JOURNAL OF THEORETICAL POLITICS

Article

Journal of Theoretical Politics
25(2) 214–233

DOI:10.1177/0951629812460120
jtp.sagepub.com
SAGE

Electoral agency in the lab: Learning to throw out the rascals

Leif Helland
Department of Economics, Norwegian Business School, Norway

Lars Monkerud
Department of Economics, Norwegian Business School, Norway

Abstract
Models of electoral agency address the levels of discipline and selection that voters can achieve in elections. The models are demanding in terms of individual belief formation and consistency of behavior. We investigate a baseline model of electoral agency in a controlled laboratory environment. This baseline model, although simple, forms the central plank of more complex electoral agency models. Our design seeks to limit the behavioral impact of social preferences. We find little support for the baseline model in our data. However, simple (non-rational) learning rules explain behavioral patterns well. Simulations indicate that non-rational learning drives behavior most forcefully towards equilibrium in situations that are favorable to Bayesian updating.

Keywords
electoral agency; experiment; learning

1. Introduction

In representative democracies voters have the option to condition reelection on the observed performance of incumbents. Regular access to elections allows voters to 'throw out the rascals' peacefully. Arguably this is the defining characteristic of democratic government (Hayek, 1979; Popper, 1989; Riker, 1982; Schumpeter, 1996). Identifying conditions that enable voters to retain or replace incumbent rulers in intelligent ways is therefore an exercise worth under taking.

Current electoral agency models assign voters the twin tasks of disciplining bad incumbents and selecting good incumbents (Austen-Smith and Banks, 1989; Banks and

Corresponding author:
Leif Helland, Department of Economics, Norwegian Business School, Nydalsveien 37, 0442 Oslo, Norway.
Email: leif.helland@bi.no

Sundaram, 1993; Besley, 2006; Fearon, 1999; Maskin and Tirole, 2004).[1] There is a trade-off between the two; better selection comes only at the cost of weakened discipline and vice versa. This trade-off places significant demands on voters. Beliefs are required to be consistently updated given observed outcomes, and votes are required to be optimal given such beliefs.

Taken literally, electoral agency models make unrealistic behavioral claims on voters. While figuring out the strategic complexities will carry significant cognitive costs, the expected return of informed voting is at best marginal in realistically sized electorates (Downs, 1957; cf. Caplan, 2007). In short, there are few incentives for voters to behave in rationally informed ways. Politicians, on the other hand, who stand to win or loose positions in high office, have high powered incentives to acquire relevant information and act rationally on it.

In this article we investigate voter behavior in a simple electoral agency experiment. Taking Besley's core model (Besley, 2006) as our starting point, we seek to identify conditions under which less than rational learning rules converge on, or diverge from, the equilibrium of this model.

Our design captures the asymmetric incentives of voters and politicians with regard to making informed choices. In the experiment politicians are automatons, while voters are human subjects.[2] The automatons are programmed to mimic the equilibrium behavior of politicians in the electoral agency model that is tested. Thus, we explore the behavior of subjects as voters in a (highly artificial) situation where it is public knowledge that politicians behave in a manner fully consistent with the electoral agency model.

As is well known, systematic deviations from self-interested equilibrium are explained by fairness preferences and intentions-based reciprocity in a number of simple games (see Fehr (2009) and Bolton et al. (2009) for reviews). Insofar as such motivations are not part of the agency model we explore, we wish to minimize their behavioral impact. The use of politician-automatons renders intention-based subject responses unlikely (since a computer program does not have intentions). By design there is no payoff variation among voters belonging to the same electorate in our experiments; fairness concerns are therefore not likely to come into play.[3] By controlling for fairness preferences and intentions in this way, our design allows us to explore the learning rules in an environment in which self-regarding voter responses are cultivated. It also allows us to draw firmer conclusions about learning in an environment that closely resembles that of the agency model we are studying.

While there have been numerous field data tests of electoral agency models (Besley et al., 2010; Helland and Sørensen, 2011; Petterson-Lidbomm, 2006; Svaleryd and Vlachos, 2009; Svensson, 1999, and more indirectly by Alesina et al., 1999; Easterly and Levine, 1997; cf. Persson and Tabellini, 2000), it is difficult to draw clear-cut causal inferences from them (due to, for example, institutional heterogeneity of polities, measurement problems on key variables, thorny questions of reversed causation, endogeneity, and selection bias). The experimental method allows a greater degree of direct control of the central building blocks of agency models (for example, voter preferences, beliefs, electoral institutions, and incumbency performance), and (partly for this reason) facilitates inferences about causal mechanisms. The price paid is uncertainty with respect to external validity. In our opinion experiments can provide a useful supplement to field data studies, not least, we believe, in the study of electoral agency.

A number of experiments deal with elections.[4] Only a minority investigate agency problems. Markussen and Tyran (2009) use an agency framework to study selection of politicians, given two kinds of signals about potential candidates (contributions in a public goods game, and score on an IQ-test). Discipline is not an issue in their experiment, since there is no reelection; rather it was designed to explore the impact of fairness preferences on selection.

Aragones and Palfrey (2005) and Houser et al. (2008) run experiments on selection of politicians who differ in (exogenously given) types. These experiments do not depart from an explicit agency framework (but utilize variants of prospective voting models). Elections are not repeated, so issues of discipline do not enter here either.

Dasgupta and Williams' (2002) study comes closest to our own, in that they also approach the twin challenges of discipline and selection. In their set-up voters are subdivided into two groups; one group is informed about policy outcomes produced by the current incumbent, the other group is not. Informed voters observe outcomes with noise (since outcomes are a function of randomly drawn competence and incumbent effort choice). After the noisy signals have been transmitted to informed voters, all voters participate in a fixed number of polls. Each aggregate polling result is made public knowledge once it is concluded. Voters thereafter either reelect or oust the current incumbent. Since getting reelected is valuable, discipline therefore enters as a relevant concern. Selection is also a concern because the incumbent and challenger will have different policy preferences, and possibly also differ in terms of their qualities. Given the sequence of polls, uninformed voters may update by observing the poll results.[5] Voters and politicians are human subjects in the experiment.

Now, although the set-up is one of incomplete voter information, Dasgupta and Williams (2002) find that voters behave as if they were fully informed. There are two reasons for this: informed voters are able to extract information from their noisy signal; and uninformed voters in turn are able to extract this information from aggregate polls. Two alternative 'attention rules' are explored: (a) no learning related to output or polls; (b) rational learning related to output but no learning related to polls. None of these alternatives explain the data as well as the alternative in which voters learn rationally from both sources.

The possibility of voters learning from both polls and observed outcomes, but in ways other than by consistently applying Bayes' rule is not considered. Our experiment shows that non-rational learning rules may, but need not, converge on a perfect Bayesian equilibrium in an agency setting. This suggests that the results in Dasgupta and Williams (2002) need not necessarily have been produced by the mechanism underlying a rational expectations equilibrium, and that convergence on equilibrium need not happen for other parameters in their experiment. In short, while Dasgupta and Williams (2002) ask *if* voters learn to play the equilibrium in an agency environment, we ask *how* and *when* voters learn to play an equilibrium in such an environment.[6]

The explanatory force of the electoral agency model we are studying hinges critically on voters being able to update beliefs in accordance with Bayes' rule. There is an experimental literature on individuals' ability to perform such updates. What one finds is that individuals perform Bayesian calculations significantly better when the problem is presented in terms of frequencies, rather than probabilities (see Gigerenzer et al., 2009, for an overview). To give Bayesian updating a fair chance, therefore, we gave the subjects in

our experiment their decision problems in frequency terms. The frequencies versus probability literature, however, focuses on one-shot individual decision problems. We explore a richer environment, in which learning takes place in a strategic context.

With Bayesian learning as a point of departure, we check the explanatory power of two non-rational learning rules: fictitious play learning (Brown, 1951) and payoff reinforcement learning (Erev and Roth, 1995). As Camerer and Ho (1999) have demonstrated, fictitious play and payoff reinforcement learning are both special cases of a more general learning model ('experience weighted attraction'). Rather than relying on Camerer and Ho's rather heavily parametrized model, we follow the simpler twin rule approach. This choice is grounded in improved tractability, as well as previous explanatory success of these simpler rules.

While we find little support for the electoral agency model in our data, simple (non-rational) learning rules do explain behavioral patterns well. Moreover, simulations indicate that non-rational learning drives behavior most forcefully toward equilibrium in situations that are favorable to Bayesian updating. In situations that are less favorable to Bayesian updating, behavior stabilizes away from equilibrium.

A pertinent question to ask would be whether foresighted politicians facing non-rational voters would continue to behave as stipulated by the electoral agency model. Departing from our experimental results, we argue that selection pressures limit the extent to which sophisticated politicians can take advantage of non-rational voters.

The paper is organized as follows. The model is presented in the next section, followed by an outline of the design. Then, results are presented in some detail. Limits on sophisticated politicians are discussed prior to a brief conclusion.[7]

2. Model

Write individual utility as $w_t = (1-\tau)y + \alpha x_t$, where y is pre-tax income, τ a given tax rate, x_t public output in stage $t \in \{1,2\}$, and $\alpha > 1$.

The public budget is required to balance in each stage, so that $\theta(\tau y - r_t) = x_t$. In the budget restriction $\theta \in \{s,1\}$, $0 < s < 1$ is a persistent productivity shock, and $0 < r_t \leq R \leq \tau y$ is rent extraction in stage t.[8] It is also required that $R > (1-s)\tau y$, as a technical assumption. Productivity is drawn from the distribution $\Pr(\theta = s) = q$, and $\Pr(\theta = 1) = (1-q)$. Only the case with $q \geq \frac{1}{2}$ is analyzed.[9]

Let there be two types of politicians $\iota \in \{g,b\}$, referred to as 'good' and 'bad', respectively. The types have objective functions $v_g = w_1 + \delta w_2$ and $v_b = r_1 + \delta r_2$, where $\delta < 1$ is a common discount factor. Let the prior distribution of types be given by $\Pr(\iota = g) = \pi$, and $\Pr(\iota = b) = (1-\pi)$.

The following timeline applies.

1. Incumbent-type and productivity are drawn, and observed by the incumbent only.
2. The incumbent sets rent extraction for stage one. Production is determined residually.
3. Stage one payoffs are distributed, and observed by all players.
4. Elections are held at the end of stage one; they determine whether the current incumbent is kept or replaced by the challenger. If the challenger wins, he observes the productivity draw from the first stage, and his type is drawn and observed by him only.

5. The (re)elected politician sets rent extraction for stage two.
6. Stage two payoffs are distributed and the game ends.

As is immediately clear, in stage two (the last stage), a bad politician extracts maximal rents while a good politician extracts zero rents. Since good politicians never steal, they will either produce $x_1 = \tau y$ or $x_1 = s\tau y$, depending on the realization of the productivity draw. For any other level of production, it must be the case that $\Pr(g|x_1) = 0$. Since future rents are discounted, $r_1 = R$ dominates $r_1 = 0$ for a bad politician (irrespective of productivity $r_1 = R$ pays $(1 + \delta)R$ if reelected and R if not, while $r_1 = 0$ pays δR if reelected and 0 if not).

A bad politician facing $\theta = 1$ may nevertheless find it worthwhile to mimic a good politician facing $\theta = s$. This will net the bad politician $r_1 = (1 - s)\tau y$, in addition to δR if he is reelected. Denote the probability that a bad politician extracts $r_1 = (1 - s)\tau y$ by λ.

Attention is limited to the use of pure cut-off strategies by the voters. These strategies instruct the voter to reelect if and only if the updated belief in the current incumbent being good is at least as high as the probability of the challenger being good. The voter's updated belief of having a good incumbent after observing $x_1 = s\tau y$ follows from Bayes' rule

$$\Pi = \frac{q\pi}{q\pi + (1 - q)(1 - \pi)\lambda} \tag{1}$$

After observing $x_1 = s\tau y$ the voter follows his pure cut-off strategy and reelects if and only if $\Pi \geq \pi$, or equivalently if and only if $\lambda \leq \frac{q}{(1-q)}$. Thus, for $q \geq \frac{1}{2}$ (which is the case analyzed), reelection is certain after $x_1 = s\tau y$ has been observed. Let $0 < \rho < 1$ signify the probability that the incumbent is reelected.

It is easy to see that a separating equilibrium (with $\rho = 1$ and $\lambda = 0$) exists if $\tau y(1 - s) + \delta R < R$, and that a pooling equilibrium (with $\rho = \lambda = 1$) exists if $\tau y(1 - s) + \delta R \geq R$.

3. Design

In all sessions of the experiment we held the following parameters constant: $s = 0.5$ ('low productivity draw'); $y = 100$ schillings ('endowment per stage'); $\tau = 0.5$ ('tax rate'); $\alpha = 1.1$ ('marginal value of public production'); $\pi = 0.2$ ('a priory probability of a good politician').

Sessions were conducted with electorate size 1 ('decisive voter') and 3 ('deciding by simple majority'), to check for learning effects due to group decision making (which should be nil according to the model), and differences with respect to electorate size (which also should be nil according to the model).

For each electorate size we performed a session with marginal updating in which $q = 0.55 \Rightarrow (\Pi - \pi) = 0.03$, and a session with substantial updating in which $q = 0.85 \Rightarrow (\Pi - \pi) = 0.39$. The idea was to check whether equilibrium behavior requires substantial updating.

In each game, after observing first-stage production, subjects were required to register their subjective probability assessment that the first-stage politician was bad. Within each

Table 1. Design.

	Session 1	Session 2	Session 3	Session 4
Date	26 Nov 07	27 Nov 07	26 Oct 08	29 Oct 08
Electorate size	3	3	1	1
Number of subjects	18	18	20	20
Updating	Marginal $\Delta\Pi = 0.03$	Substantial $\Delta\Pi = 0.39$	Marginal $\Delta\Pi = 0.03$	Substantial $\Delta\Pi = 0.39$
Matching	Absolute stranger	Absolute stranger	Not relevant	Not relevant
Number of games played	7	7	20	20

session subjects were informed that the minimal absolute deviation between registered beliefs and actual draws would win a price of 500 Norwegian Kroner (NOK), and that a fair lottery would pick a winner in case of a non-unique minimum.

Incumbent behavior was programmed in the computer (rather than having voters face humans in the role of politicians). The programmed behavior was as follows: if good type and $\theta = 1$, allocate 50 schillings to public production in both stages; if good type and $\theta = 0.5$ allocate 25 schillings to public production in both stages; if bad type and $\theta = 1$ allocate 25 schillings to public production in stage one and nothing in stage two; if bad type and $\theta = 0.5$ allocate nothing to public production in either stage.

In the three-subject electorates we employed an absolute stranger design (in which no subject was matched with subjects whom this subject had been matched with in previous games). This imposed a limit on the number of feasible repetitions with the three-subject electorates, which is 7. With one-subject electorates no such limit is imposed; this allows for more repetitions to check whether behavior settles down over time. In sessions with one-subject electorates, therefore, we ran the game with 20 repetitions. The design is summarized in Table 1.

The design had two desiderata: (a) root out social preferences; and (b) produce statistically independent observations. Since incumbents are machines, not humans, there is no sense in punishing or rewarding past behavior. In sessions 1 and 2 every electorate is unique due to an absolute stranger design. There is no sense in trying to punish or reward other subjects for previous play, since this can not possibly have any disciplining effects that the subject may benefit from (he or she does not meet future subjects who have met subjects that he or she has punished or rewarded in the past). In sessions 1 and 2 majority decision ensures that all subjects in the same electorate earn the same amount in a specific game. No subject belonging to the same electorate is therefore ever ahead or behind any other subject. In the decisive voter treatments (sessions 3 and 4) no information on other subjects earnings was made available. Social preferences based on inequality aversion (or more generally, preferences for final earnings distributions) should consequently have no effect in the experiment. Due to the use of an absolute stranger design in sessions 1 and 2 we can also be confident that observations of electorates are statistically independent.

The experiment was programmed in z-tree (Fischbacher, 1999). After subjects had entered the lab, instructions were read out loud (to ensure public knowledge of the structure of the interaction). Each session started with two non-paying test games to

familiarize subjects with the game and the screens. All communication between subjects during the experiment took place through the computers. After concluding a session, subjects left the lab one at a time and received their earnings.

The experimental 'schillings' were converted at a fixed rate to NOK at the conclusion of the experiment, and the subjects were paid in cash. There was no show-up fee, and the average pay over all treatments was 207 NOK.[10] A session lasted on average 45 minutes, so average pay is slightly above the going optional hourly wage of a typical BA student.

In equilibrium voters should oust first-stage incumbents that do not allocate tax revenues to public production, and (given the update) should keep first-stage incumbents that do allocate tax revenues to public production. Behaviorally, one would expect voters to be quite good at keeping incumbents after observing 50 schillings of first-stage public production, and to throw incumbents out after observing 0 schillings of first-stage production. The case of a first-stage production equal to 25 schillings could either be due to a good incumbent facing a low productivity draw, or to a bad incumbent facing a high productivity draw (and mimicking a good incumbent). The conjecture is that the size of the update will determine the extent to which voters keep the incumbent when first-stage production was 25 schillings. One should also expect voters to approach equilibrium over time, possibly through non-rational forms of learning. Indeed non-rational forms of learning may arguably have a greater impact on behavior for marginal updates than for substantial updates. Lastly, decisions made by subjects in electorates are expected to be closer to equilibrium than decisions made by subjects operating as decisive voters. Electorates provide a richer learning environment, in which subjects may correct their behavior based on observing whether they were in the minority or not.

4. Results

Results are presented in four sections. First,we provide some descriptive statistics. This is followed by non-parametric tests for the effects of learning in groups versus learning alone, and for the effects of playing early games versus late games. Third, we present a set of regressions that evaluate the effects of two non-rational learning rules; fictitious play and simple payoff reinforcement. Lastly, we explore the effects of payoff reinforcement learning in the experiment by running some simulations.

4.1. Descriptive statistics

Due to the stochastic nature of the game, the distribution of first-stage production is not balanced. Table 2 shows how observations are distributed on treatments and first-stage production (denoted by $P1 = 0$, $P1 = 25$ and $P1 = 50$). Note also that the fraction of non-equilibrium decisions varies between 28 percent and 20 percent in the data, depending on treatments. While larger updates result in more equilibrium behavior for decisive voters, the opposite is the case for voters in electorates.

4.2. Rational learning and group learning

We start by looking at the effects of update size on decisive voters, and on voters in electorates. The relevant data are displayed in Figure 1.

Table 2. Descriptive statistics.

	Session 1	Session 2	Session 3	Session 4
P1 = 0	72	90	190	272
P1 = 25	36	36	182	115
P1 = 50	18	0	28	13
Share of equilibrium decisions	.80	.75	.72	.75

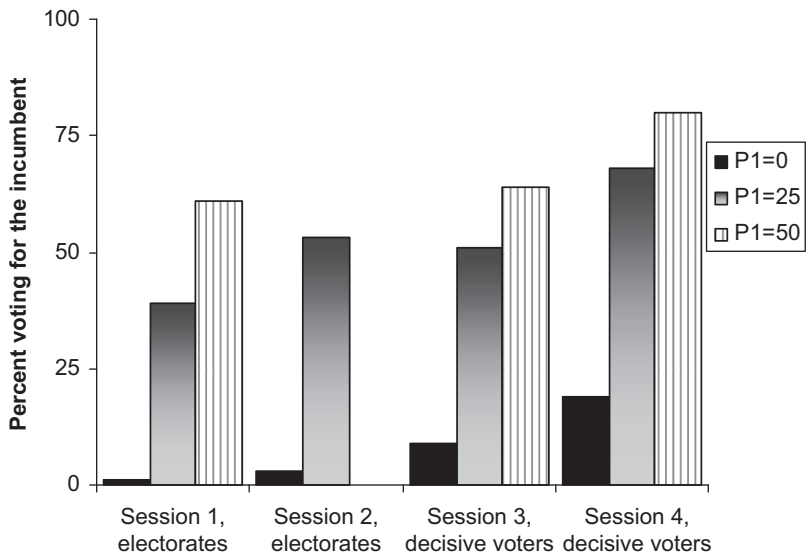

Figure 1. Decisive voters and electorates; games 1–7.

The bars show the fraction of decisive voters and the fraction of voters in electorates who decided to keep the incumbent, contingent on first-stage production.[11] Due to few observations at first-stage production level equal to 50, we do not comment on the patterns in this state.

When P1 = 25, decisive voters are generally more likely than voters in electorates to keep incumbents. This holds for both marginal update (12 percentage points difference) and substantial update (15 percentage points difference). However, none of these differences are statistically different from zero in a two-sided Mann–Whitney U-test ($z = -1.12, p > 0.26$ for marginal update; $z = -1.36, p > 0.17$ for substantial update).

In both group treatments (decisive voters versus electorates) increasing the update (from marginal to substantial) increases the fraction of incumbents kept at P1 = 25. In the treatment with decisive voters the increase is 17 percentage points, against 14 percentage points in the treatment with electorates. However, only the former difference is statistically different from zero in a two-sided Mann–Whitney U-test ($z = 1.68, p > 0.09$ for decisive voters; $z = -1.17, p > 0.24$ for electorates).

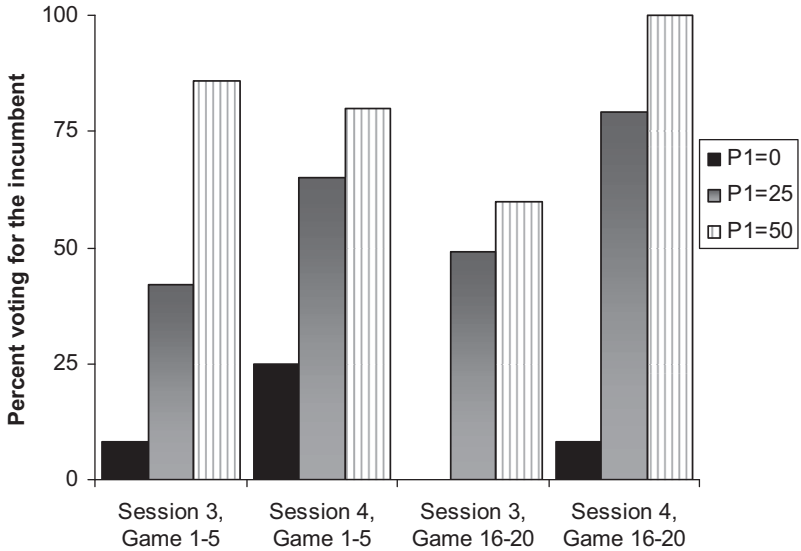

Figure 2. Decisive voters; games 1–5 and games 16–20.

On the other hand, decisive voters are significantly worse at ousting incumbents when P1 = 0, than voters operating in electorates. This holds for both marginal and substantial updates ($z = 1.98, p > 0.05$ for marginal update; $z = -3.28, p > 0.000$ for substantial update).

We turn now to the effects of update size in early versus late games. The analysis is confined to voting decisions of decisive voters, since a greater number of games was played in this group treatment. The relevant data are displayed in Figure 2.

We focus on voting contingent on having observed P1 = 25. The difference between substantial and marginal update in the first five games is 0.24, compared to 0.26 in the five last games. These differences are significant in both cases ($z = 1.89, p > 0.06$ first five games; $z = 2.48, p > 0.01$ last five games).

Now consider learning effects. The difference in votes to keep the incumbent between the last five and the first five games is 8 percentage points when the update is marginal, and 14 percentage points when it is substantial. None of these movements towards equilibrium, however, is significantly different from zero ($z = -0.73, p > 0.46$ for marginal update; $z = -1.20, p > 0.23$ for substantial update).

As is also evident from Figure 2, more incumbents are ousted after P1 = 0 in the last five games (for both marginal and substantial update), than in the first five games. This pattern is significantly different from zero at conventional levels ($z = 1.77, p > 0.08$ for marginal update; $z = -2.60, p > 0.01$ for substantial update)

All in all, these non-parametric tests indicate that allowing for substantial updating does facilitate movement towards equilibrium, but not greatly, and not always in statistically significant ways. Second, membership of an electorate does not seem to induce group learning that improves the ability to make equilibrium choices. Third, the learning effects of decisive voters are modest, and not significantly different from zero at conventional levels.

Table 3. Absolute deviations between registered beliefs and equilibrium beliefs. Averaged over registered beliefs.

		Decisive voters	
	Substantial update	Marginal update	Mann–Whitney *U*-test
P1=0	12.1	8.3	z=–2.56, $p > 0.01$
P1=25	25.9	28.6	z=0.02, $p > 0.99$
P1=50	30.4	37.5	z=0.53, $p > 0.57$
		Electorates	
	Substantial update	Marginal update	Mann–Whitney *U*-test
P1=00	11.3	1.0	z=–4.65, $p > 0.00$
P1=25	25.0	23.4	z=–1.61, $p > 0.11$
P1=50	–	27.2	–

Table 3 displays the absolute deviance between registered beliefs and equilibrium beliefs, contingent on update size, first-stage production, and group treatment. As can be seen, the differences between decisive voters and voters in electorates are small for P1 = 25. The average deviation between registered beliefs and equilibrium beliefs, however, is quite large for this production level. For decisive voters average mistakes at P1 = 25 increases on transition from marginal to substantial update. For voters in electorates the opposite is the case. As can be seen, the first effect is not significantly different from zero in a two-sided test, while the last one is. For observed P1 = 0, average mistakes are smaller. However, these mistakes grow with the size of the update for both decisive voters and voters in electorates, and the differences are significantly different from zero at conventional levels. All in all, the pattern of registered beliefs seriously challenges the conjecture that subjects form beliefs in accordance with the perfect Bayesian equilibrium of the electoral agency model. It also challenges the conjecture that Bayesian belief formation is more pronounced when observations give rise to substantial updates.

Can behavior based on payoff reinforcements and fictitious play updating help us understand behavior better?

4.3. Non-rational learning

The fictitious play update is a continuous variable constructed as follows. The belief that one is facing a good incumbent in game t equals

$$Pr_t(G) = \frac{w_{t-1}(G)}{w_{t-1}(G) + w_{t-1}(B)} \tag{2}$$

Let $w_0(G) = \pi = 0.2$ and $w_0(B) = (1-\pi) = 0.8$. The following counting rules are used. (i) If P1 = 50 (0) this counts as a good (bad) incumbent. (ii) If P1 = 25 and the incumbent was reelected, a positive (negative) stage two production counts as a good (bad) incumbent. (iii) If the incumbent was ousted a positive (negative) stage two production counts

Table 4. Initial choice probabilities (initial attractions).

First-stage production	PI = 00	PI = 25	PI = 50
Keep incumbent	.167	.481	.600
	(9.19)	(26.46)	(33.00)
Oust incumbent	.833	.519	.400
	(45.82)	(28.55)	(22.00)

as a good (bad) incumbent. Note that two counts of a good politician, two counts of a bad politician, or one count of a bad and one count of a good politician, are all possible if the stage one incumbent was ousted. Employing these counting rules, the weighting function follows the formula $w_t(G) = w_{t-1}(G) + f, f = (1, 2)$ if conclusive evidence of a good incumbent in game t was observed, and $w_t(G) = w_{t-1}(G) + 0$ if no such evidence was observed in game t. The weighting function for a bad incumbent ($w_t(B)$) is defined similarly. Now define a dummy variable that takes the value 1 if $Pr_t(G|P1 > 0) > 0.2$ or $Pr_t(G|P1 = 0) \leq 0.2$, and zero otherwise. In words, the dummy takes value 1 if fictitious beliefs favor equilibrium actions, and zero otherwise. This dummy is denoted $d(Fictitious)$.

The payoff reinforcement variable is calculated on state dependent actions. That is, reelecting or throwing the incumbent has numerical attractions that depend on the information set the subject is in. There are six attractions, given by the two possible actions (reelect or oust) in each of the three states (first-stage production 0, 25, or 50, respectively). Let $q_{i,t}(a, s)$ denote the attraction action a has for player i at time t, given that the realized state was s. Let the payoff to player i of choosing action a in state s be b. The attraction of action a in state s is updated according to the following rule: $q_{i,t+1}(a, s)$=$q_{i,t}(a, s) + b$. The probability that player i chooses action a' in state s at time t is simply

$$p_{i,t}(a', s) = \frac{q_{i,t}(a', s)}{\sum_{a \in A}[q_{i,t}(a, s)]} \tag{3}$$

To facilitate interpretation of reinforcement toward equilibrium, the variable used is coded as follows: $Reinforcement = (p_{i,t}(Keep, P1 > 0)$ and $(1 - p_{i,t}(Keep, P1 = 0))$.

The sketched learning rule raises two important questions. How are initial attractions to be determined (i.e. in period $t = 1$)? At what level should the 'strength' of initial attractions be set? The strength of initial attractions is defined as $\sum_{a \in A}[q_{i,t}(a, s)]$.

Following Roth and Erev (1995), initial attractions are estimated from data using only the first two games, with the strength of initial attractions set at the same order of magnitude as the maximal value of periodic production in the game, i.e. 55. Table 4 provides the fraction of votes for and against the incumbent in the two first games, in the three different states. These correspond to the estimates of (state-contingent) initial choice probabilities. (The initial attractions follow readily from this estimate, and the strength of attraction.)

Table 5 presents a set of logistical regression. We estimate the (log odds) that decisions are in equilibrium. That is, the dependent is a dummy that takes the value zero if first-stage production was zero and the incumbent was voted out of office, or if first-stage production was positive and the incumbent was kept. Otherwise the dependent

Table 5. Dependent: correspondence with equilibrium. Logistical regressions. Decisive voters. Coefficients (*p*-values).

	Model 1a	Model 1b	Model 2a	Model 2b
Sex	.13	.12	.13	.11
	(.66)	(.37)	(.63)	(.38)
d(Update)	−1.12	−1.19	−1.11	−1.14
	(.003)	(.006)	(.003)	(.005)
d(P1=25)	−2.83	−3.07	−1.12	−1.52
	(.000)	(.000)	(.023)	(.003)
d(P1=50)	−2.16	−2.28	−1.14	−1.38
	(.000)	(.000)	(.044)	(.014)
d(P1=25)×d(Update)	1.62	1.74	1.28	1.42
	(.000)	(.000)	(.005)	(.002)
d(P1=50)×d(Update)	2.08	2.26	2.15	2.34
	(.030)	(.017)	(.024)	(.012)
d(Fictitious)	–	–	.12	.13
			(.545)	(.535)
Reinforcement	–	–	4.51	3.81
			(.000)	(.000)
Constant	2.80	3.03	−1.18	−.044
	(.000)	(.000)	(.170)	(.636)
Deviance	749.8	650.8	720.4	650.3
	(.000)	(.000)	(.000)	(.000)
Random effects	NO	YES	NO	YES
Subjects	–	.53	–	.33
Games	–	.15	–	.14
Number of observations	800	800	800	800

has the value 1. The following explanatory variables are used: subject's sex; a dummy variable ($d(Update)$) that takes the value zero in session 3 (marginal updating) and 1 in session 4 (substantial updating); dummy variables for first-stage production equal to 25 ($d(P1 = 25)$) and 50 ($d(P1 = 50)$) respectively; the dummy that captures fictitious play; the variable that captures reinforcement learning. In addition, the production dummies are interacted with the dummy for sessions ($d(Update)$).[12]

The first two models (1a and 1b) do not account for non-rational learning rules. Results broadly confirm one of the findings shown in Figures 1 and 2. Consider model 1b, where we control for random effects of subjects and game (i.e. time). The regression says that going from marginal to substantial update at P1 = 25 drives behavior toward equilibrium. The combined effect (taking the interactive term into account) is 0.55, corresponding to an increase in the probability of keeping the incumbent of 13 percentage points if sex is set to one (from 52 percent for marginal update, to 65 percent for substantial update). This combined effect is also close to significantly different from zero at the 10 percent level, with a *p*-value of .106.[13] A similar analysis of the update effect at P1 = 50 reveals a positive effect that is far from significantly different from zero at conventional levels.

The two last models (2a and 2b) control for non-rational learning rules. Consider model 2b, where random effects are controlled for. Going from marginal to substantial update at P1 = 25, pushes behavior toward equilibrium again. The combined effect in this case is 0.26, or roughly half before control for non-rational learning rules. If reinforcement is held at its mean value, while the dummies for fictitious play, update, and sex are set to 1, it corresponds to an increase in the probability of equilibrium voting of 7 percentage points. However, this combined effect is far from significantly different from zero at conventional levels, with a p-value of .38. The dummy for fictitious play is small in magnitude, and clearly insignificant.

However, the reinforcement variable has a positive, large, and strongly significant effect on the probability of making equilibrium choices. Substantially, the probability of making an equilibrium choice after observing P1 = 25 is 45 percent if reinforcement is set at its mean minus two standard deviations, it is 68 percent if reinforcement is set at its mean (0.58), and 85 percent if reinforcement is set at its mean plus two standard deviations. In the calculation, the dummies for fictitious play, update, and sex are once again set to 1. The standard deviation of reinforcement for these controls is 0.13.[14]

The main message boils down to this: the last trace of rational updating disappears when it is controlled for the very simplistic (almost Pavlovian) learning rule in which past payoffs reinforce current choices.[15]

4.4. Long-run effects of payoff reinforcement

We explore in this section the long-run behavior of payoff reinforcement learning in the game. Figures 3 shows simulation results for substantial update (left panel) and marginal update (right panel), when behavior is driven by payoff reinforcement only. The randomness and payoffs of these simulations are identical to those of the experiments. We used the same initial choice probabilities, and the same strength of attractions as in the regressions. An individual that plays a sequence of 1000 independent games is simulated. The simulations are averages of 10,000 draws of such sequences. The figure maps the average fraction of decisions in equilibrium (y-axis) for the sequence of games (x-axis) and for each of the three states (top to bottom: P1 = 0, P1 = 25 and P1 = 50). Ninety percent confidence intervals are attached to the behavioral paths (grey curves).

The main insight from these simulations is that behavior moves (asymptotically) toward equilibrium (but never quite reaches it for P1 $>$ 0) in the substantial update condition (Figure 3, left panel). This is very different from the marginal update condition (Figure 3, right panel), in which behavior contingent on observing P1 = 25 diverges (slowly) from equilibrium. The reason for this is quite simple: in the substantial update condition keeping the incumbent provides a (posterior) probability (after observing P1 = 25) of 59 percent for a positive second stage payoff. In the marginal update condition keeping the incumbent provides a (posterior) probability (after observing P1 = 25) of only 23 percent for a positive second stage payoff. So, high posterior probability of a good incumbent reinforces the choice of keeping the incumbent, and pushes behavior towards equilibrium.

In the event then, observing higher levels of equilibrium behavior at P1 = 25 for substantial updates by no means implies that the mechanism (perfect Bayesian equilibrium) identified in the basic electoral agency model is at work. Exactly the same conditions – clear and strong Bayesian updating – will also reinforce behavior towards equilibrium in

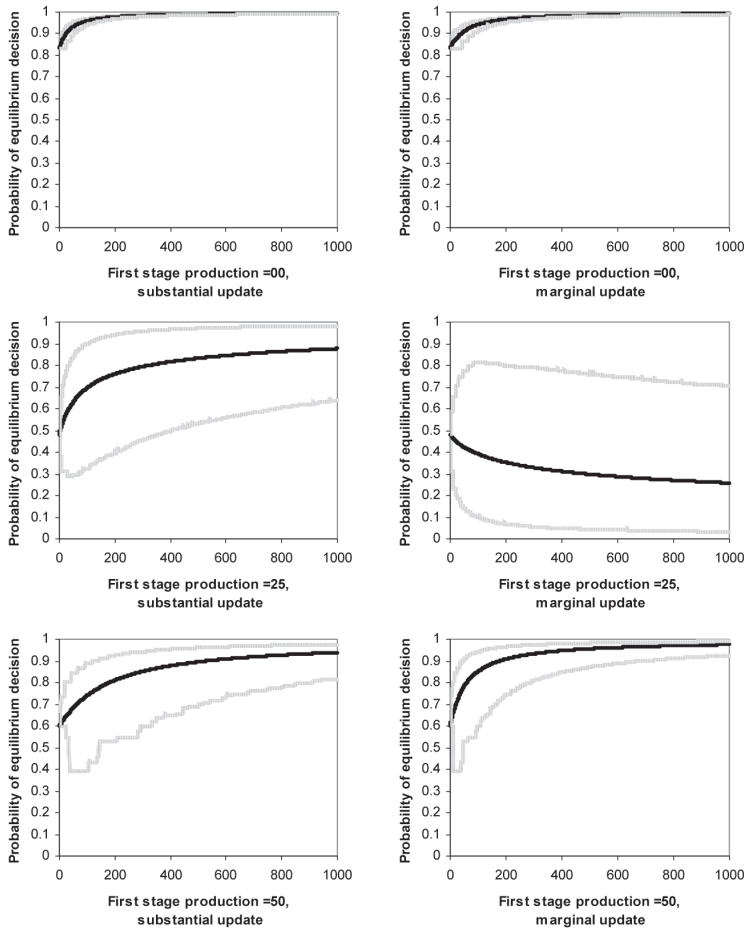

Figure 3. Simulation results; substantial update left panel, marginal update right panel (90 percent confidence intervals in grey)

a Pavlovian manner. The confidence intervals around the behavioral path at P1 = 25 indicate the experiment had too few rounds for this effect to be pronounced. Still, as noted, the regressions weed out all trace of Bayesian updating after control for reinforcement learning.

In the Appendix we show evolving choice probabilities based on simple payoff reinforcement for the 40 subjects in sessions 3 (marginal update) and 4 (substantial update). As can be seen, the choice probabilities, given production equal to 25 schillings, tend to converge more towards equilibrium in session 4 than in session 3.

5. Limits on sophisticated politicians

Assume voter behavior is fully described by the simple payoff reinforcement rule. Assume also that the utility of bad incumbents increases linearly in rents. Consider a bad type of incumbent who has drawn high productivity. If she takes maximal rents, her

profit is 50 schillings off the bat. According to our findings, in the long run voters will not reelect incumbents with zero value in their public production. Thus the incumbent's profit is 50 schillings. Alternatively, the bad-type incumbent might mimic a good-type incumbent facing a low productivity draw. In this case, her first period rent is 25 schillings, while her second period expected rent is $Q\delta 50$ schillings, where Q signifies the reelection probability in this case, and $\delta < 1$ is a discount factor. The bad incumbent only mimics if $Q \geq \frac{1}{2\delta}$. For the sake of the argument, simplify by letting δ be arbitrarily close to unity.

Consider first the case pictured in Figure 3, left panel. In this case, the mimic condition is always satisfied, and behavior approaches the equilibrium of the electoral agency model in the long run.

Consider now the case pictured in Figure 3, right panel. Assume the mimic condition is not satisfied. Given first-stage production equalling 25, all second-stage outcomes following reelection will be good ones. This drives up reelection probability, Q, through the payoff reinforcement rule. Eventually, Q will reach the threshold of 0.5 where bad-type incumbents start to mimic good ones. A fraction $((1 - q)(1 - \pi))$ of bad second-stage outcomes will follow reelection after observing first-stage production equalling 25 schillings. This lowers Q through the payoff reinforcement rule. If the mimic condition is satisfied initially, a mirror argument ensures that Q is driven towards $\frac{1}{2\delta}$ from above.

In the upshot, two forces limit the ability of bad incumbents to take advantage of non-rational voters. First, payoff reinforcement learning is not exposed to exploitation in the long run if low productivity is common enough, and/or bad incumbents are rare enough. Second, even if payoff reinforcement learning can be exploited by bad incumbents, learning will take place among voters and sophisticated incumbents will adjust to this learning. The result is a selection dynamics that places a cap on exploitation of non-rational voters. This constraint will be weaker the more bad incumbents value immediate rents. In the stable state, behavior deviates from the equilibrium of the electoral agency model, but not as much as with unsophisticated incumbents.

6. Conclusion

In his great book on capitalism and democracy, Joseph Schumpeter (1996: 262) notes how '...the typical citizen drops down to a lower level of mental performance as soon as he enters the political field. He argues and analyzes in a way which he would readily recognize as infantile within the sphere of his real interests. He becomes primitive again. His thinking is associative and affective ... [This] may prove fatal to the nation.' Current models of electoral agency derive from a radically different idea; utilizing the standard assumptions of rational and self-regarding behavior.

Our experimental design has sought to eliminate the impact of social preferences and intentions ('affections') on voting behavior, in order to focus more clearly on non-rational ('associative') forms of learning. We found that simple payoff reinforcement learning explains subjects voting behavior well in our electoral agency experiment.[16]

Our simulations indicate that situations in which Bayesian updating is strong and clear also make payoff reinforcement push behavior towards the equilibrium (which is 'good for the nation' since selection and discipline tends towards optimality). When

Bayesian updating produces a less clear-cut answer, on the other hand, payoff reinforcement pushes behavior away from equilibrium (which may be 'fatal to the nation', since selection and discipline does not work optimally). This movement away from equilibrium is limited by sophisticated incumbents responding optimally to non-rational voters. The less farsighted bad incumbents are, the further from equilibrium the behavior stabilizes.

For reasons such as these, we believe that observing voting patterns that approach equilibrium behavior in field data or, for that matter, in experimental data does not justify strong conclusions about data being generated by a perfect Bayesian equilibrium. Our results are limited to the basic electoral agency model. However, the core mechanism of this simple model is shared by more complex electoral agency models. Exploring the implications of associative and affective thinking more systematically could well benefit the development of electoral agency models.

Appendix

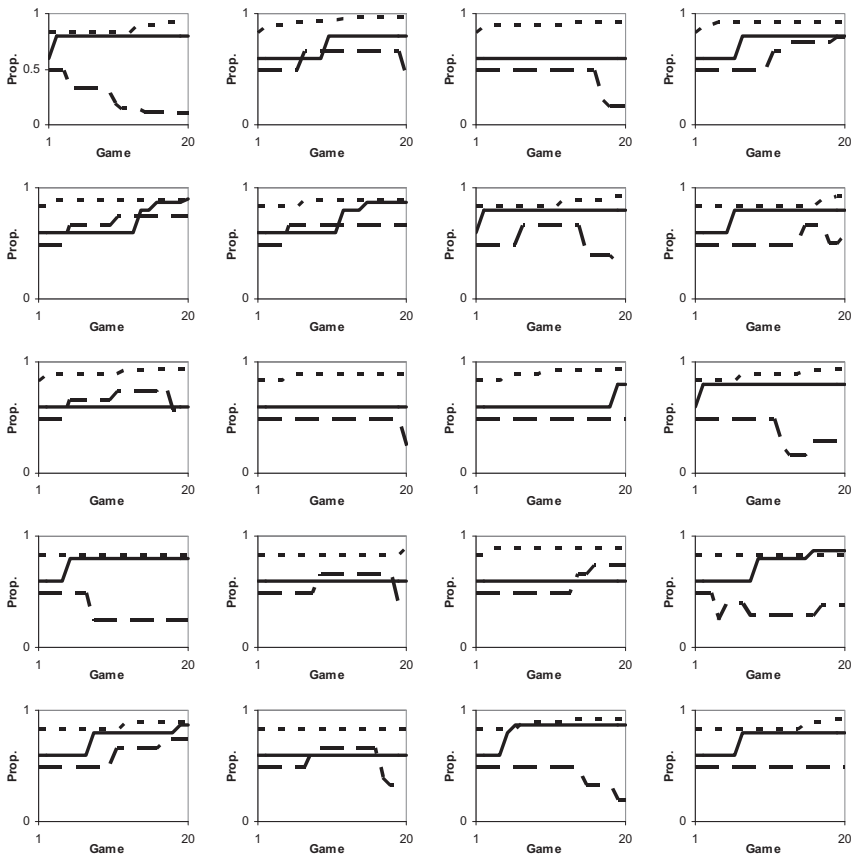

Figure 4. Figures A1 – A20: Marginal update. Dotted line, production = 0, dashed line, production = 25, solid line, production = 50.

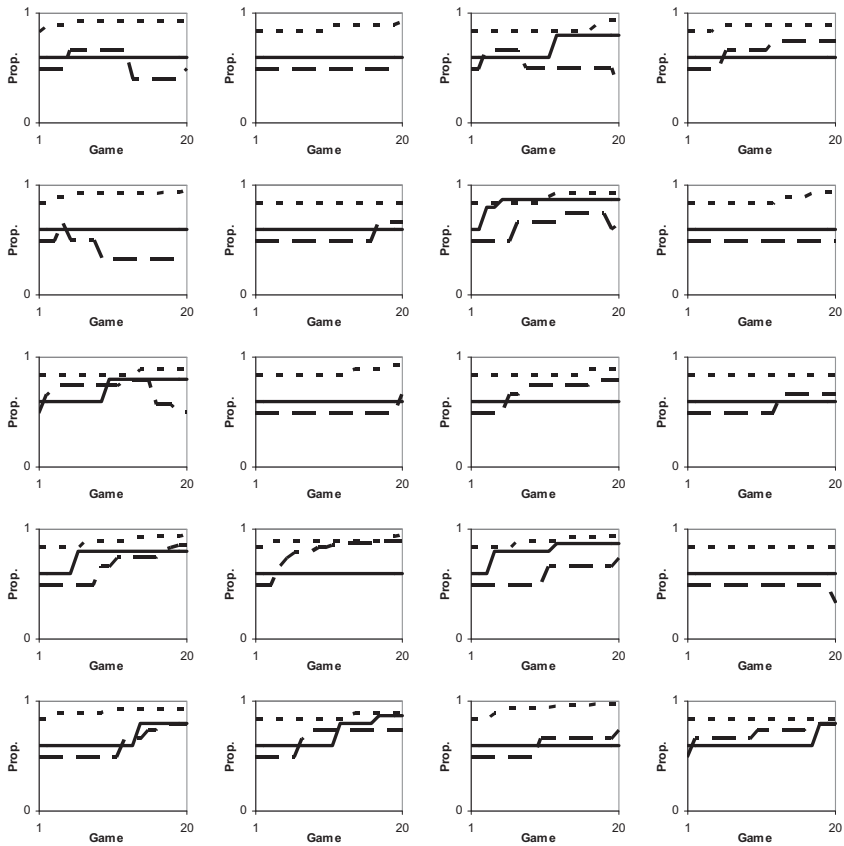

Figure 5. Figures A21 – A40: Substantial update. Dotted line, production = 0, dashed line, production = 25, solid line, production = 50.

Funding

The Norwegian Business School has generously financed the project.

Acknowledgements

Useful comments were provided by participants at the Oslo Seminar in Behavioral and Experimental Economics, November 2008; the 3rd Nordic Conference on Behavioral and Experimental Economics, November 2008; and the Workshop on Electoral Politics in the Lab and the Field arranged by the Department of Political Science/University of Oslo, May 2011.

Notes

1. First-generation models of electoral agency, such as those of Barro (1973) and Ferejohn (1986), focus exclusively on discipline, since all politicians are 'bad' (in the sense of maximizing a combination of rents and office).
2. The use of automatons is common in election experiments, and abundant in the literature on positioning in two-candidate contests (in which voters are habitually programmed to vote for

the candidate closest to their own ideal point). For an overview see Ordeshook (1997). A more recent example can be found in Aragones and Palfrey (2005), which we discuss below.

3. We do not provide information on earnings in other groups; inequality concerns based on the population of subject do not, therefore, enter into the frame.
4. Beginning in the late 1980s, a number of papers have explored convergence towards the median voter in two-candidate majoritarian elections (see Ordeshook 1997 for an overview). From the early 1990s onwards, a series of studies demonstrated how various pre-election signals can help voters eliminate Condorcet losers in three-candidate contests under single-member majority rule (see Rietz, 2009, for an overview). There is also a growing experimental literature on turnout in elections (a brief overview is provided by Sonnemans and Schram, 2009).
5. In the manner suggested by McKelvey and Ordeshook (1985).
6. A possible weakness in the Dasgupta and Williams (2002) design is its failure to (attempt to) control for social preferences (which we know are forceful drivers of behavior in a large class of other experiments).
7. Instructions for participants in the experiment are available at: http://home.bi.no/a0111218/EA_JTP_Instructions.pdf.
8. Persistent in the sense of persisting through the election.
9. This eliminates a hybrid equilibrium, in which voters and incumbents randomize over pure actions.
10. The exchange was reduced in long sessions, to produce an expected pay of 200 NOK in all sessions.
11. Results remains qualitatively similar if we instead analyze average voting decisions of subjects.
12. We also ran regressions in which we let the non-rational belief formation variables interact with the update dummy. The exercise did not affect our results.
13. The t-test in this case is a joint test of $d(Update)$, $d(P1 = 25)$, and the interaction between these two variables. Details about these tests can be found in Kam and Franzese (2007).
14. All the conclusions drawn so far would also follow had we interpreted regressions 1a and 2a instead of 1b and 2b.
15. Comparing regressions 1b and 2b, we also see that random time effects are fairly constant (0.15 compared to 0.14), and that the main difference is captured by within-subjects variation as we control for non-rational learning rules (0.52 compared to 0.33). This should come as no surprise. In regression 2b we introduced randomly generated histories at the subject level, which gave rise to random variation in the non-rational learning rules.
16. The explanatory force of simple payoff reinforcement has been well documented for market games, ultimatum bargaining and contribution games in Roth and Erev (1995).

References

Alesina A, Baqir R and Easterly W (1999) Public goods and ethnic divisions. *Quarterly Journal of Economics* 114(4): 1243–1284.

Aragones E and Palfrey T (2005) The effect of candidate quality on electoral equilibrium: An experimental study. In: Austen-Smith D and Duggan J (eds) *Social Choice and Strategic Behavior: Essays in Honor of Jeffrey Scot Banks*. Springer: Berlin.

Austen-Smith D and Banks J (1989) Electoral accountability and incumbency. In: Ordeshook P (ed.) *Models of Strategic Choice in Politics*. Ann Arbor, MI: University of Michigan Press.

Banks J and Sundaram R (1993) Adverse selection and moral hazard in a repeated elections model. In: Barnett W, Schofield N and Hinich M (eds) *Political Economy: Institutions, Information, Competition and Representation.* New York: Cambridge University Press.

Barro R (1973) The control of politicians: An economic model. *Public Choice* 14: 19–42.

Besley T (2006) *Principled Agents? The Political Economy of Good Government*. Oxford: Oxford University Press.

Besley T, Persson T and Sturm D (2010) Political competition, policy and growth: Theory and evidence from the US. *Review of Economic Studies* 77(4): 1329–1352.

Bolton G, Brandts J, Katok E et al. (2009) Testing theories of other-regarding behavior: A sequence of four laboratory studies. In: Plott C and Smith V (eds) *Handbook of Experimental Economics Results. Volume 1.* Amsterdam: North-Holland.

Brown G (1951) Iterative solutions of games by fictitious play. In: Koopmans TC (ed.) *Activity Analysis of Production and Allocation*. New York: Wiley.

Camerer C and Ho T (1999) Experience weighted attraction learning in games: Estimates from weak link games. In: Budescu D, Erev I and Zwick R (eds) *Games and Human Behavior: Essays in Honour of Amnon Rapoport*. Mahwah, NJ: Lawrence Erlbaum.

Caplan B (2007) *The Myth of the Rational Voter: Why Democracies Choose Bad Policies*. Princeton, NJ: Princeton University Press.

Dasgupta S and Williams K (2002) A principal–agent model of elections with novice incumbents: Some experimental results. *Journal of Theoretical Politics*14(4): 409–438.

Downs A (1957) *An Economic Theory of Democracy*. New York: Harper & Row.

Easterly W and Levine R (1997) Africa's growth tragedy: Policies and ethnic divisions. *Quarterly Journal of Economics* 112(4): 1203–1250.

Erev I and Roth A (1995) Learning in extensive form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior* 8: 164–212.

Fearon J (1999) Electoral accountability and the control of politicians: Selecting good types versus sanctioning poor performance. In: Przeworski A, Stokes S and Manin B (eds) *Democracy, Accountability, and Representation*. Cambridge: Cambridge University Press.

Fehr E (2009) Social preferences and the brain. In: Glimcher P, Camerer C, Fehr E, et al. (eds) *Neuroeconomics: Decision Making and the Brain*. London: Academic Press.

Ferejohn J (1986) Incumbent performance and electoral control. *Public Choice* 50: 5–26.

Fischbacher U (1999) *Z-tree. Zurich Toolbox for Readymade Economic Experiments*. Zurich: University of Zurich.

Gigerenzer G, Hertwig R, Hoffrage U et al. (2009) Cognitive illusions reconsidered. In: Plott C and Smith V (eds) *Handbook of Experimental Economic Results. Volume 1*. Amsterdam: North-Holland.

Hayek F (1979) *Law Legislation and Liberty. Volume 3: The Political Order of a Free People*. Chicago: University of Chicago Press.

Helland L and Sørensen R (2011) Biased competition, popularity shocks, and government inefficiency. Working Paper, BI Norwegian Business School.

Houser D, Morton R and Stratman T (2008) Turned off or turned out? Campaign advertising, information and voting. Discussion Paper, Interdisciplinary Center for Economic Science, George Mason University.

Kam C and Franzese R (2007) *Modeling and Interpreting Interactive Hypotheses in Regression Analysis*. Ann Arbor, MI: University of Michigan Press.

McKelvey R and Ordeshook P (1985) Sequential elections with limited information. *American Journal of Political Science* 29: 480–512.

Markussen T and Tyran JR (2009) Serving the public interest. Centre for Experimental Economics, University of Copenhagen (unpublished).

Maskin E and Tirole J (2004) The politician and the judge: Accountability in government. *American Economic Review* 94(4): 1034–1054.

Ordeshook P (1997) The spatial analysis of committees and elections. In: Mueller D (ed) *Perspectives on Public Choice*. Cambridge: Cambridge University Press.

Persson T and Tabellini G (2000) *Political Economics: Explaining Economic Policy*. Cambridge, MA: The MIT Press.

Petterson-Lidbom P (2006) Testing political agency models. Working Paper, Department of Economics, Stockholm University.

Popper K (1989) *Conjectures and Refutations: The Growth of Scientific Knowledge*. London: Routledge.

Rietz T (2009) Three-way experimental election results: Strategic voting, coordinated outcomes and Duverger's law. In: Plott C and Smith V (eds) *Handbook of Experimental Economics Results. Volume 1*. Amsterdam: North-Holland.

Riker W (1982) *Liberalism against Populism: A Confrontation between the Theory of Democracy and the Theory of Social Choice*. San Francisco: Freeman.

Roth A and Erev I (1995) Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior* 8: 164–212.

Schumpeter J (1996) [1943]: *Capitalism, Socialism, and Democracy*. London: Routledge.

Sonnemans J and Schram A (2009) Participation game experiments: Explaining voter turnout. In: Plott C and Smith V (eds) *Handbook of Experimental Economics Results. Volume 1*. Amsterdam: North-Holland.

Svaleryd H and Vlachos J (2009) Political rents in a non-corrupt democracy. *Journal of Public Economics* 93: 355–372.

Svensson J (1999) Controlling spending: Electoral competition, polarization and endogenous platforms, IIES/Stockholm University (unpublished).