An Expert's Effect on Social Influence: An Application to Genome Editing in Domestic Livestock

Joseph P. Navelski*

School of Economic Sciences

Washington State University

October 15, 2022

Abstract

This paper investigates how an agent's latent position in an outside network heterogeneously affects the social influence of agents in another network. To do this, I first predict the probability of n agents simultaneously connecting with m other expert agents using their latent positions in space, and then use these predictions to form a $n \times n$ social network through their weighted common connections they have with the m experts. I then use DeGroot (1974)'s social learning model to estimate individual social influence in the $n \times n$ network, and conduct comparative statics to see how social influence changes with respect to a change in an experts latent position. An applied analysis to $m = 46$ experts in the genome editing in domestic livestock (GEDL) field on Twitter shows that if the most popular expert, @NonGMOProject, shifts their latent position from $-.63$ to 0 (positions are mapped to a $(-1, 1)$ scale) they will experience a 39.9% increase in the amount of follows they get, if they change their position from $-.63$ to .45 they will make social influence in the $n \times n$ network 40.62% more equal, and if they change their position from $-.63$ to .475 they will see a 21.4% gain in social influence for the least influential quantile of followers. Policy implication show that expert rhetoric can affect social influence in an outside network, similar to that of an "echo chamber," and that policy makers can regulate this rhetoric to make influence in this network more or less equal.

Key Words: Latent variable MCMC estimation, opinion formation, network formation, social influence, social network analysis.

*Address: 301H Hulbert Hall, Washington State University, Pullman, WA 99164. E-mail: joseph.navelski@wsu.edu.

1 Introduction

Social networks facilitate information transmission and this allows agents in a society to gain insight from other agent's opinions and behaviors. This process is known as social learning and was mathematically modeled by DeGroot (1974) in a non-Bayesian setting.[1] They theorize an agent "learns" from their peers by repeatedly updating their beliefs by averaging all agents opinions, including their own. DeGroot (1974)'s model led researchers to become more engaged in the dynamics of opinion formation through social learning, including those in the field of economics, where DeMarzo et al. (2003) and Golub and Jackson (2010) propose that if agents are in a strongly connected network, a network where no agents are completely isolated from the other, then certain graphical properties of te matrix are satisfied where the social learning process will converge to a consensus, or a steady state. This proposition is important because given a social network converges to a steady state, this learning process yields a unique vector of social influence weights where each weight summarizes the social influence each agent has in the learning process (DeGroot (1974), DeMarzo et al. (2003), Golub and Jackson (2010)). This result is powerful because it is one of the first to characterize a theoretical measurement of social influence through the dynamic learning process.

Many others studies investigate DeGroot (1974)'s learning model where they study how learning happens over time given agents have some initial endowment, or belief, $p_i^{(0)}$ about the state of nature and an already constructed precision π_{ij} level in which agent i weighs the opinion of agent j (DeMarzo et al. (2003), Golub and Jackson (2010)). This precision weight can be seen as a weighted link in a $n \times n$ network and is usually difficult to measure since the literature on empirical network formation is at it's infancy in the field of economics. Graham (2017)'s recent econometric work in network formation has given researchers a base for how to identify the parameters that govern the formation of a network and how the marginal effects of these parameters can alter how a network is formed, but assumes agents do not build a connection with themselves, or in the case of DeGroot (1974)'s model, does not weigh their own beliefs against others. Chandrasekhar et al. (2020)'s tests to see if agents use a mixture of learning types, either Bayesian or DeGroot (1974)'s Non-Bayesian learning model, and see what network structures lead to failures in asymptotic learning

[1] Many other economist use learning in a Bayesian setting where opinions update as new evidence or information becomes available (Banerjee (1992), Bikhchandani et al. (1992), Smith and Sørensen (2000), Acemoglu and Ozdaglar (2011), and Glass and Glass (2021)). For this paper, I focus on the non-Bayesian setting.

convergence (i.e., drawing a consensus) and test real-life social networks to see how much they mix in their learning methods. They also use a mixture of Penrose (2003)'s random geometric graph theory and Erdős et al. (1960)'s methods to synthetically form a network where random geometric graph theory, much like the work of Graham (2017), assumes that links between agents π_{ij} are formed according to how close they are in latent space where as Erdős et al. (1960)'s method assumes links are independently realized. Chandrasekhar et al. (2020) then test how the structure of these sparse and non-sparse graphs affect the mixture of learning over time, but do not investigate how the latent variables, used to form the synthetic network, affect this learning process. They also do not investigate how the parameters that govern the formation of a network, for example an agent's connections outside of the network, affect the social influence of each agent within the network.

This paper seeks to fill these gaps in the literature assuming a model where networks are formed based on connections outside of the network where individuals compare these connections to build a weighted or unweighted undirected network. Individuals then socially learn in a DeGroot (1974) model until convergence, where a social influence vector is realized. The main contribution of the paper is to see how the social influence vector inside the network changes with respect to an individual's characteristics outside of the network. I propose three different examples to show how the undirected social network can be formed, with one example where I estimate the latent parameters that contribute to the formation of that network. I apply this analysis to a set of Twitter users, where I investigate how their independent decisions to "follow" a set of experts outside the network dictates the social network structure an influence weights derived from this network. This application specifically investigates $m = 46$ genome editing in the domestic livestock experts, and $n = 3,383$ of their most informed followers, on Twitter. Assuming individual's decisions to follow others on Twitter are homophilic, I use this data and a random coefficient latent variable model to uncover the latent position for all $m = 46$ experts and $n = 3,383$ users, and then build a weighted undirected network based on the $n = 3,383$ users comparing who they follow out of the $m = 46$ experts. I then computationally show how a change in a popular expert's latent position affects the social influence distribution between the $n = 3,383$ agents, post social learning convergence.

This paper is organized as follows: Section 2 presents the theoretical model and examples, Section 3 presents the application data used to demonstrate the theoretical model, Section 4 presents the application results and the synthetic marginal effects, and Section 5 presents the discussion.

2 Model

2.1 Learning and Social Influence

The DeGroot model considers a finite set of $N = \{1, \ldots, n\}$ *agents* or *nodes* that interact in a social network. The interactions between agents are defined by an $n \times n$ nonnegative row stochastic matrix $\mathbf{T}$, where each element of the $\mathbf{T}$ matrix t_{ik} represents the amount of attention agent i gives to agent k's opinion. If $i = k$, this is the amount of attention agent i put on themselves, and the $\mathbf{T}$ matrix does not need to be symmetric implying that agent i could put a high weight on agent k's opinion, but agent k need not put a high weight on agent i's opinion. The $\mathbf{T}$ matrix is called an interaction matrix, and t_{ik} can be seen as the weight or trust agent i puts on the current belief of agent k in constructing their belief for next period (Jackson (2010)).[2] Each agent has belief $p_i^{(t)} \in [0, 1]$, and the vector of all n agent's beliefs is represented by $\mathbf{p}^{(t)}$. The belief updating rule is $\mathbf{p}^{(t)} = \mathbf{T}\mathbf{p}^{(t-1)}$ which implies

$$\mathbf{p}^{(t)} = \mathbf{T}^t \mathbf{p}^{(0)} \tag{1}$$

where,

$$\mathbf{T}_{n\times n} = \begin{bmatrix} t_{11} & t_{12} & \ldots & t_{1n} \\ t_{21} & t_{22} & \ldots & t_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ t_{n1} & t_{n2} & \ldots & t_{nn} \end{bmatrix} \quad \text{and} \quad \mathbf{p}_{n\times 1}^{(t)} = \begin{bmatrix} p_1^{(t)} \\ p_2^{(t)} \\ \vdots \\ p_n^{(t)} \end{bmatrix}.$$

Beliefs evolve according to the Bayesian theory presented by DeMarzo et al. (2003), where opinions are formed through social interaction and are boundedly rational. The authors propose that at time $t = 0$, each agent receives a noisy signal $p_i^{(0)} = \mu + \epsilon_i$, where $\epsilon_i \in \mathbb{R}$ is a noise term with expectation zero $E[\epsilon_i] = 0$ and μ is some state of nature. The vector $\mathbf{p}^{(0)}$ can be interpreted as a vector of initial beliefs for all agents, and each vector $\mathbf{p}^{(t)}$ can be interpreted as each agent i's updated belief at time t after each agent i interacts with all other agents. This interaction process continuously updates until $t = \infty$, and converges to the proportion of social influence agent i has

[2] DeMarzo et al. (2003) refer to $\mathbf{T}$ as the "listening" matrix where each element t_{ik} represents how much agent i listens to agent k's opinion, Golub and Jackson (2010) refer to t_{ik} as how much precision agent i puts k's opinion, Jadbabaie et al. (2012) refer $\mathbf{T}$ as the social interaction matrix where t_{ik} represents the "influence" or "persuasion power" agent i gets from agent k, and Jackson (2010) refers to $\mathbf{T}$ as the "weight" or "trust" matrix where t_{ik} represents the weight or trust the ith agent has on the current belief of agent k in forming its own belief for the next period.

in that social network. This paper investigates how individuals and their attributes outside of the social interaction network $\mathbf{T}$ affect individual social influence within the network, and I explain the methodology and properties needed to uncover this social influence vector below.

2.1.1 Convergence of Beliefs Under Naïve Updating and Social Interaction

The notion of matrix convergence is important when understanding social influence, and a matrix $\mathbf{T}$ is said to be *convergent* if $\lim_{t\to\infty} \mathbf{T}^t\mathbf{p}^{(0)}$ exists for all vectors $\mathbf{p}^{(0)} \in [0,1]^n$. This can be interpreted as if a matrix reaches a steady state, where multiplying it onto itself results in itself, from any starting vector of initial values $\mathbf{p}^{(0)}$, then it is convergent. For this convergence to happen theoretically, a matrix needs to be *row stochastic*, *strongly connected*, and *aperiodic*. A matrix $\mathbf{T}$ is said to be *row stochastic* if all rows in the matrix sum to 1. A matrix $\mathbf{T}$ is said to be *strongly connected* if there is a path in $\mathbf{T}$ from any node to any other node. A matrix $\mathbf{T}$ is said to *aperiodic* if the greatest common divisor of the lengths of its simple cycles is 1.[3] Golub and Jackson (2010) use these definitions in the DeGroot (1974) setting, showing that, if a matrix is *strongly connected*, it is also *aperiodoc*, and thus, converges to where there is a unique left-hand unit eigenvector $\mathbf{s}$, whose values sum to one $\sum_{i=1}^{n} s_i = 1$. Intuitively, this means that if a matrix is *strongly connected* the limiting beliefs are the weighted average of initial beliefs, with agent i's weight being s_i, which is known as the *social influence* of agent i (Golub and Jackson (2010)). This is an important result because the characterized $\mathbf{s}$ holds for *every* $\mathbf{p}^{(0)} \in [0,1]$, entailing that if the matrix is *strongly connected* the left-hand unit eigenvector exists, and the social influence of agent i is then characterized and unique. I derive the initial "belief" matrix $\mathbf{T}$ and its left-hand unit eigenvector $\mathbf{s}$ below, and in very simulation I use the Depth First Search (DFS) algorithm and eigenvector property that $\mathbf{s}_{1\times n}\mathbf{T}_{n\times n} = \mathbf{s}_{1\times n}$ to verify strong connectivity, convergence and uniqueness in each example.

2.2 A Transition Matrix through Common Connections

Societal interaction can be seen as $\mathbf{T}_{n\times n}$ matrix, but this matrix is not always trivial to obtain. Most of the time this matrix is acquired through surveys that ask agents who they have close ties with (Sampson (1968), Banerjee et al. (2013) and Krackhardt (1987)). Collecting this data is usually

[3] Golub and Jackson (2010) provide more detailed examples to explain these properties, so I do not reiterate the discussion here.

tedious and costly, and results can be inaccurate since these types of questions are usually intrusive, and thus, Hawthorn effects may contaminate and bias results similar to a measurement error in the econometric literature (Landsberger (1958)). Similar problems arise when collecting personal information on individuals to help explain the reason why these "close ties" exist in the first place. To circumvent these difficulties, I propose a general framework, with supporting examples, where n agents first connect with m other agents outside of the network, and then build a $\mathbf{T}_{n\times n}$ matrix based on how similar their connections are outside of the network. Intuitively, this process resembles a "who do you know" type of scenario where agents build social ties based on their commonalities outside of the network. The social interaction process is the following:

1. A set of n and m different agents are realized, where the ith agent in n can choose to connect with as many j agents (one or more) in m as they would like. It is assumed that all m agents are open to connecting, and that the connection represents a one-way connection where agent i is the sender and agent j is the receiver.

2. All agents independently and simultaneously make their connections, and a matrix $\mathbf{A}_{n\times m}$ is realized, as follows

$$\mathbf{A}_{n\times m} = \begin{bmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{im} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nm} \end{bmatrix}$$

   where a_{ij} represents a connection, weighted or unweighted, for agents i and j.

3. All n agents then have a initial interaction with each other (including themselves) about the connections they have within the group of m agents. This initial interaction takes the general form of

$$\mathbf{T}_{n\times n} = \mathbf{A}_{n\times m} \cdot (\mathbf{A}_{n\times m})^T$$
$$= \begin{bmatrix} t_{11} & \dots & t_{1k} & \dots & t_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ t_{i1} & \dots & t_{ik} & \dots & t_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ t_{n1} & \dots & t_{nk} & \dots & t_{nn} \end{bmatrix}$$

where each element in matrix $\mathbf{T}_{n\times n}$ is $t_{ik} = \sum_{j=1}^{m} f(a_{ij}, a_{kj})$, and the functional form $f()$ represents how agents initially interact. This process can be interpreted as individuals building a connection based on their "common connections" within the set of m agents, and mimics that of the "who do you know behavior."

4. Agents then reevaluate each interaction and weigh them relative to all the other initial interactions $t^*_{ik} = t_{ik} / \sum_{i=1}^{n} t_{ik}$. This process yields a row stochastic transition matrix that is structurally equivalent to the matrix presented in DeGroot (1974)

$$\mathbf{T}^*_{n\times n} = \begin{bmatrix} t^*_{11} & \dots & t^*_{1k} & \dots & t^*_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ t^*_{i1} & \dots & t^*_{ik} & \dots & t^*_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ t^*_{n1} & \dots & t^*_{nk} & \dots & t^*_{nn} \end{bmatrix}$$

where each t^*_{ik} represents the weight agent i puts on agent k's opinion relative to the weight agent i puts on all other agent opinions.

5. Given $\mathbf{T}^*_{n\times n}$ is *strongly connected*, the left-hand unit eigenvector $\mathbf{s}_{1\times n}$ of $\mathbf{T}^*_{n\times n}$ is calculated using the convergence results from Golub and Jackson (2010)'s Proposition 1, presented in

Section 2.1.1, such that

$$\lim_{t\to\infty} \mathbf{T}^{*^t}\mathbf{p}^{(0)} = \mathbf{s}\mathbf{p}^{(0)}$$

for any initial starting belief vector $\mathbf{p}^{(0)}$, and where each element s_i of $\mathbf{s}$ is the social influence of agent $i \in n$. Note that this result can be verified by $\mathbf{s}_{1\times n}\mathbf{T}_{n\times n} = \mathbf{s}_{1\times n}$, which is done for all examples below.

Intuitively, agents reveal their types through outside connections, and these connections dictate how closely related the n agents are to each other based on their commonalities.[4] The relationship between the n and m agents can also be applied to other contexts and represent different behaviors including: n agents revealing their connections to m different media outlets, m different political candidates, or m different recruiters at a networking event. For the theme of this paper, it can be interpreted as n agents revealing their connections, through their following decisions, to m different experts that disseminate information about a particular topic on Twitter. This interaction can take on many different interpretations depending on the functional form of $f(a_{ij}, a_{kj})$, and I provide examples of different functional forms in the following section.

2.3 Example 1: Any Common Connection

Consider an indicator matrix $\mathbf{A}_{n\times m}$ where $a_{ij} = 1$ if agent i and agent j mutually connect, and zero otherwise. A matrix $\mathbf{T}_{n\times n}$ represents a matrix of common connections if agent i and agent k have at least one the jth connection in common. More formally, each element t_{ik} in matrix $\mathbf{T}_{n\times n}$ is defined as

$$t_{ik} = \begin{cases} 1, & \text{if } \sum_{j=1}^{m} a_{ij}a_{kj} \geq 1 \\ 0, & \text{otherwise} \end{cases}$$

[4] This interaction can also be interpreted as a "first impression" where agents reveal "who they know" and then interact based on how close their preferences are to each other.

where $t^*_{ik} = t_{ik} / \sum_{i=1}^{n} t_{ik}$ is then the normalized vector for each agent i in $\mathbf{T}^*$. Intuitively, $\mathbf{T}^*$ represents the level of precision agent i puts on agent k's opinion when agents have *at least one* connection in common. Agents then learn from each other based on this precision, and assuming $\mathbf{T}^*$ is strongly connected, convergence happens and a social influence vector $\mathbf{s}$ is realized.

2.4 Example 2: Weighted Common Connections

Consider the same indicator matrix $\mathbf{A}_{n\times m}$ in Example 1 (Section 2.3), but where each element in $\mathbf{T}_{n\times n}$ is defined as

$$t_{ik} = \sum_{j=1}^{m} a_{ij} a_{kj}$$

where $t^*_{ik} = t_{ik} / \sum_{i=1}^{n} t_{ik}$ still represents the normalized vector for each agent i in $\mathbf{T}^*$. Intuitively, this matrix represents agents weighing their connections they have in common where the more commonalities they have the higher precision weight agent i puts on agent k's opinion.

2.5 Example 3: Estimated Weighted Common Connections

Consider the same indicator matrix $\mathbf{A}_{n\times m}$ in Example 1 (Section 2.3), but where each dichotomous choice can be predicted by a variation of the generalized linear mixed model (GLMM)

$$Pr(A_{ij} = 1 | \mu, \alpha_j, \beta_i, \gamma, \theta_i, \phi_j) = [1 + exp(-\pi_{ij})]^{-1} \tag{2}$$

where $\pi_{ij} = \mu + \alpha_j + \beta_i - \gamma|\theta_i - \phi_j|$, $\mu \in \mathbb{R}$ is a fixed effect intercept, $\alpha_j \in \mathbb{R}$ and $\beta_i \in \mathbb{R}$ represent individual random effects observed in the link matrix, $\phi_j \in \mathbb{R}^k$ and $\theta_i \in \mathbb{R}^k$ are latent variables representing spatial positions in the network structure, and γ is the parameter that governs this weight. α_j and β_i are parameters characterized by the the number of links each individual has in matrix $A_{n\times m}$, and ϕ_j and θ_i are parameters in k-dimensional space where the probability of realizing a link depends on the similarity between these parameters. The specification in Equation 2 follows that of Hoff et al. (2002), Barberá (2015), and Navelski and Pascual (2022), which all propose models that assume links are formed based on how closely related their latent positions are in space. The euclidean distant specification of these latent parameters implies that social networks

are "homophilic." I use Navelski and Pascual (2022) method to map these latent positions on a spectrum between -1 and 1 for ease of interpretation and comparison,

Data is used to fit the GLMM and each predictive element is estimated where

$$\widehat{a}_{ij} = [1 + exp(-\widehat{\pi}_{ij})]^{-1}$$

and the initial social interaction rule is defined as

$$t_{ik} = \sum_{j=1}^{m} \widehat{a}_{ij}\widehat{a}_{kj}$$

where each $\widehat{a}_{ij}$ is an index on how similar agent i is to agent j, and t_{ik} defines a weighted relationship between agents i and k based on their compared similarities to agent j. Like all other examples, each agent i reevaluates their relationship with agent k relative to all n relationships $t^*_{ik} = t_{ik} / \sum_{i=1}^{n} t_{ik}$ to form a vector for agent i in $\mathbf{T}^*$. Intuitively, this social interaction matrix $\mathbf{T}^*$ represents agents comparing how close they are to each agent j, $\widehat{a}_{ij}$ and $\widehat{a}_{kj}$, and then deriving a linear relationship based on these weighted comparisons.

2.5.1 MC-MC Estimation

The model parameters are unknown and the statistical problem is to take inference on $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m)'$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)'$, $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_m)'$, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)'$, γ, and μ. Under the assumption of logical independence (i.e. individual following decisions are independent across all users n and m, conditional on the parameters), the likelihood function to maximize is defined as:

$$p(\mathbf{y}|\boldsymbol{\mu}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma, \boldsymbol{\theta}, \boldsymbol{\phi}) = \prod_{i=1}^{n}\prod_{j=1}^{m} \text{logit}^{-1}(\pi_{ij})^{y_{ij}} \left[1 - \text{logit}^{-1}(\pi_{ij})\right]^{1-y_{ij}} \tag{3}$$

where $\pi_{ij} = \mu + \alpha_j + \beta_i - \gamma|\theta_i - \phi_j|$. The complexity of this equation makes direct estimation using maximum likelihood highly intractable because there is no analytical solution to this maximization problem and because there are an infinite number of θ_i and ϕ_j combinations that will produce the same Euclidean distance. Also, maximum likelihood becomes even more difficult as datasets become larger because any matrix with m expert agents and n following agents yields a total of

$m \times n$ simultaneous models to fit with $2 \times (m + n) + 2$ parameters to estimates since there are m α_j's, m ϕ_j's, n β_i's, n θ_i's, 1 intercept μ, and 1 normalizing constant γ. To overcome the tractability problem, I follow Navelski and Pascual (2022)'s Markov-Chain Monte Carlo (MC-MC) method with a hierarchical setup to generate samples from the posterior distribution where each of the parameters μ, α_j, β_i, γ, θ_i, and ϕ_j are assumed to be drawn from respective common prior population distributions. I provide more detail on how to implement this estimation method in the Appendix Section 6.1.

3 Data: Genome Editing in Domestic Livestock (GEDL)

Social media data is a natural setting for this analysis because of the network of followers it provides where individuals can choose to connect with or "follow" an expert in the field they are interested in. I use network data from Twitter because about one in four US adults (23%) say they use this social network and Twitter allows academics to conduct research using information from their public accounts (Pew Research Center (2021) Twitter Inc. (2022a)). I apply my model to a panel of $m = 46$ experts, and their followers, from the genome editing in domestic livestock (GEDL) field with the assumption that individuals follow these experts to gain information about GEDL and because their preferences are homophilic. McPherson et al. (2001) makes a strong and well documented case that this behavior is true in practice for social networks, and this assumption has been applied in other empirical network formation models (Graham (2017),Chandrasekhar et al. (2020)). With these assumptions, I create a choice matrix $\mathbf{A}_{n \times m}$ with $m = 46$ experts and $n = 187,209$ followers. The $m = 46$ experts were chosen based on how informed they are about GEDL and on their ideological position about GEDL.[5] The experts are initially labeled as being "Anti" or "Pro" GEDL based on the message they are trying to disseminate from their account, but this labeling is only used for preliminary analyses to motivate the data structure.[6]

[5]The list of $m = 46$ experts is supported by large players in the genome editing industry, like Synthego, who lists many of these accounts on their web pages as accounts to follow on Twitter for information about genome editing (Prabhune, M. (2019); Synthego (2022)). Additionally, many of these accounts are officially verified by Twitter, which implies the account is authentic, notable, and active (Twitter Inc. (2022b)). I use Twitter's Academic Research's application programming interface (API) to obtain all of the followers' unique user IDs that follow each expert account on Twitter, and merge them based on these IDs to create $\mathbf{A}_{n \times m}$ (Twitter Inc. (2022a)).

[6]This labeling will not affect future analyses as the estimation technique allows individuals to move on a continuous spectrum from -1 to 1 which represents being the most Anti GEDL vs. the most Pro GEDL.

The full dataset has 46 experts ($m = 46$), 12 perceived to be Anti and 34 perceived to be Pro GEDL, and Figure 12 in the Appendix presents their account information from a Twitter API query conducted in January 2022. OrganicConsumer and NonGMOProject have the most amount of followers at 187,209 and 125,516, respectively, and Recombinetics has the least amount of followers at 1,238. Table 1 presents the Anti and Pro group averages where the average number of followers per Anti account is greater than Pro with 59,596 followers compared to 15,223 followers. Additionally, the Anti GEDL experts are more active on Twitter than the Pro experts because they tend to follow more accounts on Twitter and Tweet more on average. These results imply the average Anti GEDL experts are more connected and engaged to the Twitter network than the Pro GEDL experts.

Table 1: Mean Statistics and Standard Deviations by Viewpoint (Raw)

Viewpoint	Followers	Following	Tweets
Anti Genome Editing	59,596 (51,190)	6,373 (8,381)	28,214 (31,220)
Not Opposed to Genome Editing	15,223 (15,536)	2,665 (3,274)	15,041 (18,700)

Once the full dataset is collected, I reduce the choice matrix to those followers that follow greater than 8 out of the 46 accounts to ensure all t_{ij}'s are strongly connected and help with estimation tractability. This reduces the set of followers to $n = 3,383$, and this reduction slightly alters the research question to be focused on the "more informed" followers in regards to GEDL. The decision to only look at the more informed followers does not seem to be a strong assumption because the mean percentage of followers per account in each group does not change much with percentages moving from 79.65% to 73.79% in the reduced dataset for the Anti GEDL accounts. Similarly, the percentage of total follows does not change much from 58.93% to 49.86% for the Anti GEDL accounts. The number of followers per account in the reduced dataset $\mathbf{A}_{n\times m}$ is presented in Table 2, and Table 3 presents the Anti and Pro GEDL group averages where the Anti accounts still have, as a group, more followers (59,596) per expert account than the Pro account experts (15,223).

Table 2: Total Number of Followers per Influencer Account (Reduced)

Screen Name	Followers	Position	Screen Name	Followers	Position
GMOEvidence	1184	Anti	Recombinetics	139	Pro
USRightToKnow	582	Anti	AzMilkProducers	62	Pro
JoelSalatin	330	Anti	CRISPRchef	474	Pro
nongmoreport	1795	Anti	AquaBountyFarms	131	Pro
GMWatch	1824	Anti	joeBondyDenomy	431	Pro
RachelsNews	959	Anti	FrancoiseBaylis	158	Pro
CFSTrueFood	1825	Anti	jcornlab	746	Pro
GMOFreeUSA	1811	Anti	AprilPawluk	514	Pro
OrganicTrade	1665	Anti	jsherkow	236	Pro
OrganicValley	1583	Anti	shsternberg	735	Pro
NonGMOProject	1978	Anti	JKamens	326	Pro
OrganicConsumer	1936	Anti	mem_somerville	402	Pro
TysonFoods	305	Pro	Synthego	708	Pro
Cargill	309	Pro	pcronald	523	Pro
Kevin_Faulconer	45	Pro	JonEntine	441	Pro
doudna_lab	1119	Pro	ELS_Genetics	119	Pro
CRISPR_News	1008	Pro	KevinADavies	659	Pro
SynBioBeta	873	Pro	BioBeef	665	Pro
AgBioWorld	675	Pro	igisci	881	Pro
pknoepfler	718	Pro	CamiDRyan	395	Pro
GeneticLiteracy	736	Pro	nmpf	231	Pro
CRISPRjournal	1038	Pro	pdhsu	814	Pro
GaetanBurgio	760	Pro	NPPC	196	Pro

Table 3: Mean Follower Accounts and Standard Deviations by Group (Reduced)

Group	Followers
Anti	1456 (558)
Pro	517 (302)

The reduced dataset will be the choice matrix used in all subsequent analyses, and to give readers a better understanding about the structure of matrix $\mathbf{A}_{n\times m}$ used in the analysis, I plot a heat-map of $\mathbf{A}_{n\times m}$ in Figure 1. Figure 1 lists all m expert accounts in columns while the followers are represented by rows. A black "dash mark" indicates that user following that expert account, where in the data it is represented as a one.[7] One thing to observe about the data structure is that many of the followers in the north-west quadrant are following a large portion of the Anti

[7]The scale is presented in continuous form because we will reference this scale in future heat-maps.

GEDL accounts, indicated by the dark black mass, while the followers in the south-east quadrant are more sparse and not following a large proportion of the Pro accounts, indicated by the patchy black and white area. This indicates Anti GEDL account followers tend to be more "loyal" whereas Pro account followers tend to have more of a "smattering," or mix, in their following structure instead of following many or all of them at once. Another key observation about the structure of the adjacency matrix is the lack of extremely informed followers that follow many of the 46 accounts. These followers are presented in the middle of the matrix, and it is apparent that out of these followers many of them avidly follow the Anti accounts while following a spread of the Pro accounts. This indicates that even though some followers are extremely informed about GEDL, most of them are more informed through the Anti experts rather than the Pro experts. Lastly, and perhaps most important, note that the choice matrix appears to be *strongly connected* where there is a path, even if it is not direct, from any follower i to expert j, and vice versa. This implies that when the dot product of $\mathbf{A}_{n\times m}$ is taken, $\mathbf{T}_{n\times n} = \mathbf{A}_{n\times m} \cdot (\mathbf{A}_{n\times m})^T$, it will satisfy the properties needed to derive the unique social influence vector of $\mathbf{s}$. I also use a Depth First Search (DFS) algorithm to verify that I can arrive at any node starting from every node.[8]

[8]The Depth First Search (DFS) algorithm checks to see if any node in a matrix can be reached starting from every other node in the matrix, and vice-versa

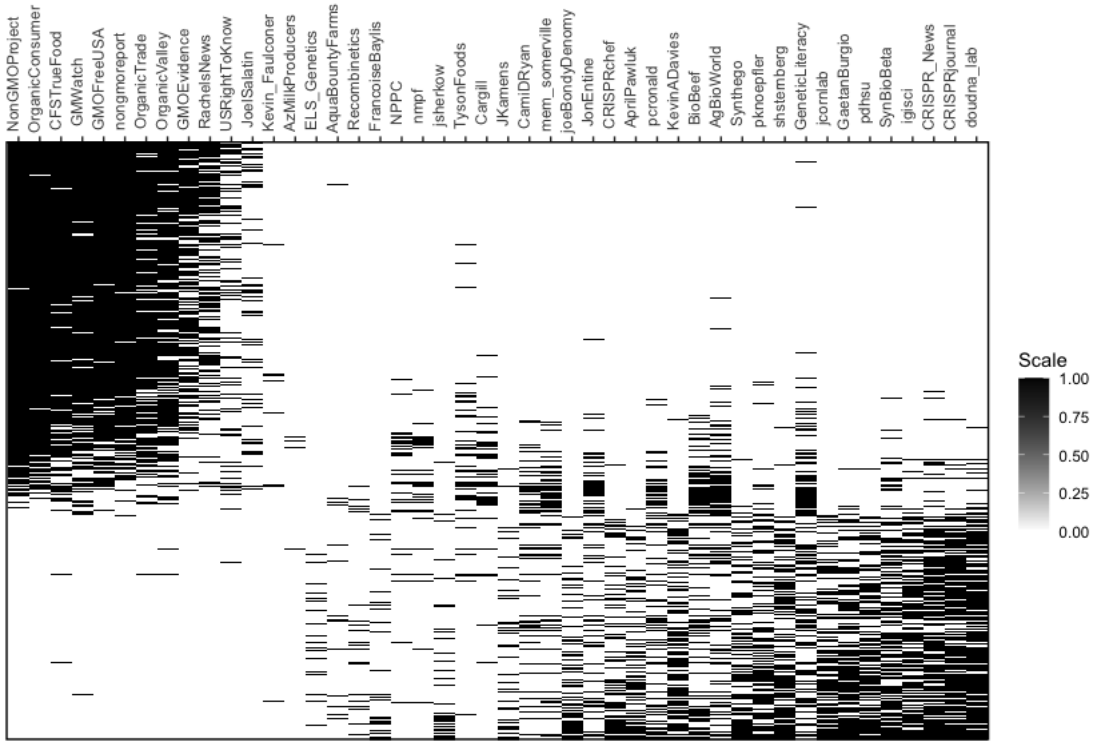

Figure 1: Heat Map of Adjacency Matrix

3.1 Estimating the Data Structure

I estimate the connection matrix $\mathbf{A}_{n\times m}$ from Figure 1 using the model and methodology outlined in Section 2.5. Barberá (2015) and Navelski and Pascual (2022) also use Twitter data in their analyses, one dataset based on political elites and their followers and one dataset based on US media influencers and their followers, and interpret these estimated latent parameters as "ideal" points.[9] This interpretation is deeply ingrained in the probabilistic voting model literature where agents vote for candidates based on how close their "ideology" or "reputations" align (Enelow and Hinich (1984), Coughlin and Nitzan (1981), Enelow and Hinich (1989), Coughlin (1992)). This literature also proposes that these ideology parameters can be functions of many other parameters that explain the position of this ideological parameter. For example, the ideological position of a candidate can be a function of an array of political positions and/or it can be a function of other individual characteristics like charisma or athleticism. For this, I use the ideal point terminology for

[9] An "influencer" is defined by Merriam-Webster as "a person who is able to generate interest in something (such as a consumer product) by posting about it on social media," and a "follower" is an individual that follows that influencer because they have similar interests or ideologies (Merriam-Webster, 2021).

ease of reference and because of its repeatability to the probabilistic voting literature. Additionally, since the random effects of α_j and β_i are estimated using the observed individual effects that explain the connection between each agent j and i (i.e., the number of total followers for each agent j and the number of total follows for each agent i), these parameters can be interpreted as agent j's popularity and agent i's engagement (Barberá (2015) and Navelski and Pascual (2022)).

All MC-MC diagnostics yield desirable results. All $\widehat{R}$ values are less than 1.1, which is standard practice, implying that all chains have converged to the same posterior distribution, and thus, there is no divergence in the MC-MC estimation process. This diagnostic is important because if convergence is satisfied and the likliehood function is in the same form as Equation 3, then the estimates are consistent and hypothesis testing can be conducted. The MC-MC diagnostics are supported by model fit diagnostics where, using all 155,618 individual decisions as observations, the prediction rate is 88.5% accurate.[10] To further motivate estimation results, Figure 2 plots the predicted probabilities in a heat-map, which closely resembles the raw data structure presented in Figure 1.

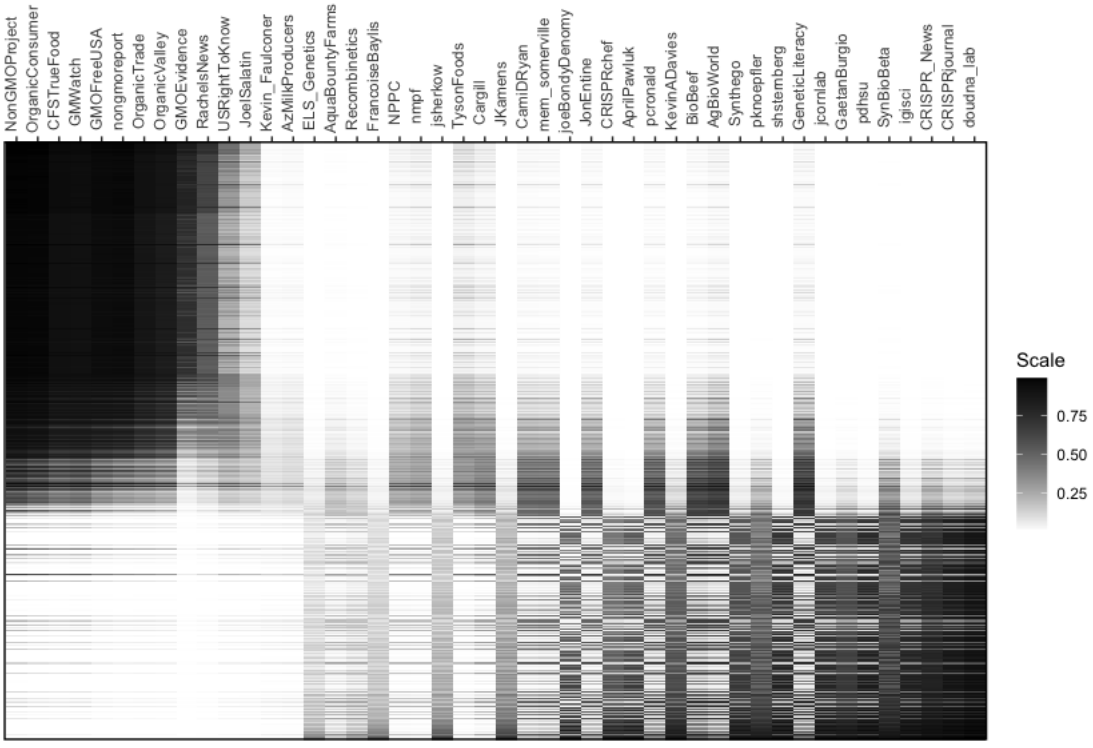

Figure 2: Example 3: Heat Map of Adjacency Matrix (Predicted Data)

[10] I provide a more detailed explanation about prediction diagnostics in Appendix Section 6.4.

Results show that @NonGMOProject is the most popular expert with an estimate of $\widehat{\alpha}_2 = 3.78$, and that @Kevin_Faulconer is the least popular expert with an estimate of have $\widehat{\alpha}_{15} = -3.30$. These results are interesting, but also expected since @NonGMOProject and @Kevin_Faulconer have the most and lease amount of followers in the data. The more interesting result about the experts are their estimated ideal positions where the most Anti GEDL expert is @GMOEvidence ($\widehat{\phi}_{12} = -.687$) and the most Pro GEDL expert is @joeBondyDenomy ($\widehat{\phi}_{42} = .847$). Figure 3 plots all of the expert's estimated ideologies on a line segment ranging from -1 to 1, with a segment in the middle representing a zero line, where I highlight some experts to show how their estimates align with the detailed information in their profile. For example, @NPPC, @Cargill and @TysonFoods are all representatives of the meat producing industry, where it could be that they were more Pro GEDL to reduce production costs, but they are seen to have ideologies that are more moderate. To further support these results, @doudna_lab is the official Twitter account for Dr. Jennifer Doudna's lab, and Dr. Doudna was awarded the 2020 Nobel Prize in Chemistry, with Emmanuelle Charpentier, for their methodological developments in genome editing. These developments were essentially, the first discovery of CRISPR, and it is not surprising that her lab's ideal point is positioned on the more Pro GEDL side of the spectrum at 0.72.[11]

[11] A summary and discussion about expert popularity and ideology is presented in the Appendix Section 6.3.

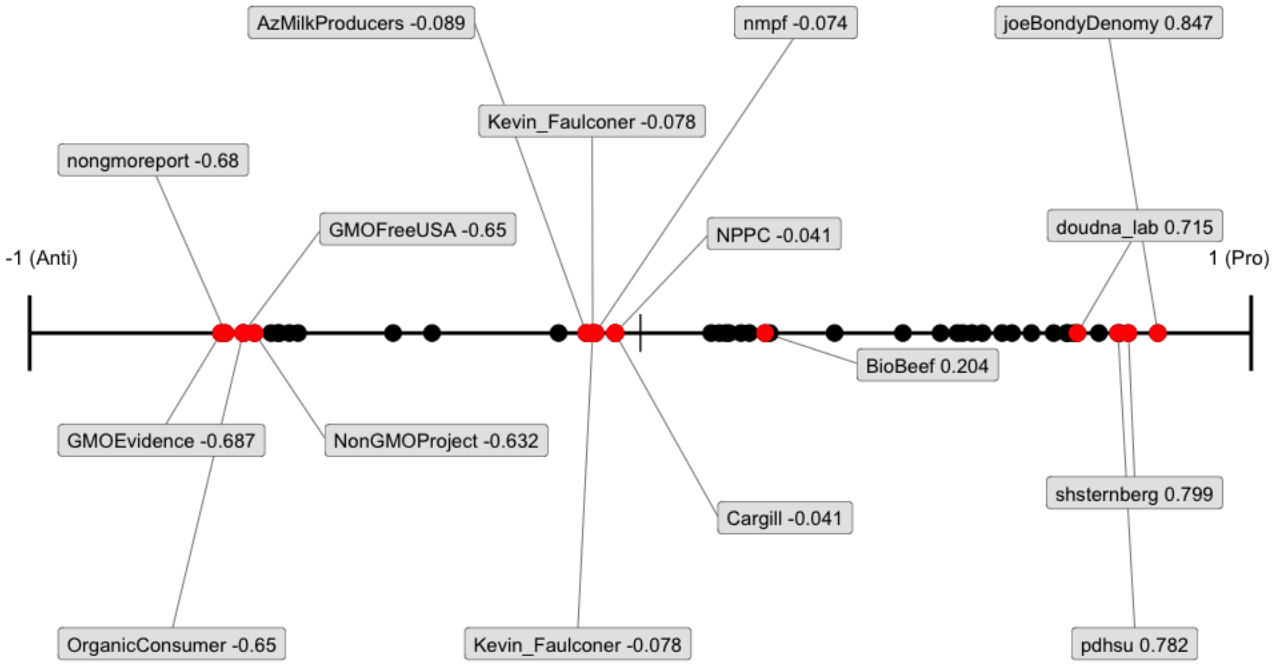

Figure 3: Predicted Ideology of Experts $\widehat{\phi}$

The estimated engagement parameters for all $n = 3,383$ followers $\widehat{\beta}_i$ are plotted in Figure 4 (a) as a histogram. Figure 4 (a) has a short "fat" tail on the negative side of the distribution and a long "thin" tail on the positive side indicating that overall engagement in GEDL is really high for some followers, but low for most. Figure 4 (b) shows the histogram of all $n = 3,383$ ideal points $\widehat{\theta}_i$ where many individuals are polarized about GEDL. The mean and median ideology are -0.074 and -0.19 implying that the average informed follower about GEDL will have an ideology that is more Anti GEDL. These metrics are represented by the solid and dotted lines in Figure 4 (b), and this result is even more apparent when observing the large "spike" in individuals on the Anti GEDL side. 36.54% of the followers are between the scale of -0.5 to 0.5, implying that there are not as many followers with moderate ideologies compared to the extreme ideology. This observation is further supported by where the middle 50% of the data is located because both the 25th and 75th quantiles are below and above these thresholds at $-.64$ and .51, respectively. These two metrics are represented by the dotted lines on the edges of Figure 4 (b)'s histogram. This implies that there are

some followers that have moderate ideologies, and this is an interesting result because one would think a polarizing topic like GEDL would create a stark separation in the ideological distribution, which is does.

(a) (b)

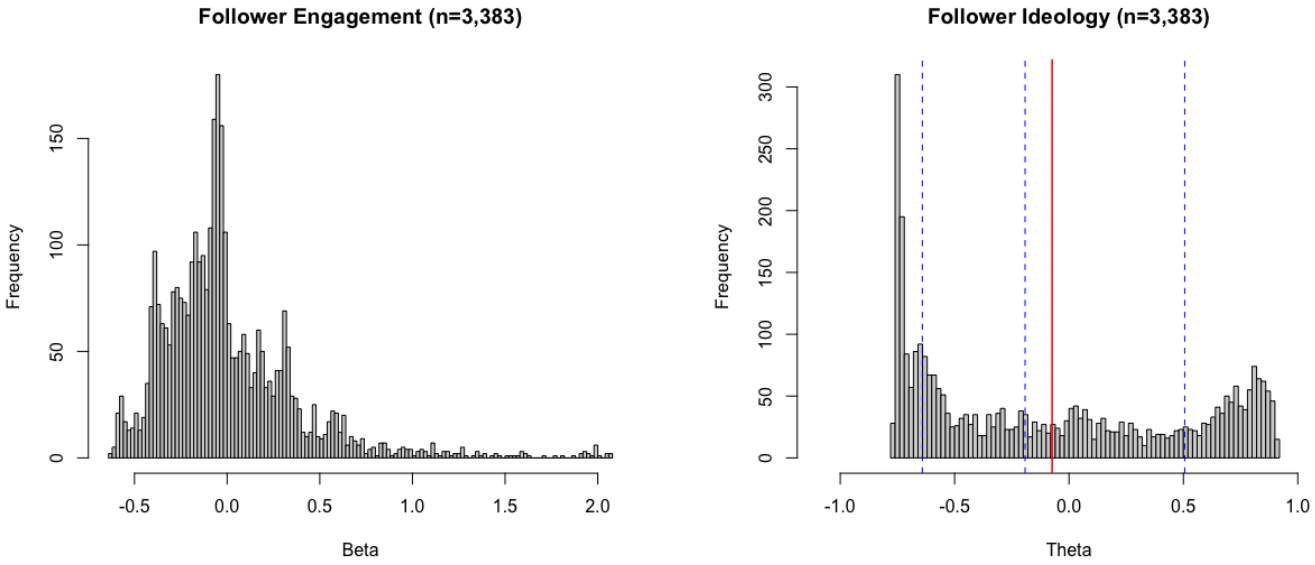

Figure 4: Ideology $\widehat{\beta}_i$ (a) and Engagement $\widehat{\theta}_i$ (b) of $n = 3,383$ Followers

Figure 1 and Figure 1 are used as the initial choice matrices $\mathbf{A}_{n\times m}$ for Example 1 and 2, and in Example 3, respectively. In each example, a trust matrix $\mathbf{T}^*$ is realize and I plot a heat map of these matrices in Figures 17, 18, and 19 in the Appendix Section 6.5 for a visual representation. Note that individuals in the upper north-west and south-east quadrants have strong connections, and these strong connections are a result of them having similar interests with each other, which are represented by maintaining high levels of trust, or probabilities of interacting, when comparing the experts they follow. I derive the social influence vectors $\mathbf{s}$ for each of these matrices and present the results below.

4 Results

4.1 Social Influence

The social influence vectors $\mathbf{s}$, derived from each 'common connection' example in Section 2, all maintain the same location paramter at $\mu_{SI} = 0.0295\%$ but vary in terms of other distributional statistics.[12] Figure 5 and Table 4 show this difference where the CDF of each social influence vector $\mathbf{s}$ varies in shape and where the summary statistics of each social influence vector change about the mean. The difference in CDF shape implies that influence changes depending on how individuals weigh their common connections, and these results are supported in Table 4 where standard deviations, quartile values, maximum and minimum are different. Results show that Example 1 has the lowest standard deviation at 0.007% in social influence, while Example 2 has the highest at 0.0093%. This implies that influence varies more when individuals are allowed to weigh each of their common connections rather than just building a connection off of them having anything any connection in common.

[12] All location paramter estimates are equivalent because by mathematical theory the *left-hand unit eigenvector* $\boldsymbol{s}$ of the matrix $\mathbf{T}^*$ must sum to 1 $\sum_{i=1}^{n} s_i = 1$.

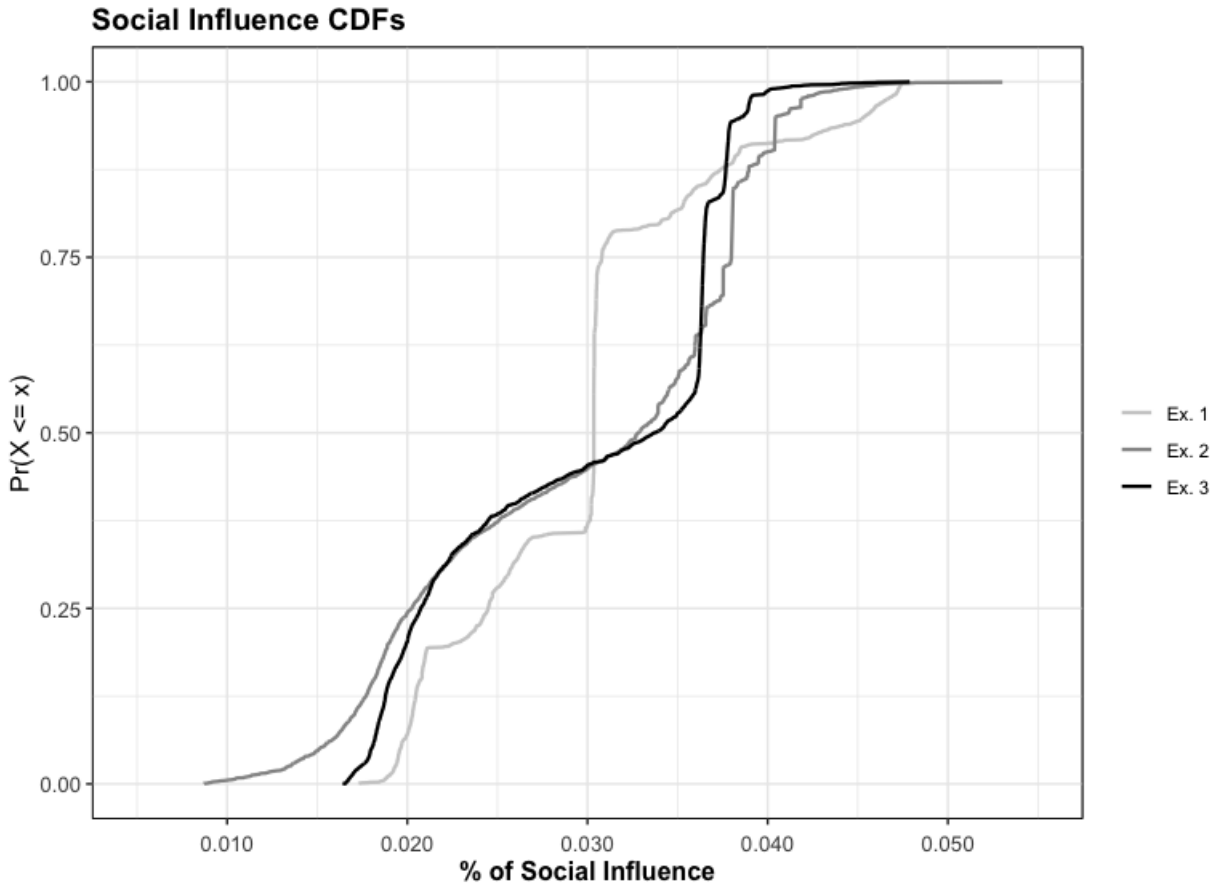

Figure 5: CDF of Relative Social Influence as Influence Ideology Changes.

An increase in variance also implies that influence becomes more unequal because if influence was equal, all s_i values would be at 0.0295%, which is presented in Equal SI row in Table 4 for reference. This result is emphasized by the median percentage of social influence in all three examples at 0.0304 %, 0.0328 %, and 0.0337 %, respectively, where all values are higher than the mean at 0.0295 %. This implies that the social influence distribution is negatively skewed and that there are more individuals with high social influence relative to low. This is an interesting result because it also shows a that there is a trade-off between individuals with high and low social influence.

I use $\mathbf{T}^*$ from Figure 19 to estimate each individuals social influence s_i, and I find the follower with the highest social influence about GEDL is @careygillam, with an influence level of approximately 0.0594%. This result is supported by their Twitter biography where they write that they

Table 4: Summary Statistics for Social All Influence Vectors **s**

Example	Min.	1st Qu.	Median	Mean	SD	3rd Qu.	Max.
1 (Any)	0.0173 %	0.0245 %	0.0303 %	0.0295 %	0.007 %	0.0308 %	0.0474 %
2 (Weighted)	0.00871 %	0.0203 %	0.0328 %	0.0295 %	0.0093 %	0.0381 %	0.0577 %
3 (Est. Weighted)	0.0164 %	0.0207 %	0.0337 %	0.0295 %	0.008 %	0.0364 %	0.0479 %
Equal SI	0.0295 %	0.0295 %	0.0295 %	0.0295 %	0.0295 %		0.0295 %

have been a journalist at Reuters for 17 years.[13] I present this follower's demographics from Twitter in Table 5, along with four other followers with high social influence, and find that the social influence metric does well at identifying followers that would be share information about GEDL. This is apparent because the top five most influential followers all have a lot of individuals that follow them, have profile descriptions that relate to the topic of GEDL, and two out of the five users are officially verified by Twitter.

[13] @careygillam has written two books about the US agriculture industry, and is a correspondent for the guardian. These credentials would indicate that they are very connected in this industry, and that they are disseminating information about their opinion on the topic of genome editing in domestic livestock.

Table 5: Top Five Most Influential Followers

Username	Followers	Description	Verified	Social Influence
careygillam	15,139	Journalist 17 years @Reuters; Author of Whitewash & The Monsanto Papers, and @Guardian contributor	✓	0.0594%
NewFoodMag	14,102	New Food is the leading bi-monthly FREE publication. Keeping the food & beverage industry up-to-date with the latest developments & technology. Sign up today.		0.0583%
EcoWatch	22,6944	Environmental News for a Healthier Planet and Life		0.0582%
KarenJeanHood	187,554	Author, Poet, Freelance Writer, Researcher, Child Advocate, Feminist, Teacher, Literacy Advocate, Cook, & Food Writer, Wife & Mom/16 Great Kids/13 Foster Kids		0.058%
zerowastehome	35,294	"Mother of the zero waste lifestyle movement" CNN - Speaker - Author of Zero Waste Home (in 20+ languages)	✓	0.0573%

4.2 Synthetic Marginal Effects

One benefit to estimating the $\mathbf{T}^*$ matrix is that I can then see how the $\mathbf{T}^*$ heterogeneously changes as I change the ideology of an expert on the -1 to 1 spectrum. This analysis is equivalent to the synthetic version of a marginal effect $\frac{\partial s_i}{\partial \phi_j}$ where I hold everything constant in the linear model from Equation 2 and change a ϕ_j parameter. This change will have an individual heterogenous effect on the probability of following for each follower i in the adjacency matrix, and it will change the empirical moments that describe that distribution. I synthetically show this effect where I change the most popular expert's, @NonGMOProject, latent ideal point from $\phi_2 = -.63$ to a range of ideal points from -1 to 1.

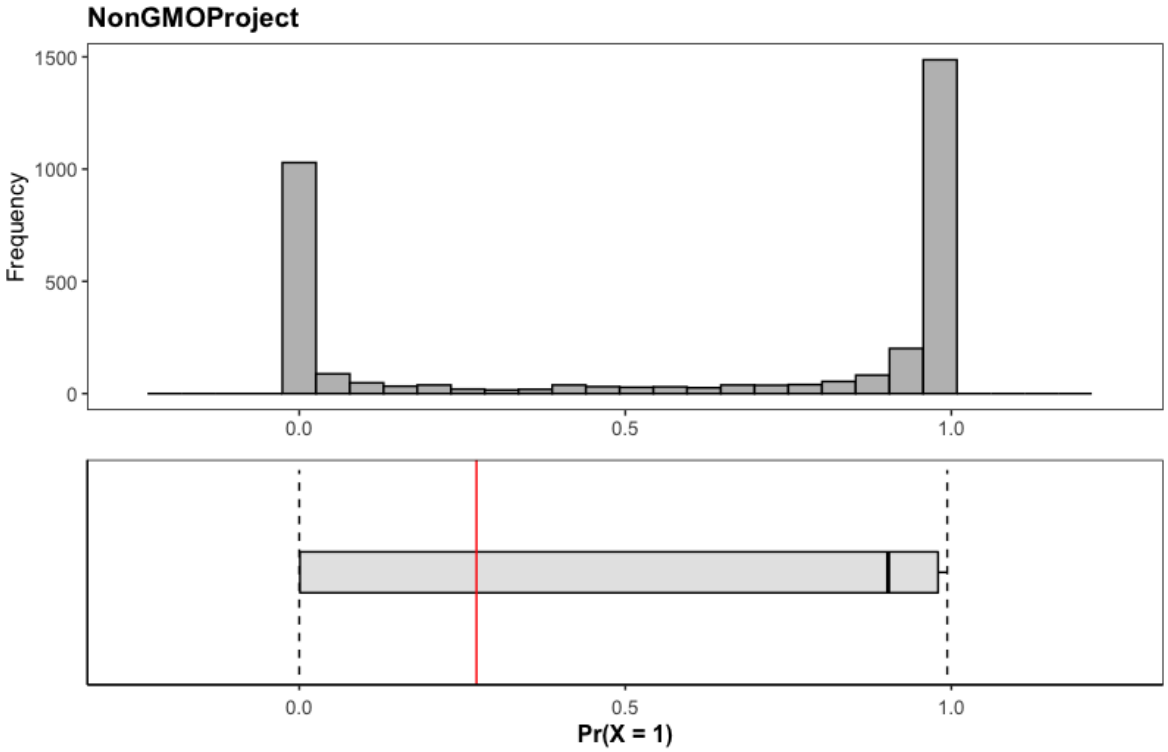

Figure 6: Box-Plot of Probability of Following

Figure 6 shows the distribution of the probability of following @NonGMOProject $\widehat{\mathbf{a}}_{i2}$ when $\widehat{\phi}_2 = -.63$ for reference, and Figure 7 shows a range of box-plots that represent the distribution of the predicted probability of following at each synthetic ideal point. The x-axis is the range of ideology values @NonGMOProject can have, which ranges from -1 to 1 and increases by .1, and the y-axis is the probability of following @NonGMOProject, which indicates the different levels of quantile measurement for each box-plot. Each box-plot shows the minimum (bottom of the whisker in the box-plot), lower quartile (the bottom of the box in the box-plot), the median (the solid black line in the box-plot), the upper quartile (top of the box in the box-plot), and the maximum (the top whisker in the box plot). I also plot the prediction threshold as a solid horizontal line at .272 to give a reference on how many individuals would follow that expert at that ideological position.

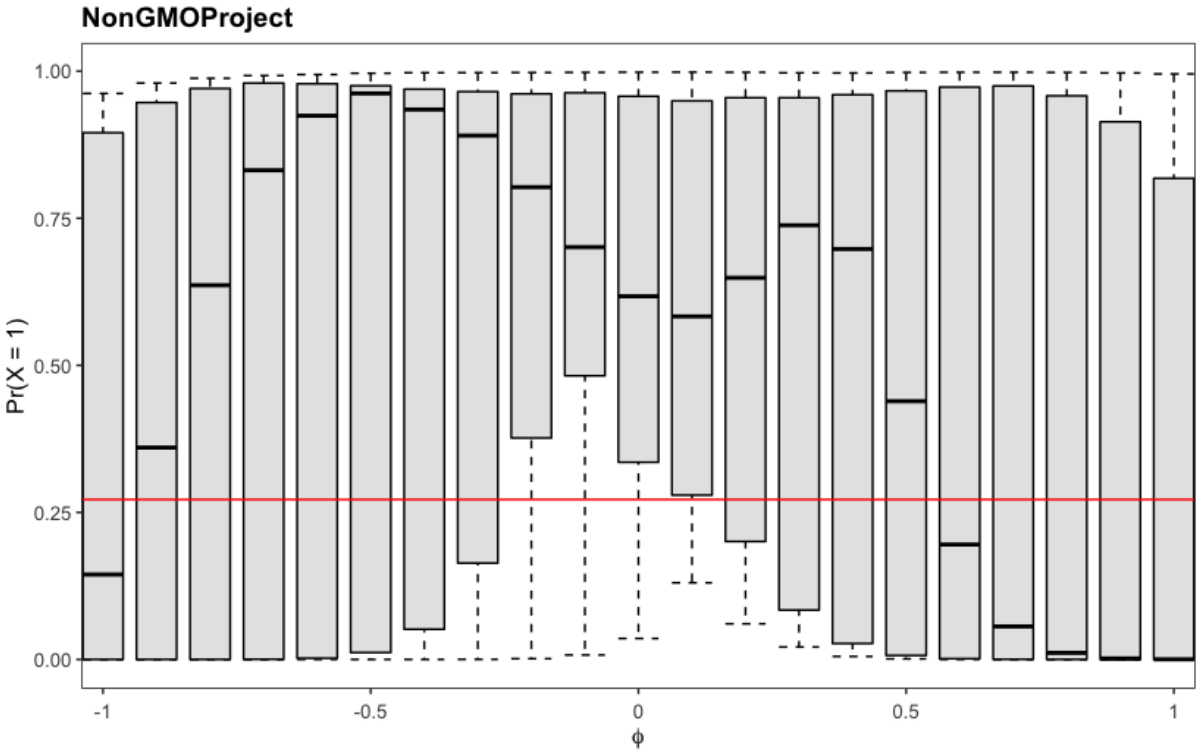

Figure 7: Box-Plot of Probability of Following

It is apparent that the median change in the probability of following is not linear across the ideological values because the median value first starts very low, increases, decreases, increases and then decreases again. This would imply that the median user with that probability moves significantly over @NonGMOProject's ideology, and that the distribution of followers changes significantly with this change. This is supported with all of the other distributional metrics that are presented with each box plot.

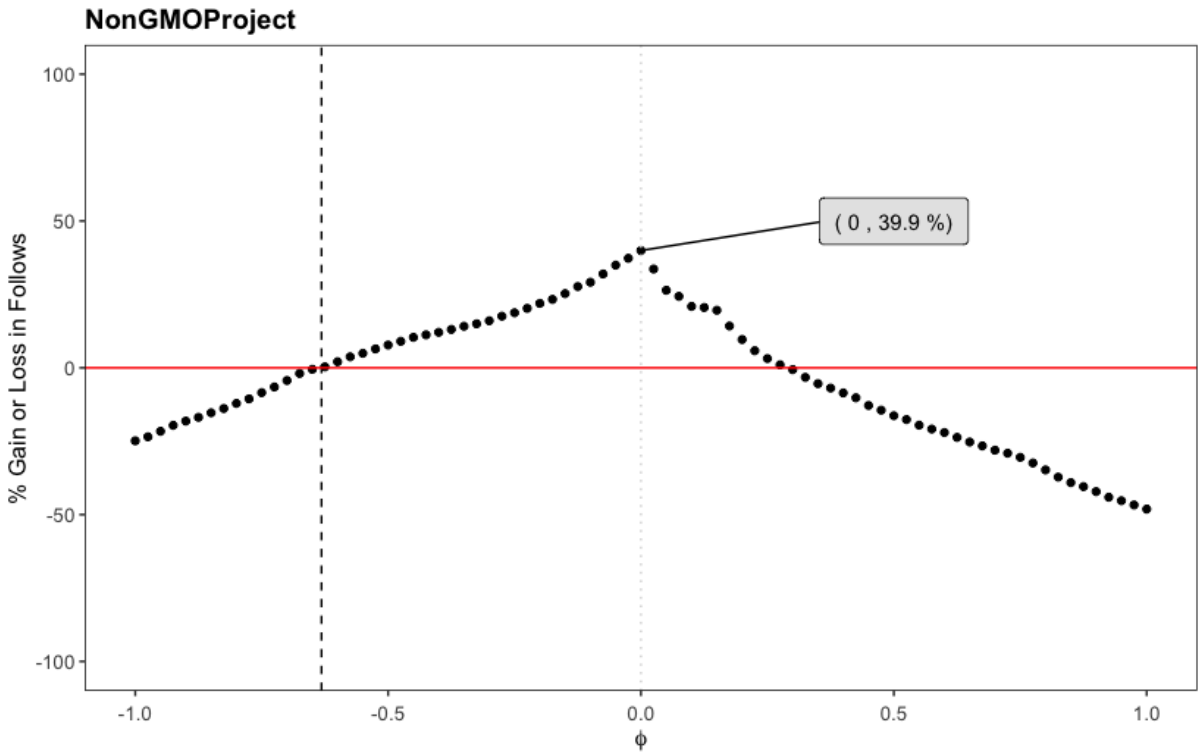

Figure 8: Percent Gain or Loss in Followers (b) by Ideology.

Another interesting question this analysis can help answer is "Given the current following landscape, what ideological position would give the expert the most followers?" The analysis in Figure 8 can help answer this question where I use the original threshold of .272 to predict the percent gain (or loss) in followers across each ideological interval. The x-axis in Figure 8, similar to Figure 7, is the range of ideological values @NonGMOProject can take, while the y-axis is the percent gain in followers at each ideological interval. The zero line implies that there is no gain, and a positive percentage implies there is a positive gain while a negative indicates a percentage loss. The dotted vertical line plots the expert's, in this case @NonGMOProject's, initial ideological point for reference, and the label shows the maximum percent gain in followers if they were to change their ideological position to that point. Results from this analysis show that if @NonGMOProject were to position their ideology at zero, ironically, they would maximize their percent gain in followers to 39.9%. Interesting enough, there is a trade-off for @NonGMOProject where if they were to increase or decrease their ideology on GEDL, they would either lose or gain more followers.

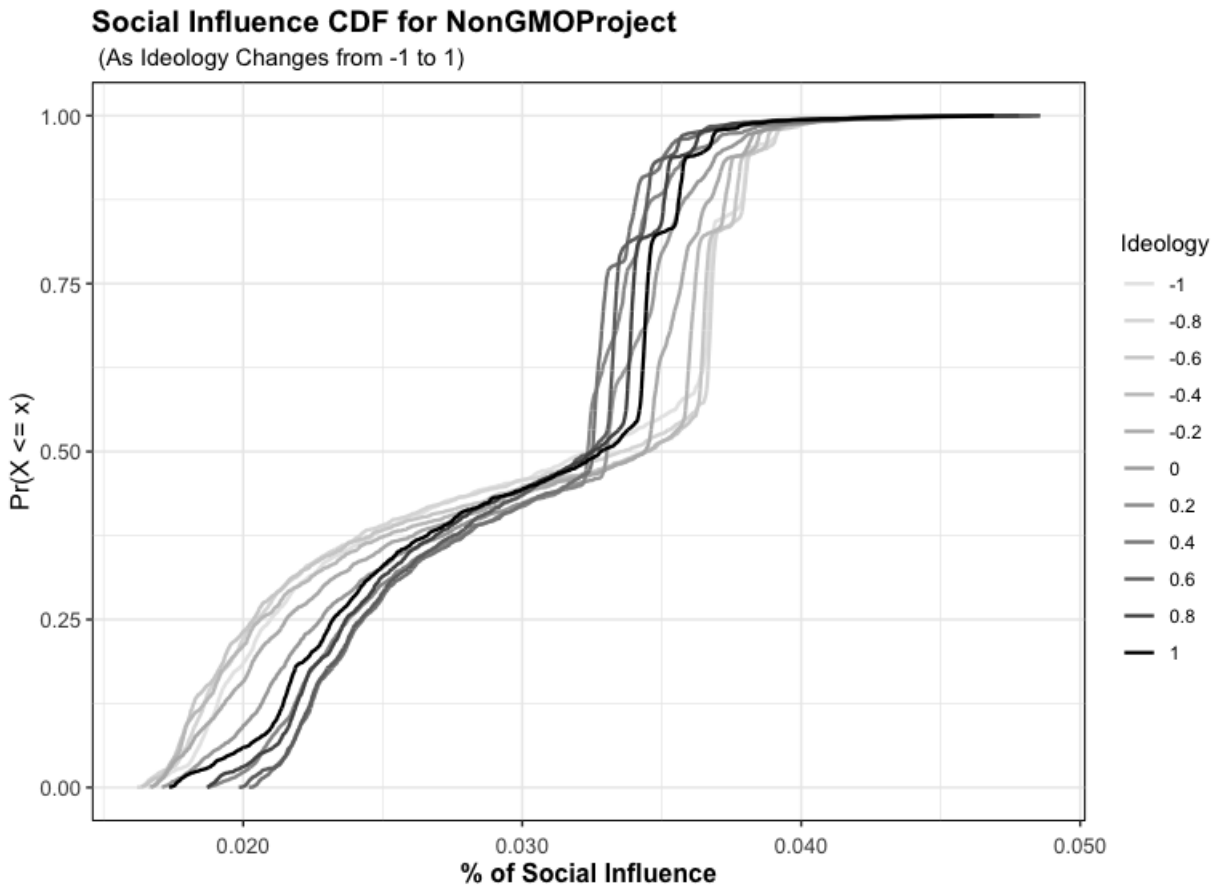

Figure 9: CDF of Follower Social Influence as Expert Ideology Changes

The previous result suggests that if experts change their ideological position, they the total number of their followers and the overall distribution of followers. Similarly, this heterogenous effect of ideological change implicitly affects social interaction and social influence in the sub-network of followers. For example, if an expert changes their ideology, this can affect which individuals follow that expert, and thus change the social influence a follower has on other followers because they are less informed and less connected. Figure 9 depicts this structural change by showing 11 different social influence CDFs, where each CDF represents the social influence distribution realized from changing @NonGMOProject's ideology from -1 to 1 in increments of .1. It is apparent the CDF changes, but the question is "how does it change?" Figures 10 and 11 investigate this change by plotting the percent gain (or loss) in the standard deviation (SD) and quantiles of the social influence vector as @NonGMOProject's ideology changes from -1 to 1. Figures 10 and 11 have

the same interpretation as Figure 8, where there is no percent gain or loss in the SD and any of the quartiles at the vertical dashed line because it is @NonGMOProject's originally estimated ideological position. The percent gain is relative to the social influence estimates at this ideological point, and as I increase @NonGMOProject's ideology towards 1 there is a percent loss in SD until it reaches .45 where the percent loss in SD is 40.62%. This result implies that if @NonGMOProject shifts their ideology to .45 from $-.63$, the social influence will become 40.62% more equal since there is a trade off between individuals with high social influence and low social influence. This trade off can be seen in Figure 11 where there is a percent gain in social influence for the lower two quartiles of social influence (Q1 and Q2) and a percent loss for the upper two quantiles (Q3 and Q4). This is represented by the divergence in the series for each quantile, where the two lighter series (groups Q1 and Q2) increase when moving away from the original ideology while the two darker series (groups Q3 and Q4) decrease. Intuitively, if @NonGMOProject wanted to increase their ideology to maximize the percent gain in a social influence quantile, they would position themselves at .475 to give the lowest quantile (Q1) the greatest gain in social influence, at a gain of 21.4%. Additionally, this ideological shift would gouge into the social influence of the upper two quantiles, since there is a trade-off in social influence, where the largest percent loss would be from the third quantile (Q3) at a loss of 9.3 %.

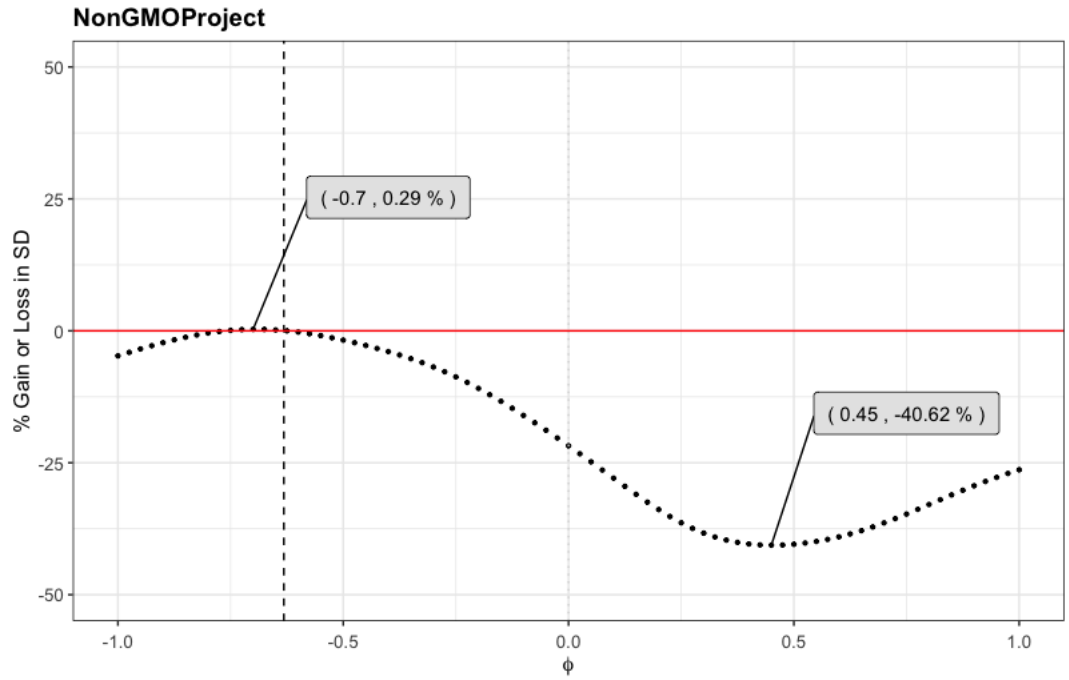

Figure 10: Social Influence as Ideology Changes

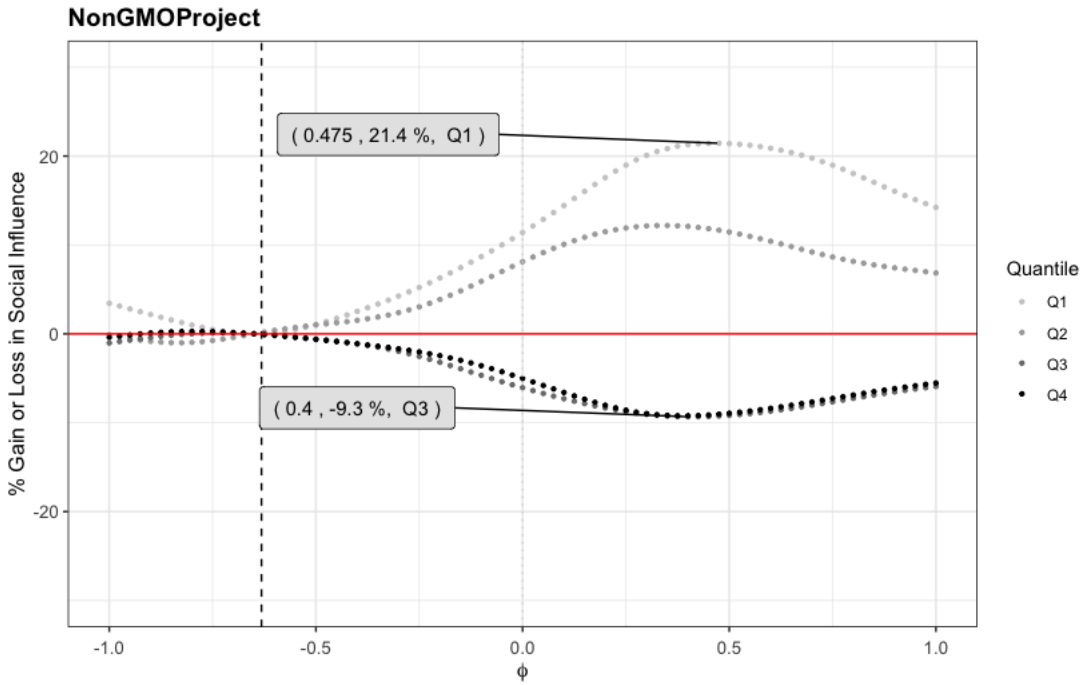

Figure 11: Social Influence as Ideology Changes

5 Discussion

In this paper, I investigate how social influence changes with respect to an agent's position change outside of the social network $\mathbf{T}^*$ in DeGroot (1974)'s leaning model. This method uses a "common connections" assumption where individuals form a network connect based on who they know outside of the network. This yields a network where individual assigns a precision weight all of the opinions they hear, and a vector of influence $\mathbf{s}_n$, with a social influence weight for each individual s_i, is realized using learning convergence theory. To motivate this process, I propose three examples on how the "common connections" interaction can happen, including a method for estimating the heterogeneous weight each individual i puts on connection j. All three examples yield similar social influence distributions, in an application to genome editing in domestic livestock (GEDL), where the location parameter is the same, but the spreads of the distributions differ. This implies that the way agents interact and compare their common connections matter, and can lead to a more or less equal levels of individual social influence. This is a powerful result because it shows that social influence is a function of how the initial weighted social network $\mathbf{T}^*$ is created.

I further support this result by showing, in an application, that agents outside the social network can have an effect on the influence of agents within the social network. To do this, I synthetically adjust the estimated latent parameter of the most popular agent outside of the network, and show how this heterogenous effect changes the social influence distribution. The most popular expert an Anti GEDL expert on Twitter named @NonGMOProject, with a popularity estimate of $\widehat{\alpha}_2 = 3.78$, and I show how changing their estimated latent "ideal" point can heterogeneously effect who follows them and the influence these followers have outside the network. Results show that if @NonGMOProject were to shift their ideology of $\widehat{\phi}_2 = -.63$ to zero (i.e., being completely neutral about the subject), they would have the highest percent gain in followers at a 39.9% gain. This is an important result because it shows there is a trade-off for @NonGMOProject, which is to either gaining more followers or stick to their ideals, and that if they choose to not shift their stance on GEDL more towrds the Pro GEDL side, then they value their message more than their Twitter presence.

This heterogeneous effect also bleeds into the social influence of the agents that follow or don't follow them because if @NonGMOProject were to shift their ideology point to .475 from $\widehat{\phi}_2 = -.63$,

which is an extreme shift from being Anti GEDL to Pro GEDL, they would increase the influence of the lowest quantile of social influence distribution by 21.4%. This implies that they could drastically change the influence within their following network because there is a trade-off between individuals with high influence and low influence. This is an interesting result, even outside of this specific application, because it shows that agents outside the social network have the power to control what the social influence distribution looks like in another network. In other words, an outside agent's position, or stance, on a topic matters and changing this stance can have effects on the social influence in networks they are not a part of.

References

Acemoglu, D. and Ozdaglar, A. (2011). Opinion dynamics and learning in social networks. *Dynamic Games and Applications*, 1(1):3–49.

Banerjee, A., Chandrasekhar, A. G., Duflo, E., and Jackson, M. O. (2013). The diffusion of microfinance. *Science*, 341(6144):1236498.

Banerjee, A. V. (1992). A simple model of herd behavior. *The quarterly journal of economics*, 107(3):797–817.

Barberá, P. (2015). Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data. *Political Analysis*, 23:76–91.

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., and Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science*, 26(10):1531–1542. PMID: 26297377.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, 100(5):992–1026.

Chandrasekhar, A. G., Larreguy, H., and Xandri, J. P. (2020). Testing models of social learning on networks: Evidence from two experiments. *Econometrica*, 88(1):1–32.

Coughlin, P. and Nitzan, S. (1981). Electoral outcomes with probabilistic voting and nash social welfare maxima. *Journal of Public Economics*, 15(1):113–121.

Coughlin, P. J. (1992). *Probabilistic voting theory*. Cambridge University Press.

DeGroot, M. H. (1974). Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121.

DeMarzo, P. M., Vayanos, D., and Zwiebel, J. (2003). Persuasion bias, social influence, and unidimensional opinions. *The Quarterly journal of economics*, 118(3):909–968.

Enelow, J. M. and Hinich, M. J. (1984). *The spatial theory of voting: An introduction*. CUP Archive.

Enelow, J. M. and Hinich, M. J. (1989). A general probabilistic spatial theory of elections. *Public Choice*, 61(2):101–113.

Erdős, P., Rényi, A., et al. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (1995). *Bayesian data analysis*. Chapman and Hall/CRC.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical science*, 7(4):457–472.

Glass, C. A. and Glass, D. H. (2021). Social influence of competing groups and leaders in opinion dynamics. *Computational Economics*, 58(3):799–823.

Golub, B. and Jackson, M. O. (2010). Naive learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, 2(1):112–49.

Graham, B. S. (2017). An econometric model of network formation with degree heterogeneity. *Econometrica*, 85(4):1033–1063.

Hoff, P. D. (2003). Random effects models for network data.

Hoff, P. D., Raftery, A. E., and Handcock, M. S. (2002). Latent space approaches to social network analysis. *Journal of the American Statistical Association*, 97(460):1090–1098.

Jackson, M. O. (2010). *Social and economic networks*. Princeton university press.

Jadbabaie, A., Molavi, P., Sandroni, A., and Tahbaz-Salehi, A. (2012). Non-bayesian social learning. *Games and Economic Behavior*, 76(1):210–225.

Krackhardt, D. (1987). Cognitive social structures. *Social networks*, 9(2):109–134.

Landsberger, H. A. (1958). Hawthorne revisited: Management and the worker, its critics, and developments in human relations in industry.

McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1):415–444.

Merriam-Webster (2021). Definition of "Influencer". Retrieved from: https://www.merriam-webster.com/dictionary/influencer.

Navelski, J. and Pascual, J. (2022). New methods in estimating the latent space network model with an application to the us media industry on twitter. *Working Paper*.

Penrose, M. (2003). *Random geometric graphs*, volume 5. OUP Oxford.

Pew Research Center (2021). Social Media Use in 2021. Retrieved from: https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/.

Prabhune, M. (2019). Top 20 Twitter Accounts to Follow for the Latest CRISPR News. Retrieved from: https://www.synthego.com/blog/10-crispr-accounts-twitter.

Sampson, S. F. (1968). *A novitiate in a period of change: An experimental and case study of social relationships*. Cornell University.

Smith, L. and Sørensen, P. (2000). Pathological outcomes of observational learning. *Econometrica*, 68(2):371–398.

Stan Development Team (2021). RStan: the R interface to Stan. R package version 2.21.3.

Synthego (2022). 60 Best CRISPR News Sources: Podcasts, Blogs, Journals, Twitter, and More. Retrieved from: https://www.synthego.com/learn/best-crispr-news-sources.

Twitter Inc. (2022a). Twitter Academic Research API. Accessed through Twitter API v2: https://developer.twitter.com/en/products/twitter-api/academic-research/.

Twitter Inc. (2022b). Twitter Documentation. Retrieved from: https://help.twitter.com/en/managing-your-account/about-twitter-verified-accounts.

6 Appendix

View	Username	Location	Created At	Name	Description	Verified	Follower Count	Following Count	Tweet Count	Listed Count	Followers Scraped	Diff
Anti	GMOEvidence	United Kingdom	9/9/12 11:10	GMO Evidence	GMO Evidence is a worldwide user-friendly library of evidence of harm caused by GMOs to animals and humans.	FALSE	6785	2516	1609	138	6000	785
Anti	USRightToKnow	United States	8/1/13 3:43	U.S. Right To Know	Investigative public health group working globally to expose corporate wrongdoing and government failures threatening our food, environment and health.	FALSE	10151	1910	6118	231	10000	151
Anti	JoelSalatin	Swoope, VA	7/12/11 17:05	Joel Salatin	OFFICIAL TWITTER: Healing the earth with good food. Author of 8 books. Featured in National Geographic, Smithsonian, Ominvore,Äôs Dillemma, Food Inc.	FALSE	29272	12	35	876	29000	272
Anti	nongmoreport	NA	3/25/09 22:28	Non-GMO Report	Monthly news magazine focusing on the risks of genetically modified foods and the non-GMO food trend.	FALSE	31293	842	21014	460	37000	5707
Anti	GMWatch	NA	8/17/09 16:46	GMWatch	Countering the propaganda of the biotech industry. Subscribe to free email newsletters: https://t.co/4GoxRVvA0L	FALSE	42109	2334	96161	1423	41997	112
Anti	RachelsNews	Made in Canada	9/14/11 16:58	Rachel Parent	Food Climate #Youth Speaker @SXSW Debated TV Host @CBC https://t.co/dnOzBSYKwB @TEDx https://t.co/3OkImzLnch Organic ,ï፟ü±@GenEarthEvent	FALSE	44885	29232	83899	825	43999	886
Anti	CFSTrueFood	DC | CA | OR | HI	1/6/09 20:56	Center 4 Food Safety	Protecting our food, farms & environment! Get the latest: https://t.co/3G9qZWDnR5	FALSE	44891	9267	16007	1264	44000	891
Anti	GMOFreeUSA	NA	7/16/12 2:14	GMO Free USA	We are a national group, educating consumers about the potential hazards of Genetically Engineered foods.	FALSE	52621	14613	41595	715	55998	3377
Anti	OrganicTrade	Washington, D.C.	4/8/09 15:22	Organic Trade	The Organic Trade Association (OTA) is the membership-based business association for organic agriculture and products in North America.	FALSE	59563	4214	20771	1160	58998	565
Anti	OrganicValley	La Farge, WI	7/17/08 14:29	Organic Valley	We,Äôre a cooperative of family farmers on a mission to produce the world,Äôs best organic foods for you and your family.	TRUE	80863	8094	16402	1394	79999	864
Anti	NonGMOProject	NA	6/11/09 23:20	Non-GMO Project	A nonprofit organization committed to preserving and building sources of #nonGMO products, educating consumers and providing verified non-GMO choices.	FALSE	125516	1100	7274	1676	127995	2479
Anti	OrganicConsumer	Finland, Minnesota	2/3/09 16:42	Organic Consumers Association	Want to take action on the issues you care about,Äîpesticides, factory farms and GMOs? Subscribe to our newsletter, #OrganicBytes üëâ https://t.co/cauFyafpsi	FALSE	187209	2339	27686	3298	186995	214
Pro	Recombinetics	Saint Paul, MN USA	1/23/11 15:37	Recombinetics	Biotech leader in precision breeding (Acceligen), biomed swine models for preclinical research (Surrogen) & oinkubators for regenerative medicine (Regenevida)	FALSE	1238	1080	1524	46	1000	238
Pro	AzMilkProducers	Arizona	11/12/09 17:23	AZ Milk Producers	Official Twitter account for AZ Milk Producers and Dairy Council of AZ. Encouraging everyone to get their 3 servings of delicious dairy everyday.	FALSE	2459	1175	6350	75	2000	459
Pro	CRISPRchef	NC State University	11/10/17 20:43	Rodolphe Barrangou	Distinguished professor of Food Science at @NCState. Leader of The @CRISPRlab. Editor-in-Chief of The @CRISPRjournal. #EverythingCRISPR	FALSE	2511	132	190	50	2000	511
Pro	AquaBountyFarms	Maynard, Massachusetts	12/26/12 16:24	AquaBounty	We provide fresh, nutritious Atlantic Salmon, genetically engineered once 30 years ago, and raised in safe, secure & sustainable land-based aquaculture farms.	FALSE	3341	638	1634	54	3000	341
Pro	joeBondyDenomy	San Francisco, CA	12/25/12 1:54	Joe Bondy-Denomy	BLACK LIVES MATTER. Associate Prof. @UCSF, Dept. of Micro & Imm. Toronto and Waterloo alum. Acrigen Biosciences co-founder. he/him/his	FALSE	3851	2120	2919	41	3000	851
Pro	FrancoiseBaylis	Halifax, Canada	11/7/12 14:57	Francoise Baylis, PhD	Author "Altered Inheritance", Philosopher, University Research Professor, Dalhousie University	FALSE	3919	656	12963	112	3000	919
Pro	jcornlab	Zurich, Switzerland	2/23/14 21:47	Jacob Corn	Genome editing, functional genomics, and cells figuring out how to eat themselves without dying. Professor of Genome Biology at ETH Zürich.	FALSE	4248	438	2505	78	4000	248
Pro	AprilPawluk	Toronto, Ontario	5/20/10 1:36	Dr. April Pawluk	Strategic Program Manager at Harvard Chan Microbiome in Public Health Center. üá®üá¶ PhD. Former Cell Press editor. Perpetually fascinated by microbes. (she/her)	FALSE	4262	1799	7654	62	4000	262
Pro	jsherkow	Champaign, IL	8/17/12 0:12	Jacob S. Sherkow üë¶Üüíäüè∏,ö´	Prof @UIllinoisLaw & @IGBIllinois; also @CeBIL_Center #patents, #IP, #biotech, #pharma, #bioethics. Ducks, drugs, and nicks that roll.	FALSE	4840	299	16887	173	4000	840
Pro	shsternberg	New York, NY	4/23/15 4:04	Sam Sternberg	Assistant professor at Columbia University / CRISPR aficionado / Protein,ÄìRNA biochemist / Co-author with Jennifer Doudna of "A Crack In Creation"	FALSE	4946	495	1788	64	4000	946
Pro	JKamens	Boston, MA	5/25/11 14:05	Dr. Joanne Kamens	Mom, Grandmother, Interim ED @BentleyCWB Proud founder @MassAWIS. Career, workplace and #Diversity rabble rouser, especially in STEM. Mentoring expert. She/Her	TRUE	5826	4808	19086	175	5000	826
Pro	mem_somerville	NA	6/22/08 21:36	Dr. mem_somerville | Wossamotta U	The kind of independent scientist people claim they want to hear from, until they dislike the conclusions they hear. Spectacularly combustible. "a real person?"	FALSE	6279	4485	83965	259	5999	280
Pro	Synthego	California, USA	8/16/12 1:17	Synthego	The Genome Engineering Platform Company. üß¨ #CRISPR #GenomeEngineer	TRUE	6327	2178	3876	145	6000	327
Pro	pcronald	In the lab or garden in Davis	5/20/08 23:15	@pcronald	UC Davis #rice geneticist & author of Tomorrow's Table: Organic #Farming Genetics & the Future of #Food. TED talk. My lab studies #immunity & #stress. #scicomm	FALSE	7096	2024	9108	321	6999	97
Pro	JonEntine	Cincinnati	3/25/09 15:44	Jon Entine	Dad, journalist, iconoclast, founder https://t.co/VaCWq74soa - ag biotech/biomedicine/sustainability/ethics. Emmy-winning TV news producer, book author	FALSE	7569	5961	23207	190	6994	575
Pro	ELS_Genetics	Amsterdam, Boston, Oxford	10/11/11 18:59	Genetics & Genomics	Official page for Elsevier Genetics & Genomics books & journals. General genetics & genomics news, call for papers, special issues, content innovations & more!	FALSE	8166	7976	3939	140	8000	166
Pro	KevinADavies	Washington DC	4/17/09 22:22	Kevin Davies	Author "EDITING HUMANITY: The CRISPR Revolution and the New Era of Genome Editing"; Exec Ed @CRISPRJournal, @GENbiotechjrnl; @hwaspeakers	FALSE	8419	1675	10022	312	8000	419
Pro	BioBeef	Davis, California	9/30/10 19:24	Prof Alison Van Eenennaam (@BioBeef)	Animal üêÑ biotechnology/genetics, Uni of CA Davis, Mom, Consumer, Ag Science, üá¶üáƒüáƒüá∏ Melbourne Uni, Food Evolution https://t.co/HGmEgkDYwX	FALSE	10059	1124	17188	236	10000	59
Pro	igisci	Berkeley, CA	2/24/16 20:33	Innovative Genomics Institute	The Innovative Genomics Institute (IGI) is dedicated to advancing genome research for a better world. Partnership between @UCBerkeley and @UCSF	FALSE	10655	750	3374	164	10000	655
Pro	CamiDRyan	St Louis County, Missouri	3/6/09 23:58	Cami Ryan, PhD	Information ,ÑñÔ∏è wrangler üåæ | Social & behavioral üß† sciences üë• @Bayer4Crops | @GenomeCanada | @agbiousask | artist üé® | more interesting on IG and TikTok	FALSE	10797	5012	57347	307	10000	797
Pro	nmpf	Arlington, VA, United States	7/31/09 19:02	National Milk Producers Federation	Representing farmer-owned #dairy cooperatives on Capitol Hill since 1916	FALSE	11824	1471	15678	251	12000	176
Pro	pdhsu	hsu.berkeley.edu	3/17/09 3:05	Patrick Hsu	Assistant Professor @BerkeleyBioE, Cofounder @ArcInstitute | CRISPR & biotech to improve human health | üá®üá¶ to @harvard @broadinstitute @editasmed, Fast Grants	FALSE	12272	1554	1538	148	12000	272
Pro	NPPC	Washington, DC	1/13/09 16:53	NPPC	The National Pork Producers Council is the global voice of the U.S. pork industry. #TeamPork #PorkPriorities	FALSE	12422	2409	7423	272	12000	422
Pro	GaetanBurgio	Canberra, Australia	8/14/14 10:50	Dr Gaetan Burgio, MD, PhD.	Group leader, Australian National University ANU, Australia. Geneticist working on infections & #CRISPR. Opinions and views	FALSE	12430	1589	22856	415	12000	430
Pro	CRISPRjournal	New Rochelle, NY	5/18/17 20:15	The CRISPR Journal	Cutting-edge peer-reviewed research and analysis on genome editing for the global CRISPR community. EIC @crisprchef; Exc Ed @kevinadavies; Publisher @liebertpub	FALSE	18303	912	4967	230	18000	303
Pro	GeneticLiteracy	NA	1/20/12 17:34	Genetic Literacy Project	The Genetic Literacy Project fosters dialogue about the scientific, social and ethical implications of human and agricultural genetics.	FALSE	18880	7412	36732	495	17999	881
Pro	pknoepfler	Davis, CA	2/16/10 20:49	Paul Knoepfler	Prof @UCDAVIS #Stemcells & Kids Brain Cancer Epigenomics #CRISPR TED talk https://t.co/QmSyU0534S Reedie. Author of The Niche https://t.co/oXW4Pis1Cl YouTube channel https://t.co/s96VUCXf6G	TRUE	21595	4750	26285	537	20999	596
Pro	AgBioWorld	Alabama, USA	4/14/11 14:09	Channa Prakash	Professor, biotech guru. I am curious about science, farming, food security, innovation, plants, history. Come for the dog/cat videos, but stay for science.	FALSE	25412	2440	54695	514	24999	413
Pro	SynBioBeta	San Francisco, CA.	8/29/12 2:06	SynBioBeta	The leading community of innovators in #synbio. | Our Twitter Spaces show #BuiltWithBioSpaces goes LIVE every Weds 8am PDT: https://t.co/wo5hj5uBAk	FALSE	31192	4320	27606	666	30998	194
Pro	CRISPR_News	San Francisco, CA	6/8/15 5:42	CRISPR News	The latest CRISPR breakthroughs. Also at https://t.co/oSD5xoDl5y #biotechnology #cas9 #genetics #raredisease #genomics #germline	FALSE	32463	539	1479	391	32000	463
Pro	doudna_lab	Berkeley, CA	12/3/15 21:01	Doudna Lab	News from Jennifer Doudna's lab at @UCBerkeley. Tweets from lab members and not Jennifer Doudna unless signed JD. Tweets represent personal views only.	FALSE	42945	55	304	430	41999	946
Pro	Kevin_Faulconer	San Diego	6/18/09 23:35	Kevin Faulconer	Candidate for California Governor. #CaliforniaComeback Husband, father, former Mayor of San Diego, and visiting professor at @PepperdineSPP	TRUE	45607	1194	6666	614	44998	609
Pro	Cargill	Minneapolis, Minnesota (HQ)	4/28/09 20:48	Cargill	Official Twitter account for Cargill Inc.	TRUE	57468	258	7497	948	58000	532
Pro	TysonFoods	Global - HQ in Arkansas	8/13/08 12:42	Tyson Foods	Tyson Foods, Inc. founded in 1935.	TRUE	57952	16878	12130	913	57999	47

Figure 12: Expert Accounts Demographics (Raw Data)

6.1 Estimation Methodology and Prior Distributions for Parameters

I use `RStan`'s No U-Turn Sampling algorithm, developed by Gelman et al. (1995), and simulate two chains with 1,000 draws and a burn-in of 500 samples (Stan Development Team, 2021). The assumed prior distributions for the population are

$$\begin{aligned} \mu &\sim \mathcal{N}(\mu_\mu, \sigma_\mu) & \gamma &\sim \mathcal{N}(\mu_\gamma, \sigma_\gamma) \\ \alpha_j &\sim \mathcal{N}(\mu_\alpha, \sigma_\alpha) & \beta_i &\sim \mathcal{N}(\mu_\beta, \sigma_\beta) \\ \theta_i &\sim \mathcal{N}(\mu_\theta, \sigma_\theta) & \phi_j &\sim \mathcal{N}(\mu_\phi, \sigma_\phi), \end{aligned}$$

and the full joint posterior distribution is thus defined as:

$$\begin{aligned} p(\mu, \boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma, \boldsymbol{\theta}, \boldsymbol{\phi} | \mathbf{y}) &\propto p(\mu, \boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma, \boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{\mu}, \boldsymbol{\sigma}) \\ &\propto \prod_{i=1}^{n}\prod_{j=1}^{m} \text{logit}^{-1}(\pi_{ij})^{y_{ij}}(1 - \text{logit}^{-1}(\pi_{ij}))^{1-y_{ij}} \\ &\quad \prod_{j=1}^{m\times n} \mathcal{N}(\mu|\mu_\mu, \sigma_\mu) \prod_{i=1}^{m\times n} \mathcal{N}(\gamma|\mu_\gamma, \sigma_\gamma) \\ &\quad \prod_{j=1}^{m} \mathcal{N}(\alpha_j|\mu_\alpha, \sigma_\alpha) \prod_{i=1}^{n} \mathcal{N}(\beta_i|\mu_\beta, \sigma_\beta) \\ &\quad \prod_{i=1}^{n} \mathcal{N}(\theta_i|\mu_\theta, \sigma_\theta) \prod_{j=1}^{m} \mathcal{N}(\phi_j|\mu_\phi, \sigma_\phi) \end{aligned} \tag{4}$$

where $\pi_{ij} = \mu + \alpha_j + \beta_i - \gamma|\theta_i - \phi_j|$, and the latent prior parameters are $\boldsymbol{\mu} = (\mu_\theta, \mu_\phi)'$ and $\boldsymbol{\sigma} = (\sigma_\theta, \sigma_\phi)'$. While Navelski and Pascual (2022) develop a new set of Jeffery's priors, I use the prior specification from Barberá et al. (2015), which is supported by Hoff (2003).

6.2 Identification Strategy

The model in Equation (2) is still unidentified due to "additive aliasing" and "scaling invariance" since there are an infinite number of combinations between the parameters that will give the same probability of following. An example of additive aliasing is $\mu = 0, \alpha_j = -1, \beta_i = 1, \phi_j = 1, \theta_i = -1$, which gives the same probability of $\mu = 0, \alpha_j = 1, \beta_i = -1, \phi_j = -1, \theta_i = 1$, and an example of scaling invariance is multiplying the distance $-\gamma|\theta_i - \phi_j|$ by any constant k where γ will absorb part of the constant $-\frac{\gamma}{k}|(\theta_i - \phi_j)k|$ (Barberá (2015)).[14] These problems are usually solved by restricting one of the j^{th} or i^{th} parameters in each parameter set, but becomes difficult to do when working with the distance between two latent parameters. Barberá (2015) deals with this problem by making two of the priors for ϕ_j extremely informative, with very low variances , and close to -1 for a "liberal" leaning expert j and close $+1$ for "conservative" leaning expert j. This methodology essentially

[14] Barberá (2015) does a great job at explain this in the Supplementary Materials if a deeper explanation is of interest.

"anchors" these two estimates while all other estimates are jointly estimated from this specification. Navelski and Pascual (2022) suggest a different identification strategy where all priors are treated equally for ϕ_j and θ_i, and are transformed using an invariant transformation.

6.2.1 Latent Positions (ϕ_j)

I employ a modification of a standard practice in the field of statistics called Fisher's inverse arc tangent transformation, which was developed and applied by Navelski and Pascual (2022) in this setting. The general formulation I use is:

$$x_1' = \arctan\left(e^{(x_1+x_{center})\frac{2}{\pi}} - x_{cut}\right) \qquad x_2' = \arctan\left(x_{cut} - e^{(x_2+x_{center})\frac{2}{\pi}}\right) \tag{5}$$

where, $x_1 \in \mathbb{R}$ and $x_2 \in \mathbb{R}$ are mapped to $x_1' \in (x_{cut}, 1)$ and $x_2' \in (-1, x_{cut})$, and

- x_1 and x_2 are the initially estimated values
- x_1' and x_2' are the transformed values
- x_{cut} is hyperparameter and a constraint put on the lower bound for each side of the estimation
- $x_{center} = \ln(\tan\left(\frac{\pi}{2} * 0\right) + x_{cut})$ and is a value that centers the transformation.

I apply this transformation to the expert latent parameters to constrain the parameter estimates and to map estimates to an intuitive scale. More specifically, I map all samples drawn from the priors to a constrained parameter space where $\phi_j \in (-1, 1)$ for all $j \in \{1, \ldots, m\}$. This transformation is intuitive because is allows researchers to analyze relative latent positions on a -1 (Anti) to 1 (Pro) scale. In Equation 5 x_{cut} is a hyperparameter that allows latent parameters to "switch" to the other side of the spectrum if that is the true location of the latent parameter. For example, in the GEDL application the @AzMilkProducers expert was originally classified as Pro GEDL and were given an initial value of .8 to initiate the MC-MC estimation, and the mean of their latent posterior distribution converged to -0.089, which is more Anti leaning than Pro. This intuitive transformation

An additional difficulty when using a distance model is "reflection invariance" where the resulting scale between θ_i and ϕ_j could lead to estimates that are misinterpreted since their signs could be flipped. To combat "reflection invariance," I assume the aforementioned transformation in Equation 5 for $\boldsymbol{\phi}$, and use $\phi_{cut} = .6$ to guide the parameter estimates to their anticipated estimates. To aid this transformation, I use a technique similar to Barberá (2015) and Navelski and Pascual (2022), where I assume -0.8 as the starting value for ϕ_j for the Anti GEDL experts and $+0.8$ for the Pro GEDL experts. These assumptions are not strong as Bayesian theory allows posterior draws to converge to their theoretically correct distribution, and if estimates were to diverge from their anticipated underlying distribution, I would see estimates trying to converge to the posterior distribution on the other side of the spectrum, unsatisfactory conversion diagnostics, and high prediction errors. In this analysis, which is not implemented in Barberá (2015) and Navelski and Pascual (2022), I use the standard practice of constraining the m random effects to sum up to zero (i.e., $\sum_{j=1}^{m} \alpha_j = 0$) and the same for the n random effects (i.e., $\sum_{i=1}^{n} \beta = 0$). This allows all estimates to be identified relative to their expectation, which is zero.

6.3 Summary of Estimation Results

Figures 13 (a) and (b) show the distribution of the posterior means of expert popularity and follower engagement estimates, respectively, and Figures 14 (a) and (b) show the distribution of the posterior means for the expert and follower ideology estimates. Expert popularity is centered at zero (i.e., $\mathrm{E}\left(\widehat{\alpha}_j\right) = 0$), and even tough the distribution seems to be symmetric overall, it is clear that the more popular experts, experts with values greater than zero, are more "intensely" popular than those on the negative side. This indicates that a popular expert has more of an effect on a follower's decision to follow than and unpopular expert since the change in the probability of following has a greater increase for a popular expert than a decrease in an unpopular expert. This result is motivated in Table 6 where @NonGMOProject and @OrganicConsumer both have popularity estimates of 3.78 and 3.69, respectively, and @Kevin_Faulconer and @AzMilkProducers have popularity estimates of -3.30 and -3.07, respectively.

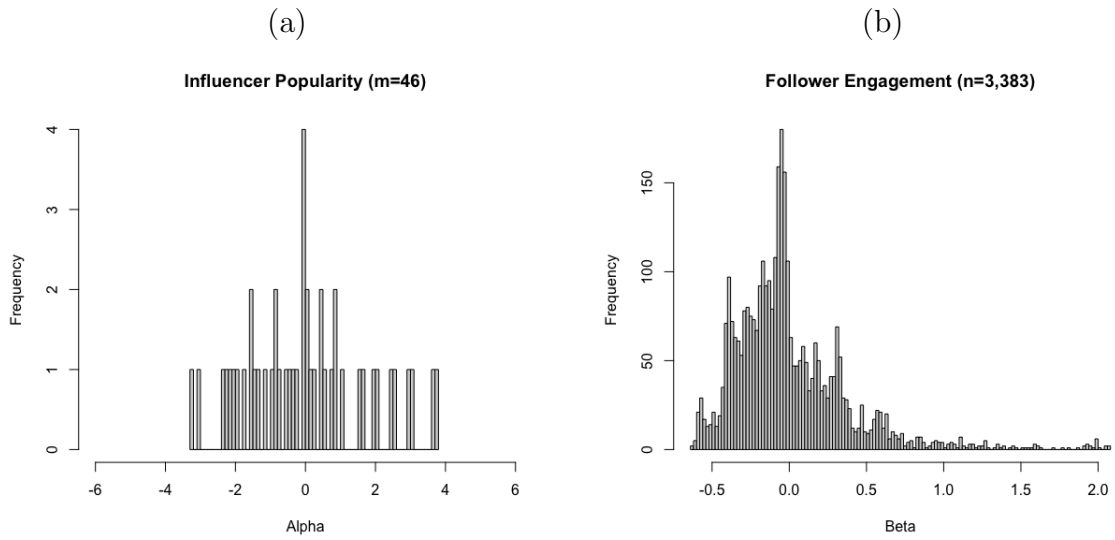

Figure 13: Popularity for Experts $\widehat{\alpha}_j$ (a) and Engagement of Followers $\widehat{\beta}_i$ (b).

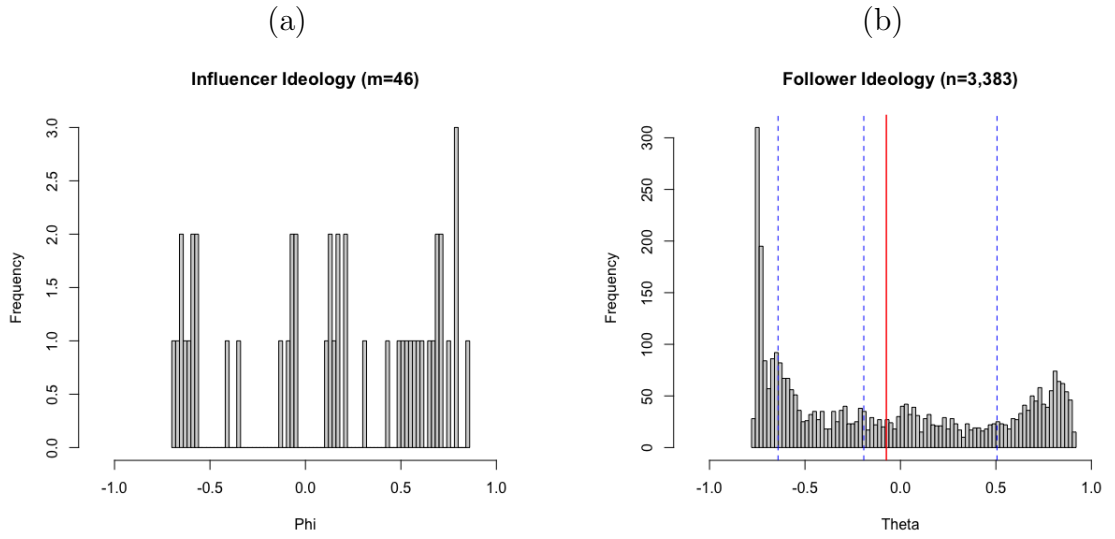

Figure 14: Ideology for Experts $\widehat{\phi}_j$ (a) and Followers $\widehat{\theta}_i$ (b).

Table 6: Examples of Popularity and Ideology Extremes

Parameter (Name - Initializing View)	$\widehat{R}$	Mean	SD	2.5%	97.5%
Most Popular					
$\widehat{\alpha}_2$ (@NonGMOProject - Anti)	1.00	3.78	0.16	3.49	4.11
$\widehat{\alpha}_1$ (@OrganicConsumer - Anti)	1.00	3.69	0.16	3.38	4.02
Least Popular					
$\widehat{\alpha}_{15}$ (@Kevin_Faulconer - Pro)	1.00	-3.30	0.12	-3.51	-3.06
$\widehat{\alpha}_{45}$ (@AzMilkProducers - Pro)	1.00	-3.07	0.12	-3.29	-2.84
Most Extreme Anti Genome Editing Ideology					
$\widehat{\phi}_{12}$ (@GMOEvidence - Anti)	1.03	-0.69	0.02	-0.72	-0.66
$\widehat{\phi}_4$ (@nongmoreport - Anti)	1.01	-0.68	0.02	-0.71	-0.65
$\widehat{\phi}_1$ (@OrganicConsumer - Anti)	1.01	-0.65	0.02	-0.68	-0.62
$\widehat{\phi}_5$ (@GMOFreeUSA - Anti)	1.03	-0.65	0.01	-0.68	-0.62
Most Extreme Pro Genome Editing Ideology					
$\widehat{\phi}_{42}$ (@joeBondyDenomy - Pro)	1.03	0.85	0.03	0.79	0.91
$\widehat{\phi}_{37}$ (@shsternberg - Pro)	1.00	0.80	0.02	0.77	0.84
$\widehat{\phi}_{25}$ (@pdhsu - Pro)	1.01	0.78	0.02	0.75	0.81
$\widehat{\phi}_{39}$ (@AprilPawluk - Pro)	1.01	0.78	0.02	0.74	0.83
Moderate Genome Editing Ideology					
$\widehat{\phi}_{24}$ (@NPPC - Pro)	1.01	-0.04	0.02	-0.08	-0.00
$\widehat{\phi}_{14}$ (@Cargill - Pro)	1.01	-0.04	0.02	-0.07	-0.01
$\widehat{\phi}_{26}$ (@nmpf - Pro)	1.01	-0.07	0.02	-0.11	-0.04
$\widehat{\phi}_{15}$ (@Kevin_Faulconer - Pro)	1.00	-0.08	0.04	-0.15	0.00

Ideologies of the experts and followers exhibit opposite distributional patterns, where the negative side of distributional values in Figure 14 (a) and (b) represent those that are Anti GEDL, and the positive side are those that are Pro. For clarity, an ideology value of -1 indicates the most extreme Anti GEDL ideology, while a value of 1 indicates the most extreme Pro GEDL ideology. Both the ideology of the experts and followers tend to be polarized since a large majority of the estimates are concentrated at the end of the spectrum $(-1, 1)$. The experts with the most extreme ideologies are presented in Table 6 where @GMOEvidence, @nongmoreport, @OrganicConsumer and @GMOFreeUSA all have the lowest ideal points at -0.69, -0.68, -0.65, and -0.65, respectively, and @joeBondyDenomy, @shsternberg and @pdhsu, and @AprilPawluk have the highest ideal points at 0.85, 0.80, 0.78 and 0.78, respectively. The polarization between experts ideology is interesting because the Anti GEDL accounts are very extreme while some Pro accounts are polarized and some are seen to be more moderate. For example, @NPPC, @Cargill, @nmpf and @Kevin_Faulconer were all initially considered experts on the Pro GEDL side, but in reality, they

have ideologies that are more moderate at −0.04, −0.04, −0.07 and −0.08, respectively. These point estimates are presented in Table 6, and this is an interesting result because these experts initially started as Pro GEDL experts while their uncovered ideological estimates are moderate to Anti-moderate GEDL. This type of result could be indicative of these accounts having followers that have more moderate ideologies, or that have a spread of Anti and Pro ideologies.

6.4 Estimation Diagnostics

Gelman and Rubin (1992) recommend a $\widehat{R}$ statistic at 1.1, implying that there are no divergent transitions in the estimation process, and this is the benchmark most researchers follow in practice. To support this intuition, Figure 15 plots all $\widehat{R}$ values, showing that all values are below the 1.1 line. The optimal classification threshold was derived by maximizing the area under the ROC curve (i.e., maximizing the sensitivity and specificity of the prediction diagnostics), and Figure 16 (a) and (b) show the ROC curve and confusion matrix, respectively.

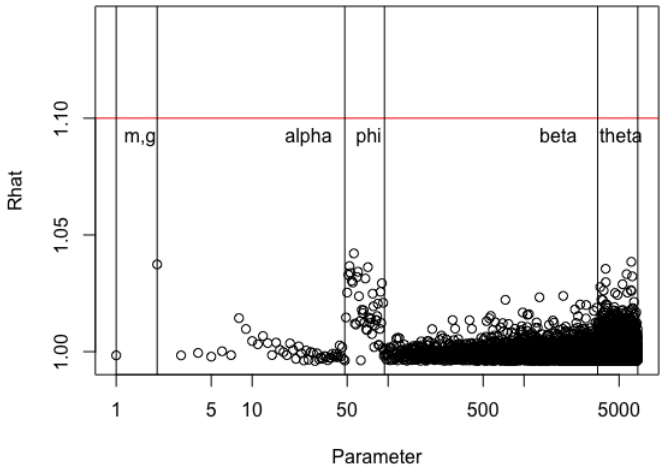

Figure 15: $\hat{R}$ Plot of All 6,860 Parameters (MCMC Convergence Diagnostics)

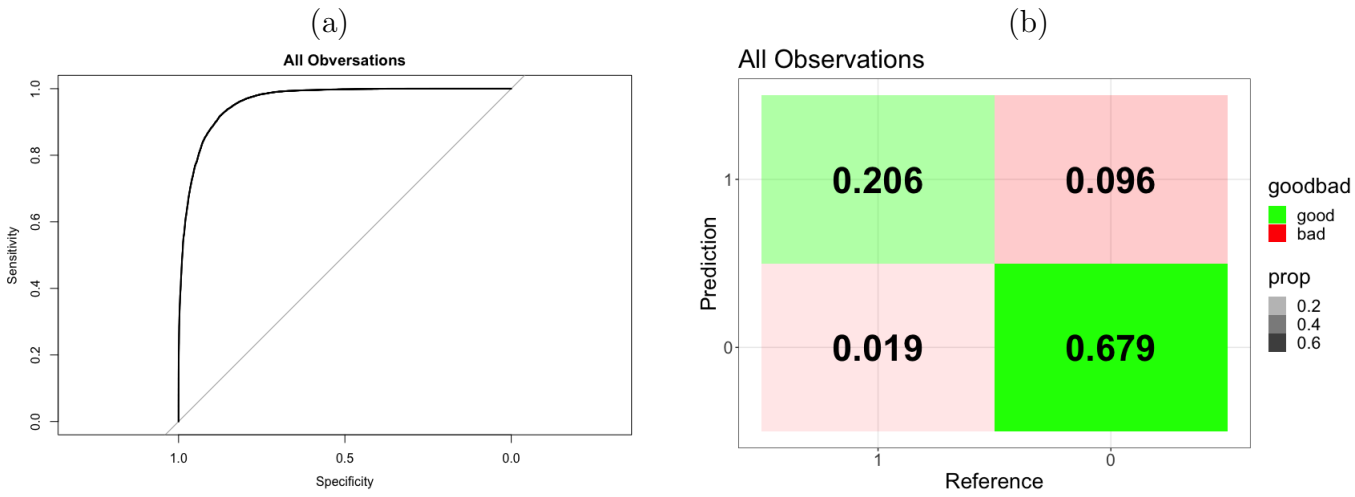

Figure 16: ROC Curve (a) and Confusion Matrix (b) for All Observations

6.5 $\mathbf{T}^*$ Matrix for All Examples

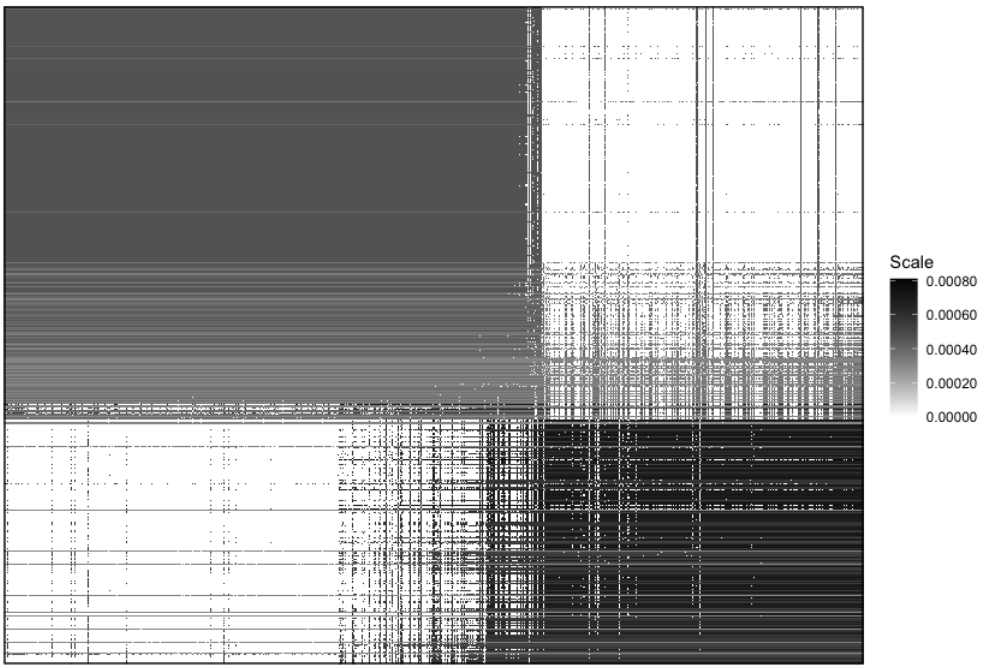

Figure 17: Example 1: Heat Map of $\mathbf{T}^*$ Matrix (Common Interests)

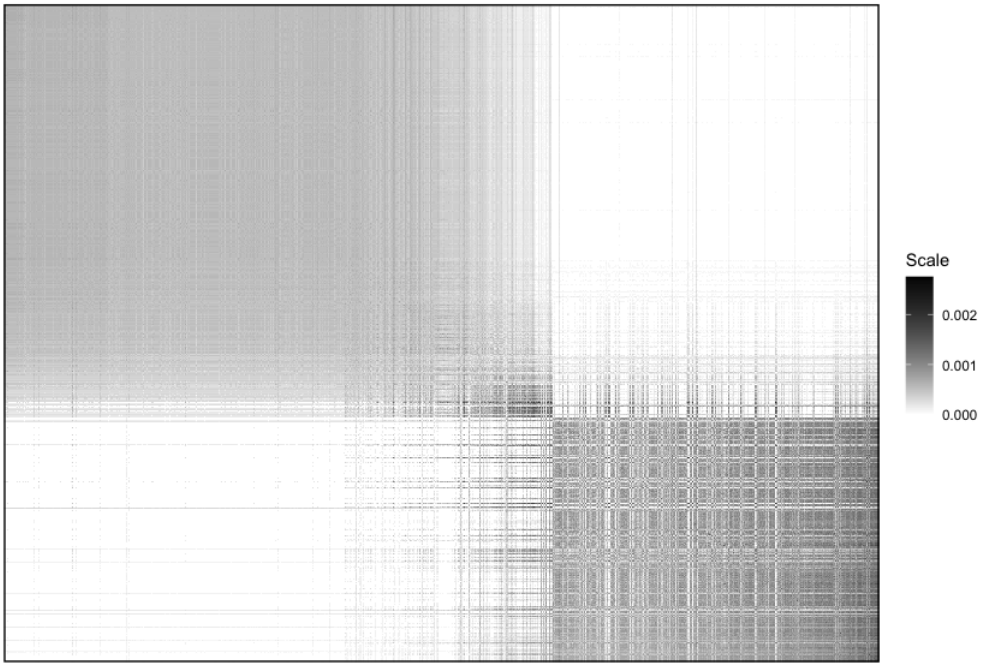

Figure 18: Example 2: Heat Map of $\mathbf{T}^*$ Matrix (Weighted Common Interests)

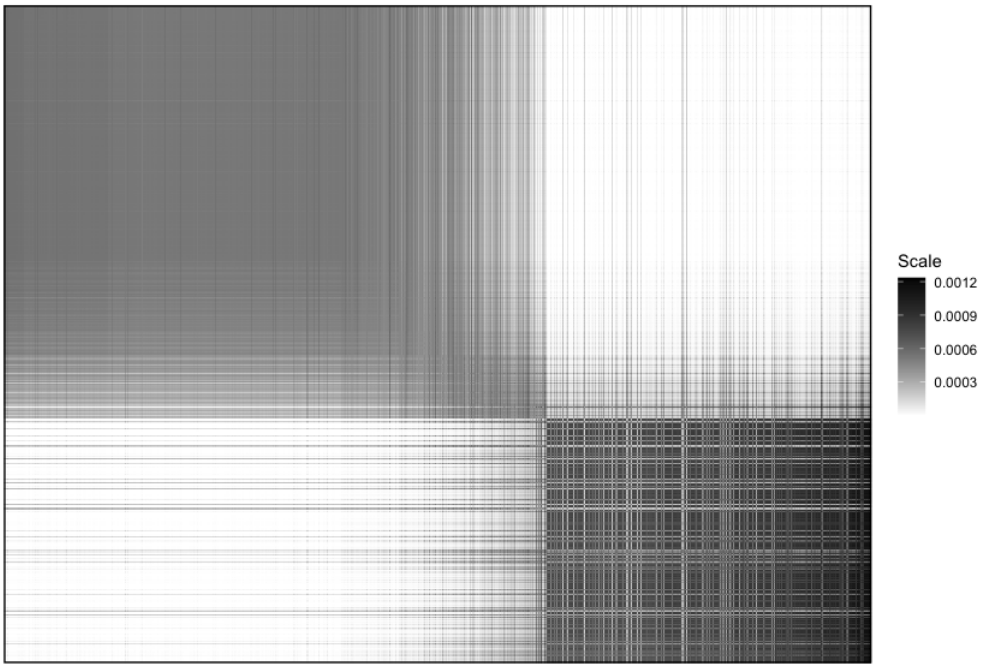

Figure 19: Example 3: Heat Map of $\mathbf{T}^*$ Matrix (Est. Weighted Common Interests)