Legislative Informational Lobbying[1]

Arnaud Dellis
Dept. of Economics, Université du Québec à Montréal
E-mail: dellis.arnaud@uqam.ca

Version: July 2021

Whom should an interest group lobby in a legislature? I develop a model of informational lobbying in which a legislature must decide on the allocation of district-specific goods and projects. An interest group chooses to search and provide information on districts' valuations of the goods. The setting is one of distributive politics with an allocation of goods and projects that is endogeneous to the information provided by the interest group. I characterize the information search strategy of the interest group. I determine who gains and who loses from lobbying, identifying circumstances in which legislators would unanimously prefer to ban informational lobbying. I also identify two empirical and institutional implications. First, I establish a relationship between informational lobbying and legislative majority requirement. Second, I provide an informational rationale for friendly lobbying (that is, the interest group lobbying legislative allies).

Key Words: Lobbying; Interest group; Information; Persuasion; Legislature; Policy-making; Distributive politics; Majority requirement.

Subject Classification: C72; D72; D78.

1. INTRODUCTION

Two features pertain to interest group influence. One is the prevalence of legislative policymaking. Policies are chosen not by a single policymaker, but by a legislature composed of representatives elected in a number of districts.[2] The other feature is the prevalence of lobbying as an instrument of interest group influence,

[1] I thank Emiel Awad, Maria Gallego, Tom Groll, Nicolas Klein, David Miller, Clement Minaudier, Jacob Morrier, Alain Paquet, Maik Schneider, Arthur Silve and Jan Stuckatz, as well as seminar and conference participants for helpful comments and suggestions.

[2] Huneeus and Kim (2021) shows that over the period 1999-2018, almost all lobbying activities by publicly traded firms in the US were targeting the Congress.

where lobbying is defined as the act of providing information. The offices of lobbying firms are an integral fixture of capital cities in many developed countries. Multiple accounts of policymaking in the U.S. (e.g., Bauer, Dexter and De Sola Pool, 1963; Hansen, 1991; Drutman, 2015) assert that interest groups' influence often takes the form of persuasion through information provision. These accounts are substantiated by the Center for Responsive Politics (opensecrets.org), which reports that about 7 billion US dollars were spent on lobbying in the years 2019 and 2020, compared to 6.46 billion US dollars of monetary contributions. Moreover, multiple studies (e.g., Wright, 1990; Ansolabehere, Snyder and Tripathi, 2002) provide evidence suggesting that campaign contributions serve to gain access to legislators in order for interest groups to communicate their information. The prevalence of lobbying is associated with legislators' reliance on the information provided by interest groups, as noted in Hansen (1991: 5): "Lawmakers operate in highly uncertain electoral environments. They have an idea of the positions they need to take to gain reelection, but they do not know for sure. Interest groups offer to help ... They provide political intelligence about the preferences of congressional constituents."[3] Because of legislators' reliance on interest groups' information, lobbying can be an effective instrument of interest group influence (e.g., Gawande, Maloney and Montes-Rojas, 2009; Igan and Mishra, 2014; Belloc, 2015). As Baumgartner et al. (2009: 124) writes: "There is evidence that organizational advocates are often successful in getting Congress to make policy decisions that are informed by research and the technical expertise that they provide."

This paper investigates the question of which legislators an interest group should lobby.[4] This question is empirically relevant as evidence suggests that interest

[3] Since then U.S. Congressmen seem to have become even more reliant on information provision by interest groups. In an article (dated March 10th, 2015) published in *The Atlantic* and titled "Why Congress Relies on Lobbyists Instead of Thinking for Itself", Lee Drutman and Steven M. Teles write: "The federal government across all its branches has experienced a deterioration in its ability to acquire, process, and analyze information. But the problem is especially urgent in Congress." They further write: "Lacking their own capacity, government will have no other choice, but to increasingly turn to them [corporate lobbyists] for help."

[4] There is a large literature on informational lobbying – e.g., Potters and van Winden (1992), Austen-Smith and Wright (1992), Austen-Smith (1995, 1998), Lohmann (1995) and Cotton (2009). However, these contributions consider settings with a single policymaker, not a legislative assembly. There is also a smaller literature on interest groups seeking to influence a legislative assembly – e.g., Snyder (1991), Groseclose and Snyder (1996), Diermeier and Myerson (1999), Baron (2006), Dekel,

groups engage in selective persuasion, targeting some legislators while ignoring others (e.g., Bombardini and Trebbi, 2020). As You (2020: 6) writes: "... identifying the *sequence* and the *set of politicians*[5] who are targeted is crucial to understanding interest groups' lobbying strategies when groups try to persuade politicians in a collective decision making environment." To investigate this question, I construct a model of interest group influence that includes both informational lobbying and legislative policymaking. The model is set in the context of distributive politics,[6] where the legislative proposal is endogenous to the information provided by the interest group.[7]

Specifically, the model considers a legislative assembly consisting of $N+1$ legislators, each representing a different district. The legislature must decide on the allocation of district-specific goods and projects (hereafter referred to as goods) that can be local public goods or pork-barrel projects such as road construction, mass-transit projects, grants-in-aid, or recreational projects such as sports arenas and public libraries. Goods are financed by a national tax base. One legislator serves as the agenda setter, proposing an allocation of goods across the $N+1$ districts. The agenda setter can be, for instance, the chair of the Appropriations committee. Adoption of the agenda setter's proposed allocation requires the approval of at least M other legislators, where M can take value between 0 (dictatorship of the agenda setter) and N (unanimity). Each district has a valuation of the goods which, to keep things simple, is either low or high. Districts are ex ante heterogeneous, varying in their prospects of high valuation. Districts' valuations are ex ante unknown to all, but an interest group that benefits from the provision of goods can search infor-

Jackson and Wolinsky (2009), Schneider (2014), Battaglini and Patacchini (2018) and Fehrler and Schneider (2021). However, these contributions look at vote buying, not informational lobbying.

[5] Italics in the original.

[6] Distributive politics is defined as "those projects, programs, and grants that concentrate the benefits in geographically specific constituencies, while spreading their costs across all constituencies through generalized taxation." (Weingast, Shepsle and Johnsen, 1981: 643).

[7] Bennedsen and Feldmann (2002), Schnakenberg (2015, 2017) and Awad (2020) are other formal contributions studying legislative informational lobbying. In contrast with this paper, Bennedsen and Feldmann (2002) does not investigate the question of legislator targeting. Schnakenberg (2015, 2017) and Awad (2020) study this question, but, in contrast with the present paper, do so in the context of regulatory politics, where the legislative proposal is exogenous and does not respond to the information provided by interest groups. I discuss further these papers in the next section.

mation on districts' valuations. The interest group can be, for instance, the union of road builders advocating infrastructure projects (e.g., the American Association of State Highway and Transportation Officials – AASHTO), a national association for the promotion of the arts advocating state financing of local arts organizations, a sports league advocating state financing for the construction of sports arenas, or a municipal league advocating earmarks for local projects.[8] Lobbying is modelled as persuasion, where information takes the form of verifiable evidence. Search is costly and sequential; the interest group searches one district at a time, observing the outcome of its search on a district's valuation before deciding whether to search another district.[9]

In equilibrium, the agenda setter forms a legislative coalition consisting of himself and M other legislators whose districts have amongst the highest expected valuations. Districts in the legislative coalition are offered goods, those outside are not. Two features of the legislative choice are critical drivers of legislator targeting by the interest group. First, the proposed allocation of goods and the composition of the legislative coalition are endogenous to the information provided by the interest group. Second, the goods function, that specifies the total quantity of goods as a function of the (expected) valuations of districts in the legislative coalition, is strictly increasing in each of its arguments and, depending on the provision cost of goods, can be concave or convex.[10]

[8] U.S. municipalities are joining forces in 49 state municipal leagues (e.g., League of California Cities, New Jersey State League of Municipalities, Texas Municipal League) and a federal league (National League of Cities) in order to lobby Congress and state legislatures, advocating federal and state funding for city projects such as local infrastructure or local public transportation projects. As Jensen (2018: 274) writes: "lobbying is indeed valuable to members ... they rely on their associations to do the lobbying that they cannot do themselves." Shannon Najmabadi writes in an article published in *The Texas Tribune* on May 14th, 2019 and titled ""People Were Giving Us Lip Service": Texas Cities' Legislative Efforts Have Struggled This Year": "The interest group representing Texas cities used to be one of the most powerful legislative forces at the Capitol. This session, it has become the GOP's most prominent adversary." Canadian municipalities are likewise organized in municipal associations (e.g., Union of Canadian Municipalities, Towns of New Brunswick Association, Union des Municipalités du Québec) that lobby federal and provincial MPs in support of specific local public projects.

[9] I take sequential lobbying as given and investigates its implications. Sequential lobbying applies to situations where, for example, producing information can be relatively quick (e.g., providing poll results on voters' support for the realization of local public projects).

[10] The concave case is arguably more relevant than the convex case. The motivation for pre-

The interest group engages in selective persuasion, choosing strategically the districts on which it searches information.

When the provision cost of goods is such that the goods function is concave, the interest group starts by searching districts with 'moderate' prospects of high valuation, and then moves towards districts with 'extreme' prospects of high valuation, alternating between districts with better and worse prospects depending on the received information. To be more specific, let's label districts (other than the agenda setter's) in a decreasing order, with district 1 having the best prospects of high valuation and district N having the worst prospects. The interest group starts by searching district $M+1$, that is, the district with the best prospects of high valuation *among the districts that otherwise would not be included in the legislative coalition.* If the interest group obtains favorable information (that is, information of high valuation), it then moves to district M, that is, the next district with better prospects. If the interest group rather obtains unfavorable information (that is, information of low valuation), it moves instead to district $M+2$, that is, the next district with worse prospects. This process is iterated, with the interest group moving gradually towards districts 1 and N, alternating between districts with better prospects of high valuation and districts with worse prospects depending on the information it receives. The equilibrium stopping rule prescribes that the interest group continues searching until one of two things happens. Either the interest group has obtained favorable information for M districts, so that the agenda setter will form a legislative coalition consisting exclusively of known high-valuation districts. Or the interest group has obtained unfavorable information for $N-M$ districts, so that if the interest group were to search one more district, the agenda setter might have to include a known low-valuation district in the legislative coalition.

When the provision cost of goods is such that the goods function is convex, the interest group starts by searching districts with the best prospects of high valuation, and then moves gradually to districts with worse prospects. The equilibrium stopping rule prescribes the interest group to continue searching until it has obtained favorable information for M districts, so that the agenda setter will form a

senting both cases is twofold. One reason is for completeness sake. The other is that the results and the underlying intuition from the convex case are similar to those in an extended version of the model, presented in a supplementary online appendix, where districts differ in the quality of information that can be obtained by the interest group.

legislative coalition consisting exclusively of known high-valuation districts.

The analysis delivers three interesting implications. First, I identify who gains and who loses from lobbying. Informational lobbying has two implications in the context of distributive politics: (i) induce a change in the composition of the legislative coalition, and (ii) induce an increase in the total quantity of goods provided. While the first implication appears also in the context of regulatory politics, the second implication is specific to distributive politics, where the policy proposal responds to the information provided by the interest group. Both the agenda setter and the interest group gain from lobbying since they benefit from the increase in the total quantity of goods provided. The legislators from districts 1 through M lose from lobbying, since their district may no longer be included in the legislative coalition and, when this happens, their payoff drops further with the increase in the quantity of goods provided. The legislators from districts $M + 1$ through N may gain or lose from lobbying, since they benefit from the possibility that their district is now included in the legislative coalition, but lose from the increase in the quantity of goods when their district is kept out of the legislative coalition. The analysis identifies circumstances in which all legislators (other than the agenda setter) lose from the possibility of lobbying and, therefore, would unanimously prefer to ban informational lobbying.

Second, I relate legislator targeting to the nature of lobbying. Empirical studies (see de Figueiredo and Richter, 2014) find evidence that interest groups lobby both legislative allies (friendly lobbying) and opponents (confrontational lobbying), and that friendly lobbying tends to prevail over confrontational lobbying. While confrontational lobbying makes sense in terms of persuasion, friendly lobbying is more difficult to rationalize. The present analysis shows that both friendly and confrontational lobbying can be rationalized in the context of distributive politics where the legislative proposal is endogenous to the information provided by interest groups. The analysis further identifies conditions under which friendly lobbying prevails over confrontational lobbying.

Third, I relate information provision to the legislative majority requirement M. When the goods function is concave, the relationship between legislative majority requirement and expected number of searched districts is single-peaked, with the location of the peak depending on districts' prospects of high valuation. When the goods function is convex, the (expected) number of districts for which the interest group provides information increases with the majority requirement, reaching a maximum when the adoption of a proposal requires unanimity ($M = N$). In either case, this yields an interesting institutional implication: while Diermeier and Myerson (1999) suggests that legislators in a unicameral legislature may want to adopt infra-majority requirements (by delegating authority to a leader) if they seek to maximize the expected amount of *monetary contributions* they receive, my analysis suggests that legislators may be better off adopting simple- or super-majority requirements if they seek to maximize the expected amount of *information* they receive.

While the present analysis is developed in the context of lobbying, it applies more generally to collective decisionmaking, with the interest group acting as the information sender and the legislators as the receivers. Institutions with collective decisionmaking to which the model could be applied are boards of directors or shareholders in firms, boards of governors in professional sports leagues (e.g., NHL), or coalition governments.

In a supplementary online appendix, I present two extensions of the model. In one extension the quality of the information that the interest group can obtain varies across districts. In the other extension each legislator can decide to search information on her district's valuation, following the interest group's search.

The remainder of the paper is organized as follows. Section 2 discusses the related literature. Section 3 describes the model. Section 4 characterizes the legislative choice. Section 5 analyses the interest group's search. Section 6 discusses three key implications from the analysis. Section 7 concludes. All proofs are in the appendix. A supplementary online appendix contains additional material.

2. RELATED LITERATURE

The contribution of this paper is to study selective persuasion by an interest group in the context of legislative policymaking. In this section, I relate my paper to the most relevant existing contributions in the literature.

This paper contributes to the literatures on legislative lobbying and group persuasion. Bennedsen and Feldmann (2002), Schnakenberg (2015, 2017) and Awad (2020) all study legislative informational lobbying.

As in Bennedsen and Feldmann (2002), I consider a setting of distributive politics, involving a decision on allocating district-specific goods, in which an interest group seeks to persuade legislators by producing verifiable evidence on districts' valuations of their goods. However, Bennedsen and Feldmann do not study legislator targeting, as I do here, but rather study how the vote of confidence procedure affects the incentives to lobby. Accordingly, they assume ex ante homogeneous districts. To study legislator targeting, I relax this assumption and consider ex ante heterogeneous districts.

Schnakenberg (2015, 2017) and Awad (2020) consider a setting of regulatory politics in which an interest group seeks to induce legislators to vote in favor of a policy proposal. Schnakenberg (2015, 2017) use this setup to demonstrate that, in situations of collective choice instability, cheap talk messaging can affect negatively the ex ante expected utilities of all legislators. Schnakenberg (2017) and Awad (2020) use this setting to rationalize friendly lobbying, showing that an interest group may choose to enroll legislative allies as intermediaries who help persuade opposed legislators. In Schnakenberg (2017), friendly lobbying arises because of costly access to legislators and the possibility for legislators to lobby their colleagues (indirect lobbying). In Awad (2020), friendly lobbying arises because of the correlation between legislators' preferences, which allows the emergence of persuasion cascades. My paper differs from these in several key ways. First, while these papers consider a setting of regulatory politics with an exogenous binary policy proposal, I consider a setting of distributive politics with an allocation proposal that is endogenous to the information provided by the interest group. Second, the rationales for friendly lobbying in Schnakenberg (2017) and Awad (2020) are not present in my analysis since access to legislators can be seen as costless and districts' valuations are drawn independently, thereby precluding persuasion cascades. Third, while Schnakenberg (2015, 2017) consider a privately informed interest group sending cheap talk messages, I consider an uninformed interest group producing verifiable evidence.

Caillaud and Tirole (2007) and Alonso and Câmara (2016) are two key contributions in the literature on group persuasion, in which a sender seeks to persuade multiple receivers who have to make a collective decision. Caillaud and Tirole (2007) studies the emergence of persuasion cascades which, as noted above, cannot arise in my model. Alonso and Câmara (2016) shows how Bayesian persuasion by an uninformed politician can reduce the welfare of a majority of voters. My paper differs from these papers in several ways, notably in that they consider a group deciding on an exogeneous binary proposal, while I consider a group deciding on a proposal that is endogenous to the information provided by a sender.

The present paper is also related to the literature on Weitzman's (1979) Pandora box problem. In the seminal Pandora box problem, an agent is presented with a finite set of boxes. Each box contains a prize. Boxes are ex ante heterogeneous, differing in their probability distributions over the value of the prize contained in the box. The agent can open boxes to reveal the prizes they contain. Search is sequential and costly. At the end of the search process, the agent chooses one of the opened boxes.[11] Legislative lobbying exhibits several features of the Pandora box problem. Districts can be interpreted as boxes, the interest group as the agent, and districts' valuations as prizes. At the same time, the legislative lobbying problem differs in multiple ways from the seminal Pandora box problem. First, Weitzman's Pandora box problem applies to an individual choice problem, with the agent selecting one box at the end of the search process. By contrast, the legislative lobbying problem applies to a collective choice problem, where a legislature decides on an allocation of goods that involves selecting multiple districts. Second, in the Pandora box problem, the agent is responsible for both the search and the selection of a box. By contrast, in the legislative lobbying problem, the interest group is responsible for the search, while the legislators are responsible for the choice of allocation. Third, in Weitzman's Pandora box problem, the order of the search is history independent, which allows for an index characterization (related to Gittins index for bandit processes). By contrast, the order of the search is history dependent in the legislative lobbying problem.

3. MODEL

Consider a country which is divided into a finite number $N+1$ of districts, with $N \in \mathbb{N}$. I denote the set of districts by $\mathcal{N}_0 = \{0, 1, ..., N\}$, with typical element n. Each district is represented by a legislator. The legislature decides on the allocation

[11] Weitzman's setup has been generalized (i) to allow the agent's payoff to depend not just on the prize in the chosen box but on all uncovered prizes (Olszewski and Weber, 2015), and (ii) to allow the agent to choose an unopened box (Doval, 2018).

of district-specific goods. Let $g_n \in \mathbb{R}_+$ be the quantity of good provided to district n. I denote the allocation of goods across districts by $\mathbf{g} = (g_0, g_1, ..., g_N)$. The total quantity of goods provided in the country is equal to $G = \sum_{n=0}^{N} g_n$. It costs $c(G)$ to provide a quantity of goods G, where $c(\cdot)$ is a differentiable, strictly increasing and strictly convex function. I let $c(0) = \lim_{G\downarrow 0} c'(G) = 0$. The cost of providing goods is financed by a national tax base, and is divided equally across districts, that is, each district bears a cost $c(G)/(N+1)$.[12]

Each district n has a valuation for the good r_n, which can take on two values: $r_n = \underline{r}$ (low valuation) or $r_n = \overline{r}$ (high valuation), where $\overline{r} > \underline{r} > 0$. District $n \in \mathcal{N}_0$ has a high valuation, $r_n = \overline{r}$, with commonly-known probability $p_n \in (0, 1)$. Districts valuations are drawn independently and are ex ante unknown to all. I denote district n's ex ante expected valuation by $Er_n = p_n\overline{r} + (1 - p_n)\underline{r}$. Following Bennedsen and Feldmann (2002), the legislator from district n gets a payoff $u_n(\mathbf{g}) = r_n g_n - \frac{c(G)}{N+1}$ from allocation $\mathbf{g} = (g_0, g_1, ..., g_N)$.

The legislator from district 0 is the proposer or agenda setter (henceforth, AS). I denote the set of legislators other than AS by $\mathcal{N} = \{1, ..., N\}$. The legislature operates under a closed rule, with AS proposing an allocation $\mathbf{g} = (g_0, g_1, ..., g_N)$ and the legislators voting for or against the proposal. The proposal is adopted, and each district $n \in \mathcal{N}_0$ receives the proposed quantity of good g_n, if at least $M \in \{0, ..., N\}$ legislators in $\mathcal{N}$ vote in favor of the proposal. Otherwise, the proposal is defeated and $g_n = 0$ for each district n. In order to study the effect of majority requirements on informational lobbying, I let M take value between $M = 0$, which corresponds to a dictatorship of AS, and $M = N$, where unanimity is required to pass a proposal. Simple majority corresponds to $M = \lfloor N/2 \rfloor$.

An interest group (henceforth, IG) benefits from the provision of goods. IG can produce information on districts' valuations. If IG searches district n, it receives a signal $\sigma_n = r_n \in \{\underline{r}, \overline{r}\}$ that reveals district n's valuation.[13] Signals are publicly observed.[14] IG gets a payoff $v(\mathbf{g}, I) = G - I\varepsilon$ when it searches I districts and the

[12] One can interpret g_n as a grant to district n. In this case, G is the total amount of grants, and $c(G)$ is equal to the total amount of grants plus, for example, the deadweight loss of taxation.

[13] In a supplementary online appendix, I allow for signals to be partially informative and for signal informativeness to vary across districts. Specifically, I assume signal σ_n reveals district n's valuation ($\sigma_n = r_n \in \{\underline{r}, \overline{r}\}$) with probability $q_n \in (0, 1]$, and conveys no information ($\sigma_n = \emptyset$) with probability $1 - q_n$.

[14] Results would be qualitatively unchanged if signals were private to IG, who could then choose

resulting allocation of goods is $\mathbf{g}$, where $\varepsilon > 0$ is the cost of searching a district. To simplify the analysis, I assume that ε is sufficiently small so that IG is willing to search a district whenever it expects a favorable signal will trigger an increase in the total quantity of goods.[15]

The policymaking process has four stages. At stage 0, Nature chooses districts' valuations. Realized valuations are unknown to IG and the legislators. At stage 1, IG chooses on information collection. It proceeds sequentially, observing the signal just obtained before deciding whether to search yet another district. The search process stops when IG chooses to not search one more district. The game then moves to stage 2, where AS proposes an allocation $\mathbf{g}$. At stage 3, legislators vote on AS' proposal.

Legislator n's (pure) voting strategy on AS' proposed allocation is $\upsilon_n : \mathbb{R}_+^{N+1} \times \{\underline{r}, \overline{r}, \emptyset\} \rightarrow \{0,1\}$, where $\upsilon_n(\mathbf{g}, \sigma_n) = 1$ if legislator n votes in favor of proposal $\mathbf{g}$ and $\upsilon_n(\mathbf{g}, \sigma_n) = 0$ if she votes against the proposal, where $\sigma_n = r_n$ if IG searched district n and $\sigma_n = \emptyset$ otherwise.

AS' (pure) proposal strategy is $\gamma : \{\underline{r}, \overline{r}, \emptyset\}^{N+1} \rightarrow \mathbb{R}_+^{N+1}$, where $\gamma(\sigma_0, \sigma_1, ..., \sigma_N) = \mathbf{g}$ is the allocation AS proposes given a signal profile $(\sigma_0, \sigma_1, ..., \sigma_N)$ on $(r_0, r_1, ..., r_N)$.

IG's (pure) search strategy specifies for each round $t = 1, ..., N+1$ of the search process a search decision as a function of the signals received in the previous rounds of the search process. I denote round-t search history by h^t, where $h^1 = \emptyset$ and $h^t = ((s^\tau, \sigma^\tau))_{\tau=1}^{t-1}$ for $t = 2, ..., N+1$, where $s^\tau = n \in \mathcal{N}_0$ and $\sigma^\tau = r_n \in \{\underline{r}, \overline{r}\}$ if IG searched district n at round τ, and $s^\tau = \sigma^\tau = \emptyset$ if IG did not search at round τ. IG's search strategy, s, is a sequence of functions $\{s^t(\cdot)\}_{t=1}^{N+1}$, where $s^t(h^t)$ is IG's search decision at round t given history h^t. At each round t we have $s^t(h^t) \in (\mathcal{N}_0 \backslash \mathcal{I}^t(h^t)) \cup \{\emptyset\}$, where $\mathcal{I}^t(h^t) \equiv \{n \in \mathcal{N}_0 : s^\tau = n \text{ for some } \tau \leq t-1\}$ is the set of districts that have been searched by round t. Since the search process stops when IG chooses to not search at one round, we have that $s^t(h^t) = \emptyset$ implies $s^{t+1}\left(h^{t+1}\right) = \emptyset$.

Beliefs are derived using Bayes' rule. Let $E\widetilde{r}_n$ be district n's posterior expected

whether to reveal a signal or not. Likewise, results would remain unchanged if instead of being publicly observed, a signal on district n would be observed only by IG, AS and the legislator from district n.

[15] There is also an endogenous implicit cost of searching a district that is associated with the effect of a search on the composition of the legislative coalition. The assumption of a small explicit cost of searching ε allows me to focus on this implicit cost of searching.

valuation, where $E\widetilde{r}_n = r_n$ if IG searched district n and $E\widetilde{r}_n = Er_n$ otherwise.

The solution concept is (pure-strategy) Perfect Bayesian equilibrium, with the standard refinement of weakly undominated voting strategies for the legislators. Throughout the analysis, I order districts such that $p_1 > p_2 > ... > p_N$.[16]

4. LEGISLATIVE PROCESS

To characterize equilibria, I proceed backwards, analyzing each stage of the game in reverse order. I first consider the legislative process.

I start by characterizing legislators' voting strategies at Stage 3 of the game. Suppose AS has proposed an allocation $\mathbf{g}' = (g'_0, g'_1, ..., g'_N)$. A legislator votes in favor of the proposal if she is better off with the proposal being adopted than with the proposal being rejected. Thus, legislator $n \in \mathcal{N}$ votes in favor of the proposal $(\upsilon_n(\mathbf{g}', \sigma_n) = 1)$ when

$$E\widetilde{r}_n \; g'_n - \frac{c(G')}{N+1} \geq 0. \tag{1}$$

Let $\mathcal{L}(\mathbf{g}') \subseteq \mathcal{N}$ denote the set of districts (other than AS' district) which legislators vote in favor of proposal $\mathbf{g}'$. I call $\mathcal{L}(\mathbf{g}')$ the legislative coalition associated with proposal $\mathbf{g}'$. Proposal $\mathbf{g}'$ is adopted if and only if $\#\mathcal{L}(\mathbf{g}') \geq M$.

I now move backwards and characterize AS' proposal strategy at Stage 2 of the game. AS proposes an allocation $\mathbf{g}$ that maximizes his expected utility $E\widetilde{r}_0 g_0 - c(G)/(N+1)$ subject to the constraint that $\#\mathcal{L}(\mathbf{g}) \geq M$.

Since the provision of goods is costly, AS seeks to provide districts other than his own with as few goods as possible. This has two implications for AS' equilibrium proposal $\mathbf{g}^*$. First, AS offers $g^*_n = 0$ to any $n \notin \mathcal{L}(\mathbf{g}^*)$, that is, AS does not offer goods to districts outside the legislative coalition. Second, AS offers $g^*_n = c(G^*)/(N+1)E\widetilde{r}_n$ to any $n \in \mathcal{L}(\mathbf{g}^*)$, that is, AS offers to any district in the legislative coalition a quantity of good that binds the participation constraint of the district's legislator (equation (1)).

[16] Considering a strict, rather than weak, ordering of districts is made to simplify exposition; it avoids a multiplicity of equilibrium search sequences that would result from IG being indifferent between searching two a priori identical districts. Also, I do not include AS's district (district 0) in the ordering of districts since, as we shall see, p_0 does not matter for the equilibrium characterization.

Given these two implications, AS' problem can be written as

$$\begin{aligned} \max_{\mathbf{g}\in\mathbb{R}_+^{N+1}} \quad & E\widetilde{r}_0 \; g_0 - \tfrac{c(G)}{N+1} \\ \text{s.t.} \quad & (i): \#\mathcal{L}(\mathbf{g}) \geq M, \text{ and} \\ & (ii): g_n = \begin{cases} \frac{c(G)}{(N+1)\; E\widetilde{r}_n} & \text{if } n \in \mathcal{L}(\mathbf{g}) \\ 0 & \text{if } n \notin \mathcal{L}(\mathbf{g}) \end{cases} \quad \text{for any } n \in \mathcal{N}\,. \end{aligned}$$

Observe that AS' payoff decreases with the quantity of goods offered to districts in $\mathcal{N}$, and that this quantity increases with the number of districts included in the legislative coalition. Furthermore, the quantity of good g_n offered to a district $n \in \mathcal{L}(\mathbf{g})$ decreases with the district's expected valuation, $E\widetilde{r}_n$. As a result, AS forms a legislative coalition that consists of exactly M districts whose expected valuations are among the highest ones.

Given that $g_0 = G - \sum_{n\in\mathcal{N}} g_n$, AS chooses a total quantity of goods

$$G^* \in \arg\max_{G\in\mathbb{R}_+} E\widetilde{r}_0 \cdot \left[G - \frac{c(G)}{N+1} \sum_{n\in\mathcal{L}(\mathbf{g})\cup\{0\}} \frac{1}{E\widetilde{r}_n} \right].$$

The solution to this problem is given by

$$G^* = c'^{-1}\left(\frac{N+1}{\sum_{n\in\mathcal{L}(\mathbf{g}^*)\cup\{0\}} (1/E\widetilde{r}_n)} \right). \tag{2}$$

Relabelling districts in $\mathcal{N}$ such that $E\widetilde{r}_1 \geq ... \geq E\widetilde{r}_N$, we get that AS proposes an allocation of goods ($\gamma(\sigma_0, \sigma_1, ..., \sigma_N) = \mathbf{g}^*$) such that

$$\begin{aligned} g_n^* &= \begin{cases} \frac{c(G^*)}{(N+1)\; E\widetilde{r}_n} & \text{if } n \in \mathcal{L}(\mathbf{g}^*) \\ 0 & \text{if } n \notin \mathcal{L}(\mathbf{g}^*) \end{cases} \quad \text{for any } n \in \mathcal{N} \\ g_0^* &= G^* - \sum_{n\in\mathcal{N}} g_n^*, \end{aligned}$$

where $\mathcal{L}(\mathbf{g}^*) = \{1, ..., M\}$ and $G^* = \Gamma(E\widetilde{r}_0, E\widetilde{r}_1, ..., E\widetilde{r}_M)$ is given in (2).[17]

[17] AS keeping each district in the legislative coalition at its reservation utility ($g_n^* = \frac{c(G^*)}{(N+1)\; E\widetilde{r}_n}$ for all $n \in \mathcal{L}(\mathbf{g}^*)$) is a consequence of the take-it-or-leave-it protocol, that gives all the bargaining power to AS. Following Bennedsen and Feldmann (2002) I have adopted this simple bargaining protocol in order to focus on informational lobbying. Other bargaining protocols would give some bargaining power to the legislators in $\mathcal{N}$. While the total quantity of goods may then be different, the main results from the analysis would be qualitatively similar if AS still seeks to form a legislative coalition consisting of districts with higher (expected) valuations and the total quantity of goods increases with the (expected) valuations of the districts in the legislative coalition.

Two properties of the goods function $\Gamma(\cdot)$ are key for IG's search decision. First, the goods function $\Gamma(\cdot)$ is strictly increasing in $E\widetilde{r}_n$ for each $n \in \mathcal{L}(\mathbf{g}^*) \cup \{0\}$, that is, the total quantity of goods increases with the expected valuation of each of the districts in the legislative coalition (including AS's district). Second, the goods function $\Gamma(\cdot)$ can be concave or convex in each of its arguments, depending on the third derivative of the cost function $c(\cdot)$. I shall distinguish in the analysis the case where $\Gamma(\cdot)$ is strictly concave in each of its arguments (hereafter referred to as the concave case) from the case where $\Gamma(\cdot)$ is strictly convex (hereafter referred to as the convex case).[18]

5. INTEREST GROUP'S SEARCH PROCESS

In this section, I analyze IG's search process, that is, Stage 1 of the game. I characterize first IG's search decision on AS's district, then the equilibrium stopping rule that specifies when IG stops searching and, finally, the equilibrium search sequence that specifies the order in which IG searches districts.

5.1. Searching AS's district

AS's district (district 0) is always part of the legislative coalition. As a result, IG never (resp. always) searches district 0 in the concave case (resp. convex case)

[18] To see this, observe that for a cost function $c(G) = G^{\beta}/\beta$ for $\beta > 1$, equation (2) writes as

$$G^* = \left[\frac{N+1}{\sum_{n \in \mathcal{L}(\mathbf{g}^*) \cup \{0\}} (1/E\widetilde{r}_n)} \right]^{1/(\beta-1)}.$$

We then get

$$\frac{\partial^2 \Gamma}{\partial E\widetilde{r}_i^2} = H \cdot K$$

where

$$H = \frac{G^*}{\beta - 1} \frac{1}{(E\widetilde{r}_i)^4} \frac{1}{\left[\sum_{n \in \mathcal{L}(\mathbf{g}^*) \cup \{0\}} (1/E\widetilde{r}_n)\right]^2}$$

and

$$K = \left(\frac{2-\beta}{\beta-1}\right) - 2 \sum_{n \in (\mathcal{L}(\mathbf{g}^*) \cup \{0\}) \setminus \{i\}} \left(\frac{E\widetilde{r}_i}{E\widetilde{r}_n}\right).$$

Observe that $H > 0$ for every $\beta > 1$. At the same time, K is strictly decreasing and continuous in β, with $\lim_{\beta \downarrow 1} K = +\infty$ and $K < 0$ for $\beta = 2$. Hence $K > 0$ (resp. $K < 0$) and, therefore $\Gamma(\cdot)$ strictly convex (resp. concave) in each of its arguments, when β is close to 1 (resp. $\beta \geq 2$). Bennedsen and Feldmann (2002) considers the concave case where $\beta = 2$. The concave case is arguably more interesting than the convex case as the range of values for β to which the concave case applies is much larger than the range of values to which the convex case applies.

since it would yield a lower (resp. higher) expected quantity of goods.[19]

From now on, I shall focus on IG's search among districts in $\mathcal{N}$, that is, districts other than AS's. I can then limit the number of search rounds to N, assuming, without loss of generality, that in the convex case IG searches AS's district at a preliminary round.

5.2. Equilibrium stopping rule

I now characterize the equilibrium stopping rule, which specifies when IG stops searching. I first introduce extra notation. Given round-t search history h^t, let

$$i^{t+}\left(h^t\right) \equiv \begin{cases} \#\left\{\tau \in \{1,...,t-1\} : \sigma^\tau = \overline{r}\right\} & \text{for } t = 2,...,N \\ 0 & \text{for } t = 1 \end{cases}$$

be the number of favorable signals received prior to round t. Likewise, let

$$i^{t-}\left(h^t\right) \equiv \begin{cases} \#\left\{\tau \in \{1,...,t-1\} : \sigma^\tau = \underline{r}\right\} & \text{for } t = 2,...,N \\ 0 & \text{for } t = 1 \end{cases}$$

be the number of unfavorable signals received prior to round t.

Proposition 1. *Consider round $t \in \{1,...,N\}$ and search history h^t.*

1. *In the concave case, we have that*

$$s^t\left(h^t\right) = \emptyset \textit{ if and only if } i^{t+}\left(h^t\right) \geq M \textit{ or } i^{t-}\left(h^t\right) \geq N - M.$$

2. *In the convex case, we have that*

$$s^t\left(h^t\right) = \emptyset \textit{ if and only if } i^{t+}\left(h^t\right) \geq M.$$

Thus, in the concave case IG stops searching if and only if either the search process has already produced enough favorable signals for AS to form a legislative

[19] Empirical studies find evidence that committee chairs in the U.S. Congress are more likely to be lobbied than rank-and-file members (e.g., Hojnacki and Kimball, 1998; You, 2020). This observation is consistent with the convex case. There are several ways to reconcile this observation with the concave case. First, IG's search in my model can be interpreted as a form of legislative subsidy with IG lobbying AS by providing him information about districts' valuations in order to help AS form a legislative coalition. Second, committee chairs' control over the agenda can explain why they are lobbied more often than rank-and-file members. This rationale for lobbying AS is absent from my model where the agenda is a singleton (namely, the decision on an allocation of goods). Third, one may consider AS not as a congressional committee chair but as a member of the executive branch of governement (e.g., cabinet minister or secretary).

coalition composed only of known high valuation districts ($i^{t+}(h^t) \geq M$), or the search process has already produced enough unfavorable signals such that one more unfavorable signal would force AS to include a known low valuation district in the legislative coalition ($i^{t-}(h^t) \geq N - M$).

The intuition underlying the *sufficiency* of these conditions runs as follows. As soon as IG receives M favorable signals ($i^{t+}(h^t) = M$), it no longer wants to search since searching is costly and the total quantity of goods would remain unchanged at $\Gamma(Er_0, \overline{r}, ..., \overline{r})$. Likewise, as soon as IG receives $N - M$ unfavorable signals ($i^{t-}(h^t) = N - M$), it stops searching since continuing its search would yield a lower expected total quantity of goods due to the strict concavity of $\Gamma(\cdot)$ in each of its arguments.

The intuition underlying the *necessity* of the two conditions runs as follows. Consider round t at which IG stops searching. Suppose the search history h^t is such that $i^{t+}(h^t) \leq M - 1$ and $i^{t-}(h^t) \leq N - M - 1$. Pick a yet unsearched district n that, if search were to stop at round t, would not be included in the legislative coalition. Suppose IG deviates, searching district n at round t and then stopping its search at round $t + 1$. If the signal on r_n is favorable, AS will include district n in the legislative coalition in lieu of an unsearched district j. This will yield an increase in the total quantity of goods given that (i) $E\widetilde{r}_n = \overline{r}$ will replace $E\widetilde{r}_j = Er_j < \overline{r}$ in $\Gamma(\cdot)$ and (ii) $\Gamma(\cdot)$ is strictly increasing in each of its arguments. If instead IG receives an unfavorable signal on r_n, AS will keep district n outside the legislative coalition, and the total quantity of goods will be the same as if IG had stopped searching at round t. Hence, the expected total quantity of goods will be strictly bigger than if IG had stopped searching at round t. IG is then better off deviating.

In the convex case IG stops searching if and only if the search process has already produced enough favorable signals for AS to form a legislative coalition composed only of known high valuation districts ($i^{t+}(h^t) \geq M$). The intuition underlying this condition is the same as in the concave case. To understand why the condition on the number of unfavorable signals ($i^{t-}(h^t) \geq N - M$) does not apply in the convex case, consider a round t at which $i^{t+}(h^t) \leq M - 1$ and $i^{t-}(h^t) = N - M$. If IG were to stop searching at round t, then the legislative coalition would consist of all M districts for which no unfavorable signal has been received. Pick a yet unsearched district n (which would then be included in the legislative coalition if IG were to stop searching at round t), and suppose that IG searches district n at round t and then stops searching at round $t+1$. With probability p_n, IG will receive a favorable signal ($r_n = \overline{r}$), in which case $\overline{r}$ will replace Er_n in the goods function $\Gamma(\cdot)$. With probability $(1 - p_n)$, IG will receive an unfavorable signal ($r_n = \underline{r}$), in which case $\underline{r}$ will replace Er_n in the goods function $\Gamma(\cdot)$. Given the strict convexity of $\Gamma(\cdot)$ in each of its arguments, the expected total quantity of goods will then be strictly bigger than if IG had stopped searching at round t, meaning IG is better off continuing searching at round t.

5.3. Equilibrium search sequence

I now characterize the order in which IG searches districts in $\mathcal{N}$.

5.3.1. The concave case

The following proposition characterizes the equilibrium search sequence for the concave case.

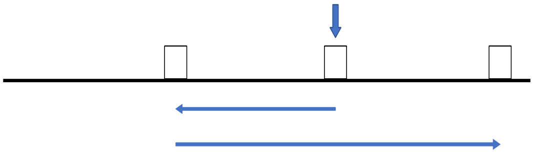
PROPOSITION 2. *Consider the concave case. At round $t \in \{1, ..., N\}$ and given search history h^t with $i^{t+}(h^t) < M$ and $i^{t-}(h^t) < N - M$, we have*

$$s^t\left(h^t\right) = \begin{cases} M + 1 - i^{t+}\left(h^t\right) & \text{if } \sigma^{t-1} = \overline{r} \text{ or } t = 1 \\ M + 1 + i^{t-}\left(h^t\right) & \text{if } \sigma^{t-1} = \underline{r}. \end{cases}$$

Thus, IG starts by searching district $M + 1$. After a favorable signal, IG moves to the closest unsearched district with higher probability of high valuation. After an unfavorable signal, IG moves to the closest unsearched district with lower probability of high valuation.

The following example illustrates the equilibrium search strategy.

EXAMPLE 1. Consider a country with five districts ($\mathcal{N}_0 = \{0, 1, 2, 3, 4\}$) where the legislature takes its decision by simple majority ($M = 2$, implying $N - M = 2$). Suppose the realized profile of valuations for districts in $\mathcal{N} = \{1, 2, 3, 4\}$ is $(\overline{r}, \underline{r}, \overline{r}, \underline{r})$, that is, districts 1 and 3 have high valuations, while districts 2 and 4 have low valuations. Following Propositions 1 and 2, IG starts with district 3 ($= M + 1$). It receives a favorable signal, and then searches district 2. This time, IG receives an unfavorable signal, and then searches district 4. IG receives a second unfavorable signal, and then stops searching ($i^{4-}\left(h^4\right) = 2$). AS forms a legislative coalition $\mathcal{L}^* = \{1, 3\}$, and the total quantity of goods is equal to $G^* = \Gamma\left(Er_0, Er_1, \overline{r}\right)$. $\square$



The intuition underlying the equilibrium search sequence runs as follows. On the one hand, the costly search implies IG wants to minimize the number of districts it searches (while maximizing the prospects of receiving favorable signals). This induces IG to search districts with the highest probabilities of high valuation (Lemma 2 in the Appendix). On the other hand, the concavity of $\Gamma\left(\cdot\right)$ implies IG wants to keep unsearched those districts with the highest expected valuations, and thus the highest probabilities of high valuation (Lemma 1 in the Appendix). This is because AS will choose to include those districts in the legislative coalition in the event IG would receive $N - M$ unfavorable signals. Taken together, these two effects imply that at each round t of the search process, IG searches the district for which the probability of receiving a favorable signal is the highest while keeping unsearched the $M - i^{t+}\left(h^t\right)$ districts with highest expected valuations (which AS will include in the legislative coalition in the event IG will not receive any favorable signal from round t on[20]).

The following example illustrates this intuition.

Example 2. Consider a country with four districts ($\mathcal{N}_0 = \{0, 1, 2, 3\}$) and a legislative majority requirement $M = 1$. We know that IG searches until it has received either one favorable signal ($M = 1$) or two unfavorable signals ($N - M = 2$). I am going to show that IG starts its search with district 2 and, conditional on receiving an unfavorable signal, continues its search with district 3.

Let's start with round 2 of the search process. Suppose IG had searched district 2 at round 1 and received an unfavorable signal. If IG searches district 1 at round 2, it will receive a favorable signal with probability p_1, in which case AS will form a legislative coalition $\mathcal{L}^* = \{1\}$ and the total quantity of goods will be equal to $\Gamma\left(Er_0, \overline{r}\right)$. With probability $1 - p_1$ IG will receive an unfavorable signal, in which

[20] If this event occurs, AS will form a legislative coalition consisting of the $i^{t+}\left(h^t\right)$ districts with known high valuation and the $M - i^{t+}\left(h^t\right)$ unsearched districts.

case $\mathcal{L}^* = \{3\}$ and the total quantity of goods will be equal to $\Gamma(Er_0, Er_3)$. Thus, the expected total quantity of goods at round 2 if IG searches district 1 is given by

$$EG_{1|2} = p_1 \cdot \Gamma(Er_0, \overline{r}) + (1 - p_1) \cdot \Gamma(Er_0, Er_3).$$

Likewise, the expected total quantity of goods at round 2 if IG searches district 3, instead of district 1, is given by

$$EG_{3|2} = p_3 \cdot \Gamma(Er_0, \overline{r}) + (1 - p_3) \cdot \Gamma(Er_0, Er_1).$$

Given the strict concavity of $\Gamma(Er_0, \cdot)$ and $Er_1 = \frac{1-p_1}{1-p_3} Er_3 + \frac{p_1-p_3}{1-p_3}\overline{r}$, we get $EG_{3|2} > EG_{1|2}$, that is, the concavity of the goods function implies IG is better off searching district 3 at round 2 and keeping district 1 unsearched.[21]

Let's now move to round 1 of the search process. On the one hand, the concavity of the goods function implies IG is better off starting its search with district 2 than with district 1.[22] On the other hand, the costly search implies IG is better off starting the search process with district 2 than with district 3. To see this, suppose IG searches district 3 at round 1. With probability p_3, IG will receive a favorable signal, in which case AS will form legislative coalition $\mathcal{L}^* = \{3\}$ and the total quantity of goods will be $\Gamma(Er_0, \overline{r})$. With probability $1 - p_3$, IG will receive an unfavorable signal and then search district 2 at round 2. The expected total quantity of goods is then given by

$$EG_3 = p_3 \cdot \Gamma(Er_0, \overline{r}) + (1 - p_3) \cdot [p_2 \cdot \Gamma(Er_0, \overline{r}) + (1 - p_2) \cdot \Gamma(Er_0, Er_1)].$$

The expected total quantity of goods is the same whether IG starts its search process with district 2 or with district 3, that is, $EG_2 = EG_3$ (see EG_2 in footnote

[21] Likewise, the concavity of the goods function implies that if IG seaches district 1 (resp. district 3) at round 1 and receives an unfavorable signal, IG will be better off searching district 3 (resp. district 2) at round 2.

[22] If IG searches district 2 at round 1, it receives a favorable signal with probability p_2, in which case AS forms a legislative coalition $\mathcal{L}^* = \{2\}$ and the total quantity of goods is equal to $\Gamma(Er_0, \overline{r})$. With probability $1 - p_2$, IG receives an unfavorable signal and the expected total quantity of goods is given by $EG_{3|2}$. Thus, the expected total quantity of goods if IG starts its search process with district 2 is given by

$$EG_2 = p_2 \cdot \Gamma(Er_0, \overline{r}) + (1 - p_2) \cdot [p_3 \cdot \Gamma(Er_0, \overline{r}) + (1 - p_3) \cdot \Gamma(Er_0, Er_1)].$$

If instead IG searches district 1 at round 1, the expected total quantity of goods is given by

$$EG_1 = p_1 \cdot \Gamma(Er_0, \overline{r}) + (1 - p_1) \cdot [p_3 \cdot \Gamma(Er_0, \overline{r}) + (1 - p_3) \cdot \Gamma(Er_0, Er_2)].$$

The strict concavity of $\Gamma(Er_0, \cdot)$ and $Er_1 = \frac{1-p_1}{1-p_2} Er_2 + \frac{p_1-p_2}{1-p_2}\overline{r}$ imply $EG_2 > EG_1$.

22). At the same time, $p_2 > p_3$ implies IG is more likely to stop its search after only one round (instead of two) if it starts its search process with district 2 than with district 3. Hence, the costly search (together with $EG_2 = EG_3$) implies that at round 1, IG is better off searching district 2 than district 3. $\square$

5.3.2. The convex case

I now characterize the equilibrium search sequence for the convex case. I start by introducing extra notation. Given round-t search history h^t, let $\mathcal{C}^t(h^t) \equiv \mathcal{N} \backslash \mathcal{I}^t(h^t)$ be the set of districts in $\mathcal{N}$ still unsearched at round t. I relabel districts in $\mathcal{C}^t(h^t)$ such that $\mathcal{C}^t(h^t) = \{1, ..., N+1-t\}$ with $p_1 > ... > p_{N+1-t}$. Also, let $K^t \equiv \min\{M - i^{t+}(h^t), N+1-t\}$ be the minimum number of districts that IG will search from round t on. This number is equal to the minimum of: (i) $M - i^{t+}(h^t)$, that is, the number of additional favorable signals necessary for the condition of the equilibrium stopping rule to be met; and (ii) $N+1-t$, that is, the number of still unsearched districts in $\mathcal{N}$ $(= \#\mathcal{C}^t(h^t))$.

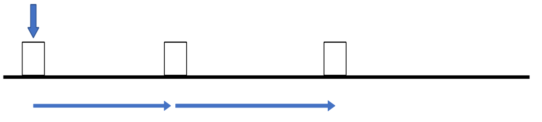
Proposition 3. *Consider the convex case. At round $t \in \{1, ..., N\}$ and given search history h^t with $K^t \geq 1$, we have*

$$s^t\left(h^t\right) \in \left\{1, ..., K^t\right\} \subseteq \mathcal{C}^t\left(h^t\right).$$

Thus, at each round t IG is indifferent searching any of the still unsearched districts among those with the K^t best prospects of high valuation. This means that IG starts the search process with any district n in $\{1, ..., M\}$. After a favorable signal, IG moves to any of the districts in $\{1, ..., M\} \setminus \{n\}$, that is, any of the $M-1$ still unsearched districts in $\{1, ..., M\}$. After an unfavorable signal, IG moves to any of the districts in $\{1, ..., M+1\} \setminus \{n\}$, that is, either any of the $M-1$ still unsearched districts in $\{1, ..., M\}$ or district $M+1$. This process is repeated until IG reaches a round t at which $K^t = 0$.

The following example illustrates the equilibrium search strategy.

Example 3. Consider the country described in Example 1, where there are five districts ($\mathcal{N}_0 = \{0, 1, 2, 3, 4\}$), the legislature takes its decision by simple majority ($M = 2$), and the realized profile of valuations for districts in $\mathcal{N} = \{1, 2, 3, 4\}$ is $(\overline{r}, \underline{r}, \overline{r}, \underline{r})$. Following Propositions 1 and 3, IG is indifferent starting the search process with district 1 or district 2.[23] Without loss of generality, let IG start with district 1. IG receives a favorable signal, and then searches district 2. IG receives an unfavorable signal, and then searches district 3. IG now receives a second favorable signal, and then stops searching ($i^{3+}\left(h^3\right) = 2$).[24] AS forms legislative coalition $\mathcal{L}^* = \{1, 3\}$ and the total quantity of goods $G^* = \Gamma\left(r_0, \overline{r}, \overline{r}\right)$. $\square$



The intuition underlying the equilibrium search sequence runs as follows. On the one hand, the equilibrium stopping rule prescribes that IG continues searching as long as it has not yet received M favorable signals (Proposition 1). The implication of this stopping rule is that AS will form a legislative coalition that includes only searched districts, either because M favorable signals have been received or because all districts have been searched. It thus follows that any search sequence satisfying the equilibrium stopping rule yields the same total quantity of goods (Lemma 3 in the Appendix). Hence, in terms of the total quantity of goods, IG is indifferent between any search sequence that satisfies the equilibrium stopping rule.[25] On the other hand, the costly search implies IG wants to minimize the number of districts it searches (while maximizing the prospects of receiving favorable signals). This induces IG to search districts with the highest probabilities of high valuation. That at each round t IG is indifferent searching any of the K^t still unsearched districts with the highest probabilities of high valuation follows because IG will anyway end up searching all these districts.

The following example illustrates this intuition.

[23] Recall IG searches district 0 at a preliminary stage of the search process.

[24] If instead IG had started its search with district 2, it would have received an unfavorable signal. At round 2 of the search process, IG would then have been indifferent searching district 1 or district 3. Searching either of these two districts, IG would have received a favorable signal and, at round 3, would have searched the other of the latter two districts. Receiving again a favorable signal, IG would have stopped its search. Thus, IG would have searched the same set of districts as when it starts the search process with district 1.

[25] This contrasts with the concave case where the equilibrium stopping rule prescribes IG to stop searching as soon as it has received $N - M$ unfavorable signals, meaning AS may include unsearched districts in the legislative coalition (see Examples 1 and 2). This explains why in the concave case, IG wants at each round t to keep unsearched the $M - i^{t+}\left(h^t\right)$ districts with the highest expected valuations.

EXAMPLE 4. Consider the country described in Example 2, where there are four districts ($\mathcal{N}_0 = \{0, 1, 2, 3\}$) and the legislative majority requirement is $M = 1$. We know from Proposition 1 that IG searches until it receives one favorable signal. I am going to show that IG starts its search with district 1 and, conditional on receiving unfavorable signals, continues with district 2 and then district 3.

If IG reaches round 3 of the search process, this means it has already received two unfavorable signals, and there is only one district left for IG to search.

Let's consider round 2 of the search process. Suppose IG had searched district 1 at round 1 and received an unfavorable signal. If IG searches district 2 at round 2, it will receive a favorable signal with probability p_2, in which case AS will form legislative coalition $\mathcal{L}^* = \{2\}$ and the total quantity of goods will be equal to $\Gamma(r_0, \overline{r})$. With probability $1 - p_2$ IG will receive instead an unfavorable signal, in which case it will then search district 3 and the total quantity of goods will be determined by the signal IG will receive at round 3. Thus, the expected total quantity of goods at round 2 if IG searches district 2 is given by

$$EG_{2|1} = [p_2 + (1 - p_2) \cdot p_3] \cdot \Gamma(r_0, \overline{r}) + (1 - p_2)(1 - p_3) \cdot \Gamma(r_0, \underline{r}).$$

Likewise, the expected total quantity of goods at round 2 if IG searches district 3, instead of district 2, is given by

$$EG_{3|1} = [p_3 + (1 - p_3) \cdot p_2] \cdot \Gamma(r_0, \overline{r}) + (1 - p_3)(1 - p_2) \cdot \Gamma(r_0, \underline{r}).$$

The expected total quantity of goods is then the same whether IG searches district 2 or district 3, that is, $EG_{2|1} = EG_{3|1}$. At the same time, $p_2 > p_3$ implies IG is more likely to stop its search after round 2 if it searches district 2 than if it searches district 3. Hence, the costly search, together with $EG_{2|1} = EG_{3|1}$, implies IG is better off searching district 2 than district 3.[26]

Let's now move to round 1 of the search process. On the one hand, we get, by the same argument as for round 2, that the total quantity of goods will be the same whether IG starts the search process with district 1, district 2 or district 3:

$$EG_1 = EG_2 = EG_3 = \begin{cases} \Gamma(r_0, \underline{r}) & \text{if } r_1 = r_2 = r_3 = \underline{r} \\ \Gamma(r_0, \overline{r}) & \text{otherwise.} \end{cases}$$

[26] Likewise, IG is better off searching district 1 at round 2 if it starts the search process with district 2 or with district 3 and receives an unfavorable signal.

On the other hand, $p_1 > p_2 > p_3$ implies IG is more likely to stop searching after only one round if it starts its search process with district 1 than with either district 2 or district 3. Hence, IG is better off starting with district 1. $\square$

5.4. Comparing the concave and convex cases

I underline four interesting differences between the concave and convex cases.

First, while there is a unique equilibrium search sequence in the concave case, there can be multiple ones in the convex case. However, all equilibrium search sequences in the convex case yield the same set of searched districts (Lemma 4 in the Appendix) and the same total quantity of goods (Lemma 3 in the Appendix).[27]

Second, in the concave case IG starts its search process with district $M+1$ and then moves *non-monotonically* towards districts 1 and N, alternating between districts with better prospects of high valuation and districts with worse prospects depending on the received signals. By contrast, in the convex case an equilibrium search sequence exists in which IG starts with district 1 and then moves *monotonically* towards district N, searching districts with worse prospects.

Third, IG may search all districts in $\mathcal{N}$ in the convex case, but never does so in the concave case. This difference follows from the equilibrium stopping rule. In the convex case, the equilibrium stopping rule prescribes IG to continue searching as long as it has not yet received M favorable signals. By contrast, in the concave case, the equilibrium stopping rule prescribes IG to stop searching as soon as it has received either M favorable signals or $N-M$ unfavorable signals. This implies that IG never searches at round N of the search process (since for every search history, either $i^{N+}\left(h^N\right) \geq M$ or $i^{N-}\left(h^N\right) \geq N-M$) and, therefore, never searches all districts in $\mathcal{N}$.

Finally, IG always searches district 1 in the convex case (except when $M=0$, where IG never searches districts in $\mathcal{N}$ in both the concave and convex cases), but never does so in the concave case. We get from Proposition 3 that in the convex case, IG searches all districts in $\{1,...,M\}$, including district 1. We get from Proposition 2 that in the concave case, IG starts the search process with district $M+1$, and then moves towards district 1 each time it receives a favorable signal. Since the

[27] Lemma 4 in the Appendix establishes that for a realized valuations profile $\mathbf{r}=(r_1,...,r_N)$ and majority requirement $M \in \{1,...,N\}$, there exists $\eta_{\mathbf{r},M} \in \{1,...,N\}$ such that the set of searched districts is $\left\{1,...,\eta_{\mathbf{r},M}\right\}$ for any equilibrium search sequence.

equilibrium stopping rule prescribes IG to stop searching as soon as it has received M favorable signals, IG always stops its search before reaching district 1.

6. DISCUSSION

I now discuss three implications from the model.

6.1. Welfare implications of lobbying

Who benefits and who loses from informational lobbying? The standard, optimistic view is that informational lobbying benefits a policymaker by relaxing his resource constraint (Hall and Deardorff, 2006; Groll and Ellis, 2020) and by helping him make better-informed policy choices (Austen-Smith and Wright, 1992; Cotton, 2009). Recent contributions have offered a more pessimistic view, showing that informational lobbying can be detrimental to legislators by triggering a change in the composition of the legislative coalition (Schnakenberg, 2015; Alonso and Camara, 2016). Both the standard, optimistic view and the more recent, pessimistic view apply to the context of regulatory politics, where the policy proposal does not respond to the information provided by IG.

I revisit this question in the context of distributive politics, where the policy proposal is endogenous to the information provided. I compare each player's equilibrium expected payoff in two games, one where IG has the possibility to search districts (the game described in Section 3) and another game where IG cannot search districts (i.e., without stage 1 of the game described in Section 3).

Informational lobbying has two key implications in the context of distributive politics. First, as in the context of regulatory politics, the composition of the legislative coalition responds to the information provided by IG. Second, informational lobbying triggers on average an increase in the total quantity of goods. This second implication is specific to the context of distributive politics, where the policy proposal is endogenous to the information provided.

In the equilibrium of the game without informational lobbying, AS forms a legislative coalition consisting of districts 1 through M, that is, the M districts with the best prospects of high valuation. AS offers an allocation of goods that binds the participation constraints of the legislators from those districts, meaning that each of them gets zero payoff. In addition, AS does not offer goods to districts

$M+1$ through N, meaning each of those legislators gets a negative payoff (equal to $-c(G)/N+1$).

IG and AS never lose from the possibility of lobbying as they both benefit from the increase in the total quantity of goods. Specifically, IG has always the possibility to not search any district, meaning its expected payoff is at least as large with the possibility of lobbying than without. Also, AS's interests are aligned with those of IG since his equilibrium payoff increases with the total quantity of goods.

Legislators from districts 1 through M never gain from lobbying. This is because they lose from the change in the composition of the legislative coalition and from the increase in the total quantity of goods. To see this, recall that in the absence of lobbying, each of those districts is included in the legislative coalition. This need not be the case with lobbying. Indeed, IG may receive unfavorable signals on the valuations of some of those districts or receive favorable signals on the valuations of districts in $\{M+1, ..., N\}$. AS may react to this information by replacing in the legislative coalition some of the districts in $\{1, ..., M\}$ with some of the districts in $\{M+1, ..., N\}$. Hence the expected welfare loss for the legislators from districts 1 through M since they get zero payoff when their district is in the legislative coalition, but a negative payoff when their district is not in the legislative coalition. This expected welfare loss is further enhanced by the expected increase in the total quantity of goods. Indeed, when a legislator's district is not included in the legislative coalition, her payoff decreases with the total quantity of goods as the provision cost of goods increases.

Legislators from districts $M+1$ through N may gain or lose from lobbying. This is because they benefit from the change in the composition of the legislative coalition, but lose from the increase in the total quantity of goods. To see this, recall that in the absence of lobbying, neither of those districts is included in the legislative coalition. This need not be the case with lobbying. Indeed, IG may receive favorable signals on the valuations of some of those districts or receive unfavorable signals on the valuations of districts in $\{1, ..., M\}$, inducing AS to replace in the legislative coalition some of the districts in $\{1, ..., M\}$ with some of the districts in $\{M+1, ..., N\}$. Hence the expected welfare gain for the legislators from districts $M+1$ through N since they get a negative payoff when their district is not in the legislative coalition, but a zero payoff when their district is included in the legislative coalition. On the other hand, the expected increase in the total quantity of goods generates a welfare loss for those legislators since their payoff decreases with the total quantity of goods when their district is kept out of the legislative coalition. Whether legislators from districts $M+1$ through N gain or lose from the possibility of lobbying thus depends on which of these two effects dominates the other. They gain if the first effect (change in the composition of the legislative coalition) dominates the second effect (increase in the total quantity of goods), and lose otherwise. The first effect is more likely to dominate the second effect the better the district's prospects of high valuation are.[28]

The next example shows that all legislators (other than AS) may lose from the possibility of lobbying and thus unanimously prefer to ban informational lobbying.

Example 5. Consider the country described in Example 2, where there are four districts ($\mathcal{N}_0 = \{0,1,2,3\}$) and the majority requirement is $M=1$. Suppose the cost function is given by $c(G) = G^2/2$ (implying the goods function $\Gamma(\cdot)$ is strictly concave in each of its arguments). For simplicity, let $\underline{r}=0$.[29] I further let districts' probabilities of high valuation be given by $(p_0,p_1,p_2,p_3) = \left(\frac{1}{2},\frac{1}{4},\frac{1}{10},\frac{1}{20}\right)$.

AS forms legislative coalition $\mathcal{L} = \{i\}$ where $i \in \arg\max_{n\in\mathcal{N}} E\widetilde{r}_n$, and proposes a goods allocation $\mathbf{g}$ with a total quantity of goods

$$G\left(E\widetilde{r}_i\right) = \frac{4}{\left(\frac{1}{Er_0} + \frac{1}{E\widetilde{r}_i}\right)}.$$

The provision cost of goods is then equal to $8/\left(\frac{1}{Er_0} + \frac{1}{E\widetilde{r}_i}\right)^2$.

In the game without lobbying (hereafter NL-game), AS forms legislative coalition $\mathcal{L}^{NL} = \{1\}$, and legislators' payoffs are equal to

$$\begin{cases} u_1^{NL} = 0 \\ u_2^{NL} = u_3^{NL} = -\frac{c(G(Er_1))}{4} = -2\overline{r}^2\left(\frac{p_0p_1}{p_0+p_1}\right)^2 = -\frac{\overline{r}^2}{18}. \end{cases}$$

In the game with lobbying (hereafter L-game), IG starts its search process with district 2. If IG receives a favorable signal, it stops searching, AS forms legislative coalition $\mathcal{L}^L = \{2\}$, and the total quantity of goods is given by $G(\overline{r})$. If IG receives

[28] Observe that under unanimity ($M=N$), each district is included in the legislative coalition, both with and without the possibility of lobbying. Hence neither legislator in $\mathcal{N}$ gains or loses from the possibility of lobbying.

[29] A similar example can be constructed with $\underline{r}>0$.

an unfavorable signal, it then searches district 3 and then stops. If IG receives a favorable signal, AS forms legislative coalition $\mathcal{L}^L = \{3\}$ and the total quantity of goods is given by $G(\overline{r})$. If IG receives an unfavorable signal, then AS forms legislative coalition $\mathcal{L}^L = \{1\}$ and the total quantity of goods is given by $G(Er_1)$. Legislators' expected payoffs are then equal to

$$\begin{cases} u_1^L = -\left[1-(1-p_2)(1-p_3)\right]\frac{c(G(\overline{r}))}{4} = -\frac{29}{900}\overline{r}^2 \\ u_2^L = -(1-p_2)\left[p_3\frac{c(G(\overline{r}))}{4} + (1-p_3)\frac{c(G(Er_1))}{4}\right] = -\frac{23}{400}\overline{r}^2 \\ u_3^L = -\left[p_2\frac{c(G(\overline{r}))}{4} + (1-p_2)(1-p_3)\frac{c(G(Er_1))}{4}\right] = -\frac{251}{3600}\overline{r}^2. \end{cases}$$

Hence $u_1^{NL} = 0 > u_1^L$. Likewise, $u_2^{NL} > u_2^L$ since

$$\underbrace{(1-p_2)\,p_3\frac{c(G(\overline{r})) - c(G(Er_1))}{4}}_{\text{Expected loss from increase in } G} > \underbrace{p_2\frac{c(G(Er_1))}{4}}_{\text{Expected gain from prospects of } \mathcal{L}=\{2\}},$$

and $u_3^{NL} > u_3^L$ since

$$\underbrace{p_2\frac{c(G(\overline{r})) - c(G(Er_1))}{4}}_{\text{Expected loss from increase in } G} > \underbrace{(1-p_2)\,p_3\frac{c(G(Er_1))}{4}}_{\text{Expected gain from prospects of } \mathcal{L}=\{3\}}.$$

Hence, all three legislators in $\mathcal{N}$ are ex ante strictly better off without than with informational lobbying. $\square$

6.2. Information provision and majority requirement

The relationship between information provision and majority requirement is non-monotonic in the concave case. This follows because the equilibrium stopping rule prescribes IG to continue searching if and only if it has not yet received either M favorable signals or $N - M$ unfavorable signals. This means that IG provides no information in the two polar cases of a dictatorship of AS ($M = 0$) and of unanimity ($M = N$). At the same time, IG searches at least one district in $\mathcal{N}$ for every majority requirement $M \in \{1, ..., N-1\}$. Hence the non-monotonicity since the number of searched districts increases between $M = 0$ and $M = 1$, and decreases between $M = N-1$ and $M = N$.

We can furthermore establish that the relationship between M and the *minimum* number of searched districts is single-peaked in the concave case. This happens because the equilibrium stopping rule implies IG searches at least $\min\{M, N-M\}$ districts in $\mathcal{N}$. Hence the minimum number of searched districts: (i) increases monotonically with M for infra majorities (where $\min\{M, N-M\} = M$); (*ii*) reaches a maximum around simple majority, that is, $M = \lfloor N/2 \rfloor$;[30] and (*iii*) decreases monotonically with M for super majorities (where $\min\{M, N-M\} = N-M$).

The relationship between M and the *expected* number of searched districts is likewise single-peaked in the concave case, with the majority requirement M at which the peak is reached depending on the probability profile $\mathbf{p} = (p_2, ..., p_N)$.[31] More formally, let $\#\mathcal{I}_{M,\mathbf{p}}$ be the expected equilibrium number of searched districts given majority requirement M and probability profile $\mathbf{p}$. For every $\mathbf{p}$, there exists $\mu_{\mathbf{p}} \in \{1, ..., N-1\}$ such that $\#\mathcal{I}_{M,\mathbf{p}}$ increases with M when $M < \mu_{\mathbf{p}}$, and decreases with M when $M > \mu_{\mathbf{p}}$.[32] The following example provides an illustration.

Example 6. Consider a country with five districts ($\mathcal{N}_0 = \{0, 1, 2, 3, 4\}$). We already know that $\#\mathcal{I}_{0,\mathbf{p}} = \#\mathcal{I}_{4,\mathbf{p}} = 0$ and $\#\mathcal{I}_{M,\mathbf{p}} \geq 1$ for every $M \in \{1, 2, 3\}$. Furthermore, simple computations give

$$\#\mathcal{I}_{2,\mathbf{p}} - \#\mathcal{I}_{1,\mathbf{p}} = p_3 + p_2(1-p_3) + (1-p_3)\left[p_2p_4 - (1-p_2)(1-p_4)\right] \tag{3}$$

$$\#\mathcal{I}_{2,\mathbf{p}} - \#\mathcal{I}_{3,\mathbf{p}} = (1-p_3) + p_3(1-p_4) + p_3\left[(1-p_2)(1-p_4) - p_2p_4\right]. \tag{4}$$

Consider first a probability profile $\mathbf{p}$ for which $\#\mathcal{I}_{1,\mathbf{p}} > \#\mathcal{I}_{2,\mathbf{p}}$. It follows from (3) that $(1-p_2)(1-p_4) > p_2p_4$, in which case (4) implies $\#\mathcal{I}_{2,\mathbf{p}} > \#\mathcal{I}_{3,\mathbf{p}}$. Thus, if the expected equilibrium number of searched districts decreases between $M = 1$ and $M = 2$, so does it between $M = 2$ and $M = 3$. Hence, $\#\mathcal{I}_{M,\mathbf{p}}$ increases between $M = 0$ and $M = 1$, and decreases monotonically thereafter.

[30] When N is an odd integer, there are two adjacent, equal peaks at $(N-1)/2$ and $(N+1)/2$.

[31] Observe that we can ignore p_1 since IG never searches district 1 in the concave case.

[32] We can easily characterize $\mu_{\mathbf{p}}$ for some probability profiles $\mathbf{p} = (p_2, ..., p_N)$.

- For $\mathbf{p}$ where $p_n \approx 1$ for all $n \in \{2, ..., N\}$, $\mu_{\mathbf{p}} \approx N-1$.
- For $\mathbf{p}$ where $p_n \approx 0$ for all $n \in \{2, ..., N\}$, $\mu_{\mathbf{p}} \approx 1$.
- For $\mathbf{p}$ where $p_n \approx 1/2$ for all $n \in \{2, ..., N\}$, $\mu_{\mathbf{p}} \approx N/2$ or $\mu_{\mathbf{p}} \approx (N-1)/2, (N+1)/2$ depending on whether N is an even or odd integer, that is, $\#\mathcal{I}_{M,\mathbf{p}}$ reaches its peak around simple majority. Moreover, $\#\mathcal{I}_{M,\mathbf{p}}$ is distributed symmetrically around simple majority, that is, $\#\mathcal{I}_{m,\mathbf{p}} = \#\mathcal{I}_{N-m,\mathbf{p}} < \#\mathcal{I}_{m+1,\mathbf{p}} = \#\mathcal{I}_{N-m-1,\mathbf{p}}$ for $m \in \{0, ..., \mu_{\mathbf{p}} - 1\}$.
- For $\mathbf{p}$ where $p_2, ..., p_N$ are distributed symmetrically around $1/2$, $\mu_{\mathbf{p}} = N/2$ or $\mu_{\mathbf{p}} = (N-1)/2, (N+1)/2$ depending on whether N is an even or odd integer. Moreover, $\#\mathcal{I}_{M,\mathbf{p}}$ is distributed symmetrically around simple majority.

Consider a probability profile $\mathbf{p}$ for which $\#\mathcal{I}_{2,\mathbf{p}} < \#\mathcal{I}_{3,\mathbf{p}}$. It follows from (4) that $p_2 p_4 > (1 - p_2)(1 - p_4)$, in which case (3) implies $\#\mathcal{I}_{1,\mathbf{p}} < \#\mathcal{I}_{2,\mathbf{p}}$. Thus, if the expected equilibrium number of searched districts increases between $M = 2$ and $M = 3$, so does it between $M = 1$ and $M = 2$. Hence, $\#\mathcal{I}_{M,\mathbf{p}}$ increases monotonically between $M = 0$ and $M = 3$, and then decreases between $M = 3$ and $M = 4$. □

By contrast, in the convex case the relationship between information provision and majority requirement (M) is monotonic. This follows because the equilibrium stopping rule prescribes IG to continue searching if and only if it has not yet received M favorable signals. This means that IG provides no information on districts in $\mathcal{N}$ when $M = 0$, that is, in the polar case of a dictatorship of AS. This means furthermore that IG provides information on every district when $M = N$, that is, in the polar case of unanimity. Finally, this also means that given a profile of realized valuations $\mathbf{r} = (r_1, ..., r_N)$, the number of districts on which IG provides information increases with M. Hence, the (expected) number of districts in $\mathcal{N}$ for which IG provides information increases monotonically with M, from none when $M = 0$ up to all when $M = N$.

I conclude this section with a comparison between two instruments of interest group influence, namely, information provision and monetary contributions. On the one hand, we have seen that the expected equilibrium number of searched districts may be maximized under a simple or a super majority requirement, depending on districts' prospects of high valuation and on the curvature of the goods function. On the other hand, in a different setting where interest groups buy votes and legislators seek to maximize the total amount of monetary contributions they receive, Diermeier and Myerson (1999) finds incentives for legislators in a unicameral legislature to delegate authority to a leader, which in my setting could be associated with an infra-majority requirement or, in the limit, a dictatorship of AS. This suggests that legislators in a unicameral legislature might want to adopt an infra-majority requirement if they seek to maximize the amount of monetary contributions they receive, while they may prefer to adopt a simple- or super-majority requirement if they care about information instead of monetary contributions.

6.3. Nature of lobbying

Lobbying is said to be *friendly* when an interest group lobbies allies, that is, legislators who, in the absence of lobbying, would vote along the interest group's position. Instead, lobbying is said to be *confrontational* when an interest group lobbies opponents, that is, legislators who, in the absence of lobbying, would vote against the interest group's position. Does an interest group lobby allies or opponents? Answering this question matters since, as Kollman (1997: 520) writes, people argue that "[i]f interest groups lobby their friends (the friendly model), the influence of lobbying may not be as large as many people think because lobbyists merely reinforce existing policy preferences among legislators."

Empirical studies (e.g., Hojnacki and Kimball, 1998; Hall and Miler, 2008; You, 2020) find evidence that, at least in the U.S., interest groups lobby both allies and opponents, and that they are more likely to lobby allies than opponents.

Lobbying opponents is rather intuitive: an interest group provides information to its opponents in order to persuade them to support the group's position. By contrast, lobbying allies is rather unintuitive. Why would an interest group waste resources on seeking to persuade legislators who already support the group's position? Scholars have proposed several explanations to rationalize friendly lobbying. Austen-Smith and Wright (1992, 1994) explain, and provide empirical evidence in support of, friendly lobbying as an effort to counteract the influence of groups with opposite interests. Bauer, Dexter and De Sola Pool (1963) argues that interest groups serve mainly as 'service bureaus' to resource-constrained legislators. Based on this argument, Hall and Deardorff (2006) and Groll and Ellis (2020) explain friendly lobbying as a legislative subsidy, whereby interest groups provide information to legislative allies in order to relax their resource constraints. In the same spirit, Cotton and Dellis (2016) explains friendly lobbying as an effort by interest groups to push their issues at the top of the agenda. Likewise, Groll and Prummer (2016), Schnakenberg (2017) and Awad (2020) explain friendly lobbying as a choice to enroll legislative allies as intermediaries who help convince legislative opponents.[33] In other words, by directly lobbying allies, interest groups indirectly

[33] In Groll and Prummer (2016) friendly lobbying arises because of the network structure among legislators. In Schnakenberg (2017) it arises as a way to save on the cost of accessing legislators. In Awad (2020) it arises because of the possibility of persuasion cascades.

lobby opponents. In Minaudier (2021) friendly lobbying is more likely to occur than confrontational lobbying because an ally is easier to re-convince than an opponent after he acquires information on his own.

All these explanations apply to the context of regulatory politics. By contrast, my model applies to the context of distributive politics where the policy proposal is endogenous to lobbying. Applying the definitions of friendly and confrontational lobbying to this context, we get that lobbying is friendly (resp. confrontational) when IG searches a district which, in the absence of lobbying, would (resp. would not) be included in the legislative coalition and which legislator would then vote in favor of (resp. against) the allocation AS would propose. Given that in the absence of lobbying AS would form a legislative coalition consisting of districts $1, ..., M$, lobbying is friendly when it targets a district in $\{1, ..., M\}$ and is confrontational when it targets a district in $\{M + 1, ..., N\}$.

In the concave case, IG starts by searching district $M + 1$, that is, the district with the best prospects of high valuation among districts that otherwise would not be included in the legislative coalition. Thus, IG starts with confrontational lobbying. Afterwards, and conditional on still searching, each time IG receives a favorable signal, it continues its search with a district in $\{1, ..., M\}$, thus moving to or keeping with friendly lobbying. Each time IG receives an unfavorable signal, it continues its search with a district in $\{M + 2, ..., N\}$, thus moving back to or keeping with confrontational lobbying.[34]

In the convex case, IG starts by searching districts in $\{1, ..., M\}$, that is, districts which, in the absence of lobbying, would have been included in the legislative coalition. Thus, IG starts with friendly lobbying. Afterwards, and conditional on still searching, IG continues its search with districts in $\{M + 1, ..., N\}$, thus moving to confrontational lobbying.

Thus, my model predicts that more lobbying would be friendly in circumstances where many districts have a high valuation or where the provision cost of goods is such that the goods function is convex.

This discussion offers an interesting contrast between distributive and regulatory politics. Suppose we were to apply the model to the context of regulatory

[34] This search sequence is consistent with Miller's (2020) finding that lobbyists prefer targeting weak allies and opponents than strong ones.

politics, where a policy proposal would be made prior to lobbying and IG would seek to persuade at least M legislators to vote in favor of the proposal. Suppose a legislator gains from, and votes in favor of, the adoption of the proposal if she believes her district's valuation is more likely to be high than low (i.e., $\widetilde{p}_n \geq 1/2$). In this setting, either there would be no lobbying or all lobbying would be confrontational. Specifically, there would be no lobbying in circumstances where at least M legislators are ex ante in favor of the policy proposal, that is, $p_M \geq 1/2$. Otherwise (i.e., when $p_M < 1/2$), IG would start by searching districts with the highest p_n below 1/2 (weakest opponents) and then move to districts with smaller p_n (stronger opponents) until M legislators favor the policy proposal. Hence, while some lobbying can be friendly in the context of distributive politics, all lobbying (if any) is confrontational in the context of regulatory politics. To look at it from a different perspective, this suggests that in the context of regulatory politics, some links between legislators are necessary to rationalize friendly lobbying, as in Schnakenberg (2017) and Awad (2020).

7. CONCLUSION

This paper has investigated the question of whom an interest group should lobby in a legislative assembly, where lobbying is defined as the act of providing information. To investigate this question, I have developed a model of informational lobbying set in the context of distributive politics, where a legislature must decide on the allocation of district-specific goods and projects. Districts are ex ante heterogeneous, differing in their prospects of having a high valuation of goods. An interest group, which benefits from the provision of goods, chooses sequentially to search information on districts' valuations. The agenda setter's allocation proposal is endogenous to the information provided by the interest group.

I have characterized the equilibrium stopping rule, which specifies when the interest group should stop searching, as well as the equilibrium search sequence, which specifies the order in which the interest group should search districts. The analysis generates three key implications. First, legislators (other than the agenda setter) may unanimously prefer to ban informational lobbying since they may all be ex ante worse off with than without the possibility of informational lobbying. Second, Diermeier and Myerson (1999) suggests that legislators in a unicameral legislature may want to adopt *infra-majority* requirements (by delegating authority to a leader) if they seek to maximize the expected amount of *monetary contributions* they receive. By contrast, my analysis suggests that legislators may want to adopt *simple- or super-majority* requirements if they seek to maximize the expected amount of *information* they receive. Third, empirical studies find evidence that interest groups lobby both their legislative opponents (confrontational lobbying) as well as their legislative allies (friendly lobbying). My analysis provides an information rationale both for friendly lobbying and for confrontational lobbying.

In a supplementary online appendix, I extend the model to a situation where districts vary in the quality of information that can be obtained on their valuation. Specifically, if the interest group chooses to search district n, it will receive a signal revealing this district's valuation with probability $q_n \in (0,1]$, and will receive no signal (or an uninformative signal) with probability $1 - q_n$. I show that the equilibrium stopping rule is the same as when districts differ in their prospects of high valuation. I show furthermore that the interest group should then start by searching districts for which it has a higher probability of receiving an informative signal, and then moves to districts which signals are less likely to be informative.[35]

I have made several assumptions to keep the analysis simple. First, the model involves a single interest group that benefits from the provision of goods. This assumption is standard in models of lobbying and is consistent with empirical evidence on many issues. For example, Baumgartner et al. (2009: 57) writes: "... a surprisingly large number of issues ... consist of a single side attempting to achieve

[35] In another extension, I consider the possibility for legislators to search their own district at the same small cost $\varepsilon > 0$ as IG. Specifically, I add to the game a stage between IG's information collection (stage 1) and AS' proposal (stage 2), where legislators decide simultaneously whether to search their own district. As signals are perfectly informative, a legislator may choose to search her district only if it has not yet been searched by IG. I find that the equilibrium total quantity of goods is the same as in the baseline model. In the concave case, I further find that IG follows the same search sequence as in the baseline model, with the exception that it stops searching once it has received $M-1$ (instead of M) favorable signals and anticipates all legislators from the still unsearched districts, except the legislator from district 1, will search their districts. In the convex case, I find that in equilibrium, each district is searched, either by IG or by its legislator. Furthermore, IG searches districts until it has received M favorable signals or it anticipates that every legislator from a still unsearched district will search information on her own. The latter implies that in the convex case, IG may deviate from the search sequence it follows in the baseline model if this can trigger more legislators to search information on their own.

a goal to which no one objects or in response to which no one bothers to mobilize." This being said, Baumgartner et al. (2009: 58) further writes: "A majority of cases [58 out of 98] had two sides." Likewise, Huneeus and Kim (2021) finds that most congressional bills are lobbied by only one or two interest groups. In the context of my analysis, one could imagine a group benefiting from the provision of goods (e.g., a road builders' association) opposed to another group pushing for cuts in the provision of goods (e.g., an environmental group advocating against building new roads). It would be interesting to investigate how the presence of an opposite interest group affects legislative informational lobbying. Second, districts' valuations are drawn independently. This means that the information obtained on a district's valuation provides no information on other districts' valuations. However, one could imagine situations where valuations are fairly similar in geographically adjacent legislative districts. A variant of the model could include some degree of correlation across districts' valuations. Third, I have adopted a persuasion setting, where signals are public information. This setting allows to focus on strategic information *production*. It would be interesting to combine strategic information production with strategic information *transmission*. Milgrom and Roberts' (1986) unravelling theorem implies that simply adding the possibility for the interest group to not reveal signals would not affect the main qualitative conclusions of the analysis. But relaxing furthermore the assumption that legislators observe the interest group's search decision could yield interesting insights. Fourth, the interest group has only one instrument for influencing policy, namely, information provision. It would be interesting to allow the interest group to make monetary contributions in addition to providing information. One could then analyze, in a way similar to Bennedsen and Feldmann (2006) and Dahm and Porteiro (2008), but in the context of a collective decisionmaking, how the possibility for the interest group to use monetary contributions as a control-damage device following unfavorable signals affects information provision by the interest group. These extensions go beyond the scope of the present analysis and are left for future research.

REFERENCES

[1] **Alonso, Ricardo, and Odilon Câmara.** 2016. "Persuading voters." *American Economic Review*, 106(11): 3590-3605.

[2] **Ansolabehere, Stephen, James Snyder, and Micky Tripathi.** 2002. "Are PAC contributions and lobbying linked? New evidence from the 1995 Lobby Disclosure Act." *Business and Politics*, 4(2): 131-155.

[3] **Austen-Smith, David.** 1995. "Campaign contributions and access." *American Political Science Review*, 89(3): 566-581.

[4] **Austen-Smith, David.** 1998. "Allocating access for information and contributions." *Journal of Law, Economics, and Organization*, 14(2): 277-303.

[5] **Austen-Smith, David, and John R. Wright.** 1992. "Competitive lobbying for a legislator's vote." *Social Choice and Welfare*, 9: 229-257.

[6] **Austen-Smith, David, and John R. Wright.** 1994. "Counteractive lobbying." *American Journal of Political Science*, 38(1): 25-44.

[7] **Awad, Emiel.** 2020. "Persuasive lobbying with allied legislators." *American Journal of Political Science*, 64(4): 938-951.

[8] **Baron, David.** 2006. "Competitive lobbying and supermajorities in a majority-rule institution." *Scandinavian Journal of Economics*, 108(4): 607-642.

[9] **Battaglini, Marco, and Eleonora Patacchini.** 2018. "Influencing connected legislators." *Journal of Political Economy*, 126(6): 2277-2322.

[10] **Bauer, Raymond, Lewis Dexter, and Ithiel de Sola Pool.** 1963. *American Business and Public Policy: The Politics of Foreign Trade.* Atherton Press: New York City, NY.

[11] **Baumgartner, Frank, Jeffrey Berry, Marie Hojnacki, and Beth Leech.** 2009. *Lobbying and Policy Change. Who Wins, Who Loses, and Why.* Chicago University Press: Chicago, IL.

[12] **Belloc, Marianna.** 2015. "Information for sale in the European Union." *Journal of Economic Behavior & Organization*, 120: 130-144.

[13] **Bennedsen, Morten, and Sven Feldmann.** 2002. "Lobbying legislatures." *Journal of Political Economy*, 110(4): 919-946.

[14] **Bennedsen, Morten, and Sven Feldmann.** 2006. "Informational lobbying and political contributions." *Journal of Public Economics*, 90: 631-656.

[15] **Bombardini, Matilde, and Francesco Trebbi.** 2020. "Empirical models of lobbying." *Annual Review of Economics*, 12: 391-413.

[16] **Caillaud, Bernard, and Jean Tirole.** 2007. "Consensus building: How to persuade a group." *American Economic Review*, 97(5): 1877-1900.

[17] **Cotton, Christopher.** 2009. "Should we tax or cap political contributions? A lobbying model with policy favors and access." *Journal of Public Economics*, 93(7/8): 831-842.

[18] **Cotton, Christopher, and Arnaud Dellis.** 2016. "Informational lobbying and agenda distortion." *Journal of Law, Economics, and Organization*, 32(4): 762-793.

[19] **Dahm, Matthias, and Nicolas Porteiro.** 2008. "Side effects of campaign finance reform." *Journal of the European Economic Association*, 6(5): 1057-1077.

[20] **de Figueiredo, John, and Brian Kelleher Richter.** 2014. "Advancing the empirical research on lobbying." *Annual Review of Political Science*, 17: 163-185.

[21] **Dekel, Eddie, Matthew Jackson, and Asher Wolinsky.** 2009. "Vote buying: Legislatures and lobbying." *Quarterly Journal of Political Science*, 4(2): 103-128.

[22] **Diermeier, Daniel, and Roger Myerson.** 1999. "Bicameralism and its consequences for the internal organization of legislatures." *American Economic Review*, 89(5): 1182-1196.

[23] **Doval, Laura.** 2018. "Whether or not to open Pandora's box." *Journal of Economic Theory*, 175: 127-158.

[24] **Drutman, Lee.** 2015. *The Business of America is Lobbying: How Corporations Became Politicized and Politics Became More Corporate.* Oxford University Press: Oxford.

[25] **Fehrler, Sebastian, and Maik T. Schneider**. 2021. "Buying supermajorities in the lab." *Games and Economic Behavior*, forthcoming.

[26] **Gawande, Kishore, William Maloney, and Gabriel Montes-Rojas.** 2009. "Foreign informational lobbying can enhance tourism: Evidence from the Carribean." *Journal of Development Economics*, 90(2): 267-275.

[27] **Groll, Thomas, and Christopher Ellis.** 2020. "Strategic legislative subsidies: Informational lobbying and the cost of policy." *American Political Science Review*, 114(1): 179-205.

[28] **Groll, Thomas, and Anja Prummer.** 2016. "Whom to lobby? Targeting in political networks." Unpublished.

[29] **Groseclose, Tim, and James Snyder.** 1996. "Buying supermajorities." *American Political Science Review*, 90(2): 303-315.

[30] **Hall, Richard, and Alan Deardorff.** 2006. "Lobbying as legislative subsidy." *American Political Science Review*, 100(1): 69-84.

[31] **Hall, Richard, and Kristina Miler.** 2008. "What happens after the alarm? Interest group subsidies to legislative overseers." *Journal of Politics*, 70(4): 990-1005.

[32] **Hansen, John.** 1991. *Gaining Access. Congress and the Farm Lobby, 1919-1981.* University of Chicago Press: Chicago, IL.

[33] **Hojnacki, Marie, and David Kimball.** 1998. "Organized interests and the decision of whom to lobby in Congress." *American Political Science Review*, 92(4): 775-790.

[34] **Huneeus, Federico, and In Song Kim.** 2021. "The effects of firms' lobbying on resource misallocation." Unpublished.

[35] **Igan, Deniz, and Prachi Mishra.** 2014. "Wall Street, Capitol Hill, and K Street: Political influence and financial regulation." *Journal of Law and Economics*, 57(4): 1063-1084.

[36] **Jensen, Jennifer.** 2018. "Intergovernmental lobbying in the United States: Assessing the benefits of accumulated knowledge." *State and Local Government Review*, 50(4): 270-281.

[37] **Kollman, Ken.** 1997. "Inviting friends to lobby: Interest groups, ideological bias, and Congressional committees." *American Journal of Political Science*, 41(2): 519-544.

[38] **Lohmann, Susanne.** 1995. "Information, access, and contributions: A signaling model of lobbying." *Public Choice*, 85(3/4): 267-284.

[39] **Milgrom, Paul, and John Roberts.** 1986. "Relying on the information of interested parties." *RAND Journal of Economics*, 17(1): 18-32.

[40] **Miller, David.** 2020. "On whose door to knock? Organized interests' strategic pursuit of access to members of Congress." *Legislative Studies Quarterly*, forthcoming.

[41] **Minaudier, Clement.** 2021. "The value of confidential policy information: Persuasion, transparency, and influence." *Journal of Law, Economics, and Organization*, forthcoming.

[42] **Olszewski, Wojciech, and Richard Weber.** 2015. "A more general Pandora rule?" *Journal of Economic Theory*, 160: 429-437.

[43] **Potters, Jan, and Frans van Winden.** 1992. "Lobbying and asymmetric information." *Public Choice*, 74(3): 269-292.

[44] **Schnakenberg, Keith.** 2015. "Expert advice to a voting body." *Journal of Economic Theory*, 160: 102-113.

[45] **Schnakenberg, Keith.** 2017. "Informational lobbying and legislative voting." *American Journal of Political Science*, 61(1): 129-145.

[46] **Schneider, Maik.** 2014. "Interest-group size and legislative lobbying." *Journal of Economic Behavior & Organization*, 106: 29-41.

[47] **Snyder, James.** 1991. "On buying legislatures." *Economics and Politics*, 3(2): 93-109.

[48] **Weingast, Barry, Kenneth Shepsle, and Christopher Johnsen**. 1981. "The political economy of benefits and costs: A neoclassical approach to distributive politics." *Journal of Political Economy*, 89(4): 642-664.

[49] ¸**Weitzman, Martin.** 1979. “Optimal search for the best alternative.” *Econometrica*, 47(3): 641-654.

[50] **Wright, John R.** 1990. “Contributions, lobbying, and committee voting in the U.S. House of Representatives.” *American Political Science Review*, 84(2): 417-438.

[51] **You, Hye Young.** 2020. “Dynamic lobbying: How to persuade Congress.” Unpublished.

APPENDIX

Given round $t \in \{1, ..., N\}$ and search history h^t, define $\mathcal{C}^t(h^t) \equiv \mathcal{N} \backslash \mathcal{I}^t(h^t)$ as the set of districts in $\mathcal{N}$ that are still unsearched at the beginning of round t. Given search strategy s, denote the expected quantity of goods by $EG(s)$, IG's expected utility by $EV(s)$ and the legislative coalition by $\mathcal{L}(s)$.

To ease notation, I shall omit $E\widetilde{r}_0$ from $\Gamma(\cdot)$ and the argument h^t whenever it does not create confusion.

Proof of Proposition 1. **(Sufficiency)** I proceed by contradiction. Let $\mathbf{r} \in \{\underline{r}, \overline{r}\}^N$ be any vector of valuations for districts in $\mathcal{N}$. Let s be an equilibrium search sequence. Consider round $t \in \{1, ..., N\}$ at which $i^{t+} \geq M$ or, in the concave case, $i^{t-} \geq N - M$.[36] Assume by way of contradiction that $s^t = h \in \mathcal{C}^t$. Construct search sequence s' such that

$$s^{\tau\prime} = \begin{cases} s^{\tau} & \text{for all } \tau < t \\ \emptyset & \text{for all } \tau \geq t. \end{cases}$$

If $i^{t+} \geq M$, we have $EG(s') = EG(s) = \Gamma(\overline{r}, ..., \overline{r})$. This, together with $\varepsilon > 0$, implies $EV(s') > EV(s)$.

If $i^{t-} \geq N - M$, then $Er_k > \underline{r}$ for all $k \in \mathcal{C}^t$ implies $\mathcal{C}^t \subseteq \mathcal{L}(s')$. There is no loss of generality in letting $\mathcal{L}(s) = \mathcal{L}(s')$. In the concave case, we then have

$$\begin{aligned} EG(s') &= \Gamma(Er_h; R) \\ &> p_h \cdot \Gamma(\overline{r}; R) + (1 - p_h) \cdot \Gamma(\underline{r}; R) \geq EG(s) \end{aligned}$$

[36] Observe $i^{t+} + i^{t-} = t - 1$, meaning $i^{t+} \geq M$ and $i^{t-} \geq N - M$ cannot both be true.

where R is the $(M-1)$-dimensional vector of expected valuations of districts in $\mathcal{L}(s) \setminus \{h\}$. Hence $EV(s') > EV(s)$.

In either case, $EV(s') > EV(s)$ contradicts s is an equilibrium search sequence.

(Necessity) I prove the contrapositive. Pick any $\mathbf{r} \in \{\underline{r}, \overline{r}\}^N$. Consider round t at which $i^{t+} \leq M-1$ and, in the concave case, $i^{t-} \leq N-M-1$. Assume by way of contradiction that $s^t = \emptyset$ and, if $t \geq 2$, that $s^{t-1} \neq \emptyset$ (that is, IG stops searching at round t). Observe that $i^{t+} \leq M-1$ implies $\mathcal{L}(s) \cap \mathcal{C}^t \neq \emptyset$.

I start with the convex case. Pick some $h \in \mathcal{L}(s) \cap \mathcal{C}^t$, and construct search sequence s' such that

$$s'_\tau = \begin{cases} s_\tau & \text{for all } \tau < t \\ h & \text{for } \tau = t \\ \emptyset & \text{for all } \tau > t. \end{cases}$$

The expected total quantity of goods at round t is given by

$$\begin{aligned} EG(s') &\geq p_h \cdot \Gamma(\overline{r}; R) + (1-p_h) \cdot \Gamma(\underline{r}; R) \\ &> \Gamma(Er_h; R) = EG(s) \end{aligned}$$

where R is the $(M-1)$-vector containing $E\widetilde{r}_n$ for every $n \in \mathcal{L}(s) \setminus \{h\}$. The weak inequality follows from the possibility that $h \notin \mathcal{L}(s')$ when $r_h = \underline{r}$. The strict inequality follows from the strict convexity of $\Gamma(\cdot; R)$.

I now consider the concave case. Observe that $i^{t+} + i^{t-} \leq N-2$ implies $\#\mathcal{C}^t \geq 2$ and that $N - i^{t-} \geq M+1$ implies $\mathcal{C}^t \nsubseteq \mathcal{L}(s)$. Construct search sequence s' such that

$$s^{\tau\prime} = \begin{cases} s^\tau & \text{for all } \tau < t \\ h & \text{for } \tau = t \\ \emptyset & \text{for all } \tau > t \end{cases}$$

where $h = \underset{i \in \mathcal{C}^t}{\arg\min}\, p_i$. Observe that $\mathcal{C}^t \nsubseteq \mathcal{L}(s)$ and $p_h < p_k$ for all $k \in \mathcal{C}^t \setminus \{h\}$ implies $h \notin \mathcal{L}(s)$.

If $r_h = \overline{r}$, then $\mathcal{L}(s') = (\mathcal{L}(s) \setminus \{j\}) \cup \{h\}$ where $j = \underset{i \in \mathcal{L}(s) \cap \mathcal{C}^t}{\arg\min}\, p_i$. If $r_h = \underline{r}$, then $\mathcal{L}(s') = \mathcal{L}(s)$. Since $\Gamma(\cdot)$ is strictly increasing in each of its arguments and $\overline{r} > Er_j$, we get

$$\begin{aligned} EG(s') &= p_h \cdot \Gamma(\overline{r}; R) + (1-p_h) \cdot \Gamma(Er_j; R) \\ &> \Gamma(Er_j; R) = EG(s) \end{aligned}$$

where R is the $(M-1)$-dimensional vector of expected valuations of districts in $\mathcal{L}(s) \setminus \{j\}$.

Hence $EV(s') > EV(s)$ whether $\Gamma(\cdot)$ is strictly concave or convex, which contradicts that s is an equilibrium search sequence. ■

Proof of Proposition 2. The proof is inductive. I am going to show that at each round t where $i^{t+} < M$ and $i^{t-} < N-M$, we have $s^t = K^t + 1$ where $K^t = M - i^{t+}$ and districts in $\mathcal{C}^t$ are relabelled such that $\mathcal{C}^t = \{1, ..., N+1-t\}$ with $p_1 > ... > p_{N+1-t}$.

Proposition 1 implies $s^N = \emptyset$. Consider round $t = N-1$ where $i^{t+} < M$ and $i^{t-} < N-M$. It follows from Proposition 1 that $s^{N-1} \neq \emptyset$. Observe that $\mathcal{C}^{N-1} = \{1, 2\}$, $i^{t+} = M-1$ and $i^{t-} = N-M-1$. Hence $K^t = 1$. I am going to show that $s^{N-1} = 2$.

For $s^{N-1} = i \in \mathcal{C}^{N-1}$, the expected total quantity of goods is given by

$$EG_i = p_i \cdot \Gamma(\overline{r}, ..., \overline{r}) + (1-p_i) \cdot \Gamma(Er_j, \overline{r}, ..., \overline{r})$$

where $\{j\} = \mathcal{C}^{N-1} \setminus \{i\}$. We then have

$$EG_2 - EG_1 = (1-p_2) \cdot \left[\Gamma(Er_1, \overline{r}, ..., \overline{r}) - \frac{1-p_1}{1-p_2} \cdot \Gamma(Er_2, \overline{r}, ..., \overline{r}) - \frac{p_1 - p_2}{1-p_2} \cdot \Gamma(\overline{r}, \overline{r}, ..., \overline{r}) \right] > 0,$$

the strict inequality since $Er_1 = \frac{1-p_1}{1-p_2} Er_2 + \frac{p_1-p_2}{1-p_2} \overline{r}$ and $\Gamma(\cdot, \overline{r}, ..., \overline{r})$ is strictly concave. Hence $s^{N-1} = 2$.

Assume the statement is true at round $(t+1) \in \{2, ..., N-1\}$. Consider round t with $i^{t+} < M$ and $i^{t-} < N-M$. We have $K^t \in \{1, ..., M\}$ and $\#\mathcal{C}^t \geq K^t + 1$. I am going to show that $s^t = K^t + 1$. I first establish in Lemma 1 that

$$\max_{h \in \{1, ..., K^t\}} EG_h < EG_{K^t+1} = ... = EG_{N+1-t},$$

implying $s^t \geq K^t + 1$. I then establish in Lemma 2 that

$$E\varepsilon_{K^t+1} < E\varepsilon_{K^t+2} < ... < E\varepsilon_{N+1-t},$$

where $E\varepsilon_i$ denotes the expected search cost for the continuation search sequence starting at $s^t = i$, which implies $s^t = K^t + 1$.

LEMMA 1. *Consider round t at which $i^{t+} < M$ and $i^{t-} < N - M$. Letting $K^t \equiv M - i^{t+}$, we have that*

$$\max_{h \in \{1,...,K^t\}} EG_h < EG_{K^t+1} = ... = EG_{N+1-t}.$$

PROOF OF LEMMA 1. Pick $i \in \mathcal{C}^t$, and suppose $s^t = i$. Recall that $\#\mathcal{C}^t = (N + 1 - t) \in \{K^t + 1, ..., N\}$.

Define $h(i;1) \equiv \max \mathcal{C}^t \backslash \{i\}$ and $h(i;k) \equiv \max \mathcal{C}^t \backslash \{i, h(i;1), ..., h(i;k-1)\}$ for $k = 2, ..., K^t$ as the district in $\mathcal{C}^t \backslash \{i\}$ with the k^{th} highest expected valuation.

The expected total quantity of goods at round t is given by

$$EG_i = \sum_{j=0}^{K^t-1} \alpha_j^i \cdot \Gamma\left(Er_{h(i;1)}, ..., Er_{h(i;K^t-j)}, \overline{r}, ..., \overline{r}\right) + \alpha_{K^t}^i \cdot \Gamma(\overline{r}, ..., \overline{r}),$$

where $\Gamma(x_1, ..., x_{K^t})$ is a shorthand for $\Gamma(x_1, ..., x_{K^t}, \overline{r}, ..., \overline{r})$, $\alpha_{K^t}^i = \left(1 - \sum_{j=0}^{K^t-1} \alpha_j^i\right)$ and

$$\alpha_j^i = \sum_{S \in S_j(h(i;1),...,h(i;K^t+1-j))} \left[\prod_{\ell \in S} p_\ell\right] \left[\prod_{\ell \in \mathcal{C}^t \backslash (S \cup \{h(i;1),...,h(i;K^t-j)\})} (1 - p_\ell)\right]$$

for $j \in \{0, ..., K^t - 1\}$, where $S_j(H)$ is the collection of cardinality-j subsets of $\mathcal{C}^t \backslash H$. α_j^i is the probability IG will receive j $\overline{r}$-signals starting from round t onwards, and $\alpha_{K^t}^i$ is the probability IG will stop searching after having received K^t $\overline{r}$-signals starting from round t onwards.

Observe that $\{h(i;1), ..., h(i;K^t)\}$ is the same for all $i \in \{K^t + 1, ..., N + 1 - t\}$, which implies $EG_{K^t+1} = ... = EG_{N+1-t}$.

Pick some $i \in \{1, ..., K^t\}$. Observe that

$$h(i;k) = \begin{cases} h(K^t + 1; k) = k & \text{for } k \in \{1, ..., i - 1\} \\ h(K^t + 1; k) + 1 = k + 1 & \text{for } k \in \{i, ..., K^t\}. \end{cases}$$

Observe also that if $i \neq 1$, we have $\alpha_j^{K^t+1} = \alpha_j^i$ for all $j \in \{K^t + 2 - i, ..., K^t\}$.[37] We then get[38]

$$\begin{aligned} EG_{K^t+1} - EG_i \quad = \quad & \sum_{j=0}^{K^t+1-i} [\alpha_j^{K^t+1} \cdot \Gamma(Er_1, ..., Er_{K^t-j}, \overline{r}, ..., \overline{r}) \\ & \quad -\alpha_j^i \cdot \Gamma\left(Er_{h(i;1)}, ..., Er_{h(i;K^t-j)}, \overline{r}, ..., \overline{r}\right)]. \end{aligned}$$

---

[37] For $i = 1$, we have $\alpha_j^{K^t+1} = \alpha_j^1$ for all $j \in \{0, ..., K^t\}$.

[38] For $i = 1$ and $j = K^t$, I write loosely $\Gamma\left(Er_1, ..., Er_{K^t-j}, \overline{r}, ..., \overline{r}\right)$ and $\Gamma\left(Er_{h(i;1)}, ..., Er_{h(i;K^t-j)}, \overline{r}, ..., \overline{r}\right)$ as $\Gamma(\overline{r}, ..., \overline{r})$.

Define

$$D_j \equiv (p_i - p_{K^t+1-j}) \cdot \left\{ \sum_{S \in S_j(1,...,K^t+1-j)} \left[\prod_{\ell \in S} p_\ell \right] \left[\prod_{\ell \in \mathcal{C}^t \backslash (S \cup \{1,...,K^t+1-j\})} (1-p_\ell) \right] \right\}$$

for each $j \in \{0, ..., K^t - i - 1\}$. We can rewrite $EG_{K^t+1} - EG_i$ as follows

$$\begin{aligned}
&\alpha_0^{K^t+1} \cdot \Gamma(Er_i; H_0) - D_0 \cdot \Gamma(\overline{r}; H_0) - \alpha_0^i \cdot \Gamma\left(Er_{h(i;K^t)}; H_0\right) \\
&+ \sum_{j=1}^{K^t-i-1} \left\{ \alpha_j^{K^t+1} \cdot \Gamma(Er_i; H_j) - D_j \cdot \Gamma(\overline{r}; H_j) - \left(\alpha_j^i - D_{j-1}\right) \cdot \Gamma\left(Er_{h(i;K^t-j)}; H_j\right) \right\} \\
&+ \alpha_{K^t-i}^{K^t+1} \cdot \Gamma(Er_i; H_{K^t-i}) - \left(\alpha_{K^t-i}^i - D_{K^t-1-i}\right) \cdot \Gamma\left(Er_{h(i;i)}; H_{K^t-i}\right) \\
&+ \left(\alpha_{K^t+1-i}^{K^t+1} - \alpha_{K^t+1-i}^i\right) \cdot \Gamma(\overline{r}; H_{K^t-i}),
\end{aligned}$$

where H_j is the (K^t-1)-dimensional vector $\left(Er_{h(i;1)}, ..., Er_{h(i;K^t-1-j)}, \overline{r}, ..., \overline{r}\right)$ containing j $\overline{r}$-entries.

Tedious algebra (see Supplementary online appendix) gives

$$\begin{aligned}
EG_{K^t+1} - EG_i \quad = \quad & \sum_{j=0}^{K^t-i} \left\{ \sum_{S \in S_j(1,...,K^t+1-j)} \left[\prod_{\ell \in S} p_\ell \right] \left[\prod_{\ell \in \mathcal{C}^t \backslash (S \cup \{1,...,K^t-j\})} (1-p_\ell) \right] \right\} \\
& \cdot [\Gamma(Er_i; H_j) - \left(\frac{p_i - p_{h(i;K^t-j)}}{1 - p_{h(i;K^t-j)}} \right) \cdot \Gamma(\overline{r}; H_j) - \\
& \qquad \left(\frac{1-p_i}{1-p_{h(i;K^t-j)}} \right) \cdot \Gamma\left(Er_{h(i;K^t-j)}; H_j\right)]. \qquad (5)
\end{aligned}$$

We then have $EG_{K^t+1} > EG_i$ since $p_n \in (0,1)$ for all $n \in \mathcal{N}$, $\Gamma(\cdot; H_j)$ is strictly concave, and $Er_i = \frac{p_i - p_k}{1-p_k}\overline{r} + \frac{1-p_i}{1-p_k}Er_k$ for all $k > i$. ■

Lemma 2. *Consider round t at which $i^{t+} < M$ and $i^{t-} < N - M$. Let $K^t \equiv M - i^{t+}$. For $K^t \leq (N-t-1)$, we have that*

$$E\varepsilon_{K^t+1} < ... < E\varepsilon_{N+1-t}.$$

Proof of Lemma 2. Pick $i \in \{K^t+1, ..., N-t\}$. I shall show that $E\varepsilon_i < E\varepsilon_{i+1}$.

Observe that $i^{t+} + i^{t-} = t - 1$ and $K^t \leq (N-t-1)$ imply $i^{t-} \leq N - M - 2$. Let $E\varepsilon_h(r_i, r_{i+1})$ denote the expected search cost when the continuation search sequence starts at $s^t = h \in \{i, i+1\}$ and districts i's and $(i+1)$'s valuations are r_i and r_{i+1}, respectively. Relabel districts in $\mathcal{C}^{t+1}$ such that $\mathcal{C}^{t+1} = \{1, ..., N-t\}$, with $p_1 > ... > p_{N-t}$. There are two types of cases to consider.

$\underline{(1)\ r_i = r_{i+1} \equiv r \in \{\underline{r}, \overline{r}\}.}$ We have

$$\begin{cases} i^{(t+1)+} = i^{t+} & \&\ i^{(t+1)-} = i^{t-} + 1 \quad \text{if } r = \underline{r} \\ i^{(t+1)+} = i^{t+} + 1 & \&\ i^{(t+1)-} = i^{t-} \quad \text{if } r = \overline{r} \end{cases}$$

whether $s^t = i$ or $s^t = i+1$. It follows that $s^{t+1} \in \mathcal{C}^{t+1} \cup \{\emptyset\}$ is the same whether $s^t = i$ or $s^t = i+1$. Moreover, the valuations profile of districts in $\mathcal{C}^{t+1}$, $(r_1, ..., r_{i-1}, r, r_{i+1}, ..., r_{N-t})$, is the same whether $s^t = i$ or $s^t = i+1$. It follows that for all $k = 2, ..., N-t$, we have s^{t+k} is the same whether $s^t = i$ or $s^t = i+1$. We then have $E\varepsilon_i(r, r) = E\varepsilon_{i+1}(r, r)$.

$\underline{(2)\ r_i \neq r_{i+1}.}$ I start with the case where $r_i = \overline{r}$ and $r_{i+1} = \underline{r}$. Observe that

$$\begin{cases} i^{(t+1)+} = i^{t+} + 1 & \& \ i^{(t+1)-} = i^{t-} & \text{if } s^t = i \\ i^{(t+1)+} = i^{t+} & \& \ i^{(t+1)-} = i^{t-} + 1 & \text{if } s^t = i+1 \end{cases}$$

and the valuations profile of districts in $\mathcal{C}^{t+1}$ is

$$\begin{cases} (r_1, ..., r_{i-1}, \underline{r}, r_{i+1}, ..., r_{N-t}) & \text{if } s^t = i \\ (r_1, ..., r_{i-1}, \overline{r}, r_{i+1}, ..., r_{N-t}) & \text{if } s^t = i+1. \end{cases}$$

We can make three observations. First, let $s^{t+1,h}$ be the search decision at round $t+1$ following $s^t = h$. We have $s^{t+1,i} \in \{\emptyset, K^t\}$ (with $s^{t+1,i} = \emptyset$ if and only if $K^t = 1$) while $s^{t+1,i+1} = K^t + 1$ in $\mathcal{C}^{t+1}$. Second, IG needs one less $\overline{r}$-signal to reach $i^+ = M$ following $s^t = i$ than following $s^t = i+1$. Finally, since $i > K^t$ IG will never reache $i^- = N - M$ before $s^{t+k,i} = i+1$ and $s^{t+h,i+1} = i$ for some $h, k \geq 1$.

I am going to show that the search cost following $s^t = i+1$ is at least as large as the one following $s^t = i$ for every valuations profile of districts in $\mathcal{C}^t$ where $r^i = \overline{r}$ and $r^{i+1} = \underline{r}$, and is strictly larger for some profiles. Pick one such profile. There are two possible cases.

Either $s^{t+k,i} = i+1$ for some $k \geq 1$, that is, the continuation search sequence following $s^t = i$ reaches district $i+1$. Given that $i^{(t+1)+}$ following $s^t = i+1$ is one below $i^{(t+1)+}$ following $s^t = i$ and that IG does not reach $i^- = N - M$ before having searched both districts i and $i+1$, we have $s^{t+h,i+1} = i$ for some $h \geq 1$, that is, the continuation search sequence following $s^t = i+1$ reaches district i. As both continuation search sequences involve searching both districts i and $i+1$, the set of districts IG searches is the same whether $s^t = i$ or $s^t = i+1$. For those valuations profiles, the search cost is thus the same whether $s^t = i$ or $s^t = i+1$.

Or $s^{t+k,i} \neq i+1$ for all $k \geq 1$, that is, the continuation search sequence following $s^t = i$ does not reach district $i+1$. This means that there is some $h \geq 1$ at which $i^{(t+h)+} = M$ following $s^t = i$ (meaning also that IG does not stop searching because

i^- reaches $N-M$). Since IG needs one less $\overline{r}$-signal to reach $i^+ = M$ following $s^t = i$ than following $s^t = i+1$, we get that the continuation search sequence involves searching at least one district less when $s^t = i$ than when $s^t = i+1$. For those valuations profiles, the search cost is thus bigger when $s^t = i+1$ than when $s^t = i$.

As these two cases exhaust all possibilities, we have that

$$E\varepsilon_{i+1}\left(\overline{r},\underline{r}\right) - E\varepsilon_{i}\left(\overline{r},\underline{r}\right) = T > 0.$$

I now continue with the case where $r_i = \underline{r}$ and $r_{i+1} = \overline{r}$. By the same argument, we get

$$E\varepsilon_{i+1}\left(\underline{r},\overline{r}\right) - E\varepsilon_{i}\left(\underline{r},\overline{r}\right) = -T < 0.$$

Profiles where $r_i = r_{i+1}$ occur with probability $[p_i \cdot p_{i+1} + (1-p_i)\cdot(1-p_{i+1})]$. Profiles where $r_i = \overline{r}$ and $r_{i+1} = \underline{r}$ occur with probability $p_i \cdot (1-p_{i+1})$. Profiles where $r_i = \underline{r}$ and $r_{i+1} = \overline{r}$ occur with probability $(1-p_i)\cdot p_{i+1}$. We then have

$$\begin{aligned} E\varepsilon_{i+1} - E\varepsilon_{i} &= [p_i \cdot p_{i+1} + (1-p_i)\cdot(1-p_{i+1})]\cdot 0 + p_i\cdot(1-p_{i+1})\cdot T \\ &\quad + (1-p_i)\cdot p_{i+1}\cdot(-T) \\ &= (p_i - p_{i+1})\cdot T > 0. \end{aligned}$$

Hence $E\varepsilon_{i+1} > E\varepsilon_i$. ■

To prove Proposition 3, I first state and prove two results. Lemma 3 establishes that all search sequences that satisfy the stopping rule stated in Proposition 1 generate the same (expected) total quantity of goods. Lemma 4 characterizes the set of districts in $\mathcal{N}$ IG searches in equilibrium.

Given valuations profile $\mathbf{r} = (r_1, ..., r_N)$, define $\mathcal{P}^{N+}(\mathbf{r}) \equiv \{n \in \mathcal{N} : r_n = \overline{r}\}$ as the set of high-valuation districts in $\mathcal{N}$. Likewise, define $\mathcal{P}^{N-}(\mathbf{r}) \equiv \{n \in \mathcal{N} : r_n = \underline{r}\}$ as the set of low-valuation districts in $\mathcal{N}$.

Lemma 3. *Consider the convex case. Let $\mathbf{r} = (r_1, ..., r_N)$ be any valuations profile of districts in $\mathcal{N}$. For a search sequence s satisfying the equilibrium stopping rule, the total quantity of goods is given by*

$$EG(s) = \begin{cases} \Gamma\left(\overline{r}, ..., \overline{r}\right) & \text{if } \#\mathcal{P}^{N+}(\mathbf{r}) \geq M \\ \Gamma\left(\underbrace{\overline{r}, ..., \overline{r}}_{\#\mathcal{P}^{N+}(\mathbf{r})}, \underbrace{\underline{r}, ..., \underline{r}}_{M-\#\mathcal{P}^{N+}(\mathbf{r})}\right) & \text{if } \#\mathcal{P}^{N+}(\mathbf{r}) \leq M-1. \end{cases}$$

PROOF OF LEMMA 3. If $\#\mathcal{P}^{N+}(\mathbf{r}) \geq M$, then there exists $t \in \{1, ..., N\}$ at which $i^{t+} \geq M$. In this case, we have $E\widetilde{r}_n = \overline{r}$ for every district $n \in \mathcal{L}(s)$, meaning $EG(s) = \Gamma(\overline{r}, ..., \overline{r})$.

If $\#\mathcal{P}^{N+}(\mathbf{r}) \leq M - 1$, then $s^t \neq \emptyset$ at every round $t \in \{1, ..., N\}$. In this case, we have $\mathcal{P}^{N+}(r) \subset \mathcal{L}(s)$ and $\mathcal{L}(s) \backslash \mathcal{P}^{N+}(r) \subseteq \mathcal{P}^{N-}(r)$, meaning

$$EG(s) = \Gamma\left(\underbrace{\overline{r}, ..., \overline{r}}_{\#\mathcal{P}^{N+}(r)}, \underbrace{\underline{r}, ..., \underline{r}}_{\#(\mathcal{L}(s)\backslash\mathcal{P}^{N+}(r))}\right). \qquad \blacksquare$$

LEMMA 4. *Consider the convex case. Let* $\mathbf{r} = (r_1, ..., r_N)$ *be any valuations profile of districts in* $\mathcal{N}$*. At any round* $t \in \{1, ..., N\}$*, relabel districts in* $\mathcal{C}^t$ *such that* $\mathcal{C}^t = \{1, ..., N + 1 - t\}$ *with* $p_1 > p_2 > ... > p_{N+1-t}$*. Pick a search sequence* s *where at any round* $t \in \{1, ..., N\}$*, we have*

$$\begin{cases} s^t \in \{1, ..., K^t\} & \textit{if } i^{t+} < M \\ s^t = \emptyset & \textit{if } i^{t+} \geq M \end{cases}$$

where $K^t \equiv \min\{M - i^{t+}, N + 1 - t\}$*. Then, the set of districts IG searches is*

$$\mathcal{I}(\mathbf{r}) = \begin{cases} \mathcal{N} & \textit{if } \#\mathcal{P}^{N+}(\mathbf{r}) < M \\ \mathcal{P}^T(\mathbf{r}) & \textit{if } \#\mathcal{P}^{N+}(\mathbf{r}) \geq M \end{cases}$$

where $T \equiv \min\{\tau : \#\mathcal{P}^{\tau+}(\mathbf{r}) \geq M\}$*.*

PROOF OF LEMMA 4. I start with a valuations profile $\mathbf{r}$ for which $\#\mathcal{P}^{N+}(\mathbf{r}) < M$. At any round $t \in \{1, ..., N\}$, we have $K^t \geq 1$, meaning $s^t \neq \emptyset$. Hence $\mathcal{I}(\mathbf{r}) = \mathcal{N}$.

I continue with a valuations profile $\mathbf{r}$ for which $\#\mathcal{P}^{N+}(\mathbf{r}) \geq M$. Pick any round $t \in \{1, ..., N\}$ at which $K^t \geq 1$. Let $C^t \equiv \{1, ..., K^t\}$ and $s^t = i \in C^t$. Then, the set of districts C^{t+1} that IG may choose to search at round $t + 1$ is

$$C^{t+1} = \begin{cases} (C^t \backslash \{i\}) \cup \{K^t + 1\} & \text{if } r_i = \underline{r} \text{ and } K^t \leq N - t \\ C^t \backslash \{i\} & \text{otherwise.} \end{cases}$$

Observe that $\#C^{t+1} \leq \#C^t$. Furthermore, $\#\mathcal{P}^{N+}(\mathbf{r}) \geq M$ implies $K^\tau = 1$ and, therefore, $\#C^\tau = 1$ at some round $\tau \in \{t, ..., N\}$.

Pick any two districts h and i with $p_h > p_i$. Suppose $s^{t'} = i$ at some round $t' \in \{1, ..., N\}$. There are two cases to consider:

1. $s^t = h$ at some round $t \in \{1, ..., t' - 1\}$.

2. $s^t \neq h$ at every round $t \in \{1, ..., t' - 1\}$. Observe that $i \in C^{t'}$ and $p_h > p_i$ imply $h \in \left(C^{t'} \cap C^{t'+1}\right)$. Since $C^t \backslash \{s^t\} \subseteq C^{t+1}$ and $\#C^\tau = 1$ at some round $\tau \in \{t' + 1, ..., N\}$, we then have $s^{t''} = h$ at some round $t'' \in \{t' + 1, ..., \tau\}$.

Since these two cases exhaust all possibilities, we then have $s^t = h$ at some round t. Hence there exists $k \in \mathcal{N}$ such that

$$\begin{cases} h \in \mathcal{I}(\mathbf{r}) & \text{for all } h \in \mathcal{N} \text{ with } p_h \geq p_k \\ h \notin \mathcal{I}(\mathbf{r}) & \text{for all } h \in \mathcal{N} \text{ with } p_h < p_k. \end{cases}$$

Given $C^1 = \{1, ..., M\}$ and the sequencing of C^t for $t = 1, ..., N$, we have $k = \min\{\tau : \#\mathcal{P}^{\tau+}(\mathbf{r}) \geq M\}$. ■

Proof of Proposition 3. The proof is inductive.

Consider round $t = N$. Proposition 1 implies $s^N = \emptyset$ if and only if $K^N = 1$. Consider round $t = N - 1$ with $K^{N-1} \geq 1$. There are two cases to consider:

1. $K^{N-1} = 2$: Since $i^{(N-1)+} \leq M - 2$, Proposition 1 implies the continuation search sequence is either $s^{N-1} = 1$ and $s^N = 2$, or $s^{N-1} = 2$ and $s^N = 1$. Both continuation sequences result in the same total quantity of goods and the same continuation search cost, implying IG is indifferent between $s^{N-1} = 1$ and $s^{N-1} = 2$. Hence $s^{N-1} \in \{1, 2\} = \mathcal{C}^{N-1}$.
2. $K^{N-1} = 1$: We know from Lemma 3 that the expected total quantity of goods is the same whether $s^{N-1} = 1$ or $s^{N-1} = 2$, that is, $EG_1 = EG_2$. The expected continuation search cost for $s^{N-1} = i \in \mathcal{C}^{N-1}$ is

$$E\varepsilon_i = p_i + 2 \cdot (1 - p_i) = 2 - p_i.$$

   Since $p_1 > p_2$, we have $E\varepsilon_1 < E\varepsilon_2$. This, together with $EG_1 = EG_2$, implies $EV_1 > EV_2$. Hence $s^{N-1} = 1$.

Assume the statement is true at round $(t + 1) \in \{2, ..., N - 1\}$. Consider round t with $K^t \geq 1$. We know from Lemma 3 that the expected total quantity of goods is the same for any continuation search sequence satisfying the equilibrium stopping rule: $EG_1 = ... = EG_{N+1-t}$.

There are two cases to consider:

1. $K^t = N + 1 - t$: Proposition 1 implies $\mathcal{I} = \mathcal{N}$, that is, IG searches all districts. IG is therefore indifferent between any $s^t \in \mathcal{C}^t$ since they all yield the same expected total quantity of goods ($EG_1 = ... = EG_{N+1-t}$) and

the same (continuation) search cost ($E\varepsilon_1 = ... = E\varepsilon_{N+1-t}$). Hence $s^t \in \{1, ..., N+1-t\} = \mathcal{C}^t$.

2. $K^t < N+1-t$: Let $s^t = h \in \mathcal{C}^t$. Define

$$\alpha_x(Y) \equiv \sum_{S \in S_x(Y)} \left[\prod_{\ell \in S} p_\ell\right] \left[\prod_{\ell \in Y \setminus S} (1-p_\ell)\right]$$

as the probability that exactly x districts in $Y \subseteq \mathcal{N}$ have a high valuation.

Using Lemma 4, we get the expected continuation search cost is

$$E\varepsilon_h = K^t \cdot \left[\prod_{\ell \in \mathcal{P}^{K^t}} p_\ell\right] + \sum_{i=K^t+1}^{N+1-t} i \cdot p_i \cdot \alpha_{K^t-1}\left(\mathcal{P}^{i-1}\right) + (N+1-t) \sum_{i=1}^{K^t} \alpha_{i-1}\left(\mathcal{C}^t\right)$$

if $h \in \{1, ..., K^t\}$, and

$$\begin{aligned} E\varepsilon_h &= K^t \cdot p_h \cdot \left[\prod_{\ell \in \mathcal{P}^{K^t-1}} p_\ell\right] + \sum_{i=K^t}^{h-1} (i+1) \cdot p_i \cdot \alpha_{K^t-1}\left(\mathcal{P}^{i-1} \cup \{h\}\right) \\ &\quad + \sum_{i=h+1}^{N+1-t} i \cdot p_i \cdot \alpha_{K^t-1}\left(\mathcal{P}^{i-1}\right) + (N+1-t) \sum_{i=1}^{K^t} \alpha_{i-1}\left(\mathcal{C}^t\right) \end{aligned}$$

if $h \in \{K^t+1, ..., N+1-t\}$. Observe that $E\varepsilon_1 = ... = E\varepsilon_{K^t}$.

Pick $k \in \{K^t, ..., N-t\}$. I show that $E\varepsilon_{k+1} > E\varepsilon_k$. Tedious algebra (see Supplementary online appendix) gives

$$\frac{E\varepsilon_{k+1} - E\varepsilon_k}{p_k - p_{k+1}} =$$

$$\begin{aligned} &\sum_{h=K^t-1}^{k-2} p_h \cdot \alpha_{K^t-2}\left(\mathcal{P}_{h-1}\right) \cdot \\ &\quad \left\{\sum_{i=h+1}^{k+1} (i-h) \cdot p_i \cdot \left[\prod_{\ell \in \mathcal{P}^{i-1} \setminus \mathcal{P}^h} (1-p_\ell)\right] + (k-h) \cdot \left[\prod_{\ell \in \mathcal{P}^{k-1} \setminus \mathcal{P}^h} (1-p_\ell)\right]\right\} \\ &+ p_{k-1} \cdot \alpha_{K^t-2}\left(\mathcal{P}^{k-2}\right). \end{aligned} \tag{6}$$

Since $p_k > p_{k+1}$ and the right-hand side is strictly positive, we get $E\varepsilon_{k+1} > E\varepsilon_k$.

To sum up, we have

$$\begin{cases} EG_1 = ... = EG_{N+1-t} \\ E\varepsilon_1 = ... = E\varepsilon_{K^t} < E\varepsilon_{K^t+1} < ... < E\varepsilon_{N+1-t} \end{cases}$$

which implies $EV_1 = ... = EV_{K^t} > EV_{K^t+1} > ... > EV_{N+1-t}$. Hence $s^t \in \{1, ..., K^t\}$. ■

SUPPLEMENTARY ONLINE APPENDIX

DERIVATION OF (5)

Equation (5) is obtained by rearranging

$$\begin{aligned}
&\alpha_0^{K^t+1} \cdot \Gamma\left(Er_i; H_0\right) - D_0 \cdot \Gamma\left(\overline{r}; H_0\right) - \alpha_0^i \cdot \Gamma\left(Er_{h(i;K^t)}; H_0\right) \\
&+ \sum_{j=1}^{K^t-i-1} \left\{\alpha_j^{K^t+1} \cdot \Gamma\left(Er_i; H_j\right) - D_j \cdot \Gamma\left(\overline{r}; H_j\right) - \left(\alpha_j^i - D_{j-1}\right) \cdot \Gamma\left(Er_{h(i;K^t-j)}; H_j\right)\right\} \\
&+\alpha_{K^t-i}^{K^t+1} \cdot \Gamma\left(Er_i; H_{K^t-i}\right) - \left(\alpha_{K^t-i}^i - D_{K^t-1-i}\right) \cdot \Gamma\left(Er_{h(i;i)}; H_{K^t-i}\right) \\
&+ \left(\alpha_{K^t+1-i}^{K^t+1} - \alpha_{K^t+1-i}^i\right) \cdot \Gamma\left(\overline{r}; H_{K^t-i}\right).
\end{aligned}$$

I decompose this expression into three parts:

$(*)$ $\alpha_0^{K^t+1} \cdot \Gamma\left(Er_i; H_0\right) - D_0 \cdot \Gamma\left(\overline{r}; H_0\right) - \alpha_0^i \cdot \Gamma\left(Er_{h(i;K^t)}; H_0\right)$.

$(**)$ $\sum_{j=1}^{K^t-i-1} \left\{\alpha_j^{K^t+1} \cdot \Gamma\left(Er_i; H_j\right) - D_j \cdot \Gamma\left(\overline{r}; H_j\right) - \left(\alpha_j^i - D_{j-1}\right) \cdot \Gamma\left(Er_{h(i;K^t-j)}; H_j\right)\right\}$.

$(***)$ $\alpha_{K^t-i}^{K^t+1} \cdot \Gamma\left(Er_i; H_{K^t-i}\right) - \left(\alpha_{K^t-i}^i - D_{K^t-1-i}\right) \cdot \Gamma\left(Er_{h(i;i)}; H_{K^t-i}\right) + \left(\alpha_{K^t+1-i}^{K^t+1} - \alpha_{K^t+1-i}^i\right) \cdot \Gamma\left(\overline{r}; H_{K^t-i}\right)$.

I first rearrange $(*)$:

$$\begin{aligned}
&\left[\prod_{\ell \in \mathcal{C}^t \setminus \{1,\ldots,K^t\}} \left(1-p_\ell\right)\right] \cdot \Gamma\left(Er_i; H_0\right) \\
&- \left(p_i - p_{K^t+1}\right) \cdot \left[\prod_{\ell \in \mathcal{C}^t \setminus \{1,\ldots,K^{t+1}\}} \left(1-p_\ell\right)\right] \cdot \Gamma\left(\overline{r}; H_0\right) \\
&- \left[\prod_{\ell \in \mathcal{C}^t \setminus \{1,\ldots,i-1,i+1,\ldots,K^t+1\}} \left(1-p_\ell\right)\right] \cdot \Gamma\left(Er_{h(i;K^t)}; H_0\right)
\end{aligned}$$

$$= \left[\prod_{\ell \in \mathcal{C}^t \setminus \{1,\ldots,K^t\}} \left(1-p_\ell\right)\right] \cdot \left\{\Gamma\left(Er_i; H_0\right) - \frac{p_i - p_{K^t+1}}{1-p_{K^t+1}} \cdot \Gamma\left(\overline{r}; H_0\right) - \frac{1-p_i}{1-p_{K^t+1}} \cdot \Gamma\left(Er_{h(i;K^t)}; H_0\right)\right\}$$

$$\begin{aligned}
= \quad &\left[\prod_{\ell \in \mathcal{C}^t \setminus \{1,\ldots,K^t\}} \left(1-p_\ell\right)\right] \cdot \left\{\Gamma\left(Er_i; H_0\right) - \frac{p_i - p_{h(i;K^t)}}{1-p_{h(i;K^t)}} \cdot \Gamma\left(\overline{r}; H_0\right)\right. \\
&\qquad\qquad \left. - \frac{1-p_i}{1-p_{h(i;K^t)}} \cdot \Gamma\left(Er_{h(i;K^t)}; H_0\right)\right\}.
\end{aligned}$$

I now rearrange $(**)$. I first rewrite $\alpha_j^i - D_{j-1}$ as

$$\sum_{S\in S_j(1,\dots,i-1,i+1,\dots,K^t+2-j)} \left[\prod_{\ell\in S} p_\ell\right]\left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\dots,i-1,i+1,\dots,K^t+1-j\})}(1-p_\ell)\right]$$
$$-(p_i - p_{K^t+1-j}) \sum_{S\in S_{j-i}(1,\dots,K^t+2-j)} \left[\prod_{\ell\in S} p_\ell\right]\left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\dots,K^t+2-j\})}(1-p_\ell)\right]$$

$$= p_i\cdot(1-p_{K^t+2-j}) \sum_{S\in S_{j-1}(1,\dots,K^t+2-j)} \left[\prod_{\ell\in S} p_\ell\right]\left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\dots,K^t+2-j\})}(1-p_\ell)\right]$$
$$+(1-p_i)(1-p_{K^t+2-j}) \sum_{S\in S_j(1,\dots,K^t+2-j)} \left[\prod_{\ell\in S} p_\ell\right]\left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\dots,K^t+2-j\})}(1-p_\ell)\right]$$
$$-(p_i - p_{K^t+2-j}) \sum_{S\in S_{j-1}(1,\dots,K^t+2-j)} \left[\prod_{\ell\in S} p_\ell\right]\left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\dots,K^t+2-j\})}(1-p_\ell)\right]$$

$$= (1-p_i)\cdot p_{K^t+2-j} \sum_{S\in S_{j-1}(1,\dots,K^t+2-j)} \left[\prod_{\ell\in S} p_\ell\right]\left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\dots,K^t+2-j\})}(1-p_\ell)\right]$$
$$+(1-p_i)(1-p_{K^t+2-j}) \sum_{S\in S_j(1,\dots,K^t+2-j)} \left[\prod_{\ell\in S} p_\ell\right]\left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\dots,K^t+2-j\})}(1-p_\ell)\right]$$
$$= \frac{1-p_i}{1-p_{K^t+1-j}} \sum_{S\in S_j(1,\dots,K^t+1-j)} \left[\prod_{\ell\in S} p_\ell\right]\left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\dots,K^t-j\})}(1-p_\ell)\right].$$

The term in curly brackets in $(**)$ can then be rewritten as

$$\left\{\sum_{S\in S_j(1,\dots,K^t+1-j)} \left[\prod_{\ell\in S} p_\ell\right]\left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\dots,K^t-j\})}(1-p_\ell)\right]\right\}\cdot\Gamma\left(Er_i;H_j\right)$$
$$-\frac{p_i - p_{K^t+1-j}}{1-p_{K^t+1-j}}\cdot\left\{\sum_{S\in S_j(1,\dots,K^t+1-j)} \left[\prod_{\ell\in S} p_\ell\right]\left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\dots,K^t-j\})}(1-p_\ell)\right]\right\}\cdot\Gamma\left(\overline{r};H_j\right)$$
$$-\frac{1-p_i}{1-p_{K^t+1-j}}\cdot\left\{\sum_{S\in S_j(1,\dots,K^t+1-j)} \left[\prod_{\ell\in S} p_\ell\right]\left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\dots,K^t-j\})}(1-p_\ell)\right]\right\}$$
$$\cdot\Gamma\left(Er_{h(i;K^t-j)};H_j\right)$$

$$= \left\{\sum_{S\in S_j(1,\dots,K^t+1-j)} \left[\prod_{\ell\in S} p_\ell\right]\left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\dots,K^t-j\})}(1-p_\ell)\right]\right\}$$
$$\left\{\Gamma\left(Er_i;H_j\right) - \frac{p_i - p_{K^t+1-j}}{1-p_{K^t+1-j}}\cdot\Gamma\left(\overline{r};H_j\right) - \frac{1-p_i}{1-p_{K^t+1-j}}\cdot\Gamma\left(Er_{h(i;K^t-j)};H_j\right).\right\}.$$

It remains to rearrange $(***)$. I first rewrite $\alpha^i_{K^t-i} - D_{K^t-1-i}$ as

$$\sum_{S\in S_{K^t-i}(1,\ldots,i-1,i+1,i+2)} \left[\prod_{\ell\in S} p_\ell\right] \left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\ldots,i-1,i+2\})} (1-p_\ell)\right]$$
$$-(p_i - p_{i+2}) \cdot \sum_{S\in S_{K^t-i-1}(1,\ldots,i+2)} \left[\prod_{\ell\in S} p_\ell\right] \left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\ldots,i+2\})} (1-p_\ell)\right]$$

$$= \; p_i \cdot (1-p_{i+2}) \sum_{S\in S_{K^t-i-1}(1,\ldots,i+2)} \left[\prod_{\ell\in S} p_\ell\right] \left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\ldots,i+2\})} (1-p_\ell)\right]$$
$$+(1-p_i)\cdot(1-p_{i+2}) \sum_{S\in S_{K^t-i}(1,\ldots,i+2)} \left[\prod_{\ell\in S} p_\ell\right] \left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\ldots,i+2\})} (1-p_\ell)\right]$$
$$-(p_i - p_{i+2}) \cdot \sum_{S\in S_{K^t-i-1}(1,\ldots,i+2)} \left[\prod_{\ell\in S} p_\ell\right] \left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\ldots,i+2\})} (1-p_\ell)\right]$$

$$= \; (1-p_i)\cdot p_{i+2} \sum_{S\in S_{K^t-i-1}(1,\ldots,i+2)} \left[\prod_{\ell\in S} p_\ell\right] \left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\ldots,i+2\})} (1-p_\ell)\right]$$
$$+(1-p_i)(1-p_{i+2}) \sum_{S\in S_{K^t-i}(1,\ldots,i+2)} \left[\prod_{\ell\in S} p_\ell\right] \left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\ldots,i+2\})} (1-p_\ell)\right]$$

$$=(1-p_i) \sum_{S\in S_{K^t-i}(1,\ldots,i+1)} \left[\prod_{\ell\in S} p_\ell\right] \left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\ldots,i+1\})} (1-p_\ell)\right].$$

I now rewrite $\alpha^{K^t+1}_{K^t+1-i} - \alpha^i_{K^t+1-i}$ as

$$\sum_{S\in S_{K^t+1-i}(1,\ldots,i)} \left[\prod_{\ell\in S} p_\ell\right] \left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\ldots,i-1\})} (1-p_\ell)\right]$$
$$-\sum_{S\in S_{K^t+1-i}(1,\ldots,i-1,i+1)} \left[\prod_{\ell\in S} p_\ell\right] \left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\ldots,i-1\})} (1-p_\ell)\right]$$

$$
\begin{aligned}
= \quad & (1-p_i)\Bigg\{p_{i+1}\sum_{S\in S_{K^t-i}(1,\dots,i+1)}\left[\prod_{\ell\in S}p_\ell\right]\left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\dots,i+1\})}(1-p_\ell)\right] \\
& +(1-p_{i+1})\sum_{S\in S_{K^t+1-i}(1,\dots,i+1)}\left[\prod_{\ell\in S}p_\ell\right]\left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\dots,i+1\})}(1-p_\ell)\right]\Bigg\} \\
& -(1-p_{i+1})\Bigg\{p_i\sum_{S\in S_{K^t-i}(1,\dots,i+1)}\left[\prod_{\ell\in S}p_\ell\right]\left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\dots,i+1\})}(1-p_\ell)\right] \\
& +(1-p_i)\sum_{S\in S_{K^t+1-i}(1,\dots,i+1)}\left[\prod_{\ell\in S}p_\ell\right]\left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\dots,i+1\})}(1-p_\ell)\right]\Bigg\} \\
= & -(p_i-p_{i+1})\sum_{S\in S_{K^t-i}(1,\dots,i+1)}\left[\prod_{\ell\in S}p_\ell\right]\left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\dots,i+1\})}(1-p_\ell)\right].
\end{aligned}
$$

Part $(***)$ can thus be rewritten as

$$
\begin{aligned}
& \left\{\sum_{S\in S_{K^t-i}(1,\dots,i+1)}\left[\prod_{\ell\in S}p_\ell\right]\left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\dots,i\})}(1-p_\ell)\right]\right\}\cdot\Gamma\left(Er_i;H_{K^t-i}\right) \\
& -\frac{1-p_i}{1-p_{i+1}}\left\{\sum_{S\in S_{K^t-i}(1,\dots,i+1)}\left[\prod_{\ell\in S}p_\ell\right]\left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\dots,i\})}(1-p_\ell)\right]\right\}\cdot\Gamma\left(Er_{h(i;i)};H_{K^t-i}\right) \\
& -\frac{p_i-p_{i+1}}{1-p_{i+1}}\left\{\sum_{S\in S_{K^t-i}(1,\dots,i+1)}\left[\prod_{\ell\in S}p_\ell\right]\left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\dots,i\})}(1-p_\ell)\right]\right\}\cdot\Gamma\left(\overline{r};H_{K^t-i}\right)
\end{aligned}
$$

$$
\begin{aligned}
= \quad & \left\{\sum_{S\in S_{K^t-i}(1,\dots,i+1)}\left[\prod_{\ell\in S}p_\ell\right]\left[\prod_{\ell\in\mathcal{C}^t\setminus(S\cup\{1,\dots,i\})}(1-p_\ell)\right]\right\}\cdot \\
& \cdot\left[\Gamma\left(Er_i;H_{K^t-i}\right)-\frac{p_i-p_{h(i;i)}}{1-p_{h(i;i)}}\cdot\Gamma\left(\overline{r};H_{K^t-i}\right)-\frac{1-p_i}{1-p_{h(i;i)}}\cdot\Gamma\left(Er_{h(i;i)};H_{K^t-i}\right)\right].
\end{aligned}
$$

DERIVATION OF (6)

We have that $E\varepsilon_{k+1}-E\varepsilon_k$ is equal to

$$
\begin{aligned}
& K^t\cdot p_{k+1}\cdot\left[\prod_{\ell\in\mathcal{P}^{K^t-1}}p_\ell\right]+\sum_{i=K^t}^{k}(i+1)\cdot p_i\cdot\alpha_{K^t-1}\left(\mathcal{P}^{i-1}\cup\{k+1\}\right) \\
& +\sum_{i=k+2}^{N+1-t}i\cdot p_i\cdot\alpha_{K^t-1}\left(\mathcal{P}^{i-1}\right)-K^t\cdot p_k\cdot\left[\prod_{\ell\in\mathcal{P}^{K^t-1}}p_\ell\right]-\sum_{i=K^t}^{k-1}(i+1)\cdot p_i\cdot\alpha_{K^t-1}\left(\mathcal{P}^{i-1}\cup\{k\}\right) \\
& -\sum_{i=k+1}^{N+1-t}i\cdot p_i\cdot\alpha_{K^t-1}\left(\mathcal{P}^{i-1}\right).
\end{aligned}
$$

$$
\begin{aligned}
= \quad & K^t \cdot (p_{k+1} - p_k) \cdot \left[\prod_{\ell \in \mathcal{P}^{K^t-1}} p_\ell \right] + (k+1) \cdot p_k \cdot \alpha_{K^t-1}\left(\mathcal{P}^{k-1} \cup \{k+1\}\right) \\
& + \sum_{i=K^t}^{k-1} (i-1) \cdot p_i \cdot \left[p_{k+1} \cdot \alpha_{K^t-2}\left(\mathcal{P}^{i-1}\right) + (1-p_{k+1}) \cdot \alpha_{K^t-1}\left(\mathcal{P}^{i-1}\right) \right] \\
& - \sum_{i=K^t}^{k-1} (i-1) \cdot p_i \cdot \left[p_k \cdot \alpha_{K^t-2}\left(\mathcal{P}^{i-1}\right) + (1-p_k) \cdot \alpha_{K^t-1}\left(\mathcal{P}^{i-1}\right) \right] \\
& - (k+1) \cdot p_{k+1} \cdot \alpha_{K^t-1}\left(\mathcal{P}^k\right).
\end{aligned}
$$

$$
\begin{aligned}
= \quad & K^t \cdot (p_{k+1} - p_k) \cdot \left[\prod_{\ell \in \mathcal{P}^{K^t-1}} p_\ell \right] + (p_{k+1} - p_k) \cdot \sum_{i=K^t}^{k-1} (i+1) \cdot p_i \cdot \alpha_{K^t-2}\left(\mathcal{P}^{i-1}\right) \\
& - (p_k - p_{k+1}) \cdot \sum_{i=K^t}^{k-1} (i+1) \cdot p_i \cdot \alpha_{K^t-1}\left(\mathcal{P}^{i-1}\right) + (k+1) \cdot (p_k - p_{k+1}) \cdot \alpha_{K^t-1}\left(\mathcal{P}^{k-1}\right).
\end{aligned}
$$

Dividing both sides by $(p_k - p_{k+1})$ and rearranging the right hand side, we get that $\frac{E\varepsilon_{k+1} - E\varepsilon_k}{p_k - p_{k+1}}$ is equal to

$$
\begin{aligned}
& -K^t \cdot \left[\prod_{\ell \in \mathcal{P}^{K^t-1}} p_\ell \right] - \sum_{i=K^t}^{k-1} (i+1) \cdot p_i \cdot \alpha_{K^t-2}\left(\mathcal{P}^{i-1}\right) \\
& + \sum_{i=K^t}^{k-1} (i+1) \cdot p_i \cdot \alpha_{K^t-1}\left(\mathcal{P}^{i-1}\right) + (k+1) \cdot \alpha_{K^t-1}\left(\mathcal{P}^{k-1}\right).
\end{aligned}
$$

$$
\begin{aligned}
= \quad & \left[\prod_{\ell \in \mathcal{P}^{K^t-1}} p_\ell \right] \Bigg\{ -K^t + \sum_{i=K^t}^{k-1} (i+1) \cdot p_i \cdot \left[\prod_{\ell \in \mathcal{P}^{i-1} \setminus \mathcal{P}^{K^t-1}} (1-p_\ell) \right] \\
& \qquad + (k+1) \left[\prod_{\ell \in \mathcal{P}^{k-1} \setminus \mathcal{P}^{K^t-1}} (1-p_\ell) \right] \Bigg\} \\
& + p_{K^t} \cdot \alpha_{K^t-2}\left(\mathcal{P}^{K^t-1}\right) \Bigg\{ -\left(K^t+1\right) + \sum_{i=K^t+1}^{k-1} (i+1) \cdot p_i \cdot \left[\prod_{\ell \in \mathcal{P}^{i-1} \setminus \mathcal{P}^{K^t}} (1-p_\ell) \right] \\
& \qquad + (k+1) \left[\prod_{\ell \in \mathcal{P}^{k-1} \setminus \mathcal{P}^{K^t}} (1-p_\ell) \right] \Bigg\} \\
& + p_{K^t+1} \cdot \alpha_{K^t-2}\left(\mathcal{P}^{K^t}\right) \Bigg\{ -\left(K^t+2\right) + \sum_{i=K^t+2}^{k-1} (i+1) \cdot p_i \cdot \left[\prod_{\ell \in \mathcal{P}^{i-1} \setminus \mathcal{P}^{K^t+1}} (1-p_\ell) \right] \\
& \qquad + (k+1) \left[\prod_{\ell \in \mathcal{P}^{k-1} \setminus \mathcal{P}^{K^t+1}} (1-p_\ell) \right] \Bigg\} \\
& + \ldots + p_{k-1} \cdot \alpha_{K^t-2}\left(\mathcal{P}^{k-2}\right) \cdot \left[-k + (k+1)\right].
\end{aligned}
$$

$$
\begin{aligned}
= & \left[\prod_{\ell\in\mathcal{P}^{K^t-1}} p_\ell\right]\left\{\sum_{i=K^t}^{k-1}\left(i+1-K^t\right)\cdot p_i\cdot\left[\prod_{\ell\in\mathcal{P}^{i-1}\setminus\mathcal{P}^{K^t-1}}(1-p_\ell)\right]\right.\\
& \qquad\left.+\left(k+1-K^t\right)\left[\prod_{\ell\in\mathcal{P}^{k-1}\setminus\mathcal{P}^{K^t-1}}(1-p_\ell)\right]\right\}\\
& +p_{K^t}\cdot\alpha_{K^t-2}\left(\mathcal{P}^{K^t-1}\right)\cdot\left\{\sum_{i=K^t+1}^{k-1}\left(i-K^t\right)\cdot p_i\cdot\left[\prod_{\ell\in\mathcal{P}^{i-1}\setminus\mathcal{P}^{K^t}}(1-p_\ell)\right]\right.\\
& \qquad\left.+\left(k-K^t\right)\left[\prod_{\ell\in\mathcal{P}^{k-1}\setminus\mathcal{P}^{K^t}}(1-p_\ell)\right]\right\}\\
& +p_{K^t+1}\cdot\alpha_{K^t-2}\left(\mathcal{P}^{K^t}\right)\cdot\left\{\sum_{i=K^t+2}^{k-1}\left(i-K^t-1\right)\cdot p_i\cdot\left[\prod_{\ell\in\mathcal{P}^{i-1}\setminus\mathcal{P}^{K^t+1}}(1-p_\ell)\right]\right.\\
& \qquad\left.+\left(k-K^t-1\right)\left[\prod_{\ell\in\mathcal{P}^{k-1}\setminus\mathcal{P}^{K^t+1}}(1-p_\ell)\right]\right\}\\
& +...+p_{k-1}\cdot\alpha_{K^t-2}\left(\mathcal{P}^{k-2}\right).
\end{aligned}
$$

$$
\begin{aligned}
= & \sum_{h=K^t-1}^{k-2} p_k\cdot\alpha_{K^t-2}\left(\mathcal{P}^{h-1}\right)\left\{\sum_{i=h+1}^{k-1}(i-h)\cdot p_i\cdot\left[\prod_{\ell\in\mathcal{P}^{i-1}\setminus\mathcal{P}^{h}}(1-p_\ell)\right]\right.\\
& \qquad\left.+(k-h)\left[\prod_{\ell\in\mathcal{P}^{k-1}\setminus\mathcal{P}^{h}}(1-p_\ell)\right]\right\}+p_{k-1}\cdot\alpha_{K^t-2}\left(\mathcal{P}^{k-2}\right).
\end{aligned}
$$

HETEROGENEOUS INFORMATION QUALITY

I consider an extension of the model in which districts vary in the probability IG receives a signal.

If IG searches district $n\in\mathcal{N}_0$, it receives a signal $\sigma_n\in\{\underline{r},\overline{r},\emptyset\}$. The signal reveals district n's valuation ($\sigma_n=r_n\in\{\underline{r},\overline{r}\}$) with probability $q_n\in(0,1]$, and does not convey any information about r_n ($\sigma_n=\emptyset$) with probability $1-q_n$. For simplicity, I assume IG can search a district at most once or would receive the same signal if it were to search a district multiple times. I label districts such that $q_1>q_2>...>q_N$ and let $p_n=p\in(0,1)$ for every district $n\in\mathcal{N}$, that is, every district faces the same prospects of high valuation. We then have $Er_n=Er$ for every $n\in\mathcal{N}$.

In order to shorten the exposition, I shall consider only the case where the goods function $\Gamma(\cdot)$ is strictly concave in each of its arguments.

As in the case considered in the paper (hereafter, p-case), IG never searches AS's district in the case considered here (hereafter, q-case).

Also, the equilibrium stopping rule and the intuition underlying it are the same as in the p-case, that is, IG searches districts until it has received either M favorable signals or $N-M$ unfavorable signals (Proposition 1). Formally, we have:[39]

Proposition 4. *Suppose $q_1 > q_2 > ... > q_N$ and $p_n = p \in (0,1)$ for every district $n \in \mathcal{N}$. At round $t \in \{1,...,N\}$ and search history h^t, we have that*

$$s^t\left(h^t\right) = \emptyset \text{ if and only if } i^{t+}\left(h^t\right) \geq M \text{ or } i^{t-}\left(h^t\right) \geq N-M.$$

It remains to characterize the equilibrium search sequence. Let

$$K^t \equiv \min\left\{M - i^{t+}\left(h^t\right), N-M-i^{t-}\left(h^t\right), N+1-t\right\}$$

be the minimum number of districts IG will search from round t on. This number is equal to the minimum of: (i) $M - i^{t+}(h^t)$, which is the number of additional favorable signals that would trigger IG to stop searching because it has received M favorable signals; (ii) $N-M-i^{t-}(h^t)$, which is the number of additional unfavorable signals that would trigger IG to stop searching because it has received $N-M$ unfavorable signals; and (iii) $N+1-t$, which is the number of yet unsearched districts in $\mathcal{N}$, that is, $\#\mathcal{C}^t(h^t)$.

Proposition 5. *Suppose $q_1 > q_2 > ... > q_N$ and $p_n = p \in (0,1)$ for every district $n \in \mathcal{N}$. Consider a round $t \in \{1,...,N\}$ and search history h^t with $K^t \geq 1$. Relabelling districts in $\mathcal{C}^t(h^t)$ such that $\mathcal{C}^t(h^t) = \{1,...,N+1-t\}$ with $q_1 > ... > q_{N+1-t}$, we have*

$$s^t\left(h^t\right) \in \left\{1,...,K^t\right\} \subseteq \mathcal{C}^t\left(h^t\right).$$

The proof of Proposition 5 appears in the next section.

At each round t, IG is indifferent searching any district in the relabelled set $\{1,...,K^t\}$, that is, any of the yet unsearched districts with one of the K^t highest q_ns. Thus, IG starts by searching one district among those with the min $\{M, N-M\}$

[39] I omit here the proof since it is very similar to the proof of Proposition 1. It is however available upon request.

highest q_ns. At each round t where IG receives a favorable signal (resp. unfavorable signal) when $M - i^{t+}(h^t) > N - M - i^{t-}(h^t)$ (resp. $M - i^{t+}(h^t) < N - M - i^{t-}(h^t)$), IG chooses at round $t+1$ to search one among the yet unsearched districts in the round-t-relabelled set $\{1, ..., K^t, K^t + 1\} \backslash s^t(h^t)$. Conversely, at each round t where IG receives a favorable signal (resp. unfavorable signal) when $M - i^{t+}(h^t) \leq N - M - i^{t-}(h^t)$ (resp. $M - i^{t+}(h^t) \geq N - M - i^{t-}(h^t)$), IG chooses at round $t+1$ to search one among the yet unsearched districts in the round-t-relabelled set $\{1, ..., K^t\} \backslash s^t(h^t)$. This process continues until IG has received M favorable signals or $N - M$ unfavorable signals or, if neither of the two happens by round N, until all districts have been searched.

The intuition underlying Proposition 5 runs as follows. Given that the ex ante expected valuation Er is the same for all districts in $\mathcal{N}$ and that the probability q_n of receiving a signal when IG searches district $n \in \mathcal{N}$ is independent of the district's valuation r_n, any search sequence satisfying the equilibrium stopping rule stated in Proposition 4 yields the same expected total quantity of goods (Lemma 5 in the proof of Proposition 5). At the same time, the costly search causes IG to follow a search sequence that minimizes the number of districts it searches. This induces IG to search districts for which the probability of receiving a signal is the highest. At each round t IG anticipates it will ultimately search the K^t districts with highest q_ns before it will stop searching, which explains why IG is indifferent searching at round t any of the districts in the relabelled set $\{1, ..., K^t\}$.

The following example illustrates IG's equilibrium search strategy.

Example 7. Consider a country with five districts ($\mathcal{N}_0 = \{0, 1, 2, 3, 4\}$) where the legislature takes its decisions by simple majority ($M = 2$, implying $N - M = 2$). Suppose the realized profile of valuations for districts in $\mathcal{N} = \{1, 2, 3, 4\}$ is $(\overline{r}, \underline{r}, \overline{r}, \underline{r})$.

I start with a situation where a signal is received for any district that IG chooses to search, that is, the profile of potential signal realisations is given by $(\overline{r}, \underline{r}, \overline{r}, \underline{r})$. Following Propositions 4 and 5, IG is initially indifferent searching district 1 or district 2. Without loss of generality, suppose IG searches district 1. Receiving a favorable signal, IG chooses to search district 2, from which it receives an unfavorable signal, and then moves to search district 3. Receiving a second favorable signal, IG stops searching. Thus, IG searches districts 1, 2 and 3, AS forms a legislative coalition $\mathcal{L}^* = \{1, 3\}$, and the total quantity of goods is $G^* = \Gamma(Er_0, \overline{r}, \overline{r})$.[40]

I continue with a situation where the profile of potential signal realisations is $(\emptyset, \underline{r}, \overline{r}, \underline{r})$, that is, searching district 1 yields no signal while searching district $n \in \{2, 3, 4\}$ yields a signal r_n. IG is again indifferent starting with district 1 or district 2. As in the previous case, suppose IG searches district 1. Receiving no signal, at round 2 IG is indifferent searching district 2 or district 3. Without loss of generality, suppose IG searches district 2. IG then receives an unfavorable signal and, at round 3, searches district 3. This time, IG receives a favorable signal and then moves to search district 4, for which it receives an unfavorable signal. AS forms a legislative coalition $\mathcal{L}^* = \{1, 3\}$, and the total quantity of goods is $G^* = \Gamma(Er_0, Er, \overline{r})$. $\square$

I conclude by underlining two interesting differences between the p-case and the q-case. First, there is a unique equilibrium search sequence in the p-case, while there are multiple equilibrium search sequences in the q-case. However, it is important to note that in the q-case, all equilibrium search sequences generate the same set of searched districts (Lemma 6 in the proof of Proposition 5) and the same total quantity of goods (Lemma 5 in the proof of Proposition 5), meaning the equilibrium outcome is unique. Second, in the p-case IG starts with district $M + 1$ and then moves non-monotonically towards districts 1 and N. By contrast, in the q-case an equilibrium sequence exists where IG starts with district 1 and then moves monotonically towards district N. More generally, given a profile of potential signal realisations and a majority requirement $M \in \{1, ..., N - 1\}$, in the q-case there exists $\eta \in \{1, ..., N\}$ such that for any equilibrium search sequence, the set of searched districts is given by $\{1, ..., \eta\}$ (Lemma 6 in the proof of Proposition 5).[41]

[40] If instead IG searches district 2 at round 1, it will receive an unfavorable signal and move to search district 1 followed by district 3. Thus, although the order in which districts are searched is different, the set of searched districts as well as the legislative coalition and the total quantity of goods are the same as when IG starts its search process with district 1.

[41] Observe that for $M = 0$ or $M = N$, Propositions 1 and 4 imply that the equilibrium set of searched districts is empty in both the p-case and the q-case.

PROOF OF PROPOSITION 5

I first state and prove two results. Lemma 5 establishes that all search sequences satisfying the equilibrium stopping rule generate the same expected total quantity of goods. Lemma 6 characterizes the set of districts IG searches in equilibrium.

Lemma 5. *For any two search sequences, s and s', that satisfy the condition stated in Proposition 4, we have $EG(s) = EG(s')$.*

Proof of Lemma 5. Pick any sequence η that orders all elements in $\mathcal{N}$. Construct a sequence η' such that

$$\eta^{\tau\prime} = \begin{cases} \eta^{\tau+1} & \text{if } \tau = t \\ \eta^{\tau-1} & \text{if } \tau = t+1 \\ \eta^{\tau} & \text{if } \tau \neq t, t+1 \end{cases}$$

for some $t \in \{1, ..., N-1\}$, that is, η' is obtained from η by switching two consecutive districts at rounds t and $t+1$. Let $s(\eta)$ and $s(\eta')$ be the search sequences that are obtained from η and η', respectively, and that satisfy the equilibrium stopping rule.

Consider $\sigma(\eta)$ a N-tuple of signals associated with sequence η. We have $G(s(\eta)) \neq G(s(\eta'))$ if and only if $\sigma(\eta)$ is such that $i^{t+}(\sigma(\eta)) = M-1$ and $i^{t-}(\sigma(\eta)) = N-M-1$, and

(*i*) $\{\sigma^t(\eta) = \overline{r} \text{ and } \sigma^{t+1}(\eta) = \underline{r}\}$, in which case $G(s(\eta)) = \Gamma(\overline{r}, ..., \overline{r})$ and $G(s(\eta')) = \Gamma(Er, \overline{r}, ..., \overline{r})$, or

(*ii*) $\{\sigma^t(\eta) = \underline{r} \text{ and } \sigma^{t+1}(\eta) = \overline{r}\}$, in which case $G(s(\eta)) = \Gamma(Er, \overline{r}, ..., \overline{r})$ and $G(s(\eta')) = \Gamma(\overline{r}, ..., \overline{r})$.

Since cases (*i*) and (*ii*) occur with the same probability, we get $EG(s(\eta)) = EG(s(\eta'))$.

By repeatedly switching two consecutive districts, one can cover the whole set of possible sequences, thereby establishing $EG(s) = EG(s')$ for any two sequences that satisfy the equilibrium stopping rule. ■

Define $\mathcal{P}^{\tau} \equiv \{1, ..., \tau\} \subseteq \mathcal{N}$ as the set of districts with the τ highest q_ns. Given an N-tuple of signals over $\mathcal{N}$, $\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_N)$, define

- $\mathcal{P}^{\tau+} \equiv \{n \in \mathcal{P}^{\tau} : \sigma_n = \overline{r}\}$ as the set of districts in $\mathcal{P}^{\tau}$ with favorable signal.
- $\mathcal{P}^{\tau-} \equiv \{n \in \mathcal{P}^{\tau} : \sigma_n = \underline{r}\}$ as the set of districts in $\mathcal{P}^{\tau}$ with unfavorable signal.

Lemma 6. *Let* $\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_N)$ *be an* N*-tuple of signals* $\sigma_n \in \{\underline{r}, \overline{r}, \emptyset\}$ *over* $\mathcal{N}$*. At any round* $t \in \{1, ..., N\}$*, relabel districts such that* $\mathcal{C}^t = \{1, ..., N+1-t\}$ *with* $q_1 > q_2 > ... > q_{N+1-t}$*. Pick a search sequence* s *where at any round* $t \in \{1, ..., N\}$*, we have*

$$\begin{cases} s^t \in \{1, ..., K^t\} & \text{if } i^{t+} < M \text{ and } i^{t-} < N - M \\ s^t = \emptyset & \text{if } i^{t+} \geq M \text{ or } i^{t-} \geq N - M \end{cases}$$

where $K^t \equiv \min\{M - i^{t+}, N - M - i^{t-}, N + 1 - t\}$*. Then, the set of districts IG searches is*

$$\mathcal{I}(\boldsymbol{\sigma}) = \begin{cases} \mathcal{N} & \text{if } \#\mathcal{P}^{N+}(\boldsymbol{\sigma}) < M \text{ and } \#\mathcal{P}^{N-}(\boldsymbol{\sigma}) < N - M \\ \mathcal{P}_T(\boldsymbol{\sigma}) & \text{if } \#\mathcal{P}^{N+}(\boldsymbol{\sigma}) \geq M \text{ or } \#\mathcal{P}^{N-}(\boldsymbol{\sigma}) \geq N - M \end{cases}$$

where $T \equiv \min\{t \in \{1, ..., N\} : i^{t+} = M \text{ or } i^{t-} = N - M\}$*.*

Proof of Lemma 6. I start with a signal profile $\boldsymbol{\sigma}$ for which $\#\mathcal{P}^{N+}(\boldsymbol{\sigma}) < M$ and $\#\mathcal{P}^{N-}(\boldsymbol{\sigma}) < N - M$. At any round $t \in \{1, ..., N\}$, we have $K^t \geq 1$, implying $s^t \neq \emptyset$. Hence $\mathcal{I}(\boldsymbol{\sigma}) = \mathcal{N}$.

I continue with a signal profile $\boldsymbol{\sigma}$ for which $\#\mathcal{P}^{N+}(\boldsymbol{\sigma}) \geq M$ or $\#\mathcal{P}^{N-}(\boldsymbol{\sigma}) \geq N - M$. Pick any round $t \in \{1, ..., N-1\}$ at which $K^t \geq 1$. Let $C^t \equiv \{1, ..., K^t\}$ and $s^t = i \in C^t$. Then, the set C^{t+1} among which IG chooses one district to search at round $t + 1$ is

- if $i^{t+} < i^{t-}$,

$$C^{t+1} = \begin{cases} C^t \backslash \{i\} & \text{if } \sigma_i = \overline{r} \\ (C^t \cup \{h\}) \backslash \{i\} & \text{if } \sigma_i \in \{\underline{r}, \emptyset\} \end{cases}$$

- if $i^{t+} > i^{t-}$,

$$C^{t+1} = \begin{cases} C^t \backslash \{i\} & \text{if } \sigma_i = \underline{r} \\ (C^t \cup \{h\}) \backslash \{i\} & \text{if } \sigma_i \in \{\overline{r}, \emptyset\} \end{cases}$$

- if $i^{t+} = i^{t-}$,

$$C^{t+1} = \begin{cases} C^t \backslash \{i\} & \text{if } \sigma_i \in \{\overline{r}, \underline{r}\} \\ (C^t \cup \{h\}) \backslash \{i\} & \text{if } \sigma_i = \emptyset \end{cases}$$

where $h = 1 + \max C^t$. Observe that $\#C^{t+1} \leq \#C^t$. Moreover, since $\#\mathcal{P}^{N+}(\boldsymbol{\sigma}) \geq M$ or $\#\mathcal{P}^{N-}(\boldsymbol{\sigma}) \geq N - M$, there is a round $\tau \geq t$ at which $K^\tau = 1$ and, therefore, $\#C^\tau = 1$.

Pick two districts h and i with $q_h > q_i$. Suppose $s^{t'} = i$ at some round t'. There are two cases to consider:

(i) $s^t = h$ at some round $t < t'$, or (*ii*) $s^t \neq h$ at every round $t < t'$. Observe that $q_h > q_i$ and $i \in C^{t'}$ imply $h \in C^{t'}$. Moreover, $\{h, i\} \subseteq C^{t'}$ implies $K^{t'} \geq 2$ and, therefore, $h \in C^{t'+1}$. Since $\#C^t$ is weakly decreasing with t and $\#C^\tau = 1$ at some round τ, we have $s^{t''} = h$ at some round $t'' \in \{t'+1, ..., \tau\}$.

Since these two cases exhaust all possibilities, we have $s^t = h$ at some round t. Hence there exists $k \in \mathcal{N}$ such that

$$\begin{cases} h \in \mathcal{I}(\boldsymbol{\sigma}) & \text{for all } h \in \mathcal{N} \text{ with } q_h \geq q_k \\ h \notin \mathcal{I}(\boldsymbol{\sigma}) & \text{for all } h \in \mathcal{N} \text{ with } q_h < q_k. \end{cases}$$

Given $C^1 = \{1, ..., \min\{M, N-M\}\}$ and the sequencing of C^t for $t = 1, ..., N$, we have $k = \min\{\tau : \#\mathcal{P}^{\tau+}(\boldsymbol{\sigma}) \geq M \text{ or } \#\mathcal{P}^{\tau-}(\boldsymbol{\sigma}) \geq N-M\}$. ■

I am now ready to prove Proposition 5. Let

$$\begin{cases} \overline{Q}_{S_k(\mathcal{P}^\tau)} \equiv \sum_{S \in S_k(\mathcal{P}^\tau)} \left[\prod_{\ell \in S} q_\ell\right] \left[\prod_{\ell \in \mathcal{P}^\tau \setminus S} (1 - q_\ell p)\right] \\ \underline{Q}_{S_k(\mathcal{P}^\tau)} \equiv \sum_{S \in S_k(\mathcal{P}^\tau)} \left[\prod_{\ell \in S} q_\ell\right] \left[\prod_{\ell \in \mathcal{P}^\tau \setminus S} (1 - q_\ell \cdot (1-p))\right]. \end{cases}$$

Proof of Proposition 5. The proof is inductive.

Proposition 4 implies $s^N = \emptyset$. Consider round $t = N-1$ with $K^{N-1} \geq 1$. There are two cases to consider:

(*i*) $K^{N-1} = 2$: Since $i^{t+} \leq M-2$ and $i^{t-} \leq N-M-2$, Proposition 4 implies the continuation search sequence is either $s^{N-1} = 1$ and $s^N = 2$, or $s^{N-1} = 2$ and $s^N = 1$. Both continuation sequences result in the same total quantity of goods and the same continuation search cost, implying IG is indifferent between $s^{N-1} = 1$ and $s^{N-1} = 2$. Hence $s^{N-1} \in \{1, 2\} = \mathcal{C}^{N-1}$.

(*ii*) $K^{N-1} = 1$: We know from Lemma 5 that the total expected quantity of goods is the same whether $s^{N-1} = 1$ or $s^{N-1} = 2$, that is, $EG_1 = EG_2$. The expected continuation search cost for $s^{N-1} = i \in \mathcal{C}^{N-1}$ is

$$E\varepsilon_i = \begin{cases} q_i p + 2(1 - q_i p) & \text{if } i^{(N-1)+} = M-1 \ \& \ i^{(N-1)-} < N-M-1 \\ q_i \cdot (1-p) + 2[1 - q_i \cdot (1-p)] & \text{if } i^{(N-1)+} < M-1 \ \& \ i^{(N-1)-} = N-M-1 \\ q_i + 2(1-q_i) & \text{if } i^{(N-1)+} = M-1 \ \& \ i^{(N-1)-} = N-M-1. \end{cases}$$

Since $q_1 > q_2$, we have $E\varepsilon_1 < E\varepsilon_2$. This, together with $EG_1 = EG_2$, implies $EV_1 > EV_2$. Hence $s^{N-1} = 1$.

Assume the statement is true at round $(t+1) \in \{2, ..., N-1\}$. Consider round t with $K^t \geq 1$. We know from Lemma 5 that the expected total quantity of goods is the same for any continuation search sequence satisfying the equilibrium stopping rule: $EG_1 = ... = EG_{N+1-t}$.

There are two cases to consider:

(*i*) $K^t = N+1-t$: We know $s^t \in \mathcal{C}^t$ (Proposition 4) and $\mathcal{I} = \mathcal{N}$ (Lemma 6 and the statement being true at round $t+1$). IG is then indifferent between any $s^t \in \mathcal{C}^t$ since they all yield the same total quantity of goods ($EG_1 = ... = EG_{N+1-t}$) and the same (continuation) search cost ($E\varepsilon_1 = ... = E\varepsilon_{N+1-t}$). Hence $s^t \in \{1, ..., N+1-t\} = \mathcal{C}^t$.

(*ii*) $K^t < N+1-t$: Let $s^t = h \in \mathcal{C}^t$. Using Lemma 6, we get that the expected continuation search cost is

$$E\varepsilon_h = (N+1-t) - \sum_{i=K^t}^{N-t} (N-i) \left(\overline{\alpha}_i^h + \underline{\alpha}_i^h\right)$$

where

$$\overline{\alpha}_i^h = \begin{cases} 0 & \text{if } i < M - i^{t+} \\ p^{M-i^{t+}} \cdot q_{i-1} \cdot \overline{Q}_{S_{M-i^{t+}-1}(\mathcal{P}^{i-2} \cup \{h\})} & \text{if } h > M - i^{t+} \ \& \\ & \quad i = M - i^{t+}, ..., h \\ p^{M-i^{t+}} \cdot q_i \cdot \overline{Q}_{S_{M-i^{t+}-1}(\mathcal{P}^{i-1})} & \text{otherwise} \end{cases}$$

is the probability $i^{i+} = M$ and $i^{(i-1)+} = M-1$, and

$$\underline{\alpha}_i^h = \begin{cases} 0 & \text{if } i < N - M - i^{t-} \\ (1-p)^{N-M-i^{t-}} \cdot q_{i-1} \cdot \underline{Q}_{S_{N-M-i^{t-}-1}(\mathcal{P}^{i-2} \cup \{h\})} & \text{if } h > N - M - i^{t-} \ \& \\ & \quad i = N - M - i^{t-}, ..., h \\ (1-p)^{N-M-i^{t-}} \cdot q_i \cdot \underline{Q}_{S_{N-M-i^{t-}-1}(\mathcal{P}^{i-1})} & \text{otherwise} \end{cases}$$

is the probability $i^{i-} = N-M$ and $i^{(i-1)-} = N-M-1$.

I write

$$E\varepsilon_{h+1} - E\varepsilon_h = \left(\overline{E\varepsilon}_{h+1} - \overline{E\varepsilon}_h\right) + \left(\underline{E\varepsilon}_{h+1} - \underline{E\varepsilon}_h\right)$$

where

$$\begin{cases} \overline{E\varepsilon}_{h+1} - \overline{E\varepsilon}_h \equiv \sum_{i=K^t}^{N-1} (N-i) \cdot \left(\overline{\alpha}_i^h - \overline{\alpha}_i^{h+1}\right) \\ \underline{E\varepsilon}_{h+1} - \underline{E\varepsilon}_h \equiv \sum_{i=K^t}^{N-1} (N-i) \cdot \left(\underline{\alpha}_i^h - \underline{\alpha}_i^{h+1}\right). \end{cases}$$

Tedious algebra (see next section), gives

$$\frac{\overline{E\varepsilon}_{h+1} - \overline{E\varepsilon}_h}{p^{M-i^{t+}} \cdot (q_h - q_{h+1})} = \tag{7}$$

$$\begin{cases} 0 & \text{if } h < M - i^{t+} \\ \sum_{k=1}^{\overline{\nu}+1} \sum_{i=M-i^{t+}-1}^{M-i^{t+}+\overline{\nu}-k} q_i \cdot \Lambda_i^+ \cdot \overline{Q}_{S_{M-i^{t+}-2}(\mathcal{P}^{i-1})} & \begin{array}{c} \text{if } h = M - i^{t+} + \overline{\nu} \ \& \ i^{t+} < M-1 \\ \text{for } \overline{\nu} \in \{0, ..., i^{t+} + N - M - t\} \end{array} \\ 1 + 1_{h \geq 1} \cdot \sum_{j=1}^{h-1} \left[\prod_{\ell \in \mathcal{P}^j} (1 - q_j p) \right] & \text{if } i^{t+} = M - 1 \end{cases}$$

and

$$\frac{\underline{E\varepsilon}_{h+1} - \underline{E\varepsilon}_h}{(1-p)^{N-M-i^{t-}} \cdot (q_h - q_{h+1})} =$$

$$\begin{cases} 0 & \text{if } h < N - M - i^{t-} \\ \sum_{k=1}^{\underline{\nu}+1} \sum_{i=N-M-i^{t-}-1}^{N-M-i^{t-}+\underline{\nu}-k} q_i \cdot \Lambda_i^- \cdot \underline{Q}_{S_{N-M-i^{t-}-2}(\mathcal{P}^{i-1})} & \begin{array}{c} \text{if } h = N - M - i^{t-} + \underline{\nu} \ \& \\ i^{t-} < N - M - 1 \\ \text{for } \underline{\nu} \in \{0, ..., i^{t-} + M - t\} \end{array} \\ 1 + 1_{h \geq 1} \cdot \sum_{j=1}^{h-1} \left[\prod_{\ell \in \mathcal{P}^j} (1 - q_j \cdot (1-p)) \right] & \text{if } i^{t-} = N - M - 1 \end{cases}$$

where

$$\Lambda_i^+ \equiv \begin{cases} 1 & \text{for } k = 1 \\ \prod_{j=1}^{k-1} (1 - q_{i+j} p) & \text{for } k > 1 \end{cases}$$

$$\Lambda_i^- \equiv \begin{cases} 1 & \text{for } k = 1 \\ \prod_{j=1}^{k-1} (1 - q_{i+j} \cdot (1-p)) & \text{for } k > 1. \end{cases}$$

Hence

$$\overline{E\varepsilon}_{h+1} - \overline{E\varepsilon}_h \begin{cases} = 0 & \text{if } h < M - i^{t+} \\ > 0 & \text{if } h \geq M - i^{t+} \end{cases}$$

$$\underline{E\varepsilon}_{h+1} - \underline{E\varepsilon}_h \begin{cases} = 0 & \text{if } h < N - M - i^{t-} \\ > 0 & \text{if } h \geq N - M - i^{t-}, \end{cases}$$

implying

$$\begin{cases} E\varepsilon_{h+1} = E\varepsilon_h & \text{for } h \in \{1, .., K^t - 1\} \\ E\varepsilon_{h+1} > E\varepsilon_h & \text{for } h \in \{K^t, ..., N - t\}. \end{cases}$$

To sum up, we have

$$\begin{cases} EG_1 = ... = EG_{N-t+1} \\ E\varepsilon_1 = ... = E\varepsilon_{K^t} < E\varepsilon_{K^t+1} < ... < E\varepsilon_{N-t+1}, \end{cases}$$

which implies $EV_1 = ... = EV_{K^t} > EV_{K^t+1} > ... > EV_{N+1-t}$. Hence $s^t \in \{1, ..., K^t\}$. ■

DERIVATION OF (7)

Equation (7) is obtained by rearranging

$$\overline{E\varepsilon}_{h+1} - \overline{E\varepsilon}_h \equiv \sum_{i=K^t}^{N-1} (N-i) \cdot \left(\overline{\alpha}_i^h - \overline{\alpha}_i^{h+1}\right).$$

I am going to develop the computations for the case where $h = M - i^{t+} + \overline{\nu}$ and $i^{t+} < M-1$ for $\overline{\nu} \in \{0, ..., i^{t+} + N - M - t\}$, with $h \in \{M+1-i^{t+}, ..., N-2-i^{t+}\}$‘; the computations are simpler or similar for the other cases.

We can rewrite $\overline{E\varepsilon}_{h+1} - \overline{E\varepsilon}_h$ as

$$\begin{aligned}
&\sum_{i=M-i^{t+}}^{h+1} (N-i) \cdot \left(\overline{\alpha}_i^h - \overline{\alpha}_i^{h+1}\right) \\
= \; & \left(N - M + i^{t+}\right) \cdot \left(\overline{\alpha}_{M-i^{t+}}^h - \overline{\alpha}_{M-i^{t+}}^{h+1}\right) \\
& + \sum_{i=M+1-i^{t+}}^{h} (N-i) \cdot \left(\overline{\alpha}_i^h - \overline{\alpha}_i^{h+1}\right) \\
& + (N-h-1) \cdot \left(\overline{\alpha}_{h+1}^h - \overline{\alpha}_{h+1}^{h+1}\right)
\end{aligned}$$

$$\begin{aligned}
= \; & \left(N - M + i^{t+}\right) \cdot p^{M-i^{t+}} \cdot q_{M-1-i^{t+}} \cdot \Big[\overline{Q}_{S_{M-1-i^{t+}}\left(\mathcal{P}^{M-2-i^{t+}} \cup \{h\}\right)} \\
& \qquad - \overline{Q}_{S_{M-1-i^{t+}}\left(\mathcal{P}^{M-2-i^{t+}} \cup \{h+1\}\right)}\Big] \\
& + \sum_{i=M+1-i^{t+}}^{h} (N-i) \cdot p^{M-i^{t+}} \cdot q_{i-1} \cdot \left[\overline{Q}_{S_{M-1-i^{t+}}(\mathcal{P}^{i-2} \cup \{h\})} - \overline{Q}_{S_{M-1-i^{t+}}(\mathcal{P}^{i-2} \cup \{h+1\})}\right] \\
& + (N-h-1) \cdot p^{M-i^{t+}} \cdot \left[q_{h+1} \cdot \overline{Q}_{S_{M-1-i^{t+}}(\mathcal{P}^h)} - q_h \cdot \overline{Q}_{S_{M-1-i^{t+}}(\mathcal{P}^{h-1} \cup \{h+1\})}\right].
\end{aligned}$$

Dividing both sides by $p^{M-i^{t+}} \cdot (q_h - q_{h+1})$ and rearranging the right hand side, we get that $\frac{\overline{E\varepsilon}_{h+1} - \overline{E\varepsilon}_h}{p^{M-i^{t+}} \cdot (q_h - q_{h+1})}$ is equal to

$$\begin{aligned}
& \left(N - M + i^{t+}\right) \cdot q_{M-1-i^{t+}} \cdot \overline{Q}_{S_{M-2-i^{t+}}\left(\mathcal{P}_{M-2-i^{t+}}\right)} \\
& + \sum_{i=M+1-i^{t+}}^{h} (N-i) \cdot q_{i-1} \cdot \left[\overline{Q}_{S_{M-2-i^{t+}}(\mathcal{P}^{i-2})} - p \cdot \overline{Q}_{S_{M-1-i^{t+}}(\mathcal{P}^{i-2})}\right] \\
& - (N-h-1) \cdot \overline{Q}_{S_{M-1-i^{t+}}(\mathcal{P}^{h-1})}
\end{aligned}$$

$$
\begin{aligned}
= \quad & \sum_{k=M-1-i^{t+}}^{h-1} q_k \cdot \overline{Q}_{S_{M-2-i^{t+}}(\mathcal{P}^{k-1})} \\
& + \left(N-M+i^{t+}-1\right) \cdot \Big[q_{M-1-i^{t+}} \cdot \overline{Q}_{S_{M-2-i^{t+}}\left(\mathcal{P}^{M-2-i^{t+}}\right)} \\
& \qquad\qquad - q_{M-i^{t+}} \cdot p \cdot \overline{Q}_{S_{M-1-i^{t+}}\left(\mathcal{P}^{M-1-i^{t+}}\right)} \Big] \\
& + \sum_{i=M+1-i^{t+}}^{h-1} (N-i-1) \cdot \Big[q_{i-1} \cdot \overline{Q}_{S_{M-2-i^{t+}}(\mathcal{P}^{i-2})} - q_i \cdot p \cdot \overline{Q}_{S_{M-1-i^{t+}}(\mathcal{P}^{i-1})} \Big] \\
& + (N-h-1) \cdot \Big[q_{h-1} \cdot \overline{Q}_{S_{M-2-i^{t+}}(\mathcal{P}^{h-2})} - \overline{Q}_{S_{M-1-i^{t+}}(\mathcal{P}^{h-1})} \Big]
\end{aligned}
$$

$$
\begin{aligned}
= \quad & \sum_{k=M-1-i^{t+}}^{h-1} q_k \cdot \overline{Q}_{S_{M-2-i^{t+}}(\mathcal{P}^{k-1})} \\
& + \left(N-M+i^{t+}-1\right) \cdot q_{M-1-i^{t+}} \cdot (1-q_{M-i^{t+}} \cdot p) \cdot \overline{Q}_{S_{M-2-i^{t+}}\left(\mathcal{P}^{M-2-i^{t+}}\right)} \\
& + \sum_{i=M+1-i^{t+}}^{h-1} (N-i-1) \cdot [q_{i-1} \cdot (1-q_i \cdot p) \cdot \overline{Q}_{S_{M-2-i^{t+}}(\mathcal{P}^{i-2})} \\
& \qquad\qquad - q_i \cdot p \cdot (1-q_{i-1} \cdot p) \cdot \overline{Q}_{S_{M-1-i^{t+}}(\mathcal{P}^{i-2})}] \\
& - (N-h-1) \cdot (1-q_{h-1} \cdot p) \cdot \overline{Q}_{S_{M-1-i^{t+}}(\mathcal{P}^{h-2})}.
\end{aligned}
$$

$$
\begin{aligned}
= \quad & \sum_{k=M-1-i^{t+}}^{h-1} q_k \cdot \overline{Q}_{S_{M-2-i^{t+}}(\mathcal{P}^{k-1})} + \sum_{k=M-1-i^{t+}}^{h-2} q_k \cdot (1-q_{k+1} \cdot p) \cdot \overline{Q}_{S_{M-2-i^{t+}}(\mathcal{P}^{h-1})} \\
& + \left(N-M+i^{t+}-2\right) \cdot [q_{M-1-i^{t+}} \cdot (1-q_{M-i^{t+}} \cdot p) \cdot \overline{Q}_{S_{M-2-i^{t+}}\left(\mathcal{P}^{M-2-i^{t+}}\right)} - \\
& \qquad - q_{M+1-i^{t+}} \cdot p \cdot (1-q_{M-i^{t+}} \cdot p) \cdot \overline{Q}_{S_{M-1-i^{t+}}\left(\mathcal{P}^{M-1-i^{t+}}\right)}] \\
& + \sum_{i=M+1-i^{t+}}^{h-2} (N-i-2) \cdot (1-q_i \cdot p) \cdot \Big[q_{i-1} \cdot \overline{Q}_{S_{M-2-i^{t+}}(\mathcal{P}^{i-2})} - q_{i+1} \cdot p \cdot \overline{Q}_{S_{M-1-i^{t+}}(\mathcal{P}^{i-1})} \Big] \\
& + (N-h-1) \cdot (1-q_{h-1} \cdot p) \cdot \Big[q_{h-2} \cdot \overline{Q}_{S_{M-2-i^{t+}}(\mathcal{P}^{h-3})} - \overline{Q}_{S_{M-1-i^{t+}}(\mathcal{P}^{h-2})} \Big].
\end{aligned}
$$

By repeating this process, we obtain (7).