Article

Opinion Models, Election Data, and Political Theory

Matthias Gsänger [1] , **Volker Hösel** [2] , **Christoph Mohamad-Klotzbach** [1] and **Johannes Müller** [2,3,*]

1 Institute of Political Science and Sociology, Julius-Maximilians-University (JMU), 97074 Würzburg, Germany; matthias.gsaenger@uni-wuerzburg.de (M.G.); christoph.mohamad-klotzbach@uni-wuerzburg.de (C.M.-K.)
2 School for Computation, Information and Technology, TU München (TUM), 80333 Munich, Germany; volker.hoesel@tum.de
3 Institute for Computational Biology, Helmholtz Center Munich, 85764 Munich, Germany
* Correspondence: johannes.mueller@mytum.de

Abstract: A unifying setup for opinion models originating in statistical physics and stochastic opinion dynamics are developed and used to analyze election data. The results are interpreted in the light of political theory. We investigate the connection between Potts (Curie–Weiss) models and stochastic opinion models in the view of the Boltzmann distribution and stochastic Glauber dynamics. We particularly find that the q-voter model can be considered as a natural extension of the Zealot model, which is adapted by Lagrangian parameters. We also discuss weak and strong effects (also called extensive and nonextensive) continuum limits for the models. The results are used to compare the Curie–Weiss model, two q-voter models (weak and strong effects), and a reinforcement model (weak effects) in explaining electoral outcomes in four western democracies (United States, Great Britain, France, and Germany). We find that particularly the weak effects models are able to fit the data (Kolmogorov–Smirnov test) where the weak effects reinforcement model performs best (AIC). Additionally, we show how the institutional structure shapes the process of opinion formation. By focusing on the dynamics of opinion formation preceding the act of voting, the models discussed in this paper give insights both into the empirical explanation of elections as such, as well as important aspects of the theory of democracy. Therefore, this paper shows the usefulness of an interdisciplinary approach in studying real world political outcomes by using mathematical models.

Keywords: opinion dynamics; Potts models; Glauber dynamics; q-voter model; reinforcement model; weak and strong effects continuum limit; data analysis and model comparison; elections; voting behavior; interdisciplinarity

1. Introduction

The search for common compromises in a discursive social process is at the heart of all democracies. In these debates, citizens seek their standpoint on the basis of information and news, but also based on discussions with family, colleagues, and acquaintances. Opinion dynamics aims to model the basic structure of precisely this process [1–3].

Many opinion dynamics models are constructed on communications graphs, where the nodes represent persons, who only interact with neighboring persons [4]. Basically, there are two groups of models: either individuals are equipped with one of a finite number of opinions (typically pro and contra) or their opinions are characterized by a continuous spectrum of possibilities (typically the interval $[0, 1]$, see [5]). In particular, the models with a continuous state space are often used to investigate whether a consensus can be reached in the long run [6–8].

Most papers do not aim to validate their models in analyzing empirical data in a quantitative way [9]. Those papers that do address empirical data mainly focus on the dynamics of two opinions within homogeneous groups and do not use an underlying graph structure [10–17], as many interesting aspects, like different interaction patterns or phase transitions, already appear in models describing homogeneous populations and do

not require interaction graphs. Here, we find a parallel in mathematical epidemiology, where it is clear that infections spread via contact graphs, but most quantitative models and methods aiming at the description and prediction of the dynamics of a real world outbreak are based on models that assume homogeneous mixing [18,19]. Also in the present work, we find that these simple models are sufficiently rich to meet the structure of empirical data. Technically, the main objective of those more empirically oriented papers is to obtain the stationary distribution of the underlying stochastic process and then to use this result in order to fit model parameters to data. In that, these papers are able discuss possible mechanisms that generate striking patterns in data [11–15], aim to reveal changes in communication patterns in the course of time [16], or address spatial communication distances [17]. Not only in political processes, but also in other fields such as vaccination hesitancy, opinion models contribute to an adequate description of the underlying communication mechanisms and, in that, potentially open up ways to handle (in this case) public health problems [9,20,21]. It is, however, noticeable that all these aforementioned approaches from socio-physics and socio-mathematics up to now only have a small or no echo in the social and political sciences.

In the present paper, we aim to achieve three goals: in the first part, we connect two different approaches to opinion models: one driven by stochastic dynamics and the other one originating in statistical physics; in the second part, we target a model comparison to find out which models are able to explain empirical data. The third part then discusses the usefulness of mathematical models of this type for political sciences.

For the first part of the paper: if we review the literature, we find two main approaches to describe opinion models—one approach formulates mechanisms about how people change their minds when interacting with others in the form of stochastic dynamics. In the simplest case, the voter model [22], it is assumed that a person faithfully copies the opinion of some randomly chosen other person. A slightly more refined version of this idea introduces zealots (also called stubborns or activists) who never change their mind [23,24], which results in the zealot (or noisy voter) model. In the long run, we find an equilibrium in opinion dynamics, which is termed stationary distribution. This distribution can be used to analyze empirical data.

A completely different approach originates in the statistical physics of spin systems. Herein, the dynamics of the opinions are not considered, but it is assumed that only a little is known about the state of the population, e.g., surveys could inform us about the abundance of some opinions. To express this partial knowledge appropriately, a distribution is constructed that maximizes the Shannon entropy under the constraint of the known information. The models constructed in this way are called Potts models, and the distribution is called the Boltzmann distribution [25,26]. As the stationary distribution of the dynamical opinion models, the Boltzmann distribution is also used to analyze data [26].

We connect both approaches in identifying a dynamical stochastic model (first approach), which generates a stationary distribution that coincides with a Boltzmann distribution (second approach). In this way, we associate a dynamical model with a statistical physics Potts model and vice versa. The advantage of this procedure is the construction of a unifying framework for both approaches. Based on this framework, the q-voter model [27–29] appears as a natural extension of the zealot model, not in view of the mechanisms modeled by the q-voter model, but in view of the underlying mathematical structure.

We aim to apply the models to data. Election data usually aggregate information about a large number of people, as constituencies usually comprise thousands to hundreds of thousands of voters. Therefore, a continuum limit is of interest. Here, we note that the zealot model is identical to the Moran model with mutation [30], which forms the basis of population genetics [31]. In population genetics, two different diffusion limits have been established—the weak and the strong effects limit. These two approaches differ in the assumptions regarding the scaling of the parameters with respect to the population size. We investigate ways to transfer this idea to general opinion models.

In the second part, we analyze election data from the United States (US), United Kingdom (UK), France (FRA), and Germany (GER) based on four opinion dynamics models (the Curie–Weiss model, the weak and strong effects continuum limit of the q-voter model, and, additionally, the weak effects limit of the reinforcement model [21]). We use this analysis to test the models to find out to what extent they are able to describe the data not only qualitatively, but also quantitatively. The central finding is that models derived by a weak effects limit perform much better than their strong version. Potentially, these findings will be useful for future empirical studies. In the last part, we again change the focus and turn to the interpretation of our results in the light of political science. In particular, political processes and changes in social interactions potentially leave their traces in the election data, and therewith in the estimated parameters. However, this conjecture can only be confirmed based on in-depth political research. These considerations lead us to another important aspect, which is the question of the extent to which socio-mathematical models of this kind can also be a fruitful instrument as an integral part of political science, or whether the methods, objectives, and research questions of socio-mathematics and socio-physics on the one hand and political science on the other are too different.

2. General Structure

In this section, we first introduce the notation before introducing the two different types of models that we will investigate.

We consider a population of N individuals numbered $1, \ldots, N$, where each of the individuals supports either opinion A or opinion B. The opinions are coded by 1 (for A) and -1 (for B). The state space is given by $\Sigma_2 = \{\pm 1\}^N$, such that the i'th component $\sigma_i \in \{\pm 1\}$ of state $\sigma \in \Sigma_2$ indicates the opinion of individual i. We consider Cannings models [32], that is, all individuals are exchangeable and the population is homogeneous. Particularly, we do not have an interaction graph; respectively, the interaction graph is the full graph. We introduce the functions

$$n_+(\sigma) = \sum_{i=1}^{N} \mathbb{1}(\sigma_i = 1), \qquad n_-(\sigma) = N - n_+(\sigma),$$

which count the supporters of opinion A (function $n_+(\sigma)$) and the supporters of opinion B (function $n_-(\sigma)$), respectively. Our knowledge on the opinion distribution in the population will be expressed by random measures on Σ_2. Due to the assumption of Cannings models, the random measures, necessarily, are invariant with respect to the permutations of individuals: two states $\sigma^1, \sigma^2 \in \Sigma_2$ with $n_+(\sigma^1) = n_+(\sigma^2)$ have the same probability.

That is, any random measure $Q : \Sigma_2 \to [0, 1]$ describing the state of the population induces a random measure P on the state space $V_N := \{0, .., N\}$, which indicates the number of opinion-A supporters. Let $\sigma^{(k)} \in \Sigma_2$ be a given state with $n_+(\sigma^{(k)}) = k$; then, for combinatorial reasons:

$$P(k) = \binom{N}{k} Q(\sigma^{(k)}). \tag{1}$$

We will find out later that the binomial coefficient, which appears here for symmetry reasons, plays a distinct role in the theory developed below.

In the next two sections, we introduce two very different ways used in the literature to construct random measures for the opinion state of the population, that is, on Σ_2, respectively, V_N. The first approach is based on stochastic processes, while the second is statistical in design. We should keep in mind that—due to the symmetry discussed—every rate and every function used to define the models can be constructed in such a way that it depends on $\sigma \in \Sigma_2$ only via $n_\pm(\sigma)$ and the population size N.

2.1. Opinion Process

A stochastic opinion process is a Σ_2-valued Markov process $\hat{\sigma}_t$, where single persons reconsider their opinion at rate ν. There is a certain probability that this person indeed changes her mind. These probabilities depend on the opinion distribution in the population; as mentioned above, this dependency is established via $n_\pm(\sigma)$ and not via the fine-structure of the state (which individual is an A- and which individual is a B-supporter). For mathematical convenience, but without loss of generality, we assume that the rate for $\sigma_i i$ to switch from -1 to 1 is a function of $n_+(\sigma)$ and N, while the rate to switch from 1 to -1 depends on $n_-(\sigma)$ and N,

$$\sigma_i = -1 \rightarrow \ 1 \quad \text{at rate} \quad \nu f^+(n_+(\sigma); N) \tag{2a}$$
$$\sigma_i = \ 1 \rightarrow -1 \quad \text{at rate} \quad \nu f^-(n_-(\sigma); N). \tag{2b}$$

To obtain a feeling for which terms to use for $f^\pm$, we can look at the voter model. Herein, ν represents the rate at which a person reconsiders her opinion and interacts with some randomly chosen person in the population (with so-called selfing, that is, the person might also choose herself). The functions $f^\pm(n_\pm(\sigma)) = n_\pm(\sigma)/N$ simply specify the probability of interacting with a person of the other opinion. Below, we also consider other examples, where the probability to change the mind is slightly more involved, but the overall structure of the terms will be similar. All other entries in state σ are not affected by a flip of the i'th person's opinion.

As mentioned above, the Σ_2-valued process $\hat{\sigma}_t$ induces a $V_N = \{0, .., N\}$-valued process X_t via $X_t = n_+(\hat{\sigma}_t)$. The transition rates of X_t are given by

$$X_t = k \rightarrow k+1 \quad \text{at rate} \quad \nu(N-k)\, f^+(k; N) \tag{3a}$$
$$X_t = k \rightarrow k-1 \quad \text{at rate} \quad \nu k \ f^-(N-k; N). \tag{3b}$$

We call a Markov process with transition rates given in this form an opinion process. It is straightforward to determine the stationary distribution of an opinion process if we have no absorbing states. As we aim at a specific notation that parallels the usual notation of Potts models, we derive the stationary distribution step by step. In what follows, we suppress the dependency of $f^\pm$ on N.

Proposition 1 (stationary distribution). *Assume $f^\pm(k) > 0$ for $k \in V_N = \{0, .., N\}$. Furthermore, let $F_\pm : V_N \rightarrow \mathbb{R}$ be defined by*

$$F_\pm(0) = 0, \qquad F_\pm(k) = \frac{1}{k} \sum_{\ell=0}^{k-1} \ln(f^\pm(\ell)) \quad \text{for } k \in V_N \setminus \{0\}. \tag{4}$$

Denoting the probability of state $k \in V_N$ in the stationary distribution by p_k, we have

$$p_k = \frac{e^{-\tilde{H}(k)}}{\tilde{Z}}, \quad \tilde{H}(k) = -\ln\left(\binom{N}{k}\right) - k\,F_+(k) - (N-k)F_-(N-k), \quad \tilde{Z} = \sum_{k=0}^{N} e^{-\tilde{H}(k)}. \tag{5}$$

Proof. The detailed balance equation for the stationary distribution $p_k = P(X = k)$ yields

$$p_k\, k\, f^-(N-k) = p_{k-1}(N-k+1)\, f^+(k-1). \tag{6}$$

Hence,

$$
\begin{aligned}
p_k &= p_0 \prod_{\ell=1}^{k} \frac{(N-\ell+1)f^+(\ell-1)}{\ell\, f^-(N-\ell)} = p_0 \frac{\prod_{\ell=1}^{k} N-\ell+1}{k!} \prod_{\ell=1}^{k} \frac{f^+(\ell-1)}{f^-(N-\ell)} \\
&= p_0 \binom{N}{k} \prod_{\ell=1}^{k} \frac{f^+(\ell-1)}{f^-(N-\ell)} = \frac{p_0}{\prod_{\ell=1}^{N} f^-(N-\ell)} \binom{N}{k} \left(\prod_{\ell=1}^{k} f^+(\ell-1) \right) \left(\prod_{\ell=k+1}^{N} f^-(N-\ell) \right) \\
&= \frac{p_0}{\prod_{\ell=1}^{N} f^-(\ell)} \binom{N}{k} \left(\prod_{\ell=1}^{k} f^+(\ell-1) \right) \left(\prod_{\ell=1}^{N-k} f^-(\ell-1) \right) \\
&= C \binom{N}{k} \exp\left(\sum_{\ell=0}^{k-1} \ln(f^+(\ell)) + \sum_{\ell=0}^{N-k-1} \ln(f^-(\ell)) \right),
\end{aligned}
$$

where C is determined by $\sum_{k=0}^{N} p_k = 1$. Together with the definition of $F_\pm$, $\tilde{H}(k)$, and $\tilde{Z}$, this formula proves the proposition. $\square$

Interestingly, the corresponding stationary distribution on Σ_2 can be written as

$$
Q(\sigma) = Z^{-1} \exp\left\{ n_+(\sigma)\, F_+(n_+(\sigma)) + n_-(\sigma)\, F_-(n_-(\sigma)) \right\} = Z^{-1} \prod_{i=1}^{N} \exp\left(F_{\sigma_i}(n_{\sigma_i}(\sigma)) \right), \tag{7}
$$

where $F_{\sigma_i} = F_+$ if $\sigma_i = 1$, and $F_{\sigma_i} = F_-$ if $\sigma_i = -1$; this is similar for n_{σ_i}. Each individual has an independent contribution $\exp(F_{\sigma_i}(.))$ to the probability of state σ, where $F_\pm(.)$ depend on the global statistics of the state via $n_\pm(\sigma)$. That is, $F_{\sigma_i}(n_{\sigma_i}(\sigma))$ can be regarded as the environment of individual i, which determines the probability of the opinion that individual i has adopted. It is, furthermore, interesting to observe that the population size N does not explicitly appear in the expression $\exp(F_{\sigma_i}(n_{\sigma_i}(\sigma)))$. The number of the opposite-opinion-supporters only comes in indirectly, as $n_+(\sigma)$ and $n_-(\sigma)$ add up to the given population size N.

For obvious reasons, we call the functions $F_\pm(.)$ the environmental conditions, or simply the environments, of the opinion process.

We have a degree of freedom in Equation (5). We might add a real constant A in the exponent, $p_k = \tilde{Z}^{-1} \binom{N}{k} e^{-\tilde{H}(k)+A}$. Then, $\tilde{Z}$ is still defined as the normalizing constant, guaranteeing that $\sum_{k=0}^{N} p_k = 1$. Thus, also in $\tilde{Z}$, the term e^A appears such that A cancels out and does not affect the value of p_k. Below, in the definition of the Boltzmann distribution, we will find a similar invariance. This invariance will be used later to eliminate singularities appearing in Section 3.2, where we investigate the large population limit with weak effects.

For now, let us discuss the freedom given by this invariance more in detail and, in particular, explore the implication for the choice of the environments. If we replace $F_\pm(k)$ by $\tilde{F}_\pm(k)$,

$$
\tilde{F}_\pm(k) = F_\pm(k) + U_\pm(k)/k, \qquad k = 1, \dots, N,
$$

where we choose $U_\pm(0) = 0$ in accordance with $F_\pm(0) = 0$ and require $U_\pm(k)$ to satisfy

$$
U_+(k) + U_-(N-k) = A, \qquad k \in V_N.
$$

Then, the stationary distribution p_k is not affected. We now go backward and determine transition rates $k\tilde{f}^\pm(N-k)$ that produce the new environments. Here, we use $\ln(\tilde{f}^\pm(k)) = (k+1)\tilde{F}_\pm(k+1) - k\tilde{F}_\pm(k) = \ln(f^\pm(k)) + U_\pm(k+1) - U_\pm(k)$ such that

$$
\tilde{f}^\pm(k) = f^\pm(k) \frac{e^{U_\pm(k+1)}}{e^{U_\pm(k)}}.
$$

Since $U_-(N-k) = A - U_+(k)$, we have

$$\tilde{f}^-(N-k) = f^-(N-k)\frac{e^{U_-(N-k+1)}}{e^{U_-(N-k)}} = f^-(N-k)\frac{e^{A-U_+(k-1)}}{e^{A-U_+(k)}} = f^-(N-k)\frac{e^{U_+(k)}}{e^{U_+(k-1)}}.$$

The function U_+ indeed cancels out in the detailed balance of Equation (6).

Please note that we are free to choose $U_+(k)$ (with the understanding that $U_+(0) = 0$). Next, we defined $A = U_+(N)$. Then, $U_-(k)$ is already determined via $U_-(N-k) = A - U_+(k)$ for $k \in V_N$. In particular, the choice $A = U_+(N)$ is equivalent with the requirement $U_-(0) = 0$ such that A is not free. The choice $A = U_+(N)$ is unique. All in all, the freedom we have is completely characterized by the choice of U_+.

Corollary 1. *The rates $k\,f^\pm(k)$ (or, alternatively, the environments $F_\pm(k) = \frac{1}{k}\sum_{\ell=0}^{k-1}\ln(f^\pm(\ell))$) determine the stationary distribution of an opinion process. The set of all opinion processes with stationary distributions coinciding with that given distribution is characterized by*

$$\tilde{f}^\pm(k) = f^\pm(k)\frac{e^{U_\pm(k+1)}}{e^{U_\pm(k)}} \qquad resp. \qquad \tilde{F}_\pm(k) = F_\pm(k) + U_\pm(k)/k, \tag{8}$$

where $U_\pm$ are functions satisfying $U_+(0) = U_-(0) = 0$ and

$$U_+(k) + U_-(N-k) = A \qquad for\ k \in V_N \quad and \quad A = U_+(N) \in \mathbb{R}. \tag{9}$$

2.2. Potts Machinery

Next, we introduce Potts models, where we again consider Potts models only on a full graph. Potts models on a full graph are often termed mean-field or Curie–Weiss models. In agreement with the literature, we return for the moment to the individual-based formulation of the opinion model, that is, we use Σ_2 as a state space.

We do not know the state of the population. The knowledge we assume to have are the results of some polls. For example, we could observe/measure the fraction of individuals with opinion $+1$. Consequently, we will know the expected number of persons in state 1, that is, $n_+(.)$. As the Potts models originate in physics, we follow the tradition and call these polls "observations" and, accordingly, the function $n_+(.)$ an "observable".

In general, observables are defined as functions $\hat{F} : \Sigma_2 \to \mathbb{R}$ without further restrictions. Another example of an observable that is often used is the number of pairs with identical opinions minus the number of pairs with different opinions,

$$\hat{F}(\sigma) = \sum_{k=1}^{N-1}\sum_{\ell=k+1}^{N}\sigma_k\sigma_\ell.$$

This observable incorporates information about correlations in the population.

Let us assume that we have m observables $\hat{F}_1, \ldots, \hat{F}_m$. Our knowledge about the state of the population is restricted to the knowledge of $\hat{f}_\ell := \hat{F}_\ell(\sigma)$, $\ell = 1, \ldots, m$. We represent our knowledge, and, particularly, the absence of complete knowledge about the state, in the form of a random measure Q on Σ_2. We, of course, require that $E(\hat{F}_\ell(.)) = \hat{f}_\ell$ such that Q does express our partial knowledge appropriately. However, there are many random measures that will satisfy this requirement. We express this lack of complete knowledge by uniquely defining Q as the random measure that maximizes the Shannon entropy $S(Q) = -\sum_{\sigma \in \Sigma_2} Q(\sigma)\ln(Q(\sigma))$, under the constraints $E(\hat{F}_\ell(.)) = \hat{f}_\ell$. The random measure $Q(.)$ we construct in this way is called the Boltzmann distribution.

We restrict ourselves to Cannings models such that the observables $\hat{F}_\ell$ only depend on Σ_2 via $n_\pm(\sigma)$ and thus can be defined as maps $\hat{F}_\ell : V_N \to \mathbb{R}$. Also, the Boltzmann distribution can be defined directly on V_N via $P(n_+(\sigma)) = \binom{N}{n_+(\sigma)}Q(\sigma)$. As $Q(\sigma) = Q(\tilde{\sigma})$ if $n_+(\sigma) = n_+(\tilde{\sigma})$, this formula defines P consistently. However, as the lack of knowledge

still concerns the state of the population in Σ_2, we do not use the original Shannon entropy for P, but measure the entropy for P by the entropy for the associated measure Q on Σ_2,

$$
\begin{aligned}
S_{\Sigma_2}(P) \quad &:= \quad -\sum_{\sigma \in \Sigma_2} Q(\sigma) \ln(Q(\sigma)) \\
&= \quad -\sum_{k \in V_N} \binom{N}{k} \left\{ P(k) \binom{N}{k}^{-1} \ln\left(P(k) \binom{N}{k}^{-1} \right) \right\} \\
&= \quad -\sum_{k \in V_N} P(k) \left(\ln(P(k)) - \ln(\binom{N}{k}) \right).
\end{aligned}
$$

Proposition 2. *Let $m \in \mathbb{N}$ denote the number of observables and $\hat{F}_\ell : V_N \to \mathbb{R}$, $\ell \in \{1, .., m\}$ the observables themselves. The Boltzmann distribution is the distribution $P : V_N \to \mathbb{R}_+$ that maximizes the Shannon entropy $S_{\Sigma_2}(P) = -\sum_{k \in V_N} P(k) \left(\ln(P(k)) - \ln(\binom{N}{k}) \right)$ under the constraint $E(\hat{F}_\ell(.)) = \hat{f}_\ell \in \mathbb{R}$. If the Boltzmann distribution P exists, then*

$$
P(k) = \frac{e^{-H(k)}}{Z}, \qquad H(k) = -\ln(\binom{N}{k}) - \sum_{\ell=1}^{m} \lambda_\ell \hat{F}_\ell(k), \qquad Z = \sum_{k \in V_N} e^{-H(k)}, \tag{10}
$$

where λ_ℓ, $\ell \in \{1, .., m\}$ are Langrange multipliers.

Proof. The proof consists of a short computation (see, e.g., [25,26]), based on the standard Lagrangian approach for the maximization of $S(P)$ under the constraints $E(\hat{F}_\ell(.)) = \sum_{k \in V_N} \hat{F}_\ell(k) P(k) = \hat{f}_\ell$, $\ell = 1, \ldots, m$, and $\sum_{k \in V_N} P(k) = 1$. Let $P(.)$ denote the set of probabilities $P(k) \in [0, 1]$ for $k \in V_N = \{0, \ldots, N\}$. We determine all values $P(k)$ by maximizing the function

$$
L(P(.), \lambda_1, .., \lambda_{m+1}) = S_{\Sigma_2}(P) + \sum_{\ell=1}^{m} \lambda_\ell \left(E(\hat{F}_\ell(.)) - \hat{f}_\ell \right) + \lambda_{m+1} \left(\sum_{k \in V_N} P(k) - 1 \right).
$$

Fix $\tilde{k} \in V_N$. If we equate the derivative of L with respect to $P(\tilde{k})$ to zero, we find

$$
0 = \frac{\partial L(P(.))}{\partial P(\tilde{k})} = -\ln(P(\tilde{k})) + \ln(\binom{N}{\tilde{k}}) - 1 + \sum_{\ell=1}^{m} \lambda_\ell \hat{F}_\ell(\tilde{k}) + \lambda_{m+1}
$$

and $P(\tilde{k}) = e^{-H(\tilde{k})} e^{\lambda_{m+1}-1}$. We obtain $e^{\lambda_{m+1}-1} = 1/Z$ by the condition that the probabilities add up to 1, that is, $\sum_{k \in V_N} P(k) = 1$. $\square$

Remark 1. *(a) In accordance with the literature, the Lagrangian multipliers $\lambda_1, \ldots, \lambda_m$ are not specified to actually determine a Boltzmann distribution that indeed satisfies $E(\hat{F}_\ell(.)) = \hat{f}_\ell \in \mathbb{R}$, but instead, the Lagrangian multipliers are from now on considered as parameters of the Boltzmann distribution.*

(b) We note that additive constants in the Hamiltonian do not affect the stationary distribution, as these constants appear in a multiplicative way in numerator $e^{-H(\tilde{k})}$ as well as in the denominator Z of the stationary distribution. We make use of this observation below to get rid of singularities appearing in the weak effects limit.

The Curie–Weiss model in sensu stricto is defined by observables that are polynomials of the second order, as we will discuss in Section 4.1. For the time being (Sections 2.3 and 3), we allow for more general functions as observables, as this is necessary to obtain and utilize a connection between the Curie–Weiss models and the opinion processes, which we discuss next.

2.3. Connection between Opinion Processes and the Curie–Weiss Model

An opinion process has a stationary distribution, and a Potts model a Boltzmann distribution. In order to connect opinion processes and Potts models, we ask which conditions the observables (Potts models), respectively, the environments (opinion processes) need to satisfy such that the stationary distribution of an opinion model coincides with the Boltzmann distribution. If we combine Propositions 1 and 2, we find the following corollary.

Corollary 2. *The stationary distribution of an opinion model with population size N and environment $F_\pm : V_N \to \mathbb{R}$ and the Boltzmann distribution for observables $\hat{F}_\pm : V_N \to \mathbb{R}$ coincide, if*

$$\hat{F}_\pm(k) = k\, F_\pm(k) \qquad and \qquad \lambda_\pm = 1. \tag{11}$$

Note that this corollary only states sufficient but not necessary conditions. Corollary 1 allows, for example, for constructing more observables that generate the same distribution. Considering the settings of this corollary for general $\lambda_\pm \in \mathbb{R}$, we observe

$$\lambda_\pm \hat{F}_\pm(k) = \lambda_\pm k\, F_\pm(k) = \sum_{\ell=0}^{k-1} \ln\big(\, f^\pm(\ell)^{\lambda_\pm} \,\big).$$

Therefore, a Boltzmann distribution with observables derived from environments of an opinion process is, for general $\lambda_\pm \in \mathbb{R}$, the stationary distribution of the opinion process with transition rates

$$X_t \to X_t + 1 \quad \text{at rate} \quad \nu\,(N-k)\, f^+(k)^{\lambda_+} \tag{12a}$$
$$X_t \to X_t - 1 \quad \text{at rate} \quad \nu\;k\; f^-(N-k)^{\lambda_-}. \tag{12b}$$

We call this family of opinion models the Glauber family for the given observables/environments. Please note that the standard Glauber dynamics for the spin up/spin down mean-field Ising model [25] utilized the freedom discussed in Corollary 1 in choosing a non-trivial function $U_+(k)$.

3. Large Population Limits

Below, we consider applications of opinion processes to data, where, often, the population size is in the magnitude of $N \approx 10^5$. Therefore, a continuity limit is of interest. As above, we might consider the dynamics (opinion process) and work out a diffusion limit, or we might focus on the Boltzmann distribution, and consider a continuum limit for that distribution. The interesting point is to understand the connection between the two resulting objects.

Herein, we note that the rates $f^\pm$, the corresponding environments $F_\pm$, respectively, and the observables $\hat{F}_\pm$ incorporate parameters. Two different assumptions about the dependencies of these parameters on the population size N (which itself is a parameter) lead to sensible continuum limits for $N \to \infty$: one is the strong effects limit and the other the weak effects limit, as introduced for the Moran model as part of population genetics [33–35] (where the Moran model with mutation is mathematically identical to the zealot model). In physics, the terms "extensive" and "nonextensive" models are used to describe the same idea [36,37].

To explain the difference between strong and weak effects, let us focus on the zealot model. In the strong effects models, the number of zealots is assumed to increase proportional with the population size. In that, the probability for a person to interact with a zealot is independent of the population size, and zealots clearly will affect the invariant distribution and also in the limiting case, $N \to \infty$.

In the weak effects setting, however, the number of zealots is constant and independent of the population size. A person's probability to interact with a zealot tends toward zero if the population size tends toward infinity. The effects of zealots become weaker and weaker

if N increases, which is why these models are termed "weak effects models". Mobilia, indeed, asked in his paper [23]: "Does a single zealot affect an infinite group of voters?" Though it sounds unlikely that this single zealot has some noteworthy effect on a huge population, it turns out that this is the case.

We define the weak/strong effects limit assumptions in a mathematically precise way in Sections 3.1 and 3.2 below (please do not confuse the weak effects limit, which refers to the scaling of the parameters, with a weak limit, which refers to the topology of convergence).

We assume throughout the current section that $f^{\pm}$ and $\hat{F}_{\pm}$ depend on k/N for $k \in V_N$, that is, on the sharing of an opinion $x = k/N$,

$$
\begin{aligned}
f^{\pm}(k; N) \quad &\text{is replaced by} \quad f^{\pm}(k/N; N) \quad k \in V_N, \\
\hat{F}_{\pm}(k; N) \quad &\text{is replaced by} \quad \hat{F}_{\pm}(k/N; N) \quad k \in V_N,
\end{aligned}
$$

where

$$
f^{\pm} : [0, 1] \times \mathbb{N} \to \mathbb{R}, \qquad \hat{F}_{\pm} : [0, 1] \times \mathbb{N} \to \mathbb{R}
$$

such that $f^{\pm}(x; N)$, $F_{\pm}(x; N)$ are well-defined for all $x \in [0, 1]$. Since the observables are defined via $f^{\pm}$, we rewrite them separately; see Section 3.1 below.

For given N, the old and the new scaling are mathematically equivalent. However, if we aim at a limit $N \to \infty$, the new scaling is a reasonable and very helpful assumption that the most commonly used models actually fulfill.

3.1. Strong Effects Limit

In the strong effects limit, we assume that $f^{\pm}(x; N)$ approximate functions $\mathbf{f}^{\pm}(x) \in C^2$ for $N \to \infty$ such that

$$
\lim_{N \to \infty} f^{\pm}(x; N) = \mathbf{f}^{\pm}(x) \text{ in } C^2[0, 1] \tag{13}
$$

is well-defined. For clarity of notation, we write the limits of rate functions, environments, and observables for $N \to \infty$ in bold. To adapt the definition region of the associated environments from the discrete state space V_N to the continuous state space $x \in [0, 1]$, we re-define the environments as

$$
F_{\pm} : [0, 1] \times \mathbb{N} \to \mathbb{R}, \quad F_{\pm}(x; N) := \frac{1}{N x} \sum_{\ell=0}^{\lfloor N x \rfloor - 1} \ln(f^{\pm}(\ell/N; N)).
$$

For $x = k/N, k \in V_N$, we recover the environments as defined in Proposition 1. Therewith, the environments also satisfy a proper limit for $N \to \infty$,

$$
\mathbf{F}_{\pm}(x) := \lim_{N \to \infty} F_{\pm}(x; N) = \lim_{N \to \infty} \frac{1}{x} \sum_{\ell=0}^{\lfloor N x \rfloor - 1} \ln(f^{\pm}(\ell/N; N)) \frac{1}{N} = x^{-1} \int_0^x \ln(\mathbf{f}^{\pm}(y)) \, dy. \tag{14}
$$

We also introduce a limit for the observables. Here, a certain subtlety appears: we have $\hat{F}(k/N; N) = kF(k/N; N)$. As $x = k/N$, this formula becomes $\hat{F}(k/N; N) = N x F(x; N)$. In leading order, the observables are $\mathcal{O}(N)$. We thus scale them by $1/N$ and define

$$
\hat{\mathbf{F}}_{\pm}(x) := \lim_{N \to \infty} \frac{1}{N} \hat{F}(k/N; N) = x \mathbf{F}_{\pm}(x) = \int_0^x \ln(\mathbf{f}^{\pm}(y)) \, dy. \tag{15}
$$

We emphasize at this point that when using $\hat{\mathbf{F}}_{\pm}(x)$ below, we need to take the scale $1/N$, introduced at this point, into account.

To obtain the behavior of the opinion process under the strong effects limit, we briefly sketch the Kramers Moyal [38] expansion of the model, which is—as usual—truncated at the second order to obtain a Fokker–Planck (or Kolmogorov forward) equation.

Proposition 3. *The Kramers Moyal expansion up to the second order for the Glauber family with limiting observables* $\hat{\mathbf{F}}_{\pm}(x) = \int_0^x \ln(\mathbf{f}^{\pm}(y))\, dy$ *is given by*

$$\partial_t u(x,t) = -\nu\, \partial_x\left(\left((1-x)\,\mathbf{f}^+(x)^{\lambda_+} - x\,\mathbf{f}^-(1-x)^{\lambda_-} \right) u(x,t) \right) \tag{16}$$

$$+ \frac{\nu}{2N}\, \partial_{xx}\left(\left((1-x)\,\mathbf{f}^+(x)^{\lambda_+} + x\,\mathbf{f}^-(1-x)^{\lambda_-} \right) u(x,t) \right).$$

Proof. We start with the master equations

$$\dot{p}_k = -\nu\left((N-k)f^+(k/N;N)^{\lambda_+} + k\,f^-(1-k/N;N)^{\lambda_-} \right) p_k$$

$$+ \nu(N-(k-1))f^+((k-1)/N)^{\lambda_+} p_{k-1} + \nu(k+1)f^-(1-(k+1)/N)^{\lambda_-} p_{k+1}.$$

If we now assume that $p_k \approx h\, u(x,t)$ for some smooth probability density $u(x,t)$, where $x = k\,h$ and $h = 1/N$, we have $\partial_t u(x,t) \approx \dot{p}_k$, and

$$\partial_t h\, u(x,t) = -\nu\, h^{-1}((1-x)f^+(x;N)^{\lambda_+} + xf^-(1-x;N)^{\lambda_-})\, u(x,t)$$
$$+ \nu\, h^{-1}\,(1-x-h)\, f^+(x-h;N)^{\lambda_+} u(x-h,t)$$
$$+ \nu\, h^{-1}\,(x+h)\, f^-(1-x-h;N)^{\lambda_-} u(x+h,t).$$

The Taylor expansion of the last two terms up to the second order neglecting the error term and using the limit of $f^{\pm}(\,.\,;N)$ for $N \to \infty$ yields the Kramers Moyal expansion. $\square$

Next, we turn to the Boltzmann distribution. In the continuum limit, we will denote the Boltzmann distribution (and later the stationary distribution of a limiting stochastic opinion process as well) by $\varphi(x)$.

Proposition 4. *The Boltzmann distribution for observables* $\hat{\mathbf{F}}_{\pm}(x) = \int_0^x \ln(\mathbf{f}^{\pm}(y))\, dy$ *and large N is given in leading order by the Hamiltonian*

$$H(x) = N\left(-H_2(x) - \lambda_+ \int_0^x \ln(\mathbf{f}^+(y))\, dy - \lambda_- \int_0^{1-x} \ln(\mathbf{f}^-(y))\, dy \right), \tag{17}$$

where $H_2(x) = -x\ln(x) - (1-x)ln(1-x)$ *is the binary entropy. The limiting Boltzmann distribution reads* $\varphi(x) = e^{-H(x)}/Z(x)$, *where* $Z(x) = \int_0^1 e^{-H(x)}\, dx$.

The result is a consequence of Equation (10), the scale of $\hat{\mathbf{F}}_{\pm}$ with respect to N introduced in (15), and the well-known approximation of the binomial coefficient by means of the binary entropy

$$\ln\left(\binom{N}{n} \right) = NH_2(n/N) - \frac{1}{2}\ln(2\pi N) - \frac{1}{2}\ln(x(1-x)) + \mathcal{O}(N^{-1}). \tag{18}$$

The connection between the stationary distributions of the Kramers Moyal expansion Equation (16) and the stationary distribution Equation (17) is not clear; though they are derived from the associated Potts models and opinion processes, they look rather different. The next proposition clarifies the connection.

Proposition 5. *Assume that N is large. Let* $\mathbf{f}(x) = (1-x)\,\mathbf{f}^+(x)^{\lambda_+} - x\,\mathbf{f}^-(1-x)^{\lambda_-}$, *and assume* $\mathbf{f}(\mu) = 0$, $\mathbf{f}'(\mu) < 0$ *for some* $\mu \in (0,1)$. *Then, the local normal approximation of the Boltzmann distribution (17) coincides with the stationary distribution of the Ornstein–Uhlenbeck approximation of the Kramers Moyal expansion (16) at $x = \mu$ in leading order in N.*

Proof. We first investigate the Boltzmann distribution. Since

$$\frac{H'(\mu)}{N} = \ln(\mu) - \ln(1-\mu) - \lambda_+ \ln(\mathbf{f}^+(\mu)) + \lambda_- \ln(\mathbf{f}^-(1-\mu)) = \ln\left(\frac{\mu\,\mathbf{f}^-(1-\mu)^{\lambda_-}}{(1-\mu)\,\mathbf{f}^+(\mu)^{\lambda_+}}\right) = 0,$$

we have a critical point of H(x) at $x = \mu$. The second derivative reads, again using $\mathbf{f}(\mu) = 0$,

$$
\begin{aligned}
\frac{H''(\mu)}{N} \\
&= \frac{(1-\mu)\,\mathbf{f}^+(\mu)^{\lambda_+}}{\mu\,\mathbf{f}^-(1-\mu)^{\lambda_-}} \frac{(1-\mu)\,\mathbf{f}^+(\mu)^{\lambda_+}\frac{d}{d\mu}(\mu\,\mathbf{f}^-(1-\mu)^{\lambda_-}) - \mu\,\mathbf{f}^-(1-\mu)^{\lambda_-}\frac{d}{d\mu}((1-\mu)\,\mathbf{f}^+(\mu)^{\lambda_+})}{((1-\mu)\,\mathbf{f}^+(\mu)^{\lambda_+})^2} \\
&= \frac{\frac{d}{d\mu}(\mu\,\mathbf{f}^-(1-\mu)^{\lambda_-}) - \frac{d}{d\mu}((1-\mu)\,\mathbf{f}^+(\mu)^{\lambda_+})}{(1-\mu)\,\mathbf{f}^+(\mu)^{\lambda_+}} = \frac{-\mathbf{f}'(\mu)}{(1-\mu)\,\mathbf{f}^+(\mu)^{\lambda_+}} = \frac{|\mathbf{f}'(\mu)|}{(1-\mu)\,\mathbf{f}^+(\mu)^{\lambda_+}}.
\end{aligned}
$$

Hence, $H(x) = H(\mu) + \frac{1}{2}(x-\mu)^2 H''(\mu) + \mathcal{O}((x-\mu)^3)$. Locally, at $x = \mu$, the stationary distribution $e^{-H(x)}/Z$ behaves as $N(\mu, \sigma^2)$ with

$$\sigma^2 = H''(\mu)^{-1} = \frac{(1-\mu)\,\mathbf{f}^+(\mu)^{\lambda_+}}{N\,|\mathbf{f}'(\mu)|} = \frac{\mu\,\mathbf{f}^-(\mu)^{\lambda_-}}{N\,|\mathbf{f}'(\mu)|}.$$

Next, we proceed to the stationary distribution based on the Kramers Moyal expansion (16). The leading order terms of the linearization of the drift and the noise term at $x = \mu$ yield the Ornstein–Uhlenbeck approximation

$$
\begin{aligned}
\partial_t u(x,t) &= -\nu\,\partial_x\left((x-\mu)\mathbf{f}'(\mu)\,u(x,t)\right) + \frac{\nu}{N}\,\partial_{xx}\left((1-\mu)\,\mathbf{f}^+(\mu)^{\lambda_+}\,u(x,t)\right) \quad (19) \\
&= \partial_x\left\{\nu\left((x-\mu)\mathbf{f}'(\mu)\,u(x,t)\right) + \frac{\nu}{N}\,\partial_x\left((1-\mu)\,\mathbf{f}^+(\mu)^{\lambda_+}\,u(x,t)\right)\right\}
\end{aligned}
$$

where we used $(1-\mu)\,\mathbf{f}^+(\mu)^{\lambda_+} = \mu\,\mathbf{f}^-(1-\mu)^{\lambda_-}$ in the noise term. In order to identify the stationary distribution, we substitute $u(x,t) = \varphi(x)$ with $\varphi(x) = e^{-a(x-\mu)^2}$ into the term bracketed with curly brackets,

$$
\begin{aligned}
&\nu\left((x-\mu)\mathbf{f}'(\mu)\,e^{a(x-\mu)^2}\right) + \frac{\nu}{N}\,\partial_x\left((1-\mu)\,\mathbf{f}^+(\mu)^{\lambda_+}\,e^{-a(x-\mu)^2}\right) \\
&= \left\{\mathbf{f}'(\mu) - 2a\frac{\nu}{N}(1-\mu)\,\mathbf{f}^+(\mu)^{\lambda_+}\right\}(x-\mu)\,e^{-a(x-\mu)^2}.
\end{aligned}
$$

This term becomes zero and, therewith, $\varphi(x)$ an invariant measure, if $a^{-1} = 2\frac{(1-\mu)\,\mathbf{f}^+(\mu)^{\lambda_+}}{N\,|\mathbf{f}'(\mu)|}$. Therefore, the stationary distribution of this approximate Fokker–Planck equation is a normal distribution $N(\mu, \sigma^2)$, where

$$\sigma^2 = \frac{(1-\mu)\,\mathbf{f}^+(\mu)^{\lambda_+}}{N\,|\mathbf{f}'(\mu)|} = \frac{\mu\,\mathbf{f}^-(1-\mu)^{\lambda_-}}{N\,|\mathbf{f}'(\mu)|}.$$

$\square$

For large N, the stationary distribution will be concentrated close to μ and, hence, in the relevant regions both stationary distributions, that of Equation (19) and of Equation (17), coincide. In that, we consider the Kramers Moyal expansion (16) as the Glauber dynamics of the Boltzmann distribution.

3.2. Weak Effects Limit

The idea of the weak effects limit is to take a simple reference model as a basis and to perturb this model in such a way that for $N \to \infty$, the transition rates converge back to that of the reference model. For good reasons, as we will find out later, we use the voter

model as the reference model. In the voter model, at rate ν, a person copies the opinion of a randomly chosen person in the population (inclusive "selfing"—that does mean that the focal person might by chance copy her opinion from herself). Therewith,

$$X_t = k \to k+1 \quad \text{at rate} \quad \nu(N-k)\, f_{voter}^+(k/N) = \nu(N-k)\,\frac{k}{N} \tag{20a}$$

$$X_t = k \to k-1 \quad \text{at rate} \quad \nu k\; f_{voter}^-(1-k/N) = \nu k\;\frac{N-k}{N}, \tag{20b}$$

and hence, $f_{voter}^\pm(x) = x$. For the weak effects limit, we allow for rates $f^\pm(x; N)$, which depend on N, as long as $\lim_{N\to\infty} f^\pm(x; N) = f_{voter}^\pm(x)$. Also, the Lagrangian parameters are allowed to depend on N, $\lambda_\pm = \lambda_\pm(N)$, and the paradigm of weak effects requires again $\lim_{N\to\infty} f^\pm(x; N)^{\lambda_\pm(N)} = f_{voter}^\pm(x)$. Therefore, the expansion of $f^\pm$ with respect to $1/N$ has $f_{voter}^\pm(x)$ as the zero order term and some arbitrary (well-behaved) function $g^\pm(x)$ as the first order coefficient. Similarly, the zero order term of the expansion of $\lambda^\pm(N)$ is 1, while the first order terms are some parameters $\kappa^\pm \in \mathbb{R}$, which we are free to choose. All in all, the Glauber family suited for the weak effects limit assumes the form

$$f^\pm(k/N; N)^{\lambda_\pm(N)} = \left(f_{voter}^\pm(k/N) + \frac{1}{N}g^\pm(k/N) + \mathcal{O}(N^{-2}) \right)^{1+\kappa_\pm/N}. \tag{21}$$

As we will see, for the weak effects limit, we not only assume the appropriate scaling of the parameters, but also rescale time. In that, even for $N \to \infty$, we still obtain a non-trivial limiting process and do not simply return to the voter model.

As a consequence, we need to re-consider the large population limit conducted in (14) and (15) more in detail and also pay attention to the terms of order $\mathcal{O}(N^{-1})$.

Proposition 6. *Consider the observables for the opinion process defined by* (21),

$$\hat{F}_\pm(x; N) = \sum_{\ell=0}^{\lfloor Nx \rfloor - 1} \ln\left(\frac{\ell}{N} + \frac{1}{N}g^\pm(\ell/N) + \mathcal{O}(N^{-2}) \right)\frac{1}{N}.$$

With the definition

$$G^\pm(x) = \int^x g^\pm(y)/y\, dy, \tag{22}$$

the leading order terms of the expansion of $\hat{F}_\pm(x; N)$ *in* $1/N$ *reads*

$$\hat{F}_\pm(x; N) = x\left(\ln(x) - 1\right) - \frac{1}{2N}\ln(x) + \frac{1}{N}G^\pm(x) + \mathcal{O}(N^{-2}).$$

Proof. We first rewrite the sum such that it extends to $\lfloor Nx \rfloor$ instead of $\lfloor Nx \rfloor - 1$,

$$
\begin{aligned}
\hat{F}_\pm(x; N) &= \sum_{\ell=0}^{\lfloor Nx \rfloor - 1} \ln\left(\frac{\ell}{N} + \frac{1}{N}g^\pm(\ell/N) + \mathcal{O}(N^{-2}) \right)\frac{1}{N} \\
&= \sum_{\ell=0}^{\lfloor Nx \rfloor} \ln\left(\frac{\ell}{N} + \frac{1}{N}g^\pm(\ell/N) + \mathcal{O}(N^{-2}) \right)\frac{1}{N} \\
&\quad - \ln\left(\frac{\lfloor Nx \rfloor}{N} + \frac{1}{N}g^\pm(\lfloor Nx \rfloor/N) + \mathcal{O}(N^{-2}) \right)\frac{1}{N}.
\end{aligned}
$$

Next, we replace the sum by an integral. Here, we take the Euler–McLaurin correction terms into account in the step from sum to integral. Furthermore, we note that an additive constant in the Hamiltonian does not change the stationary distribution. Instead of $\sum_{\ell=0}^{\lfloor Nx \rfloor}(\dots)$, we can change the lower starting value to any value independent on x and might, e.g., consider

$\sum_{\ell=\lfloor N/2 \rfloor}^{\lfloor Nx \rfloor}(\ldots)$ instead. Only the upper bound of the sum matters; we only need an anti-derivative. To express this fact, we skip the lower bound of the integral and proceed (where the first equal sign has to be interpreted with the knowledge that we did drop some irrelevant term):

$$
\begin{aligned}
\hat{F}_{\pm}(x;N) &= \int^{\lfloor Nx \rfloor} \ln\left(\frac{\ell}{N} + \frac{1}{N}g^{\pm}(\ell/N) + \mathcal{O}(N^{-2})\right) d\ell \frac{1}{N} \\
&\quad - \frac{1}{2}\ln\left(\frac{\lfloor Nx \rfloor}{N} + \frac{1}{N}g^{\pm}((\lfloor Nx \rfloor)/N) + \mathcal{O}(N^{-2})\right)\frac{1}{N} + \mathcal{O}(N^{-2}) \\
&= \int^{x} \ln\left(y + \frac{1}{N}g^{\pm}(y)\right) dy - \frac{1}{2}\ln\left(x + \frac{1}{N}g^{\pm}(x)\right)\frac{1}{N} + \mathcal{O}(N^{-2}) \\
&= \int^{x} \ln\left(y + \frac{1}{N}g^{\pm}(y)\right) dy - \frac{\ln(x)}{2N} - \frac{1}{2}\ln\left(1 + \frac{1}{N}g^{\pm}(x)/x\right)\frac{1}{N} + \mathcal{O}(N^{-2}) \\
&= \frac{-1}{2N}\ln(x) + \int^{x} \ln\left(y\left\{1 + \frac{1}{Ny}g^{\pm}(y)\right\}\right) dy + \mathcal{O}(N^{-2}) \\
&= x(\ln(x) - 1) - \frac{1}{2N}\ln(x) + \int^{x} \ln\left(1 + \frac{1}{Ny}g^{\pm}(y)\right) dy + \mathcal{O}(N^{-2}).
\end{aligned}
$$

Note that $\lim_{N\to\infty}\int^{x}\ln\left(1 + \frac{1}{yN}g^{\pm}(y)\right)dy = 0$ (in the sense that 0 is a possible limiting anti-derivative) such that this integral only contributes to terms of order $\mathcal{O}(N^{-1})$ or higher. We introduce

$$
G^{\pm}(x) = \lim_{N\to\infty} N\int^{x}\ln\left(1 + \frac{1}{yN}g^{\pm}(y)\right)dy = \int^{x} g^{\pm}(y)/y\, dy
$$

and obtain the result. □

Therewith, we are in the position to establish the following proposition.

Proposition 7. *Assume that the functions $f^{\pm}(x)$ scale with N as described above,*

$$
f^{\pm}(x;N) = \left(x + \frac{1}{N}g^{\pm}(x) + \mathcal{O}(N^{-2})\right)
$$

and scale the Lagrangian parameters by $\lambda_{\pm} = 1 + \kappa_{\pm}/N$. Then, in leading order, the Hamiltonian reads

$$
H(x) = \ln(x(1-x)) - G^{+}(x) - G^{-}(1-x) - \kappa_{+}\zeta(x) - \kappa_{-}\zeta(1-x) \tag{23}
$$

with $\zeta(x) = -x(1 - \ln(x))$.

Proof. Recall $H_2(x) = -x\ln(x) - (1-x)\ln(1-x)$. With the scaling $\lambda_{\pm} = 1 + \kappa_{\pm}/N$, we obtain

$$
\begin{aligned}
&\lambda_{+}\hat{F}_{+}(x;N) + \lambda_{-}\hat{F}_{-}(x;N) \\
&= -N H_2(x) - N - \frac{1}{2}\ln(x(1-x)) + \kappa_{+}x\ln(x) + \kappa_{-}(1-x)\ln(1-x) \\
&\quad -\kappa_{+}x - \kappa_{-}(1-x) + G^{+}(x) + G^{-}(1-x) + \mathcal{O}(N^{-1}).
\end{aligned}
$$

We use the approximation of the binomial coefficient (18) to obtain the result with $H(x) = -\ln(\binom{N}{Nx}) - \lambda_{+}\hat{F}_{+}(x;N) - \lambda_{-}\hat{F}_{-}(x;N)$, where we drop terms independent of x and terms of a higher order in N^{-1}. □

It is remarkable and typical for the weak effects scaling that the Hamiltonian and, in that, also the Boltzmann distribution become independent of N. We now turn to the underlying opinion process and discuss the Kramers Moyal expansion under the scaling assumed.

Proposition 8. *The Kramers Moyal expansion of the Glauber family under the weak effects scaling in rescaled time $T = \nu t/N$ is given by*

$$u_T(x, T) = -\partial_x \Big(\big((1-x)g^+(x) - xg^-(x) + \kappa_+(1-x)\ln(x) + \kappa_- x\ln(1-x)\big) u(x, T) \Big) \quad (24)$$
$$+ \partial_{xx}\Big(x(1-x)\, u(x, T) \Big).$$

Proof. The Glauber family with the weak effects scaling is defined by

$$f^{\pm}(k/N) = \left(\frac{k}{N} + \frac{1}{N}g^+(k/N) + \mathcal{O}(N^{-2}) \right)^{1+\kappa_{\pm}/N},$$

while the rate $X_t = k \to k+1$ reads $\nu(1-x)f^+(..)|_{k=xN}$ and that for $X_t = k \to k-1$ is $\nu\, x\, f^-(..)|_{k=(1-x)N}$. The drift term becomes in leading order

$$\nu(1-x)\left(x + \frac{1}{N}g^+(x) + \mathcal{O}(N^{-2})\right)^{1+\kappa_+/N} - \nu x\left((1-x) + \frac{1}{N}g^-(1-x) + \mathcal{O}(N^{-2})\right)^{1+\kappa_-/N}$$
$$= \frac{\nu}{N}\Big((1-x)g^+(x) - xg^-(x) + \kappa_+(1-x)x\ln(x) + \kappa_- x(1-x)\ln(1-x)\Big) + \text{h.o.t.}$$

(where h.o.t. is a placeholder for higher order terms), and the coefficient of the noise term becomes in leading order

$$\nu(1-x)\left(x + \frac{1}{N}g^+(x) + \mathcal{O}(N^{-2})\right)^{1+\kappa_+/N} + \nu x\left((1-x) + \frac{1}{N}g^+(1-x) + \mathcal{O}(N^{-2})\right)^{1+\kappa_-/N}$$
$$= 2\nu x(1-x) + \text{h.o.t.}$$

If we rescale time $T = \nu\, t/N$, we obtain the result. $\square$

The Kramers Moyal expansion in rescaled time becomes, as the Hamiltonian, independent of N.

Proposition 9. *The stationary distribution of the Kramers Moyal expansion is identical with the Boltzmann distribution $\varphi(x) = \exp(-H(x))/Z$, where the Hamiltonian $H(x)$ is given in (23).*

Proof. We plug the Boltzmann distribution into the right-hand side of the Kramers Moyal expansion. Thereto, we note that

$$\varphi'(x) = -\varphi(x)\, H'(x) = -\left(\frac{1}{x} - \frac{1}{1-x} - (G^+)'(x) + (G^-)'(1-x) - \kappa_+\ln(x) - \kappa_-\ln(1-x) \right)\varphi(x).$$

Furthermore, $\frac{d}{dx}G^{\pm}(x) = g^{\pm}(x)/x$ such that

$$\partial_x\Big(x(1-x)\varphi(x) \Big) = -x\varphi(x) + (1-x)\varphi(x) + x(1-x)\varphi'(x)$$
$$= \Big((1-x)g^+(x) - xg^-(x) + x(1-x)\kappa_+\ln(x) + \kappa_- x(1-x)\ln(1-x) \Big)\varphi(x).$$

If we take the derivative of this equation with respect to x, we indeed find that $\varphi(x)$ is a stationary solution of Equation (24). $\square$

Remark 2. *For the weak effects limit, we have chosen the voter model with $f^{\pm}(x) = x$ as the reference model. If $N \to \infty$, the rates of the model at hand converge back to this model. This choice looks, at first glance, arbitrary. It is, however, up to the freedom characterized in Corollary 1,*

a unique choice: the binary entropy $H_2(x)$ generates in the Hamiltonian terms $Nx \ln(x)$ and $N(1-x) \ln(1-x)$. For a weak effects limit to exist, these terms of order $\mathcal{O}(N)$ need to be balanced and annihilated by the environments of the reference model, which already forces the reference model to be the voter model (or some model, which is, according to Corollary 1, equivalent to the voter model).

4. Four Models: Curie–Weiss, Weak and Strong q-Voter Model, and Reinforcement

We use the framework introduced above to briefly introduce four opinion models we intend to apply to data.

4.1. Curie–Weiss Model

To introduce the classical Curie–Weiss model, we start with the Potts machinery. Recall that the central ingredients are observables, that is, functions $\hat{F}_\pm : V_N \to \mathbb{R}$, which form constraints when determining the random measure maximizing the Shannon entropy. Perhaps the most simple, non-trivial case is given by observables, which are polynomials of the second order,

$$F_\pm(k/N; N) = a_\pm k + b_\pm k^2/N = N\left(a_\pm x + b_\pm x^2\right)_{x=k/N}.$$

We do not need a zero order term, as additive constants in the Hamiltonian do not influence the Boltzmann distribution. Furthermore, we scale the quadratic term by $1/N$ to balance the squared terms in case of large N. With this setting, the Hamilton defined in (10) reads

$$\begin{aligned} H(x) &= -\ln(\binom{N}{Nx})) - \lambda_+ F_+(x; N) - \lambda_- F_-(1-x; N) \\ &= -\ln(\binom{N}{Nx})) - N\left(\lambda_+ b_+ x^2 + \lambda_- b_- (1-x)^2 + \lambda_+ a_+ x + \lambda_- a_- (1-x)\right). \end{aligned}$$

We might rewrite the quadratic terms as

$$\lambda_+ b_+ x^2 + \lambda_- b_- (1-x)^2 = \frac{(\lambda_+ b_+ + \lambda_- b_-)}{2}(x^2 + (1-x)^2) + (\lambda_+ b_+ - \lambda_- b_-)x + \frac{\lambda_- b_- - \lambda_+ b_+}{2}$$

We define $J = \lambda_+ b_+ + \lambda_- b_-$ and $h_\pm$ appropriately (e.g., $h_+ = \lambda_+ b_+ - \lambda_- b_-$). Furthermore, for historical reasons, we introduce $h_- = 0$ and drop terms independent of x. Therewith, we obtain

$$H(x) = -\ln(\binom{N}{xN})) - N\left\{\frac{J}{2}\left(x^2 + (1-x)^2\right) + h_+ x + h_- (1-x)\right\}, \tag{25}$$

which is the standard form of the model on the state space V_N [25]. We can still reduce the number of parameters from four (N, J, h_+, h_-) to three (N, J, h) with $h = h_+ - h_-$ as the additive constant, which, apparently, can again be dropped.

Last, we check the existence of the strong and the weak effects limit. The strong effects limit can be derived trivially, simply by replacing the binomial coefficient by $NH_2(x)$, cf. (18),

$$H(x) = -N\left(H_2(x) + \frac{J}{2}(x^2 + (1-x)^2) + h_+ x + h_- (1-x)\right).$$

The binary entropy introduces logarithmic terms of order N into the Hamiltonian $H(x)$. As the observables of the Curie–Weiss model consist of polynomial terms, they cannot cancel these logarithmic terms such that the Hamiltonian always incorporates nontrivial terms of order $\mathcal{O}(N)$. In the proper weak effects limit, however, terms of this order are not present. Thus, a weak effects limit for the Curie–Weiss model is not possible (cf. Remark 2). This will be different for the other models we shall discuss next.

4.2. Two Flavors of the q-Voter Model

For the q-voter model, we do not start with the stationary distribution but the transition rates. Perhaps the most simple extension of the voter model where no opinion can die out is the zealot model, where we have $N^\pm$ zealots for the opinion ± 1. Palombi and Toti ([24] p. 337) call zealots "stubborn agents [...] who never change political preference". In our case, zealots are not real persons but represent sources of information that stand for a specific opinion. These could be politicians, friends, newspapers, or social media channels. As in the voter model, individuals copy their opinion from a randomly chosen person, now also from the zealots, which leads to

$$f^\pm(k; N) = \frac{N^\pm + k}{N + N^- + N^+}.$$

The corresponding Glauber family is given by

$$X_t = k \to k+1 \quad \text{at rate} \quad \nu(N-k)\left(\frac{N^+ + k}{N + N^- + N^+}\right)^{\lambda_+} \tag{26a}$$

$$X_t = k \to k-1 \quad \text{at rate} \quad \nu k\left(\frac{N^- + (N-k)}{N + N^- + N^+}\right)^{\lambda_-}. \tag{26b}$$

For $\lambda_+ = \lambda_- > 1$, this is the q-voter model for a homogeneous population [27]. In the case of $\lambda_\pm = 1$, we are back in the zealot model: a person simply copies the opinion of another person or a zealot. If $\lambda_\pm > 1$, the model can be interpreted as follows: the person will ask $\lambda_\pm$ other persons for their opinion and will only change her opinion if all these other persons have the identical opinion.

4.2.1. q-Voter Model—Strong Effects

We consider the strong effects limit: if $N^\pm = \eta^\pm N$, that is, if the number of zealots scales linearly with the population size where $\eta^\pm$ are the proportionality constants,

$$f^\pm(x; N) = \frac{N\eta^\pm + Nx}{N + N\eta^- + N\eta^+} = \frac{\eta^\pm + x}{1 + \eta^- + \eta^+},$$

such that $f^\pm(x; N)$ becomes independent of N, and $f^\pm(x; N) \equiv \mathbf{f}^\pm(x)$. Therewith, the limiting observables are given by

$$\hat{\mathbf{F}}_\pm(x) = \int_0^x \ln\left(\frac{\eta^\pm + y}{1 + \eta^- + \eta^+}\right) dy = (x + \eta^\pm)\ln(\eta^\pm + x) - x - x\ln(1 + \eta^- + \eta^+) + C$$

where C is a constant. The stationary distribution in the strong effects limit is given by $\varphi(x) = Z^{-1}\exp(-H(x))$, where

$$H(x) = -N\left(H_2(x) + \lambda_+\left[(x + \eta^+)\ln(\eta^+ + x) - x(1 + \ln(1 + \eta^- + \eta^+))\right]\right. \tag{27}$$

$$\left. + \lambda_-\left[((1-x) + \eta^-)\ln(\eta^- + (1-x)) - (1-x)(1 + \ln(1 + \eta^- + \eta^+))\right]\right).$$

In the last formula, we again made use of the fact that we are allowed to drop constant terms from the Hamiltonian.

4.2.2. q-Voter Model—Weak Effects

We now go into the weak effects limit for the q-voter model. The basis is the zealot model with $f^\pm(k; N) = \frac{N^\pm + k}{N + N^- + N^-}$. For the weak effects limit, the number of zealots $N^\pm$ does not scale with the population size, and in that, zealots become rare if N becomes large. We can choose the time units, and in this, we have a degree of freedom in the form

of a multiplicative positive constant. This freedom can be used to replace the original denominator $N + N^- + N^-$ by N and work with

$$f^{\pm}(x; N) = x + \frac{1}{N} g^{\pm}(x), \qquad g^{\pm}(x) \equiv N^{\pm}.$$

That is, for our particular choice, the functions $g^{\pm}(x)$ are independent of x. Recall that we also expand the Lagrangian parameters in terms of N and write $\lambda_{\pm} = 1 + \kappa_{\pm}/N$.

To obtain the weak effects limit, we note that $G^{\pm}(x) = \int^x g^{\pm}(y)/y \, dy = \ln(x) \, N^{\pm}$ and obtain the Hamiltonian

$$H(x) = -\ln(x)(N^+ - 1) - \ln(1 - x)(N^- - 1) - \kappa_+ x \ln(x) - \kappa_-(1 - x) \ln(1 - x) + (\kappa_+ - \kappa_-) \, x.$$

with the corresponding stationary distribution

$$\varphi(x) = C \, x^{N^+ - 1} \, (1 - x)^{N^- - 1} \, e^{-\kappa_+ x(1 - \ln(x))} \, e^{-\kappa_-(1 - x)(1 - \ln(1 - x))}. \tag{28}$$

The stationary distribution becomes a beta distribution in the case of the zealot model ($\kappa_{\pm} = 0$); this result was first derived in the context of population genetics, where the zealot model is termed the Moran model ([33] page 108). The extension to the weak effects limit of the Glauber family presented here is novel.

4.3. Reinforcement Model—Weak Effects

We add one more model, which also allows for a strong as well as weak effects limit and which is, as the q-voter model, also a descendant from the zealot model: the reinforcement model [21]. The idea of the reinforcement model is to express the psychological mechanisms that lead to filter bubbles and echo chambers: several kinds of cognitive biases let individuals communicate with persons of the opposite opinion with less awareness than with individuals of their own opinion. Some interactions with the opposite group are ignored. In that, the effective size of the opposite group is reduced by a factor $\theta_{\pm} \in (0, 1]$. The zealot model is described by

$$f^{\pm}(k/N; N) = \frac{\theta_{\pm}(N^{\pm} + k)}{N^{\pm} + N - k + \theta_{\pm}(N^{\mp} + k)}.$$

We focus on the weak effects limit and hence keep (as in the weak limit of the q-voter model) $N^{\pm}$ independent of N. Furthermore, we choose $\theta_{\pm} = 1 - \vartheta_{\pm}/N$ such that we return to the voter model if N becomes large; the Lagrangian parameters are taken to be $\lambda_{\pm} = 1$ and are not scaled. It turns out that the computations assume a simpler form (additive constants will vanish below in the first order term of the expansion) if we use a trivial time scale such that a multiplicative term $N^+/N + 1 + N^-/N$ appears in the rates,

$$f^{\pm}(k/N; N) = (N^+/N + 1 + N^-/N) \frac{(1 - \vartheta_{\pm}/N)(N^{\pm} + k)}{N^{\pm} + N - k + (1 - \vartheta_{\pm}/N)(N^{\mp} + k)}.$$

Therewith, we obtain the expansion of the rates with respect to $1/N$,

$$f^{\pm}(x; N) = x + \frac{1}{N} g^{\pm}(x) + \mathcal{O}(N^{-2}), \qquad g^{\pm}(x) = \vartheta_{\pm} x^2 - \vartheta_{\pm} x + N^{\pm}.$$

Consequently, we obtain

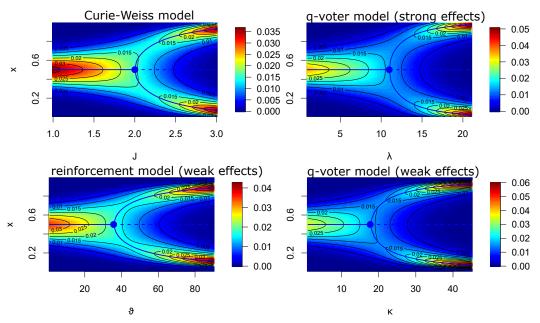$$G^{\pm}(x) = \int^x \frac{g^{\pm}(y)}{y} \, dy = \frac{1}{2} \vartheta_{\pm} x^2 - \vartheta_{\pm} x + N^{\pm} \ln(x) = -\frac{1}{2} \vartheta_{\pm} \left(2x - x^2 \right) + N^{\pm} \ln(x)$$

and the stationary distribution (we use Equation (23) with $\kappa_{\pm} = 0$)

$$\varphi(x) = C \, x^{N^+ - 1} \, (1 - x)^{N^- - 1} \, e^{-\vartheta_+ x(2 - x)/2} \, e^{-\vartheta_-(1 - x)(2 - (1 - x))/2}. \tag{29}$$

If we compare the stationary distribution of the weak effects q-voter model and the weak effects reinforcement model, we find a striking similarity: in both cases, the measure is an adaptation of the beta distribution, $\varphi(x) = C\, x^{N^+-1}\,(1-x)^{N^--1}\, e^{a_+\zeta(x)}\, e^{a_-\zeta(1-x)}$ where $a_\pm = \kappa_\pm$ and $\zeta(x) = -x(1 - \ln(x))$ in the q-voter case, while $a_\pm = \vartheta_\pm/2$ and $\zeta(x) = -x(2-x)$ in the reinforcement case. As both functions $\zeta(x)$ resemble each other in that $\zeta(0) = 0$, $\zeta(1) = 1$, and both are convex, we expect very similar behavior for the two models if we take $\vartheta_\pm = 2\kappa_\pm$ (also inspect Figure 1).

Figure 1. Phase transitions of the four models in the symmetric case. For a given parameter (x-axis), the density of the distribution is indicated (by the heat and contour plot) over x (y-axis). The blue lines indicate local maxima (solid lines) and local minima (dashed line) if the parameter (given at the x-axis) is fixed, while the dot marks the phase transition. (Curie–Weiss: $h_\pm = 0$; strong q-voter: $\eta_\pm = 5$ and $N = 20$; weak q-voter: $N^\pm = 10$, $\kappa_\pm = \kappa$; weak reinforcement: $N^\pm = 10$, $\vartheta_\pm = \vartheta$).

As a last remark, we note that we do allow, in the weak q-voter model and the weak reinforcement model, not only for positive parameter values $\kappa_\pm$ and $\vartheta_\pm$, but also for negative values. While positive values for these parameters lead to filter bubbles and echo chambers (a person hesitates to change her mind), negative values have the interpretation that the person iss open-minded and pays particular attention to the opposite opinion. In the case of the functioning of democracies and their institutions, open-minded people are much more preferable because they are easier to find compromises for solving political problems.

4.4. Model Behavior

We will not go deeper into the analysis (which can be found, e.g., for the Curie–Weiss model in [25,26], for the q-voter model in the strong effects limit in [27], and for the reinforcement model in [21]), but simply refer to Figure 1, which shows that all four models undergo a phase transition if the coupling between the individuals is sufficiently large. We also emphasize that the models based on the voter and zealot model (which do not allow for phase transitions per se), the two kinds of q-voter model, and the reinforcement model exhibit phase transitions. The mechanisms modifying the effects of zealots target in/outgroup communication. If in/outgroup communication is sufficiently strong, a bimodal distribution appears via a phase transition. Also, the behavior under non-symmetric conditions (parameters) leads to similar behavior of all four models.

It is interesting to note that the Curie–Weiss and the strong effects q-voter model incorporate the population size N explicitly, while the (weak effects limit of) the q-voter

and the reinforcement model become independent of N. Usually, in applications, N is very large, and if we naively take N to the population size, the variance generated by the model is much smaller than the variance that is present in empirical data. The way out is to assume that individuals cluster together and to estimate an effective population size $N = N_{eff}$ along with the other parameters, which, of course, is slightly dubious but pragmatic [26]. The weak effects models elegantly circumvent this difficulty.

5. Data Analysis

We use data from four different Western democracies that represent different types of government and electoral systems (see [39] pp. 145–161, 271, and [40] and Table 1 for further details). Concerning the governmental system, we have one presidential (US), one semi-presidential (France), and two parliamentary systems (UK, Germany). In the case of the electoral systems, we have two majority systems with a first-past-the-post design and relative majority (US, UK) and one majority system in France with a two-round system and absolute majority. In Germany, we have a mixed member proportional system that combines both a first-past-the-post vote and proportional representation. Finally, for each country, we study different numbers of elections (US: six presidential elections, 2000–2020; UK: 20 parliamentary elections, 1945–2019; FRA: second round of five presidential elections, 2002–2022; GER: two parliamentary elections, 2017–2021; the data sources are indicated in the data availability statement).

Table 1. Design of the study: data set used.

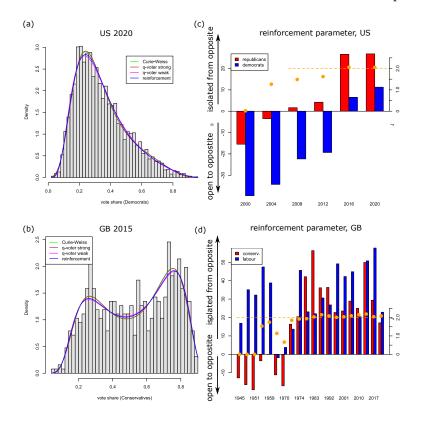|	**US**	**UK**	**France**	**Germany**
Government system	presidential	parliamentary	semi-presidential	parliamentary
Electoral system	Plurality/majority: single-member districts, first-past-the-post, relative majority	Plurality/majority: single-member districts, first-past-the-post, relative majority	Plurality/majority: two-round system, absolute majority	Mixed: mixed member proportional
Election years covered in the analysis	2000–2020 (6 presidential elections)	1945–2019 (20 parliamentary elections)	2002–2022 (5 presidential elections, second round)	2017–2021 (2 parliamentary elections)

Such a design is useful in understanding how the models used can explain the dynamics in different institutional settings with different political cultures and in varying periods of time. This comparative approach gives us more information about the functioning of the mechanisms in different contexts (e.g., [12,13]) and contributes to the existing research often based on single-case studies (e.g., [24,41]). In the analysis, we consider each election district as an i.i.d. repetition of the election. Herein, we obviously reduce the complexity of the data by neglecting social co-factors and spatial effects. In that, we obtain an empirical distribution of vote shares and can use a maximum likelihood estimator. Please find the technical details, particularly the algorithm used to perform the maximum-likelihood estimation, in Appendix A.

The present approach to data analysis is based on a steady state assumption, that is, the opinion formation process is assumed to be approximately in equilibrium. If there is a huge shift in the vote share of a candidate or party in recent times, this steady state assumption might not be met. The tables with estimates, p-values for the Kolmogorov–Smirnov test, and AICs can be found in Appendix A.

United States data: The densities of the four models (Curie–Weiss, weak and strong q-voter model, and reinforcement model) are rather similar (Figure 2a), and the Kolmogorov–Smirnov tests also resemble each other (see table in Appendix A.1). In the year 2000 only, the p-values of this test were small (between 0.018 and 0.04); in that year, apart from the Democrats and the Republicans, the green candidate did win a small but reasonable fraction

of votes such that the dichotomous models might not be completely suited. In 2016, we also have about four percent of third-party votes, but at that point, the models fit better. Maybe this is due to the fact that polarization in the American society has already grown during these past 16 years. The AIC always selects the (weak) reinforcement model as the best-suited model, but in the year 2020, the weak q-voter model fit best. However, in this year, the reinforcement model and the weak q-voter model are very close. The strong models always perform worse (Figure A1). We clearly find a trend in the parameters, which shows that the model moves in time more and more towards a phase transition.

Figure 2. (**a**,**b**): Histograms of the vote shares for (**a**) the presidential US elections in 2020 and (**b**) the parliamentary elections in 2015 in UK, together with the probability density for the four models. (**c**,**d**): Bars indicate the reinforcement parameters $\vartheta_{\pm}$ for (**c**) Democrats and Republicans, and (**d**) the conservatives and the Labour Party (left axis); orange bullets indicate parameter J of the Curie–Weiss model (right axis), while the horizontal dashed orange line indicates the threshold for the phase transition of that model for $h = 0$.
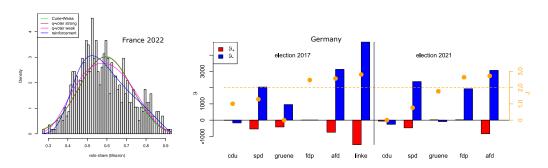
United Kingdom data: The trend in the reinforcement parameters/coupling is particularly interesting (Figure 2d). As can be clearly visualized in the empirical and the estimated distributions for the election from 2015, the UK indeed became super-critical. We observe a bimodal distribution in 2015. It is most interesting to see that the theoretical prediction of possible phase transitions is realized in the UK.

France data: Particularly in 2022, the empirical distribution of vote shares is skewed. The strong effects models have difficulties dealing with this result, while the weak effects models are more flexible; in particular, the reinforcement model still performs well.

German data: As we have a proportional electoral system, the dichotomous model requires adaptation—for each party, we distinguish between the votes in favor of this party versus the votes for all other parties. In Figure 3, we find the reinforcement parameters together with the coupling J of the Curie–Weiss model for those parties in the 2017 and 2021 elections that did receive at least 5% of the votes (where we did disregard the CSU, as this is a Bavarian local party that only stands for election in few election districts). If we
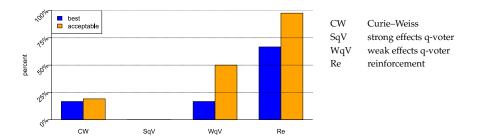
focus on the best models (lowest AIC), these models indeed are able to meet the empirical vote share distributions quite well (Kolmogorov–Smirnov test), with only one exception: the left-wing party Left Party (die linke) in the 2017 election. Due to historical reasons, this party performs very differently in the federal states coming from the former East and West Germany, respectively. Though these historical effects can also be understood to be based on opinion dynamics and in/out-group behavior, all models have difficulties in capturing this data structure. The assumption of a homogeneously mixed population may no longer be appropriate; instead, a two-island model would capture the communication structure better. Also, the (relatively recent) right-wing party AfD, which also performs rather differently in the two regions (former East/West Germany), poses a problem for all models but the reinforcement model, at least according to the Kolmogorov–Smirnov test. However, the difference in the two regions is less pronounced than in the case of the Left Party, which might also be the reason why the reinforcement model is still able to handle the vote share data of the AfD.

Figure 3. Right: Histogram of the vote shares for the presidential elections in France (2022), together with the probability density for the four models. Left: Histogram of $\vartheta_{\pm}$ and J (together with critical level, orange) for the parties in 2017, 2021, which hold more than 5% of the votes (threshold), apart from the CSU, which is a Bavarian local.

Summary of the estimations: In almost all elections, at least some of the models describe the empirical data adequately (according to the Kolmogorov–Smirnov test). If we compare the performance of the models according to the AIC, we find that the weak effects models outperform the strong effects models, and the weak effects reinforcement model is superior to the weak effects q-voter model (Figure 4). It is interesting that although the structures of the weak effects models are very similar, we nevertheless find a difference in their suitability for practical applications.

Figure 4. Performance of the four models. For the 42 elections we consider in the present paper, we indicate the percentage for which the models performed best according to the AIC (yellow) and acceptable (AIC has a maximum distance of 2 to the best model).

We should keep in mind that we work with aggregated data (only the outcome of election in election districts that typically have a population of hundreds of thousands of individuals). We might consider the model parameters as a reduction of the data complexity to a low-dimensional parameter space, which allows us to better interpret the data. As the Kolmogorov–Smirnov test indicates the appropriateness of the models, we can be

confident that the models capture at least some fundamental structure in the data. In that, the interpretation of the parameters suggested by the models will be appropriate.

Political Science Interpretations

United States: Looking at the parameters and measuring the actual strength of reinforcement in the reinforcement model, we can see that since the year 2000, there has been a continuous trend for a much more polarized voting behavior in the United States. For the voters of the Republican Party, the switch could already be seen in 2008 with the election of Barack Obama, and it continued in his re-election, while it became dominant in 2016 and 2020 when Donald Trump became candidate of the Republican Party. The voters for the Democratic Party clearly also changed from open- to closed-mindedness in the political realm. In fact, what we can see is that during those 20 years, the political discourse became so polarized in the US public arena that now both the parties and their voters are becoming clearly more and more separated from each other. For example, Binder [42] shows that the frequency of legislative gridlock has risen since the 1990s, and other studies show that polarization in the US citizenry is not going down [43].

Great Britain: Due to the electoral system (first-past-the-post) and the governmental system (parliamentary), it is clearly useful for the two dominant parties—the Labour Party on the left and the Conservative Party (Tories) on the right—to keep a certain or even a high degree of polarization so that they can form the government on their own. The bars of the reinforcement model show this clearly for Labour for almost all elections since 1945, but for the Tories, this strategy started only in the 1970s and became very dominant since the second election in 1974 and the Thatcher years, 1979–1990. Additionally, we can see that in 2019, the reinforcement parameters had not been so strong. This might be the result of the Brexit decision in 2016 and its political aftermath in which the Conservatives had become the party of the Leavers (in support of Brexit) and the Remainers have split between Labour and the Liberal Democrats [44].

France: The French Party System has undergone a political change in the last few years since 2017, so some already argue that we may see the rise of a new French Party System [45]. In 2017, both presidential candidates of the traditional left-wing (Parti Socialiste) and right-wing parties (Les Républicains) were disqualified in the first round of the presidential elections [46]. And this happened again in 2022, when both Emmanuel Macron of the centrist "La République En Marche" and Marine Le Pen of the new-named right-wing populist "Rassemblement National" were, for the second time, the political opponents in the second round of the election [47]. That said, we can see in Figure A3 for the previous elections that the models did not explain so much for the second round.

Germany: Here, we can see that the right-wing populist party Alternative for Germany (AfD) was able, in both elections, to create their own space for resonating with their voters. In 2017, both the Social Democrats (SPD), the Left Party, and the Greens had been able to mobilize their voters against the AfD, but in 2021, it was the SPD and the liberal party (FDP). Compared to the left parties, the FDP tried to position itself as the party of Freedom, where some parts of the party also raised their critical voice against the political means of the former government of the two conservative parties CDU and CSU and the Social Democrats used during the Covid-19 pandemic. This can be observed in their parliamentary work, where they used so called "Kleine Anfragen" to question the government parties the most compared to all the other remaining opposition parties in the German Bundestag [48]. Therefore, they had been a competitor both to the AfD and the other left and center parties and also gained a lot of support by young voters [49].

6. Discussion of the Findings and Their Interpretation in Political Sciences

In the present paper, we first discussed the connection between Potts models and stochastic opinion models, both for well mixing populations. In particular, we did provide an alternative approach to the q-voter model as a natural extension of the zealot model in the view of Glauber dynamics. Consequently, motivated by similar constructions in

population genetics, we introduced a strong and also a particularly weak effects continuum limit. While the strong continuum limit is generically possible (and locally approximates a normal distribution), the weak effects limit requires additional structure. In that, the Curie–Weiss model only allows for the strong limit, while the q-voter model has both limits. Afterwards, we additionally introduced the reinforcement model, which has its foundation not in the Potts machinery, but is derived based on considerations of the impact of several kinds of cognitive biases on communication, especially on the resulting in/out-group communication strategies. Also, that model allows for both continuum limits, where it turned out that the weak q-voter and the weak reinforcement model are very similar in their mathematical structure. Basically, both are an adaptation of the beta distribution, which is the well-known weak effects limit of the zealot (or Moran) model [33]. We also found that only models that are based on the voter model allow for a weak effects limit, which indicates that these models are in some sense special.

After these theoretical considerations, we turned to test the models based on election data. Herein, we used each election district as an i.i.d. repetition of the election, neglecting social co-factors, which vary between election districts, as well as spatial factors. We also assume the opinion process to be approximately in equilibrium such that the stationary distribution is an adequate description of the data; in case of a large shift in the vote share of a candidate or party, this assumption can also be called into question. Though the approach was simple, we mostly found that the models meet the data quite well, where the weak models performed better than the strong models, and the weak reinforcement model outperformed the weak q-voter model. It is interesting that the weak effects models seem to be better suited to describe the data appropriately. The comparison of models and data is always challenging, but it also seems that population genetics, where the weak effects such as weak selection are often used, rather supports that these kind of models are well suited for real world applications [50]. The background could be that striking and immediately disruptive events are rare. Most stimuli are weak and require time to unfold their effect. If this observation is correct, weak effects models with their slow time scale might indeed be a better description of reality than strong effects models with a fast time scale. As a practical consequence, we propose to use rather weak effects opinion models than strong effects models in empirical studies, which also has the advantage of not needing to choose an appropriate population size, which is a well-known problem in itself [26,33]. Though the models allowing for a weak effects limit are rather special, they seem to be a powerful description of reality and still have sufficient flexibility to address different mechanisms and different real world (electorate) systems.

Elections are at the center stage of modern representative democracies. Correspondingly, research on elections and attempts to explain the formation of their results are also central. Prominent and established approaches use statistical data concerning the social characteristics of voters to determine their voting behavior (e.g., [51]). But, there is an ongoing debate among scholars that the correlation between social characteristics and voting behavior has diminished over the last few decades (e.g., [52–54]). Additionally, party membership is also in decline, which also has consequences for voter turnout and voting behavior (e.g., [55,56]). As Clarke et al. [57] bluntly declared, for understanding electoral choice, one has to look elsewhere. It is not that the classic approaches have lost all their explanatory power, but it makes sense to look for explanations that are less context-dependent.

By focusing on the dynamics of opinion formation preceding the act of voting, the models discussed in this paper promise insights both into the empirical explanation of elections as such, as well as important aspects of the theory of democracy.

Our leading assumption, ensuring a larger independence of specific social contexts, is opinion formation via frequently contacting social sources of information, constituting a ubiquitous mechanism of collective decision-making. For sure, this assumption also holds for elections. The sources of information here may be real persons or media of all kinds. Pamphlets, newspapers, magazines, radio, TV, and the diversity of social media have

accompanied political discussion since the early modern period. Albeit the basic mechanism is taken to be the same everywhere, its effects may be modulated by the impact of or the interaction with other mechanisms of other sections within an "organized complexity". Electoral systems in their specific forms are nested within the broader construction of a political system. The effects of opinion formation processes concerning elections are shaped by specific institutional settings as well as the political culture of the respective countries.

The model's reduction to only two opinions may look like too crude a simplification. But it is not as implausible as it may appear at first glance. As Denver and Johns [58] stress, when preparing their decisions, voters don't "sit down before an election to comb through the parties' manifestos and make detailed calculations of the costs and benefits of voting for each party. (…) Such a process is neither realistic nor particularly rational ([58] p. 294)". Rather, voters base their decision on just a few subjects dominating the discussion. "Issue Voting" and "Valence Voting" denominate the approaches based on that assumption. Where Issue Voting stresses the main issues of the election campaign like economic questions or social policy, Valence Voting focuses, for example, on the performance of the incumbent government. "Neither involves complex calculations; indeed, the simpler versions of both approaches have fared better when confronted with the empirical evidence (ibid.)". And both aspects fare better than explaining results with respect to voters' social characteristics. Therefore, looking at opinion and opinion formation in this simple form offers a promising starting point to delve deeper into campaigning and voting.

In this view, the models in their present form may be interpreted as to suppose a stage of the election process, where the main issues are already settled. From here, the model may be enhanced by stepwise nesting the basic opinion formation process within a whole set of similar modeled ones. The determination of the salient issues and valences may be the next step. The party that succeeds in putting its topics in place may be at an advantage. A variant of the Issue Voting approach stresses that it is not only the preferences on the specific topics of the very election of today that affect the voter's decision, but general values and principles ([58] p. 295). We can think of processes concerning ideological backgrounds running on a larger time scale spanning across two or more elections, affecting the probability with which a voter makes up his mind. Another perspective would look at the developments within the zealots in particular. This would mean looking at the development of party programs and strategies, also in the form of opinion formation among party members. The possibilities are manifold.

The reinforcement model in particular also highlights important aspects concerning the theory of democracies. Especially in the liberal tradition of democracy, it is a common view to interpret campaigning and elections as a market analogous competition, where votes are exchanged for programs and personnel (see [59]). Competition appears as a form of regulated and limited conflict. The opponent's purpose is not to harm the antagonist, but only to be better. The idea behind this is that the aspiration to trump the adversary leads to the advancement of a common good, qualitatively better or cheaper products and processes in economy, better theories and methods in science, and better programs and personnel in politics.

The zealot model, as a predecessor of the reinforcement model, was used before in economics for market analysis [60]. But behind this application lies a model designed to explain foraging processes of ant colonies [61]. It describes a form of collective information processing against an uncertain environment. Time and again the colony has to leave an established feeding ground and look for another in time. It is inspiring to see the similarities. Political communities also have to alter their processes and organization because of altering circumstances, for example, transforming their way of living to a more sustainable way. In this way, an open political process, defining problems and looking for solutions, is a form of collective information processing, too. This idea was emphasized by Karl Popper [62], for example, and further developed by John Dryzek [63]. It is not implausible to assume that a part of the success of democracies in general is their dealing with the world's shakiness in an analogous way to modern science.

Whatever makes the workers of the ant colony change their paths, the driving force behind the parliamentary process, at least from the perspective of the liberal standard model of democracy, is party competition. Since party competition is itself affected by special interests and personal ambition of politicians, democracies need additional features to balance these forces and bring the wanted effects of the competition to the fore. This has been part of the considerations from Harrington [64] to Tocqueville [65] to Dewey [66] but shall not be the point here.

What the reinforcement model enables us to see is the possible polarization between the opponents. It is important to note that polarization indicating a higher grade of conflict is not a problem per se, as higher grades of conflict are not as well. As sociologist Lewis Coser (1967) [67] argued for, conflicts in the first place point to societal problems within a society urging to deal with their causes. If the community is productively addressing that challenge, society reintegrates on a new level. We could see such effects, for example, in the course of the environmental movement in the 1970s to 1990s in Germany. The polarization on the side of the Greens was high in the beginning when they cracked open the consensus of the established parties on the use of nuclear power, and became lower again when environmental issues were successfully established on the agenda.

Polarization may become problematic when the reinforcement effects are strong on both sides of the debate. Conflicts can be disruptive too. American philosopher of law Ronald Dworkin was already asking in 2006 [68], looking at the polarization in the US "Is Democracy possible here?". The polarization in the US has not been in decline since then (Pew Research Center 2022, [43]). Polarization may also become problematic when actors show no tendency to consent or compromise, enabling reintegration. This appears to be the case with the populist movements and parties of the last decade. On the other hand, party polarization may generate stronger party attachments, which could also be a desirable strategy for political competitors ([69] p. 350). And here is probably the point where political scientists (at least at the moment) have to reach out for other methods than mathematical modeling, too. Qualitative analysis of texts or focus groups may be an appropriate means here. However, it should be ascertained that the reinforcement model supplies us with a strong indicator concerning an important variable of political processes. And since the claim that a society is polarized is also often used in an alarmist way, impeding compromise, the more it is helpful to have this indicator. It should be ascertained further that because of its context independence, the model will be useful when we look not only at well-established democracies of the West, but also at young ones or democracies in other world regions.

Appendix A. Data Analysis

The maximum likelihood estimates were performed in R [70], based on the optimizer `nlm`. In the case of the Curie–Weiss and the strong q-voter model, all parameters were optimized in one run. The weak effects models have two parameters related to the zealots ($N^\pm$ and $\eta^\pm$, respectively), and two parameters related to ingroup/outgroup communication structures ($\lambda_\pm$, $\kappa_\pm$, and $\vartheta_\pm$). Here, we used an EM-like algorithm by alternately optimizing the parameters for the zealots and holding the in/outgroup parameters, and then optimizing the parameters for the in/outgroup while holding the zealot parameters until the vector of parameters became approximately constant.

The Kolmogorov–Smirnov test was performed to test for the appropriateness of the current model, and the AIC was determined in order to compare the models. The columns of the tables are to be read specifically for the model, e.g., the $J/\eta^+/N^+$ column holds J in case of the Curie–Weiss model, η^+ for the strong q-voter model, and N^+ for the weak q-voter model and reinforcement model.

In the resulting tables, we mark the models that have the minimal AIC (best performing models) for the corresponding elections in yellow. Of course, if the AIC of two models is close, it is not sensible to dismiss one of the two models. Burnham et al. [71] suggest that models that have a difference of at most two do perform comparably. These models are marked in gray. Additionally, we provide a graphic for the differences in the AICs. Here, we use for each given election the AIC of the Curie–Weiss model as a reference and provide

$$\Delta\text{AIC} = \text{AIC other model} - \text{AIC Curie–Weiss}.$$

That is, if this value is positive, the "other model" performs better than the Curie–Weiss model; if it is negative, the Curie–Weiss model performs better. Moreover, if ΔAIC is, e.g., larger for the weak q-voter model than for the reinforcement model, then the weak q-voter model performs better than the reinforcement model.

Appendix A.1. United States Data

US presidential elections in the years 2000–2020 were analyzed, where we only considered Democrats and Republicans and the vote share of the Democrats (e.g., votes for Democrats/votes for Democrats or Republicans) and dismiss the votes for any other candidate. The performance of the AICs for the models is shown in Figure A1.
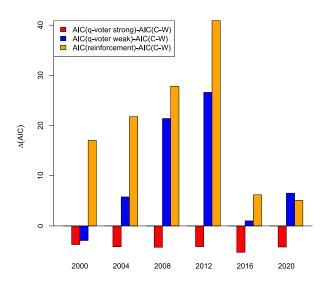

Figure A1. Comparison of the models' AICs for the US.

Year	Model	$J/\eta^+/N^+$	$h/\eta^-/N^-$	$\lambda_+/\kappa_+/\vartheta_+$	$\lambda_-/\kappa_-/\vartheta_-$	N	AIC	p_{KS}
2000	Curie–Weiss	0.015	−0.38	−	−	15.1	−4299.2	0.04
2000	q-voter strong	7	19.5	−2.7	−12.7	12.2	−4295.5	0.046
2000	q-voter weak	3.7	0.00031	−3.2	−22.1	−	−4296.3	0.023
2000	reinforcement	3	6.9×10^{-6}	−15.5	−39.6	−	−4316.3	0.1
2004	Curie–Weiss	1.3	−0.2	−	−	31.8	−4150.6	0.39
2004	q-voter strong	31.9	50.2	36.6	69.4	31.5	−4146.4	0.43
2004	q-voter weak	6.8	2.8×10^{-6}	7.5	−21.2	−	−4156.4	0.29
2004	reinforcement	4.4	0.12	−3.5	−34.3	−	−4172.4	0.48
2008	Curie–Weiss	1.5	−0.11	−	−	34.2	−3489.9	0.31
2008	q-voter strong	15.8	20.3	23.3	32.7	35	−3485.6	0.23
2008	q-voter weak	7.9	0.04	11.1	−17.3	−	−3511.3	0.49
2008	reinforcement	5	1.3	1.7	−22.4	−	−3517.7	0.28
2012	Curie–Weiss	1.6	−0.13	−	−	31.7	−3074.7	0.15
2012	q-voter strong	22	25.2	35.5	42.8	31.8	−3070.6	0.19
2012	q-voter weak	6.5	0.018	10.5	−14.7	−	−3101.3	0.3
2012	reinforcement	4.3	0.92	4.2	−19.3	−	−3115.6	0.88
2016	Curie–Weiss	2.1	−0.085	−	−	56.8	−3167.7	0.71
2016	q-voter strong	4.9	8	10.1	19.3	58.2	−3162.5	0.79
2016	q-voter weak	9.3	2.4	22.3	−5	−	−3168.7	0.3
2016	reinforcement	6.4	4.1	26.7	6.5	−	−3173.9	0.59
2020	Curie–Weiss	2.1	−0.084	−	−	55.7	−3066	0.51
2020	q-voter strong	13.8	25.3	25.7	59.7	56.6	−3061.8	0.59
2020	q-voter weak	9.2	3.5	21.5	−2	−	−3072.5	0.84
2020	reinforcement	6.5	5.2	26.9	11.3	−	−3071.1	0.69

Appendix A.2. UK Data

We investigated the vote share of the conservative among the conservative and Labour votes. The performance of the AICs for the models is shown in Figure A2.

Year	Model	$J/\eta^+/N^+$	$h/\eta^-/N^-$	$\lambda_+/\kappa_+/\vartheta_+$	$\lambda_-/\kappa_-/\vartheta_-$	N	AIC	p_{KS}
1945	Curie–Weiss	3.3×10^{-8}	−0.26	−	−	12.1	−593.4	0.86
1945	q-voter strong	7.6	0.67	−6.1	−0.56	9.5	−589.6	0.84
1945	q-voter weak	1.1	11.2	−13.7	10.6	−	−592.6	0.16
1945	reinforcement	3.3	11.7	−12.8	16.9	−	−596.7	0.64
1950	Curie–Weiss	7.1×10^{-8}	−0.039	−	−	12.2	−628.5	0.64
1950	q-voter strong	4.2	4.1	−24	−23	3.7	−624.9	0.61
1950	q-voter weak	0.098	21.3	−21.7	40.3	−	−643.4	0.98
1950	reinforcement	2.8	15	−16.5	35.1	−	−646.2	0.96
1951	Curie–Weiss	6.4×10^{-7}	0.013	−	−	13.1	−652.9	0.34
1951	q-voter strong	2.6	−0.0066	−9.4	−0.68	6.5	−651.4	0.45
1951	q-voter weak	3.3×10^{-8}	29.3	−26.1	60.6	−	−666.4	0.92
1951	reinforcement	2.6	14.4	−19.3	32.2	−	−669.4	0.94
1955	Curie–Weiss	1.5	0.018	−	−	43.8	−715.8	0.65
1955	q-voter strong	17.9	16.5	28.5	25.5	43.4	−711.8	0.74
1955	q-voter weak	1.7	17.7	−16.5	31.9	−	−730	0.96
1955	reinforcement	5.3	16.8	−3.5	47.4	−	−734.2	0.87
1959	Curie–Weiss	1.7	0.022	−	−	63.2	−699.5	0.47
1959	q-voter strong	27.2	24	49.4	41.4	61.6	−695.4	0.52
1959	q-voter weak	3.3	17.9	−11.8	33.4	−	−709.4	0.82
1959	reinforcement	5.7	14	0.018	38.8	−	−710.9	0.88
1964	Curie–Weiss	1.1	−0.011	−	−	22.4	−639	0.72
1964	q-voter strong	9.6	9.5	11.3	11.1	21.9	−635	0.75
1964	q-voter weak	2	5.5	−9.1	0.89	−	−638.4	0.87
1964	reinforcement	3.1	5.3	−11.2	−1.9	−	−639.2	0.95

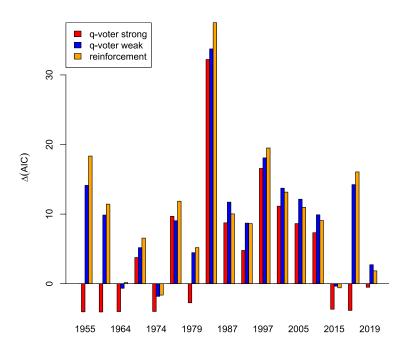| Year	Model	$J/\eta^+/N^+$	$h/\eta^-/N^-$	$\lambda_+/\kappa_+/\vartheta_+$	$\lambda_-/\kappa_-/\vartheta_-$	N	AIC	p_{KS}
1970	Curie–Weiss	0.66	0.017	—	—	15.1	−625.1	0.41
1970	q-voter strong	−0.62	0.56	0.73	2.2	67.9	−628.9	0.96
1970	q-voter weak	0.35	10	−15.3	13.8	—	−630.3	0.91
1970	reinforcement	1.9	6.5	−17.1	3.8	—	−631.7	0.89
1974	Curie–Weiss	1.9	0.0039	—	—	50.5	−497.5	0.7
1974	q-voter strong	5.7	5.6	11.6	11.2	50.6	−493.5	0.69
1974	q-voter weak	8.2	7.7	10	8.8	—	−495.7	0.65
1974	reinforcement	7.3	6.6	16.4	13.7	—	−495.9	0.74
1974	Curie–Weiss	1.9	−0.018	—	—	51.7	−450.8	0.15
1974	q-voter strong	0.042	0.22	1	1.5	178.1	−460.5	0.88
1974	q-voter weak	7.1	14.2	6.8	26.8	—	−459.8	0.34
1974	reinforcement	7.1	13.2	20.5	45.6	—	−462.6	0.89
1979	Curie–Weiss	1.9	0.021	—	—	57.5	−473.5	0.31
1979	q-voter strong	10.4	3.8	25.4	6.7	59.7	−470.8	0.37
1979	q-voter weak	18.4	7.8	37	7.1	—	−478	0.63
1979	reinforcement	12.6	7.7	42.2	23.1	—	−478.7	0.65
1983	Curie–Weiss	2	0.072	—	—	32.8	−327.7	0.044
1983	q-voter strong	1.4	0.04	5	0.81	108.1	−360	0.54
1983	q-voter weak	22.2	5.3	51.4	3.8	—	−361.5	0.51
1983	reinforcement	14.2	5.6	56.3	22	—	−365.3	0.67
1987	Curie–Weiss	2.1	0.029	—	—	51.9	−293.6	0.21
1987	q-voter strong	0.016	0.011	1	1	692.9	−302.4	0.61
1987	q-voter weak	11.1	8.6	23.6	17.9	—	−305.4	0.55
1987	reinforcement	8.6	6.7	36.1	30.6	—	−303.7	0.57
1992	Curie–Weiss	2.1	0.017	—	—	52.5	−309.1	0.2
1992	q-voter strong	0.2	0.099	1.5	1.2	114.1	−313.9	0.44
1992	q-voter weak	12.3	8.1	25.8	14.3	—	−317.8	0.47
1992	reinforcement	9.2	6.7	36.3	26.8	—	−317.7	0.6
1997	Curie–Weiss	2	−0.056	—	—	40.4	−376	0.34
1997	q-voter strong	0.07	1.8	0.83	6	96.3	−392.6	0.52
1997	q-voter weak	6.2	16	7.3	33.8	—	−394.1	0.9
1997	reinforcement	6.2	13.1	22.8	49.1	—	−395.5	0.54
2001	Curie–Weiss	2	−0.044	—	—	44.6	−363.2	0.48
2001	q-voter strong	0.039	0.15	1	1.4	183.1	−374.3	0.22
2001	q-voter weak	6.7	15.2	9.1	32.2	—	−376.9	0.37
2001	reinforcement	6.2	11.3	23.5	42.3	—	−376.3	0.22
2005	Curie–Weiss	2.1	−0.018	—	—	56.5	−316.5	0.2
2005	q-voter strong	0.052	0.13	1.1	1.3	173.7	−325.2	0.17
2005	q-voter weak	7.9	16.6	12.2	37.1	—	−328.7	0.21
2005	reinforcement	7.1	11	28.9	44.9	—	−327.5	0.18
2010	Curie–Weiss	2.1	0.021	—	—	39.4	−211.1	0.061
2010	q-voter strong	0.0093	0.0073	1	1	709.6	−218.4	0.15
2010	q-voter weak	7.7	6.1	15.6	11.9	—	−221	0.12
2010	reinforcement	6.2	4.9	25	20.4	—	−220.1	0.17
2015	Curie–Weiss	2.2	0.0076	—	—	83.4	−307.9	0.71
2015	q-voter strong	1.7	1.6	4.9	4.4	86.9	−304.2	0.71
2015	q-voter weak	13.8	14.1	30.8	32.4	—	−307.5	0.79
2015	reinforcement	10.5	10.4	49.9	50.8	—	−307.3	0.79
2017	Curie–Weiss	2	0.0039	—	—	76	−416.3	0.47
2017	q-voter strong	4.9	5.1	10.8	11.6	77	−412.5	0.5
2017	q-voter weak	8.7	17.8	10.9	38.8	—	−430.6	0.45
2017	reinforcement	8.2	14.4	29.4	57.7	—	−432.4	0.61
2019	Curie–Weiss	2.1	0.036	—	—	45.5	−314.8	0.77
2019	q-voter strong	0.043	0.065	1.1	1.2	143.9	−314.2	0.57
2019	q-voter weak	6.3	7.6	9.2	15.5	—	−317.5	0.8
2019	reinforcement	5.3	5.9	17.1	22.9	—	−316.6	0.77

Figure A2. Comparison of the models' AICs for the UK.

Appendix A.3. France Data

This is the second round of the presidential election, and we considered the vote share of the winning candidate. The performance of the AICs for the models is shown in Figure A3.

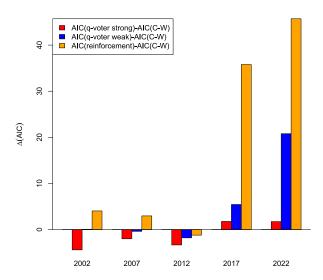| Year	Model	$J/\eta^+/N^+$	$h/\eta^-/N^-$	$\lambda_+/\kappa_+/\vartheta_+$	$\lambda_-/\kappa_-/\vartheta_-$	N	AIC	p_{KS}
2002	Curie–Weiss	0.95	0.98	—	—	58	−1672.5	0.11
2002	q-voter strong	35.5	6	14	2	47	−1668.1	0.099
2002	q-voter weak	23.5	5	−56.5	−7	—	−1672.5	0.16
2002	reinforcement	533	9	1275.5	−452.7	—	−1676.5	0.32
2007	Curie–Weiss	0.097	0.11	—	—	41.8	−1297.7	0.58
2007	q-voter strong	3.2	2.9	−280.7	−241.6	1.1	−1295.7	0.66
2007	q-voter weak	41.1	5.3×10^{-5}	56.2	−57	—	−1297.3	0.48
2007	reinforcement	23.2	6.1×10^{-5}	−0.32	−86.2	—	−1300.7	0.29
2012	Curie–Weiss	1.4×10^{-7}	0.082	—	—	29.9	−1138	0.2
2012	q-voter strong	2.8	2.6	−162.4	−145.5	1.2	−1134.7	0.23
2012	q-voter weak	0.006	0.00031	−33.9	−30.4	—	−1136.2	0.22
2012	reinforcement	0.54	0.0014	−66.8	−64.1	—	−1136.8	0.24
2017	Curie–Weiss	1.2×10^{-7}	0.74	—	—	14.3	−819.2	0.18
2017	q-voter strong	8.1	4.5	−817	−343.8	0.41	−821	0.25
2017	q-voter weak	0.064	1	−27.6	−8.8	—	−824.7	0.24
2017	reinforcement	109.6	6.3	325.2	−61.2	—	−855	0.95
2022	Curie–Weiss	1.5×10^{-9}	0.38	—	—	13.4	−718.8	0.0077
2022	q-voter strong	0.009	0.029	−206.6	−129.5	0.085	−720.5	0.003
2022	q-voter weak	85.1	0.11	192	−49.2	—	−739.6	0.097
2022	reinforcement	65.6	7.4	211.9	−17.2	—	−764.5	0.96

Figure A3. Comparison of the models' AICs for France.

Appendix A.4. German Data

We include all parties that did reach a vote share of at least 5% except the CSU, which is a local party and thus only stands for election in a few districts. Note that "die linke" is present in the parliament of 2021, though this party did not reach 5%, and in that, we did exclude this party in 2021. In order to fit our dichotomy model, we focused on the vote share of the focal party, essentially distinguishing between supporters of this party and supporters of any other party. The performance of the AICs for the models is shown in Figure A4.

Year	Model	Party	$J/\eta^+/N^+$	$h/\eta^-/N^-$	$\lambda_+/\kappa_+/\vartheta_+$	$\lambda_-/\kappa_-/\vartheta_-$	N	AIC	p_{KS}
2017	Curie–Weiss	cdu	1	−0.41	—	—	123.1	−745.2	0.81
2017	q-voter strong	cdu	22.6	48.9	19.8	57.9	123.4	−741.2	0.81
2017	q-voter weak	cdu	14	26.7	−15	−56.3	—	−742.5	0.87
2017	reinforcement	cdu	14.4	9.4×10^{-5}	−14.7	−173	—	−743.1	0.83
2017	Curie–Weiss	spd	1.3	−0.63	—	—	68.7	−814.1	0.23
2017	q-voter strong	spd	7.2	31.3	5	33.1	55.5	−809.8	0.17
2017	q-voter weak	spd	10.1	0.27	18.9	−80.6	—	−815.3	0.18
2017	reinforcement	spd	19.6	752.9	−545.5	2039.8	—	−824.8	0.4
2017	Curie–Weiss	linke	2.8	−0.11	—	—	81.9	−1068.6	3.6×10^{-8}
2017	q-voter strong	linke	0.06	20.5	0.6	82	108.6	−1071.6	2.7×10^{-7}
2017	q-voter weak	linke	9.1	2.7	42.8	9.3	—	−1112.1	5.7×10^{-7}
2017	reinforcement	linke	14.2	1800.1	−1509.8	4806	—	−1152.6	6×10^{-5}
2017	Curie–Weiss	gruene	2×10^{-6}	−2.8	—	—	19.2	−1047.3	2.6×10^{-6}
2017	q-voter strong	gruene	0.95	45.2	0.72	12.5	50.8	−1103.8	0.26
2017	q-voter weak	gruene	4.7	0.063	15.3	−164.1	—	−1113.4	0.7
2017	reinforcement	gruene	4.4	472.5	−421.4	961.3	—	−1113.8	0.71
2017	Curie–Weiss	fdp	2.5	−0.21	—	—	244.3	−1320.7	0.5
2017	q-voter strong	fdp	11.9	116.1	1.3	8.4	131.5	−1314.9	0.34
2017	q-voter weak	fdp	17.2	5.4	53.7	−308.8	—	−1320.5	0.57
2017	reinforcement	fdp	14	117.1	0.14	1.1	—	−1320.2	0.57
2017	Curie–Weiss	afd	2.6	−0.088	—	—	110	−985.3	0.0082
2017	q-voter strong	afd	1.1	23	2.2	86.7	117.5	−983.7	0.011
2017	q-voter weak	afd	14.5	8.9	54.5	23.7	—	−1009.4	0.094
2017	reinforcement	afd	22.4	1081.9	−747.5	3127.1	—	−1035	0.77
2021	Curie–Weiss	cdu	2.6×10^{-7}	−1.3	—	—	64	−778.3	0.6
2021	q-voter strong	cdu	16.6	53	−40	−215.2	30.8	−774.6	0.75
2021	q-voter weak	cdu	1.8	8.8	−31	−156.9	—	−776.6	0.76
2021	reinforcement	cdu	3.5	0.8	−60.7	−261.1	—	−776.6	0.75
2021	Curie–Weiss	spd	0.76	−0.7	—	—	72.8	−829.6	0.74
2021	q-voter strong	spd	13.1	38.5	5.3	21.2	63	−825.5	0.74
2021	q-voter weak	spd	13.1	1.2	14	−97.2	—	−828.9	0.74
2021	reinforcement	spd	38.7	827.5	−477	2376.9	—	−836.6	0.93

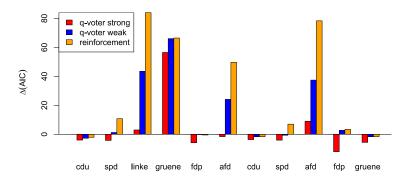| Year	Model	Party	$J/\eta^+/N^+$	$h/\eta^-/N^-$	$\lambda_+/\kappa_+/\vartheta_+$	$\lambda_-/\kappa_-/\vartheta_-$	N	AIC	p_{KS}
2021	Curie–Weiss	afd	2.7	−0.12	−	−	71.9	−971.4	1.3×10^{-5}
2021	q-voter strong	afd	0.044	14.9	0.59	56.4	122.8	−980.4	7.7×10^{-5}
2021	q-voter weak	afd	8.6	4.1	37.3	11	−	−1008.8	0.0014
2021	reinforcement	afd	13.7	1092.2	−841	3064.4	−	−1049.7	0.21
2021	Curie–Weiss	fdp	2.6	−0.032	−	−	430.3	−1424.6	0.37
2021	q-voter strong	fdp	19.7	170.3	1.5	11.5	198.3	−1412.6	0.15
2021	q-voter weak	fdp	40.9	254	74.7	703.9	−	−1427.3	0.64
2021	reinforcement	fdp	43.3	522.6	19	1933.9	−	−1428	0.79
2021	Curie–Weiss	gruene	1.8	−0.61	−	−	52.2	−813.8	0.18
2021	q-voter strong	gruene	1.2	24.7	0.98	21.2	36.2	−808.2	0.081
2021	q-voter weak	gruene	3.8	22.1	−2.1	−21.3	−	−812.1	0.13
2021	reinforcement	gruene	3.8	0.3	13.9	−87.8	−	−812.2	0.13

Figure A4. Comparison of the models' AICs for Germany.

References

1. Galam, S. Sociophysics: A review of Galam models. *Int. J. Mod. Phys. C* **2008**, *19*, 409–444. [CrossRef]
2. Mimkes, J. A thermodynamic formulation of social science. In *Econophysics and Sociophysics: Trends and Perspectives*; John Wiley & Sons: Hoboken, NJ, USA, 2006; pp. 279–309.
3. Castellano, C.; Fortunato, S.; Loreto, V. Statistical physics of social dynamics. *Rev. Mod. Phys.* **2009**, *81*, 591–646. [CrossRef]
4. Ye, M. *Opinion Dynamics and the Evolution of Social Power in Social Networks*; Springer: Berlin/Heidelberg, Germany, 2019; p. 209.
5. Sîrbu, A.; Loreto, V.; Servedio, V.D.P.; Tria, F. Opinion Dynamics: Models, Extensions and External Effects. In *Participatory Sensing, Opinions and Collective Awareness*; Springer International Publishing: Berlin/Heidelberg, Germany, 2016; pp. 363–401. [CrossRef]
6. Dong, Y.; Zhan, M.; Kou, G.; Ding, Z.; Liang, H. A survey on the fusion process in opinion dynamics. *Inf. Fusion* **2018**, *43*, 57–65. [CrossRef]
7. Anderson, B.D.O.; Ye, M. Recent Advances in the Modelling and Analysis of Opinion Dynamics on Influence Networks. *Int. J. Autom. Comput.* **2019**, *16*, 129–149. [CrossRef]
8. Bernardo, C.; Altafini, C.; Proskurnikov, A.; Vasca, F. Bounded confidence opinion dynamics: A survey. *Automatica* **2024**, *159*, 111302. [CrossRef]
9. Sobkowicz, P. Whither Now, Opinion Modelers? *Front. Phys.* **2020**, *8*, 587009. [CrossRef]
10. Chinellato, D.D.; Epstein, I.R.; Braha, D.; Bar-Yam, Y.; de Aguiar, M.A. Dynamical response of networks under external perturbations: Exact results. *J. Stat. Phys.* **2015**, *159*, 221–230. [CrossRef]
11. Costa Filho, R.N.; Almeida, M.P.; Andrade, J.S.; Moreira, J.E. Scaling behavior in a proportional voting process. *Phys. Rev. E* **1999**, *60*, 1067. [CrossRef]
12. Fortunato, S.; Castellano, C. Scaling and universality in proportional elections. *Phys. Rev. Lett.* **2007**, *99*, 138701. [CrossRef]
13. Chatterjee, A.; Mitrović, M.; Fortunato, S. Universality in voting behavior: An empirical analysis. *Sci. Rep.* **2013**, *3*, 1049. [CrossRef]
14. Kononovicius, A. Empirical analysis and agent-based modeling of the Lithuanian parliamentary elections. *Complexity* **2017**, *2017*, 7354642. [CrossRef]
15. Hösel, V.; Müller, J.; Tellier, A. Universality of neutral models: Decision process in politics. *Palgrave Commun.* **2019**, *5*, 14. [CrossRef]
16. Braha, D.; de Aguiar, M.A. Voting contagion: Modeling and analysis of a century of US presidential elections. *PLoS ONE* **2017**, *12*, e0177970. [CrossRef] [PubMed]
17. Fernández-Gracia, J.; Suchecki, K.; Ramasco, J.J.; San Miguel, M.; Eguíluz, V.M. Is the voter model a model for voters? *Phys. Rev. Lett.* **2014**, *112*, 158701. [CrossRef]
18. Andersson, H.; Britton, T. *Stochastic Epidemic Models and Their Statistical Analysis*; Springer: New York, NY, USA, 2000. [CrossRef]
19. Diekmann, O.; Heesterbeek, H.; Britton, T. *Mathematical Tools for Understanding Infectious Disease Dynamics*; Princeton University Press: Princeton, NJ, USA, 2012. [CrossRef]

20. Funk, S.; Salathé, M.; Jansen, V.A.A. Modelling the influence of human behaviour on the spread of infectious diseases: A review. *J. R. Soc. Interface* **2010**, *7*, 1247–1256. [CrossRef] [PubMed]

21. Müller, J.; Tellier, A.; Kurschilgen, M. Echo chambers and opinion dynamics explain the occurrence of vaccination hesitancy. *R. Soc. Open Sci.* **2022**, *9*, 220367. [CrossRef]

22. Liggett, T. *Interacting Particle Systems*; Springer: Berlin/Heidelberg, Germany, 1985.

23. Mobilia, M. Does a single zealot affect an infinite group of voters? *Phys. Rev. Lett.* **2003**, *91*, 028701. [CrossRef]

24. Palombi, F.; Toti, S. Stochastic Dynamics of the Multi-State Voter Model Over a Network Based on Interacting Cliques and Zealot Candidates. *J. Stat. Phys.* **2014**, *156*, 336–367. [CrossRef]

25. Krapivsky, P.L.; Redner, S.; Ben-Naim, E. *A Kinetic View of Statistical Physics*; Cambridge University Press: Cambridge, UK, 2017.

26. Nicolao, L.; Ostilli, M. Critical states in Political Trends. How much reliable is a poll on Twitter? The Potts model and the inverse problem in Social Science. *Phys. A Stat. Mech. Its Appl.* **2019**, *533*, 121920. [CrossRef]

27. Castellano, C.; Muñoz, M.A.; Pastor-Satorras, R. Nonlinear q-voter model. *Phys. Rev. E* **2009**, *80*, 041129. [CrossRef]

28. Mobilia, M. Nonlinear q-voter model with inflexible zealots. *Phys. Rev. E* **2015**, *92*, 012803. [CrossRef] [PubMed]

29. Nyczka, P.; Sznajd-Weron, K.; Cisło, J. Phase transitions in the q-voter model with two types of stochastic driving. *Phys. Rev. E* **2012**, *86*, 011105. [CrossRef] [PubMed]

30. Moran, P. Random processes in genetics. *Math. Proc. Camb. Philos. Soc.* **1958**, *54*, 60–71. [CrossRef]

31. Hösel, V.; Kuttler, C.; Müller, J. *Mathematical Population Genetics and Evolution of Bacterial Cooperation*; World Scientific: Singapore, 2020. [CrossRef]

32. Cannings, C. The latent roots of certain Markov chains arising in genetics: A new approach, I. Haploid models. *Adv. Appl. Probab.* **1974**, *6*, 260–290. [CrossRef]

33. Ewens, W.J. *Mathematical Population Genetics*; Springer: New York, NY, USA, 2004. [CrossRef]

34. Etheridge, A. *Some Mathematical Models from Population Genetics*; LNM 2012; Springer: Berlin/Heidelberg, Germany, 2011.

35. Tavaré, S. Ancestral Inference in Population Genetics. In *Proceedings of the Lectures on Probability Theory and Statistics*; Picard, J., Ed.; Lecture Notes in Mathematics 1837; Springer: Berlin/Heidelberg, Germany, 2004; pp. 3–180.

36. Kononovicius, A.; Ruseckas, J. Continuous transition from the extensive to the non-extensive statistics in an agent-based herding model. *Eur. Phys. J. B* **2014**, *87*, 169. [CrossRef]

37. Alfarano, S.; Milaković, M. Network structure and N-dependence in agent-based herding models. *J. Econ. Dyn. Control* **2009**, *33*, 78–92. [CrossRef]

38. Gardiner, C.W. *Stochastic Methods*; Springer: Berlin/Heidelberg, Germany, 2009; p. 447.

39. Hague, R.; McCormick, J.; Harrop, M. *Comparative Government and Politics An Introduction*; Palgrave: London, UK, 2016.

40. Reynolds, A. *Electoral System Design*; International Institute for Democracy and Electoral Assistance: Stockholm, Sweden, 2005.

41. Galam, S. The Trump phenomenon: An explanation from sociophysics. *Int. J. Mod. Phys. B* **2017**, *31*, 1742015. [CrossRef]

42. Binder, S. The Dysfunctional Congress. *Annu. Rev. Political Sci.* **2015**, *18*, 85–101. [CrossRef]

43. Pew Research Center, August 2022, "As Partisan Hostility Grows, Signs of Frustration with the Two-Party System". Available online: https://www.pewresearch.org/politics/2022/08/09/as-partisan-hostility-grows-signs-of-frustration-with-the-two-party-system/ (accessed on 8 January 2024).

44. Prosser, C. The end of the EU affair: The UK general election of 2019. *West Eur. Polit.* **2020**, *44*, 450–461. [CrossRef]

45. Gougou, F.; Persico, S. A new party system in the making? The 2017 French presidential election. *Fr. Polit.* **2017**, *15*, 303–321. [CrossRef]

46. Durovic, A. The French elections of 2017: Shaking the disease? *West Eur. Polit.* **2019**, *42*, 1487–1503. [CrossRef]

47. Durovic, A. Rising electoral fragmentation and abstention: The French elections of 2022. *West Eur. Polit.* **2022**, *46*, 614–629. [CrossRef]

48. Donovan, B. The Political Exploitation of Covid-19: The AfD as Challenger Party and the Impact on Parliament. *Z. Für Parlam.* **2021**, *52*, 824–843. [CrossRef]

49. Faas, T.; Klingelhöfer, T. German politics at the traffic light: New beginnings in the election of 2021. *West Eur. Polit.* **2022**, *45*, 1506–1521. [CrossRef]

50. Loewe, L.; Hill, W.G. The population genetics of mutations: Good, bad and indifferent. *Philos. Trans. R. Soc. B Biol. Sci.* **2010**, *365*, 1153–1167. [CrossRef] [PubMed]

51. Lipset, S.M. *Party Systems and Voter Alignments*; Free Press: Glencoe, IL, USA, 1967.

52. Dalton, R.J.; Flanagan, S.E. *Electoral Change in Advanced Industrial Democracies Realignment or Dealignment?*; Princeton University Press: Princeton, NJ, USA, 1984.

53. Karvonen, L.; Kuhnle, S. *Party Systems and Voter Alignments Revisited*; Taylor & Francis Group: Abingdon, UK, 2001.

54. Kitschelt, H.; Rehm, P. Party Alignments: Change and Continuity. In *The Politics of Advanced Capitalism*; Beramendi, P., Häusermann, S., Kitschelt, H., Kriesi, H., Eds.; Cambridge University Press: Cambridge, UK, 2015; pp. 179–201. [CrossRef]

55. Biezen, I.v.; Poguntke, T. The decline of membership-based politics. *Party Polit.* **2014**, *20*, 205–216. [CrossRef]

56. Siaroff, A. The Decline of Political Participation: An Empirical Overview of Voter Turnout and Party Membership. In *Activating the Citizen: Dilemmas of Participation in Europe and Canada*; DeBardeleben, J., Pammett, J.H., Eds.; Palgrave Macmillan UK: London, UK, 2009; pp. 41–59. [CrossRef]

57.   Clarke, H.D.; Sanders, D.; Stewart, M.C.; Whiteley, P. *Political Choice in Britain*; Oxford University Press: Oxford, UK, 2004. [CrossRef]
58.   Denver, D.; Johns, R. *Elections and Voters in Britain*; Palgrave Macmillan: Cham, Switerland, 2020. [CrossRef]
59.   Dahl, R.A. *Polyarchy. Participation and Opposition*; Yale University Press: New Haven, CT, USA, 1972.
60.   Kirman, A. Ants, Rationality, and Recruitment. *Q. J. Econ.* **1993**, *108*, 137–156. [CrossRef]
61.   Pasteels, J.; Deneubourg, J.; Goss, S.; Prigogine, I.; Sanglier, M. Transmission and amplification of information in a changing environment: The case of insect societies. In *Law of Nature and Human Conduct*; Prigogine, I., Sanglier, M., Eds.; Bruxelles: Cordes, Belgium, 1987.
62.   Popper, K.R. *The Open Society and Its Enemies*; Princeton University Press: Princeton, NJ, USA, 1966.
63.   Dryzek, J.S. *Discursive Democracy*; Cambridge University Press: Cambridge, UK, 1990.
64.   Harrington, J. *The Commonwealth of Oceana and a System of Politics*; Cambridge University Press: Cambridge, UK, 1992. [CrossRef]
65.   De Tocqueville, A.; Mansfield, H.C.; Winthrop, D. *Democracy in America*; University of Chicago Press: Chicago, IL, USA, 2000.
66.   Dewey, J. *The Public and Its Problems an Essay in Political Inquiry*; Swallow Press: Athens, OH, USA, 2016.
67.   Coser, L.A. *Continuities in the Study of Social Conflict*; Free Press: Glencoe, IL, USA, 1967.
68.   Dworkin, R.M. *Is Democracy Possible Here? Principles for a New Political Debate*; Princeton University Press: Princeton, NJ, USA, 2006.
69.   Lupu, N. Party Polarization and Mass Partisanship: A Comparative Perspective. *Polit. Behav.* **2014**, *37*, 331–356. [CrossRef]
70.   R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2023.
71.   Burnham, K.P.; Anderson, D.R. Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociol. Methods Res.* **2004**, *33*, 261–304. [CrossRef]