Dynamics of Voting Propensity:
Experimental Tests of Adaptive Learning Models

Susumu Shikano* & Bernhard Kittel†

Version June 22, 2014

Abstract

This paper aims to deliver experimental evidence on the dispute between two behavioral models of electoral turnout (Bendor, Diermeier & Ting, APSR 2003; Fowler,JoP 2006). Both models share the idea that the subjects’ voting propensities are updated from their past propensities, aspirations and realized payoffs. However, they differ in the exact specification of the feedback mechanism. The first model has a strong feedback mechanism toward 50%, while the other has only moderate feedback. This difference leads to two distinct distributions of voter types: the first model generates more casual voters who vote and abstain from time to time. The latter generates more habitual voting behavior. Thus far, the latter model seemed to be better supported empirically since survey data reveal more habitual voters and abstainers than casual voters. Given that the two models differ in their propensity updating mechanism in dynamic processes, a more direct test of their assumptions as well as implications with survey data is still pending. We designed a laboratory experiment in which subjects repeatedly make turnout and voting decisions. The results from experimental data is mixed, but more supportive of the second model with habitual voters and abstainers.

1 Introduction

Electoral turnout is one of the most widely investigated phenomena in political science and is also known as a paradox for rational choice approaches based on cost benefit calculation (see for literature reviews, e.g. Aldrich, 1993; Laver, 1997; Dewan and Shepsle, 2011; Kittel and Marcinkiewicz, 2012). Criticizing previous solutions of the voting paradox which are based on pivot probabilities, Bendor, Diermeier and Ting (2003, hereafter BDT) proposed an alternative behavioural model (see also Bendor et al., 2011). Although their model is also based on a cost-benefit core, the actors do not optimize

*University of Konstanz, Department of Political Science, Fellow at the Hanse Institute for Advanced Study
†University of Vienna, Department of Economic Sociology
their decision based on subjective expectations, but adapt their voting propensity retrospectively. More specifically, actors are assumed to possess a certain aspiration level which is compared to payoffs (benefit minus cost). If an agent’s behavior in an election (vote or abstain) results in higher payoffs than her aspiration level, the agent’s corresponding propensity to either vote or abstain is reinforced, and vice versa. Simultaneously, her aspiration level is updated by current payoffs. By simulating multiple agents deciding to vote or not at sequential elections, in which their aspiration and propensity continue to be updated, BDT achieved stable turnout rate of roughly 50% on average even with substantial voting cost.

In response to this model, Fowler (2006) proposed an alternative approach with a different updating mechanism. He criticized the BDT model because of its strong tendency toward a 50% voting propensity. Because of the moderating feedback mechanism built into that model, any departure from 50% is temporary and most agents are casual voters who vote and abstain from time to time. Fowler argues, in contrast, that most real citizens are either habitual voters or habitual abstainers. For empirical evidence, he referred to data from South Bend (Huckfeldt and Sprague, 1991) which resulted in a bimodal distribution of turnout at the level of individual respondents for general elections. In order to generate a distribution with more habitual voters and habitual abstainers, Fowler proposed an alternative updating mechanism. Similar in motivation to the mechanism used by BDT, agents compare their aspiration and actual payoffs to update their voting propensity. However, the updating mechanism has a much weaker moderating feedback so that the voting propensity is not concentrated around 50%. In a simulation, Fowler showed that this mechanism in fact realizes more habitual voters and abstainers than casual voters.

Empirical support for the superiority of Fowler’s against BDT’s model, however, exists only in terms of individual turnout rates. A more direct test of the two competing propensity updating mechanisms in dynamic processes is still pending. Such a test can hardly be conducted by using conventional survey data. Therefore, we design a simple laboratory experiment in which subjects make repeated decisions whether to vote or not.

In the remainder of this paper, we proceed as follows: In the next section, we compare and rerun the simulation models corresponding to the Fowler and BDT-model as well as to our experimental setting. This is, in particular, necessary since our experimental setting allows for only a small number of subjects compared to the original simulations of Fowler and BDT. The third section introduces the experimental design in detail, the fourth section presents the experimental results. This section begins with tests of core assumptions and the implications of the model. Subsequently we look into dynamics of experimental subjects’ voting decisions. In particular, we construct a full Bayesian probability model corresponding to both theoretical models and obtain the posterior information of vote propensity dynamics using Markov-Chain-Monte-Carlo methods. The last section summarizes
and discusses the results.

2 A Comparison of Simulation Models

Both the Bendor, Diermeier and Ting (2003) and the Fowler (2006) models are characterized by aspiration-based adaptive rules (ABAR) for the agents’ decisions. The basic idea of ABAR is simple: An agent’s actions are based on her action propensities. If the agent’s payoff from the previous action is larger than her aspiration level, then the corresponding propensity increases and the actors becomes more likely to take the same action again. This is called positive feedback. If the action has an unsatisfactory result, that is, the agent’s payoff is less than her aspiration level, then the corresponding action propensity decreases and the agents becomes less likely to take the same action again (negative feedback). Besides these two feedback mechanisms, the third important element of both models is the adjustment of aspirations. An agent’s aspiration level is constantly adjusted by its past level and the payoffs of the previous period. If an agent receives higher payoffs than her aspiration level, then her aspiration level increases, and vice versa.

The BDT and Fowler models only differ in their feedback mechanisms. To formalize, we use the following notation:

- $I_{i,t}$: citizen $i$’s action in period $t$ (vote=1; abstain=0)
- $p_{i,t}(I)$: citizen $i$’s propensity to take action $V$ in period $t$.
- $a_{i,t}$: citizen $i$’s aspiration in period $t$.
- $\pi_{i,t}$: citizen $i$’s payoff in period $t$.

According to the BDT model, citizens update their propensity after their action $I$ as follows:

$$p_{i,t+1}(I) = \begin{cases} 
  p_{i,t}(I) + \alpha(1 - p_{i,t}(I)) & \text{if } \pi_{it} > a_{it} \\
  p_{i,t}(I) & \text{if } \pi_{it} = a_{it} \\
  p_{i,t}(I) - \alpha p_{i,t}(I) & \text{if } \pi_{it} < a_{it}
\end{cases}$$

(1)

The parameter $\alpha \in (0, 1]$ controls the updating speed from the current propensity $p_{i,t}$ to the propensity in the next period. The aspiration level is updated by a weighted average of current aspiration and payoff, yielding $a_{i,t+1} = \lambda a_{i,t} + (1 - \lambda)\pi_{i,t}$ where $\lambda \in (0, 1)$.

According to the Fowler model, in contrast, the magnitude of the update parameter is solely controlled by $\alpha$:

---

1This model is called the Bush-Mosteller rule by BDT since it was originally proposed by Bush and Mosteller (1955).
Figure 1: Impact of feedback mechanisms on propensities ($\alpha = 0.05$)

The difference is depicted in Figure 1. The left panel shows the update magnitude ($\Delta p_{i,t+1}$) on the vertical axis as a function of $p_{i,t}$ on the horizontal axis according to BDT model. If a citizen obtains a higher payoff than her aspiration after voting, her vote propensity increases up to the level shown by the solid line. The update parameter $\alpha$ is highest at $p_{i,t} = 0$ and lowest (zero) at $p_{i,t} = 1$. If we assume the same probability for $\pi_{it} > a_{it}$ and $\pi_{it} < a_{it}$, then her vote propensity change as a function of $p_{i,t}$ is given by the dashed line. The mean of the solid and the dashed lines represents the expected change of propensity for $p_{i,t} = 0 \in [0, 1]$. Accordingly, if an individual’s propensity is less than 0.5, she is expected to increase her propensity, and vice versa. This is the moderating feedback which drives the average propensity toward 50%. This phenomenon is clearly absent in the Fowler model. Except for the two boundaries, the mid of the solid and dashed lines corresponds to zero, which means that individuals can stay at a certain propensity level for a longer time.

We first reproduced the simulation results obtained by BDT and Fowler in order to obtain reference values for a small constituency to be realized in a laboratory experiment. For both updating mechanisms, we ran 1000 simulations. We set voting costs higher (0.30) for one half and lower (0.18) for the other half of these 1000 simulation runs. In each simulation run, 10 supporters of one party and 10 of the other party repeated 1000 election periods. It corresponds exactly to the number of experimental persons in individual experimental sessions described below. For the other simulation
Table 1: Predicted turnout for different levels of voting cost

<table>
<thead>
<tr>
<th></th>
<th>c=.70</th>
<th>c=.50</th>
<th>c=.30</th>
<th>c=.18</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDT</td>
<td>.49</td>
<td>.52</td>
<td>.57</td>
<td>.58</td>
</tr>
<tr>
<td>Fowler</td>
<td>.33</td>
<td>.41</td>
<td>.45</td>
<td>.49</td>
</tr>
</tbody>
</table>

parameters, we used the following set of parameters, which are almost identical with those of BDT and Fowler:

- $\alpha = 0.1$: updating speed of propensity level
- $\lambda = 0.95$: updating speed of aspiration level
- $\epsilon = 0.05$: inertia\(^2\)
- support = 0.2\(^3\)

Table 1 presents the turnout rates of 500 simulation runs over 1000 periods. As a reference, we also gives the predicted value based on the expected utility model in the style of Palfrey and Rosenthal (1985). The simulated turnout rates are consistent with the results of BDT and Fowler, revealing only a small impact of voting costs in comparison with the expected utility model. Among the two simulation models, the Fowler model is more sensitive to the cost due to the lack of the moderating feedback mechanism (Fowler, 2006, 341). The impact of voting costs is, however, smaller than the results in Fowler (2006) due to the small population size ($n = 20$) in our simulation runs.

The more relevant difference between implications of the two simulation models, however, relates to the individual turnout rates. Figure 2 gives the distribution of turnout rates of individual voters for the last 30 periods. As Fowler has already shown, most agents in the BDT model are casual voters who vote and abstain from time to time. Therefore, the mode of the individual turnout lies around 50%. In contrast, agents in the Fowler model tend to be habitual voters or abstainers so that the turnout distribution is bimodal. Voting costs do not seem to substantively affect turnout rates, with the expectable but noteworthy exception of a smaller number of agents who always vote or abstain (i.e., habitual voters and abstainers) in the case of higher voting costs.

The distributions presented in Figure 2 are generated by different system dynamics. Figure 3 presents the variation in propensity changes in the simulations as a function of the previous propensity level and shows that the simulations actually generated the patterns derived by the model assumptions in Figure 1. The grey tone represents the frequency, with a darker tone indicating larger numbers, and the solid lines are the mean propensity changes conditional on the propensities in the previous period.

\(^2\)This parameter specifies the probability that an agent does not update its propensity or aspiration.

\(^3\)This parameter specifies a random shock to the payoff. More specifically, this parameter determines the width of the uniform distribution around the payoff from the previous round. From the distribution a payoff is randomly drawn and compared with the aspiration level.
These mean changes are clearly different in the two system dynamics. In the BDT model, high voting propensities in the previous period tend to reduce current propensities on average and vice versa. This mechanism generates the means-reverting processes resulting in the observed cluster of propensities around 50%. In the Fowler model, especially with $c = 0.18$, agents with high voting propensities tend to increase their propensities and vice versa. This generates system dynamics toward the extreme high and low level of propensities.

The centrifugal dynamics of the Fowler model, however, cannot be explained solely by its propensity updating mechanism. As can be seen in Figure 1, the propensity can increase or decrease at the same rate for a wide range of previous propensity levels. As such, this updating mechanism would produce a uniform distribution of propensities rather than a bimodal one. The generating process of the bimodal distribution of Figure 2 can be found by inspecting patterns in individual simulation runs, which aggregate into the distribution exhibited in Figure 3. Figure 4 exemplifies the observed pattern. The top left panel shows turnout levels for the two groups, $A$ and $B$. The simulation shows a regular pattern of two phases. In the first phase, the two groups compete in a close rally by mobilizing their members. This rally is decided by relative turnout. After this phase of increasing turnout, the turnout rate of the group which has lost the rally begins to decline while the other group maintains a high turnout level. However, after some iterations, the leading group’s turnout level begins to decrease as well, although it remains above the other groups turnout and hence continues to win. When both groups touch the bottom, their turnout levels begin to rise and a new rally begins. The two different phases, an increasing trend in a close rally and a decreasing trend with a time lag, is also manifest in the two groups’ mean payoffs and aspiration levels (the bottom left panel). In the
close rally with rising turnout, the payoffs of both groups have a larger variance and both group’s aspiration converge to the mean payoff in a longer run. In the one-sided phase with declining turnout, the winning group’s payoffs remain at a higher level while the aspiration level begins to increase at a slower rate. Therefore the group can continue to mobilize its members. The other group, in contrast, continues to obtain lower payoffs while its mean aspiration level is still higher. As a result, more and more members shift to abstaining and the group continues to lose against the other group. In this one-sided phase, the voting propensities of the winning and the losing groups show a interesting contrast (the top and bottom right panels). While most members of the losing group have very low propensities, the members of the winning group are divided in two clusters, those with very high propensities and those with very low propensities. In total, this results in a division of the agents into two groups, habitual voters and abstainers. In short, the Fowler model produces a phase of one-sided turnout which generates, in turn, habitual voters and abstainers.  

We can compare the dynamics of the Fowler model described above with that of the BDT model,
which is exemplified by Figure 5. In this simulation run, which is typical of the BDT model, no clear distinction between two phases like in Fowler’s model is observable. Since the propensity updating mechanism of the BDT model has a feedback effect toward 50% the race is very close in almost iterations and one-sided games are only exceptionally observed.

### 3 Experimental Design

A test separating the BDT mechanism from the one proposed by Fowler can hardly be conducted by using conventional survey data. First, these studies can track respondents’ turnout only for a small number of elections. Second, turnout rates based on surveys in general suffers from over-reporting due to social desirability (McDonald, 2003; Holbrook, Green and Krosnick, 2003; Holbrook and Krosnick, 2010; Karp and Brockington, 2005) and/or lower response rates of abstainers (Clausen, 1968; Yalch, 1976; Burden, 2000). The South Bend data cited by Fowler tracked seven elections with official
Figure 5: An example simulation run based on the BDT model.

turnout records, but also this data suffers from the second problem. Third, measures of benefit, cost and aspiration in survey data are of low quality. This prevents a direct comparison of the two different propensity updating mechanisms.

In order to study whether real behavioural dynamics can be described by the propensity updating process proposed by BDT or by Fowler we set up a laboratory experiment testing the two different propensity updating mechanisms in a more direct way than other methods would allow.

More specifically, our laboratory experiment used the following setup. Each session is conducted with 20 subjects. 10 subjects are randomly assigned to one group (A) and the other 10 to the other group (B). Subjects participate in 30 rounds, repeating a voting task in which they have to decide between two alternatives, A and B. In each period, subjects decide to cast their vote or not. The voting process is set up as a two-step procedure. In the first stage, subjects are asked if they wished to cast their vote or abstain. If a subject decides to participate in the vote she has to pay a voting

5The experiments were programmed and conducted in z-Tree (Fischbacher, 2007), which is partly inspired by the design chosen by (Duffy and Tavits, 2008).
cost and is admitted to the second stage in which she is offered a choice between A and B. Subjects deciding to abstain are shown a waiting screen until the second stage is finished. After the second stage, the number of votes for A and B are summed and the alternative with more votes is declared winner. In case of a tie, the winning alternative is decided by a random draw.\(^6\) Individual payoffs in each period depend on the vote result as well as on the subject's first decision to participate or not. Depending on the vote results, individuals receive payoffs according to Table 2.

Table 2: Payoff depending on the vote result

<table>
<thead>
<tr>
<th></th>
<th>Group A</th>
<th>Group B</th>
</tr>
</thead>
<tbody>
<tr>
<td>A wins</td>
<td>1 Euro</td>
<td>0 Euro</td>
</tr>
<tr>
<td>B wins</td>
<td>0 Euro</td>
<td>1 Euro</td>
</tr>
</tbody>
</table>

Given that our core question is whether the BDT or the Fowler propensity updating mechanism is better able to describe the actual dynamics of voting behaviour, our experiment does not differentiate between a treatment and a control group. We do not test a causal effect but whether the observed pattern matches one of two theoretically predicted patterns. In addition, we test for the effect of voting costs by differentiating between two cost levels. This is 0.30 Euro in the high-cost treatment and 0.18 Euro in the low-cost treatment. Costs are subtracted from the payoff.

We conduct six sessions with 20 subjects each, three of which were high-cost treatments and the other three low-cost treatment (for more detail see below). In the first stage, subjects have to make their decisions within 20 seconds, otherwise they are assumed to abstain and do not participate in the second stage. Those who do not take a decision in the second stage within 20 seconds are assumed to vote for the alternative with the same label (A for A and B for B) as their group affiliation.\(^7\)

In each round, after the decisions and before the announcement of the result, subjects are also asked to estimate the likelihood of the result using a 11 point scale. The leftmost point on this scale corresponds to a certain victory of A and the rightmost point a certain victory of B. The mid-point of the scale represents an equal likelihood of victory of both alternatives. In the following analysis, we interpret this expected result as an approximation of aspirations.

After 30 voting rounds, participants are asked to fill a questionnaire. After that, they received their payoff which is the sum of the payoffs of 10 randomly selected rounds. On average, a session lasted about 45 minutes and average payment was 7.96 Euro. 120 subjects from the subject pool of the laboratory at the University of (NAME INSERTED AFTER REVIEW) took part in the experiment.\(^8\)

\(^6\)Out of a total of 180 rounds, 25 (13.9\%) resulted in a tie. These ties were rather uniformly distributed over sessions.

\(^7\)99.9 percent of the decisions in stage 1 and 99.98 percent of the decisions in stage 2 were made within the 20 seconds limit.

\(^8\)We utilized ORSEE for the recruitment of participants (Greiner, 2004).
4 Experimental Results

In this section we report our experimental results in terms of the model assumptions and implications. As has been seen above, both the Fowler and the BDT models are based on certain behavioural assumptions. We first study the assumptions on the updating mechanisms of voting propensities and aspirations. Then we turn to the implications of the two models. We report on the aggregate turnout level, on the distribution of individual turnout levels, and on the dynamics of group competition.

In reporting results we group our data, which are measurements of individual behaviour of 120 subjects for 30 rounds, into three periods of ten rounds each in order to compare differences in behaviour over time which may be the result of learning from experiences during the session. One piece of evidence pointing toward learning processes is presented in Figure 6 which gives average response times for turnout decision in each round. The time needed to decide whether to vote or not decreases rapidly during the first five rounds. Afterwards, the delay still continues to decline at a very slow and decreasing rate approaching some limit around two seconds. This indicates that over time the subjects develop certain response heuristics which may correspond to the mechanisms of the BDT and/or Fowler model.

Figure 6: Average response time for turnout decision in each round (point estimates with 90% confidence intervals)

4.1 Test of the Core Model Assumptions

The core assumptions shared by the BDT from the Fowler model are the updating parameters of voting propensities $\lambda$ and aspiration $\alpha$. This will be tested by using a relatively simple statistical model. For updating propensities, both the BDT and the Fowler models assume that citizens would
change their behaviour if they are not satisfied with their payoffs from the previous period. This assumption is captured by the following statistical model:

\[ C_{i,t} = \beta_0 + \beta_1 S_{i,t-1} \]  
\[ \hat{a}_{i,t} = \pi_{i,t-1} \geq \hat{a}_{i,t-1} 
\]

\[ \hat{a}_{i,t} = e_{it} \times 1 - c_{it}, \]  

where \( e_{it} \) is the expectation of subject \( i \) in round \( t \) that the preferred alternative wins the vote. As stated above, subject would receive payoff of 1 if the preferred alternative wins. \( c_{it}, \) voting cost, equals either 0.30 (High cost) or 0.18 (Low cost), which is subtracted from the payoff if the subject voted. If the subject abstained, \( c_{it} = 0. \)

If the assumption of propensity updating holds the estimated value for \( \beta_1 \) should be negative. This is because positive feedback at \( t - 1 \) should decrease the likelihood of changing one’s action at \( t \) relative to \( t - 1 \). This prediction is, however, only partially confirmed by our data. According to Figure 7 subjects in the low cost treatment behaved as expected in the first and last ten rounds. For subjects in the high cost treatment this only holds for the last ten rounds. During the second period, the coefficient is positive for both the high cost and the low cost treatments. That is, subjects tend to change behaviour if they are satisfied with their payoffs, and vice versa.

One possible explanation for this unexpected result in the second period is that subjects who voted and won have an incentive to abstain in the next round because their payoff increases by free-riding on the other group members’ turnout. Another possible explanation focuses on subjects who abstained in the previous election. Accordingly, they expected their party to lose but their group won.
the poll. Thus, they received more payoff than expected. These subjects may tend to vote in the next round since they have been convinced of the other group members’ engagement at the previous round, thereby generating a bandwagon effect. We tested both explanations by estimating the same statistical model separately on decisions after a vote and on decisions after abstention (Figure 8). The results clearly support the latter explanation referring to abstainers who were convinced to vote by the outcome of the previous round. In contrast, decisions made after a vote are consistent with the core assumption of propensity updating. That is, if subjects voted in the previous round and are satisfied with the result, they tend to continue to vote rather than changing their behaviour. This holds for all three periods. That is, ABAR works partially conditioned by which decision was made in the previous round.

Turning to the aspiration updating process assumed in the BDT and Fowler models, we proceed in an analogous way and estimate a simple statistical model:

\[ \hat{a}_{i,t} = \gamma_0 + \gamma_1 \hat{a}_{i,t-1} + \gamma_2 \pi_{it-1} \]  \hspace{1cm} (5)

Figure 9 shows that both estimated aspirations and realized payoffs in the previous round have a positive impact on current expected payoffs. This corresponds to the mechanism assumed by BDT and Fowler. Further, estimated aspirations at \( t \) are more strongly affected by those at \( t - 1 \) than by payoffs at \( t - 1 \). This result indicates that laboratory results are close to the parameter choice in the BDT and the Fowler model, as well as our simulation (\( \lambda = 0.95 \)).\(^9\)

\(^9\)Since estimated aspirations based on expected payoffs can include some idiosyncratic elements, we have also estimate random intercept models. Controlling for this factor does not affect the substantive results presented in Figure 9.
Figure 8: Parameter estimates of the propensity updating model. Point estimates with 95% confidence intervals. (separate analyses for voters and abstainers in the previous round)
Figure 9: Parameter estimates of the aspiration updating model

Cost=0.3

Cost=0.18
4.2 Test of Implications: Aggregate and Individual Turnout Levels

We begin our analysis of model implications by inspecting aggregate-level turnout. Table 3 compares turnout in each treatment and each round. We observe at least two important findings: First, there are no clear differences in turnout between the high-cost (solid lines) and low-cost (dashed lines) treatment. The null hypothesis that the two treatments yield identical turnout levels cannot be rejected at $\alpha = 5\%$. This is also the case when we compare the two treatments in the first, second, and third period of ten rounds separately. Second, in general, turnout has a decreasing trend over time.

<table>
<thead>
<tr>
<th>Table 3: Average turnout of each treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Periods</td>
</tr>
<tr>
<td>$c = 0.30$</td>
</tr>
<tr>
<td>overall</td>
</tr>
<tr>
<td>1-10</td>
</tr>
<tr>
<td>11-20</td>
</tr>
<tr>
<td>21-30</td>
</tr>
</tbody>
</table>

Comparing this result with the predictions in Table 1, we conclude that the expected utility model shows the least congruence with the data with regard to both the predicted substantive differences between the two treatments and the predicted level of turnout. In contrast, these aggregate-level results do not allow us to rank the fit of the two aspiration-based models, since the level of turnout depends on the periods under study. In the first ten rounds, the observed data seem to be right between the predictions of BDT and Fowler. In the second ten rounds, the Fowler model seems to fit the data best. In the last ten rounds, however, the level of turnout further decreases.

As discussed above, the aggregate level results contain only limited information for differentiating between the BDT and the Fowler models. The difference between the two should be more remarkable at the individual level. More specifically, the simulated data suggested that the BDT-model produces a unimodal distribution of individual turnout rates over the rounds with a peak at 50%, while the Fowler-model realizes a bimodal distribution with habitual voters and abstainers. Figure 10 shows the developments of the corresponding distribution over time. In the first period of ten rounds (shown in the first row), there is a more or less unimodal distribution with a mode near 50% for both cost treatments, which corresponds more to the BDT model. In the second period, the number of habitual abstainers begins to grow and this group becomes a mode in the last period. This result can be interpreted as evidence in favour of the Fowler model. We also have to note, however, that our simulation results based on the Fowler model predicted a clearer bimodal distribution while our experimental data contain less habitual voters than predicted by the Fowler model.
Figure 10: Distribution of individual turnout rate in different periods

\[ c = 0.30: \text{Period 1–10} \]

\[ c = 0.18: \text{Period 1–10} \]

\[ c = 0.30: \text{Period 11–20} \]

\[ c = 0.18: \text{Period 11–20} \]

\[ c = 0.30: \text{Period 21–30} \]

\[ c = 0.18: \text{Period 21–30} \]
4.3 Test of Implications: Dynamics in the Data

In this subsection we observe the dynamics in our experimental data more directly. As a reference we keep in mind the different patterns of both models described in Figures 4 and 5.

Figure 11: Dynamics of each group’s turnout

Figure 11 presents the turnout level of both groups in each of the six sessions. Remember that the simulation models based on the BDT and Fowler models showed a clear contrast: While both groups in the BDT model continue to compete for victory in a close rally, the Fowler model generates two distinct phases of a close rally with increasing turnout and a one-sided phase with decreasing turnout. Each of the six sessions provides somewhat mixed results. However, longer sequences of close rallies such as generated by the BDT model are barely observed. Rather, rallies are one-sided with alternating winning groups and some of the groups’ turnout is extremely high or low, as generated by
the Fowler model. However, we do not observe a clear distinction into the two phases characteristic of the Fowler model.

Figure 12: Dynamics of average payoffs and estimated aspiration (\(\hat{a}\))

Figure 12 presents the average payoffs and estimated aspiration levels of the two groups in each of the six sessions. The trends of the payoffs resemble the pattern generated by the Fowler model, and the pattern can be clearly attributed to the one-sided rally with alternating winning groups. In contrast, the trends in the aspiration levels seem to correspond to neither of the two models. While both theoretical models generate larger discrepancies of the mean aspiration levels between groups, our experimental results point toward more moderate discrepancies between the aspiration levels. This might be due to measurement bias of estimated aspirations based on the expected payoffs. That is, the evaluation of the expected results is a relatively difficult task for subjects, so that their answer
may be biased toward the midpoint of the scale (a 50-50 chance), which results in more moderate values of this proxy of aspiration levels.

Despite this concern about our measure of aspirations we can still say that the dynamics of turnout, payoffs and aspiration are directly observable. In contrast, the trend of voting propensity cannot be directly found in our data set. In order to study voting propensities, we therefore set up a Bayesian full probability model in which voting propensity and aspiration are integrated as latent variables:

\[
\begin{align*}
    V_{i,t} & \sim BR(p_{i,t}) \\
    \hat{a}_{i,t} & \sim N(a_{i,t}, \sigma_e^2) \\
    p_{i,t} & = f(p_{i,t}, \alpha, \hat{a}_{i,t}, \pi_{i,t}) \\
    a_{i,t} & = \lambda a_{i,t-1} + (1 - \lambda) \pi_{i,t-1} \\
    \alpha & \sim Unif(0,1) \\
    \lambda & \sim Unif(0,1) \\
    p_{i,1} & \sim Unif(0,1) \\
    a_{i,1} & \sim N(0,1000)
\end{align*}
\]

The first two equations define vote choice and expected payoffs in terms of probabilistic processes. Voting behaviour is modelled as a series of Bernoulli trials with vote propensities as probabilities of success (Equation 6). In contrast, estimated aspiration based on the expected payoffs is assumed to be distributed normally with aspiration as expected value (Equation 7). Both parameters, voting propensities and aspirations, are updated using their past values, past behaviour, and past payoffs by using either the BDT or the Fowler model (Equations 8 and 9). For these updating processes, we estimate two parameters, the magnitude of the voting propensity update \(\alpha\) and aspiration inertia \(\lambda\). For these two parameters, we set flat priors between 0 and 1. Furthermore, we also set uninformative priors for the voting propensity and aspiration in the first period (Equations 12 and 13).

The posterior information of the parameters are obtained using Markov Chain Monte Carlo techniques. The corresponding Markov Chains were run separately for individual subjects for three sets of 10 period for several reasons. First, the adaptive process is individual-specific and there is no plausible reason to assume a common parameter value for all subjects. Second, the analysis in the previous subsections revealed some differences in the first, second and third period of ten rounds. Thus, the parameters changed their values in the course of the experimental sessions. The third reason is a more practical one: The estimation of the model above is highly computation-intensive.
and it is not feasible to run the Markov Chain for multiple individuals for all 30 rounds. Although the Markov Chains were run separately we used the posterior of $p$ and $a$ in the last round as prior of the subsequent round of the same subject. By doing so, we mimic an estimation with all rounds with varying parameter of $\alpha$ and $\lambda$. Below, we report only the posterior information from the Markov Chains which converged. 24% of 360 estimations (3 periods of ten rounds for each of 120 subjects) for the BDT model and 21% of those for the Fowler model had convergence problems and the corresponding posterior cannot be reported. Many of these problematic cases are person-periods with no or extremely low variance in voting behaviour and/or estimated aspiration. In total, the person-periods with convergence problems are distributed almost uniformly across sessions and groups. Therefore, reporting only the posteriors of the converged Markov Chains will not seriously bias the conclusions.

Figures 13-15 report the distribution of overall propensities of groups in grey tone for session 1 and 2, 3 and 4, and 5 and 6, respectively. The left column of graphs in the figures report the results of group $A$, the left panel contains those of group $B$. For each session, the upper row presents the predictions based on the model containing the updating mechanism used by Fowler, the lower row presents those based on the BDT model.

The trends of the mean propensity levels of individual subjects are indicated by white lines. At the first look, the individual voting propensity levels have large discrepancies in most groups, which corresponds to the dynamics generated by the Fowler model. This is even the case in the results based on the BDT-model. The most striking result in this regard is a group of session 5 (right panel), which shows a clear bimodal distribution of propensities in both models. That is, even if one relies on a statistical model based on the BDT-model, the data forces the estimates of voting propensities into a bimodal distribution. This result in favour of the Fowler model is also reflected in the deviance measures which give information about the extent to which the model fits the data. According to this measure, the Fowler model fits the data better than the BDT model with probability 61%. Further, comparing deviances for individual person-periods confirms the better fit of the Fowler model in the third period of ten rounds. This is also consistent with the visual impression obtained from Figures 13-15, suggesting a bimodal tendency of voting propensity or a convergence at extreme values (0 or 1) in the the third period. As has been seen above, the test of one of the core assumption, propensity updating (Figure 7 and 8) revealed that the propensity updating mechanism holds better in the third period than the other two periods. Considering this result, and bearing in mind that subjects often need some experience before they find their strategy, the better fit of the Fowler model in the third period should be more emphasized than the results for the first period.

\[^{10}\text{The deviance measure is also collected from the posterior distribution per Markov-Chain-Monte-Carlo. That is, in 61% of overall iterations of all person-periods we obtained a better fit for the Fowler model.}\]
5 Conclusion

This paper’s aim has been to provide a more direct empirical test of two competing adaptive behavioural models of turnout. The two models differ only in their feedback mechanisms which, however, leads to interesting implications for several aspects of the voting dynamics. Our experimental data based on a lab experiment allowed us to test the core assumptions of the models including the propensity updating mechanisms and to evaluate the fit of model implications to the data.

Concerning the model assumptions, there is mixed evidence with regard to the voting propensity updating mechanisms. More specifically, subjects who voted in the previous round tended to follow the updating principles common to the BDT and Fowler models, while subjects who abstained in the previous round did not respect the principles. Analyzing all voters together we find the expected propensity updating mechanism only in the third period of ten rounds. In contrast, the aspiration mechanism on which the model build is fully supported by our data.

Results concerning the model implications are also somewhat mixed. One important exception is the trends in estimated voting propensities based on the BDT and Fowler model. Individual voting propensities are likely to take extreme values corresponding to the Fowler model while no supporting evidence for the BDT model can be found in this respect. Moreover, the model fit favours the Fowler model over the BDT model in particular in the third period.

While our experimental data tend to produce results that are more in line with the Fowler model, we also have to note that not all dynamics predicted by the Fowler model could be observed. Perhaps the most relevant element missing in the data is the prediction of two distinct phases: the close rally with increasing turnout levels and the one-sided phase with decreasing turnout levels. The data showed some one-sided rallies with alternating winning parties, but no cycles with two distinct phases. One possible reason is that our sessions contained an insufficient number of rounds so that subjects had insufficient time to adjust their aspiration levels and regain hope of winning in the next round after lengthy sequences of continuing defeats so that no cycles were generated. Note that the simulations using similar parameters as the ones estimated from the data needed a much larger number of rounds than the number implemented in the laboratory. This is also consistent with the result that the third period had a better fit with respect to both the core assumptions and the model dynamics.

Finally, let us briefly speculate about the overall comparison of the dynamics in the experimental data and in the simulations. The pattern observed in the individual simulations based on the Fowler model suggests that the downward phase in which one group starts shifting to abstaining, followed by the other group with a delay of a few rounds, suggests an interpretation in terms of collective action theory: As soon as their own vote does not seem necessary anymore to win the next competition, collective action theory expects individuals to turn to free riding on the participation of other group
members. On the other hand, if members of the losing group in this phase observe turnout rates in the winning group decline, they may believe that it is more likely that they may win in the next round if they participate. Hence linking aspiration-based adjustment rules with the literature on participation games (Sonnemans and Schram, 2008) may be a worthwhile avenue for further research.

From the reported results it seems like testing for aspiration-based adjustment rules using field data on consecutive elections, even with panel studies, is unlikely to generate more conclusive results. Instead, one might hope to find even clearer evidence in longer experimental sessions.

References


Figure 13: Dynamics of each group’s propensities in Session 1-2. White lines are mean propensities of individual subjects.
Figure 14: Dynamics of each group’s propensities in Session 3-4. White lines are mean propensities of individual subjects.
Figure 15: Dynamics of each group’s propensities in Session 5-6. White lines are mean propensities of individual subjects.

Session 5 (Cost=0.3); Fowler Model

Session 5 (Cost=0.3); BDT Model

Session 6 (Cost=0.18); Fowler Model

Session 6 (Cost=0.18); BDT Model