_L - C(\theta_L - \beta^k_L) \} \]

subject to \( t^k_L - c(|x^k - \beta^k_L|) \geq \bar{U}^k_R, t^k_L \geq 0 \), and \( Q - t^k_L - C(\theta_L - \beta^k_L) \geq 0 \), \( \beta^k_L \)

where \( \bar{U}^k_R \) is the median voter’s utility level from \( R \)’s offer. Notice that \( t^k_L \geq 0 \) and \( Q - t^k_L - C(\theta_L - \beta^k_L) \geq 0 \) may or may not be binding while \( t^k_L - c(|x^k - \beta^k_L|) \geq \bar{U}^k_R \) must be binding. The solution for this maximization problem is straightforward. Define \( \hat{\beta}(x^k, \theta) \) by the following equation

\[ c'(\beta^k_L) = C'\left(\theta - \hat{\beta}(x^k, \theta)\right). \]

Notice that (7) is simply the first order condition of optimization problem (6) after substituting \( t^k_L = c(|x^k - \beta^k_L|) + U^k_R \) into the objective function. The optimal policy \( \beta^k_L = \hat{\beta}(x^k, \theta_L) \) when \( -c(|x^k - \hat{\beta}(x^k, \theta_L)|) \leq \bar{U}^k_R \). That is, it is not enough for the winning party to win just by only using the policy platform. Therefore, in this case, it is clear that the optimal pork barrel promise is

\[ t^k_L = \bar{U}^k_R + c(|x^k - \hat{\beta}(x^k, \theta_L)|) \]

Although it seems not clear that \( t^k_L \) defined above is positive or not, it turns out that \( t^k_L \) is always positive, since the similar optimization problem applies for the losing party and Assumption 2.

Now, it is clear that the winning party’s decision is related to what the losing part proposes in equilibrium. The following lemma shows that the losing party cannot lose with positive surplus.

**Lemma 1.** Suppose \( R \) is the losing party in district \( k \). In equilibrium, \( R \) proposes the policy pair \( (\beta^k_R, t^k_R) \) which is the solution of the following problem

\[ \max_{\beta^k_R, t^k_R} U_{x^k}(R) = t^k_R - c(|x^k - \beta^k_R|) \]

subject to \( t^k_R \geq 0 \) and \( Q - t^k_R - C(\theta_R - \beta^k_R) \geq 0 \)

That is, the losing party leader offers a policy position and pork-barrel promise that leave herself zero surplus in equilibrium.

\[ \beta^k_R = \hat{\beta}(x^k, \theta_R) \]

\[ t^k_R = Q - C(\theta_R - \hat{\beta}(x^k, \theta_R)) \]

\[ 11 \]
Moreover, this policy pair is the best she can offer for the median voter $x^k$.

The intuition of this lemma is straightforward. If the losing party does not offer the median voter the best one, then since the winning party will provide the median voter the same utility level, the losing can always offer the median voter something better than her original offer and win the district. This cannot happen in equilibrium. Therefore, for the losing party $R$, the equilibrium strategy is $\beta^*_R = \hat{\beta}(x^k, \theta_R)$ and $t^*_R = Q - C(|\theta_R - \beta^*_R|)$. The policy pair provides the median voter with the utility $U^*_R = Q - C(|\theta_R - \beta^*_R|) - c(|x^k - \beta^*_R|)$. Using this $U^*_R$, one can solve the winning party’s equilibrium pork-barrel promise $t^*_L = Q - C(|\theta_L - \beta^*_L|) - c(|x^k - \beta^*_L|) + c(|x^k - \beta^*_L|)$.

One thing left to decide is which party should be the winning party. Notice that, by Lemma 1, the losing party always propose the best offer by depleting all her surplus. Therefore, the party can potentially provide the median voter with a higher utility level is the winner. Notice that $j$ party’s pork-barrel promise is bounded above by $j$ party leader’s payoff evaluated at $\beta^*_j$ (otherwise, the leader gets a negative utility):

$$Q - C(|\theta_R - \beta^*_R|).$$

Substituting this into median voter’s utility, we obtain

$$W^*_R = Q - C(|\theta_R - \beta^*_R|) - c(|x^k - \beta^*_R|),$$

similarly, for party $L$,

$$W^*_L = Q - C(|\theta_L - \beta^*_L|) - c(|x^k - \beta^*_L|),$$

where $W^*_R$ and $W^*_L$ are the (potential) maximum utilities that the median voter gets from the corresponding party’s offer. Therefore, party $L$ wins in the third stage if and only if

$$c(|x^k - \beta^*_R|) + C(|\theta_R - \beta^*_R|) > c(|x^k - \beta^*_L|) + C(|\theta_L - \beta^*_L|),$$

which simply means that $L$ wins if and only if

$$|\theta_L - x^k| < |\theta_R - x^k|.$$

Proposition 1. Suppose that Assumptions 1 and 2 are satisfied. Define $\hat{\beta}(x^k, \theta)$ by (7). We have
1. For the losing party \( j \), the optimal choice \( \beta_j^k(x^k, \theta_j) = \hat{\beta}(x^k, \theta_j) \) which lies in the interval \((x^k, \theta_j)\) (or \((\theta_j, x^k)\)).

2. For the winning party \( i \), \( \beta_i^k = \hat{\beta}(x^k, \theta_i) \), and \( t_i^k = Q - C(|\theta_j - \beta_j^k|) - c(|x^k - \beta_j^k|) + c(|x^k - \beta_i^k|) \).

3. Party \( i \) wins in the \( k \)th district if and only if \( |\theta_i - x^k| < |\theta_j - x^k| \).
   Party’s winning probability \( I_i(x^k) \) is described by
   \[
   I_i(x^k) = \begin{cases} 
   1 & \text{if } |\theta_i - x^k| < |\theta_j - x^k| \\
   \frac{1}{2} & \text{if } |\theta_i - x^k| = |\theta_j - x^k| \\
   0 & \text{if } |\theta_i - x^k| > |\theta_j - x^k| 
   \end{cases}
   \]

4. The winning payoff for the party leader \( i \) from district \( k \) is
   \[
   \tilde{V}_i^k(x^k, \theta_i, \theta_j) = C(\theta_j, x^k) - C(\theta_i, x^k),
   \]
   where \( C(\theta_i, x^k) \equiv C(|\theta_i - \hat{\beta}(x^k, \theta_i)|) + c(|x^k - \hat{\beta}(x^k, \theta_j)|) \).

5. The winning payoff for party \( i \), \( \tilde{V}_i^k \), increases as \( x^k \) moves away from \( \theta_j \).

Note that party leaders’ policy position choices do not strategically depend on the position decisions of the other party. This property of our model simplifies the analysis. Since there is no strategic issue, the winner of the district is determined by which party leader’s position is closer to the median voter’s position. Moreover, party leader’s winning payoff from the district is simply the difference in the two parties’ sums of ideology cost functions of the median voter and the party leader.

3.2 Tractable Cost Function

We introduce a convenient special ideology cost function such that both voters’ and party leaders’ cost functions have common constant elasticity. Let \( C(d) = a^C d^\gamma \) and \( c(d) = a^c d^\gamma \), where \( \gamma > 1 \), \( a^C > 0 \), and \( a^c > 0 \) are parameters. In this case both party leaders and voters have the same elasticity that is constant \( \gamma \). In this case, we have the following convenient formula.
Claim. Denote $A = A(a^C, a^c) = a^C \left( \frac{a}{1+a} \right)^{\gamma} + a^C \left( \frac{1}{1+a} \right)^{\gamma} > 0$ where $\alpha = \left( \frac{a^c}{a^c} \right)^{\frac{1}{\gamma-1}}$. Suppose that $(\frac{Q}{A})^\frac{1}{\gamma} \geq \max \{|x^k - \theta_L|, |x^k - \theta_R|\}$ holds to assure Assumption 2. Normalizing $\theta_L = -1$ and $\theta_R = 1$, the winning party leader’s utility from district $k$, $\bar{V}_L^k(x^k, \theta_i, \theta_j)$, can be described by

$$
\bar{V}(|x^k|) = \begin{cases} 
A \left[ \left( |x^k| + 1 \right)^\gamma - \left( 1 - |x^k| \right)^\gamma \right] & \text{if } |x^k| \leq 1 \\
A \left[ \left( |x^k| + 1 \right)^\gamma - \left( |x^k| - 1 \right)^\gamma \right] & \text{if } |x^k| \geq 1 
\end{cases}
$$

where (i) $\bar{V}'' > 0$ and (ii) $\bar{V}'' \leq 0 \iff \gamma \geq 2$. When $\gamma = 2$, the formula becomes

$$
\bar{V}(|x^k|) = 4A |x^k|
$$

Since $A$ is a constant, a parameter $\gamma$ is the only relevant one to decide the curvature of payoff function. This convenient formula provides a clear insight in the gerrymandering stage. Also, one can see the importance of political competition on pork-barrel dimension. Suppose pork-barrel promise is forbidden in stage 2, one can see that the only equilibrium is Downsian equilibrium. Therefore, the winning payoff is $Q - C(|\theta_i - x^k|)$ which is always a concave function in $x^k$. However, by considering the pork-barrel promise, the party has extra spending saving benefit when $x^k$ move closer to extreme which makes our payoff function different from the previous one in the literature.

### 3.3 Gerrymandering Problem

Now, we can formalize the gerrymandering party leader’s optimization problem. Proposition 1 shows that $x^k = x^k(\pi)$ is the sufficient statistics to determine the outcome of the $k$th district. Without loss of generality, assume that party $L$ is in charge of redistricting. Notice that the indirect utility of $L$, $\bar{V}_L^k(x^k, \theta_L, \theta_R)$, is relevant only when party $L$ wins in district $k$. The choice of $\pi = (D^1, ..., D^K)$ affects the party leader $L$’s payoff through $(x^1(D^1), ..., x^K(D^K))$ via winning chance $I(x^k(D^k))$ in each district $k$ and its indirect utility $\bar{V}_L^k(x^k(D^k), \theta_L, \theta_R)$: the party $L$ leader’s utility function is written as

$$
\bar{V}_L(\pi, \theta_L, \theta_R) \equiv \sum_{k=1}^{K} I_L(x^k(D^k))\bar{V}_L^k(x^k(D^k), \theta_L, \theta_R)
$$

From now on, we suppress $\theta_L$ and $\theta_R$ in indirect utility $\bar{V}_L^k$ and $\bar{V}_L$. Although $x^k$ is solely determined by $D^k$, we can write $x^k = x^k(D^k(\pi)) = x^k(\pi)$ for all
\( k = 1, \ldots, K \), since \( \pi \) uniquely determines \((D^1, \ldots, D^K)\). Thus, we can rewrite the party leader \( L \)'s gerrymandering choice to be the result of the following maximization problem

\[
\pi^* \in \arg \max_{\pi \in \Pi} \tilde{V}_L(\pi)
\]

The SPNE of this game is \((\pi^*, (\beta^k_L)^K_{k=1}, (\beta^k_R)^K_{k=1}, (t^k_L)^K_{k=1}, (t^k_R)^K_{k=1})\).\(^{14}\)

## 4 Gerrymandering with Complete Freedom

As a limit case, let us consider the ideal situation for the gerrymanderer (Friedman and Holden, 2008): there is a large number of infinitesimal localities with politically homogeneous population: for all position \( x \in (-\infty, \infty) \), there are localities \( \ell \)s with \( F_\ell(x - \delta) = 0 \) and \( F_\ell(x + \delta) = 1 \) for a small \( \delta > 0 \).

That is, the gerrymanderer can create any kind of population distributions for \( K \) districts freely as long as they sum up to the total population distribution. We ask what strategy the gerrymanderer should take. Especially, we want to ask that whether the gerrymanderer wants to have homogeneous or heterogeneous winning districts in terms of district median locations. First, if the gerrymander loses some districts, she does not care how she loses those districts. Therefore, there are multiple optimal redistricting plans. In order to pin down the solution, we assume that party \( L \) leader tries to make the \( R \) party candidates’ positions to the left as much as possible whenever she cannot win in some districts. Assume that the optimal number of winning districts is \( K^* \leq K \) without loss of generality. The question is now how to choose \( x^1, \ldots, x^K \) to maximize by allocating population to \( K \) districts.

In order to highlight the similarity of the result with Friedman and Holden (2008), \emph{we will first assume} \( F(-1) = 0 \) and \( F(1) = 1 \) (the support of voter distribution is limited to \([-1, 1]\)), and \( \theta_L = -1 \) and \( \theta_R = 1 \) so that the party leaders’ political positions the most extreme of all.\(^{15}\)

First imagine the situation that the optimal number of winning districts satisfies \( K^* = 1 \). By Proposition 1.5, \( x^1 \) should be as extreme as possible, and then \( x^1 \) should satisfy \( F(x^1) = \frac{1}{2K} \) (\( x^1 \) is the median voter of the district: the most extreme district achievable with population \( \frac{1}{K} \)). However,

---

\(^{14}\)The existence of SPNE is guaranteed by the convexity of function \( c \) and \( C \) and a finite \( \Pi \).

\(^{15}\)If we drop the assumption that the party leaders have the most extreme preferences by Friedman and Holden (2008), the result is modified to involve cardinality of cost functions. See Appendix C.
the remaining population to the right of $x^1$ can be anything. For example, if the $L$'s party leader cares about minimizing $R$ party leader's payoff, then she will create a district by combining sets $\{\theta \leq \theta^1_L : F(\theta^1_L) = \frac{1}{2K} + \frac{\epsilon}{K}\}$ and $\{\theta \geq \theta^1_R : 1 - F(\theta^1_R) = \frac{1}{2K} - \frac{\epsilon}{K}\}$ where $\epsilon > 0$ is arbitrarily small (what is needed is to make sure that $x^1$ becomes the median voter of the first district). If these two sets compose a district together, then $x^1$ will be defined by $F(x^1) = \frac{1}{2K}$. This is a slice-and-mix strategy for one district to waste $R$'s strong supporters' votes. When there are multiple winning districts, the same logic is applicable. Let $\theta^k_L \in [-1, 1]$ be such that $F(\theta^k_L) = \frac{k}{2K} + \frac{k^*}{K}$ for all $k = 1, ..., K^*$, and let $\theta^k_R \in [-1, 1]$ be such that $1 - F(\theta^k_R) = \frac{k}{2K} - \frac{k^*}{K}$. For $\epsilon > 0$ small enough, we have

$$-1 < \theta^1_L < \theta^{K^*}_L < \theta^{K^*}_R < ... < \theta^1_R < 1.$$ 

If $K^* > 1$, the first district is composed of all voters with $\theta \in [-1, \theta^1_L] \cup [\theta^1_R, 1]$, district $k = 2, ..., K^*$ is composed of all voters with $\theta \in (\theta^1_L, \theta^1_R) \cup (\theta^{K^*}_L, \theta^{K^*}_R)$. That is, let $\mu^k$ be the population measure at district $k$; then $\mu^k((\theta^1_L, \theta^1_R)) = F(\theta^1_L) - F(\theta^k_L - 1) = \frac{1}{2K} = F(\theta^k_R) - F(\theta^1_R) = \mu^k((\theta^{K^*}_R, \theta^{K^*}_L))$. All voters with $\theta \in [\theta^1_L, \theta^{K^*}_L)$ are packed in losing districts $k = K^* + 1, ..., K$. In this case, party leader $L$ can pick median voters $x^k \simeq \theta^k_L$ approximately for all $k = 1, ..., K^*$. We call this redistricting plan a slice-and-mix strategy, which is proposed in Friedman and Holden (2008). Under the slice-and-mix strategy, the resulting policy allocations in winning districts is $(x^1, ..., x^{K^*}) \simeq (\theta^1_L, ..., \theta^{K^*}_L)$. We will show that this is an optimal policy of choosing winning district candidates' positions, i.e., gerrymandering. First notice the following result.

**Lemma 2.** Suppose that gerrymander can create districts with complete freedom. Suppose that $K^* < K$ (there is at least one losing district) and that $x^1 \leq x^2 \leq ... \leq x^{K^*}$. Also suppose that party $L$ minimizes party $R$'s payoff in the losing districts (as a refinement). Then, under the optimal gerrymandering, we have

1. $x^k < x^{k+1}$ for any $k$ and $k + 1$,

2. $\mu^k((x^k, 0)) \simeq 0$ holds in all $k = 1, ..., K^*$, i.e., there is no party $L$ supporters to the right of the district median $x^k$.

Let $\text{Supp}(\mu^k)$ denote the support of $\mu^k$. Next we can show that, under optimal gerrymandering, slicing of $L$ supporters occurs.
Lemma 3. Suppose that gerrymanderer can create districts with complete freedom. Suppose $K^* \leq K$ with $x^1 \leq x^2 \leq \ldots \leq x^K$. Then, $(\text{Supp}(\mu^k) \cap [-1, x^k]) \cap (\text{Supp}(\mu^{k'}) \cap [-1, x^{k'}]) = \emptyset$ for all $k < k' \leq K^*$.

The two lemmas above show that the above slice-and-mix policy is one of the optimal policy.

Lemma 4. Suppose that gerrymanderer can create districts with complete freedom, and that $\theta_L = -1$ and $\theta_R = 1$. Suppose that $K^* \leq K$. Then, labeling districts by the district median voters in ascending order, we have $x^k = \theta^k_L$ for all $k = 1, \ldots, K^*$ and $\text{Supp}(\mu^k) \cap [-1, x^k] = (\theta^{k-1}_L, \theta^k_L]$ for $k = 1, \ldots, K^*$, where $\theta^0_L = -1$. Moreover, $\bigcup_{k=1}^{K^*} (\text{Supp}(\mu^k) \cap (x^k, 1]) = [\theta^{K^*}_R, 1]$ holds. Thus, the slice-and-mix strategy is one of the optimal gerrymandering strategies, which minimizes party R’s payoff in the losing districts, and all optimal strategies generate the same outcome $x^k = \theta^k_L$ for all $k = 1, \ldots, K^*$ in terms of district median voters.

In Lemma 4, $K^*$ is a given number. How does this $K^*$ be determined? One can see that, for any $k \leq K^*$, the position of $\theta^k_L$ is independent of the magnitude of $K^*$. Therefore, Lemma 4 implies that the gerrymanderer only needs to win as many districts as possible by using slice-and-mix districting plan. Let

$$\bar{K} = \max \{k | \theta^k_L \leq 0\}$$

Since party L’s leader does not care about the losing districts, there are potentially multiple optimal gerrymandering strategies. The following proposition characterizes the optimal gerrymandering policy that minimizes party R’s leader’s payoff as a refinement.

Proposition 2. Suppose that gerrymanderer can create districts with complete freedom, and that $\theta_L = -1$ and $\theta_R = 1$. The slice-and-mix strategy is one of the optimal gerrymandering strategies, which minimizes party R’s payoff in the losing districts. Moreover, $K^* = \bar{K}$.

An interesting observation from this proposition is that $K^* < K$ means that party L is actually the minority party in terms of the state population: the global median $x^m > 0$. Otherwise, party L can win in all districts by using the slice-and-mix strategy. However, one party’s monopolizing all districts is
rare in US politics partly because of the presence of majority-minority district requirement (see Shotts (2001)).\textsuperscript{16} In the case of $K^* = K$, it is obvious that winning all districts by the slice-and-mix strategy is optimal for gerrymanderer, but the majority-minority requirement forces the gerrymanderer seeking for second-best districting plan as a result. It is worthwhile to note that the slice-and-mix strategy is identical to the optimal policy analyzed in Friedman and Holden (2008). Friedman and Holden (2008) and our paper share the features that (i) the party leader prefers a more extreme median voter’s position than a moderate one, (ii) complete freedom in gerrymandering, and (iii) there is no uncertainty in the median voter’s type unlike in Owen and Grofman (1988) and in the basic model of Gul and Pesendorfer (2010).\textsuperscript{17} However, there are big differences between our paper and Friedman and Holden (2008) have neither element in their model. Nonetheless, we can say that the above three common conditions are the keys for getting the same results.

5 Gerrymandering with Indivisible Localities

In this section, we will explore how the “slice-and-mix” result would be modified if we drop "complete freedom in gerrymandering" (and "no uncertainty in district median voter’s position" in some cases). We will consider indivisible locality case to accommodate actual geographic elements in redistricting, following the spirits of Owen and Grofman (1988) and Gul and Pesendorfer (2010). In this section, we will largely work on normal distribution case that justifies the paper in Owen and Grofman (1988). Also we use party leaders and voters have cost functions with constraint elasticity denoted by $\gamma$, although we did not need to assume any additional condition on cost functions of party leaders and voters in the complete freedom case.

Let the party leaders’ and voters’ cost functions be $C(d) = a^C d^\gamma$ and $c(d) = a^c d^\gamma$, where $\gamma > 1$, $a^C > 0$, and $a^c > 0$, and let us normalize $\theta_L$ and $\theta_R$.

\textsuperscript{16}In fact, even though either one of the two parties must be the majority in a state, the majority party usually does not win all districts. This can be attributed to Section 2 of the Voting Act Rights (accompanied with other United States Supreme Court cases) which essentially avoid the minority votes being diluted in the voting process similar to our slice-and-mix strategy.

\textsuperscript{17}Gul and Pesendorfer (2010) also include aggregate uncertainty, generalizing Owen and Grofman (1988).
at $-1$ and $1$, respectively. As we have seen before from Claim (the properties of constant elasticity cost functions), if $(\frac{2}{\gamma})^{\frac{1}{\gamma}} \geq \max \{|x^k - \theta_L|, |x^k - \theta_R|\}$ (a sufficient condition for interior solutions), the winning party $i$’s leader’s utility from district $k$ is

$$\hat{V}_i^k(x^k) = \hat{V}(|x^k|) = \mathcal{C}(\theta_j, x^k) - \mathcal{C}(\theta_i, x^k) = A \left[(|x^k| + 1)^\gamma - (1 - |x^k|)^\gamma \right]$$

where $A$ is a constant. As is shown in Claim, we have (i) $\hat{V}' > 0$, and (ii) $\hat{V}'' \gtrless 0$ if and only if $\gamma \gtrless 2$. This result (ii) may require some explanation. The readers may be puzzled by having concave $\hat{V}$ function under $1 < \gamma < 2$, since cost functions $\mathcal{C}(\theta_j, x^k)$ and $\mathcal{C}(\theta_i, x^k)$ are always strictly convex in $|x^k|$. The reason for (ii) to happen is that $\hat{V}$ subtracts a strictly convex function from another strictly convex function, and such a function can be anything in general. If the cost function is a constant elasticity type, however, we happen to be able to say that $\hat{V}$ is strictly concave or strictly convex depending on $\gamma \gtrless 2$.

Now, consider two partitions $\pi$ and $\pi'$ such that $D^{k'}(\pi) = D^{k'}(\pi')$ for all $k' \neq \{k, \tilde{k}\}$. Thus, $D^k(\pi')$ and $D^{\tilde{k}}(\pi')$ are created by swapping some localities between $D^k(\pi)$ and $D^{\tilde{k}}(\pi)$: i.e., $S = D^k(\pi) \setminus D^k(\pi') = D^k(\pi') \setminus D^{\tilde{k}}(\pi)$ and $T = D^{\tilde{k}}(\pi) \setminus D^{\tilde{k}}(\pi') = D^{\tilde{k}}(\pi') \setminus D^k(\pi)$ are swapped. Is such swapping beneficial to the gerrymandering party leader? By swapping localities $S$ and $T$, $x^k$ and $x^{\tilde{k}}$ must move in the opposite directions. We call a swap “cracking” if the median voters of two districts involved moves closer to each other after swapping and a swap “anti-cracking” if the opposite happens.

We will consider the case in which voters’ distribution in each locality follows a normal distribution.

### 5.1 Normally Distributed Voters

Owen and Grofman (1988) analyzed the optimal partisan gerrymandering policy by imposing a constraint such that the sum of $x^k$s, $\sum_{k=1}^{K} x^k(\pi)$, must stay constant for all $\pi$. They obtained the famous pack-and-crack result when the party leader maximizes the number of seats under this constraint. In our indivisible locality case, if each locality has normally distributed voters, the feasibility constraint becomes exactly this constraint (the proof is obvious by noting that the median is equivalent to the mean under normality).
Lemma 5. Suppose that the voter distribution in each locality is normal distributed, i.e., \( F_\ell \sim N(\mu_\ell, \sigma_\ell) \) for each \( \ell \in \mathcal{L} \). Then, the median of district \( k \) is

\[
x^k(\pi) = \frac{1}{n} \sum_{\ell \in D^k(\pi)} \mu_\ell.
\]

This lemma automatically implies that \( \sum_{k=1}^K x^k(\pi) \) must stay constant, since \( \sum_{k=1}^K x^k(\pi) = \frac{1}{nK} \sum_{\ell \in \mathcal{L}} \mu_\ell \). Therefore, under the normal distribution assumption, swapping the sets of localities \( S \) and \( T \) between districts \( k \) and \( \bar{k} \), we have

\[
\Delta = x^k(\pi') - x^k(\pi) = \frac{\sum_{\ell \in T} \mu_\ell - \sum_{\ell \in S} \mu_\ell}{n} = \bar{x}^k(\pi) - \bar{x}^k(\pi'),
\]

while \( x^k(\pi') + x^\bar{k}(\pi') = x^k(\pi) + x^\bar{k}(\pi) \). Without loss of generality, assume \( x^\bar{k}(\pi') < x^k(\pi) < x^k(\pi') < 0 \). This means that \( \pi \) is more cracked than \( \pi' \), and we have

\[
\begin{align*}
\tilde{V}_L(\pi') - \tilde{V}_L(\pi) &= \left( \tilde{V}_L^k(x^k(\pi')) + \tilde{V}_L^{\bar{k}}(x^{\bar{k}}(\pi')) \right) - \left( \tilde{V}_L^k(x^k(\pi)) + \tilde{V}_L^{\bar{k}}(x^{\bar{k}}(\pi)) \right) \\
&= \left( \hat{V}'(\|x^k(\pi')\|) + \hat{V}'(\|x^{\bar{k}}(\pi')\|) \right) - \left( \hat{V}'(\|x^k(\pi)\|) + \hat{V}'(\|x^{\bar{k}}(\pi)\|) \right)
\end{align*}
\]

Since function \( (\cdot)^{\gamma-1} \) is strictly concave or strictly convex depending on \( \gamma < 2 \) or \( \gamma > 2 \), respectively. As is easily seen from Claim and Figure 1, if \( \gamma < 2 \) (\( \gamma > 2 \)) then cracking from \( \pi' \) to \( \pi \) improves (reduces) party \( L \)'s leader's utility, but if \( \gamma = 2 \) then there is no effect. Thus, depending on the value of \( \gamma \), "cracking" can be a good or bad strategy.\(^{18}\) When \( \gamma > 2 \) holds, concentrating supporting voters from the strongest to moderate ones, or "slicing" is optimal.

Proposition 3. Suppose that party leaders and voters have constant elasticity cost functions, and that the voter distribution is normal in each locality.

\(^{18}\)Puppe and Tasnadi (2009) consider a model in which party cares about only the number of seats and geographical constraints on feasible districting plan. They show that pack-and-crack does not generally work.
Then, cracking is preferable when $\gamma < 2$, while slicing is preferable when $\gamma > 2$. When $\gamma = 2$, cracking or slicing does not matter.  

[Figure 2 here]

The reason that the result varies depending on the value of $\gamma$ is that the party leader has policy preference. When $\gamma > 2$, $V(|x^k|) = A[|x^k| + 1]^{\gamma} - (1 - |x^k|)^{\gamma}$ is strictly convex in $|x^k|$, which means that slicing is beneficial. This is because the opponent party leader’s cost of matching $x^k$’s utility is increasingly high which implies that the winning party’s leader can save a lot of money in such a district. We can show that the optimal gerrymandering policy depends on $\gamma \geq 2$, and we obtain the famous pack-and-crack result when $\gamma < 2$. In contrast, when $\gamma > 2$, then pack-and-slice is the optimal.

In order to characterize the optimal gerrymandering policy for $\gamma > 2$ under our normality assumption, it is useful to order localities by their means. Let $\ell < \ell'$ mean $\mu_\ell \leq \mu_{\ell'}$. Then, since $x^k = \mu^k = \frac{1}{n} \sum_{\ell \in D^k} \mu_\ell$ hold, if party $L$ cannot win all districts, it is beneficial to pack party $R$ supporting localities. Let $\bar{x}^k = \frac{1}{n} \sum_{\ell = n(k-1)+1}^{nk} \mu_\ell < 0$ for all $k = 1, 2, ..., K$, and let $\bar{K} \leq K$ be a nonnegative integer such that $\bar{x}^\bar{K} < 0$ and $\bar{x}^{\bar{K}+1} = 0$. Let us call the allocation $(\bar{x}^k)_{k=1}^{\bar{K}}$ a pack-and-slice gerrymandering policy. This policy pack all opponents in unwinable districts, and slices supporting voters from the most extreme to moderate in order. This involves districts with voter group who are even more extreme than the party leader. This is because with convex $V$, the opponent party’s cost to attract the median voter of such a district is so high, and the winning party’s leader does not need to pay much in the pork-barrel politics. Thus, if $\gamma > 2$, we have the following characterization of the optimal gerrymandering policy.

**Proposition 4.** Suppose that party leaders and voters have constant elas-

---

19Given this observation, some information about optimal gerrymandering strategy can be derived. Suppose that the gerrymandering party $L$ wins $K^*$ districts and $\gamma > 2$. We rank the winning district such that $x^1 < x^2 < \ldots < x^{K^*}$. Also, we rank the localities in winning districts by $\ell^1 < \ell^2 < \ldots < \ell^{nK^*}$. Then it is optimal for the gerrymanderer to pack $n$ most extreme localities into district 1 and $n$ next most extreme localities into district 2 and so on. It is obvious that any other districting plan is not optimal since there always exists a anti-cracking swap.

20The shape of payoff function depends on the ideology and pork barrel cost functions. There is no reason to believe those cost functions following some particular forms. On the other hand, in Gul and Pesendorfer (2010) the shape of payoff functions depends on the shape of local shock distribution. It is more reasonable to assume concavity in their setup, justifying “pack-and-crack.”
ticity cost functions with $\gamma > 2$, and that the voter distribution is normal in each locality. Then, the optimal gerrymandering policy is described by pack-and-slice with \((x^k)_{k=1}^K\).

How about the case with $\gamma < 2$? By the same logic above, it is easy to see that the optimal policy must be to “pack-and-crack”, but it is hard to characterize the policy by the following two reasons: (i) there are many locally optimal cracking district plans but it is hard to say which one is the very best due to the integer problem, and (ii) the optimal number of winning districts is hard to characterize. For the second point, see the following example.

**Example 1.** Suppose $\gamma < 2$. There are 20 SL-type localities in which the voter distribution is normal and the mean voter is $-1$ (*Strong Left supporter*), 5 WR-type localities with mean voter 0.1 (*Weak Right supporter*), and 5 SR-type localities with mean voter 1 (*Strong Right supporter*). Each district must include 5 localities and there are 6 districts. If the gerrymandering party is $L$, the optimal districting plan must group the SR-type localities together. The gerrymanderer must decide between grouping 4 WR localities with 1 SL in five winning districts or grouping 5 SL localities in four winning districts. Then the decision is to win 4 districts with $x^k = 1$ or to win 5 districts with $x^k = 0.78$. The payoff difference between these two plans is

$$D = A[4 \times 2^\gamma - 5 \times ((1.78)^\gamma - (.22)^\gamma)]$$

When $\gamma < 1.226$ and $\gamma > 1.619$, $D > 0$ which means that $L$ should form strongly winning district and lose a relative moderate district. Otherwise, $D < 0$ means that the party should win more districts by forming weaker winning districts. This example shows that the decision of the number of winning districts is not monotonic and trivial even in the case under the normal distribution assumption.

### 5.2 With Uncertainty

In this subsection, we extend our discrete locality setup to include uncertainty. This extension can make an interesting comparison with the result in the gerrymandering literature. Up to this point, our model is a deterministic model. This is to say that when the gerrymandering is done, both
parties can predict the outcome of elections with perfect accuracy. However, this might oversimplifies the real world gerrymandering problem. In the U.S., since the redistricting is decennial, there is usually a time gap between redistricting and election. Even though the gerrymanderer may have a great amount of freedom and information when redistricting is conducted, the positions of district median voters can still be modified by demographic changes that take place after gerrymandering is conducted. In other words, the gerrymander faces some uncertainty in the positions of district median voters. A common approach of modeling uncertainty is to assume local uncertainty that the positions of district median voters are affected by i.i.d. district-specific shocks. These shocks realize after redistricting. How does the gerrymandering strategy change in the presence of uncertainty? An intuitive argument is that, since now a winning district can become a losing one when the shock hits, the gerrymander would prefer to win by a larger margin. Therefore, the presence of uncertainty seems to enhance the incentive to cracking the supportive localities to avoid winning with small margin in a mild (ex ante) supporting district. This intuition is the fundamental reason behind Owen and Grofman (1988). Does the same argument apply to our case with policy-motivated gerrymanderer?

Let local shock \( y^k \) be a random variable with density function \( g^k : [-\infty, \infty] \to \mathbb{R}_+ \) and \( y^k \) being a representative realization. We also assume that \( y^k \)'s are i.i.d. and denote \( g^k = g \). We also normalize the mean so that \( \mathbb{E}(y^k) = 0 \). After \( y^k \) realizes, the \( k \) district’s median voter position shifts from \( x^k(\pi) \) to \( x^k(\pi) + y^k \). We also modify the timing of the game as the following:

1. One party, say \( L \), chooses a districting partition \( \pi = \{ D^1, \ldots, D^K \} \) of \( L \).

2. Nature plays and the local shock \( y^k \)'s are realized.

3. Party leaders \( L \) and \( R \) simultaneously choose policy positions and pork-barrel promises \( (\beta^k_L, t^k_L)_{k=1}^K \) and \( (\beta^k_R, t^k_R)_{k=1}^K \) (under the above behavioral assumption), respectively.

4. All voters vote sincerely (with our tie-breaking rule). The winning party is committed to its policy position and its pork-barrel promise in each district \( k = 1, \ldots, K \). All payoffs are realized.
Notice that the result of political competition is still deterministic. Therefore, the equilibrium of election stage is similar to the one in the deterministic case. The only difference is that, now, the sufficient statistics for a district becomes $x^k + y^k$ and the winning party $L$’s payoff becomes $\hat{V}(|x^k + y^k|)$. Therefore, in the first stage, even if $x^k < 0$, $L$ can only wins when $x^k + y^k < 0$ \textit{ex post}. On the other hand, even in an \textit{ex ante} losing district, i.e., $x^k > 0$, $L$ can win when $x^k + y^k < 0$. At the first stage, the gerrymanderer maximizes the expected payoff in the future election. To isolate the effect of uncertainty, we focus on the case of $\gamma = 2$. Notice that given this case the gerrymanderer is indifferent between any swap opportunity. Using Claim on the properties of constant elasticity cost functions, the gerrymandering maximization problem in the first stage is

$$\max_{\pi \in \Pi} \bar{V}_L(\pi)$$

where

$$\bar{V}_L(\pi) = \sum_{k=1}^{K} \int_{-\infty}^{-x^k(\pi)} \hat{V}(|x^k(\pi) + y|) g(y) dy$$

$$= \sum_{k=1}^{K} \int_{-\infty}^{-x^k(\pi)} -4A (x^k(\pi) + y) g(y) dy.$$

Now, we consider a cracking swap similar to the one discussed in the previous section. Consider districts $k$ and $\tilde{k}$ with $x^k < x^\tilde{k} < 0$, and a swap that decreases $x^k$ and increases $x^\tilde{k}$ by the same amount $\Delta$ (to keep the sum of median voters’ positions constant). Notice that

$$\frac{d}{d\Delta} \left( \int_{-\infty}^{-x^k(\pi)+\Delta} -4 \left( x^k(\pi) + \Delta + y \right) g(y) dy \right)$$

$$= -4A[\hat{G}(-x^k) - x^k g(-x^k) - (-x^k g(-x^k))] = -4A\hat{G}(-x^k)$$

Thus, the effect on the party leader’s payoff of such swapping is $-4A(\hat{G}(-x^\tilde{k}) - \hat{G}(x^k)) < 0$, since $x^k < x^\tilde{k}$ implies $\hat{G}(-x^\tilde{k}) > \hat{G}(-x^k)$. Hence, even when $\gamma = 2$, which was a neutral case under determinacy, cracking is \textit{not} beneficial.

**Proposition 5.** Suppose that the number of winning districts is the same. Moreover, suppose that in each locality the voter distribution is normal. When $\gamma = 2$, slicing is beneficial under local uncertainty.
Corollary. In the same setup as Proposition 5, introduce of local uncertainty make the “pack-and-slice” strategy optimal even then $\gamma = 2$.

Strikingly, our result is the opposite of pack-and-crack strategy in Owen and Grofman (1988). The reason behind it is that in Owen and Grofman (1988), the party only cares about winning probability and do not care about the cost of winning a district. The only reason a gerrymanderer move the position of a median voter is that this operation yields a higher winning probability. In Owen and Grofman (1988), the party leader’s objective is the sum of winning probabilities (i.i.d.):

$$\max_{\pi \in \Pi} W_L(\pi, \theta_L, \theta_R)$$

where

$$W_L(\pi, \theta_L, \theta_R) = \sum_{k=1}^{K} \int_{-\infty}^{-x_k(\pi)} g(y) dy$$

Let us do the same exercise of a cracking swap between districts $k$ and $\tilde{k}$ (decreasing $x^k$ and increasing $x^{\tilde{k}}$ by the same amount $\Delta > 0$). Furthermore, we assume that the distribution $g$ is unimodal. Then, the effect on the party leader’s payoff of such swapping is $-(g(-x^{\tilde{k}}) - g(x^k)) > 0$, since $x^k < x^{\tilde{k}} < 0$ implies $g(-x^{\tilde{k}}) < g(-x^k)$ by unimodality. Hence, cracking is beneficial if the party leader does not care about policy positions. However, in our case, the gerrymanderer cares about the winning probability and the position of median voter. An extreme median voter not only gives a higher winning probability but also a low pork barrel cost. This extra incentive makes an extreme redistricting plan more attractive under $\gamma = 2$. This means that even if $\gamma < 2$, there can be cases that anti-cracking is beneficial to the party leaders, despite that the payoff function is concave in $x^k$.

5.3 Without Normality: The Case of FOSD-ordered Localities

What if we drop the linearity embedded in the normal distribution assumption? Once normality is dropped, we can no longer use the result of Lemma 5: if some localities are swapped between districts $k$ and $\tilde{k}$, then the sizes of the shifts of $x^k$ and $x^{\tilde{k}}$ are unknown though the directions of the moves are opposite. This makes the analysis very complicated, so we will consider the
case where localities are ordered in the first-order stochastic domination (FOSD) in their voter distributions over the political spectrum, and the gerrymanderer can create political districts by using any combination of localities. Formally, we assume the following:

**Assumption 3. (FOSD)** The set of localities \( \mathcal{L} \) is FOSD-ordered if for every pair of localities \( \ell, \ell' \) with \( \ell < \ell' \), we have (i) \( F_\ell(\theta) > F_{\ell'}(\theta) \) for all \( \theta \in [-1, 1] \), or (ii) \( F_\ell(\theta) = F_{\ell'}(\theta) \) for all \( \theta \in [-1, 1] \). Any combination of \( n \) localities can form a political district.

That is, smaller number localities are more leaning towards \( L \) party, and as the number increases localities are leaning more towards \( R \) party. The following property is the direct consequence of FOSD-ordered localities.

**Lemma 6.** Suppose that \( \mathcal{L} \) is FOSD-ordered. Let \( \ell \in D^k \) and \( \ell' \notin D^k \) be such that \( \ell' < \ell \) and \( F_\ell \neq F_{\ell'} \), and let \( D^{k'} = (D^k \cup \{\ell'\}) \setminus \{\ell\} \). Then, \( x^k(D^{k'}) < x^k(D^k) \) holds.

This result is a direct consequence of the FOSD-ordered localities.\(^{21}\) The qualification \( F_\ell \neq F_{\ell'} \) implies case (i) of the FOSD definition holds, and the district median moves to left by replacing locality \( \ell \) by locality \( \ell' \). Recall Proposition 1.5 that \( \bar{V}_L^k(x^k, \theta_L, \theta_R) = C(\theta_R, x^k) - C(\theta_L, x^k) \) is a decreasing function in \( x^k \), we have the following result:

**Proposition 6 (Weak Segregation).** Suppose that \( \mathcal{L} \) is FOSD-ordered. For simplicity, suppose further that \( \theta_L = -1 \) and \( \theta_R = 1 \). Then, under the optimal gerrymandering, there is a locality \( \ell^* \) such that all localities \( \ell \leq \ell^* \) belong to \( L \)'s winning districts, while all localities \( \ell > \ell^* \) belong to losing districts, unless there are completely tied districts.

This result is reminiscent to the packing gerrymandering strategy mentioned above. Thus, if localities are not fine enough to allow us slice-and-mix strategies, and localities are FOSD-ordered, the gerrymanderer packs those relatively unfavorable localities to form losing districts in \( \pi^* \). Again note that this packing result is obtained under deterministic model unlike Owen and Grofman (1988). Our result is based on the assumption of the FOSD-ordered

---

\(^{21}\) Suppose that FOSD assumption does not hold, even if \( \ell \) has a larger median than \( \ell' \) does, it is possible that \( x^k(D^{k'}) \geq x^k(D^k) \) where \( D^{k'} = (D^k \cup \{\ell'\}) \setminus \{\ell\} \). One can replace FOSD assumption by simply assuming that the median voter increases when replace a locality with any other one with higher median.
localities and the slope of winning payoff function. The FOSD assumption
may sound strong, but there are geographical constraints in redistricting
in the real world. If this were the case, the principle of the proof of the
proposition might still work even if localities are not exactly ordered according to
FOSD.

Now we turn to the other side of the traditional strategy in Owen and
Grofman (1988), cracking one’s supports to form homogeneous winning dis-


tRICTS. We consider a similar swapping localities between two districts \( k \) and
\( \~k \) as before. However, with FOSD assumption, we cannot say much about
how this swap will affect \( x^k \) and \( x^{\~k} \) except for the opposite directions of mov-
ing (say, decreasing \( x^k \) and increasing \( x^{\~k} \)). Generally speaking, the cracking
strategy creates more homogeneous districts. It increases \( \~V^k_L \) but decreases
\( \~V^k_L \). The characterization of these two effects depends on the shape of cost
functions (\( c \) and \( C \), as we have demonstrated in the normality case) and all \( F_i \)’s.
Therefore, it is difficult to derive the optimal gerrymandering strategy
in general. To compare with the previous argument, we again consider the
case of common elasticity cost functions with \( \gamma = 2 \) but consider general
locality distributions.

Example 2. Suppose that \( \theta_L = -1 \) and \( \theta_R = 1 \). We now consider a
simple case with 6 localities \( L = \{l, l, w_l, w_l, r, r\} \) where locality \( l \) stands for
left-leaning locality with the distribution \( F_l \) and the median position \( x_l \) less
than 0. A locality \( r \) is right-leaning with the median larger than 0. A locality
\( w_l \) is the weakly left-leaning with the median position equal to \(-x_{w_l} \in (x_l, 0)\). Also we assume that \( a^C = a^c = 1 \), \( \gamma = 2 \), and \( Q > 1 \).

In this case, there are two possible optimal districting plan for \( L \): \( D = \{\{l, w_l\}, \{l, w_l\}, \{r, r\}\} \) or \( D’ = \{\{l, l\}, \{w_l, w_l\}, \{r, r\}\} \). We can easily
compare these two districting plans. Notice that these two plans both packing the
unfavorable localities to form the losing district in accordance with Propo-
sition 6. The former one cracking the strong supportive localities (\( \{l, l\} \) in
this example) into the winning districts, i.e., pack-and-crack. In contrast,
the second plan slices supportive localities in the winning district, i.e., pack-
and-slice.

Proposition 7. In Example 2 with \( \gamma = 2 \), \( \{l, l, w_l, w_l, r, r\} \) and \( K = 3 \),
\( L \) (strictly) prefers cracking, i.e., districting plan \( D \) over \( D’ \), if and only if
\[ \frac{1}{2}(x^l + x^{w_l}) = \frac{1}{2}(x^{l,l} + x^{w_l, w_l}) > x^{l,w_l} \], which is equivalent to
\[ \int_{\frac{1}{2}(x^l + x^{w_l})}^{\frac{1}{2}(x^{l,l} + x^{w_l})} f_i d\theta > \int_{\frac{1}{2}(x^l + x^{w_l})} f_s d\theta \].

27
There is no reason to believe the last inequality to hold or not, since it is not related to the FOSD assumption. This inequality means that $L$ prefers $D$ if voters in his strong supporting localities are more clustered around the median than those in the weak supporting localities. If $\int_{\frac{1}{2}x^l + x^{wl}}^{x^l + x^{wl}} f_{l} d\theta > \int_{\frac{1}{2}x^l + x^{wl}}^{x^{wl}} f_{s} d\theta$ holds and $L$ crack strong supporters with weak ones in the winning district, the median voter will be pulled toward left a lot due to the fact that the strong supporting locality is more clustered. In this case, $L$ wins both districts and the costs of pork-barrel transfer is moderate. On the other hand, if the inequality fails to hold, weak winning district would demand too much money and the equilibrium policy would be too far away from $\theta_L$. Therefore, instead of pack-and-crack, the gerrymanderer should choose pack-and-slice. Note that $\gamma = 2$ is a special case that makes $V^k_L$ linear in $x^k$. Compare with the normality case, this example isolates the effect of locality distribution from the curvature effect shown in the normality case. If $\gamma > 2$ or $1 < \gamma < 2$, then although each agent’s cost function is convex, function $\tilde{V}^k_L$ becomes strictly convex and strictly concave in $x^k$, respectively. Under such situations, it becomes even harder to say if cracking is the right strategy.

6 Conclusion

In this paper, we propose a gerrymandering model with endogenous candidates’ political positions, in which two parties compete in their positions and pork-barrel politics. Even though we mostly assume deterministic outcomes, we can generate policy divergence unlike Wittman (1983). We derive the optimal partisan gerrymandering policies under different circumstances—from the case of the gerrymanderer having full freedom in redistricting to the cases with indivisible localities with various voter distributions.

Our model is flexible enough to analyze the optimal policies under various feasibility constraints in gerrymandering. The results depend on circumstances. We obtain “slice-and-mix,” “pack-and-crack,” and “pack-and-slice” strategies, some of which are obtained in different gerrymandering models in

\[22\text{To be more precise, the inequality requires that voters who is relatively right (compared to the median, i.e., } \theta \in [x^l, \frac{1}{2}(x^l + x^{wl})]) \text{ in } l \text{ have a larger population mass than voters who relatively left (} \theta \in [\frac{1}{2}(x^l + x^{wl}), x^{wl}) \text{) in } w.l.\]
There are possibly interesting extensions. The first one is to introduce the uncertainty into our model. Our conjecture is that, if the uncertainty is infinitesimal, e.g., the gerrymander can only observe that the median voter’s position belongs to the interval \([x^k - \varepsilon, x^k + \varepsilon]\) for \(\varepsilon\) being a (small) preference perturbation and if the gerrymanderer has complete freedom in redistricting, the slice-and-mix strategy may still be optimal à la Friedman and Holden (2008). However, with significant uncertainty in median voters’ positions like in Gul and Pesendorfer (2010), we do not know what can happen. Second, in this paper we concentrated on one type of pork-barrel politics: candidates “promise” transfer contingent to their winning the districts. This kind of promises are different from campaign financing. In the latter case, even if a candidate loses in a district, the spent campaign expenditure will not come back. In some circumstance, such a model may be more realistic. These extensions are subject to future research.

Appendix A: Proofs

**Proof of Lemma 1.** First, by Assumption 2, the non-negativity constraint of \(t^k_R\) is not needed. There are three cases: if \(R\) loses with \(Q - t^k_R - C(|\theta_R - \beta^k_R|) > 0\) and its offer gives the median voter utility equal to \(\bar{U}\) in equilibrium, it must be that \(L\) wins with positive indirect utility and also provides the median voter with the utility level \(\bar{U}\). However, this means that \(R\) can win the election by providing, say, \(\varepsilon\) more pork-barrel promise. This contradicts with the equilibrium condition. The second case is that \(Q - t^k_R - C(|\theta_R - \beta^k_R|) = 0\) but \(U_x(R)\) is not maximized. In this case, there must exist some points \((t', \beta')\) which satisfies \(Q - t' - C(|\theta_R - \beta'|) = 0\) but the pair provides the median voter strictly higher utility. Then any point on the segment connecting \((t^k_R, \beta^k_R)\) is strictly better off for both \(R\) and the median voter \(x^k\) by the strict convexity of the preferences. A contradiction. The third case, \(Q - t^k_R - C(|\theta_R - \beta^k_R|) < 0\), cannot happen, since the strategy that generates a negative payoff is a dominated strategy for \(R\)'s leader. \(\square\)

**Proof of Proposition 1** The proofs of part 1 to 3 are in the main text. \(\tilde{V}^k_i\) can be derive from substituting the equilibrium \(t^k_{iR}\) into the utility function of winning party \(i\). As for part 5, we first set \(i = L\) and take derivative of \(\tilde{V}^k_L\) with respect to \(x^k\). There are two cases: if \(\theta_L \leq x^k\), \(\frac{d\tilde{V}^k_L}{dx^k} = -c^k_R - x^k\) -
\[ c^k - \beta_L^{k^*} < 0. \] If \( \theta_L > x^k \), \( \frac{dV^k}{dx^k} = -c_R^{k^*} - x^k + c_L^{k^*} - x^k < 0. \) The inequality still holds since \( c(\cdot) \) is an increasing function and \( \beta_L^{k^*} - x^k < \beta_R^{k^*} - x^k. \)

**Proof of Claim.** By Proposition 1, if \( \theta_L < x^k \) then \( \beta_L^k \in (\theta_L, x^k) \) is characterized by
\[
a^C \gamma (\beta_L^k - \theta_L)^{-1} = C'(\beta_L^k - \theta_L) = c'(x^k - \beta_L^k) = a^C \gamma (x^k - \beta_L^k)^{-1},
\]
or \( \alpha = \left( \frac{a^C}{\alpha} \right)^{\frac{1}{\gamma - 1}} = \frac{x^k - \beta_L^k}{\beta_L^k - \theta_L}. \) Thus, \( \beta_L^{k^*} = \frac{\theta_L + \alpha x^k}{1 + \alpha}, \) and
\[
C(\beta_L^{k^*} - \theta_L) + c(x^k - \beta_L^{k^*}) = a^C \left( \frac{\alpha}{1 + \alpha} (x^k - \theta_L) \right)^\gamma + a^C \left( \frac{1}{1 + \alpha} (x^k - \theta_L) \right)^\gamma \\
= \left( a^C \left( \frac{\alpha}{1 + \alpha} \right)^\gamma + a^C \left( \frac{1}{1 + \alpha} \right)^\gamma \right) (x^k - \theta_L)^\gamma \\
= A (x^k - \theta_L)^\gamma \tag{10}
\]
where \( A = A(a^C, a^c) = a^C \left( \frac{\alpha}{1 + \alpha} \right)^\gamma + a^C \left( \frac{1}{1 + \alpha} \right)^\gamma > 0. \) Thus, in general, the sum of the party leader’s and the median voter’s costs \( C(\theta_i, x^k) \) can be written as
\[
C(\theta_i, x^k) = A \left( |x^k - \theta_i| \right)^\gamma.
\]

If \( A(a^C, a^c) \left( |x^k - \theta_i| \right)^\gamma \leq Q \) for both parties \( i = L, R \), then Assumption 2 is satisfied. Then, Proposition 1.4 implies that if party \( i \) is the winner in district \( k \) (i.e., \( |x^k - \theta_i| < |x^k - \theta_j| \)), then party \( i \)'s winning payoff is
\[
\hat{V}_i^k(x^k, \theta_i, \theta_j) = A \left[ \left( |x^k - \theta_j| \right)^\gamma - \left( |x^k - \theta_i| \right)^\gamma \right]
\]
Let \( \theta_L = -1 \) and \( \theta_R = 1 \) without loss of generality. \( \Box \)

**Proof of Lemma 2.** First, we show that \( x^k < x^{k+1} \) by contradiction. If \( x^k = x^{k+1} \), then by swapping \( \epsilon \) measure population of the right-most voter in \([-1, x^k]\) in district \( k \) with \( \epsilon \) measure population of the left-most voter in \([-1, x^{k+1}] \) in district \( k + 1 \), party \( L \) can move \( x^k \) to the left without changing \( x^{k+1} \) which strictly increases \( L \) leader’s payoff by Proposition 1.5. Thus, we have \( x^k < x^{k+1}. \) For the second part, suppose that \( \mu^k((x^k, x^{k+1})) \) is not. Then, suppose that \( \mu^k((x^{k+1}, 0)) > 0 \) holds. In this case, party \( L \) can move losing districts’ candidate to the left by swapping the district \( k \) population in the interval \((x^{k+1}, 0)\) with the party \( R \) supporters in losing districts. This
is again a contradiction. For the case \( k = K^* \), if \( \mu_k((x^k,0)) > 0 \) holds, then by swapping this population with the losing districts’ right-wing population, the losing districts’ candidates can be moved to the left which is not optimal. Hence, \( \mu_k((x^k,0)) = 0 \) must hold for all \( k = 1, \ldots, K^* \). □

**Proof of Lemma 3.** The proof is rather obvious. From Lemma 2, we know that \( x^1 < x^2 < \ldots < x^{K^*} \) holds, and that \( \mu_k((x^k,0)) = 0 \) for all \( k = 1, \ldots, K^* \). Thus, if \( (Supp(\mu^k) \cap [-1,x^k]) \cap (Supp(\mu^k) \cap [-1,x^{k'}]) \neq \emptyset \), then by swapping the population at the most left in district \( k' \) and those voters at the most right within \([-1,x^k] \) in the district \( k \), party \( L \) leader can move \( x^k \) to the left without \( x^{k'} \) intact. This is a contradiction. □

**Proof of Proposition 4.** When \( \gamma > 2, \hat{V}(|x^k|) = A[|(x^k| + 1)^\gamma - (1 - |x^k|)^\gamma] \) is convex in \(|x^k|\) even for the case of \(|x^k| \geq 1 \). This convexity makes slicing policy is more preferable than cracking within the winning districts, resulting in \( K^* \) winning districts. Noting the gerrymandering party leader gets zero payoff from the losing districts, the party does not want to increase the number of winning districts by moving supporting voters to another district. Given the convexity of \( \hat{V} \), it does not pay to dilute their supporters. □

**Proof of Proposition 6.** Recall that party \( L \) maximizes \( \hat{V}_L(\pi) = \sum_{k=1}^{K} I_L(x^k(D^k))\hat{V}_L^k(x^k(D^k)) \).

Thus, what is relevant is to maximize the indirect utilities of party \( L \)’s leader from her winning districts. Suppose that there is a pair \( \ell \in D^k \) and \( \ell' \in D^{k'} \) under \( \pi^* \), where \( I(x^k(D^k)) = 1 \) and \( I(x^{k'}(D^{k'})) = 0 \), respectively. Then, by swapping \( \ell \) and \( \ell' \), \( I(x^k((D^k \cup \{\ell\}) \backslash \{\ell'\})) = 1 \) and \( I(x^{k'}((D^{k'} \cup \{\ell'\}) \backslash \{\ell\})) = 0 \) still hold, while \( x^k((D^k \cup \{\ell\}) \backslash \{\ell'\}) < x^{k'}(D^{k'}) \) and \( x^{k'}((D^{k'} \cup \{\ell'\}) \backslash \{\ell\}) > x^{k'}(D^{k'}) \) must follow by Lemma 6. By Proposition 1.5, the party leader \( L \)’s payoff increases by this swapping. Thus, \( \pi^* \) is not the optimal choice for her. This is a contradiction. Hence, we conclude the weak segregation result. □

**Proof of Proposition 7.** The first claim is straightforward from \( \hat{V}_L^k = -2x^k \). For Lemma 6, we know that \( x^l = x^{\{l,l\}} < x^{\{l,w\}} < x^{wl} = x^{\{wl,wl\}} \). The
The definition of median voter implies that

$$
\frac{1}{2} = \frac{1}{2} \int_{-\infty}^{x_l} (f_l + f_wl)d\theta = \frac{1}{2} \left( \int_{-\infty}^{x_l^l} f_l d\theta + \int_{x_l^l}^{x_l^w} f_l d\theta + \int_{-\infty}^{x_w^l} f_wl d\theta + \int_{x_w^l}^{x_w^w} f_wl d\theta \right)
$$

Therefore, $\frac{1}{2} (x^l + x^w) > x^{l,w}$ implies $\int_{x_l^l}^{x_l^w} f_l d\theta > \int_{\frac{1}{2} (x^l + x^w)}^{x_w^w} f_wl d\theta$. \square

**Appendix B: Corner Solutions**

In this appendix, we discuss the situations when Assumption 2 is violated. There are several cases. First, suppose that $Q \geq C(|\theta_j - \hat{\beta}(x^k, \theta_j)|)$ for the losing party. In this case, we allow the median voter to have negative utility in the equilibrium. This is acceptable since the voter cannot choose to leave the society. So, the equilibrium utility can be negative. In this case, since the constraint $t_j^k$ is not binding for the losing party $j$, the optimal policy for the losing party is the same as the one predict in Proposition 1.

However, the winning party may not need to promise any pork barrel policy. To see this, notice that the losing party $j$ provides the median voter with utility $\hat{U}_j^k = Q - C(|\theta_j - \hat{\beta}(x^k, \theta_j)|) - c(|x^k - \hat{\beta}(x^k, \theta_j)|)$. Without Assumption 2, it is not guaranteed that $\hat{U}_j^k + c(|x^k - \hat{\beta}(x^k, \theta_i)|)$ is positive for winning party $i$. If $\hat{U}_j^k + c(|x^k - \hat{\beta}(x^k, \theta_i)|) < 0$, it means that $\hat{\beta}(x^k, \theta_i)$ is too “generous” for the median voter in order to beat the losing party candidate. Therefore, instead of offering some positive pork barrel, the winning party candidate actually needs to taxes the district $k$ to tie offer from $j$. So, to tie $j$’s offer, the winning party will propose $0$ pork barrel promise and policy position $\beta_{i}^{k*} = \hat{\beta}(x^k, \theta_j)$ which is defined by

$$
-c(|x^k - \hat{\beta}|) = \hat{U}_j^k = Q - C(|\theta_j - \hat{\beta}(x^k, \theta_j)|) - c(|x^k - \hat{\beta}(x^k, \theta_j)|).
$$

This simply means that the winning party uses only policy position to compete. This happens when $x^k$ is very close to the ideology position of winning party, $\theta_i$. Also notice that (i) $\hat{\beta}(x^k, \theta_j)$ is closer to $\theta_i$ than $\hat{\beta}(x^k, \theta_j)$, and (ii)
the nonstrategic nature of policy choice in the interior solution does not hold when corner solution happens. Moreover, the winning payoff for the party leader $i$ from district $k$ is

$$\tilde{V}^k_i(x^k, \theta_i, \theta_j) = C(\theta_j, x^k) - \tilde{C}(\theta_j, \theta_i, x^k),$$

where $\tilde{C}(\theta_j, \theta_i, x^k) \equiv C(|\theta_i - \tilde{\beta}(x^k, \theta_j)|) + c(|x^k - \tilde{\beta}(x^k, \theta_j)|)$. Mostly, the payoff function of winning party will be the one in Proposition 1.4, and has a kink at $\bar{x}^k$ where $-c(|\bar{x}^k - \hat{\beta}(\bar{x}^k, \theta_i)|) = Q - C(|\theta_j - \hat{\beta}(\bar{x}^k, \theta_j)|) - c(|\bar{x}^k - \hat{\beta}(\bar{x}^k, \theta_j)|)$.

Furthermore, if $x^k$ moves even closer to the winning party such that $\hat{\beta}(x^k, \theta_j)$ is more extreme than $\theta_i$. In this case, the winning party wins by proposing $t^*_k = 0$, $\beta^*_k = \theta_i$ and $\tilde{V}^k_i = Q$. The following graph is the when $Q = 1$ and $C(d) = c(d) = d^2$ which is an example of the discussion above.

[Figure 3 is here]

Next, we briefly comments on the case when $Q$ is small or $x^k$ is too close to the winning party so that $Q \leq C(|\theta_j - \beta^*_j|)$. In this case, even the losing party stop using pork barrel. Notice that Lemma 1 still applies in this case as long as the winning party still promise positive pork barrel or not choosing $\beta^*_j = \theta_L$. In those situations, the equilibrium exists and payoff function is well defined. Otherwise, if in equilibrium, the winning party provides 0 pork barrel and choose $\beta^*_j$, then the equilibrium is not unique and Lemma 1 does not apply. However, the winning party’s payoff is still well-defined and equal to $Q$.

The other point is that, now, it is clear that Assumption 2 is just a sufficient condition. We can replace Assumption 2 by the following necessary and sufficient condition

**Assumption 2**: For any $x^k$, $Q - C(|\theta_j - \hat{\beta}(x^k, \theta_j)|) - c(|x^k - \hat{\beta}(x^k, \theta_j)|) + c(|x^k - \hat{\beta}(x^k, \theta_j)|) > 0$.

It is clear that Assumption 2 implies Assumption 2'.

**Appendix C: Gerrymandering with Complete Freedom by a Moderate Party Leader**

In the main text, we follow Friedman and Holden (2008) to assume extreme party leaders. In this Appendix, we relax the assumption of $F(-1) = 0$ and $F(1) = 1$ by keeping $\theta_L = -1$ and $\theta_R = 1$. That is, there are voters who are
more extreme in their political positions. Note that party L’s leader would not want to create districts of which median voters are more extreme. What he can do is the following: Let \( \hat{K} \) be the largest integer such that \( \frac{\hat{K}}{2K} \) does not exceed \( F(-1) \). We consider two cases.

**Case with** \( \gamma < 2 \). Create \( \hat{K} \) homogeneous districts with median voter \( \theta_{L}^{\hat{K}} = \theta_{L} = -1 \) by combining proportionally divided \( \frac{1}{F(-1)} \times \frac{1}{2K} + \frac{\epsilon}{K} \) of voters in interval \( (-\infty, -1) \) with proportionally divided \( \frac{1}{2K} \) of \( [0, \infty) \), where \( 1 - F(\theta_{L}^{\hat{K}}) = \frac{\hat{K}}{2K} - \frac{\epsilon}{K} \) for a small \( \epsilon > 0 \). For district \( \hat{K} + 1 \), define \( \theta_{L}^{\hat{K}+1} \) and \( \theta_{R}^{\hat{K}+1} \) by \( F(\theta_{L}^{\hat{K}+1}) = \frac{\hat{K}+1}{2K} + \frac{\hat{K}+1}{K} \epsilon \) and \( 1 - F(\theta_{R}^{\hat{K}+1}) = \frac{\hat{K}+1}{2K} - \frac{\hat{K}+1}{K} \epsilon \). District \( \hat{K} + 1 \) has leftover voters in \( (-\infty, -1) \), all voters in \( [-1, \theta_{L}^{\hat{K}+1}] \) and all voters in \( [\theta_{R}^{\hat{K}+1}, \theta_{R}^{\hat{K}}] \). We continue this process up to \( \hat{K} = \max \{ k | k^2 < 0 \} \), creating \( \hat{K} \) winning districts with \( \theta_{L} \), succeeded by the standard slice-and-mix districts. \(^{23}\) In this case, we have \( \hat{K} \) “cracked” districts in a very special way — the party leader’s ideal districts.

**Case with** \( \gamma > 2 \). By the same logic of Proposition 4, party L’s leader wants to create median voters as extreme as possible despite that party L’s leader’s position is \( \theta_{L} = -1 \), since \( \hat{V} \) is convex. □

**Appendix D: The Case of** \( \gamma = 1 \)

Here, we analyze the case of \( \gamma = 1 \), which means that both party leaders’ and voters’ costs from the distance in political positions are linear: i.e., \( C(d) = \alpha \cdot d \) and \( c(d) = \alpha' \cdot d \), where \( d \) denotes the distance. The winning party leader’s optimal choice of their candidates is

\[
\beta_{i}(x^{k}) = \begin{cases} x^{k} & \text{if } \alpha' > \alpha \\ \text{anything between } x^{k} \text{ and } \theta_{i} & \text{if } \alpha' = \alpha \\ \theta_{i} & \text{if } \alpha' < \alpha \end{cases}
\]

and the resulting cost function is

\[
C_{i}(\theta_{i}, x_{i}) = A |\theta_{i} - x_{i}| \, ,
\]

where \( A = \min \{ \alpha', \alpha \} \) is just a constant.

\(^{23}\)For the problem by Friedman and Holden (2008) when party leaders’ positions are not most extreme, we get the same optimal solution.
Thus, if party $L$ is the gerrymanderer, (assuming $\theta_R = 1$ and $\theta_L = -1$) her payoff function from winning in district $k$ is

$$V^k_L(x^k) = \begin{cases} 2A & \text{if } x^k \leq -1 \\ 2A|x^k| & \text{if } -1 \leq x^k \leq 0 \\ 0 & \text{if } 0 \leq x^k \end{cases}$$

Thus, the optimal gerrymandering policy is to create as many districts with $-1 \leq x^k \leq 0$ as possible, but otherwise, the gerrymanderer is indifferent with how to allocate $x^k$'s, otherwise. Thus, under determinacy, cases $\gamma = 1$ and $\gamma = 2$ are similar to each other. When there is uncertainty, however, “pack-and-crack” would be recommended for the case of $\gamma = 1$ as in Owen and Grofman (1988).\textsuperscript{24}

\textbf{References}


\textsuperscript{24}In order to maximize the number of $x^k + \epsilon$ in interval $[-1,0]$, $x^k$ should be concentrated at the central area of the interval.


Congress Has Polarized Since the 1960s

(Source: Figure 2.2 in “Culture War? The Myth of a Polarized America,” in Fiorina, Abrams, and Pope, 2011)

Figure 1

Scores calculated with DW-Nominate procedure developed by Keith Poole and Howard Rosenthal. (Roll call data available online at voteview.uh.edu/default_nomdata.htm)
Figure 2

The Shape Winning Payoff Function

When $\gamma \geq 0$
Figure 3

The Winning Payoff Function When Corner Solutions Occur