Supplementary Information

May 27, 2020

Contents

1 Model analysis

We perform a deterministic limit of our model, and a weak.effects limit; the deterministic limit allows for a bifurcation analysis. Accordingly, in the weak-effects limit, we find phase transitions. Moreover, the invariant measure of the weak.effects limit is used in the data analysis.
Recall that the model reads (details about the meaning of parameters are stated in the main part of the paper)

$$X_t \to X_t + 1 \quad \text{at rate} \quad \mu(N - X_t)\,\frac{\vartheta_1(X_t + N_1)}{\vartheta_1(X_t + N_1) + (N - X_t + N_2)}, \tag{1}$$

$$X_t \to X_t - 1 \quad \text{at rate} \quad \mu X_t\,\frac{\vartheta_2(N - X_t + N_2)}{(X_t + N_1) + \vartheta_2(N - X_t + N_2)}. \tag{2}$$

Only the latter case we obtain a limiting ODE. In order to better understand the consequences of the mechanism proposed, we first consider the deterministic limit.

1.1 Deterministic limit

Proposition 1.1 *Let $N_i = n_i N$. Then, the deterministic limit for $x(t) = X_t/N$ reads*

$$\dot{x} \;=\; -\mu x\,\frac{\vartheta_2(1 - x + n_2)}{(x + n_1) + \vartheta_2(1 - x + n_2)} + \mu(1 - x)\,\frac{\vartheta_1(x + n_1)}{\vartheta_1(x + n_1) + (1 - x + n_2)}. \tag{3}$$

For $n_1 = n_2 = n$ and $\vartheta_1 = \vartheta_2$, $x = 1/2$ always is a stationary point; this stationary point undergoes a pitchfork bifurcation at $\vartheta_1 = \vartheta_2 = \vartheta_p$, where

$$\vartheta_p \;=\; \frac{1 - 2n}{1 + 2n}. \tag{4}$$

Proof: The rates to increase/decrease the state can be written as $f_+(X_t/N)$ resp. $f_-(X_t/N)$, where (recall that $n_i = N_i/N$)

$$f_+(x) = \mu(1 - x)\,\frac{\vartheta_1(x + n_1)}{\vartheta_1(x + n_1) + (1 - x + n_2)}, \qquad f_-(x) = \mu x\,\frac{\vartheta_2(1 - x + n_2)}{(x + n_1) + \vartheta_2(1 - x + n_2)}.$$

Therewith, the Fokker-Planck equation for the large population size (Kramers-Moyal expansion) reads

$$\partial_t u(x,t) = -\partial_x((f_+(x) - f_-(x))\,u(x,t)) + \frac{1}{2N}\partial_x^2((f_+(x) + f_-(x))\,u(x,t))$$

and the ODE due to the drift term in case of $N \to \infty$ is given by

$$\frac{d}{dt}x = f_+(x) - f_-(x).$$

This result establishes eqn. (3). For the following, let $\vartheta_1 = \vartheta_2 = \vartheta$. If we also choose $n_1 = n_2 = n$, we have a neutral model, and $x = 1/2$ is a stationary point for all $\vartheta \geq 0$, $n > 0$. We find the Taylor expansion of the r.h.s. at $x = 1/2$ (using the computer algebra package maxima [1])

$$\begin{aligned}
\mu^{-1}\frac{d}{dt}x &= -x\,\frac{\vartheta(1 - x + n_2)}{(x + n_1) + \vartheta(1 - x + n_2)} + (1 - x)\,\frac{\vartheta(x + n_1)}{\vartheta(x + n_1) + (1 - x + n_2)} \\
&= -2\,\vartheta\,\frac{(2\,n + 1)\,\vartheta + (2\,n - 1)}{(2\,n + 1)\,(\vartheta + 1)^2}\left(x - \frac{1}{2}\right) + \frac{32\,\vartheta\,(\vartheta + n - \vartheta^2(n + 1))}{(2\,n + 1)^3\,(\vartheta + 1)^4}\left(x - \frac{1}{2}\right)^3 + \mathcal{O}((x - 1/2)^4)
\end{aligned}$$

For $\vartheta \in (0,1)$, $n > 0$, the coefficient in front of the third order term always is non-zero, while the coefficient in front of the linear term becomes zero at $\vartheta = \vartheta_p$. Hence, we have a pitchfork bifurcation at that parameter.

□

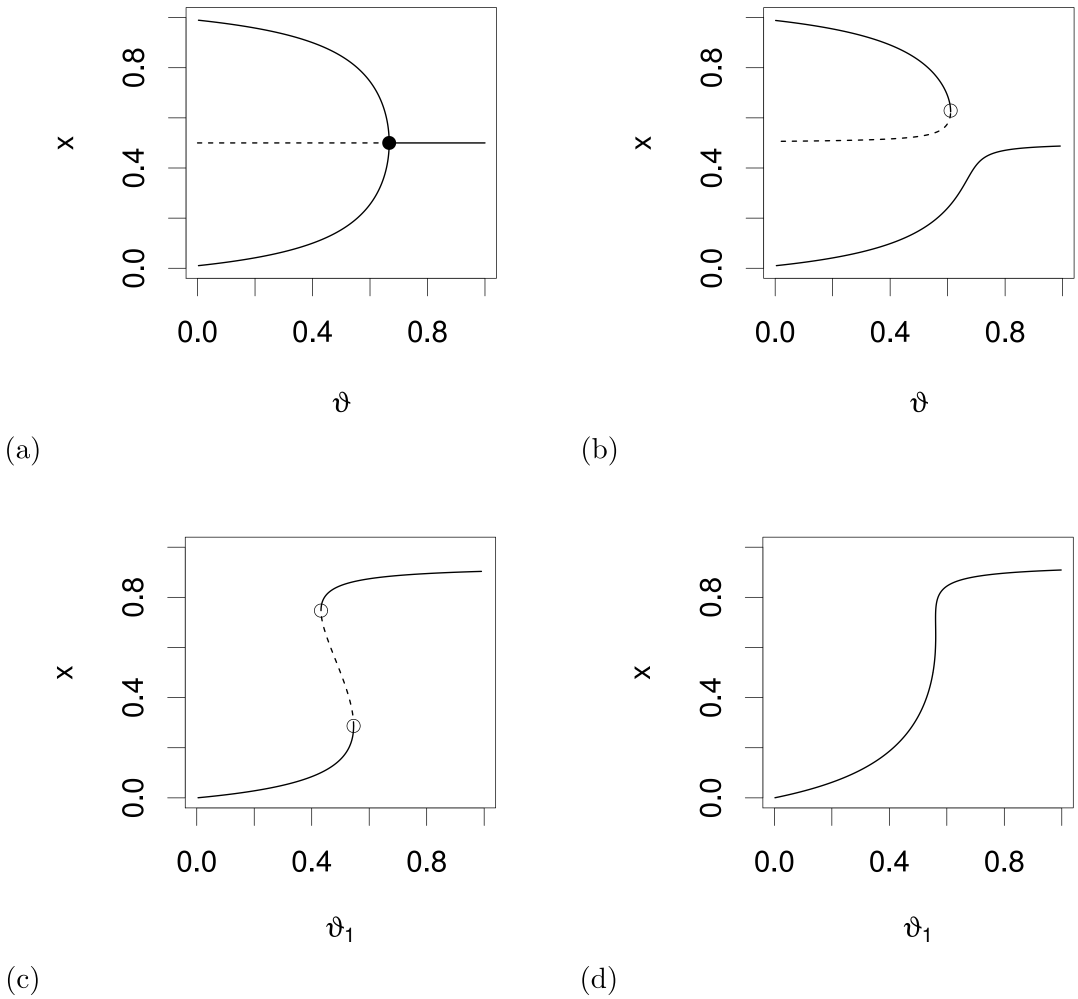

Figure 1: Stationary points of the reinforcement model over ϑ. The pitchfork bifurcation in (a) is indicated by a bullet, the saddle-node bifurcations in (b) and (c) are indicated by open circles. Stable branches of stationary points are represented by solid lines, unstable branches by dotted lines. (a) $n_1 = n_2 = 0.1$, $\vartheta_1 = \vartheta_2 = \vartheta$, (b) $n_1 = 0.1$, $n_2 = 0.105$, $\vartheta_1 = \vartheta_2 = \vartheta$, (c) $n_1 = n_2 = 0.1$, $\vartheta_2 = 0.5$, (d) $n_1 = 0.2$, $n_2 = 0.02$, $\vartheta_1 = 1.0$.

The pitchfork bifurcation is unstable against any perturbation that breaks the symmetry $x \mapsto 1-x$ (Fig. 1). In panel (a), we have the symmetric case, and find the proper pitchfork bifurcation.

Panel (b) shows the result if the number of zealots only differs slightly, where the reinforcement-parameter for both groups are assumed to be identical. We still find a reminiscent of the pitchfork bifurcation: The stable branches in (b) are close to the stable branches in (a), and also the unstable branches correspond to each other. For the limit $n_2 \to n_1$, panel (b) converges to panel (a). However, the branches are not connected any more but dissolve in two unconnected parts, and the pitchfork bifurcation is replaced by a saddle-node bifurcation.
In panel (c) and (d), the upper branch visible in panel (b) did vanish, and only the lower branch is present. As ϑ_2 is kept constant ($\vartheta_2 = 0.5$ in panel (c) and $\vartheta_2 = 0$ in panel (d)) and only ϑ_1 does vary, there is no continuous transition to panel (a).

The effect of reinforcement for a given group resembles an increase in the number of the group's zealot. Reinforcement may lead to the dominance of a group. In panel (d), the second group has only 1/10 of the zealots of the first group, but is able to take over if the members of that group do an extreme reinforcement ($\vartheta_1 \ll 1$). However, if the reinforcement of both groups is has a similar intensity and is strong, the mechanism is symmetrical, with a bistable setting as the consequence (panel (a)).

1.2 Weak effects limit

We now turn to the second scaling – the effect of zealots, and also the effect of the echo chambers, are taken to be weak. Under these circumstances, it is possible to find a limiting distribution for the invariant measure of the process.

Theorem 1.2 *Let N_i denote the number of zealots for group i, N the population size, and $\vartheta_i = 1 - \theta_i/N$ the parameter describing reinforcement. In the limit $N \to \infty$, the density of the invariant measure for the random variable $z_t = X_t/N$ is given by*

$$\varphi(x) = C\, e^{\frac{1}{2}(\theta_1+\theta_2)x^2 - \theta_1\, x}\; x^{N_1-1}\,(1-x)^{N_2-1}, \tag{5}$$

where C is determined by the condition $\int_0^1 \varphi(x)\,dx = 1$.

Proof: We again start off with the Fokker-Planck equation, obtained by the Kramers-Moyal expansion, where we use the scaling $n_i = N_i/N$, and ϑ_i constant in N. Only afterwards, we proceed to the desired scaling.
As seen above, the rates to increase/decrease the state can be written as $f_+(X_t/N)$ resp. $f_-(X_t/N)$, where (recall that $n_i = N_i/N$)

$$f_+(x) = \mu(1-x)\,\frac{\vartheta_1(x+n_1)}{\vartheta_1(x+n_1) + (1-x+n_2)}, \qquad f_-(x) = \mu x\,\frac{\vartheta_2(1-x+n_2)}{(x+n_1) + \vartheta_2(1-x+n_2)}.$$

Therewith, the limiting Fokker-Planck equation reads

$$\partial_t u(x,t) = -\partial_x((f_+(x) - f_-(x))\,u(x,t)) + \frac{1}{2N}\partial_x^2((f_+(x) + f_-(x))\,u(x,t))$$

Now we rewrite drift and noise term with the new scaling $n_i = N_i/N$, $\vartheta_i = 1 - \theta_i/N$, where we

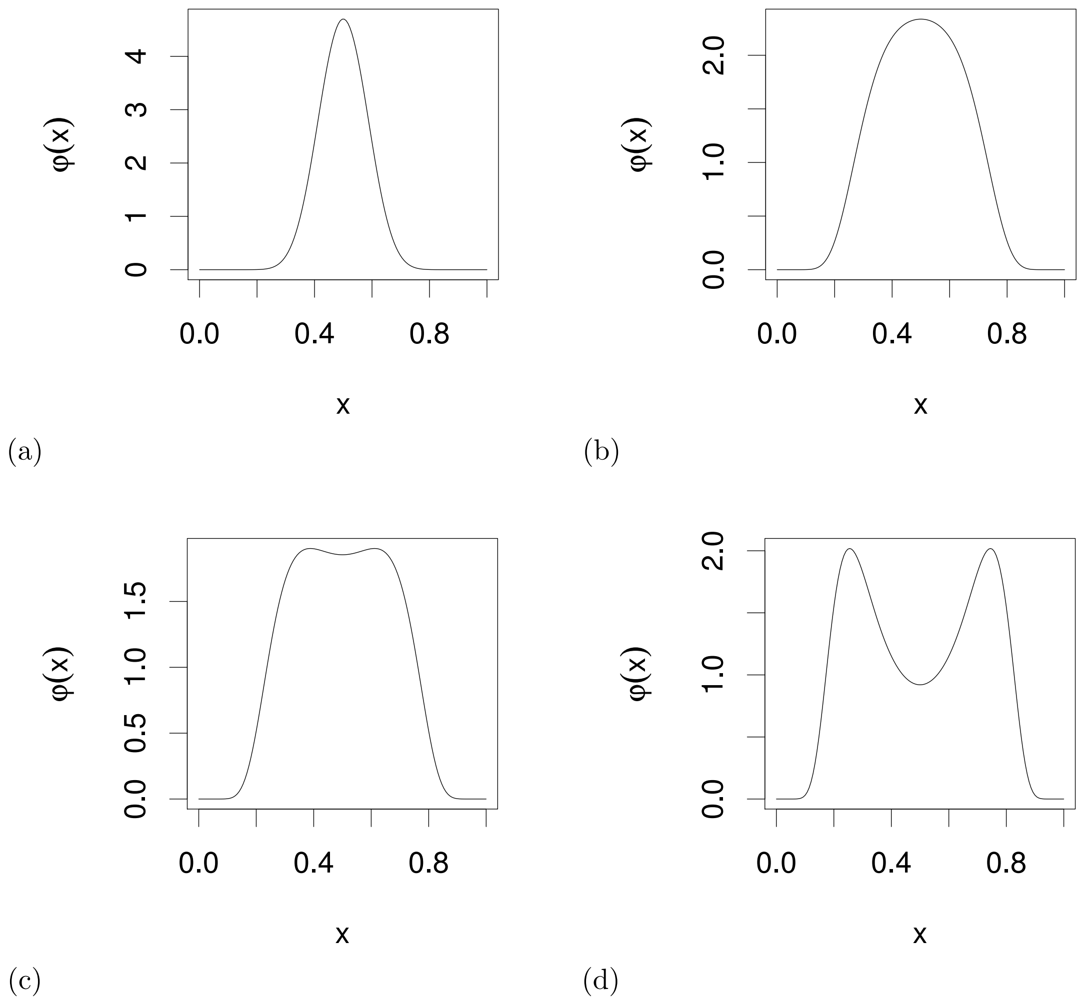

Figure 2: Invariant distribution, given in eqn. (5) for $N_1 = N_2 = 20$. We have $\theta_1 = \theta_2$, where (a) $\theta_1 = \theta_2 = 10$, (b) $\theta_1 = \theta_2 = 70$, (c) $\theta_1 = \theta_2 = 80$, (d) $\theta_1 = \theta_2 = 100$.

neglect terms of order $\mathcal{O}(N^{-2})$. We find (using maxima [1]) that ($h := 1/N$)

$$
\begin{aligned}
& f_+(x) - f_-(x) \\
= \;& \mu(1-x)\,\frac{(1-h\,\theta_1)(x+h\,N_1)}{(1-h\,\theta_1)(x+h\,N_1)+(1-x+h\,N_2)} - \mu x\,\frac{(1-h\,\theta_2)(1-x+h\,N_2)}{(x+h\,N_1)+(1-h\,\theta_2)(1-x+h\,N_2)} \\
= \;& \mu\Big([(\theta_1+\theta_2)x-\theta_1]\,x\,(1-x) - (N_1+N_2)\,x + N_1\Big)\,h + \mathcal{O}(h^2),
\end{aligned}
$$

while $h(f_+(x)+f_-(x)) = h\,2\,\mu x(1-x) + \mathcal{O}(h^2)$. If we rescale time, $T = \mu\,h\,t$, the Fokker-Planck

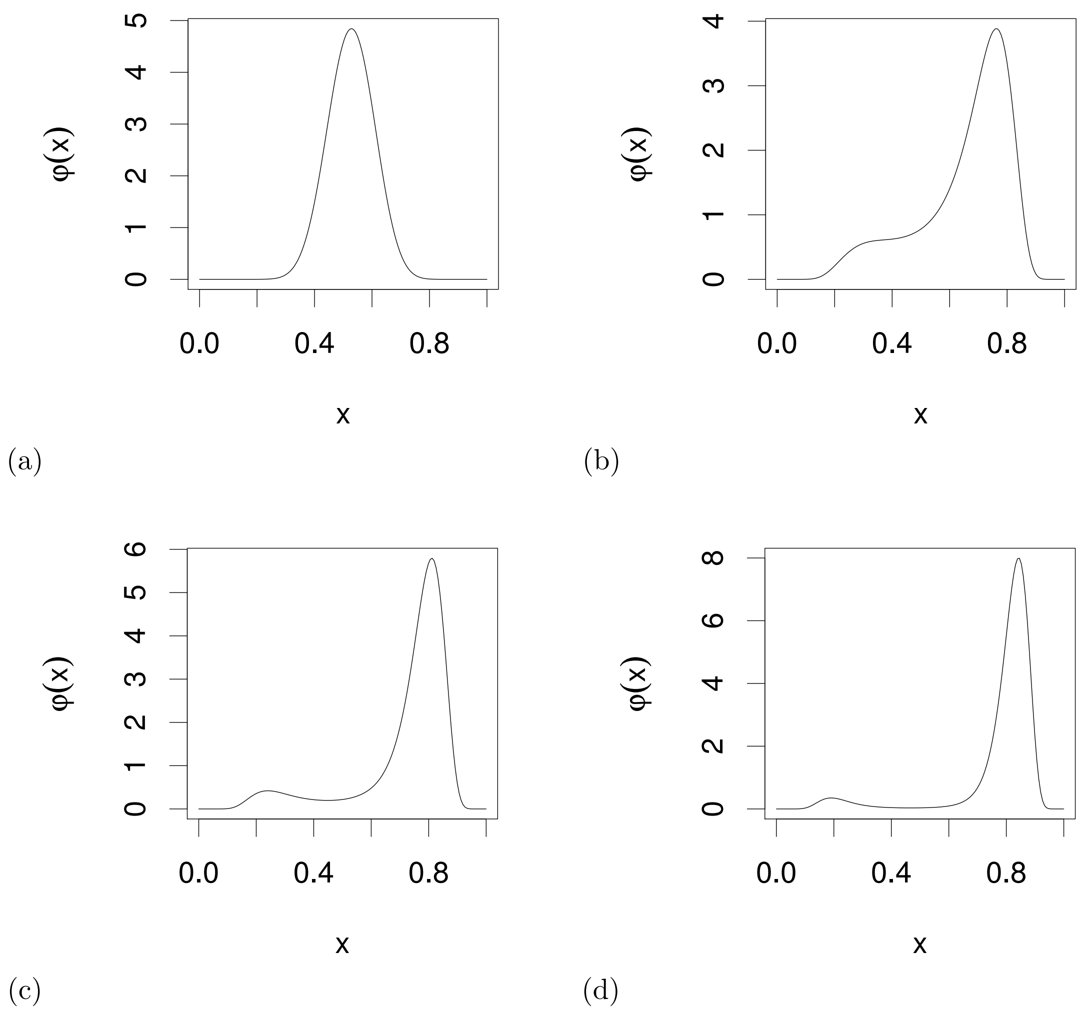

Figure 3: Invariant distribution, given in eqn. (5) for $N_1 = 22$, $N_2 = 20$. We have $\theta_1 = \theta_2$, where (a) $\theta_1 = \theta_2 = 10$, (b) $\theta_1 = \theta_2 = 100$, (c) $\theta_1 = \theta_2 = 120$, (d) $\theta_1 = \theta_2 = 140$.

equation becomes

$$\partial_T u(x,T) = -\,\partial_x\Big\{\,\Big([(\theta_1{+}\theta_2)x{-}\theta_1]\,x\,(1{-}x){-}(N_1{+}N_2)\,x{+}N_1\Big)\;u(x,T)\;\Big\}{+}\partial_x^2\Big\{x\,(1{-}x)\,u(x,T)\Big\}.$$

For the invariant distribution $\varphi(x)$, the flux of that rescaled Fokker-Planck equation is zero, that is,

$$-\Big([(\theta_1+\theta_2)x-\theta_1]\,x\,(1-x)-(N_1+N_2)\,x+N_1\Big)\varphi(x)+\frac{d}{dx}\Big(x(1-x)\,\varphi(x)\Big)=0.$$

With $v(x) = x(1-x)\varphi(x)$, we have

$$v'(x) = \left([(\theta_1 + \theta_2)x - \theta_1] + \frac{N_1}{x} - \frac{N_2}{1-x}\,x\right)v(x)$$

and hence

$$v(x) = C\,e^{\frac{1}{2}(\theta_1+\theta_2)x^2 - \theta_1\,x}\;x^{N_1}\,(1-x)^{N_2}$$

resp.

$$\varphi(x) = C\,e^{\frac{1}{2}(\theta_1+\theta_2)x^2 - \theta_1\,x}\;x^{N_1-1}\,(1-x)^{N_2-1}$$

□

For $\theta_1 = \theta_2 = 0$, we obtain the beta distribution, as we fall back to the zealot model without reinforcement. In the given scaling, the reinforcement is expressed by the exponential multiplicative factor. As $\vartheta_i = 1 - h\,\theta_i$, and $h = 1/N$ is small, one could be tempted to assume that we are in the subcritical parameter range of the reinforcement model only, s.t. the distribution does not show a phase transition. As we see next, this idea is wrong.

Let us first consider the symmetric case, $N_1 = N_2 = \underline{N}$, and $\theta_1 = \theta_2 = \underline{\theta}$ (see Fig. 2). In that case, the distribution is given by

$$\varphi(x) = C\,e^{-\underline{\theta}\,x\,(1-x)}\;x^{\underline{N}-1}\,(1-x)^{\underline{N}-1}.$$

The function always is symmetric w.r.t. $x = 1/2$. If $\underline{\theta}$ is small, and $\underline{N} > 0$, we find an unimodal function, with a maximum at $1/2$. If, however, $\hat{\theta}$ is increased, eventually a bimodal distribution appears – we find back the pitchfork bifurcation that we already known from the deterministic limit of the model (Fig. 1, panel a).

As soon as $N_1 \neq N_2$, the symmetry is broken (Fig. 3), and we have an a situation resembling Fig. 1, panel (b). In the stochastic setting, however, we have more information: the second branch concentrates only little probability mass, and will play in practice only a minor role (if any at all). Only if N_1 and N_2 are in a similar range, or the dissimilarity is balanced by appropriate reinforcement parameters, this second branch is able to concentrate sufficient probability mass to gain visibility in empirical data.

Comparison of the reinforcement model and the zealot model. We can use the zealot model or we can use the reinforcement model to fit and interpret election data. The zealot model for two parties yields the beta distribution. The density of the reinforcement model basically consist of a product, where one term is identical with the beta distribution,

$$x^{N_1-1}\,(1-x)^{N_2-1}$$

while the second term expresses the influence of reinforcement

$$e^{\frac{1}{2}(\theta_1+\theta_2)x^2 - \theta_1\,x}.$$

Only if the data have a shape that is different from that of a beta distribution, the reinforcement component leads to a significantly improved fit. This is given, e.g., in case of a bimodal shape of the data (where at least one maximum is in the interior of the interval $(0, 1)$), or if the data have heavy tails. Both properties hint to the fact that the election districts are of two different types: one, where the party under consideration is relatively strong, and one where it is relatively weak. This difference, in turn, can be interpreted as the effect of reinforcement: In some election districts voters agree that the given party is preferable, in others they agree that the party is to avoid. The population is not (spatially) homogeneous, but some segregation - most likely caused by social mechanisms - take place. In that, the data analysis of spatially structured election data (results structured by election districts) based on the reinforcement model is able to detect spatial segregation and the consequences thereof.

2 Data analysis

2.1 Methods

Parameter estimation: We aim at a maximum-likelihood estimation of the parameters of the distribution

$$\varphi(x) = C\, e^{\frac{1}{2}(\theta_1+\theta_2)x^2-\theta_1\, x}\; x^{N_1-1}\,(1-x)^{N_2-1},$$

where C is determined by $\int_0^1 \varphi(x)\,dx = 1$. The maximum likelihood parameter estimation is somewhat subtle as the distribution incorporates exponential terms - in particular, if the parameters become large, the integral $\int_0^1 e^{\frac{1}{2}(\theta_1+\theta_2)x^2-\theta_1\, x}\; x^{N_1-1}\,(1-x)^{N_2-1}\,dx$ becomes numerically unstable. Therefore, we re-parameterize the distribution, defining $\hat\nu$, $\hat s$, $\hat\theta$, and $\hat\psi$ by

$$\theta_1 = \hat s\,\hat\theta\,\hat\psi, \quad \theta_2 = \hat s\,\hat\theta\,(1-\hat\psi), \quad N_1+1 = \hat s\,(1-\hat\theta)\,\hat\nu, \quad N_2+1 = \hat s\,(1-\hat\theta)\,(1-\hat\nu), \tag{6}$$

where $\hat\theta, \hat\psi, \hat\nu \in [0,1]$, and $\hat s > 0$, with the restriction $\hat s\,(1-\hat\theta)\,\hat\nu > 1$, and $\hat s\,(1-\hat\theta)\,(1-\hat\nu) > 1$. Therewith, the distribution becomes

$$\varphi(x) = \hat C\,\exp\left[\;\hat s\;\left(\hat\theta(x^2/2-\hat\psi\, x)\; + (1-\hat\theta)\,\hat\nu\,\ln(x) + (1-\hat\theta)\,(1-\hat\nu)\,\ln(1-x) + A\right)\;\right].$$

Here, A is a constant that can be chosen in dependency on the data at hand. In practice, it is used to avoid an exponent that has a very large absolute number. The constant $\hat C$ is, as before, determined by the fact that the integral is one. This form allows for a reasonable maximum likelihood estimation, given appropriate election data.

Numerical issues: The model assumes continuous data, while the election data are discrete. Therefore, a vote share of 0 or 1 is possible in the empirical data, but the distribution may have poles for those values. We replace all empirical vote shares below 10^{-5} by 10^{-5}, and similarly, all data above $1-10^{-5}$ by $1-10^{-5}$. In order to determine the normalization constant of $\varphi(x)$, we do not integrate from 0 to 1, but only from 0.001 to 0.999. Furthermore, for numerical reasons, we restrict $\hat s$ by an upper limit, that we mostly define as 1800.

Test for reinforcement: The zealot model and the reinforcement model are nested. In that, we can use the likelihood-ratio test to check for the significance of the reinforcement component: If $\mathcal{LL}_0$ is the log-likelihood for the restricted model ($\theta_1 = \theta_2 = 0$, resp. $\hat\theta = 0$), and $\mathcal{LL}$ is that for the full reinforcement model, we have asymptotically, for a large sample size

$$2(\mathcal{LL} - \mathcal{LL}_0) \sim \chi^2_2$$

That is, twice the difference in the log-likelihoods is asymptotically χ^2 distributed, where the degree of freedom is the number of the surplus parameters (here: θ_1 and θ_2, resp. $\hat{\theta}$ and $\hat{\psi}$, that is, the degree of freedom is 2).
Additionally, we use the Kolmogorov-Smirnov test to find out if the model-distribution of either the full reinforcement mode, or the zealot model ($\theta_1 = \theta_2 = 0$, resp. $\hat{\theta} = 0$) agrees with the empirical distribution. If both models are in line with the data, then the reinforcement component will rather not add to the interpretation of the data, if only the reinforcement model approximates the data well (or, at least, much better than the zealot model), we can expect that it is sensible to take the reinforcement component into account.

2.2 Details – US

In the US, the candidates for the presidential elections are determined by the "primary elections". We do not consider them, but only the presidential elections themselves. The election of the president happens indirectly via an Electoral Collage. Each state nominates a certain number of delegates. In most states, a winner-take-all system is established. If no candidate receives the majority of the votes, he Congress will elect a candidate.

The data set used is provided by the Havard Univ., and contains the data on county-level for the elections 2000-2016,
`https://doi.org/10.7910/DVN/VOQCHQ`, file `countypres_2000-2016.csv`.

year	party	$\hat{\nu}$	$\hat{\theta}$	$\hat{\psi}$	$\hat{s}$	θ_2	p_{ll}	p_{ks} (Reinf)	p_{ks} (beta)
2000	green	0.035	0.920	0.33	196.5	120.9	<1e-10	0.043	0.0011
2000	republicans	0.515	0.323	6.61e-05	21.6	6.97	0.86	0.02	0.02
2000	democrats	0.385	0.0272	0.99993	14.0	2.51e-05	1	0.007	0.007
2004	republicans	0.485	0.4971	6.61e-05	26.6	13.2	0.003	0.20	0.17
2004	democrats	0.474	0.435	0.99993	22.9	0.00066	0.03	0.10	0.14
2008	republicans	0.399	0.588	6.61e-05	29.6	17.5	9.65e-07	0.26	0.09
2008	democrats	0.579	0.581	0.99996	27.4	0.00066	7.4e-07	0.17	0.08
2012	republicans	0.404	0.624	6.61e-05	25.3	15.8	1.5e-10	0.43	0.009
2012	democrats	0.561	0.610	0.99993	23.0	0.0009	1.6e-09	0.39	0.012
2016	republicans	0.327	0.793	0.207	60.0	37.7	<1e-10	0.88	<1e-10
2016	democrats	0.529	0.801	0.660	55.0	14.9	<1e-10	0.36	<1e-10

Table 1: Estimated parameters for the two parties in the eight elections. p_{ll} is the result of the likelihood ratio test for the significance of the reinforcement component; p_{ks} is the result of the Kolmogorov-Smirnov-test for the question of the empirical cumulative distribution differs significantly from the cumulative distribution of the model (either the reinforcement model, or zealot model with the beta distribution).

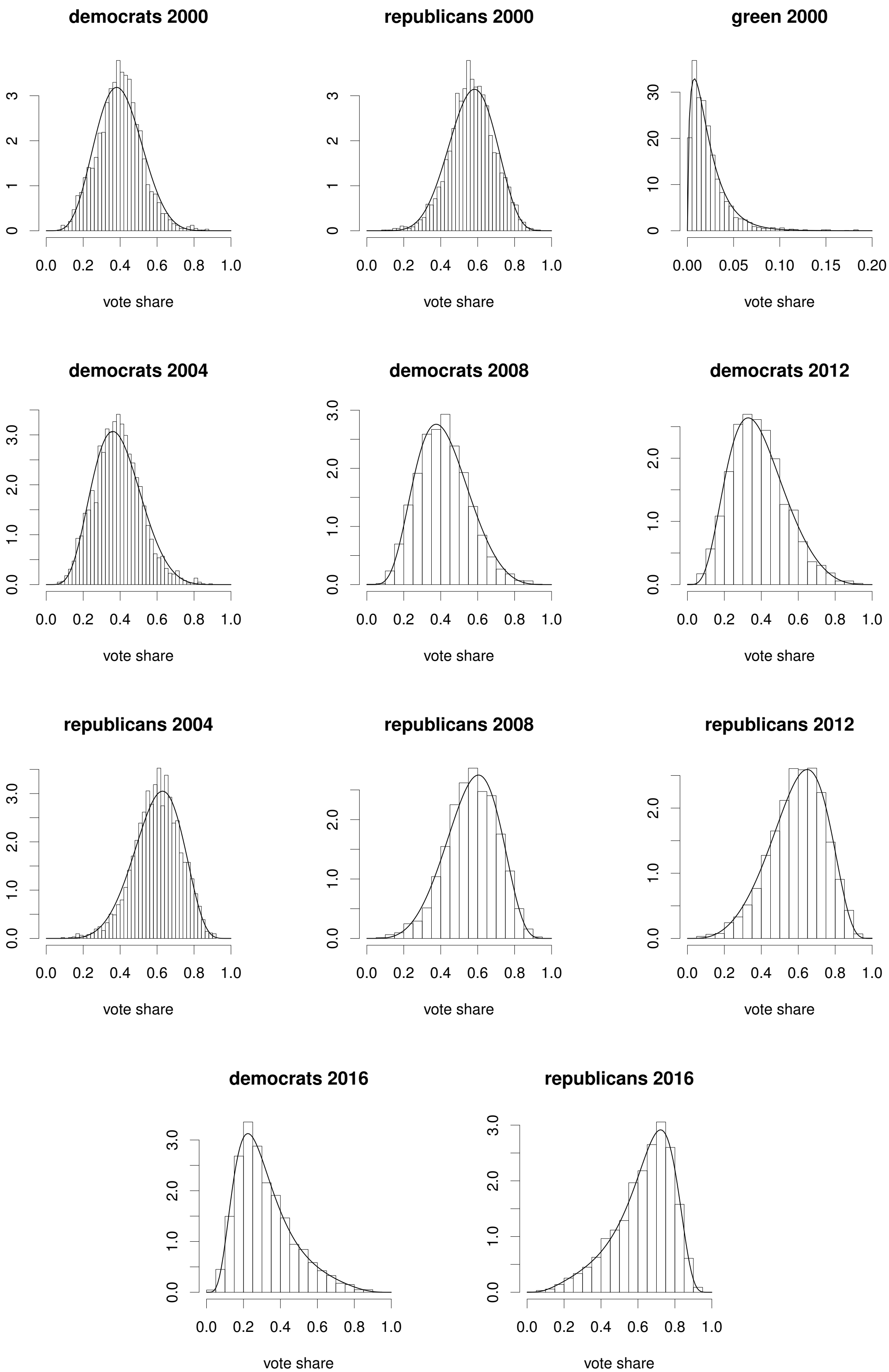

Figure 4: Distribution of the vote share of republicans/democrats in 2000-2016 presidential elections. For 2000, we also show the result for the green candidate (note the difference in the scaling of the x-axis here).

2.3 Details – Brexit

In the Brexit referendum (23 June 2016), each voter had the choice “remain” or “leave”. The election was equal, each vote was counted directly. With 51.89% (and 72.2% turnout rate), the outcome has been “leave”.

The data follow well the reinforcement model (KS, $p = 0.75$), but the zealot model does not fit nicely (KS, $p = 0.0009$). The likelihood-ratio-test on the null hypothesis that remainders and brexitiers have the same amount of reinforcement is rejected at $p <$1e-10.

The data set used is provided by the British Government, and contains data on election district level,
`https://www.electoralcommission.org.uk/who-we-are-and-what-we-do/elections-and-referendums/past-elections-and-referendums/eu-referendum/results-and-turnout-eu-referendum`, file `EU-referendum-result-data.csv`.

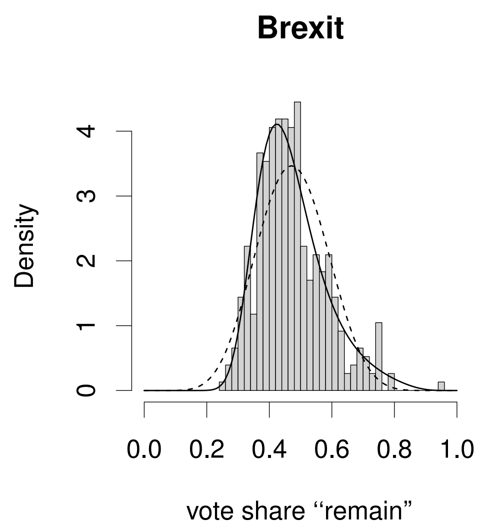

Figure 5: Distribution of the share of votes for remain in the UK Brexit election with fit of the reinforcement model (solid line) and the zealot model (dashed line).

$\hat{\nu}$	$\hat{\theta}$	$\hat{\psi}$	$\hat{s}$	θ_1	θ_2	p_{ll}	p_{ks} (Reinf.)	p_{ks} (beta)
0.87	0.76	0.99996	254.5	193.0	0.0084	<1e-10	0.75	0.0009

Table 2: Parameter for the Brexit referendum, “remainers”. For symmetry reasons, the parameters for “leave” are identical, but $\hat{\psi}_{leave} = 1 - \hat{\psi}_{remain}$, and therewith $\theta_{2,leave} = \theta_1 = 193$.

2.4 Details – Germany

Each voter has two votes in the elections for the German parliament: a "first" and a "second" vote. With the first vote, a candidate in the election district can be selected. The candidate with the most votes will be send into the parliament. With the second vote, a party is selected. The vote share of the parties determines the number of seats in the parliament, where a party has to overcome a 5% threshold. IN the analysis, only the second votes are used.

The data are provided by the German Government,
`https://www.bundeswahlleiter.de/en/bundeswahlleiter.html`, file `btw17_kerg.csv`.

party	$\hat{\nu}$	$\hat{\theta}$	$\hat{\psi}$	$\hat{s}$	θ_2	p_{ll}	p_{ks} (Reinf.)	p_{ks} (beta)
CDU	0.35	0.34	0.67	108.5	12.4	1	0.72	0.79
SPD	0.18	0.30	6.61e-05	63.5	19.0	0.99	0.19	0.20
AFD	0.07	0.77	6.61e-05	792.5	609.0	8.7e-11	0.29	0.0009
FDP	0.14	0.31	1.00	147.5	0.003	1	0.55	0.55
die linke	0.05	0.78	6.61e-05	583.5	453.7	3.4e-09	2.11e-06	2.93e-08
Gruenen	0.07	0.26	6.61e-05	69	18.2	1	0.87	0.84
CSU.	0.37	6.75e-05	6.61e-05	35.6	0.002	1	0.004	0.004

Table 3: Results for whole Germany, 2017.

party	$\hat{\nu}$	$\hat{\theta}$	$\hat{\psi}$	$\hat{s}$	θ_2	p_{ll}	p_{ks} (Reinf.)	p_{ks} (beta)
CDU	0.43	0.74	0.50	575.5	212.5	0.31	0.41	0.78
SPD	0.21	0.09	6.61e-05	53.5	4.87	1	0.17	0.17
AFD	0.13	0.29	1.00	153.5	0.003	1	0.92	0.92
FDP	0.13	0.19	0.88	198.5	4.59	0.91	0.68	0.68
die linke	0.09	0.83	0.14	1138.5	814.5	0.001	0.0007	0.0002
Gruenen	0.13	0.81	0.23	378.5	237.6	0.15	0.12	0.075
CSU	0.37	6.75e-05	6.61e-05	35.6	0.002	1	0.004	0.004

Table 4: Results for Germany, only the "old" states ,2017.

2.5 Details – France/Le Pen

The presidential elections in France have (in principle/mostly) two rounds, a first round and runoff elections, in case that no candidate receives more than 50% of votes in the first round. The elections are direct, each vote counts the same.
The data are provided by the France Government. We used data on the level of départements. `https://www.data.gouv.fr/fr/posts/les-donnees-des-elections/`.

year	candid.	$\hat{\nu}$	$\hat{\theta}$	$\hat{\psi}$	$\hat{s}$	θ_2	p_{ll}	p_{ks} (Reinf.)	p_{ks} (beta)
1965	MITTERRAND (CIR)	0.29621	0.00006	0.00007	16.70000	0.00092	1.00000	0.00264	0.00264
1965	LECANUET (MRP)	0.14076	0.20379	0.00007	64.50000	13.14346	1.00000	0.90755	0.91032
1965	DE GAULLE (UNR)	0.66489	0.78879	0.69318	873.50000	211.39934	0.00000	0.12579	0.00091
1965	TIXIER-VIGNANCOUR (EXD)	0.03379	0.79450	0.00007	700.50000	556.51034	0.00000	0.00258	0.00004
1969	DUCLOS (PCF)	0.17985	0.00005	0.00007	18.10000	0.00090	1.00000	0.10169	0.10171
1969	DEFFERRE (SFIO)	0.18107	0.91352	0.35652	473.50000	278.33585	0.00000	0.02718	0.00083
1969	POMPIDOU (UDR)	0.87801	0.76567	0.99995	446.50000	0.01612	0.00000	0.25179	0.00015
1969	POHER (CD)	0.22691	0.00008	0.00007	70.00000	0.00550	1.00000	0.12838	0.12844
1974	MITTERRAND (PS)	0.19448	0.69360	0.00007	281.50000	195.23601	0.01984	0.04230	0.00419
1974	GISCARD D'ESTAING (RI)	0.31544	0.05921	0.00007	30.70000	1.81770	1.00000	0.69891	0.69903
1974	CHABAN-DELMAS (UDR)	0.12745	0.80028	0.12755	901.50000	629.42834	0.00000	0.00108	0.00000
1974	ROYER (DVD)	0.02110	0.80942	0.00007	979.50000	792.77773	0.00000	0.00000	0.00000
1981	MARCHAIS (PCF)	0.12149	0.54736	0.13981	62.00000	29.19165	0.80061	0.25322	0.33004
1981	MITTERRAND (PS)	0.28329	0.16863	0.99993	140.50000	0.00157	1.00000	0.17275	0.17276
1981	GISCARD D'ESTAING (UDF)	0.32811	0.46247	0.61799	104.50000	18.46186	0.99834	0.14571	0.12528
1981	CHIRAC (RPR)	0.30222	0.83611	0.35581	678.50000	365.44584	0.00000	0.11781	0.00019
1988	MITTERRAND (PS)	0.33390	0.00006	0.00007	84.00000	0.00486	1.00000	0.24234	0.24247
1988	BARRE (UDF)	0.15714	0.00004	0.00007	95.50000	0.00416	1.00000	0.25615	0.25586
1988	CHIRAC (RPR)	0.31120	0.82687	0.36834	532.50000	278.12750	0.00000	0.02471	0.00059
1988	LE PEN (FN)	0.18816	0.76069	0.29438	252.50000	135.53205	0.01609	0.87942	0.23976

year	candid.	$\hat{\nu}$	$\hat{\theta}$	$\hat{\psi}$	$\hat{s}$	θ_2	p_{ll}	p_{ks} (Reinf.)	p_{ks} (beta)
1995	JOSPIN (PS)	0.27240	0.25114	0.99993	134.50000	0.00223	1.00000	0.57881	0.56790
1995	BALLADUR (UDF)	0.17659	0.00005	0.00007	85.00000	0.00395	1.00000	0.15455	0.15375
1995	CHIRAC (RPR)	0.29723	0.82776	0.33493	997.50000	549.13604	0.00000	0.06349	0.00000
1995	LE PEN (FN)	0.13366	0.16371	0.00007	57.50000	9.41255	1.00000	0.35950	0.35765
2002	JOSPIN (PS)	0.25701	0.83780	0.31043	890.50000	514.45847	0.00000	0.59056	0.00649
2002	BAYROU (UDF)	0.05982	0.00006	0.00007	114.50000	0.00740	0.99922	0.00000	0.00000
2002	CHIRAC (UMP)	0.27697	0.82723	0.31351	1069.50000	607.35573	0.00000	0.03866	0.00000
2002	LE PEN (FN)	0.14604	0.00004	0.00007	35.30000	0.00149	1.00000	0.06609	0.06608
2007	ROYAL (PS)	0.33165	0.77302	0.38618	419.50000	199.05094	0.00014	0.13556	0.03850
2007	BAYROU (UDF)	0.17800	0.00005	0.00007	108.50000	0.00504	1.00000	0.06028	0.06028
2007	SARKOZY (UMP)	0.48899	0.81562	0.50867	1135.50000	455.03893	0.00000	0.41382	0.00342
2007	LE PEN (FN)	0.08921	0.00004	0.00007	54.00000	0.00235	1.00000	0.11541	0.11533
2012	Jean-Luc MELENCHON (FG)	0.10230	0.00004	0.00007	93.00000	0.00380	1.00000	0.00051	0.00050
2012	François HOLLANDE (PS)	0.41777	0.80629	0.45421	599.50000	263.81758	0.00000	0.28500	0.00830
2012	Nicolas SARKOZY (UMP)	0.40273	0.81085	0.44187	610.50000	276.28720	0.00000	0.29963	0.00091
2012	Marine LE PEN (FN)	0.15076	0.00005	0.00007	28.60000	0.00138	1.00000	0.01465	0.01466
2017	LE PEN	0.19953	0.00006	0.00007	23.40000	0.00140	1.00000	0.05362	0.05371
2017	MLENCHON	0.23086	0.47868	0.53214	104.50000	23.40335	1.00000	0.02407	0.01695
2017	MACRON	0.28845	0.39413	0.80474	112.50000	8.65774	0.31516	0.10480	0.03947
2017	FILLON	0.29702	0.82423	0.35947	457.50000	241.53438	0.00000	0.07550	0.00000

Table 5: Results for France - departement level.

The estimates for Le Pen from 2017 indicate that the model finds no reinforcement aspects in the data of Le Pen ($\hat{\theta} \approx 0$). Accordingly, the LL-test indicates that the zealot model performs as well as the reinforcement-model. The KS-test indicates that the data (canton level) do not follow the distribution predicted by the reinforcement model, while the data on district level are close to th reinforcement model (cannot rejected at the significance level of 0.01, but only at a significance level of 0.05). Note that the number of data on canton level ($n = 35703$) is much larger than that on departments level ($n = 2090$), which explains that even small deviations from the model distribution leads to extremely significant values in the KS-test.

election(2017)	$\hat{\nu}$	$\hat{\theta}$	$\hat{\psi}$	$\hat{s}$	θ_2	p_{ll}	p_{ks} (Reinf.)	p_{ks} (beta)
first round/canton	0.23063	0.00005	0.00007	15.90000	0.00074	1.00000	0.00000	0.00000
first round/departement	0.19953	0.00006	0.00007	23.40000	0.00140	1.00000	0.05362	0.05371
second round/canton	0.39724	0.05783	0.00007	12.40000	0.71699	1.00000	0.00000	0.00000

Table 6: Estimates for Le Pen in the presidential elections (2017), different data sets: First round and second round (canton level), first round (departement level).

2.6 Details – The Netherlands/ Catholic People's Party

The parliament's election in The Netherlands do not have a threshold, but are purely proportional. That's interesting as the effect of strategic voting will be less prominent as e.g. in Germany, where a 5% threshold is implemented. The election districts, however, have a very different size as only the total number of votes, all over the country, counts. That might disturb our assumption that all election districts are i.i.d.

The Catholic People's party did play a central role after the second world war. After 1971, it did lose importance, and eventually merged with other parties.

The model indicates that the success of the party is almost exclusively due to reinforcement ($\hat{\nu} \approx 0$), which explains the peak of the distribution at a vote share of zero (Fig. 2 in the main paper).

The data can be found at the internet-pages of the Dutch government,
`https://www.verkiezingsuitslagen.nl/verkiezingen`

year	party	$\hat{\nu}$	$\hat{\theta}$	$\hat{\psi}$	$\hat{s}$	θ_2	p_{ll}	p_{ks} (Reinf.)	p_{ks} (beta)
1946	(KVP)	6.61e-05	0.98	0.46	29.9	15.9	<1e-10	<1e-10	0
1948	(KVP)	6.61e-05	0.98	0.44	31.0	16.8	<1e-10	<1e-10	0
1952	(KVP)	6.61e-05	0.98	0.39	34.0	19.7	<1e-10	<1e-10	0
1956	(KVP)	6.61e-05	0.98	0.44	30.2	16.5	<1e-10	<1e-10	0
1959	(KVP)	6.61e-05	0.96	0.39	35.4	20.7	<1e-10	<1e-10	0
1963	(KVP)	6.61e-05	0.93	0.32	40.9	26.1	<1e-10	3.21e-06	0
1967	(KVP)	6.61e-05	0.86	0.10	71	54.5	<1e-10	5.1e-05	0
1971	(KVP)	6.61e-05	0.82	6.61e-05	95	77.5	<1e-10	0.0003	0
1972	(KVP)	6.61e-05	0.80	6.61e-05	65	52.1	<1e-10	1.26e-05	0

Table 7: Results for NL/KVP. We find a transition from bimodal to unimodal during the years.

References

[1] Maxima. Maxima, a computer algebra system. version 5.34.1, 2014.