Measuring Election Frauds[*]

Walter R. Mebane, Jr.[†] Diogo Ferrari[‡] Kevin McAlister[§]

Patrick Y. Wu[¶]

March 6, 2022

*A previous version was presented at the Workshop on Quantitative Methods, University of Chicago, February 12, 2021. Work supported by NSF award SES 1523355 and IIE/USAID subaward #DFG-10-APS-UM. Thanks to Shang Zhou Xia, Preston Due, Joseph Klaver, Jonathan Wall and Jacob Berch for assistance, to Kirill Kalinin for comments and to Rob Barry, Bill Gallery, Susan Hyde, Kirill Kalinin and Tom McGinty for supplying data.

†Department of Political Science, Department of Statistics, Center for Political Studies in the Institute for Social Research, University of Michigan (E-mail: wmebane@umich.edu). Corresponding author.

‡Department of Political Science, University of California Riverside (E-mail: diogo.ferrari@ucr.edu).

§Department of Quantitative Theory and Methods, Emory University (E-mail: kevin.mcalister@emory.edu).

¶Department of Political Science, University of Michigan (E-mail: pywu@umich.edu).

Abstract

A measurement model called `eforensics` produces valid estimates of the incidence and magnitude of election frauds, using as input aggregation unit counts of eligible voters and of votes for the ballot alternatives. While valid, the estimates are imperfect: `eforensics` cannot detect procedural defects, measures only frauds that benefit one alternative and is sensitive to things—particularly elector strategic behavior and election administration weaknesses—not produced by the kinds of malevolent distortions of elector intentions that constitute genuine frauds. `eforensics` parameters support distinguishing results of malevolent distortions from results of strategic behavior. We use 32 real elections to demonstrate that `eforensics` measures the magnitude of election frauds at aggregation units such as polling stations, precincts or ballot boxes.

1 Introduction

Why is it difficult to measure election frauds? Actually we take on a limited aspect of this question: why is such measurement difficult when all one has to work with are basic summaries of the election such as the number of eligible voters and the number of votes received by ballot alternatives (say candidates) at each aggregation unit? Addressing the question is the core mission 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 (cf. Mebane 2008).[1] By referring to measurement we go beyond trying merely to detect election anomalies (e.g. Myagkov, Ordeshook and Shaikin 2009; Mebane 2014; Montgomery, Olivella, Potter and Crisp 2015; Rozenas 2017; Cantú 2019). Recently methods attempting frauds measurement have been proposed (Klimek, Yegorov, Hanel and Thurner 2012; Klimek, Jiménez, Hidalgo, Hinteregger and Thurner 2018; Zhang, Alvarez and Levin 2019). We use the new statistical model `eforensics` (Ferrari, Mebane, McAlister and Wu 2019) to argue that the key challenges to frauds measurement are unobservable information and ambiguity—and `eforensics` goes a long way towards overcoming these.

What are election frauds? Mindful of both polyarchy (Dahl 1956) and social choice theory (Riker 1982), election forensics refers to election results accurately reflecting the intentions of the electors. Frauds thwart such accurate reflection. To distinguish frauds from mere failures of election administration or other accidents, we might require such thwarting to result from undemocratic actions such as are undertaken by authoritarians (e.g. Lehoucq 2003; Schedler 2006; Svolik 2012; Simpser 2013; Norris 2014). Procedural failures might distort intentions, and procedural failures might be planned to distort. But we say the distortions are the frauds, not the procedural failures per se.

Reference to electors' intentions points to the key unobservable: we can't observe electors' intentions. In politics people act not merely based on what they prefer but also

[1] "Elector" refers to a registered or otherwise eligible voter.

based on what they expect other people to do: by acting in part based on their expectations regarding others, electors act strategically (e.g. Stephenson, Aldrich and Blais 2018). Strategic behavior must be admitted in any election system that satisfies criteria for democracy such as the Gibbard-Satterthwaite theorem expresses (Riker 1982). Acting strategically does not necessarily mean that electors' actions differ from what their sincere actions would be (Kawai and Watanabe 2013), but they might.

Unobservable intentions imply that frauds measurement faces a fundamental ambiguity. Any measurement effort focuses on empirical patterns. Malevolent distortions of intentions—frauds—produce patterns via a mechanism similar to what happens when there is strategic behavior, that is when there is normal politics. We will show that `eforensics` provides valid but imperfect measures of frauds, the imperfection largely stemming from this ambiguity. Features of `eforensics` may support discriminating the frauds.

Manifestations of election frauds are many and varied: beyond tampering with vote tallies there are deploying fake voters and votes, voter intimidation, election violence, voter suppression, misinformation and more (e.g. Birch 2011; Wang 2012; Rundlett and Svolik 2016; Jamieson 2018). Audits can detect some kinds of tampering (e.g. Electoral Complaints Commission 2010; Alvarez, Morrell, Rivest, Stark and Stewart 2019), but with `eforensics` the aim is to measure distortions that might make even procedurally accurate vote tallies not match electors' intentions.

Strategic behavior is no less diverse. Strategies we consider in relation to cases include wasted-vote strategies (Cox 1994), threshold insurance and two-vote strategies (Shikano, Herrmann and Thurner 2009; Harfst, Blais and Bol 2018), bandwagons (Berch 1989), coordinating split-ticket voting (Alesina and Rosenthal 1995; Mebane 2000) and majority-or-runoff strategies (Bouton and Gratton 2015).

We use cases of real elections to exemplify how `eforensics` estimates appear in the presence of genuine bad acts, of various kinds of strategic behavior and of weaknesses in election administration. Administrative weaknesses include bad ballots and voter identification requirements that produce lost votes. We show that the imperfections `eforensics` exhibits as a model for measuring frauds do not prevent its being used if it is used with appropriate attention to nuances linked to estimated parameter values.

After motivating and describing the Bayesian formulation of `eforensics` and its Markov Chain Monte Carlo (MCMC) estimation approach, we apply `eforensics` to a series of elections and postelection audits and court decisions to demonstrate its validity and key limitations (Afghanistan 2009, France 2017, Mexico 2003–2009). Then we present examples dominated by strategic behavior (Germany 2005, Ohio 2006, Georgia 2020) and election administration failures (Florida 2000, Wisconsin 2016). We illustrate substantively motivated covariates (California 2006). We compare to previously proposed methods (Argentina 2015, Turkey 2017, Russia) and apply to several authoritarian elections (Turkey 2011 and 2015, Russia 2000–2020, Armenia 2003), in the last case showing how `eforensics` estimates refine understanding of observer effects.

2 Model Justification and Specification

Statistical approaches based only on counts of electors and votes are challenged because neither electors' preferences, strategies nor information are observed, yet the election forensics task is to assess whether electors' intentions are accurately reflected in the election outcome. `eforensics` is based on an explict model: functional form commitments stand in place of features it is impossible to observe. What's the problem?

The `eforensics` model assumes that if there are no frauds then each elector decides whether to vote and, if so, for whom in a way that can be represented by two binary choices governed by Bernoulli probabilities. The turnout choice is between "vote" and "abstain," and the vote choice decision is between "leader" and "opposition." Conditioning on the number of electors (N_i) at aggregation unit i, the number of votes cast is then an overdispersed binomial random variable: the turnout probability averages the electors' probabilities at i with extra variation due to variation across individuals. Conditioning on N_i, the number of people voting for the leader is an overdispersed binomial with probability being the product of turnout and vote choice probabilities, with extra variation due to variation across individuals.

A key aspect of the mechanism by which frauds distort intentions is that they induce dependencies among individuals' observed votes. Imagine, for example, that fake votes are added all for the vote leader. The fake votes are dependent. Notionally similar dependence arises if many are coerced to vote for the leader, etc. The `eforensics` model measures the dependencies using a finite mixture model. One component corresponds to no fraud, one to "incremental fraud" and one to "extreme fraud." The idea for such a mixture structure comes from Klimek et al. (2012), who emphasize how frauds induce multimodality,[2] which is the manifestation of the dependencies we can observe given aggregate count data.

The problem is that other factors in politics induce dependence, and discriminating dependence that traces to frauds from dependence that originates elsewhere is challenging. Strategic behavior generically presents the most difficult problem because both strategic behavior and frauds induce dependence. For instance, in a Nash equilibrium each elector considers and responds to all others' expected votes. Generally strategic behavior implies there are dependencies among all strategic electors' behavior, with some electors acting systematically similarly and others acting systematically in opposition. Both frauds and strategic behavior can involve votes being changed—with frauds it's some malefactor that changes votes while with strategic behavior individual electors may change their own votes from what each would do if acting sincerely. How to discriminate effects of strategic behavior from effects of frauds is a question.

Features of election administration can also induce dependence among votes but not originate with malevolent actions. E.g., variation in ballot quality or voting equipment provision can induce widespread confusion or delays that lead to voting errors or decreased

[2]Klimek et al. (2012) rely on Borghesi and Bouchaud (2010), whose diffusion model represents local dependencies but not the kinds of global dependencies frauds induce.

turnout (e.g. Mebane 2004; Pettigrew 2017). Multiple sources of dependence mean `eforensics` estimates may be ambiguous as far as interpretations in terms of frauds—malevolent distortions—are concerned.

2.1 Model Specification

In `eforensics` electors either vote or abstain, and vote choices are reduced to two options: one candidate or other ballot alternative is the "leader"; the remaining alternatives are grouped as "opposition." Frauds benefit the leader. Some votes are transferred to the leader from opposition ("stolen"), and some are taken from nonvoters ("manufactured").

In the finite mixture model two types of election fraud refer to how many of the opposition and nonvoter votes are shifted: with "incremental fraud" moderate proportions and with "extreme fraud" almost all of the votes are shifted. Unconditional probabilities that each unit experiences no, incremental or extreme fraud are π_1, π_2 and π_3. The prior ensures π_1 is largest: using $U(0,1)$ for the uniform distribution,

$$\tilde{\pi}_1 \sim U(0,1); \quad \tilde{\pi}_2 \sim U(0,\tilde{\pi}_1); \quad \tilde{\pi}_3 \sim U(0,\tilde{\pi}_1) \tag{1a}$$

$$\pi_j = \frac{\tilde{\pi}_j}{\tilde{\pi}_1 + \tilde{\pi}_2 + \tilde{\pi}_3}, \quad j \in \{1,2,3\} \tag{1b}$$

The model specification features observed and unobserved quantities. Observed data for n aggregation unit observations indexed by $i = 1, \ldots, n$ are the total number of vote-eligible persons (N_i), the number of votes for the leader (W_i) and the number of votes cast (V_i). The model conditions on N_i. The number of abstentions is $A_i = N_i - V_i$, and the number of votes for opposition is $O_i = V_i - W_i$. Observed proportions are $t_i = V_i/N_i$ (turnout proportion), $a_i = 1 - t_i$ (proportion abstaining) and $w_i = W_i/N_i$ (leader proportion).

The model introduces unobserved variables: τ_i is the true turnout proportion; ν_i is the true proportion of votes cast for the leader; and $Z_i \in \{1,2,3\}$ is the fraud type indicator—$Z_i = 1$, no; $Z_i = 2$, incremental; $Z_i = 3$, extreme. Fraud magnitudes depend on unobserved proportions ι_i^M and ι_i^S (the proportions of votes manufactured from abstainers or stolen from opposition given incremental fraud), and υ_i^M and υ_i^S (the proportions manufactured or stolen given extreme fraud). These proportions depend on observed covariates and random effects: for $k = .7$

$$\nu_i = \frac{1}{1+\exp[-(\beta^\top x_i^\nu + \kappa_i^\nu)]} \tag{2a}$$

$$\tau_i = \frac{1}{1+\exp[-(\gamma^\top x_i^\tau + \kappa_i^\tau)]} \tag{2b}$$

$$\iota_i^l = \frac{k}{1+\exp[-(\rho_l^\top x_i^\iota + \kappa_i^{\iota l})]}\,, l \in \{M, S\} \tag{2c}$$

$$\upsilon_i^l = k + \frac{1-k}{1+\exp[-(\delta_l^\top x_i^\upsilon + \kappa_i^{\upsilon l})]}\,, l \in \{M, S\}\,. \tag{2d}$$

For $\xi \in \{\nu, \tau, \iota, \upsilon\}$ each x_i^ξ is a vector of observed covariates, and β, γ, ρ_M, ρ_S, δ_M, δ_S are vectors of coefficients (independent Normal priors, $N(0, 1/10000)$). Each κ_i^ξ is an unobserved variable that for unknown mean $\mu^{\kappa\xi}$ and standard deviation $\sigma^{\kappa\xi}$ is assumed to have as prior the Normal distribution $\kappa_i^\xi \sim N(\mu^{\kappa\xi}, \sigma^{\kappa\xi})$ with $\mu^{\kappa\xi} \sim N(0, 1)$, $\sigma^{\kappa\xi} \sim \text{Exp}(5)$, and likewise for $\kappa_i^{\xi M}$ and $\kappa_i^{\xi S}$. In ν_i and τ_i random effects κ_i^ν and κ_i^τ capture overdispersion, and in ι_i^M, ι_i^S, υ_i^M and υ_i^S random effects $\kappa_i^{\iota M}$, $\kappa_i^{\iota S}$, $\kappa_i^{\upsilon M}$ and $\kappa_i^{\upsilon S}$ capture extra variation in observation-level frauds.

Each aggregation unit has all its counts from either no frauds, incremental fraud or extreme fraud mixture components. The proportions of N_i that are fraudulent are

$$p_{ti} = \begin{cases} 0, & \text{if } Z_i = 1 \\ \iota_i^M(1-\tau_i), & \text{if } Z_i = 2 \\ \upsilon_i^M(1-\tau_i), & \text{if } Z_i = 3 \end{cases}, \quad p_{wi} = \begin{cases} 0, & \text{if } Z_i = 1 \\ \iota_i^M(1-\tau_i) + \iota_i^S\tau_i(1-\nu_i), & \text{if } Z_i = 2 \\ \upsilon_i^M(1-\tau_i) + \upsilon_i^S\tau_i(1-\nu_i), & \text{if } Z_i = 3\,, \end{cases} \tag{3}$$

with p_{ti} being fraudulent turnout and p_{wi} being fraudulent vote choice. The likelihood for A_i and W_i is a product of binomial distributions each having N_i "trials" and binomial probability respectively $a_i^* = 1 - \tau_i - p_{ti}$ and $w_i^* = \nu_i\tau_i + p_{wi}$.

`eforensics` implements MCMC using Metropolis-Hastings (Plummer, Stukalov and Denwood 2016) with four chains (Denwood 2016), using Markov Chain Monte Carlo Standard Error (MCMCSE) (Flegal, Haran and Jones 2008; Flegal and Hughes 2012; Gong and Flegal 2016) for a stopping rule. Chains run until $\sigma_{\theta_j} < .05$ where $\sigma^2_{\theta_j}$ is the estimated asymptotic variance of the j^{th} component of parameter vector $\boldsymbol{\theta}$, computing $\sigma^2_{\theta_j}$ using consistent nonoverlapping batch means (Jones, Haran, Caffo and Neath 2006).

Observation i is classified as type $\tilde{Z}_i \in \{1, 2, 3\}$ if a plurality of MCMC iterations have $Z_i = \tilde{Z}_i$. Using indicator function $\mathcal{I}(\cdot)$ the number of `eforensics`-fraudulent observations is $H = \sum_{i=1}^{n} \mathcal{I}(\tilde{Z}_i \in \{2, 3\})$, and the proportion is $\varphi = H/n$. Numbers of `eforensics`-fraudulent voters and votes at i are $F_{ti} = p_{ti}N_i$ and $F_{wi} = p_{wi}N_i$, with totals $F_t = \sum_{i=1}^{n} F_{ti}$ and $F_w = \sum_{i=1}^{n} F_{wi}$. The proportion of $\sum_{i=1}^{n} W_i$ that is `eforensics`-fraudulent is $\psi_{wW} = F_w / \sum_{i=1}^{n} W_i$, and the proportion manufactured is $\psi_{tW} = F_t / \sum_{i=1}^{n} W_i$. The ratio of `eforensics`-fraudulent votes to the difference M between votes counted for first and second place is $\psi_{wM} = F_w/(M + 1)$.

As we explain more fully in the context of cases, features of parameters' and `eforensics`-frauds' posterior distributions are important for interpreting `eforensics` estimates. Whether credible intervals of intercept coefficients ρ_{M0}, ρ_{S0}, δ_{M0} and δ_{S0} are negative, positive or include zero will be shown to be indicators for whether an election is distorted by bad acts or merely features strategic behavior:[3] if an interval includes zero then the parameter does not differ from the prior mean, which suggests the parameter estimate is not being materially updated from the observed data, so that the referent aspect of the `eforensics` model may be effectively not operative for the election. Whether π_2 has a multimodal posterior will be shown to be connected to the occurrence of lost votes. F_t, F_w, ψ_{tW}, ψ_{wW} and ψ_{wM} help specify the degree to which `eforensics`-fraudulent votes are decisive for election outcomes and for court judgments regarding election outcomes.

[3] We use ρ_{M0}, ρ_{S0}, δ_{M0}, δ_{S0}, F_t and F_w to refer both to the values at each MCMC iteration and to their posterior means and credible intervals.

3 Cases

We consider empirical examples to illustrate features, strengths and limitations of `eforensics` as a tool for measuring election frauds.

3.1 Afghanistan 2009 President

The Afghanistan 2009 presidential election helps demonstrate the validity of `eforensics` estimates as measures of frauds while also illustrating the method's two key limitations. Limitations are that the method is insensitive both to frauds that benefit candidates other than whoever is designated as the "leader" and to purely procedural frauds.

The election featured frauds that were assessed by an audit that manually examined a random sample of 345 polling stations (Electoral Complaints Commission 2010). The audit supplies judgment-based indicators of polling station frauds determined by the Electoral Complaints Commission [ECC] to which we can compare `eforensics` estimates. Inspection of polling stations included review of physical conditions and contents of ballot boxes as well as of results and forms (see Supplemental Information [SI] 1.2.1). Audit decisions were either Fraudulent, Valid, "Valid; but form fraud" or N/A (SI 1.2.1). Although polling stations are in the sample only if vote or turnout are extreme (SI 1.2.1), the audit found 74 Valid and 249 Fraudulent polling stations (Democracy International 2009), so ECC decisions depend on more than mere extremity. Reasons for Fraudulent decisions include, "75% uniform markings; more than 100 ballots marked with felt marker," "No forms in box; no stubs, materials cannot be reconciled; 75% uniform markings; 100% of ballots never folded" and other similar findings (Electoral Complaints Commission 2009).

Table 1 reports `eforensics` parameters estimates.[4] Extreme fraud is more likely than incremental fraud ($\pi_3 = .0579 > \pi_2 = .0351$), $H = 1622$ polling stations and $F_w = 612639.5$ of leader Karzai's votes are `eforensics`-fraudulent. Using (2) at each MCMC iteration (ignoring the random effects) produces the 99.5% credible intervals for proportions τ_0, ν_0,

[4] N_i is set equal to 600 for all polling stations i unless $V_i > 600$, in which case we set $N_i = V_i$ (SI 1.1.1).

ι_0^M, ι_0^S, υ_0^M and υ_0^S shown in Table 2. $\nu_0 \in [.470, .493]$ is an `eforensics` estimate for the probability that a nonfraudulent vote was for Karzai.

`eforensics`-fraudulent votes F_w are 20 percent of Karzai's votes (Table 3). Karzai received 3093256 of the 5662758 votes cast (Democracy International 2009). Removing F_t and F_w would leave him with vote proportion $\frac{3093256 - 612639.5}{5662758 - 483730.5} = .479$, less than the threshold of .5 needed to avoid a runoff election (Electoral Complaints Commission 2010, 37). Either that quantity or ν_0 gives a result similar to the result of the ECC procedure by which the Independent Election Commission (IEC) "certified the final results with Karzai receiving 49.67 percent of the vote" (Electoral Complaints Commission 2010, 37) (SI 1.2.1).

Agreement about the overall number of fraudulent votes does not mean that `eforensics` agrees with audit decisions for specific polling stations. Checking that agreement reveals both the strength and the main limitations of `eforensics`. First, `eforensics` generally ignores frauds that benefit non-leader candidates: the audit sample included 85 polling stations at which Karzai did not have the most votes, with 26 Valid and 54 Fraudulent, but none of these polling stations are `eforensics`-fraudulent. Second, `eforensics` is generally unable to capture purely procedural frauds. A mishap with the sampling procedure the ECC used facilitates illustrating this limitation: the ECC triggered a second round of sampling to include polling stations omitted in the initial sampling frame, and the second frame includes polling stations all of which had many votes deemed invalid before the audit for procedural reasons (SI 1.2.1). Using polling stations at which Karzai led, Table 4 shows that among stations sampled from the initial, "valid votes" frame `eforensics`-fraud classifications agree strongly with the audit decisions, while among stations from the second, "total votes" frame classifications and decisions are unrelated.

3.2 France 2017 National Assembly

Regarding the 2017 National Assembly election in France (SI 1.2.2), the *Conseil Constitutionnel* issued 505 decisions concerning 307 districts, including eight decisions to annul a district's election (SI 1.1.2). We show the decisions relate to `eforensics`-fraudulent votes. `eforensics` estimation for second-round votes uses commune aggregation units (SI 1.1.2), and W_i contains district winner votes. The model includes all communes with district fixed effects specified for turnout and choice (in x^τ and x^ν).

Across the whole country `eforensics`-frauds are rare: π_1 has posterior mean .990 with HPD 95% interval $[.989, .992]$; $\pi_2 = .00935$ $[.00785, .0109]$; $\pi_3 = .000153$ $[.0000260, .000295]$ (estimates for mixture probabilities are shown graphically in Figure 1). Among frauds parameters ρ_{M0}, ρ_{S0}, δ_{M0} and δ_{S0} (Figure 2) only $\rho_{M0} <$ and $\delta_{M0} > 0$ differ from their prior means of zero (i.e., $\rho_{S0} \approx \delta_{S0} \approx 0$). $\varphi = 76/35750$ is small, and pooled over all districts posterior means $F_t = 7154.4$ and $F_w = 10900.3$ are small in absolute terms and as a proportion of leaders' votes (Table 3). Empirical densities and rug plots in Figure 3 show outcomes by district: few districts have φ greater than .1 though in one district $\varphi > .5$ (Figure 3(1)(a)); in most districts $\psi_{wW} < .05$ but one district has $\psi_{wW} \approx .3$ (Figure 3(1)(b)); and in two districts F_w is nearly as large as margin M between first and second, while in one district $\psi_{wM} > 5$ (Figure 3(1)(c)).

We use two binomial logistic regressions to assess `eforensics`-frauds estimates' relation to *Conseil* actions. One model's outcome is the number of cases in each district. The other's is whether annulment occurs given that there is a case.[5] Regressors are manufactured (F_t) and stolen ($F_w - F_t$) votes in each district, each normalized for comparability across districts.[6] District F_t is divided by votes cast: F_t/V, where $V = \sum_i V_i$ is the district sum. Language in the annulment decisions guides how we normalize stolen votes. Frequently decisions say, "in view of the small difference in votes between the two candidates present in the second ballot, it is necessary, without there being any need to examine the other complaints, to annul the contested electoral operations."[7] M being interesting to the court, we use $(F_w - F_t)/(M + 1)$.

[5] Annulments model uses a cases model with regressor M to adjust for censoring (Maddala 1983, 277–278).

[6] To account for variation across MCMC draws, we use Normal approximation coefficient means and confidence intervals (Pemstein, Meserve and Melton 2010). The algorithm uses robust covariance matrices.

[7] Translation of paragraph 9 in Conseil Constitutionnel (2017).

Coefficient estimates in Table 5 show that probabilities of cases and annulments each increases with both F_t/V and $(F_w - F_t)/(M+1)$.

3.3 Mexico 2003–2009 Deputies and President

During 2003–2009 (SI 1.2.3) the *Tribunal Electoral del Poder Judicial de la Federación* (TEPJF) (Eisenstadt 2007) annulled results from many *casillas* (ballot boxes)—676 *casillas* in 2003, 2006 and 2009 deputies elections; 79 in 2006 senate elections; and 716 in the 2006 presidential election (SI 1.1.3)—although the proportion of *casillas* annulled is small. For example, in the 2006 presidential election annulment petitions were filed against 27,109 of the 130,788 *casillas*, and the 716 TEPJF annulled comprised .56 percent (237,736/41,791,322) of the votes—but that small proportion of annulled *casillas* contained nearly as many votes as the margin of victory, which was 243,934 votes. Unlike with the *Conseil Constitutionnel*, TEPJF annulment decisions focus not on district contests on but on specific *casillas*. To enhance comparison with *Conseil* actions, we consider both district-level regression models that have as outcomes the number of annulled *casillas* and *casilla*-level models that check whether annulment outcomes are associated with each *casilla*'s `eforensics`-frauds estimates. The *casilla*-level approach works for the presidential race, for which districts are not meaningful.

Annulment petitions occur according to Article 75 of the *Ley General del Sistema de Medios de Impugnación en Materia Eletoral* (Cámara de Diputados del H. Congreso de la Unión 2014). Two features of this particularly inform our analysis (SI 1.2.3). First, often irregularities must be "determinative for the outcome of the vote" in order for them to trigger a decision to annul: the vote distortion must be bigger than the margin between the leading and second-place parties in the *casilla*. This echoes the *Conseil*'s emphasis on margin M except now the focus is on each aggregation unit. Second, in 2007 significant changes in the constitution and law were enacted (Langston 2009). Serra (2010) asserts these reforms weakened electoral institutions among other consequences.

In `eforensics` estimation *casillas* are aggregation units except for 2003 where *seccionnes* are aggregation units (SI 1.1.3). For deputy elections we consider plurality-rule *Mayoria Relativa* votes (Kerevel 2010), and W_i contains district winner votes. In the presidential election W_i comes from the winning candidate. Each model includes all units with district fixed effects in x^τ and x^ν for the deputy elections. Mixture probabilities are shown in Figure 1: π_1 is smallest ($\pi_1 = .947$ $[.939, .965]$) and π_2 largest ($\pi_2 = .0528$ $[.0350, .0610]$) in the 2006 presidential election ("Mexico 2006 P") while π_3 is largest in 2003 ($\pi_3 = .000239$ $[.000114, .000359]$). At least one of ρ_{M0}, ρ_{S0}, δ_{M0} and δ_{S0} is negative in each election (Figure 2) (SI 1.2.3). H, ψ_{tW} and ψ_{wW} are largest in "2006 P" (Table 3) (SI 1.2.3). Figure 3 shows φ, ψ_{wW} and ψ_{wM} by district for the 2003 and 2006 deputy elections (2009 is similar): no district has $\varphi > .3$; at most $\psi_{wW} \approx .15$; and in each year one district has $\psi_{wM} > 1$.

Table 6 reports results for annulment binomial logistic regression models. M is the district difference between first- and second-place candidates' vote totals (Table 6(a)), and M_i is the vote difference between the top two candidates at each *casilla* (Table 6(b)).[8] Stolen and manufactured votes have differing relationships to annulment decisions, relationships that vary over time. In the district-level models for deputy elections (Table 6(a)) for 2006 annulments occur more when $(F_w - F_t)/(M + 1)$ is higher, but for 2009 annulments are more frequent when F_t/V is greater; for 2003 the coefficient of $(F_w - F_t)/(M + 1)$ has positive mean but the confidence interval includes zero. The coefficients of F_t/V in 2003 and 2006 and of $(F_w - F_t)/(M + 1)$ in 2009 are negative. Estimates from *casilla*-level models (Table 6(b)) that condition *casilla* annulment decisions on *casilla* `eforensics` estimates are similar: for the 2006 elections $(F_{wi} - F_{ti})/(M_i + 1)$ has positive and F_{ti}/V_i negative coefficients; for 2009 F_{ti}/V_i has positive and $(F_{wi} - F_{ti})/(M_i + 1)$ negative coefficients. Apparently before 2009 stolen votes mattered to TEPJF while in 2009 manufactured votes mattered.

[8]The cases models that are the basis for censoring adjustments have as regressor M or M_i.

3.4 Germany 2005 Legislature

Discussing examples from Afghanistan, France and Mexico we emphasized how `eforensics` provides valid albeit imperfect measures of election frauds, with actions of auditors or courts providing support. We now turn to examples that draw out ambiguities that stem from strategic behavior and administrative failures.

First is Germany, in particular the 2005 federal legislative election. Germany is known for an absence of frauds (SI 1.2.4) and for the presence of at least two kinds of strategic behavior by electors: wasted-vote strategies for single-member district (SMD) seats (*Erststimmen*), and threshold insurance for proportional representation (PR) seats (*Zweitstimmen*) (e.g. Bawn 1999; Shikano, Herrmann and Thurner 2009).

`eforensics` specifications use polling station aggregation units. W_i contains district winners' votes for *Erststimmen* and votes of the party with the most votes overall for *Zweitstimmen*. x^{τ} and x^{ν} include district (*Länder*) fixed effects for *Erststimmen* (*Zweitstimmen*) (SI 1.1.4). Both *Erststimmen* ("Germany 2005 E") and *Zweitstimmen* ("Germany 2005 Z") have appreciable π_2 estimates (Figure 1): respectively $\pi_2 = .0632$ $[.0541, .0772]$ and $\pi_2 = .0322$ $[.0156, .0456]$. $\rho_{M0} < 0$ and $\rho_{S0} < 0$ but $\delta_{M0} \approx \delta_{S0} \approx 0$. Table 3 shows many `eforensics`-fraudulent polling stations—$\varphi = .0423$ and $\varphi = .0145$—and `eforensics`-fraudulent votes: $\psi_{wW} = .0158$ for *Erststimmen* exceeds the estimates of ψ_{wW} found for the Mexican deputy elections; $\psi_{wW} = .00686$ for *Zweitstimmen* exceeds ψ_{wW} for the 2003 and 2009 elections. Figure 3 shows the distributions of φ, ψ_{wW} and ψ_{wM} for *Erststimmen* across districts resemble the distributions found for the Mexican elections, except for the German election always $\psi_{wM} < 1$, and they resemble those for France 2017 apart from one French district with high φ, ψ_{wW} and ψ_{wM} values.

The resemblances suggest strategic behavior contributes to the `eforensics` estimates for France and Mexico much as it does the estimates for Germany, even though `eforensics` also measures aspects associated with *Conseil* and TEPJF annulment decisions. Mexico's deputy elections resemble Germany's *Bundestag* elections in featuring plurality- and PR-rule votes, and both countries include many parties so that wasted-vote strategic behavior is motivated (cf. Kerevel 2010). For all these elections when incremental frauds parameters differ from zero, estimates are negative. France 2017 and Mexico 2003 have extreme frauds parameters that differ from zero, unlike the German elections.

Extreme frauds parameters tend to differ from prior mean zero when conditions that warrent annulment exist (Mexico 2006 P also has $\delta_{S0} \not\approx 0$), but incremental frauds parameters can be negative when only strategic behavior is occurring.

3.5 Ohio 2006 State House

Few `eforensics`-frauds ocur in the Ohio 2006 State House election. Across districts turnout ranges from .22 to .59 while the winner's vote share ranges from .51 to 1.0 with district-level turnout and vote share correlated $-.61$ (SI 1.1.5): probably there is the usual SMD strategic behavior (cf. Berch 1989). `eforensics` estimation includes all precinct aggregation units with W_i being each district winner's votes and x^τ and x^ν including district fixed effects. $\pi_1 = .999$ $[.999, .999997]$ and $H = F_t = F_w = 0$. Even so $\rho_{M0} < 0$ and $\rho_{S0} < 0$. Note that $Z_i = 2$ for at least one MCMC iteration for 567 precincts, while none have $Z_i = 3$.

3.6 Georgia 2020 President, Senate

An important strategy for federal elections in the United States is coordinating voting (Mebane 2000), a ticket-splitting strategy that links elections for president and for the Congress (Alesina and Rosenthal 1995). The coordinating theory suggests that if the Democrat is expected to win the presidential race, then Republican legislative candidates should benefit from votes gained in strategically split tickets. The 2020 elections in Georgia illustrate how such behavior manifests in `eforensics` estimates. Vote counts and procedures have been vetted to such an extent (e.g. Georgia Secretary of State 2020;

Investigations Division 2020) that the probability that genuine frauds occurred is very low—although there are loose ends.

Recounts in Georgia focused on presidential votes—although errors uncovered during the first, hand recount produced corrections in U.S. Senate results—and they strongly suggest that at least for the presidential race the chances of frauds due to manipulation of tabulation software is low. Unfortunately, the use of ballot-marking devices causes concerns (Totenberg 2020). Because the hand recount was "a hand-review of paper ballots that were printed from electronic voting machines, as well as absentee ballots that arrived by mail" (Lee 2020), it provides confidence only for original paper absentee ballots that were recounted. Additionally concerns remain about voter suppression (Ferriss 2020), even though voter mobilization may have helped counter such (King 2020). Voter mobilization introduces strategic considerations, as people become more likely to vote because they believe their allies and opposing voters are more likely to vote.

The two senate contests differ regarding likely strategic behavior. One race ("USS1") had three candidates and incumbent Republican Perdue received the most votes, while the other ("USS2") had 21 candidates (six Republicans and nine Democrats) and Democrat Warnock finished first. Both elections had majority rules: more than fifty percent of votes were required to avoid a runoff election; both elections went to runoffs. A majority outcome in USS1 was more likely than in USS2. Indeed, 36870 more votes were cast for USS1 than for USS2.[9] Warnock with 33.4% of the votes in USS2 received 759250 fewer votes than did Perdue in USS1 (49.0%) and 742862 fewer votes than did USS1 runner-up Ossoff (48.7%). Probably USS1 was treated as more likely to be decisive. With Democrat Biden expected to win the presidency, Perdue should have received some votes due to coordinating voting.

`eforensics` estimations use precinct aggregation units (SI 1.1.6), and for each contest W_i comes from the candidate with the most votes. In all elections π_3 is small (Figure 1): always $\pi_3 < .0019$. π_2 is largest for USS1 ($\pi_2 = .197$ $[.138, .251]$), second-largest for

[9] 44857 more votes were cast for president than for USS1.

president ($\pi_2 = .0614$ [.0240, .104]) and smallest for USS2 ($\pi_2 = .0521$ [.0285, .0846]). Wide HPD intervals for π_1 and π_2 result from apparent multimodality:[10] this indicates lost votes (see section 3.7), presumably due to voter suppression. For all three elections $\rho_{M0} < 0$ but $\delta_{M0} \approx \delta_{S0} \approx 0$ (Figure 2). $\rho_{S0} > 0$ for USS1 but $\rho_{S0} \approx 0$ for president or USS2. For all three elections $F_w - F_t > F_t$. For USS1 $\psi_{wW} = .0475$, bigger than $\psi_{wW} = .0130$ for USS2 or $\psi_{wW} = .00376$ for president (Table 3).[11] $\psi_{wW} = .0475$ comes from $F_w = 111693.4$, which is bigger than the margin between Perdue and Ossoff.

The notable `eforensics`-frauds for Perdue and estimates for Biden and Warnock can be explained as results of strategic behavior, combining mobilization, wasted-vote and coordinating strategies, with `eforensics` estimates also reflecting votes lost due to voter suppression. $\delta_{M0} \approx \delta_{S0} \approx 0$ so $\rho_{S0} > 0$ for USS1 is the only exception to a pattern of incremental frauds parameters being negative when strategic behavior is occurring.

3.7 Florida 2000 and Wisconsin 2016 President

Presidential elections in Florida 2000 and Wisconsin 2016—notorious for lost votes—illustrate how lost votes affect `eforensics` estimates. It is not clear what proportion of the lost votes result from voter suppression efforts: in Florida incompetent election administration caused tens of thousands votes to be lost (Wand, Shotts, Sekhon, Mebane, Herron and Brady 2001; Mebane 2004). Both in Florida (Berman 2015) and Wisconsin (DeCrescenzo and Mayer 2019) there were efforts to suppress voters. In Wisconsin many chose to abstain at least partly due to misinformation campaigns (Jones and Bourbeau 2016; Jamieson 2018). The extent to which the lost votes reflect effects of bad acts is difficult to specify.

`eforensics` estimates use precinct or ward aggregation units (SI 1.1.7 and 1.1.8), and W_i comes from the candidate with the most votes. In light of administrative differences

[10] E.g. across four MCMC chains president π_2 estimates are .0295 [.0223, .0362], .0629 [.0527, .0727], .0975 [.0870, .110], .0558 [.0472, .0655].

[11] For president $F_w = 9129.0$ [7554.5, 10496.7]. For F_1, F_w, ψ_{tW} and ψ_{wW} we show posterior mean [99.5% equal-tailed credible interval].

across counties (Mebane 2004; Mebane and Bernhard 2019), x^τ and x^ν include county fixed effects. As Figure 1 shows, $\pi_3 < .001$ is small while 95% HPD intervals for π_1 and π_2 are wide: e.g., Florida has $\pi_2 = .127$ $[.0648, .157]$ and Wisconsin has $\pi_2 = .253$ $[.146, .306]$; the wide intervals reflect multimodality in π_1 and π_2.[12] $\rho_{S0} < 0$ but $\delta_{M0} \approx \delta_{S0} \approx 0$. $\rho_{M0} < 0$ for Florida and ρ_{M0} is multimodal for Wisconsin (SI 1.2.5).

Both presidential elections likely featured strategic behavior—$\delta_{M0} \approx \delta_{S0} \approx 0$ while $\rho_{M0} < 0$ and $\rho_{S0} < 0$—but F_w in each state apparently includes contributions from lost votes. For Florida $\psi_{wW} = .0230$ and for Wisconsin $\psi_{wW} = .0297$, which correspond respectively to $F_w = 66915.4$ and $F_w = 41692.9$, bigger than the respective margins of victory. Almost 60,000 votes were lost due to ballots and election administration in Florida 2000 (Wand et al. 2001; Mebane 2004), and about 13,000 voters were deterred in two counties in Wisconsin 2016 (DeCrescenzo and Mayer 2019)—for these counties $F_w = 11509.3$. F_w is similar to the number of lost votes: votes Democratic candidates lost appear as `eforensics`-fraudulent votes for Republican leader candidates. The evidence from the cited and other studies is not that tens of thousands of votes were wrongly added to the winning candidates' totals. But `eforensics` apparently captures the distortion in the balance between the top candidates in each race caused by the referent problems.

3.8 California 2006 Governor

We use the California 2006 gubernatorial election to illustrate that while substantively motivated covariates may materially change `eforensics` estimates, whether the changes are revealing or masking bad acts can be unclear. Of six candidates, Republican Schwarzenegger got the most votes (SI 1.2.6). `eforensics` estimation uses precinct aggregation units (SI 1.1.9). W_i comes from Schwarzenegger's votes. Another version of the model includes Republican and Democrat precinct partisan registration variables in x^τ and x^ν (SI 1.2.6). Including the covariates produces much smaller π_2 and F_w: without the

[12] Multimodality also occurs in π_1 and π_2 if fixed effects are omitted.

covariates $\pi_2 = .0881$ [.0812, .0980], $H = 1309$ and $\psi_{wW} = .0303$ (from $F_w = 123137.8$) but with them $\pi_2 = .0113$ [.00866, .0145], $H = 81$ and $\psi_{wW} = .00284$ (from $F_w = 11543.9$). For both specifications $\rho_{S0} < 0$, $\delta_{M0} < 0$ and $\delta_{S0} < 0$ while $\rho_{M0} > 0$ without and $\rho_{M0} < 0$ with the covariates. The covariates reveal turnout tends to be higher (lower) as Republican (Democratic) registration is higher, and Schwarzenegger tends to get a higher (lower) proportion of votes as Republican (Democratic) registration is higher.

Interpreting the reduction in `eforensics`-frauds is difficult. $\delta_{M0} < 0$ and $\delta_{S0} < 0$ suggests bad acts occur. Without knowing how partisan registration relates to bad acts it's difficult to say whether introducing the covariates masks bad acts—are these as frequent as $F_w = 123137.8$ or only $F_w = 11543.9$?

3.9 Argentina 2015 President

Zhang, Alvarez and Levin (2019, 7) find frauds in the first round of the Argentina 2015 president election—their simulation-based random forest method finds only 86.3% of *mesas* are "Clean." Using *mesa* aggregation units (SI 1.1.10), `eforensics` estimates scant frauds: $\pi_1 = .997$ [.994, .9998], $H = 25$ ($\varphi = .000271$) and $\psi_{wW} = .000163$ ($F_w = 1468$). $\rho_{M0} < 0$ but $\rho_{S0} \approx \delta_{M0} \approx \delta_{S0} \approx 0$.[13] This suggests strategic behavior: there were six candidates with a majority rule, and the top three received proportions .21, .34 and .37 of the votes; the outcome resembles Bouton and Gratton (2015)'s "Duverger's hypothesis equilibrium."

3.10 Turkey 2017 Referendum and 2011 and 2015 Legislature

Studying the Constitutional Referendum in Turkey in 2017, Klimek et al. (2018, 1 and 9) "find systematic and highly significant statistical support for the presence of both ballot stuffing and voter rigging. In 11% of stations we find signs for ballot stuffing," moreover "the cumulative effect of the distortions in small stations tipped the results toward a

[13] Excluding *mesas* with less than 100 ballots cast, $\pi_1 = .995$ [.991, .9994], $H = 32$, $F_w = 2483.0$ and $\rho_{M0} \approx \rho_{S0} \approx \delta_{M0} \approx \delta_{S0} \approx 0$.

majority of Yes votes. If the small stations would have followed the trends observed in larger ones, the vote percentage would not have crossed the 50% line."

`eforensics` estimates agree that bad acts occurred but disagree about the magnitude. Using polling station aggregation units (SI 1.1.11), W_i from "Yes" votes and region fixed effects in x^τ and x^ν produces $\pi_2 = .0000724$ [2.18e-08, .000201] but $\pi_3 = .0345$ [.0335, .0354]. $\rho_{M0} < 0$, $\rho_{S0} < 0$, $\delta_{M0} < 0$, $\delta_{S0} < 0$, $H = 4980$ and $\psi_{wW} = .0130$, which corresponds to $F_w = 325548.2$. $\varphi = .0291 < .11$, and F_w is smaller than the margin in our data of 1303152 votes between "Yes" and "No" (SI 1.2.7). Most `eforensics`-fraudulent votes are stolen: $F_t = 96543.4 < F_w - F_t = 229004.7$.

More `eforensics`-frauds occur for the 2011 and 2015 legislative elections (Bakıner 2017). W_i comes from each district's vote-leading party, and x^τ and x^ν include district fixed effects (SI 1.1.11). For these elections π_2 is much greater than in 2017: for 2011 $\pi_2 = .485$ [.484, .486]. The 2015 elections have π_2 high and multimodal, which perhaps stems from lost votes: between 2011 and June 2015 domestic turnout declined from 87.2% to 86.43% (Kesgin 2012; Cop 2016); violence before November 2015 deterred two parties' mobilization efforts (Sözen 2016, 869). For all elections $\rho_{M0} < 0$, $\rho_{S0} < 0$, $\delta_{M0} < 0$ and $\delta_{S0} < 0$ except $\rho_{M0} \approx \delta_{S0} \approx 0$ for June 2015 and $\rho_{M0} \approx 0$ for November 2015. H, ψ_{tW} and ψ_{wW} are always greater than in 2017. Distributions of φ, ψ_{wW} and ψ_{wM} across districts (Figure 4) differ from those observed in Figure 3: $\varphi > 0$ and $\psi_{wW} > 0$ for every district in 2011 and November 2015, and for both elections $\psi_{wM} > 1$ for several districts.

3.11 Russia 2000–2020 President, Duma and Referendum

Russia had elections that increasingly became mere "election-type events" in and following the 2004 presidential election (e.g. Gel'man 2015). `eforensics` illustrates aspects of the degradation of these events from 2000 through 2020. With polling station aggregation units (SI 1.1.12), for each election W_i contains votes for the alternative that received the most votes; in Duma SMD elections W_i comes from each district's vote leader, and district fixed effects are in x^τ and x^ν.

The elections exhibit many `eforensics`-frauds, generally increasing over time. π_2 and π_3 go from $\pi_2 = .101$ $[.0886, .118]$ and $\pi_3 = .0254$ $[.0217, .0300]$ in 2000 to $\pi_2 = .352$ $[.335, .369]$ and $\pi_3 = .116$ $[.0867, .177]$ in 2020 (SI 1.2.8). π_2 is smaller and π_3 greater for 2003, 2007, 2011 and 2012 than are Klimek et al. (2012, Table S3)'s fraud fractions f_i and f_e, and $\varphi < f_\mathrm{i} + f_\mathrm{e}$ (SI 1.2.8). During 2000–2020 π_2 exceeds the estimate for Afghanistan 2009, and after 2003 π_3 exceeds both Afghanistan's and Turkey's estimates. $\delta_{M0} > 0$ for every election except 2020: $\delta_{S0} > 0$ for 2011 and $\delta_{S0} < 0$ for 2000, 2003 SMD, 2016 PR, 2016 SMD, 2018 and 2020. $\rho_{M0} > 0$ for every election except 2004 and 2008: $\rho_{S0} > 0$ for 2011 and $\rho_{S0} < 0$ for 2003 SMD, 2016 PR and 2016 SMD. $\delta_{M0} > 0$, $\delta_{S0} \not\approx 0$ and $\rho_{M0} > 0$ suggest bad acts occurred. Manufactured votes outnumber stolen votes; from 2000 to 2020 ψ_{tW} increased from .0149 to .113, while $\psi_{wW} - \psi_{tW}$ increased from .0146 to .0530. From 2000 to 2020 φ increased from .106 to .448.

3.12 Armenia 2003 President

Illustrating `eforensics` estimates' use as outcome measures is the Armenia 2003 presidential election for which Hyde (2007) uses as-if random assignment of election observers to polling stations to study how observers affect frauds. `eforensics` estimates use polling station aggregation units (SI 1.1.13). W_i contains votes for the candidate with the most votes. Four variables indicate whether a polling station was never observed (O1), was observed only during the first election round (O2), only during the second round (O3), or in both rounds (O4). O2, O3 and O4 are included in x^τ, x^ν, x^ι and x^υ. π_2 and π_3 are greater in round two ($\pi_2 = .257$ $[.204, .31]$, $\pi_3 = .0437$ $[.032, .0552]$) than in round one ($\pi_2 = .146$ $[.127, .168]$, $\pi_3 = .0252$ $[.0174, .0335]$). Multimodality is apparent in round two: perhaps votes were lost then due to intimidation (ODHR 2003, 1 and 11).

Figure 2 shows that for never-observed polling stations $\rho_{M0} > 0$ and $\delta_{M0} < 0$ for round one and $\rho_{S0} < 0$ and $\delta_{S0} < 0$ for round two. Of ρ_{Mj}, ρ_{Sj}, δ_{Mj} and δ_{Sj} for

$j \in \{$O2, O3, O4$\}$, the only estimates that differ from zero are $\rho_{MO3} > 0$, $\rho_{SO2} < 0$ and $\delta_{SO4} < 0$ for round one and $\rho_{SO4} < 0$ for round two. At least $\delta_{M0} < 0$ and $\delta_{S0} < 0$ suggest that bad acts occurred, and the estimates show few differences across observer statuses. φ is larger in round two (.145 versus .258), but ψ_{wW} is larger in round one (.0783 [.0752, .0809] versus .0714 [.0641, .0762]); the latter is true even though F_w is greater in round two (51845.7 [49807.7, 53571.1] versus 70216.6 [62967.5, 74854.7]).

Table 7 shows φ by observer status in each round, with tests of independence between `eforensics`-frauds classifications and observer variables. For both rounds φ is lower with O2 or O4 than with O1 or O3 and tests reject independence: as Hyde (2007, 57) finds, round one observation has persistent effects.

Another approach is to compare ψ_{tW} and $\psi_{wW} - \psi_{tW}$ across observer statuses. Only six of 24 differences (marked *) have the sign expected if observation reduces ψ_{tW} or $\psi_{wW} - \psi_{tW}$ (Table 8).[14] There is no consistent association between observer status and the proportional magnitudes of `eforensics`-frauds.

Hyde (2007) demonstrates effects of as-if random observation on vote shares, but `eforensics` supports disaggregating into effects on `eforensics`-frauds occurrence, on `eforensics`-frauds magnitudes and on true turnout and votes (SI 1.2.9).[15] Observation reduces the proportion of polling stations with `eforensics`-frauds but does not consistently reduce the proportion of leader votes that are `eforensics`-fraudulent.

4 Discussion

Using only aggregation unit summaries of an election—numbers of electors and of votes received by ballot alternatives—`eforensics` provides valid but imperfect measures of election frauds. Estimates do not measure frauds that benefit nonleading candidates nor purely procedural frauds: comparison with Afghanistan 2009 audit decisions illustrate this.

[14]For instance, the first entry in the row labeled "O1−O2" shows ψ_{tW} for O1 minus ψ_{tW} for O2 for round one.

[15]Note that $\gamma_{O3} > 0$, $\beta_{O2} < 0$ and $\beta_{O4} < 0$ for round 1; $\gamma_{O4} > 0$ and $\beta_{O4} < 0$ for round 2.

To the extent that procedural deficiencies are effectively proxies for distortions `eforensics` can measure, `eforensics` estimates should nonetheless capture the frauds.

`eforensics` manifests ambiguities related to strategic behavior and administrative failures. Both elections that have demonstrated frauds—e.g., annulled outcomes in France and Mexico—and elections that feature only elector strategic behavior (e.g. Germany 2005) have `eforensics`-frauds, meaning they have $H > 0$, $F_t > 0$ and $F_w > 0$. An election for which $H = F_t = F_w = 0$ (Ohio 2006) has $\rho_{M0} < 0$ and $\rho_{S0} < 0$. Votes lost from nonleading candidates in Florida 2000 and Wisconsin 2016 appear added to F_w.

But the posterior distributions of `eforensics` parameters—particularly of mixture probabilities π_2, π_3 and frauds parameters ρ_{M0}, ρ_{S0}, δ_{M0} and δ_{S0}—help discriminate frauds from strategies from lost votes. When votes are malevolently distorted δ_{M0} and δ_{S0} tend to differ from their prior means, and π_3 tends not to be very small. ρ_{M0} and ρ_{S0} can differ from their prior means when only strategic behavior occurs—in particular they can be negative—although such estimates (and especially positive estimates) can also signal that bad acts occur. Combination of $\rho_{M0} < 0$ or $\rho_{S0} < 0$ with $\delta_{M0} \approx \delta_{S0} \approx 0$ appears typical in elections with strategic behavior but not bad acts. Multimodality in π_2 appears to indicate votes are lost.

The prevalence and magnitude of `eforensics`-frauds are not alone sufficient indicators of whether there are bad acts, strategic behavior or administrative failures. `eforensics`-frauds can be rare while still F_t and F_w strongly associate with demonstrations of bad acts: e.g., $\pi_1 = .990$ in France 2017 yet election annulments conditionally associate with `eforensics` measures. In such cases it can be important, as apparently it was to the *Conseil Constitutionnel* and to TEPJF, to consider not the overall probability of `eforensics`-frauds as measured by mixture probabilities but the incidence of `eforensics`-fraudulent aggregation units in each legislative district and how in each race F_t compares to votes cast and F_w compares to the winning margin. A moderately large π_2 need not mean that bad acts occur: e.g., $\pi_2 = .0632$ in Germany 2005 *Erststimmen* and

$\pi_2 = .197$ in Georgia 2020 U.S. Senate, which both feature strategic behavior.[16] We find that high values of π_3 always go with nonzero values of δ_{M0} or δ_{S0} and so strongly suggest that malevolent distortions have occurred.

Not only to identify which election aggregation units have frauds but to measure the number of fraudulent votes at each unit are achievements of potentially great importance. Such measures can be especially useful in authoritarian settings. We show that `eforensics`-frauds offer perspective on the increasing scale of fraudulence in Russian elections as well as the problematic nature of some recent Turkish elections. `eforensics`-frauds are scant in the Argentina 2015 election (round one). The measures reveal aspects of as-if random observer effects in Armenia 2003.

It is important to be able to determine whether the results of an election in which frauds are alleged can be credited: did the apparent winner really win? `eforensics` can contribute to answering this question. Imperfections `eforensics` exhibits as a model for measuring malevolent distortions in votes should not prevent its being used if it is used with appropriate attention to parametric nuances. `eforensics` can complement conclusions of observer and inspection missions: e.g., `eforensics` estimates agree quantitatively with the conclusions of the audit in Afghanistan 2009. As does Georgia's Secretary of State, `eforensics` finds no evidence of serious problems in the 2020 presidential race in the state.

We have not tried to interpret every feature of `eforensics` estimates. For instance, in our examples parameters other than π_2 sometimes have multimodal posteriors. Whether such features are meaningful can be determined as part of future work to further verify when and how `eforensics` accurately measures distortions in the mapping from electors' intentions to election outcomes.

[16] Caveat: plus lost votes in Georgia.

References

Alesina, Alberto and Howard Rosenthal. 1995. *Partisan Politics, Divided Government, and the Economy.* New York: Cambridge.

Alvarez, R. Michael, Jennifer Morrell, Ronald Rivest, Philip Stark and Charles Stewart, III, eds. 2019. *Election Auditing: Key Issues and Perspectives, Summary Report, Election Audit Summit, December 7–8, 2018.* Cambridge, MA: MIT Election Data & Science Lab.

Bakıner, Onur. 2017. How Did We Get Here? Turkey's Slow Shift to Authoritarianism. In *Authoritarian Politics in Turkey: Elections, Resistance and the AKP*, ed. Bahar Başer and Ahmet Erdi Öztürk. London: I. B. Tauris pp. 21–46.

Bawn, Kathleen. 1999. "Voter Responses to Electoral Complexity: Ticket Splitting, Rational Voters and Representation in the Federal Republic of Germany." *British Journal of Political Science* 29(3):487–505.

Berch, Neil. 1989. "Another Look at Closeness and Turnout: The Case of the 1979 and 1980 Canadian National Elections." *Political Research Quarterly* 46(2):421–432.

Berman, Ari. 2015. "How the 2000 Election in Florida Led to a New Wave of Voter Disenfranchisement: A Botched Voter Purge Prevented Thousands from Voting—and Empowered a New Generation of Voting-rights Critics." *The Nation* . July 28, 2015.

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

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.

Cámara de Diputados del H. Congreso de la Unión. 2014. "Ley General del Sistema de Medios de Impugnación en Materia Electoral." URL `http://www.diputados.gob.mx/LeyesBiblio/pdf/149.pdf` (accessed July 3, 2014).

Cantú, Francisco. 2019. "The Fingerprints of Fraud: Evidence from Mexico's 1988 Presidential Election." *American Political Science Review* 113(3):710–726.

Conseil Constitutionnel. 2017. "Décision n° 2017-5098/5159 AN du 18 décembre 2017." `https://www.conseil-constitutionnel.fr/decision/2017/20175098_5159AN.htm`.

Cop, Burak. 2016. "The June 2015 Legislative Election in Turkey." *Electoral Studies* 41:213–216.

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

Dahl, Robert A. 1956. *A Preface to Democratic Theory*. Chicago: University of Chicago.

DeCrescenzo, Michael G. and Kenneth R. Mayer. 2019. "Voter Identification and Nonvoting in Wisconsin–Evidence from the 2016 Election." *Election Law Journal* 18(4).

Democracy International. 2009. "Results by PS - PDF data only - 16 Sept 2009.xlsx.".

Denwood, Matthew J. 2016. "runjags: An R Package Providing Interface Utilities, Model Templates, Parallel Computing Methods and Additional Distributions for MCMC Models in JAGS." *Journal of Statistical Software* 71(9):1–25.

Eisenstadt, Todd A. 2007. "The Origins and Rationality of the 'Legal versus Legitimate' Dichotomy Invoked in Mexico's 2006 Post-Electoral Conflict." *PS: Political Science and Politics* 40(1):39–43.

Electoral Complaints Commission. 2010. "Electoral Complaints Commission Final Report: 2009 Presidential and Provincial Council Elections." `ECC Final Report 2009.pdf`.

Electoral Complaints Commission, Islamic Republic of Afghanistan. 2009. "Annex a: Audit Results By Polling Station." Annex for 18 October 2019 Letter to Azizullah Ludin, Chairman, Independent Election Commission of Afghanistan: `ECC2009_October18ANNEXADecisionAuditandRecount_20091018.pdf`.

Ferrari, Diogo, Walter Mebane, Kevin McAlister and Patrick Wu. 2019. *Election Forensics: Positive Empirical Models of Election Fraud*. R package version 0.0.4 (Supported by NSF grant SES 1523355). Initial version: August 27, 2019. URL: `https://github.com/UMeforensics/eforensics_public`.

Ferriss, Susan. 2020. "Georgia Hotbed for Voter Suppression Tactics." October 19, 2020, URL: `https://publicintegrity.org/politics/elections/us-polling-places/georgia-hotbed-for-voter-suppression-tactics/`.

Flegal, James M. and J. Hughes. 2012. "mcmcse: Monte Carlo Standard Errors for MCMC." Riverside, CA and Minneapolis, MN. R package version.

Flegal, James M., Murali Haran and Galin L Jones. 2008. "Markov Chain Monte Carlo: Can We Trust the Third Significant Figure?" *Statistical Science* 23(2):250–260.

Gel'man, Vladimir. 2015. *Authoritarian Russia: Analyzing Post-Soviet Regime Changes*. Pittsburgh: University of Pittsburgh.

Georgia Secretary of State. 2020. "Risk-Limiting Audit Report, Georgia Presidential Contest, November 2020, November 19th, 2020." URL: `https://sos.ga.gov/admin/uploads/11.19_.20_Risk_Limiting_Audit_Report_Memo_1.pdf`.

Gong, Lei and James M. Flegal. 2016. "A Practical Sequential Stopping Rule for High-dimensional Markov Chain Monte Carlo." *Journal of Computational and Graphical Statistics* 25(3):684–700.

Harfst, Philipp, André Blais and Damien Bol. 2018. Voting Strategically in Two-Vote Elections. In *The Many Faces of Strategic Voting: Tactical Behavior in Electoral Systems Around the World*, ed. Laura B. Stephenson, John H. Aldrich and André Blais. Ann Arbor, MI: University of Michigan pp. 150–177.

Hyde, Susan D. 2007. "The Observer Effect in International Politics: Evidence from a Natural Experiment." *World Politics* 60:37–63.

Investigations Division, Georgia Secretary of State. 2020. "Georgia Secretary of State/Georgia Bureau of Investigation, ABM Signature Audit Report." Case Number: SEB2020- 257; Case Name: Cobb County- Absentee Ballot Signature Verifications; Date: 12/29/2020; URL: `https://sos.ga.gov/admin/uploads/Cobb%20County%20ABM%20Audit%20Report%2020201229.pdf`.

Jamieson, Kathleen Hall. 2018. *Cyberwar: How Russian Hackers and Trolls Helped Elect a President—What We Don't, Can't, and Do Know*. Oxford: Oxford UP.

Jones, Galin L., Murali Haran, Brian S. Caffo and Ronald Neath. 2006. "Fixed-width Output Analysis for Markov Chain Monte Carlo." *Journal of the American Statistical Association* 101(476):1537–1547.

Jones, Malia and Caitlin Bourbeau. 2016. "How And Where Trump Won Wisconsin in 2016: Lower Voter Turnout And Community Size Defined The Presidential Vote." November 21, 2016, URL: `https://www.wiscontext.org/how-and-where-trump-won-wisconsin-2016`.

Kawai, Kei and Yasutora Watanabe. 2013. "Inferring Strategic Voting." *American Economic Review* 103(2):624–662.

Kerevel, Yann. 2010. "The Legislative Consequences of Mexico's Mixed-member Electoral System, 2000–2009." *Electoral Studies* 29(4):691–703.

Kesgin, Barış. 2012. "The 2011 Parliamentary Election in Turkey." *Electoral Studies* 31(4):867–870.

King, Maya. 2020. "Georgia's Legacy of Voter Suppression is Driving Historic Black Turnout." *Politico* . October 26, 2020.

Klimek, Peter, Raúl Jiménez, Manuel Hidalgo, Abraham Hinteregger and Stefan Thurner. 2018. "Forensic Analysis of Turkish Elections in 2017–2018." *PLOS One* 13(10):e0204975.

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.

Langston, Joy. 2009. "Las reformas al COFIPE, 2007." *Política y gobierno* 16(2):245–272.

Lee, Michelle Ye Hee. 2020. "Everything You Need to Know About the Georgia Presidential Recount." *Washington Post* . November 12, 2020.

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

Maddala, G. S., ed. 1983. *Limited-dependent and Qualitative Variables in Econometrics*. New York: Cambridge.

Mebane, Jr., Walter R. 2000. "Coordination, Moderation, and Institutional Balancing in American Presidential and House Elections." *American Political Science Review* 94:37–57.

Mebane, Jr., Walter R. 2004. "The Wrong Man is President! Overvotes in the 2000 Presidential Election in Florida." *Perspectives on Politics* 2:525–535.

Mebane, Jr., Walter R. 2008. Election Forensics: The Second-digit Benford's Law Test and Recent American Presidential Elections. 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.

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. New York: Princeton pp. 206–216.

Mebane, Jr., Walter R. and Matthew Bernhard. 2019. Voting Technologies, Recount Methods and Votes in Wisconsin and Michigan in 2016. In *Financial Cryptography and Data Security: FC 2018*, ed. Aviv Zohar, Ittay Eyal, Vanessa Teague, Jeremy Clark, Andrea Bracciali, Federico Pintore and Massimiliano Sala. Berlin, Germany: Springer pp. 196–209.

Montgomery, Jacob M., Santiago Olivella, Joshua D. Potter and Brian F. Crisp. 2015. "An Informed Forensics Approach to Detecting Vote Irregularities." *Political Analysis* 23(4):488–505.

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

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

ODHR. 2003. "Republic of Armenia Presidential Election 19 February and 5 March 2003 Final Report." OSCE Office for Democratic Institutions and Human Rights, Warsaw, 28 April 2003.

Pemstein, Daniel, Stephen A. Meserve and James Melton. 2010. "Democratic Compromise: A Latent Variable Analysis of Ten Measures of Regime Type." *Political Analysis* 18(4):426–

449.

Pettigrew, Stephen. 2017. "The Racial Gap in Wait Times: Why Minority Precincts Are Underserved by Local Election Officials." *Political Science Quarterly* 132(3):527–547.

Plummer, Martyn, Alexey Stukalov and Matt Denwood. 2016. "rjags: Bayesian Graphical Models using MCMC." URL `https://cran.r-project.org/web/packages/rjags/index.html`, linked to JAGS 4.2.0.

Riker, William H. 1982. *Liberalism Against Populism: A Confrontation Between the Theory of Democracy and the Theory of Social Choice*. Prospect Heights, IL: Waveland.

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

Rundlett, Ashlea and Milan W. Svolik. 2016. "Deliver the Vote! Micromotives and Macrobehavior in Electoral Fraud." *American Political Science Review* 110(1):180–197.

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

Serra, Gilles. 2010. "Mexico's Electoral Reform of 2007: A Case of De-Democratization and Partyarchy." Working paper, URL: `http://www.nuffield.ox.ac.uk/politics/papers/2010/Serra_partyarchy.pdf`.

Shikano, Susumu, Michael Herrmann and Paul W. Thurner. 2009. "Strategic Voting under Proportional Representation: Threshold Insurance in German Elections." *West European Politics* 32(3):634–656.

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

Sözen, Yunus. 2016. "The Parliamentary Elections in Turkey, November 2015." *Electoral Studies* 43:200–203.

Stephenson, Laura B., John H. Aldrich and André Blais, eds. 2018. *The Many Faces of Strategic Voting: Tactical Behavior in Electoral Systems Around the World*. Ann Arbor, MI: University of Michigan.

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

Totenberg, Amy. 2020. "Curling v. Raffensperger, Case 1:17-cv-02989-AT Document 964 (2020)." United States District Court For The Northern District Of Georgia Atlanta Division. 10/11/20, URL: `https://www.courthousenews.com/wp-content/uploads/2020/10/ga-voting.pdf`.

Wand, Jonathan, Kenneth Shotts, Jasjeet S. Sekhon, Walter R. Mebane, Jr., Michael Herron and Henry E. Brady. 2001. "The Butterfly Did It: The Aberrant Vote for Buchanan in Palm Beach County, Florida." *American Political Science Review* 95:793–810.

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

Zhang, Mali, R. Michael Alvarez and Ines Levin. 2019. "Election Forensics: Using Machine Learning and Synthetic Data for Possible Election Anomaly Detection." *PLOS ONE* 14(10):e0223950.

Table 1: Afghanistan 2009 President Election `eforensics` Estimates

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.907	.893	.926
	π_2	Incremental Fraud	.0351	.0201	.0469
	π_3	Extreme Fraud	.0579	.0530	.0621
turnout	γ_0	(Intercept)	$-.622$	$-.667$	$-.579$
vote choice	β_0	(Intercept)	$-.107$	$-.155$	$-.0434$
incremental frauds	ρ_{M0}	(Intercept)	.431	.305	.718
	ρ_{S0}	(Intercept)	$-.775$	-1.08	$-.593$
extreme frauds	δ_{M0}	(Intercept)	3.74	3.34	4.10
	δ_{S0}	(Intercept)	2.11	1.75	2.65

units `eforensics`-fraudulent: 1622 fraudulent, 21236 not fraudulent
manufactured votes $F_t = 483730.5$ $[475471.5, 490714.9]^c$
total `eforensics`-fraudulent votes $F_w = 612639.5$ $[604684.2, 618758.9]^c$

Note: `eforensics` model parameter estimates (posterior means and credible intervals). $n = 22858$ polling station units. [a] 95% HPD lower bound. [b] 95% HPD upper bound. [c] posterior mean [99.5% credible interval].

Table 2: Afghanistan 2009 President Election `eforensics` Effects

Type	lo	up
turnout (τ_0)	.324	.342
vote choice (ν_0)	.470	.493
incremental frauds manufactured (ι_0^M)	.470	.486
incremental frauds stolen (ι_0^S)	.256	.263
extreme frauds manufactured (υ_0^M)	.990	.994
extreme frauds stolen (υ_0^S)	.965	.984

Note: 99.5% credible intervals for `eforensics` parameter estimates transformed into proportions τ_0, ν_0, ι_0^M, ι_0^S, υ_0^M, υ_0^S using equations (2a)–(2d), omitting random effects.

Table 3: National or Statewide Election `eforensics` Frauds

Election	aggregation units[a]		proportions of leader votes fraudulent[b]	
	frauds	no frauds	manufactured proportion	fraudulent proportion
Afghanistan 2009	1622	21236	.156 [.154, .159]	.198 [.195, .200]
France 2017 T2	200	35422	.000662 [.000595, .000740]	.00101 [.000933, .00110]
Mexico 2003	318	62828	.00154 [.00137, .00168]	.00205 [.00183, .00222]
Mexico 2006 D	2652	127796	.00702 [.00645, .00752]	.00994 [.00918, .0106]
Mexico 2006 P	3927	126841	.0168 [.0151, .0177]	.0296 [.0268, .0311]
Mexico 2009	426	137919	.00201 [.00185, .00214]	.00265 [.00246, .00282]
Germany 2005 E	3751	84929	.0110 [.0102, .0116]	.0158 [.0148, .0166]
Germany 2005 Z	1289	87391	.00441 [.00409, .00475]	.00686 [.00639, .00731]
Ohio 2006	0	11123	0	0
Georgia 2020 Pres	20	2557	.00113 [.000914, .00133]	.00376 [.00311, .00432]
Georgia 2020 USS1	300	2277	.0165 [.0158, .0171]	.0475 [.0448, .0499]
Georgia 2020 USS2	39	2538	.00430 [.00401, .00454]	.0130 [.0114, .0141]
Florida 2000	289	5652	.0111 [.0105, .0116]	.0230 [.0217, .0240]
Wisconsin 2016	475	2919	.00763 [.00606, .00848]	.0297 [.0251, .0333]
California 2006	1309	21511	.0245 [.0232, .0257]	.0303 [.0291, .0316]
California 2006 r	81	22739	.00210 [.00173, .00234]	.00284 [.00227, .00313]
Argentina 2015 R1	25	92179	.0000472 [.0000348, .0000569]	.000163 [.000120, .000195]
Turkey 2017	4980	166372	.00387 [.00376, .00398]	.0130 [.0126, .0134]
Turkey 2011	129400	70155	.0262 [.0234, .0317]	.136 [.121, .165]
Turkey 2015 Jun	11804	162046	.00955 [.00605, .0150]	.0281 [.0225, .0343]
Turkey 2015 Nov	75086	99533	.0222 [.0142, .0250]	.103 [.033, .128]
Russia 2000	9640	81666	.0149 [.0139, .0156]	.0295 [.0278, .0307]
Russia 2003 PR	9211	85866	.0435 [.0418, .0444]	.0811 [.0791, .0822]
Russia 2003 SMD	11610	82696	.0348 [.0333, .0361]	.0622 [.0597, .0644]
Russia 2004	17110	78314	.0394 [.0361, .0411]	.0643 [.0591, .0669]
Russia 2007	25050	70752	.0699 [.0671, .0715]	.106 [.103, .108]
Russia 2008	24446	71802	.046 [.0428, .0483]	.0772 [.0721, .0806]
Russia 2011	16641	78525	.0832 [.0808, .0845]	.142 [.139, .144]
Russia 2012	18345	77068	.0532 [.051, .0542]	.0801 [.0772, .0816]
Russia 2016 PR	18541	78325	.128 [.123, .131]	.176 [.171, .179]
Russia 2016 SMD	16245	80418	.111 [.107, .113]	.150 [.146, .153]
Russia 2018	24509	73164	.0532 [.0513, .0544]	.0727 [.07, .0742]
Russia 2020 ref	43160	53079	.113 [.108, .116]	.166 [.16, .169]
Armenia 2003 R1	256	1507	.0445 [.0423, .0465]	.0783 [.0752, .0809]
Armenia 2003 R2	454	1308	.0458 [.0404, .0494]	.0714 [.0641, .0762]

Note: [a] units classified as `eforensics`-fraudulent (H) or not `eforensics`-fraudulent ($n - H$). [b] proportions are respectively posterior mean [99.5% lower bound, 99.5% upper bound] of $\psi_{tW} = F_t / \sum_{i=1}^{n} W_i$ and $\psi_{wW} = F_w / \sum_{i=1}^{n} W_i$.

Table 4: Afghanistan 2009, Audit Decisions and `eforensics` Classifications

`eforensics` classification	Sampling Frame: "valid votes" frame decision: Valid	"valid votes" frame decision: Fraudulent	"total votes" frame decision: Valid	"total votes" frame decision: Fraudulent
not fraudulent	20	51	13	21
fraudulent	4	106	11	17
log odds ratio[a]	2.34 (.57)		−.0443 (.52)	

Note: Audited polling stations in which Karzai leads. [a] subtable log odds ratio (standard error)

Table 5: Conseil Constitutionnel cases regressed on `eforensics`-frauds pooled by district

regressor	(a) cases	(b) annulments
Intercept	-2.708	-8.386
	$[-2.81, -2.60]$	$[-8.83, -7.93]$
F_t/V	25.44	39.01
	$[1.13, 49.7]$	$[9.64, 68.2]$
$(F_w - F_t)/(M+1)$	1.067	4.874
	$[.253, 1.91]$	$[3.57, 6.28]$
n of districts	572	302

Note: binomial logistic regressions of counts of French 2017 National Assembly election Conseil Constitutionnel cases and annulments (by district) on `eforensics`-frauds estimates using second round votes. Coefficient Normal approximation mean and 95% confidence interval are shown. Annulment model adjusts for censoring. M is the vote count difference between first and second in each district.
Upper bounds of F_t/V and $(F_w - F_t)/(M+1)$ (lower bounds are always zero): .0404 and 1.02.

Table 6: Mexico TEPJF cases regressed on `eforensics` frauds

(a) number of units annulled (pooled by district)

	Year		
regressor	2003	2006	2009
Intercept	-6.361	-7.023	-8.176
	$[-6.49, -6.23]$	$[-7.16, -6.88]$	$[-8.24, -8.10]$
F_t/V	-58.09	-227.6	1406
	$[-96.2, -19.7]$	$[-284, -169]$	$[1250, 1550]$
$(F_w - F_t)/(M+1)$	1.782	12.88	-15.97
	$[-.673, 4.48]$	$[11.2, 14.4]$	$[-19.0, -12.8]$
n of districts	62	53	37

(b) *casilla* annulled

	Year and Election		
regressor	2006 D	2006 P	2009
Intercept	-7.170	-6.280	-7.839
	$[-7.18, -7.16]$	$[-6.30, -6.26]$	$[-7.84, -7.83]$
F_{ti}/V_i	-7.647	-8.952	1157
	$[-8.69, -6.61]$	$[-9.79, -8.01]$	$[7.80, 6780]$
$(F_{wi} - F_{ti})/(M_i+1)$	$.2686$	$.1779$	-2169
	$[.130, .430]$	$[.140, .227]$	$[-12300, -5.40]$
n of *casillas*	5905	26168	3716

Note: binomial logistic regressions of (a) counts by district or (b) occurrences by unit of TEPJF annulments on `eforensics`-frauds estimates. Aggregation units are *secciones* in 2003, *casillas* in other elections. Normal approximation mean and 95% confidence interval are shown. Annulment models adjust for censoring. M or M_i is the vote count difference between first and second (a) in each district or (b) at each *casilla*.
(a) Upper bounds of F_t/V and $(F_w - F_t)/(M+1)$ (lower bounds are always zero): 2003, .0216 and .265; 2006, .0445 and .272; 2009, .0255 and .218. (b) Upper bounds of F_{ti}/V_i and $(F_{wi} - F_{ti})/(M_i+1)$: 2006 D, .445 and 44.7; 2006 P, .390 and 41.3; 2009, .619 and 3.14.

Table 7: Armenia 2003 President Election `eforensics`-fraud Classifications by Observer Status

	Observer Status[a]				Independence
	O1	O2	O3	O4	Test p-value[b]
round 1	.160	.109	.208	.107	.000500
round 2	.286	.223	.273	.223	.0404
n^c	754	385	260	364	

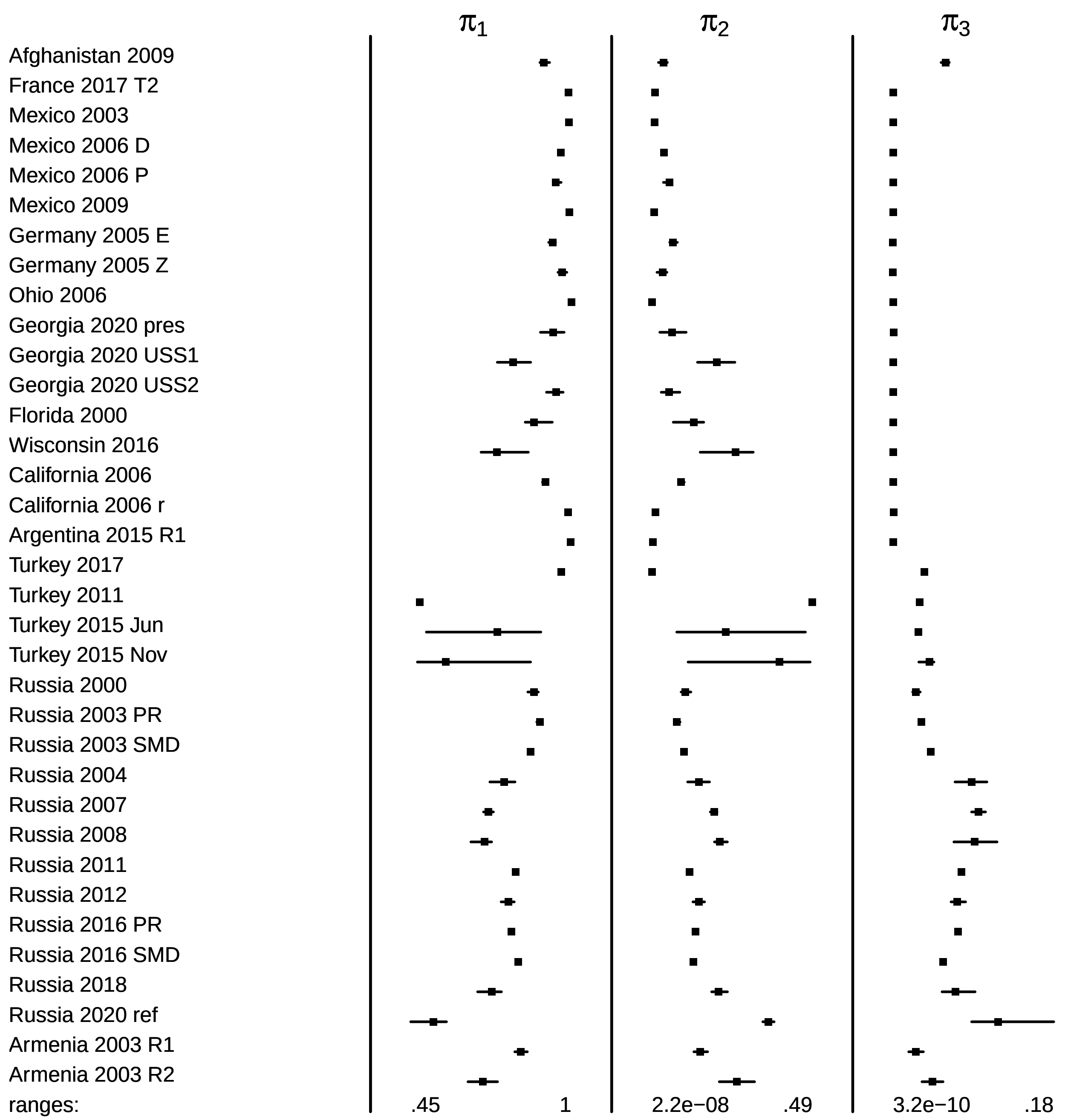
Note: proportion of polling stations `eforensics`-fraudulent (φ) for each observer status.
[a] Observer status: O1, not observed; O2, round 1 only; O3, round 2 only; O4, both rounds.
[b] p-value for likelihood ratio chi-square test of independence between observer status and frauds classification. [c] number of polling station units.

Table 8: Armenia 2003 President Election, `eforensics`-frauds Proportion Differences between Observer Statuses

	Round 1		Round 2	
Pairs[a]	manufactured[b]	stolen[c]	manufactured[b]	stolen[c]
O1−O2	$[-.000848, .0140]$	$[-.00168, .00778]$	$[-.00154, .0228]*$	$[.000211, .0112]$
O1−O3	$[-.0511, -.0357]$	$[-.0436, -.0223]$	$[-.0249, -.00622]$	$[-.0157, -.00171]$
O1−O4	$[-.0250, -.00852]$	$[-.0269, -.00742]$	$[-.0147, .00342]$	$[-.0124, .000236]$
O2−O3	$[-.0600, -.0418]*$	$[-.0477, -.0250]*$	$[-.0386, -.0164]$	$[-.0256, -.00386]$
O2−O4	$[-.0291, -.0137]$	$[-.0315, -.00994]$	$[-.0312, .00186]$	$[-.0177, -.00189]$
O3−O4	$[.0180, .0355]*$	$[-.00251, .0306]$	$[-.00218, .0228]*$	$[-.00399, .0115]*$

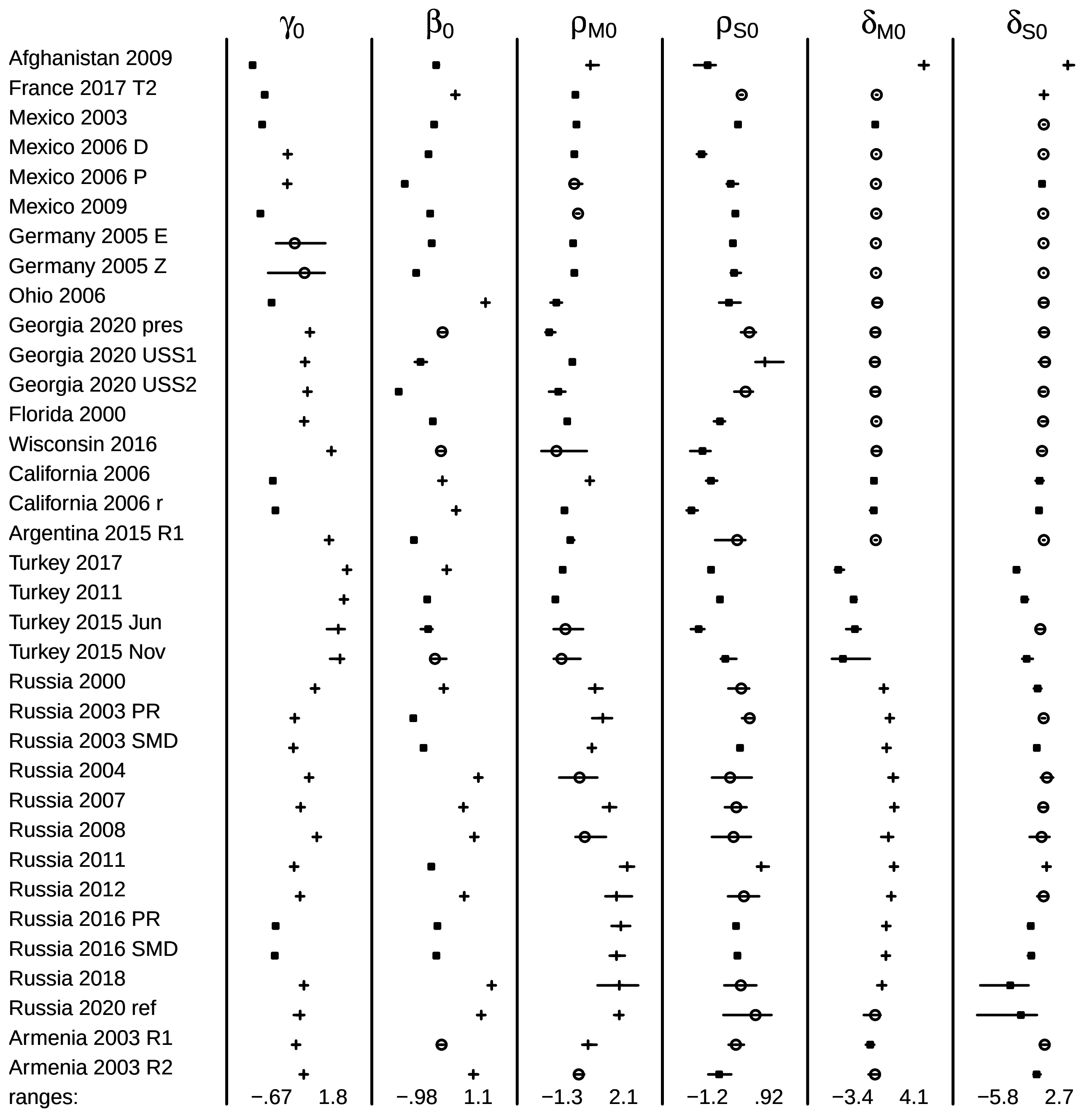
Note: 99.5% credible intervals for differences between observer statuses in ψ_{tW} and $\psi_{wW} - \psi_{tW}$. * Has sign expected if observation reduces `eforensics`-frauds magnitude.
[a] Observer status: O1, not observed; O2, round 1 only; O3, round 2 only; O4, both rounds.
[b] ψ_{tW}. [c] $\psi_{wW} - \psi_{tW}$.

Figure 1: `eforensics` Mixture Probabilities



Note: posterior means and 95% HPD intervals.

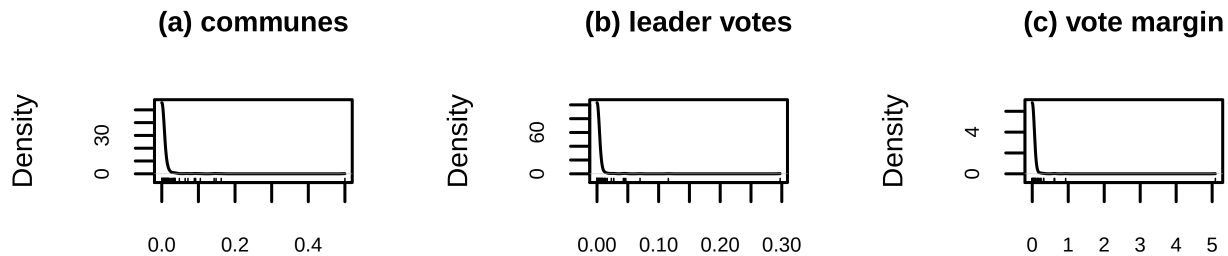
Figure 2: `eforensics` Model Coefficient Parameters



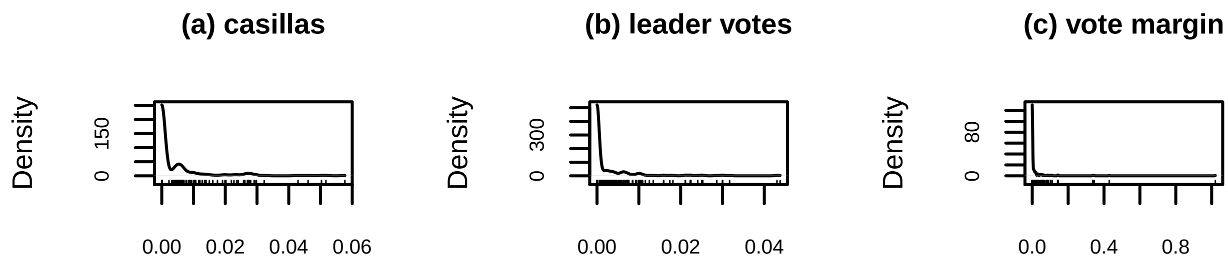
Note: posterior means and 95% HPD intervals. A square (circle, +) is used for the posterior mean if the HPD interval is less than (includes, is greater than) the prior mean.

Figure 3: Diagnostics for Legislative Districts: France, Mexico, Germany

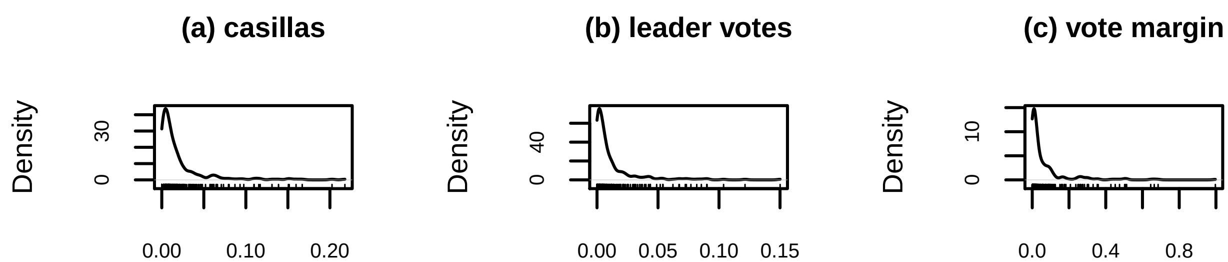
(1) France 2017



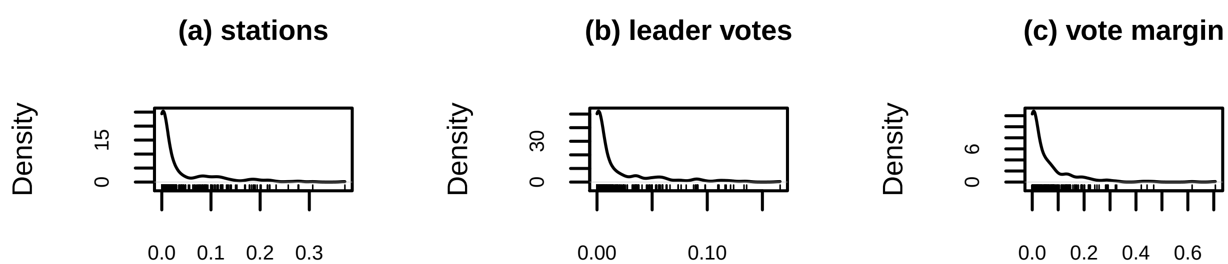
(2) Mexico 2003



(3) Mexico 2006



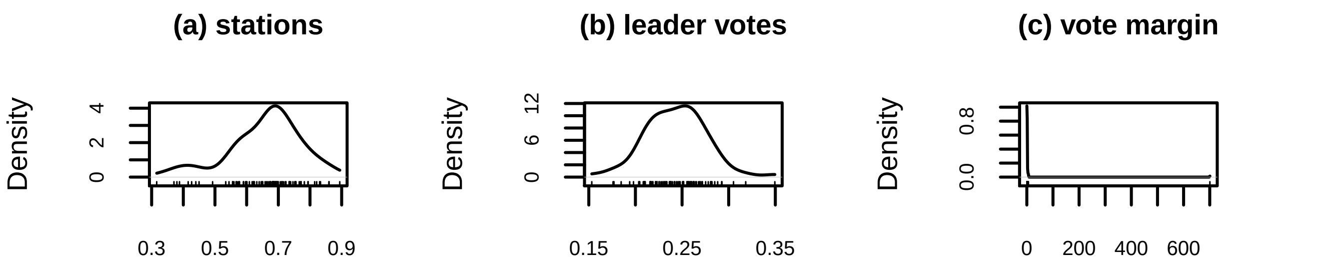
(4) Germany 2005



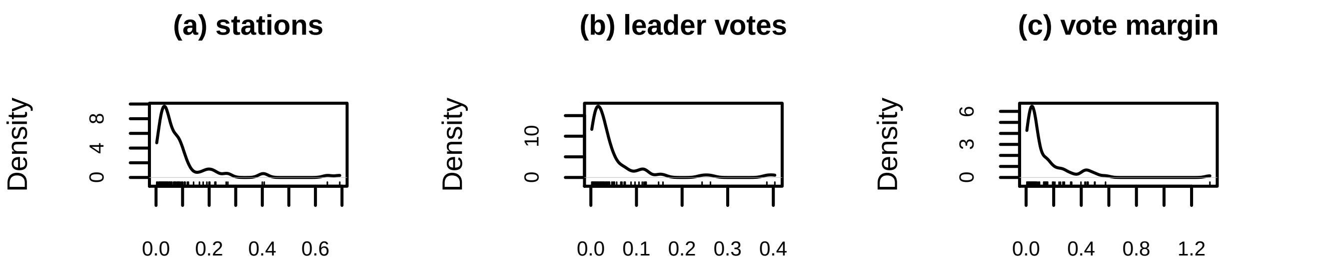
Note: empirical densities with rug plots for (a) proportion of units in each district that are `eforensics`-fraudulent (φ by district), (b) proportion of leader votes `eforensics`-fraudulent (ψ_{wW} by district), and (c) district-specific ratio of `eforensics`-frauds to the vote difference between first and second place (ψ_{wM} by district).

Figure 4: Diagnostics for Legislative Districts: Turkey

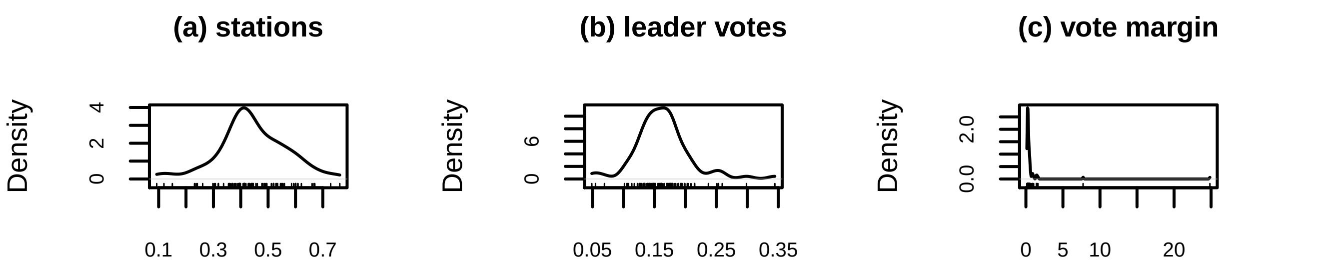
(1) Turkey 2011



(2) Turkey 2015 June



(3) Turkey 2015 November



Note: empirical densities with rug plots for (a) φ, (b) ψ_{wW}, and (c) ψ_{wM}.
The numbers of districts with $\psi_{wM} > 1$ are (1) three, (2) one, (3) seven.

Measuring Election Frauds

1 Supplemental Information

1.1 Data Notes

1.1.1 Afghanistan 2009

The data (Democracy International 2009), provided on 28 Sep 2009 by Bill Gallery, include a variable called "`IEC estimate for number of voters (released Aug 18)`," but the total of that variable over polling stations is 67,688,288, which greatly exceeds the population of Afghanistan. We have $\sum_{i=1}^{n} N_i = 13984800$, which is close to the roughly 12.8 million people aged over 15 in 2009 in Afghanistan according to National Democratic Institute (2010, 15).

1.1.2 France 2017

Decisions come from Conseil Constitutionnel (2021). In addition to the 307 districts one *Conseil Constitutionnel* decision references "*plusieurs*" districts.

Commune data come from Ministère de l'Intérieur (2021); data.gouv.fr (2021).

1.1.3 Mexico 2003, 2006, 2009

In 2015 we obtained information about cases and annulments from 397 *Juicios de Inconfmidad* (JIN) decisions reported by the courts of the TEPJF. We found HTML source documents using URLs starting with `http://portal.te.gob.mx/colecciones/sentencias/html/SUP/2006/JIN/` and using the search tool at `http://te.gob.mx/turnos-sentencias/sistema-consulta`. Information also comes from Instituto Federal Electoral (2006, 2009). Our technique for assigning *casillas* to office types relies on the lines in the HTML files that cite the "*AUTORIDAD*" for each JIN. When "*consejo local*" is mentioned we judge the file to refer to senate *casillas* and when a specific district is mentioned by number we judge the file to refer to deputy *casillas*.

casilla data come from `http://www.ife.org.mx` (on Jan 31, 2007) for 2006 and from `https://ciudadania.ife.org.mx/portalElenmex2010/exportarTablas.do?metodo=getFiltro&id=0` (on Feb 26, 2015) for 2009.

For 2003 the N_i variable, `lista_nominal`, is missing in *casilla* data obtained from `https://ciudadania.ife.org.mx/portalElenmex2010/exportarTablas.do?metodo=getFiltro&id=0` (on Feb 26, 2015). 2003 *seccion* `Lista` data come from a file `pdln_edms_2003.txt` (timestamp: Sep 13 2015).

1.1.4 Germany 2005

Polling station data (Bundeswahlleiter 2010) were purchased in September 2011. *Briefwahl* (mail ballots), which comprise 19% of the votes, come from geographic areas that usually encompass several in-person polling stations (see file `BTW05_Hinweise_Wbz.pdf`). For N_i for each mail district we use the sum of the eligible voters in all the in-person polling stations it encompasses. The 31 special polling stations altogether contain 3300 votes (overall 47,194,062 *Erststimmen* and 47,287,988 *Zweitstimmen* were cast). x^τ and x^ν include indicator variables for *Briefwahl* and special polling stations.

1.1.5 Ohio 2006

Precinct data are from `http://www.sos.state.oh.us/SOS/elections/electResultsMain/2006ElectionsResults/06-1107Precint-By-Precinct.aspx`, downloaded June 9, 2008.

1.1.6 Georgia 2020

Precinct data for the presidential election and two U.S. senate races used for the analysis come from the Georgia Secretary of State's website (Georgia Secretary of State 2020) after the second recount that was conducted by machine retabulation. Last files downloaded Nov 29, 2020.

1.1.7 Florida 2000

Precinct data are from the Florida Redistricting System (FREDS 2000), obtained on April 11, 2003. File `precinct00_pl.dbf` (timestamp 8-Sep-2001).

1.1.8 Wisconsin 2016

Ward vote data are prerecount data from `http://elections.wi.gov/sites/default/files/Ward by Ward Original and Recount President of the United States.xlsx` (on December 20, 2016), and registration data are from `http://elections.wi.gov/sites/default/files/publication/registeredvotersbywards_xlsx_19539.xlsx` (on February 4, 2017).

1.1.9 California 2006

We sum in-person and mail votes for each precinct, producing a set of $n = 22820$ combined precincts that we analyze. Precinct vote data come from file `state_G06_sov_by_g06_svprec.csv` and registration data from file `state_g06_registration_by_g06_ssprec.csv` from `http://www.acgov.org/rov/sov.htm` (on May 26, 2008).

1.1.10 Argentina 2015

Mesa data come from files `FMESPR_0101.csv`, `FMESPR_0202.csv`, `FMESPR_0313.csv`, `FMESPR_1424.csv` (timestamp Oct 26 2015) downloaded from datos.gob.ar (2017) on December 4, 2020.

Zhang, Alvarez and Levin (2019, 2) "remove *mesas* with less than 100 total ballots cast," but we include these. We exclude only 1451 *mesas* that have zero votes or missing data.

1.1.11 Turkey 2011–2017

2017 polling station data were downloaded from `https://sonuc.ysk.gov.tr` on April 28–30, 2017. 2015 data were provided on November 19, 2015, by *Wall Street Journal* staff Rob Barry and Tom McGinty. 2011 data were provided in 2016 by a scholar from Turkey.

Klimek, Jiménez, Hidalgo, Hinteregger and Thurner (2018, 3) exclude "election results from polling stations in prisons, customs authorities, or other countries," and they "removed all polling stations with an electorate of less than 100." We include such polling stations. We exclude only 2543 polling stations that have zero votes or missing data.

In 2011 (2015 June, 2015 November), 2396 (1706, 1878) polling stations have `current.vote` > `registered.voter` (*Geçerli oy* > *Kayıtlı seçmen*): N_i is set equal to `current.vote` for those polling stations and x^τ and x^ν include a dummy variable for them.

1.1.12 Russia 2000–2020

Data for 2000, 2004, 2008, 2012 and 2018 are for presidential elections, 2020 is for a constitutional referendum, and the others are for Duma elections. Data for 2003, 2004, 2007, 2008, 2011, 2012 were downloaded from URLs identified from `http://www.vybory.izbirkom.ru`. Data for 2000, 2016 SMD, 2018, 2020 were provided by Kirill Kalinin. Data for 2016 PR were provided by Josiah Augustine on November 9, 2019.

1.1.13 Armenia 2003

Data were provided by Susan Hyde on May 17, 2007.

1.2 Elaborations

1.2.1 Afghanistan 2009

Inspection of polling stations included review of physical conditions and contents of ballot boxes as well as of results and forms.[1]

[1] "For each ballot box within a sample, the IEC recorded data related to physical indicators of fraud, including: A visual inspection of the ballot box for signs of tampering; A check of the contents of the ballot box to determine whether required materials are present and whether the contents indicate signs of irregularities; An inspection of the ballots to see whether they show clear signs of fraud; and, A review of the

Audit decisions were either Fraudulent, Valid, "Valid; but form fraud" or N/A.[2]

Although polling stations are in the sample only if vote or turnout are extreme,[3]

the ECC procedure by which the Independent Election Commission (IEC) "certified the final results with Karzai receiving 49.67 percent of the vote" (Electoral Complaints Commission 2010, 37).[4]

the second frame includes polling stations all of which had many votes deemed invalid before the audit for procedural reasons.[5]

1.2.2 France 2017

For description of the election see Kuhn (2018).

1.2.3 Mexico 2003, 2006, 2009

For description of the elections see Klesner (2007, 2010). For more information about Article 75 of the *Ley General del Sistema de Medios de Impugnación en Materia Eletoral* see Corona Nakamura, De la Torre De la Torre, Gómez Torres, Grover Vaca, López Pulido, Marmolejo Gabilondo, Miranda Camarena, Orozco Montes, Rangel Jiménez, Santana Bracamontes and Torres Albarrán (2012, 215–242).

At least one of ρ_{M0}, ρ_{S0}, δ_{M0} and δ_{S0} is negative in each election (Figure 2): $\rho_{M0} < 0$, $\rho_{S0} < 0$ and $\delta_{M0} < 0$ in 2003; $\rho_{M0} < 0$ and $\rho_{S0} < 0$ in "2006 D"; $\rho_{S0} < 0$ in 2009; $\rho_{S0} < 0$

---

result and reconciliation forms to determine whether votes were recorded correctly" (Electoral Complaints Commission 2010, 36).

[2] "In accordance with ECCs Rules of Procedure, the ECC first determined whether one indicator or several indicators together provided clear and convincing evidence of fraud for a particular polling station, depending on the totality of the information available relevant to that polling station" (Electoral Complaints Commission 2010, 36). N/A refers to a station that "does not meet Order criteria" (Electoral Complaints Commission 2009).

[3] Each polling station supposedly had at most 600 electors. A polling station is in the population from which the audited sample was drawn if it features "95% or greater votes for one candidate, 600 or greater votes in a polling station, and the combination of both" (Electoral Complaints Commission 2010, 35).

[4] The ECC decided that "A finding that a ballot box was fraudulent meant that the integrity of the voting process for that box was compromised, and that all the votes contained in the box are fraudulent, in accordance with accepted electoral practice." Then they used the votes excluded by that rule to determine percentages of votes to exclude from each candidate's total (Electoral Complaints Commission 2010, 36–37).

[5] "The IEC had made its initial list of polling station based on the number of 'valid votes' equalling 600, whereas the ECC Order required the calculation to be based on the number of total votes. [...] The differences in definition amounted to hundreds of additional polling stations that needed be included in the audit and recount investigation" (Electoral Complaints Commission 2010, 36 n 37).

and $\delta_{S0} < 0$ in "2006 P." H, ψ_{tW} and ψ_{wW} are largest in "2006 P" (Table 3): in that election $\varphi = .0300$, $\psi_{tW} = .0168$ and $\psi_{wW} = .0296$; $F_w = 444053.1$ $[401318.4, 466589.8]$ is bigger than the winning margin ($F_t = 251257.8$ $[226807.7, 264804.1]$).

1.2.4 Germany 2005

For description of the election see Helms (2007).

Although not frauds, two complications affected the 2005 *Bundestagswahl*: a problem with *Briefwahl* in Dortmund; and a late election due to a candidate's death in Dresden (Mebane and Klaver 2015). In Dortmund about 10,000 ballots were mistakenly sent to absentee voters in the incorrect district. The *Bundestag* committee *Ausschuss für Wahlprüfung, Immunität und Geschäftsordnung* (Committee for Election Verification, Immunity and Rules of Procedure) acknowledged the administrative failure but did not annul the results from Dortmund because the complainants could not prove that the 10,000 ballots would have changed the outcome of either district race in Dortmund (Bundestag 2006). The other complication was that a candidate's death in Dresden's first district caused the election to be postponed (Behnke 2008), which meant that voters there could know the outcomes already produced by voting throughout the rest of the country. It was possible for a *Zweitstimme* (PR vote) cast a for their favored party to cost that party a seat (Behnke 2010). Voters could calculate the precise number of *Zweitstimmen* that would cause the CDU to lose a seat (Mebane and Klaver 2015). Voters who favored the CDU were encouraged to cast a "coalition vote," i.e., to cast their *Erststimmen* (SMD votes) for the CDU and their *Zweitstimmen* for the FDP (Behnke 2008).

The Dortmund *Briefwahl* mixup involved two districts (143 and 144, both *Dortmund, Stadt*), in which the margins SPD had over second-place CDU are respectively 42259 and 43842. `eforensics`-frauds are (posterior means) $F_t = 126.8$ and $F_w = 309.8$ in 143 and $F_t = 319.0$ and $F_w = 749.9$ in 144, much less than M. Moreover none of the `eforensics`-fraudulent polling stations are *Briefwahl*. Figure 2 shows a wide HPD interval for γ_0 that traces to apparent multimodality. Which particular fixed effects are responsible for the multimodality is unclear, but across the four MCMC chains the fixed effects for the two Dortmund districts exhibit multimodality.[6] Among the elections examined in this paper such multimodality in γ coefficients occurs uniquely for the German elections. Patterns we discuss in section 3.7 motivate speculating that the multimodality may stem from the votes lost due to the *Briefwahl* mixup.

No `eforensics`-frauds occur for Dresden first district units. Units elsewhere in the Dresden administrative district (*Regierungsbezirk*) have `eforensics`-frauds for both *Erststimmen* and *Zweitstimmen*.[7]

1.2.5 Wisconsin 2016

ρ_{M0} is multimodal for Wisconsin: In each of four MCMC chains, Wisconsin ρ_{M0} has HPD intervals $-.935$ $[-.973, -.897]$, $.277$ $[.212, .333]$, -1.23 $[-1.27, -1.18]$, -1.09 $[-1.14, -1.05]$.

1.2.6 California 2006

For description of the election see California Secretary of State (2020).

two precinct partisan registration variables in x^τ and x^ν: one variable transforms the proportion registered Republican; the other transforms the proportion Democrat. r_{ji} is the number registered for party j in precinct i and $p_{ji} = \dfrac{r_{ji} + .5}{N_i + .5}$. The registration variable is $R_{ji} = \log\left(\dfrac{p_{ji}}{1 - p_{ji}}\right)$. If $p_{ji} \geq .999$, $R_{ji} = \log(.999/(1 - .999))$.

1.2.7 Turkey 2011–2017

Removing F_t and F_w leaves "Yes" proportion

$$\frac{\sum W_i - F_w}{\sum V_i - F_t} = \frac{24965387 - 325548.2}{48627622 - 96543.4} = .508\,.$$

If fixed effects are omitted, $\pi_2 = .0469$ $[.0295, .0617]$, $\pi_3 = .0199$ $[.0127, .0296]$, only

[6] β_{143}: .314 $[.309, .320]$, .608 $[.605, .612]$, .259 $[.255, .263]$, 1.21 $[1.20, 1.21]$. β_{144}: .327 $[.322, .331]$, .0813 $[.0741, .0891]$, 1.93 $[1.92, 1.93]$, 1.00 $[.998, 1.00]$.

[7] Particularly five polling stations in district 156 in the communities of Nebelschütz, Panschwitz-Kuckau and Ralbitz-Rosenthal have between them $F_t = 170.0$ and $F_w = 500.3$ `eforensics`-fraudulent *Zweitstimmen* and $F_t = 77.13$ and $F_w = 451.0$ `eforensics`-fraudulent *Erststimmen*.

$\delta_{M0} < 0$ differs from zero, $H = 3892$, $F_t = 58566.9$ $[54431.7, 62768.8]$ and $F_w = 241992.3$ $[225390.8, 258829.2]$.

1.2.8 Russia 2000–2020

Figure 1 shows that already in 2000 π_2 and π_3 are high ($\pi_2 = .101$ $[.0886, .118]$, $\pi_3 = .0254$ $[.0217, .0300]$), increase to a peak in 2008 ($\pi_2 = .205$ $[.189, .228]$, $\pi_3 = .0901$ $[.0672, .115]$) decline back to 2004 levels in 2011 ($\pi_2 = .113$ $[.110, .116]$, $\pi_3 = .0757$ $[.0734, .0781]$) then reach a new high in 2020 ($\pi_2 = .352$ $[.335, .369]$, $\pi_3 = .116$ $[.0867, .177]$.

By election, φ is: .097, 2003 PR; .12, 2003 SMD; .26, 2007; .17, 2011; .19, 2012.

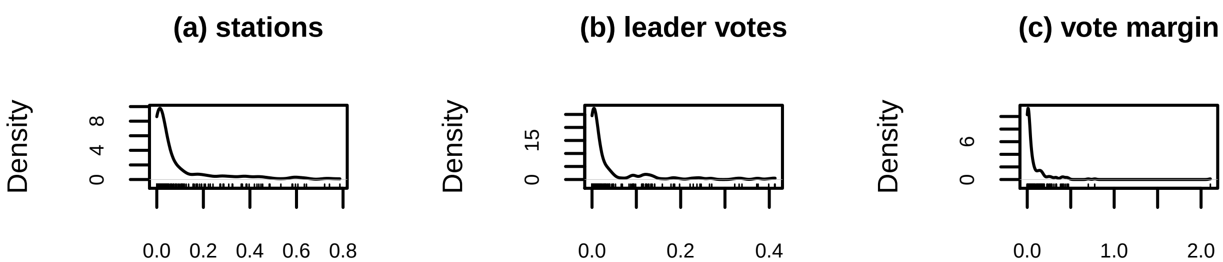
By district (Figure 5) 2003 and 2016 SMD exhibit greater `eforensics`-frauds than other legislative elections in this paper. φ and ψ_{wW} range higher than for French or Mexican elections (cf. Figure 3), and ψ_{wW} ranges higher than for Turkish elections (cf. Figure 4).

1.2.9 Armenia 2003

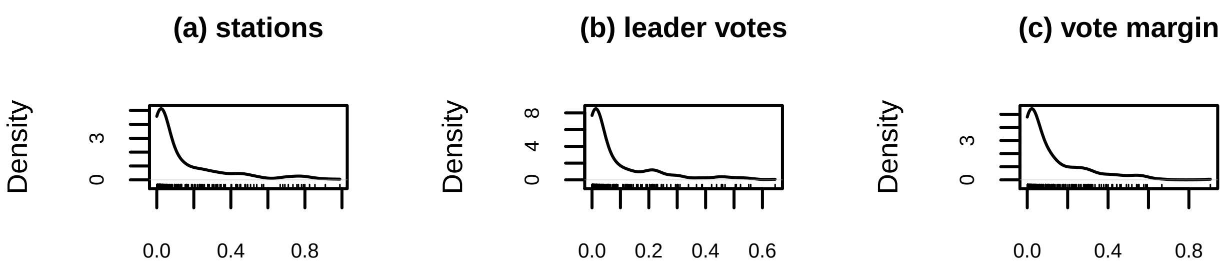
Direct comparison with Hyde (2007)'s vote share observer effects requires taking into account in addition the coefficients for O2, O3 and O4 in x^τ and x^ν that differ from zero, i.e., the estimated observer effects on `eforensics` turnout and vote choice: $\gamma_{O3} > 0$, $\beta_{O2} < 0$ and $\beta_{O4} < 0$ for round 1; $\gamma_{O4} > 0$ and $\beta_{O4} < 0$ for round 2. There are observer effects on `eforensics` estimates of turnout and vote choice.

Figure 5: Diagnostics for Legislative Districts: Russia

(1) Russia 2003



(2) Russia 2016



Note: empirical densities with rug plots for (a) proportion of units in each district that are `eforensics`-fraudulent (φ by district), (b) proportion of leader votes `eforensics`-fraudulent (ψ_{wW} by district), and (c) district-specific ratio of `eforensics`-frauds to the vote difference between first and second place (ψ_{wM} by district).

References

Behnke, Joachim. 2008. "Strategisches Wählen bei der Nachwahl in Dresden zur Bundestagswahl 2005." *Politische Vierteljahresschrift* 49(4):695–720.

Behnke, Joachim. 2010. "Negatives Stimmgewicht, Erfolgswert und Überhangmandate– einige Anmerkungen." *Kritische Vierteljahresschrift für Gesetzgebung und Rechtswissenschaft* (1):3–28.

Bundestag, Deutscher. 2006. *Erste Beschlussempfehlung des Wahlprüfungsausschusses zu 51 gegen die Gültigkeit der Wahl zum 16. Deutschen Bundestag eingegangenen Wahleinsprüchen.* Number 16/900. March 3, 2006.

Bundeswahlleiter, Der. 2010. *Wahl zum 16. Deutschen Bundestag am 18. September 2005.* Wiesbaden: im Auftrag der Herausgebergemeinschaft. Ergebnisse der Wahlbezirksstatistik: Die Wahlleiter des Bundes und der Länder, Auszugsweise Vervielfältigung und Verbreitung mit Quellenangaben gestattet.

California Secretary of State. 2020. "General Election - November 7, 2006." URL: `https://www.sos.ca.gov/elections/prior-elections/statewide-election-results/general-election-november-7-2006`.

Conseil Constitutionnel. 2021. "Les décisions du Conseil Constitutionnel." `https://www.conseil-constitutionnel.fr/decisions`.

Corona Nakamura, Luis Antonio, Jaime Benjamín De la Torre De la Torre, Carmen Sofía Gómez Torres, Jaime Grover Vaca, Mara de Jesús López Pulido, Olga Mara Marmolejo Gabilondo, Adrián Joaquín Miranda Camarena, Juan Enrique Orozco Montes, Miriam Rangel Jiménez, Ernesto Santana Bracamontes and Alejandro Torres Albarrán. 2012. "Ley General del Sistema de Medios de Impugnación en Materia Electoral: Comentada." URL `http://www.triejal.gob.mx/Publicaciones/Ley_gral_sist_medios.pdf` (accessed June 5, 2015).

data.gouv.fr. 2021. "Elections législatives des 11 et 18 juin 2017 - Résultats du 2nd tour par communes." `https://www.data.gouv.fr/fr/datasets/`

elections-legislatives-des-11-et-18-juin-2017-resultats-du-2nd-tour-par-communes/.

datos.gob.ar. 2017. "Elecciones Generales 2015: Presidente y Vice." http://datos.gob.ar/dataset/otros_5/resource/otros_5.1.

Democracy International. 2009. "Results by PS - PDF data only - 16 Sept 2009.xlsx.".

Electoral Complaints Commission. 2010. "Electoral Complaints Commission Final Report: 2009 Presidential and Provincial Council Elections." ECC Final Report 2009.pdf.

Electoral Complaints Commission, Islamic Republic of Afghanistan. 2009. "Annex a: Audit Results By Polling Station." Annex for 18 October 2019 Letter to Azizullah Ludin, Chairman, Independent Election Commission of Afghanistan: ECC2009_October18ANNEXADecisionAuditandRecount_20091018.pdf.

Georgia Secretary of State. 2020. "November 3, 2020 General Election, Results by County." URL: https://results.enr.clarityelections.com/GA/105369/web.264614/#/access-to-races.

Helms, Ludger. 2007. "The German Federal Election, September 2005." *Electoral Studies* 26(1):223–227.

Hyde, Susan D. 2007. "The Observer Effect in International Politics: Evidence from a Natural Experiment." *World Politics* 60:37–63.

Instituto Federal Electoral. 2006. "Estudio de las causales de nulidad, de las elecciones federales de 2006, relacionadas con la integración de mesas directivas de casilla y el desempeño de los funcionarios de mesas directivas de casilla, derivados de los juicios de inconfmidad." *Capacitación Electoral y Educación Civica*; in file Estudio_causales_de_nulidad_2006_optm.pdf.

Instituto Federal Electoral. 2009. "Causas de Nulidad." Dirección Ejectutiva de Capacitación Electoral y Educación Civica; URL: http://www.ine.mx/docs/IFE-v2/UTSID/UTSID-InformacionSocialmenteUtil/Docs/Estudio_Nulidad_Sentencias.pdf.

Klesner, Joseph L. 2007. "The July 2006 Presidential and Congressional Elections in Mexico." *Electoral Studies* 26(4):803–808.

Klesner, Joseph L. 2010. "The 2009 Mexican midterm congressional elections." *Electoral Studies* 29(3):537–540.

Klimek, Peter, Raúl Jiménez, Manuel Hidalgo, Abraham Hinteregger and Stefan Thurner. 2018. "Forensic Analysis of Turkish Elections in 2017–2018." *PLOS One* 13(10):e0204975.

Kuhn, Raymond. 2018. "French Revolution? The 2017 Presidential and Parliamentary Elections." *Parliamentary Affairs* 71(3):483–500.

Mebane, Jr., Walter R. and Joseph Klaver. 2015. "Election Forensics: Strategies versus Election Frauds in Germany." Paper presented at the 2015 Annual Meeting of the Midwest Political Science Association, Chicago, April 16–19, 2015.

Ministère de l'Intérieur. 2021. "Résultats des élections législatives 2017." `https://www.interieur.gouv.fr/Elections/Les-resultats/Legislatives/elecresult__legislatives-2017/(path)/legislatives-2017/index.html`.

National Democratic Institute. 2010. "The 2009 Presidential and Provincial Council Elections in Afghanistan." URL: `https://www.ndi.org/sites/default/files/Elections_in_Afghanistan_2009.pdf`.

Zhang, Mali, R. Michael Alvarez and Ines Levin. 2019. "Election Forensics: Using Machine Learning and Synthetic Data for Possible Election Anomaly Detection." *PLOS ONE* 14(10):e0223950.

Table 9: France 2017 National Assembly Election `eforensics` Estimates

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.990	.989	.992
	π_2	Incremental Fraud	.00935	.00785	.0109
	π_3	Extreme Fraud	.000153	.0000260	.000295
turnout	β_0	(Intercept)	$-.319$	$-.332$	$-.309$
vote choice	γ_0	(Intercept)	.312	.300	.325
incremental frauds	ρ_{M0}	(Intercept)	$-.0902$	$-.143$	$-.0239$
	ρ_{S0}	(Intercept)	$-.0137$	$-.0501$	.0104
extreme frauds	δ_{M0}	(Intercept)	.0555	$-.0102$	.117
	δ_{S0}	(Intercept)	.0278	.00689	.0642

units `eforensics`-fraudulent: 200 fraudulent, 35422 not fraudulent
manufactured votes $F_t = 7721.7\ [7022.5, 8747.9]^c$
total `eforensics`-fraudulent votes $F_w = 12168.3\ [10302.1, 14468.8]^c$

Note: selected `eforensics` model parameter estimates (posterior means and highest posterior density credible intervals). $n = 35622$ *commune* units. [a] 95% HPD lower bound. [b] 95% HPD upper bound. [c] posterior mean [99.5% credible interval].

Table 10: Mexico 2003, 2006, 2009 Legislative Elections `eforensics` Estimates

(a) 2003:

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.992	.991	.992
	π_2	Incremental Fraud	.00825	.00744	.00909
	π_3	Extreme Fraud	.000239	.000114	.000359
turnout	β_0	(Intercept)	$-.383$	$-.399$	$-.365$
vote choice	γ_0	(Intercept)	$-.151$	$-.162$	$-.141$
incremental frauds	ρ_{M0}	(Intercept)	$-.0509$	$-.0702$	$-.0199$
	ρ_{S0}	(Intercept)	$-.0968$	$-.122$	$-.0580$
extreme frauds	δ_{M0}	(Intercept)	$-.0652$	$-.139$	$-.00751$
	δ_{S0}	(Intercept)	.0189	$-.0508$	.119

(b) 2006:

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.964	.962	.965
	π_2	Incremental Fraud	.0363	.0350	.0375
	π_3	Extreme Fraud	.000153	8.01e-05	.000237
turnout	β_0	(Intercept)	.258	.233	.275
vote choice	γ_0	(Intercept)	$-.273$	$-.283$	$-.257$
incremental frauds	ρ_{M0}	(Intercept)	$-.125$	$-.150$	$-.106$
	ρ_{S0}	(Intercept)	$-.912$	-1.02	$-.806$
extreme frauds	δ_{M0}	(Intercept)	.0210	$-.0128$	.0415
	δ_{S0}	(Intercept)	$-.00761$	$-.0315$	.0428

(c) 2009:

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.993	.992	.994
	π_2	Incremental Fraud	.00642	.00595	.00692
	π_3	Extreme Fraud	.000115	5.92e-05	.000182
turnout	β_0	(Intercept)	$-.426$	$-.443$	$-.413$
vote choice	γ_0	(Intercept)	$-.237$	$-.258$	$-.213$
incremental frauds	ρ_{M0}	(Intercept)	.00251	$-.0623$	.0514
	ρ_{S0}	(Intercept)	$-.155$	$-.219$	$-.102$
extreme frauds	δ_{M0}	(Intercept)	.0150	$-.0290$	.0883
	δ_{S0}	(Intercept)	$-.0267$	$-.0594$	.0104

Note: selected `eforensics` model parameter estimates (posterior means and highest posterior density credible intervals). Data come from the second (machine) recount. (a) $n = 63146$ *secciones*, (b) $n = 130448$ *casillas*, (c) $n = 138345$ *casillas*. District fixed-effects for turnout and vote choice are not shown. [a] 95% HPD lower bound. [b] 95% HPD upper bound.

Table 11: Mexico 2003, 2006, 2009 Legislative Elections `eforensics` Fraudulent Vote Estimates

(a) 2003

units `eforensics`-fraudulent: 318 fraudulent, 62828 not fraudulent

manufactured votes	$F_t = 19457.89$ $[17818.6, 21066.0]^c$
total `eforensics`-fraudulent votes	$F_w = 25776.26$ $[23874.7, 27781.2]^c$

(b) 2006

units `eforensics`-fraudulent: 2652 fraudulent, 127796 not fraudulent

manufactured votes	$F_t = 127783.1$ $[117342.2, 136804.5]^c$
total `eforensics`-fraudulent votes	$F_w = 180769.7$ $[166972.3, 193649.7]^c$

(c) 2009

units `eforensics`-fraudulent: 426 fraudulent, 137919 not fraudulent

manufactured votes	$F_t = 28096.18$ $[24939.0, 31039.0]^c$
total `eforensics`-fraudulent votes	$F_w = 37191.92$ $[33017.9, 41073.9]^c$

Note: `eforensics` model `eforensics`-fraudulent vote count estimates. (a) $n = 63146$ *secciones*, (b) $n = 130448$ *casillas*, (c) $n = 138345$ *casillas*. ^{a} posterior mean [99.5% credible interval].

Table 12: Mexico 2006 President Election `eforensics` Estimates

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.947	.939	.965
	π_2	Incremental Fraud	.0528	.0350	.0610
	π_3	Extreme Fraud	.000120	5.85e-05	.000188
turnout	β_0	(Intercept)	.248	.226	.263
vote choice	γ_0	(Intercept)	$-.788$	$-.804$	$-.768$
incremental frauds	ρ_{M0}	(Intercept)	$-.127$	$-.287$	.149
	ρ_{S0}	(Intercept)	$-.260$	$-.348$	$-.0951$
extreme frauds	δ_{M0}	(Intercept)	$-.00727$	$-.0380$	.0232
	δ_{S0}	(Intercept)	$-.138$	$-.220$	$-.0667$

units `eforensics`-fraudulent: 3927 fraudulent, 126841 not fraudulent
manufactured votes $F_t = 251257.8\ [226807.7, 264804.1]^c$
total `eforensics`-fraudulent votes $F_w = 444053.1\ [401318.4, 466589.8]^c$

Note: selected `eforensics` model parameter estimates (posterior means and highest posterior density credible intervals). $n = 130768$ *casilla* units. [a] 95% HPD lower bound. [b] 95% HPD upper bound. [c] posterior mean [99.5% credible interval].

Table 13: Germany 2005 Election `eforensics` Estimates

(a) *Erststimmen* (SMD)

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.937	.923	.946
	π_2	Incremental Fraud	.0632	.0541	.0772
	π_3	Extreme Fraud	7.11e-05	9.31e-06	.000142
turnout	β_0	(Intercept)	.435	$-.0343$	1.20
	β_1	*Briefwahl*	-1.19	-1.98	$-.683$
	β_2	special	.553	-1.47	2.08
vote choice	γ_0	(Intercept)	$-.203$	$-.227$	$-.162$
	γ_1	*Briefwahl*	.0724	.0566	.0867
	γ_2	special	$-.0281$	$-.106$	.0278
incremental frauds	ρ_{M0}	(Intercept)	$-.168$	$-.202$	$-.139$
	ρ_{S0}	(Intercept)	$-.210$	$-.251$	$-.187$
extreme frauds	δ_{M0}	(Intercept)	$-.00981$	$-.0250$	.00226
	δ_{S0}	(Intercept)	$-.00987$	$-.0309$	.0184

(b) *Zweitstimmen* (PR)

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.968	.954	.984
	π_2	Incremental Fraud	.0322	.0156	.0456
	π_3	Extreme Fraud	1.48e-05	3.24e-10	4.31e-05
turnout	β_0	(Intercept)	.681	$-.234$	1.19
	β_1	*Briefwahl*	-1.56	-2.13	$-.981$
	β_2	special	$-.226$	$-.942$	.752
vote choice	γ_0	(Intercept)	$-.541$	$-.561$	$-.521$
	γ_1	*Briefwahl*	.0262	.00574	.0466
	γ_2	special	.00756	$-.0562$	.0769
incremental frauds	ρ_{M0}	(Intercept)	$-.126$	$-.183$	$-.0370$
	ρ_{S0}	(Intercept)	$-.183$	$-.258$	$-.0320$
extreme frauds	δ_{M0}	(Intercept)	$-.00116$	$-.0124$	.00990
	δ_{S0}	(Intercept)	$-.00335$	$-.122$	.00443

Note: selected `eforensics` model parameter estimates (posterior means and highest posterior density credible intervals). (a, b) $n = 88680$ polling station and *Briefwahl* units. (a) District and (b) *Länder* fixed-effects for turnout and vote choice are not shown. [a] 95% HPD lower bound. [b] 95% HPD upper bound.

Table 14: Germany 2005 Election `eforensics` Fraudulent Vote Estimates

(a) *Erststimmen* (SMD)
units `eforensics`-fraudulent: 3751 fraudulent, 84929 not fraudulent
manufactured votes $F_t = 241921.0$ $[225001.8, 254916.2]^a$
total `eforensics`-fraudulent votes $F_w = 349025.5$ $[325935.4, 366712.7]^a$

(b) *Zweitstimmen* (PR)
units `eforensics`-fraudulent: 1289 fraudulent, 87391 not fraudulent
manufactured votes $F_t = 82166.31$ $[76071.50, 88391.40]^a$
total `eforensics`-fraudulent votes $F_w = 127698.39$ $[119012.41, 136126.35]^a$

Note: `eforensics` model `eforensics`-fraudulent vote count estimates. (a, b) $n = 88680$ polling station and *Briefwahl* units. [a] posterior mean [99.5% credible interval].

Table 15: Ohio 2006 State House Election `eforensics` Estimates

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.999	.999	1.00
	π_2	Incremental Fraud	.000424	9.32e-08	.00139
	π_3	Extreme Fraud	.0000926	1.22e-08	.000277
turnout	β_0	(Intercept)	−.146	−.178	−.120
vote choice	γ_0	(Intercept)	.967	.926	1.01
incremental frauds	ρ_{M0}	(Intercept)	−.748	−.949	−.552
	ρ_{S0}	(Intercept)	−.299	−.518	−.0398
extreme frauds	δ_{M0}	(Intercept)	.0951	−.272	.423
	δ_{S0}	(Intercept)	.00868	−.227	.247

Note: selected `eforensics` model parameter estimates (posterior means and highest posterior density credible intervals). $n = 11123$ precinct units. District fixed-effects for turnout and vote choice are not shown. [a] 95% HPD lower bound. [b] 95% HPD upper bound.

Table 16: Georgia 2020 Elections `eforensics` Estimates

(a) President, Biden-Harris:

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.938	.895	.975
	π_2	Incremental Fraud	.0614	.0240	.104
	π_3	Extreme Fraud	.000630	7.27e-07	.00186
turnout	β_0	(Intercept)	.816	.787	.845
vote choice	γ_0	(Intercept)	.0338	$-.0510$	.114
incremental frauds	ρ_{M0}	(Intercept)	$-.992$	-1.12	$-.779$
	ρ_{S0}	(Intercept)	.157	$-.0310$	.305
extreme frauds	δ_{M0}	(Intercept)	$-.0506$	$-.350$	.178
	δ_{S0}	(Intercept)	.0389	$-.417$	.234

(b) Senate, Perdue:

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.803	.749	.862
	π_2	Incremental Fraud	.197	.138	.251
	π_3	Extreme Fraud	.000458	1.93e-08	.00139
turnout	β_0	(Intercept)	.694	.655	.734
vote choice	γ_0	(Intercept)	$-.446$	$-.571$	$-.308$
incremental frauds	ρ_{M0}	(Intercept)	$-.195$	$-.248$	$-.141$
	ρ_{S0}	(Intercept)	.504	.293	.919
extreme frauds	δ_{M0}	(Intercept)	$-.0889$	$-.473$	.152
	δ_{S0}	(Intercept)	.127	$-.416$	.416

(c) Senate, Warnock:

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.947	.915	.972
	π_2	Incremental Fraud	.0521	.0285	.0846
	π_3	Extreme Fraud	.000471	1.43e-07	.00141
turnout	β_0	(Intercept)	.755	.727	.785
vote choice	γ_0	(Intercept)	$-.921$	$-.978$	$-.864$
incremental frauds	ρ_{M0}	(Intercept)	$-.679$	-1.00	$-.433$
	ρ_{S0}	(Intercept)	.0726	$-.179$	.237
extreme frauds	δ_{M0}	(Intercept)	$-.0516$	$-.375$	.201
	δ_{S0}	(Intercept)	$-.00153$	$-.220$	.130

Note: selected `eforensics` model parameter estimates (posterior means and highest posterior density credible intervals). Data come from the second (machine) recount. (a, b, c) $n = 2577$ precincts. [a] 95% HPD lower bound. [b] 95% HPD upper bound.

Table 17: Georgia 2020 Elections `eforensics` Fraudulent Vote Estimates

(a) President: Biden-Harris	
units `eforensics`-fraudulent: 20 fraudulent, 2557 not fraudulent	
manufactured votes	$F_t = 2750.636$ $[2221.062, 3241.822]^c$
total `eforensics`-fraudulent votes	$F_w = 9129.020$ $[7554.512, 10496.689]^c$
(b) Senate: Perdue	
units `eforensics`-fraudulent: 300 fraudulent, 2277 not fraudulent	
manufactured votes	$F_t = 38750.64$ $[37126.19, 40208.43]^c$
total `eforensics`-fraudulent votes	$F_w = 111693.40$ $[105273.67, 117256.08]^c$
(c) Senate: Warnock	
units `eforensics`-fraudulent: 39 fraudulent, 2538 not fraudulent	
manufactured votes	$F_t = 6839.451$ $[6372.166, 7220.382]^c$
total `eforensics`-fraudulent votes	$F_w = 20660.655$ $[18122.537, 22455.856]^c$

Note: `eforensics` model `eforensics`-fraudulent vote count estimates. (a, b, c) $n = 2577$ precinct units. [a] posterior mean [99.5% credible interval].

Table 18: Florida 2000, Wisconsin 2016 President Elections `eforensics` Estimates

(a) Florida 2000: Bush-Cheney

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.873	.843	.935
	π_2	Incremental Fraud	.127	.0648	.157
	π_3	Extreme Fraud	.000350	3.60e-07	.000877
turnout	β_0	(Intercept)	.672	.648	.705
vote choice	γ_0	(Intercept)	$-.179$	$-.215$	$-.133$
incremental frauds	ρ_{M0}	(Intercept)	$-.371$	$-.406$	$-.336$
	ρ_{S0}	(Intercept)	$-.499$	$-.636$	$-.387$
extreme frauds	δ_{M0}	(Intercept)	.0236	$-.0410$	.076
	δ_{S0}	(Intercept)	$-.0433$	$-.193$	.183

(b) Wisconsin 2016: Trump-Pence

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.747	.693	.853
	π_2	Incremental Fraud	.253	.146	.306
	π_3	Extreme Fraud	.000321	2.61e-08	.000964
turnout	β_0	(Intercept)	1.36	1.30	1.40
vote choice	γ_0	(Intercept)	$-.00148$	$-.0688$	.0576
incremental frauds	ρ_{M0}	(Intercept)	$-.743$	-1.26	.307
	ρ_{S0}	(Intercept)	$-.889$	-1.16	$-.708$
extreme frauds	δ_{M0}	(Intercept)	.0347	$-.199$	.295
	δ_{S0}	(Intercept)	$-.137$	$-.305$	.0188

(c) Florida 2000: Bush-Cheney
units `eforensics`-fraudulent: 289 fraudulent, 5652 not fraudulent
manufactured votes $F_t = 32378.18$ $[30692.0, 33755.3]^c$
total `eforensics`-fraudulent votes $F_w = 66915.37$ $[63185.2, 69848.3]^c$

(d) Wisconsin 2016: Trump-Pence
units `eforensics`-fraudulent: 475 fraudulent, 2919 not fraudulent
manufactured votes $F_t = 10695.892$ $[8506.7, 11896.3]^c$
total `eforensics`-fraudulent votes $F_w = 41692.907$ $[35207.9, 46760.5]^c$

Note: selected `eforensics` model parameter estimates (posterior means and highest posterior density credible intervals) and `eforensics`-fraudulent vote count estimates. (a,c) $n = 5941$ precinct units. (b,d) $n = 3394$ ward units. County fixed-effects for turnout and vote choice are not shown. [a] 95% HPD lower bound. [b] 95% HPD upper bound. [c] posterior mean [99.5% credible interval].

Table 19: Florida 2000, Wisconsin 2016 President Elections `eforensics` Estimates

(a) Florida 2000: Bush-Cheney

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.940	.863	.978
	π_2	Incremental Fraud	.0596	.0223	.137
	π_3	Extreme Fraud	.000291	4.82e-08	.000746
turnout	β_0	(Intercept)	.726	.672	.761
vote choice	γ_0	(Intercept)	$-.280$	$-.385$	$-.218$
incremental frauds	ρ_{M0}	(Intercept)	$-.115$	$-.311$	.0536
	ρ_{S0}	(Intercept)	.0421	$-.192$	.230
extreme frauds	δ_{M0}	(Intercept)	$-.0872$	$-.174$	.0288
	δ_{S0}	(Intercept)	$-.0210$	$-.140$	.139

(b) Wisconsin 2016: Trump-Pence

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.823	.645	.938
	π_2	Incremental Fraud	.177	.0616	.355
	π_3	Extreme Fraud	.000304	3.86e-08	.000912
turnout	β_0	(Intercept)	1.34	1.28	1.39
vote choice	γ_0	(Intercept)	$-.319$	$-.528$	$-.179$
incremental frauds	ρ_{M0}	(Intercept)	$-.705$	$-.991$	$-.484$
	ρ_{S0}	(Intercept)	.0411	$-.258$	.357
extreme frauds	δ_{M0}	(Intercept)	$-.167$	$-.314$	$-.0436$
	δ_{S0}	(Intercept)	$-.201$	$-.469$	.0592

(c) Florida 2000: Bush-Cheney

units `eforensics`-fraudulent: 81 fraudulent, 5860 not fraudulent

manufactured votes $F_t = 9957.906\ [8789.042, 10744.182]^c$

total `eforensics`-fraudulent votes $F_w = 21964.332\ [20696.232, 22764.267]^c$

(d) Wisconsin 2016: Trump-Pence

units `eforensics`-fraudulent: 240 fraudulent, 3154 not fraudulent

manufactured votes $F_t = 7250.170\ [5767.896, 8120.814]^c$

total `eforensics`-fraudulent votes $F_w = 29854.268\ [23130.794, 33979.224]^c$

Note: selected `eforensics` model parameter estimates (posterior means and highest posterior density credible intervals) and `eforensics`-fraudulent vote count estimates. (a,c) $n = 5941$ precinct units. (b,d) $n = 3394$ ward units. [a] 95% HPD lower bound. [b] 95% HPD upper bound. [c] posterior mean [99.5% credible interval].

Table 20: California 2006 Governor Election `eforensics` Estimates

(a) Intercepts only:

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.912	.902	.918
	π_2	Incremental Fraud	.0881	.0812	.0980
	π_3	Extreme Fraud	.000440	.000151	.000758
turnout	β_0	(Intercept)	$-.114$	$-.135$	$-.0946$
vote choice	γ_0	(Intercept)	.0296	.00434	.0522
incremental frauds	ρ_{M0}	(Intercept)	.409	.358	.478
	ρ_{S0}	(Intercept)	$-.701$	$-.809$	$-.561$
extreme frauds	δ_{M0}	(Intercept)	$-.163$	$-.217$	$-.123$
	δ_{S0}	(Intercept)	$-.345$	$-.698$	$-.00207$

(b) Including partisan registration covariates:

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.988	.985	.991
	π_2	Incremental Fraud	.0113	.00866	.0145
	π_3	Extreme Fraud	.000620	.000230	.00104
turnout	β_0	(Intercept)	$-.0493$	$-.0618$	$-.0359$
	β_1	R_{REP}	.0506	.0459	.0556
	β_2	R_{DEM}	$-.116$	$-.120$	$-.110$
vote choice	γ_0	(Intercept)	.328	.301	.347
	γ_1	R_{REP}	.562	.558	.565
	γ_2	R_{DEM}	$-.575$	$-.581$	$-.569$
incremental frauds	ρ_{M0}	(Intercept)	$-.467$	$-.530$	$-.381$
	ρ_{S0}	(Intercept)	-1.14	-1.24	$-.995$
extreme frauds	δ_{M0}	(Intercept)	$-.171$	$-.468$	$-.0426$
	δ_{S0}	(Intercept)	$-.403$	$-.540$	$-.210$

Note: selected `eforensics` model parameter estimates (posterior means and highest posterior density credible intervals). (a) $n = 22820$, (b) $n = 22566$ precincts. [a] 95% HPD lower bound. [b] 95% HPD upper bound.

Table 21: California 2006 Governor Election `eforensics` Estimates

(a) Intercepts only

units `eforensics`-fraudulent: 1309 fraudulent, 21511 not fraudulent	
manufactured votes	$F_t = 99689.14$ $[94377.59, 104427.13]^a$
total `eforensics`-fraudulent votes	$F_w = 123137.79$ $[118312.76, 128627.22]^a$

(b) Including partisan registration covariates

units `eforensics`-fraudulent: 81 fraudulent, 22739 not fraudulent	
manufactured votes	$F_t = 9646.397$ $[7771.361, 11109.020]^a$
total `eforensics`-fraudulent votes	$F_w = 12366.112$ $[9601.822, 15087.711]^a$

Note: selected `eforensics` model parameter estimates (posterior means and highest posterior density credible intervals). (a) $n = 22820$, (b) $n = 22566$ precincts. [a] posterior mean [99.5% credible interval].

Table 22: Argentina 2015 President Election (Round 1) `eforensics` Estimates

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.997	.994	1.00
	π_2	Incremental Fraud	.00327	.0000654	.00636
	π_3	Extreme Fraud	.0000798	.0000101	.000168
turnout	β_0	(Intercept)	1.30	1.28	1.31
vote choice	γ_0	(Intercept)	−.591	−.605	−.577
incremental frauds	ρ_{M0}	(Intercept)	−.267	−.344	−.128
	ρ_{S0}	(Intercept)	−.114	−.608	.0678
extreme frauds	δ_{M0}	(Intercept)	−.0218	−.106	.0626
	δ_{S0}	(Intercept)	.0254	−.0546	.0849
units `eforensics`-fraudulent: 25 fraudulent, 92179 not fraudulent					
manufactured votes		$F_t = 425.2951$ $[313.7366, 512.4650]^c$			
total `eforensics`-fraudulent votes		$F_w = 1468.0793$ $[1083.8572, 1753.7006]^c$			

Note: selected `eforensics` model parameter estimates (posterior means and highest posterior density credible intervals). $n = 92204$ *mesa* units. [a] 95% HPD lower bound. [b] 95% HPD upper bound. [c] posterior mean [99.5% credible interval].

Table 23: Turkey 2017 Referendum Election `eforensics` Estimates

(a) Parameter Estimates

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.965	.965	.966
	π_2	Incremental Fraud	.0000724	2.17e-08	.000201
	π_3	Extreme Fraud	.0345	.0335	.0354
turnout	β_0	(Intercept)	1.75	1.72	1.76
vote choice	γ_0	(Intercept)	.123	.0986	.153
incremental frauds	ρ_{M0}	(Intercept)	$-.527$	$-.567$	$-.477$
	ρ_{S0}	(Intercept)	$-.699$	$-.749$	$-.654$
extreme frauds	δ_{M0}	(Intercept)	-2.94	-3.21	-2.50
	δ_{S0}	(Intercept)	-2.37	-2.57	-2.09

units `eforensics`-fraudulent: 4980 fraudulent, 166372 not fraudulent	
manufactured votes	$F_t = 96543.42$ $[93897.90, 99448.22]^c$
total `eforensics`-fraudulent votes	$F_w = 325548.15$ $[315247.46, 335682.87]^c$

(b) `eforensics`-fraudulent Polling Station and Vote Counts by Polling Station Type

	Polling Station Type			
count	abroad	customs	prison	village
`eforensics`-fraudulent polling stations	125	157	1	4697
manufactured votes	535.9	885.9	15.5	114617.7
stolen votes	18679.1	2612.7	30.1	262391.2
`eforensics`-fraudulent votes	19215.0	3498.6	45.6	377008.9

Note: selected `eforensics` model parameter estimates (posterior means and highest posterior density credible intervals). $n = 171352$ polling station units. Region fixed-effects for turnout and vote choice are not shown. [a] 95% HPD lower bound. [b] 95% HPD upper bound. [c] posterior mean [99.5% credible interval].

Table 24: Turkey 2011 and 2015 Legislative Elections `eforensics` Estimates

(a) 2011:

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.485	.485	.486
	π_2	Incremental Fraud	.485	.484	.486
	π_3	Extreme Fraud	.029	.0285	.0303
turnout	β_0	(Intercept)	1.67	1.65	1.69
	β_1	regvidx	1.79	1.70	1.89
vote choice	γ_0	(Intercept)	−.302	−.321	−.283
	γ_1	regvidx	.000630	−.0426	.0274
incremental frauds	ρ_{M0}	(Intercept)	−.780	−.862	−.722
	ρ_{S0}	(Intercept)	−.500	−.530	−.478
extreme frauds	δ_{M0}	(Intercept)	−1.75	−1.95	−1.51
	δ_{S0}	(Intercept)	−1.68	−1.93	−1.35
(b) June 2015:					
Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.748	.508	.896
	π_2	Incremental Fraud	.223	.0754	.464
	π_3	Extreme Fraud	.0282	.0252	.0305
turnout	β_0	(Intercept)	1.53	1.24	1.69
	β_1	regvidx	1.13	.668	1.33
vote choice	γ_0	(Intercept)	−.281	−.442	−.185
	γ_1	regvidx	.00185	−.0601	.0687
incremental frauds	ρ_{M0}	(Intercept)	−.433	−.842	.177
	ρ_{S0}	(Intercept)	−.975	−1.15	−.846
extreme frauds	δ_{M0}	(Intercept)	−1.64	−2.33	−1.20
	δ_{S0}	(Intercept)	−.289	−.579	.133
(c) November 2015:					
Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.574	.477	.861
	π_2	Incremental Fraud	.386	.110	.478
	π_3	Extreme Fraud	.0403	.0285	.0451
turnout	β_0	(Intercept)	1.57	1.32	1.67
	β_1	regvidx	.980	−.631	1.61
vote choice	γ_0	(Intercept)	−.133	−.235	.116
	γ_1	regvidx	.0825	−.0808	.181
incremental frauds	ρ_{M0}	(Intercept)	−.569	−.839	.086
	ρ_{S0}	(Intercept)	−.380	−.480	−.132
extreme frauds	δ_{M0}	(Intercept)	−2.58	−3.44	−.485
	δ_{S0}	(Intercept)	−1.48	−1.87	−.940

Note: selected `eforensics` model parameter estimates (posterior means and highest posterior density credible intervals). District fixed-effects for turnout and vote choice are not shown. (a, b, c) $n = 2577$. [a] 95% HPD lower bound. [b] 95% HPD upper bound.

Table 25: Turkey 2011 and 2015 Legislative Elections `eforensics` Fraudulent Vote Estimates

(a) 2011

units `eforensics`-fraudulent: 129400 fraudulent, 70155 not fraudulent	
manufactured votes	$F_t = 702530.6\ [641642.7, 748580.5]^c$
total `eforensics`-fraudulent votes	$F_w = 3639964.6\ [3485445.9, 3742546.5]^c$

(b) June 2015

units `eforensics`-fraudulent: 11804 fraudulent, 162046 not fraudulent	
manufactured votes	$F_t = 202406.4\ [128213.2, 317492.3]^c$
total `eforensics`-fraudulent votes	$F_w = 595896.6\ [476043.5, 725646.5]^c$

(c) November 2015

units `eforensics`-fraudulent: 75086 fraudulent, 99533 not fraudulent	
manufactured votes	$F_t = 555504.4\ [356172.1, 625989.2]^c$
total `eforensics`-fraudulent votes	$F_w = 2577705.9\ [825080.6, 3204703.1]^c$

Note: `eforensics` model `eforensics`-fraudulent vote count estimates. (a) $n = 199555$, (b) $n = 173850$, (c) $n = 174619$ precinct units. ^{a} posterior mean [99.5% credible interval].

Table 26: Russian Elections of 2000 and 2003 `eforensics` Estimates

Type	Parameter	Covariate	Mean	lo[a]	up[b]
2000 President Election:					
mixture probabilities	π_1	No Fraud	.873	.852	.888
	π_2	Incremental Fraud	.101	.0886	.118
	π_3	Extreme Fraud	.0254	.0217	.0300
turnout	β_0	(Intercept)	.945	.931	.963
vote choice	γ_0	(Intercept)	.0584	.0418	.0748
incremental frauds	ρ_{M0}	(Intercept)	.590	.384	.849
	ρ_{S0}	(Intercept)	$-.0228$	$-.313$	.155
extreme frauds	δ_{M0}	(Intercept)	.602	.470	.711
	δ_{S0}	(Intercept)	$-.521$	$-.851$	$-.190$
2003 Duma Election Proportional Representation:					
mixture probabilities	π_1	No Fraud	.893	.882	.901
	π_2	Incremental Fraud	.0759	.0688	.0854
	π_3	Extreme Fraud	.0312	.0296	.0328
turnout	β_0	(Intercept)	.435	.418	.456
vote choice	γ_0	(Intercept)	$-.604$	$-.629$	$-.588$
incremental frauds	ρ_{M0}	(Intercept)	.857	.491	1.18
	ρ_{S0}	(Intercept)	.164	$-.00752$	.263
extreme frauds	δ_{M0}	(Intercept)	1.07	.956	1.21
	δ_{S0}	(Intercept)	.00413	$-.156$	.135
2003 Duma Election Single-Member Districts:[c]					
mixture probabilities	π_1	No Fraud	.862	.857	.866
	π_2	Incremental Fraud	.0969	.0926	.101
	π_3	Extreme Fraud	.0415	.0399	.0431
turnout	β_0	(Intercept)	.401	.390	.410
vote choice	γ_0	(Intercept)	$-.382$	$-.405$	$-.350$
incremental frauds	ρ_{M0}	(Intercept)	.479	.349	.605
	ρ_{S0}	(Intercept)	$-.0524$	$-.0682$	$-.0369$
extreme frauds	δ_{M0}	(Intercept)	.822	.772	.873
	δ_{S0}	(Intercept)	$-.596$	$-.827$	$-.361$

Note: selected `eforensics` model parameter estimates (posterior means and highest posterior density credible intervals). $n = 22858$. [a] 95% HPD lower bound. [b] 95% HPD upper bound. [c] District fixed-effects for turnout and vote choice are not shown.

Table 27: Russian Elections of 2004, 2007 and 2008 `eforensics` Estimates

Type	Parameter	Covariate	Mean	lo[a]	up[b]
2004 President Election:					
mixture probabilities	π_1	No Fraud	.771	.723	.809
	π_2	Incremental Fraud	.142	.108	.174
	π_3	Extreme Fraud	.0870	.0684	.104
turnout	β_0	(Intercept)	.796	.771	.817
vote choice	γ_0	(Intercept)	.812	.766	.855
incremental frauds	ρ_{M0}	(Intercept)	.0562	$-.647$	.671
	ρ_{S0}	(Intercept)	$-.272$	$-.679$	.211
extreme frauds	δ_{M0}	(Intercept)	1.33	.959	1.72
	δ_{S0}	(Intercept)	.289	$-.243$	.833
2007 Duma Election:					
mixture probabilities	π_1	No Fraud	.718	.702	.735
	π_2	Incremental Fraud	.188	.177	.197
	π_3	Extreme Fraud	.0941	.0866	.102
turnout	β_0	(Intercept)	.584	.560	.608
vote choice	γ_0	(Intercept)	.486	.470	.508
incremental frauds	ρ_{M0}	(Intercept)	1.09	.860	1.32
	ρ_{S0}	(Intercept)	$-.134$	$-.390$	.0954
extreme frauds	δ_{M0}	(Intercept)	1.42	1.16	1.67
	δ_{S0}	(Intercept)	$-.0299$	$-.415$	.305
2008 President Election:					
mixture probabilities	π_1	No Fraud	.705	.659	.729
	π_2	Incremental Fraud	.205	.189	.228
	π_3	Extreme Fraud	.0901	.0672	.115
turnout	β_0	(Intercept)	.990	.961	1.03
vote choice	γ_0	(Intercept)	.723	.700	.759
incremental frauds	ρ_{M0}	(Intercept)	.236	$-.0864$	.968
	ρ_{S0}	(Intercept)	$-.195$	$-.680$	.193
extreme frauds	δ_{M0}	(Intercept)	.966	.404	1.34
	δ_{S0}	(Intercept)	$-.175$	-1.22	.520

Note: selected `eforensics` model parameter estimates (posterior means and highest posterior density credible intervals). $n = 22858$. [a] 95% HPD lower bound. [b] 95% HPD upper bound.

Table 28: Russian Elections of 2011 and 2012 `eforensics` Estimates

Type	Parameter	Covariate	Mean	lo[a]	up[b]
2011 Duma Election:					
mixture probabilities	π_1	No Fraud	.811	.807	.814
	π_2	Incremental Fraud	.113	.110	.116
	π_3	Extreme Fraud	.0757	.0734	.0781
turnout	β_0	(Intercept)	.419	.402	.434
vote choice	γ_0	(Intercept)	$-.213$	$-.234$	$-.197$
incremental frauds	ρ_{M0}	(Intercept)	1.70	1.45	1.94
	ρ_{S0}	(Intercept)	.421	.327	.593
extreme frauds	δ_{M0}	(Intercept)	1.39	1.33	1.47
	δ_{S0}	(Intercept)	.259	.157	.367
2012 President Election:					
mixture probabilities	π_1	No Fraud	.787	.762	.806
	π_2	Incremental Fraud	.142	.124	.159
	π_3	Extreme Fraud	.0709	.0640	.0801
turnout	β_0	(Intercept)	.569	.544	.594
vote choice	γ_0	(Intercept)	.504	.467	.531
incremental frauds	ρ_{M0}	(Intercept)	1.33	.948	1.86
	ρ_{S0}	(Intercept)	.0356	$-.320$	.376
extreme frauds	δ_{M0}	(Intercept)	1.18	.978	1.35
	δ_{S0}	(Intercept)	.000804	$-.556$	.361
2016 Duma Election Proportional Representation:					
mixture probabilities	π_1	No Fraud	.796	.790	.802
	π_2	Incremental Fraud	.132	.126	.139
	π_3	Extreme Fraud	.0716	.0695	.0737
turnout	β_0	(Intercept)	$-.0444$	$-.0714$	$-.0285$
vote choice	γ_0	(Intercept)	$-.0800$	$-.0895$	$-.0701$
incremental frauds	ρ_{M0}	(Intercept)	1.48	1.16	1.80
	ρ_{S0}	(Intercept)	$-.140$	$-.194$	$-.0826$
extreme frauds	δ_{M0}	(Intercept)	.797	.756	.840
	δ_{S0}	(Intercept)	-1.12	-1.38	$-.885$

Note: selected `eforensics` model parameter estimates (posterior means and highest posterior density credible intervals). $n = 22858$. [a] 95% HPD lower bound. [b] 95% HPD upper bound.

Table 29: Russian Elections of 2016, 2018 and 2020 `eforensics` Estimates

Type	Parameter	Covariate	Mean	lo[a]	up[b]
2016 Duma Election Single-Member Districts:[c]					
mixture probabilities	π_1	No Fraud	.820	.816	.824
	π_2	Incremental Fraud	.125	.121	.129
	π_3	Extreme Fraud	.0552	.0534	.0571
turnout	β_0	(Intercept)	−.0661	−.0777	−.0534
vote choice	γ_0	(Intercept)	−.102	−.127	−.0887
incremental frauds	ρ_{M0}	(Intercept)	1.33	1.09	1.62
	ρ_{S0}	(Intercept)	−.110	−.147	−.0821
extreme frauds	δ_{M0}	(Intercept)	.765	.700	.843
	δ_{S0}	(Intercept)	−1.07	−1.38	−.829
2018 President Election:					
mixture probabilities	π_1	No Fraud	.730	.682	.763
	π_2	Incremental Fraud	.201	.181	.228
	π_3	Extreme Fraud	.0691	.0541	.0905
turnout	β_0	(Intercept)	.666	.626	.698
vote choice	γ_0	(Intercept)	1.10	1.08	1.13
incremental frauds	ρ_{M0}	(Intercept)	1.43	.676	2.08
	ρ_{S0}	(Intercept)	−.0328	−.402	.313
extreme frauds	δ_{M0}	(Intercept)	.472	.0876	.748
	δ_{S0}	(Intercept)	−2.91	−5.53	−1.31
2020 Referendum:					
mixture probabilities	π_1	No Fraud	.532	.455	.576
	π_2	Incremental Fraud	.352	.335	.369
	π_3	Extreme Fraud	.116	.0867	.177
turnout	β_0	(Intercept)	.572	.411	.660
vote choice	γ_0	(Intercept)	.874	.788	.932
incremental frauds	ρ_{M0}	(Intercept)	1.43	1.25	1.55
	ρ_{S0}	(Intercept)	.295	−.423	.652
extreme frauds	δ_{M0}	(Intercept)	−.0690	−.952	.349
	δ_{S0}	(Intercept)	−1.99	−5.80	−.582

Note: selected `eforensics` model parameter estimates (posterior means and highest posterior density credible intervals). $n = 22858$. [a] 95% HPD lower bound. [b] 95% HPD upper bound. [c] District fixed-effects for turnout and vote choice are not shown.

Table 30: Armenia 2003 President Election Round 1 `eforensics` Estimates

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.829	.808	.850
	π_2	Incremental Fraud	.146	.127	.168
	π_3	Extreme Fraud	.0252	.0174	.0335
turnout	β_0	(Intercept)	.469	.430	.506
	β_1	O2	−.00290	−.0727	.0404
	β_2	O3	.0691	.0149	.112
	β_3	O4	.0215	−.0342	.0621
vote choice	γ_0	(Intercept)	.0115	−.0467	.0672
	γ_1	O2	−.246	−.332	−.133
	γ_2	O3	−.179	−.276	.000632
	γ_3	O4	−.273	−.350	−.190
incremental frauds	ρ_{M0}	(Intercept)	.349	.149	.642
	ρ_{M1}	O2	.046	−.0621	.135
	ρ_{M2}	O3	.174	.0434	.413
	ρ_{M3}	O4	.036	−.172	.197
	ρ_{S0}	(Intercept)	−.143	−.307	.037
	ρ_{S1}	O2	−.149	−.299	−.00407
	ρ_{S2}	O3	−.0639	−.307	.218
	ρ_{S3}	O4	−.0165	−.198	.0887
extreme frauds	δ_{M0}	(Intercept)	−.455	−.794	−.143
	δ_{M1}	O2	.259	−.0677	.844
	δ_{M2}	O3	.157	−.267	.718
	δ_{M3}	O4	.266	−.0188	.521
	δ_{S0}	(Intercept)	.0947	−.349	.511
	δ_{S1}	O2	−.0577	−.223	.0664
	δ_{S2}	O3	−.208	−.539	.0869
	δ_{S3}	O4	−.248	−.561	−.00196

Note: selected `eforensics` model parameter estimates (posterior means and highest posterior density credible intervals). $n = 1763$. [a] 95% HPD lower bound. [b] 95% HPD upper bound.

Table 31: Armenia 2003 President Election Round 2 `eforensics` Estimates

Type	Parameter	Covariate	Mean	lo[a]	up[b]
mixture probabilities	π_1	No Fraud	.699	.649	.749
	π_2	Incremental Fraud	.257	.204	.310
	π_3	Extreme Fraud	.0437	.0320	.0552
turnout	β_0	(Intercept)	.660	.591	.724
	β_1	O2	$-.0686$	$-.136$	.0442
	β_2	O3	.0338	$-.0704$	.180
	β_3	O4	.0694	.0151	.109
vote choice	γ_0	(Intercept)	.700	.614	.806
	γ_1	O2	$-.153$	$-.250$	.0112
	γ_2	O3	$-.0780$	$-.166$	.00528
	γ_3	O4	$-.172$	$-.290$	$-.0385$
incremental frauds	ρ_{M0}	(Intercept)	.0261	$-.0957$	.214
	ρ_{M1}	O2	$-.0421$	$-.390$	.248
	ρ_{M2}	O3	.0974	$-.184$	.308
	ρ_{M3}	O4	.0345	$-.107$	.123
	ρ_{S0}	(Intercept)	$-.518$	$-.761$	$-.247$
	ρ_{S1}	O2	$-.113$	$-.194$	.0112
	ρ_{S2}	O3	$-.120$	$-.321$	.136
	ρ_{S3}	O4	$-.169$	$-.271$	$-.0544$
extreme frauds	δ_{M0}	(Intercept)	$-.079$	$-.584$	.242
	δ_{M1}	O2	.0643	$-.0534$	.202
	δ_{M2}	O3	$-.00956$	$-.411$	.317
	δ_{M3}	O4	.372	$-.205$	.781
	δ_{S0}	(Intercept)	$-.610$	$-.887$	$-.259$
	δ_{S1}	O2	.148	$-.236$	.402
	δ_{S2}	O3	$-.0352$	$-.639$	.485
	δ_{S3}	O4	$-.147$	$-.465$	.100

Note: selected `eforensics` model parameter estimates (posterior means and highest posterior density credible intervals). $n = 1762$. [a] 95% HPD lower bound. [b] 95% HPD upper bound.

Table 32: Armenia 2003 President Election `eforensics` Fraudulent Vote Count Totals

Round 1:
units `eforensics`-fraudulent: 256 fraudulent, 1507 not fraudulent

manufactured votes	$F_t = 29501.59\ [28036.68, 30785.69]^a$
total `eforensics`-fraudulent votes	$F_w = 51845.73\ [49807.68, 53571.11]^a$

Round 2:
units `eforensics`-fraudulent: 454 fraudulent, 1308 not fraudulent

manufactured votes	$F_t = 44980.77\ [39683.11, 48515.87]^a$
total `eforensics`-fraudulent votes	$F_w = 70216.65\ [62967.52, 74854.70]^a$

Note: [a] posterior mean [99.5% credible interval].