Agent-Based Models of Strategic Electoral Behavior in Election Forensics*

Samuel Baltz† Kevin McAlister‡ Alejandro Pineda§

Fabricio Vasselai¶ Walter R. Mebane, Jr.‖

March 31, 2018

*Prepared for presentation at the 2018 Annual Meeting of the Midwest Political Science Association, Chicago, IL, April 5–8, 2018. Work supported by NSF award SES 1523355.

†Department of Political Science, University of Michigan, Haven Hall, Ann Arbor, MI 48109-1045 (E-mail: sbaltz@umich.edu).

‡Department of Political Science, University of Michigan, Haven Hall, Ann Arbor, MI 48109-1045 (E-mail: kamcal@umich.edu).

§Department of Political Science, University of Michigan, Haven Hall, Ann Arbor, MI 48109-1045 (E-mail: ajpineda@umich.edu).

¶Department of Political Science, University of Michigan, Haven Hall, Ann Arbor, MI 48109-1045 (E-mail: vasselai@umich.edu).

‖Professor, Department of Political Science and Department of Statistics, University of Michigan, Haven Hall, Ann Arbor, MI 48109-1045 (E-mail: wmebane@umich.edu).

Abstract

How to distinguish frauds from consequences of strategic behavior is the primary challenge for election forensics. Election forensics methods use empirical distributions of turnout and vote choices to identify fraudulent activity. Strategic behavior can affect these distributions. Generally, many types of election forensics methods interpret a multiplicity of modes in election data as indicators of frauds, but strategic behavior induces correlations among electors' behavior that can produce multimodalities. We present agent-based models (ABMS) designed to represent electors who decide how to act in elections based both on their own tastes and on their beliefs about what the other electors will do. In particular we define ABMs designed to match equilibria derived in models of wasted vote logic and of strategic abstention. To facilitate computing pivotal vote probabilities, our ABMs use a Poisson games formulation of the models instead of the fixed-electorate assumptions the formal models originally used. We replicate the formal models' basic equilibria.

1 Introduction

A key challenge for election forensics—the field devoted to using statistical methods to determine whether the results of an election accurately reflect the intentions of the electors—is to be able to distinguish election results caused by election frauds from results produced by strategic behavior or other normal politics (Mebane 2013*a*, 2016). Election forensics studies counts of votes, counts of eligible voters and other traces of an election—preferably at low levels of aggregation such as tallies for each polling station—to produce evidence regarding what happened in the election. By starting with the numerical results and other measures of the voting process election forensics does not address the entirety of an election or address the full range of frauds that are possible (Lehoucq and Jiménez 2002; Lehoucq 2003; Magaloni 2006; Schedler 2006; Levitsky and Way 2010; Minnite 2010; Birch 2011; Hyde and Marinov 2012; Svolik 2012; Wang 2012; Simpser 2013; Stokes, Dunning, Nazareno and Brusco 2013; Norris 2014). For example, if parties are excluded from the ballot, such an action may not produce distinctive patterns in the votes that are cast. But some violations of electoral integrity such as unfair access to campaign resources, wrongfully manipulated voter lists, vote buying, voter intimidation and other coercive actions may produce distinctive patterns in votes that statistics can detect. Ambiguities arise when such patterns might also be produced by strategic voting and other normal political activites such as forming coalitions.

Many methods for trying to detect election frauds have been proposed (e.g. Myagkov, Ordeshook and Shaikin 2009; Levin, Cohn, Ordeshook and Alvarez 2009; Shikano and Mack 2009; Mebane 2010; Breunig and Goerres 2011; Pericchi and Torres 2011; Cantu and Saiegh 2011; Deckert, Myagkov and Ordeshook 2011; Beber and Scacco 2012; Hicken and Mebane 2015; Montgomery, Olivella, Potter and Crisp 2015; Mebane 2016; Rozenas 2017; Ferrari and Mebane 2017; Ferrari, McAlister and Mebane 2018). Methods based on the second significant digits of vote counts have been shown to respond both to normal political activities (strategic behavior, district imbalances, special mobilizations, coalitions) and to frauds (Mebane 2013*a*, 2014). Methods that examine the last digit of vote counts can be fooled if malefactors have sufficient control over the numbers (Mebane 2013*b*). All the methods can effectively identify various kinds of anomalies, but assessing whether the anomalies are due to frauds presents further challenges.

Some of the methods explicitly focus on the modality of election data. Some methods in this vein emphasize that unproblematic elections feature unimodal distributions of turnout and regular flows of votes—the latter are most compatible with assumed unimodal distributions for parties' shares of the votes (Myagkov, Ordeshook and Shaikin 2008, 2009; Levin et al. 2009). Other contributions connect "spiky" (hence multimodal) distributions of turnout and vote proportions to ideas about agents committing frauds in ways that they intend to be detected (Kalinin and Mebane 2011; Mebane 2013*b*; Rundlett and Svolik 2015). The sharpest contribution featuring multimodality is a model proposed by Klimek, Yegorov, Hanel and Thurner (2012) that stipulates a particular functional form according to which frauds occur.

Review of the theory (Borghesi 2009; Borghesi and Bouchaud 2010) that motivates the Klimek et al. (2012) conception suggests, however, that multimodal distributions may be as readily produced by strategic voting, coalitions and other strategic behavior as by election frauds that stem from maleficent activity (for details see Mebane (2016)). That theory, as it has been developed so far, does not imply that the distributions produced by strategic voting and by frauds are the same: currently the theory is not specifically quantitative. But the theoretical ambiguity about the origins of multimodality may carry over to make the parameters of the Klimek et al. (2012) conception ambiguous. Many of the other statistical methods also, on inspection, are crucially triggered by multimodalities and so may be subject to similar ambiguities. An election may appear to have frauds when in fact it has only robust politics featuring strategic activity by electors—by voters and would-be voters.

Here we report on the initial steps of our plan to use agent-based models (ABMs) to produce quantitatively specific realizations of theories of strategic election behavior and of election frauds. We hope to use ABMs to simulate various election systems at realistic scales, with agents that are involved in various kinds of networks and endowed with various distributions of attributes, where agents engage in equilibrium behavior such as is studied in formal models of rational election behavior (e.g. Palfrey and Rosenthal 1985; Cox 1994). We will introduce various kinds of frauds into such systems. Among the questions will be to what extent can statistical methods for election forensics discriminate strategic actions from frauds, and can the methods measure the extent and location of the frauds accurately.

Previous efforts to use simulations motivated by strategic voting ideas to develop data to use to assess election forensics methods have not rigorously imposed equilibrium conditions, but have only roughly tried to approximate expected consequences of such behavior. For instance, simulations have shifted votes between candidates with references to the kind of behavior that happens when voters act according to wasted vote logic (e.g. Mebane 2013*a*). Perhaps such approaches capture some important qualitative aspects of what happens when voters act that way, but there is no proof of that and there is little reason to believe the quantitative details about how many votes are affected are correct. While some formal models have equilibrium ideas with notions of frauds (e.g. Simpser 2013; Rundlett and Svolik 2015), no simulations have attempted to achieve quantitative precision about what happens.

Here we take only the first steps toward such a plan. We use ABMs to replicate equilibria obtained in two important formal models of elector behavior: the Palfrey and Rosenthal (1985) model of strategic abstention, as updated by Demichelis and Dhillon (2010) to include aspects of learning; and the Cox (1994) model of wasted vote behavior. In both instances we modify the models to use the conception of Poisson games (Myerson 1998, 2000). Others have studied election equilibria using the Poisson games formulation (e.g. Bouton and Gratton 2015). We take a Poisson game approach to make it feasible to compute pivotal voting probabilities (Skellam 1946).

Game Theory and Agent Based Models

Social science is not just concerned with individual behavior, but also, with how individual *interactions* aggregate to larger scale events (Axelrod and Tesfatsion 2006). To this end, agent based models allow us to study systems with the following characteristics: "(1) the system is composed of interacting agents; and (2) the system exhibits emergent properties, that is, properties arising from the interactions of the agents that cannot be deduced simply by aggregating the properties of the agents" (Axelrod and Tesfatsion 2006). Indeed, the strength of ABMs is their ability to capture interdependent human behavior (Bruch and Atwell 2015; Epstein 1999).

That is, individuals behave in a social context that is defined by the collective behavior of individuals. There is feedback between micro level behavior on the one hand, and macro level context on the other: "In the short run, individuals respond to their environments; in the longer run, the accumulation of individuals choices or behavior changes the environment" (Bruch and Atwell 2015). Both game theory and agent based models attempt to capture this interdependence; player strategy is contingent on other players' actions and agent behavior is determined, in part, by interactions with other agents.

It is worth noting the similarities and differences between game-theoretic models and ABMs. Both techniques assume an explicit payoff function, focus on micro-level behavior, and isolate specific components or parameters while holding details fixed (Johnson 1999; de Marchi and Page 2014). Computational models focus on dynamics, while game theory focuses on equlilibria (de Marchi and Page 2014). Both types of models constrain behavior, but while game theorists do so to make equilibria analytically tractable, computational models draw constraints from cognitive, or decision-making research (de Marchi and Page 2014; Kim, Taber and Lodge 2010). Rational choice is included among the cognitive heuristics found in ABMs and as such, game theoretic and agent based models are not inherently opposed (Axelrod 1997; Johnson 1999; de Marchi and Page 2014). The immediate advantage of ABMs is the ability to move beyond the rational choice assumption to model behaviors that are analytically intractable (Axelrod 1997).

Both techniques have mechanisms to study changes in parameter space. For game theory, this is comparative statics. For ABMs, we relax assumptions about micro level behavior to test if macro level outcomes occur as expected (Epstein 1999). This is invaluable for theory testing. Robustness in a computational model is important, as we must avoid "brittleness." That is, equilibria should be invariant to changes to initial parameters (Holland 1983). Additionally, in replicating formal theoretic models in an ABM framework, we find that computational power *complements* mathematical precision, leveraging the comparative advantage of each technique to gain greater insight than either one alone can provide (Miller and Page 2004).

This synthesis of technique—analytic on the one hand and agent-based on the other—parallels behavioral game theory, an approach that also has tried to solve the "paradox of voting" (Downs 1957; Bendor and Diermeier 2003). Previous research on elections employs adaptive rationality, whereby agents learn not by optimizing utility, but rather, by trial and error. This is implemented via reinforcement learning, where agents pursue actions that have previously given satisfaction and inhibit those that have proven unsatisfactory. (Kollman, Miller and Page 1992; Bush and Mosteller 1955; Bendor and Diermeier 2003). Note that previous work has focused on *parties* as agents—acting either ideological or office-seeking—to explore how polling affects party strategy (Kollman, Miller and Page 1992). This project focuses on *voters* as agents to explore how perceptions of electoral closeness affect turnout.[1]

1.1 Strategic Abstention

Palfrey and Rosenthal (1985) consider the classic instrumental formulation for the net

[1]We recognizes that some distinguish equation based models (EBMs) from ABMs (Parunak, Savit and Riolo 1998). EBMs begin with equations that capture relationships between observable characteristics (party preference, for instance), while ABMs begin with behaviors that define how individuals *interact* with one another (Parunak, Savit and Riolo 1998). The literature is divided on whether this differentiation makes sense; for any given computational model, there is a set of equations that could (theoretically) produce the same result (Bonabeau 2002; Epstein 1999). For the purposes of this paper, we do not differentiate.

benefits R from voting in an election with two alternatives:

$$R = pB - C + D\,, \tag{1}$$

where p is the probability that an individual's vote decides the election outcome (is "pivotal"), B is the difference in benefits for the individual if the individual's more preferred choice wins instead of the alternative, C is the cost of voting and D is the direct benefit from voting. Neither C nor D depend on the election outcome but depend only on the individual deciding to participate in the election by voting. Palfrey and Rosenthal (1985) demonstrate the existence of equilibria with positive turnout given a variety of assumptions about the distributions of costs and uncertainty. Palfrey and Rosenthal (1985) argue that turnout, when positive, is usually low.

Demichelis and Dhillon (2010) consider symmetric equilibria given the same net benefits and structure as Palfrey and Rosenthal (1985), while adding learning dynamics. Demichelis and Dhillon (2010) normalize (1) by setting $B = 1$, in which case a necesssary condition for an individual to vote is $p - 2c \geq 0$, where $c = C - D$. Letting q denote the symmetric mixed strategy for all players and using $g(q, M)$ to denote the probability of being pivotal, Demichelis and Dhillon (2010) suppose that q changes over time according to

$$\frac{dq}{dt} = K(q) \tag{2}$$

where sign $K(q) = \text{sign}(g(q, M) - 2c)$. Demichelis and Dhillon (2010) investigate the dynamic stability of equilibria.

1.2 Wasted Vote Logic

Cox (1994) considers situations where each voter casts a single vote, M alternatives "win" the election by gaining seats and the M alternatives that have the M highest numbers of votes win the seats. Assuming that each voter maximizes expected utility based on pivot probabilities, that utilities are diverse,[2] that the distribution of preferences is common knowledge and that expectations about the election outcome are publicly generated (to the point that Cox (1994, 610) imposes a rational expectations condition), Cox (1994) shows that in the limit, as the number of voters grows, there are two kinds of equilibria. In both equilibria M winning alternatives have the same proportion of the votes. With "Duvergerian" equilibria only the first loser (in $M + 1$-th place) receives a nonnegligible proportion of the votes while other losers get nothing, and with "non-Duvergerian" equilibria the losing alternatives all have the same positive proportion of votes. Cox (1994) notes that the formal findings about exact ties between winning alternatives depend on particular simplifying features of the model. The empirical tests he conducts of the theory are not quite so sharp.

1.3 Pivot Probabilities

A key quantity in the considered models is the probability of an individual voter casting a pivotal vote. Demichelis and Dhillon (2010) and Cox (1994) assume that electorate sizes are fixed and commonly known. This allows computation of pivotal probabilities through exact combinatorial computation or multinomial assumptions. Though this approach can yield results, computation of these probabilities are computationally demanding. In order to derive computationally efficient pivotal probabilities, we rely on an assumption that the electorate size is only known up to the distribution. In particular, we utilize the large Poisson games framework derived by Myerson (2000).

Elector $i \in \{1, \ldots, N_e\}$ has incomplete knowledge about the number of voters in an electorate. Elector i receives information about the number of voters, μ, and assumes that this value is a Poisson distributed random variable with mean N:

$$P(\mu) \sim Pois(N) \tag{3}$$

[2] Cox (1994, 610) assumes that the electorate includes voters with all possible preference orders.

Elector i also receives information about the proportions of voters in the electorate that will take certain actions: for candidate $h \in \{1, \ldots, H\}$, she learns the proportion p_h of voters that are going to cast a vote for h. Let μ_h be the number of voters that vote for candidate h. Given p_h and the properties of Poisson random variables, μ_h is also a random variable such that:

$$P(\mu_h) \sim Pois(N_h)\,, \quad N_h = p_h N\,. \tag{4}$$

In general, we are not concerned with the raw vote counts. Rather, we want to know the probability distribution of vote counts relative to other candidates. Let X and Y be independent Poisson random variables. Then

$$P(X - Y = z) = \mathcal{S}(z; N_x, N_y) = \exp(-(N_x + N_y)) \left(\frac{N_x}{N_y}\right)^{\frac{z}{2}} I_z(2\sqrt{N_x N_y}) \tag{5}$$

where $I_z(\cdot)$ is a Bessel function of the first kind. $\mathcal{S}(\cdot)$ is the Skellam distribution (Skellam 1946).

Assuming Poisson vote counts, there are two situations in which i's vote for party h is pivotal: her vote pushes party h out of a tie and into a winning position or her vote pushes party h into a tie for a winning position. For a single winner election where the candidate with the most votes wins, this implies that her probability of casting a pivotal vote for party h (pv_h) is:

$$\begin{aligned} pv_h = \sum_{j \neq h} & [P(N_h - N_j = 0 | N_h, N_j \geq N_m \ \forall\ m \neq j, h) \\ & + P(N_h - N_j = -1 | N_h, N_j \geq N_m \ \forall\ m \neq j, h)] \\ & \prod_{m \neq j, h} P(N_h - N_m > 0) P(N_j - N_m > 0) \end{aligned} \tag{6}$$

In words, the probability that she casts a pivotal vote is the probability that she breaks or creates a tie times the probability that party h is at least in second place.

Assuming that the distribution of the difference between the number of votes for two parties is independent of the probability that they are in the top two positions,[3] we can define pv_h in terms of Skellam distributions:

$$pv_h = \sum_{j \neq h} [\mathcal{S}(0; N_h, N_j) + \mathcal{S}(-1; N_h, N_j)] \prod_{m \neq j,h} \left(\sum_{w=0}^{\infty} \mathcal{S}(w; N_h, N_m) \right) \left(\sum_{w=0}^{\infty} \mathcal{S}(w; N_j, N_m) \right) \quad (7)$$

This value can be evaluated computationally using Skellam implementations in various statistical libraries.

2 Replications

We present examples of ABMs built on a Poisson games foundation designed to replicate basic results from Demichelis and Dhillon (2010), Palfrey and Rosenthal (1985) and Cox (1994). For Demichelis and Dhillon (2010) our goal is essentially to replicate Figure 1 in their paper. For Demichelis and Dhillon (2010) we stick as closely as we can to the assumptions they used, except for our population of electors being larger than the population they used to compute their Figure 1. Cox (1994) does not provide an analogous specific replication target, so we illustrate how ABMs can produce both Duvergerian and non-Durgerian equilibria. We do not restrict ourselves to utilities on the unit interval, as Cox (1994) assumes, but the scaling of the utilities is inconsequential for our ABM. As Myerson and Weber (1993) demonstrate for voting systems generally, multiple equilibria exist in the Cox (1994) model. We show some of the equilibria that can arise even given identical voter utilities.

[3] In general, this statement is not true. By virtue of knowing the distribution of the vote difference for two parties, we also have more information about their respective rankings relative to other candidates. However, as N gets large and the relative differences between parties gets large, independence provides a reasonable approximation to the true value of pv_k. Simulations and similarity of replicated results have shown that there are no substantive changes using this approximation over exact results. This topic will be explored much more in future iterations of this project.

2.1 Strategic Abstention

Here we focus on the simplified scenario in which the two candidates have equal support among electors $i = 1, \ldots, N_e$, have the same net cost c of voting and the same initial probability $q_i = q$ of voting. More specifically, we initialize each run of the model instantiating N expected voters where half support candidate A and the other half support candidate B, so $N_A = N_B = N/2$; and we also define a value for q and c s.t. $q \sim U(0, 1)$ and $c \sim U(0, 0.5)$.

Let n_A and n_B be, respectively, the expected numbers of supporters of candidates A and B that would go vote if the election happened at the current iteration of the model. Then

$$n_j = \sum_{i}^{N_j} q_i \tag{8}$$

where $i = 1, \ldots, N_j$ are the electors that support candidate j.

At each iteration t, each elector i updates her probability q_i of voting, according to the following difference equation:

$$q_i(t+1) = q_i(t) + K_i(n_A(t), n_B(t)) \tag{9}$$

To implement the function $K(.)$ of Demichelis and Dhillon (2010) (recall (2)) we use

$$K_i(n_A, n_B) = k \ \mathrm{sign}(g_i(n_A, n_B) - 2c) \tag{10}$$

where constant $k > 0$ represents the learning speed mentioned but not specified in Demichelis and Dhillon (2010)—here k simply the convergence speed and run granularity of the model—and $g_i(n_A, n_B)$ is the function that calculates the probabilities of elector i

being pivotal to untie or to tie the election:

$$g_i(n_A, n_B) = \begin{cases} \mathcal{S}(0, n_A - 1, n_B) + \mathcal{S}(-1, n_A - 1, n_B) & \text{if } i \text{ supports A} \\ \mathcal{S}(0, n_B - 1, n_A) + \mathcal{S}(-1, n_B - 1, n_A) & \text{if } i \text{ supports B} \end{cases} \tag{11}$$

where $\mathcal{S}$ is defined in (5). $\mathcal{S}(0, ...)$ and $\mathcal{S}(-1, ...)$ represent, respectively, the probability of elector i being pivotal to untie the election and the probability of i being pivotal to tie the election. Given the support of the parameters of the Skellam density, we set $\mathcal{S} = 0$ if n_A, $n_A - 1$, n_B or $n_B - 1$ is equal to zero.

This way, as the simulation goes, the probability of any elector i going to vote decreases faster when the expected abstention is low and it decreases more slowly when the expected becomes higher. Up to a point in when $g(.)$ surpasses twice the cost c of voting, in which case the probability of i voting starts to increase. The model keeps running until convergence is reached.[4]

Figure 1 shows the final probability q of voting versus the cost of voting, subsetting according to different levels of initial q. As it can be clearly seen, when initial $q = 0$ no elector ever increases such probability and nobody is expected to go vote. For all other levels of initial q, all electors are expected to vote when $c = 0$, many electors are expected to vote when c is very small, but the expected turnout rapidly decreases as c increases, to the point of quickly approaching zero turnout. This result replicates the main equilibrium condition in Demichelis and Dhillon (2010)'s Figures 1 and 2, as well as the stable lower bound equilibrium.

*** Figure 1 about here ***

Figure 1 in Demichelis and Dhillon (2010) shows an equilibrium with full turnout for all

[4] Observed convergence in this model has been of two types, which we call hard convergence and bouncing, hence we needed two convergence criteria to stop the simulation. A given run was considered to hard-converge if the values of n_A and n_B remained the same for 100 iterations. Otherwise, a run could be stopped due to achieving stable bouncing if n_A and n_B kept alternating between two exact same values every two iterations for 100 iterations. In both cases, since we are operating with floating point arithmetic, we consider equality over iterations as being equality up to $k/100$.

values $0 \leq c \leq 1/2$, but we do not reproduce this equilibrium. We find full turnout—$q = 1$—is a stable equilibrium for small values of c, but for larger values of c using an initial value of $q = 1$ leads to the low turnout equilibria. Such a result is not surprising. As Palfrey and Rosenthal (1985, 71–73) prove, the wide range of full turnout equilibria does not exist for electorates as large as we are using. Whenever n_A or n_B is not small, the Skellam distribution function returns small values so that $g(.) < 2 \cdot c$, which forces q to decrease.

Of course, the fact that with all initial q greater than zero the relation between final q and c seems identical does not mean that the process to reach such similar equilibrium is of identical speed. As it can be seen in Figure 2, the distribution of in which iteration the model did converge follows a clear pattern. The greater the initial q, the longer the model takes to achieve the near zero turnout that it achieves as c gets further from zero.

*** Figure 2 about here ***

2.2 Wasted Vote Logic

This paper's replication of the Cox (1994) wasted vote model functions as follows. First we specify a number of candidates and a number of electors. For each elector we generate one utility value per candidate, u_i. These utilities are generated according a user-specified distribution, and each elector has a utility value for every candidate. These utilities are assumed to be von Neumann-Morgenstern consistent and indicative of at least a weak preference ordering over the candidates.[5] Voters then start by choosing any candidate other than their least preferred candidate to be their initial choice. We let voters "state" any nondominated initial choice to capture effects of variation in the initial beliefs that we have not yet modeled explicitly. Once they make this initial choice, they use public information about the distribution of all voters' current choices to make a strategically

[5]Given the double-precision pseudorandom numbers we use to generate utilities, no two utility values should be the same.

optimal choice, v_i, according to

$$u_i^*(h) = \sum_{j=1}^{H} T_{h,j}(u_i(h) - u_i(j)) \\ v_i = \underset{h}{\operatorname{argmax}}[u_i^*(h)] \tag{12}$$

where $T_{h,j}$ is the probability that an elector casts a pivotal vote for h against j. Using (7), we set $T_{h,j} = pv_h$ with N_h being the current "stated" vote counts for candidates $h, j = 1, \ldots, H$.

This procedure is iterative: each elector simultaneously updates her choice to be strategically optimal conditional upon a) her utilities and b) every other electors' current choice. But after each step the state of the world has changed, so the strategic decisions that were conditioned on other electors' previous choices are no longer strategically optimal. So in the next iteration of the model every elector simultaneously updates her choices given each other electors' current choices. The electors continue to update in this manner until no electors' strategic choices change for an arbitrary user-selected number of iterations. At this point—and not before this point—electors' choices are in a strategic equilibrium.

Between any two runs of the model, there are therefore two elements of randomness. The central source of randomness is the distribution by which electors' utilities are initialized. The second source of randomness is the arbitrary selection of voters' initial candidate choice. Different results can be obtained by changing the distribution of electors' utilities, or by holding those utilities fixed while rerunning the model with different initial choices. The first option is a completely different initial setup, which generates electors that have meaningfully different preference structures. The second option is a perturbation of the same initial electors and candidates, changing only the electors' initial selection. A run of the model with the arbitrary initial selection varied should yield an equilibrium in the set of equilibria which are obtained by perturbing a model slightly away from its initial setup.

With straightforward specifications of the model, our replication yields two different outcomes which were derived by Cox (1994): the 2-candidate or $M+1$ result (in which one candidate wins, another candidate gets second place, and every other candidate falls to near-0 votes), and a non-Duvergerian result (in which one candidate wins and all of the other candidates tie for last).[6] We demonstrate a variety of these results by rerunning the model with several different utility-generating distributions, and checking which equilibria can be obtained by perturbing the initial choice of favored candidate.

Some equilibrium vote totals are shown in Figure 3. Each row of the figure is a different distribution for generating sincere utilities, and within each row are several plots which show the final vote totals of 4 competing candidates at the different equilibria which can be obtained by varying the initial selection of candidates without altering electors' utilities. Generating sincere utilities according to a power law (exponent 2), logistic ($\mu = 2, s = 1$), or beta distribution ($\alpha = 0.01, \beta = 0.01$) typically produces an $M+1$ result. Using a standard normal distribution or a uniform distribution typically produces the non-Duvergerian result, but it is also possible to obtain the $M+1$ result for some initializations.[7] It is generally possible to obtain multiple equilibria by varying electors' initial candidate choices without altering utilities. These equilibria can be quite different; the first row shows that, by perturbing the initial choice of a group of electors that otherwise does not change, it is possible to make any candidate win. Several different equilibria can be obtained in nearly any run of the model using only initial perturbations. Given Myerson and Weber (1993)'s results about the existence of many equilibria in almost any voting system, such results are what we expect.

*** Figure 3 about here ***

[6]This code is in an Alpha stage; the results that Cox (1994) predicts make up the overwhelming results for simple parameter choices, but they are not fully robust to sweeps over the random parameters. The conditions in which the model produces unexpected results are themselves interesting, and are one of the top priorities for future work in this project.

[7]In our replication the disfavored candidates typically fall to between 0% and 5% of the vote. This is presumably because this is a finite approximation of an asymptotic phenomenon, and in the very large-elector case those trailing candidates would approach 0%.

3 Discussion

Using agents in a many-equation dynamical system and a Poisson game information structure, we replicate basic equilibrum results from Demichelis and Dhillon (2010), Palfrey and Rosenthal (1985) and Cox (1994). Symmetric mixed equilibria that Demichelis and Dhillon (2010) show are dynamically stable given their learning model based on a differential equation we find are dynamically stable using, effectively, difference equations. We find that both Duvergerian and non-Duvergerian equilibria such as Cox (1994) studies are dynamically stable in our system. Even though we do not yet explicitly model the beliefs that can support multiple equilibria in the wasted vote model, we find multiple equilibria by using different initial values for voters' vote choice probabilities (compare Bouton and Gratton 2015).

Finding these equilibria requires careful attention to numerical details. Using the Poisson game approach with Skellam distributions, pivotal probabilities are extremely small with large electorates. The electorate size we use here—$N = 1000$—is not large compared to the country-scale electorates that will ultimately concern us. Use of extended precision numerical libraries is crucial. In our simulations we still encounter some anomalous behavior that we suspect may trace to numerical imprecision. It is also possible that we will need to abandon our approach using difference equations—that is, each agent updates at discrete time steps—and move to a differential or difference-differential scenario. Perhaps agent update continuously, but information about expected election outcomes updates discretely.

We do not yet incorporate beliefs explicitly into the models. Agents also have fixed preferences. In the current implementations choices are updated without reference to any prior beliefs—other than whatever beliefs are implicit in the specification of the current state—based on public information about the currently expected outcome of the election. We plan to augment the models to make agents explicitly Bayesian, and we will investigate what happens when information is more limited. For example, agents may learn only about the intentions of their neighbors, in various senses of “neighbor.” Or agents may be only adaptively rational.

We encountered apparent anomalous behavior in scenarios, not reported here, where we moved far away from the conditions specified in the formal theory papers. For instance, we've run iterations of the strategic abstention model with initial vote probabilities q_i drawn using a uniform distribution on the unit interval. Demichelis and Dhillon (2010) say their results hold for mildly dispersed initial q_i values. When initial q_i is uniformly distributed on $(0, 1)$, phenomena very different from their equilibria often arise. A next thing we will do is explore how much initial q_i can vary across individuals before their equilibrium results break. We will also investigate variation over individuals in net costs.

After similarly exploring the boundaries of weakening the assumptions in the wasted vote simulation, we expect to combine the strategic abstention and wasted vote models. Such a combination should produce new results going beyond what has been possible using paper-and-pencil formal theory tools. After that, we'll see where it goes.

References

Axelrod, Robert. 1997. Advancing the art of simulation in the social sciences. In *Simulating social phenomena*. Springer pp. 21–40.

Axelrod, Robert and Leigh Tesfatsion. 2006. "Appendix A: A guide for newcomers to agent-based modeling in the social sciences." *Handbook of computational economics* 2:1647–1659.

Beber, Bernd and Alexandra Scacco. 2012. "What the Numbers Say: A Digit-Based Test for Election Fraud." *Political Analysis* 20(2):211–234.

Bendor, Jonathan and Daniel Diermeier. 2003. "A Behavioral Model of Turnout." 97(2).

Birch, Sarah. 2011. *Electoral Malpractice.* New York: Oxford University Press.

Bonabeau, Eric. 2002. "Agent-based modeling: Methods and techniques for simulating human systems." *Proceedings of the National Academy of Sciences* 99(suppl 3):7280–7287.

Borghesi, C. 2009. "Une étude de physique sur les élections—Régularités, prédictions et modèles sur des élections françaises." arXiv:0910.4661[physics.soc-ph].

Borghesi, C. and J. P. Bouchaud. 2010. "Spatial Correlations in Vote Statistics: A Diffusive Field Model for Decision-making." *European Physical Journal B* 75(3):395–404.

Bouton, Laurent and Gabrielle Gratton. 2015. "Majority runoff elections: Strategic voting and Duverger's hypothesis." *Theoretical Economics* 10:283–314.

Breunig, Christian and Achim Goerres. 2011. "Searching for Electoral Irregularities in an Established Democracy: Applying Benford's Law Tests to Bundestag Elections in Unified Germany." *Electoral Studies* 30(3):534–545.

Bruch, Elizabeth and Jon Atwell. 2015. "Agent-Based Models in Empirical Social Research." *Sociological Methods & Research* 44(2):186–221.

URL: *http://journals.sagepub.com/doi/10.1177/0049124113506405*

Bush, Robert R and Frederick Mosteller. 1955. "Stochastic models for learning.".

Cantu, Francisco and Sebastián M. Saiegh. 2011. "Fraudulent Democracy? An Analysis of Argentina's Infamous Decade Using Supervised Machine Learning." *Political Analysis* 19(4):409–433.

Cox, Gary W. 1994. "Strategic Voting Equilibria Under the Single Nontransferable Vote." *American Political Science Review* 88:608–621.

de Marchi, Scott and Scott E. Page. 2014. "Agent-Based Models." *Annual Review of Political Science* 17(1):1–20.

URL: *http://www.annualreviews.org/doi/10.1146/annurev-polisci-080812-191558*

Deckert, Joseph, Mikhail Myagkov and Peter C. Ordeshook. 2011. "Benford's Law and the Detection of Election Fraud." *Political Analysis* 19(3):245–268.

Demichelis, Stefano and Amrita Dhillon. 2010. "Learning in Elections and Voter Turnout." *Journal of Public Economic Theory* 12(5):871–896.

Downs, Anthony. 1957. "An economic theory of political action in a democracy." *Journal of political economy* 65(2):135–150.

Epstein, Joshua M. 1999. "Agent-based computational models and generative social science." *Complexity* 4(5):41–60.

Ferrari, Diogo, Kevin McAlister and Walter R. Mebane, Jr. 2018. "Developments in Positive Empirical Models of Election Frauds: Varying Dimensions." Paper presented at the 2018 Annual Meeting of the Midwest Political Science Association, Chicago, April 5–8, 2018.

Ferrari, Diogo and Walter R. Mebane, Jr. 2017. "Developments in Positive Empirical Models of Election Frauds." Paper presented at the 2017 Summer Meeting of the Political Methodology Society, Madison, WI, July 13–15, 2017.

Hicken, Allen and Walter R. Mebane, Jr. 2015. "A Guide to Election Forensics." Working paper for IIE/USAID subaward #DFG-10-APS-UM, "Development of an Election Forensics Toolkit: Using Subnational Data to Detect Anomalies".

Holland, John H. 1983. Escaping brittleness. In *Proceedings Second International Workshop on Machine Learning*. Citeseer pp. 92–95.

Hyde, Susan D. and Nikolay Marinov. 2012. "Which Elections Can Be Lost?" *Political Analysis* 20(2):191–201.

Johnson, Paul E. 1999. "Simulation Modeling in Political Science." *American Behavioral*

Scientist 42(10):1509–1530.

Kalinin, Kirill and Walter R. Mebane, Jr. 2011. "Understanding Electoral Frauds through Evolution of Russian Federalism: from "Bargaining Loyalty" to "Signaling Loyalty"." Paper presented at the 2011 Annual Meeting of the Midwest Political Science Association, Chicago, IL, March 31–April 2.

Kim, Sung-youn, Charles S Taber and Milton Lodge. 2010. "A computational model of the citizen as motivated reasoner: Modeling the dynamics of the 2000 presidential election." *Political Behavior* 32(1):1–28.

Klimek, Peter, Yuri Yegorov, Rudolf Hanel and Stefan Thurner. 2012. "Statistical Detection of Systematic Election Irregularities." *Proceedings of the National Academy of Sciences* 109(41):16469–16473.

Kollman, Ken, John H Miller and Scott E. Page. 1992. "Adaptive Parties in Spatial Elections." *The American Political Science Review* 86(4):929–937.

Lehoucq, Fabrice. 2003. "Electoral Fraud: Causes, Types, and Consequences." *Annual Review of Political Science* 6:233–256.

Lehoucq, Fabrice and Iván Molina Jiménez. 2002. *Stuffing the Ballot Box: Fraud, Electoral Reform and Democratization in Costa Rica*. New York: Cambridge University Press.

Levin, Inès, Gabe A. Cohn, Peter C. Ordeshook and R. Michael Alvarez. 2009. "Detecting Voter Fraud in an Electronic Voting Context: An Analysis of the Unlimited Reelection Vote in Venezuela." 2009 Electronic Voting Technology Workshop/Workshop on Trustworthy Elections, Montreal, August 10–11.

Levitsky, Steven and Lucan A. Way. 2010. *Competitive Authoritarianism: Hybrid Regimes After the Cold War*. New York: Cambridge University Press.

Magaloni, Beatriz. 2006. *Voting for Autocracy: Hegemonic Party Survival and Its Demise in Mexico*. New York: Cambridge University Press.

Mebane, Jr., Walter R. 2010. "Fraud in the 2009 Presidential Election in Iran?" *Chance* 23:6–15.

Mebane, Jr., Walter R. 2013*a*. “Election Forensics: The Meanings of Precinct Vote Counts’ Second Digits.” Paper presented at the 2013 Summer Meeting of the Political Methodology Society, University of Virginia, July 18–20, 2013.

Mebane, Jr., Walter R. 2013*b*. “Using Vote Counts’ Digits to Diagnose Strategies and Frauds: Russia.” Paper prepared for the 2013 Annual Meeting of the American Political Science Association, Chicago, August 29–September 1.

Mebane, Jr., Walter R. 2014. Can Votes Counts’ Digits and Benford’s Law Diagnose Elections? In *The Theory and Applications of Benford’s Law*, ed. Steven J. Miller. Princeton: Princeton University Press pp. 206–216.

Mebane, Jr., Walter R. 2016. “Election Forensics: Frauds Tests and Observation-level Frauds Probabilities.” Paper presented at the 2016 Annual Meeting of the Midwest Political Science Association, Chicago, April 7–10, 2016.

Miller, John H and Scott E Page. 2004. “The Standing Ovation Problem.” *Complexity* 9(5):8–16.

Minnite, Lorraine C. 2010. *The Myth of Voter Fraud*. Ithaca, NY: Cornell University Press.

Montgomery, Jacob M., Santiago Olivella, Joshua D. Potter and Brian F. Crisp. 2015. “An Informed Forensics Approach to Detecting Vote Irregularities.” *Political Analysis* forthcoming.

Myagkov, Mikhail, Peter C. Ordeshook and Dimitry Shaikin. 2009. *The Forensics of Election Fraud: With Applications to Russia and Ukraine*. New York: Cambridge University Press.

Myagkov, Misha, Peter C. Ordeshook and Dimitry Shaikin. 2008. Estimating the Trail of Votes in Russia’s Elections and the Likelihood of Fraud. In *The Art and Science of Studying Election Fraud: Detection, Prevention, and Consequences*, ed. R. Michael Alvarez, Thad E. Hall and Susan D. Hyde. Washington, DC: Brookings Institution.

Myerson, Roger B. 1998. “Population Uncertainty and Poisson Games.” *International Journal of Game Theory* 27:375–392.

Myerson, Roger B. 2000. “Large Poisson Games.” *Journal of Economic Theory* 94(1):7–45.

Myerson, Roger B. and Robert J. Weber. 1993. "A Theory of Voting Equilibria." *American Political Science Review* 87(1):102–114.

Norris, Pippa. 2014. *Why Electoral Integrity Matters.* New York: Cambridge University Press.

Palfrey, Thomas R. and Howard Rosenthal. 1985. "Voter participation and strategic uncertainty." *American Political Science Review* 79(1):62–78.

Parunak, H Van Dyke, Robert Savit and Rick L Riolo. 1998. Agent-based modeling vs. equation-based modeling: A case study and users guide. In *International Workshop on Multi-Agent Systems and Agent-Based Simulation.* Springer pp. 10–25.

Pericchi, Luis Raúl and David Torres. 2011. "Quick Anomaly Detection by the Newcomb-Benford Law, with Applications to Electoral Processes Data from the USA, Puerto Rico and Venezuela." *Statistical Science* 26(4):502–516.

Rozenas, Arturas. 2017. "Detecting Election Fraud from Irregularities in Vote-Share Distributions." *Political Analysis* 25(1):41–56.

Rundlett, Ashlea and Milan W. Svolik. 2015. "Deliver the Vote! Micromotives and Macrobehavior in Electoral Fraud." Working paper.

Schedler, Andreas, ed. 2006. *Electoral Authoritarianism: The Dynamics of Unfree Competition.* Boulder and London: Lynne Rienner Publishers.

Shikano, Susumu and Verena Mack. 2009. "When Does the Second-Digit Benford's Law-Test Signal an Election Fraud? Facts or Misleading Test Results." *Jahrbücher fur Nationalökonomie und Statistik* 231(5-6):719–732.

Simpser, Alberto. 2013. *Why Governments and Parties Manipulate Elections: Theory, Practice, and Implications.* New York: Cambridge University Press.

Skellam, J. G. 1946. "The Frequency Distribution of the Difference Between Two Poisson Variates Belonging to Different Populations." *Journal of the Royal Statistical Society. Series A (General)* 109(3):296.

Stokes, Susan C., Thad Dunning, Marcelo Nazareno and Valeria Brusco. 2013. *Brokers, Vot-*

ers, and Clientelism: The Puzzle of Distributive Politics. New York: Cambridge University Press.

Svolik, Milan. 2012. *The Politics of Authoritarian Rule.* New York: Cambridge University Press.

Wang, Tova Andrea. 2012. *The Politics of Voter Suppression: Defending and Expanding Americans' Right to Vote*. Ithaca, NY: Cornell University Press.

Figure 1: Probability of voting (q) versus cost of voting (c)

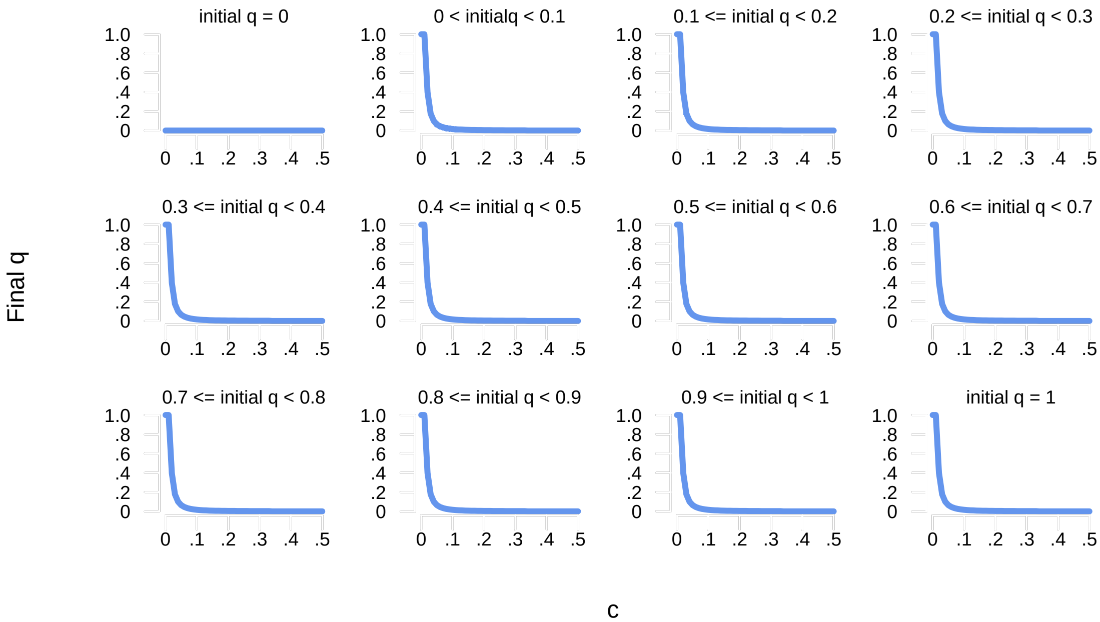

Figure 2: Iteration when model run converged

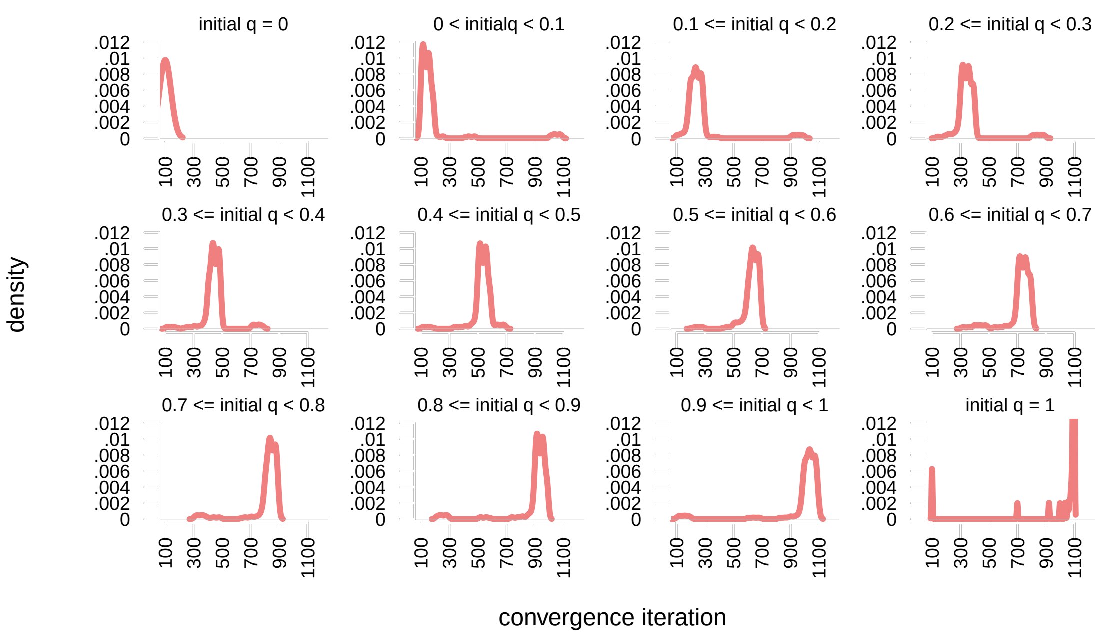

Figure 3: Wasted Vote Model with Varying Utilitiess