Modeling Lobbying Industry with Multi-Agent Multi-Armed Bandit

Suyeol Yun
Department of Political Science
Massachusetts Institute of Technology
Cambridge, MA 02139
syyun@mit.edu

Abstract

What is lobbying? To answer this question, I model the lobbying industry with a multi-agent multi-armed bandit problem. Then I provide supportive evidence for the argument that explains lobbying as a delegated information acquisition process by simulation. First, I show that interest groups can successfully explore a large legislative space by delegating information acquisition to lobbyists. Second, I show that lobbyists are hired by interest groups when lobbyists have expert knowledge on a specific issue area across legislators. These two resulsts provide supportive evidence why interest groups have incentives to delegate information acquisition process to lobbyists and why real-world lobbyists interact with both sides of the legislators in terms of interest of their clients.

1 Introduction

What is lobbying? This simple question has been answered in different literatures (Hall and Deardorff, 2006; Bertrand et al., 2014; Ansolabehere et al., 2003) but there is no clear answer with agreement among scholars. For example, *vote-buying* theory argues that lobbying is to buy votes in legislation (Grossman and Helpman, 1994). However, this theory doesn't explain why the average price of vote is too cheap compared to the expected return of lobbying (Ansolabehere et al., 2003). As another example, *persuasion* theory argue that lobbying is to persuade legislators to change their positions on legislation close to the position of the lobbyists (Young and David, 1951; Bauer et al., 2017; Milbrath, 1984). However, this theory doesn't explain why lobbyists lobby already like-minded legislators those who don't need any persuasions (Hojnacki and Kimball, 1998).

Apart from those orthodox theories, this work provides a supportive evidence to the theory that explains lobbying as a delegated information acquisition process. According to this theory, interest groups have incentives to delegate information acquisition process to lobbyists. By modeling the lobbying industry using a multi-agent multi-armed bandit problem, I simulatively show that interest groups can successfully explore a large legislative space by delegating information acquisition process to lobbyists.

In addition, I also show that lobbyists are hired by interest groups when lobbyists have expert knowledge on a specific issue area across legislators by simulation. This result explains why lobbyists interact with both sides of the legislators in terms of interest of their clients (Hojnacki and Kimball, 1998). In other words, lobbyists interact with every relevant legislator in order to acquire information on the issue area of their clients regardless of the legislator's position on the issue area.

17.804 Quantitative Research Methods III (Fall 2022) | Final Paper | https://github.com/syyunn/categorical-bandit

2 Information Acquisition: Multi-Agent Multi-Armed Bandit Problem

In this section, I explain the relationship between the U.S. lobbying industry and multi-agent multi-armed bandit problem. In U.S. lobbying industry, interest groups need to find which legislator has the biggest influence over the issue area of their interest. To do so, they need to interact with different legislators to understand the expected rewards from each legislator. However, total chances of interaction is limited because of the cost of interaction. Therefore, interest groups need to design their exploration strategy how to interact with different legislators to maximize their expected rewards. This situation can be modeled by multi-agent multi-armed bandit problem. They decide whether to keep exploring or exploiting the best legislator according to their current knowledge at each timestep. Multi-armed bandit is a metaphor of a reward-seeking agent who has multiple choices of actions over multiple timesteps. It formalizes the exploration–exploitation dilemma in sequential decision making[1]. Interest groups in the U.S. lobbying industry need to maximize their expected rewards from the exploration process by choosing the best strategy of exploration and exploitation of legislators at each timestep. Therefore, I model the U.S. lobbying industry as multi-agent multi-armed bandit problem and simulate to prove two conjectures. The First conjecture is the incentive of hiring lobbying firm increases as the size of the exploration space increases. The Second conjecture is the incentive of hiring lobbying firm is higher in case that the lobbyist has expert knowledge on a particular issue area across different legislators compared to the case that the lobbyist has expert knowledge on a particular legislator across different issue areas.

2.1 Categorical Multi-Armed Bandit Problem

Multi-armed bandit problem assumes unknown reward distributions of each arm which are unknown to agents. In this section, I explain the *categorical multi-armed bandit problem* which I used to model the environment of the U.S. lobbying industry. Categorical multi-armed bandit problem is a multi-armed bandit problem that assumes the reward distribution of each arm to be categorical distribution. Since legislators have varying level of authorities (or influences) over different issue areas, I model this situation by categorical multi-armed bandit problem. There exists $|J|$ number of interest groups $\{1, 2, \ldots, J\}$, $|K|$ number of legislators $\{1, 2, \ldots, K\}$, $|C|$ number of categories of interset $\{1, 2, \ldots, C\}$ and $|T|$ number of timesteps $\{0, 1, 2, \ldots, T\}$. Then each interst group $j \in J$ has its own category of interest $\phi(j) \in C$ and each legislator $k \in K$ has a categorical distribution $\mathrm{Cat}(\mathrm{C}, \mathbf{p_k})$ where $C = \{1, 2, \ldots, C\}$ and $\mathbf{p_k} = [p_k^{(1)}, p_k^{(2)}, \ldots, p_k^{(C)}]$ which represents the support of the distribution and the vector of probabilities for each category of interest respectively. Whenever an intesrest group $j \in J$ interact with the k-th legislator, they sample $X_k \in C$ from $\mathrm{Cat}(\mathrm{C}, \mathbf{p_k})$. If the sampled category X_k matches with the category of interest of the interest group, i.e. $X_k = \phi(j)$, it gets the reward of $r_j^k = \mathbb{1}(X_k = \phi(j))$. In other words, interest group $j \in J$ get the reward of 1 when the sampled category X_k matches with the category of interest of the interest group $\phi(j)$. Otherwise, it gets the reward of 0. Therefore, interest groups need to find the best rewarding legislator in terms of their category of interest, i.e. $\mathrm{argmax}_{k \in K}\, p_k^{\phi(j)}$.

2.2 Thompson Sampling as a Bandit Strategy

There are many bandit strategies to solve the multi-armed bandit problem such as ϵ-greedy (Mnih et al., 2015), Upper Confidence Bound (UCB) (Auer, 2002), and Thompson Sampling (Thompson, 1933). Those strategies are based on different exploration strategies.

Among those strategies, Thompson Sampling is a Bayesian bandit strategy which models agent's prior and posterior belief over the reward distribution of each arm and systematically update it based on the observations. I use Thompson Sampling as a bandit strategy in this work. This is because I can explicitly model the level of belief of each interest group over the reward distribution of each legislator. This intuitively aligns with the real-world situation where interest groups have different level of understanding (which is belief) over the each legislator.

Let's assume that each interest group hold a prior belief over the categorical distribution of each legislator. Since Dirichlet distribution is conjugate prior of categorical distribution, I model the prior belief of each interest group j over the categorical distribution of the k-th legislator at timestep t, as

[1]The problem is orginally proposed by Robbins (1952)

$\mathrm{Dir}(C, \alpha_{\mathbf{jk}}^{\mathbf{t}})$ where $\alpha_{\mathbf{jk}} = [\alpha_{jk}^{t(1)}, \alpha_{jk}^{t(2)}, \ldots, \alpha_{jk}^{t(C)}]$. At the beginning of the exploration process $t = 0$, interest group j holds the *flat prior* belief over all legislators, i.e. $\alpha_{\mathbf{jk}}^{\mathbf{t=0}} = [1, 1, \ldots, 1] \quad \forall k \in K$. Then at each timestep $t \in T$, interest group j choose the k-th legislator to interact with and sample $X_k^t \in C$ from $\mathrm{Cat}(\mathrm{C}, \mathbf{p_k})$. Then the interest group j updates its prior belief by adding 1 to $\alpha_{jk}^{t(X_k^t)}$. This yields the posterior belief at time $t+1$, $\mathrm{Dir}(C, \alpha_{jk}^{t+1})$. After update, the interest group j chooses the next legislator k_{t+1} by sampling $\hat{\mathbf{p}}_{\mathbf{1j}}, \hat{\mathbf{p}}_{\mathbf{2j}}, \ldots \hat{\mathbf{p}}_{\mathbf{Kj}}$ from posterior beliefs from $\mathrm{Dir}(C, \alpha_{\mathbf{j1}}^{\mathbf{t+1}})$, $\mathrm{Dir}(C, \alpha_{\mathbf{j2}}^{\mathbf{t+1}}), \ldots, \mathrm{Dir}(C, \alpha_{\mathbf{jK}}^{\mathbf{t+1}})$ respectively. $\hat{\mathbf{p}}_{\mathbf{kj}} = [\hat{p}_{kj}^{(1)}, \hat{p}_{kj}^{(2)}, \ldots, \hat{p}_{kj}^{(C)}]$ is the vector of probabilities assigned for each category in case of legislator k. Then j choose $k_{t+1} = \underset{K}{\mathrm{argmax}}\{\hat{p}_{1j}^{(\phi(j))}, \hat{p}_{2j}^{(\phi(j))}, \ldots, \hat{p}_{Kj}^{(\phi(j))}\}$. This means that the interest group j chooses the next legislator k_{t+1} based on samples from its own posterior belief. Among the samples from posterior beliefs, j only compares the dimension $\phi(j)$ of those samples, probability of the category $\phi(j)$ for each legislator, and chooses k which has the highest probability of giving category $\phi(j)$ when sampled. In this way, the interest group j can systematically update its prior belief and balance the exploration and exploitation based on the randomness from the sampling. This strategy is called *Thompson sampling* (Thompson, 1933).

3 Large Search Space Problem in Lobbying Industry

In this section, I provide a supportive evidence that interest groups have incentives to delegate information acquisition process to lobbyists to solve the *large search space problem*. Starting from showing that interest groups fail to find the best rewarding legislator in case of large search space, I show that they can find the best rewarding legislator by hiring lobbyists by simulation.

3.1 Simulation I: Small Search Space

First, I simulate the case of small search space. In this simulation, I used $|K| = 32, |C| = 4, |J| = 1$[2] and $|T| = 2000$. Since I don't have have a good representation of the categorical distributions of legislators, I randomly generate a hypothetical categorical distribution of each legislator for simulation [3]. Besides, I use the normalized cumulative regret as the performance metric in the following simulations. The cumulative normalized regret is defined as $\sum_t^T (p_{\phi(j)}^{\max} - p_{a_t}^{\phi(j)}) / \sum_t^T (p_{\phi(j)}^{\max} - p_{\phi(j)}^{\min})$ where $p_{\phi(j)}^{\max} = \max_{k \in K}\{p_1^{(\phi(j))}, p_2^{(\phi(j))}, \ldots, p_K^{(\phi(j))}\}$, $p_{\phi(j)}^{\min} = \min_{k \in K}\{p_1^{(\phi(j))}, p_2^{(\phi(j))}, \ldots, p_K^{(\phi(j))}\}$ and a_t represent the choice of legislator taken by the interest group at time t. Regret represents how much the agent could have done better in terms of the expected reward if it had chosen the best action. Similarly, cumulative normalized regret is sum of regret over the entire time horizon T normalized by the sum of upper bound of regret[4]. This metric is bounded in the range of $[0, 1]$ and the smaller the value is, the better the performance is.

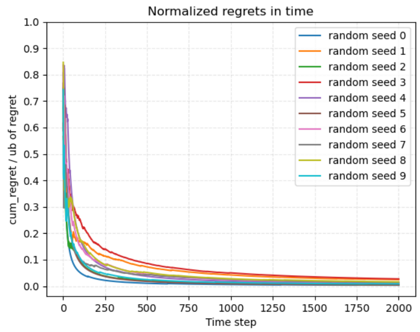
Figure 1a shows that normalized cumulative regrests of interest groups decrease as the number of timesteps increases. The mean of cumulative normalized regrets at time T from 10 different random seeds[5] is 0.012 with the variance of $5.98e-05$. This means that a interest group can successfully find the best rewarding legislator to interact in case of small search space. In this scenario, search space consists of $|K| \times |C| = 128$ number of parameters which is relatively small compared to the case of actual lobbying industry. Figure 1b shows that prior belief of the interest group well estimates the real probabilities of category of interest across legislators. It means that the mean of the posterior belief of the interest group corresponds to the real probability of category of interest across legislators. In other words, the interest group successfully reconstructed the categorical distributions of legislators within 2000 times of interaction. Therefore, figure 1c shows that the interest group successfully found the best legislator and keep interacting with the best one with the highest ratio.

[2] This is single agent setting

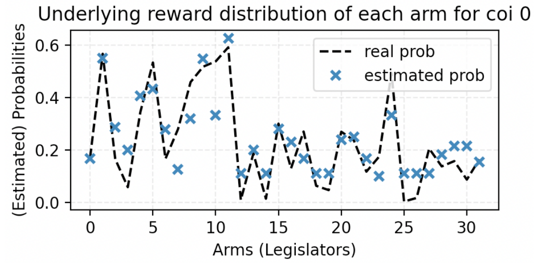
[3] By random sampling $\mathbf{p_k} = [p_k^{(1)}, p_k^{(2)}, \ldots, p_k^{(C)}]$ from $\mathrm{Dir}(C, \alpha)$ where $\alpha = [1, 1, \ldots, 1]$, I can generate hypothetical categorical distribution of each legislator k.

[4] I use this metric to compare the level of regret between different hyperparameter settings.

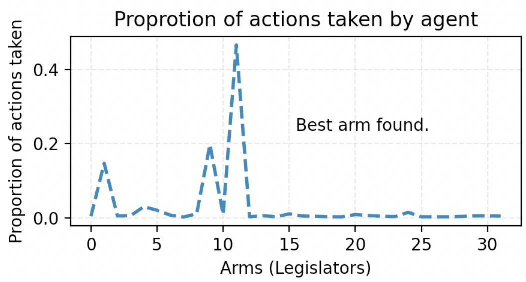
[5] Each random seed corresponds to the different set of categorical distributions of legislators.



(a) Cumulative normalized regret of interest groups in small search space with 10 random seeds.



(b) $p_1^{(\phi(j))}, p_2^{(\phi(j))} \dots p_K^{(\phi(j))}$ and $\hat{p}_1^{(\phi(j))}, \hat{p}_2^{(\phi(j))} \dots \hat{p}_K^{(\phi(j))}$ for random seed 0 at timestep T.



(c) The proportion of legislators being chosen by the interest group for random seed 0.

Figure 1: Simulation in case of Small Search Space

3.2 Simulation II: Large Search Space

In the previous simulaton, I used arbitrary small numbers for $|K|$ and $|C|$. Now I use $|K| = 112, |C| = 26, |J| = 1$ and $|T| = 2000$. $|K| = 112$ is the average number of legislators to whom top 10 lobbying firms campaign contribute in 2020. $|C| = 26$ is the average number of issue areas per clients in 2020. Both numbers are obtained from *Lobbying Disclosure Act* data[6]. I expect the simulation with this configuration to be more realistic compared to the previous simulation.

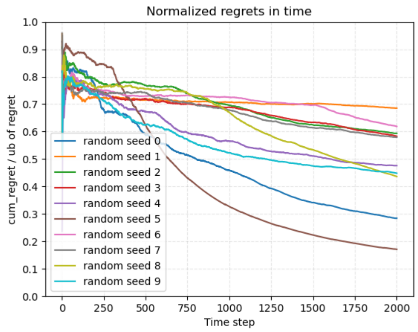
Fig 2a shows that the mean of cumulative normalized regrets at time T from 10 different random seeds[7] is 0.48 with the variance of 0.023 which is much higher and larger compared to the previous simulation, which was 0.012 with the variance of $5.98e - 05$ respectively. It implies that interset group fails to find the best rewarding legislator within 2000 times of interaction. This is because the search space is almost 23 times larger than the previous simulation. Due to the large search space, it's difficult for the interest group to explore enough and find the best legislator within the timesteps of 2000. Therefore, prior belief of the interest group fails to estimate the real world probabilities of legislators as shown in Fig 2b. This results in exploiting a worse legislator compared to the best one as shown in Figure 2c. In this case, although the best legislator k^* is the legislator 69, the interest group keeps interacting with the legislator 90 with the highest ratio.

3.3 Simulation III: Large Search Space with a Lobbyist

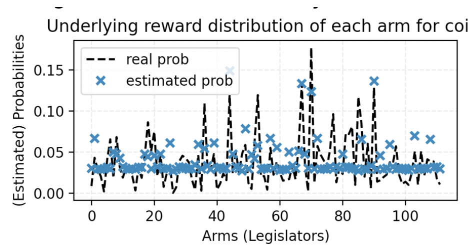
In this simulation, I introduce a lobbyist into the multi-armed bandit setting explained in Section 2.1. By showing that lobbyists can help interest groups to find the best legislators in the large search space, I provide a supportive evidence that lobbying can be explained as a delegated information acquisition process by simulation.

[6]Data is available through `https://lda.senate.gov/filings/public/filing/search/`

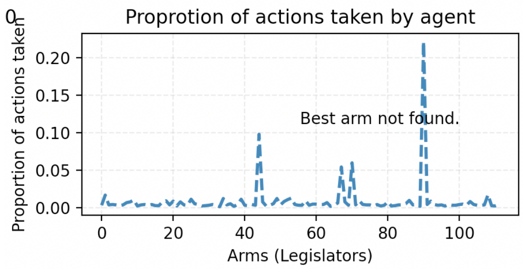
[7]Each random seed corresponds to the different set of categorical distributions of legislators.



(a) Cumulative normalized regret of interest groups in large search space



(b) $p_1^{(\phi(j))}, p_2^{(\phi(j))} \dots p_K^{(\phi(j))}$ and $\hat{p}_1^{(\phi(j))}, \hat{p}_2^{(\phi(j))} \dots \hat{p}_K^{(\phi(j))}$ for random seed 0 at timestep T.



(c) The proportion of legislators being chosen by the interest group for random seed 0.

Figure 2: Simulation in case of Large Search Space

3.3.1 Introducing Lobbyists into the Multi-Armed Bandit Setting

I conjecture that interest groups who share the same category of interest can collaborate by sharing the same lobbyist. By doing so, they can concentrate their resources on finding the best legislators in the large search space. In addition to the settings introduced in Section 2.1, I introduce a lobbyist $l \in L$ with its own set of prior beliefs $\mathrm{Dir}(C, \alpha_{\mathbf{lk}})\ \forall k \in K$ where $\alpha_{\mathbf{lk}} = [\alpha_{lk}^{(1)}, \alpha_{lk}^{(2)}, \dots, \alpha_{lk}^{(C)}]$. At each timestep t, an interest group j choose the next legislator $k_{t+1} = \underset{K}{\mathrm{argmax}} \left[\{\hat{p}_{1j}^{(\phi(j))}, \hat{p}_{2j}^{(\phi(j))}, \dots, \hat{p}_{Kj}^{(\phi(j))}\} \cup \bigcup_{l \in L} \{\hat{p}_{1l}^{(\phi(j))}, \hat{p}_{2l}^{(\phi(j))}, \dots, \hat{p}_{Kl}^{(\phi(j))}\} \right]$. This means that interest group j chooses the best rewarding legislator based on the samples from the prior beliefs of all lobbyists including that of himself. I assume complete information so that any interest groups can access to prior beliefs of all lobbyists at any timestep. In this setting, if the interest group j chooses the next legislator k_{t+1} based on the prior belief of lobbyist l, I update the prior belief of lobbyist l with the sampled observation $X_k^{t+1} \sim \mathrm{Cat}(\mathrm{C}, \mathbf{p_k})$ by adding 1 to $\alpha_{lk}^{(X_k^{t+1})}$. However, I don't update the prior belief of interest group j, $\alpha_{jk}^{(X_k^{t+1})}$. This is because if an interest group hires a lobbyist, the lobbyist explore the legislative space on behalf of the interest group. This is because an interest group doesn't accumulate any experience when they hire a lobbyist. Lobbyists interact with the legislators on behalf of the interest group and they accumulate the experience to themselves but not necessarily to the interest group. In contrast, if an interest group choose the next legislator based on their own prior belief, we update the prior belief of interest group j as usual. This modeling introduces the *pros and cons* of hiring a lobbyist. If an interest group hires a lobbyist, they may find the better legislator. However, they don't accumulate any experience to understand the reward distribution of the legislators when they hire a lobbyist.

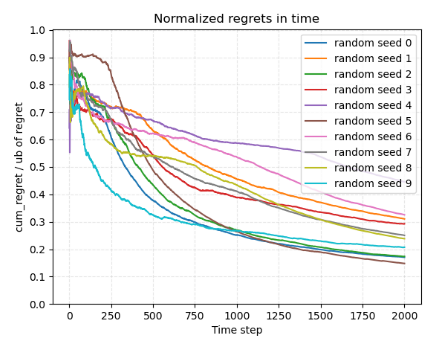
3.3.2 Simulation : Large Search Space with a Lobbyist

In this simulation, I maintain the size of the search space as same as the previous simulation by configuring $|K| = 112, |C| = 26, |T| = 2000$. However, I introduce a lobbyist $|L| = 1$ and five interest groups with $|J| = 5$. I assume that 5 different interest groups share the same category of interest. I choose $|J| = 5$ based on the average number of clients per each issue code[8] in *Lobbying Disclosure Act* data in 2020. With this configuration, I expect the simulation to be more realistic.

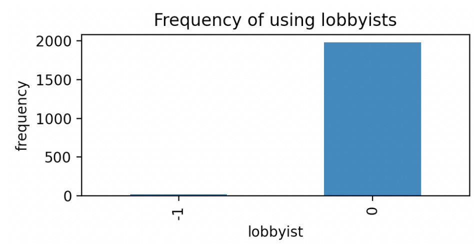
In this simulation, the mean of cumulative normalized regrets at time T from 10 different random seeds[9] records 0.25 with the variance of 0.007 which are lower and smaller than those of the previous simulation, 0.48, 0.023 respectively. (See Fig 3a). In addition, the interest groups found the best legislator as shown in Fig 3d. This is possible because multiple agents share the same lobbyist and collectively update the prior belief of the lobbyist. Fig 3b shows that the interest group keep selecting into using the lobbyist's prior rather than its own prior. This is because the lobbyist approximates the best rewarding legislator faster than interest groups. Fig 3c shows that the lobbyist approximates the true reward distribution of the legislators closely after 2000 timesteps.

3.4 Lobbying as a Delegated Information Acquisition Process

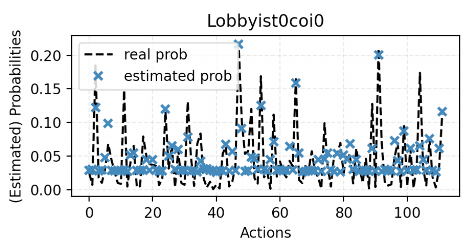
Simulations in Section 3 provides a supportive evidence that lobbyists can be used to find the best legislators when interest groups face large search space problem. If interest groups share the same category of interest, they can form a collaborative relationship via lobbyist to find the best legislator for their shared category of interest. Regardless of whether interest groups intend to form a coalition or not, I argue that interest groups can form a coalition with other interst groups via lobbyists. This finding is important because it provides a new perspective on lobbying not just as a delegated information acquisition process but as a coalition formation process.



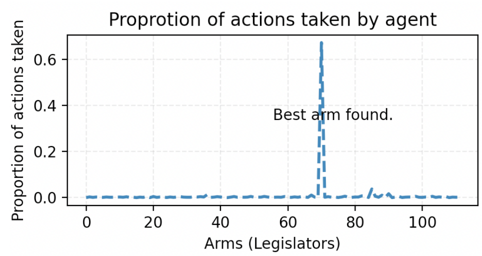
(a) Cumulative normalized regret of interest groups in large search space with lobbyist



(b) Frquency of using lobbyist. 0 represents using lobbyist and -1 represents *not* using lobbyist, relying on interst groups own prior.



(c) $p_1^{(\phi(j))}, p_2^{(\phi(j))} \dots p_K^{(\phi(j))}$ and $\hat{p}_1^{(\phi(j))}, \hat{p}_2^{(\phi(j))} \dots \hat{p}_K^{(\phi(j))}$ of *lobbyist* for random seed 0 at timestep T.



(d) The proportion of legislators being chosen by the interest group for random seed 0.

Figure 3: Simulation in case of Large Search Space with Lobbyist

4 Specialization of Lobbyists and Expert knowledge

In this section, I find the reason why lobbyists lobby both sides of legislators in reality. This phenomenon that lobbyists lobbying both sides of legislators is regarded as a puzzle in the literatures

[8] *Lobbying Disclosure Act* data provides 81 number of issue codes such as TRD (Trade) and TAX (Taxation) which are used to categorize intent of lobbying activities.

[9] Each random seed corresponds to the different set of categorical distributions of legislators.

because lobbyists don't need to lobby already supportive legislators according to the persuasion (informational) theories on lobbying (Hojnacki and Kimball, 1998).

To do so, I compare the simulation results from the two differet types of expert knowledge of lobbyists. In the section 3.3.1, I assumed that each lobbyist l has its own prior belief of legislators $\alpha_{\mathbf{l1}}, \alpha_{\mathbf{l2}}, \ldots, \alpha_{\mathbf{lK}} \in \mathbb{N}^C$ which are K number of concentration parameter of dirichlet distributions. It can be visualized as shown in Fig 4. A lobbyist can either have expert knowledge on a specific issue area across all legislators (blue in Fig 4) or have expert knowledge on a specific legislator (red in Fig 4) across all issue areas. Here,

I conjecture that lobbyists can be specialized in case lobbyists having expert knowledge on a specific issue area across all legislators rather than having expert knowledge on a specific legislator across all issue areas. Here, I use the word *specialized* to mean that lobbyists get hired by interest groups that matches the expert knowledge of lobbyists. If the conjecture is true, lobbyists will lobby both sides of legislators to acquire the expert knowledge on a specific issue area across all legislators.

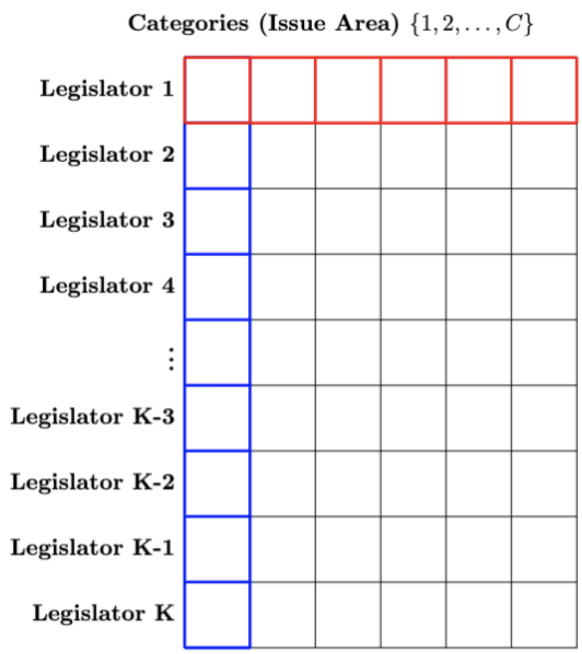


Figure 4: Prior belief of lobbyist is paramaterized by $K \times C$ matrix.

4.1 Simulation Setting

In this simulation, I use the configuration of $|K| = 112$, $|C| = 26$, $|T| = 2000$, $|J| = 10$ and $|L| = 2$. I maintain the size of search space as used in the section 3.2 and 3.3 by setting $|K| = 112$, $|C| = 26$. Then partition 10 interest groups into two groups, 5 for each group and assume that the interest groups in the same group share the same category of interest while interest groups in different groups have different categories of interest c_1 and c_2. Therefore, I add $|L| = 2$ lobbyists to represent the category of interest of each group. I define a category of expert knowledge of a lobbyist l as $\upsilon(l)$. For the two lobbyist l_1 and l_2, I assign $\upsilon(l_1) = c_1$ and $\upsilon(l_2) = c_2$ respectively.

4.1.1 Simulation Setting For Expert Knowledge

In this subsection, I describe the simulation setting for the two different types of expert knowledge of lobbyists.

First, to model the expert knowledge of lobbyists on a specific issue area across all legislators (blue in Fig 4), I preset the prior belief of lobbyists to be close to the categorical distribution of legislators. By setting $\alpha_{lk}^{(c)} = p_k^{(c)} \cdot |C| \cdot \delta$ and $\alpha_{lk}^{(-c)} = (|C| - \alpha_{lk}^{(c)} / (|C| - 1) \; \forall k$, I introduced expert knowledge on a specific issue area c across all legislators. This setting is a redistribution of the sum of concentration parameter of dirichlet distribution which is equal to $|C|$. The distribution of the quantity $|C|$ across all categories determines the expectation over the parameter of the categorical distribution of legislator k. Therefore, if I set $\alpha_{lk}^{(c)} = p_k^{(c)} \cdot |C| \cdot \delta$ and distribute remaining quantity $|C| - \alpha_{lk}^{(c)}$ across all other $|C| - 1$ number of categories , it makes the expectation of such dirichlet distribution for category c to be eqaul to $p_k^{(c)}$. In addition, by introducing $\delta \in [0, 1]$ which is a temperature parameter to control the degree of expert knowledge, I can gradually introduce expert knowledge on a specific issue area c.

Second, to model the expert knowledge of lobbyists on a specific legislator k^* across all issue areas (red in Fig 4), I set $\alpha_{lk^*}^{(c)} = p_{k^*}^{(c)} \cdot |C| \cdot \delta + \epsilon \; \forall c$ where $k^* = \mathrm{argmax}_{\mathrm{K}} \{p_1^{(c)}, p_2^{(c)}, \ldots, p_K^{(c)}\}$. ϵ is a small number to prevent concentration parameter from being zero when $\delta = 0$. This means that first I find the k^* which represent the best rewarding arm in terms of category c and copy his/her parameter of categorical distribution to dirichlet parameter of lobbyist. As same as the first setting, by introducing $\delta \in [0, 1]$ I can gradually introduce expert knowledge on a specific legislator k^* across all issue areas.

With this setting, I ran the simulation and measure the ratio of hiring matching lobbyist. Matching lobbyist is defined as the lobbyist who has the expert knowledge on the category of interest of interest groups. For example, if the category of interest of interest group j is c_1 and the expert knowledge of lobbyist l_1 is c_1, then l_1 is a "matching" lobbyist for interest group j. Then the ratio of hiring matching lobbyist is defined as the ratio of hiring "matching" lobbyist over the total number of timesteps $|T| = 2000$.

Since each group of interest groups has the shared category of interest c_1, c_2 and two lobbyist l_1, l_2 have the expert knowledge on c_1, c_2 respectively, I expect that the ratio of hiring "matching" lobbyist increases as the temperature parameter δ increases.

4.2 Simulation Result

Figure 5 shows the simulation results. The simulation is ran twice for two different types of expert knowledge of lobbyists. The figure shows that only in case of expert knowledge on a specific issue area across all legislators (blue in Fig 4), the ratio of hiring "matching" lobbyist is higher than 0.5, which is a baseline of randomly selecting a lobbyist between two available lobbyists. In case of lobbyist having expert knowledge on a specific issue area across all legislators, ratio of hiring "matching" lobbyist converges to 1 as the temperature parameter δ increases. However, in case of lobbyist having expert knowledge on a specific legislator across all issue areas (red in Fig 4), ratio of hiring "matching" lobbyist never become explicitly higher than 0.5. This result indicates that lobbyists are specialized in lobbying industry toward the direction of acquiring expert knowledge on a specific issue area across all legislators. In other words, lobbyists are selected by interest groups only when they have the expert knowledge on a specific issue area across all legislators. This is intuitively reasonable because lobbyists are a "middleman" between interest groups and legislators who explore the legislative space on behalf of interest groups. Since the meaning of exploration is to find the legislator who has the best expected retrun on a specific issue area, lobbyists need to equip themselves with expert knowledge on a specific issue area across all legislators which can provide "comparison" between legislators. Otherwise, simply knowing one best legislator without knowing other legislators' expected return is not enough to make a decision for interest groups to particularly hire such lobbyist. Therefore, for lobbyists to survive in the market, they need to accumulate expert knowledge on a specific issue area across all legislators. This result explains a puzzle in the literatures why lobbyists lobby to both sides of legislators.

5 Conclusion: Limitations and Future Directions

In this work, I modeled the lobbying industry as a multi-armed bandit problem. By modeling lobbying industry as a simulative environment based on multi-armed bandit problem, I found that interest groups are facing large search space problem and they can overcome this by using lobbyists. In addition, I found that lobbyists are hired by interest groups only when they have the expert knowledge on a specific issue area across all legislators. This finding explains a puzzle in the literatures why lobbyists lobby to both sides of legislators.

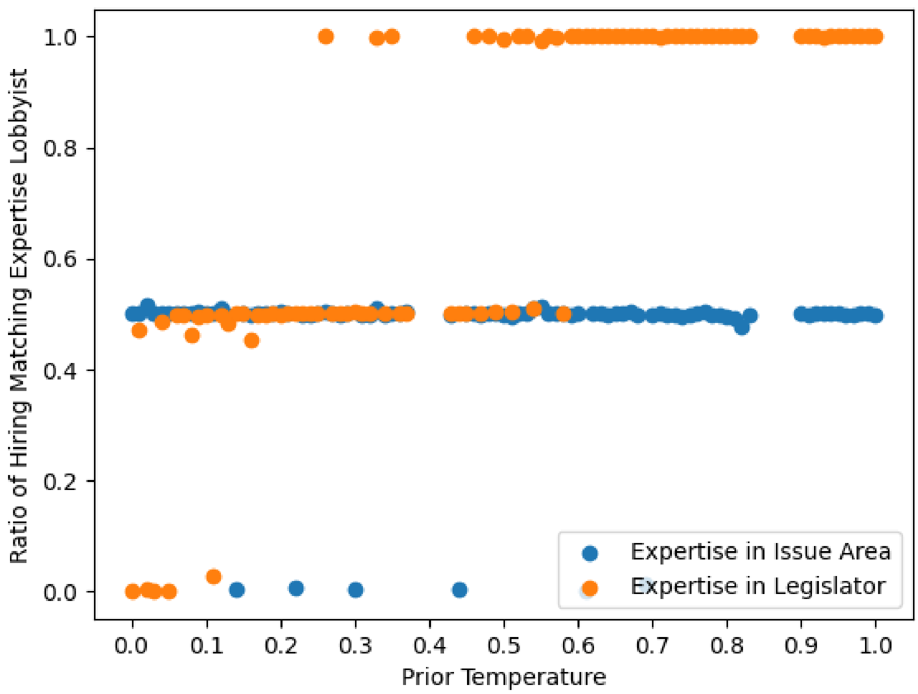


Figure 5: Mean ratio of hiring "matching" lobbyist for two different types of expert knowledge of lobbyist.

Although this work shed a light on understanding lobbying industry, there are still many limitations in this work. First, this work uses a *hypothetical* categorical distribution of legislators' influence over different issue areas. Since this distribution is hypothetical, finidngs from a few random seeds might be limited to extend to the real world. Second, this work assumes each interest group has a unique category of interest and all categories are orthogonal to each other. However, each interest group has multiple categories of interest and those categories are correlated to each other. Third, this work provides simulation results with very small number of interest groups and lobbyists. In reality, there are more than 70,000 interest groups and 20,000 lobbyists interacting with each other. Therefore, this work needs to be extended to a larger scale simulation to approximate the dynamics of lobbying industry more realistically.

Based on those limitations, I plan to extend this work in the future. First, I plan to find the good representation of legislators' influence over different issue areas. Recent development of *reinforcement learning* provides multiple methodologies to estimate the value of each action (or policy[10]) for different sitautions (Sutton and Barto, 2018). By using those methodologies, I can estimate the value of different actions which is interaction with different legislators in this case. If I model this value function as a function of interest group and legislator specific characteristics, I can approximate more realistic reward function which accomodates the real world characteristics of interest groups and legislators. Second, I plan to extend this work to a larger scale simulation. Since this simulation is based on "shared enviornment" between multiple agents, simple parallelization based on multi-processing doesn't work because it requires a *shared* memory between multiple agents. To overcome this, *decentralized* multi-agent simulation method is needed. In this method, each agent uses its own memory and relies on its own estimation of the behavior of other agents and the environment (Zhang

[10]Policy is defined as a function that maps state to optimal actions. In other words, policy represents a planning about how to act under varying situations. In *reifn learning*, we keep evaluating the value of each policy and find the optimal policy.

et al., 2018). In this way, agents can be simulated in parallel and the entire simulation can be extended to a larger scale.

In summary, this work provides a supportive evidence for a new perspective on understanding lobbying industry. In addition, this work also provides a formal way of modeling lobbying industry which can be simulatively tested with different assumptions and parameters. Since the components introduced in this modeling are replaceable with more realistic components, this work is modular so that it can be extended based on future researches that provide more realistic components. Although it's still controversial whether this kind of simulative research can be accepted as a valid knowledge about the real world, more empirical studies based on findings from this kind of simulations may indirectly prove the usefulness and validity of this simulative research in the future.

References

Ansolabehere, S., de Figueiredo, J. M., and Snyder, J. M. (2003). Why is there so little money in u.s. politics? *The Journal of Economic Perspectives*, 17(1):105–130.

Auer, P. (2002). Using confidence bounds for exploitation-exploration trade-offs. *Journal of Machine Learning Research*, 3:397–422.

Bauer, R. A., de Sola Pool, I., and Dexter, L. A. (2017). *American Business and Public Policy*. Routledge.

Bertrand, M., Bombardini, M., and Trebbi, F. (2014). Is it whom you know or what you know? an empirical assessment of the lobbying process. *The American Economic Review*, 104(12):3885–3920.

Grossman, G. M. and Helpman, E. (1994). Protection for Sale. *American Economic Review*, 84(4):833–850.

Hall, R. L. and Deardorff, A. V. (2006). Lobbying as legislative subsidy. *American Political Science Review*, 100(1):69–84.

Hojnacki, M. and Kimball, D. C. (1998). Organized interests and the decision of whom to lobby in congress. *The American Political Science Review*, 92(4):775–790.

Milbrath, L. (1984). *Washington Lobbyists Survey, 1956-1957*. ICPSR (Collection : En ligne). Inter-university Consortium for Political and Social Research.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., and Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533.

Robbins, H. E. (1952). Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society*, 58:527–535.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. The MIT Press, second edition.

Thompson, W. R. (1933). On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3/4):285–294.

Young, R. and David, T. (1951). The governmental process: Political interests and public opinion. *The ANNALS of the American Academy of Political and Social Science*, 278(1):200–201.

Zhang, K., Yang, Z., Liu, H., Zhang, T., and Başar, T. (2018). Fully decentralized multi-agent reinforcement learning with networked agents.