

eforensics Analysis of the Peru 2021 Presidential
Election

Walter R. Mebane, Jr.*

June 27, 2021

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

The recent presidential election in Peru is controversial, even as the election count process continues (Aquino 2021). I use *mesa* data obtained June 22, 2021 to estimate **eforensics**-frauds (Ferrari, Mebane, McAlister and Wu 2019; Mebane 2019*b*) in each election round. Raw data come from files: `Resultados_1ra_vuelta_Version_PCM.csv` (timestamp Jun 18 08:35) (Gobierno de Datos 2021*a*); `Resultados_2da_vuelta_Version_PCM .csv` (Jun 18 08:41) (Gobierno de Datos 2021*b*).

Table 1 shows the total counts of electors and votes in both rounds of the election in the raw data. *Número de electores* is the number of electors. *Votos en Blanco* and *Votos Nulos* are not votes for parties but are legal votes.¹ *Partido Politico Nacional Peru Libre* has the most votes in both rounds. *Votos en Blanco* has the second-most votes in the first round, while *Fuerza Popular* is the party with the next highest number of first-round votes.

eforensics operationalizes the idea that frauds occur when one party gains votes by a combination of manufacturing votes from abstentions and stealing votes from opposing parties (Mebane 2019*a*). The Bayesian specification (Mebane 2019*a*) allows posterior means and credible intervals for counts of **eforensics**-fraudulent votes to be determined both for the entire election and for individual *mesas*. The model requires that some ballot alternative be designated the “leader,” which is the alternative that the model allows to benefit from added **eforensics**-fraudulent votes. The party with the most votes in each round is this designated leader party: **eforensics**-fraudulent votes can add to the votes for *Partido Politico Nacional Peru Libre*.

Not all *mesas* have usable data. We require that each *mesa* used to estimate **eforensics** have nonmissing values for the number of electors (NVoters), the number of votes cast (NValid) and the number of votes cast for the leader (Votes). The number of electors must be weakly greater than the number of votes cast, which must be weakly greater than the number of votes for the leader. Table 2 shows the number of usable *mesas* in each round, crossed with the status of each *acta* as given by variable

¹These are included in variable `N_CVAS`, *Número de ciudadanos que votaron en la mesa según el acta de sufragio*.

Table 1: Peru 2021 President Election Elector and Vote Totals

Voters, Party or Alternative	Round	
	1	2
Número de electores hábiles de la mesa	25287954	25287954
Partido Nacionalista Peruano	230831	
El Frente Amplio por Justicia, Vida y Libertad	65300	
Partido Morado	325608	
Peru Patria Segura	55644	
Victoria Nacional	814516	
Accion Popular	1306288	
Avanza Pais - Partido de Integracion Social	1674201	
Podemos Peru	812721	
Juntos por el Peru	1132577	
Partido Popular Cristiano - PPC	286447	
Fuerza Popular	1930762	8791730
Union por el Peru	101267	
Renovacion Popular	1692279	
Renacimiento Unido Nacional	89376	
Partido Democratico Somos Peru	240234	
Partido Politico Nacional Peru Libre	2724752	8835970
Democracia Directa	50802	
Alianza para el Progreso	867025	
Votos en Blanco	2190059	121478
Votos Nulos	1123027	1107640

Note: number of voters and vote totals by party or ballot alternative.

DESCRIP_ESTADO_ACTA. In round 1 $n = 83366$ *mesas* are usable, and in round 2 $n = 86237$ *mesas* are usable.

`eforensics` estimates for round 1 are in Table 3 and for round 2 are in Table 4.² The tables show `eforensics` parameters, the total number of *mesas* classified as `eforensics`-fraudulent (H) and the number of `eforensics`-fraudulent votes. For the parameters the table reports posterior means and 95% highest-posterior density (HPD) intervals. The number of manufactured votes (F_t) is shown separately from the total number of `eforensics`-fraudulent votes (F_w); the number of stolen votes is $F_w - F_t$. For F_t and F_w the tables report posterior means and 99.5% credible intervals.

²Each Monte Carlo Markov Chain (MCMC) estimation uses four chains.

Table 2: Peru 2021 President Election *Mesa* Usability and *Acta* Status

Round	Usable	Status (DESCRIP_ESTADO_ACTA)					
		<i>Anulada</i>	<i>Anulada por Extraviada</i>	<i>Computada Resuelta</i>	<i>Contabilizada</i>	<i>En Proceso</i>	<i>Sin Instalar</i>
1	no	949	3	151	1497	23	499
	yes	0	0	2930	80427	9	0
2	no	213	0	9	17	6	6
	yes	0	0	1377	84846	14	0

Note: number of *mesas* by usability and DESCRIP_ESTADO_ACTA.

`eforensics` parameters support interpretations regarding whether the data feature malevolent distortions in votes (i.e., bona fide frauds), strategic behavior by electors or failures of election administration. An indicator for bad acts is that at least one of the extreme frauds parameters δ_{M0} and δ_{S0} differs from zero. Zero is the prior mean value for these parameters, so if a credible 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 is effectively not operative for the election. Incremental frauds parameters ρ_{M0} and ρ_{S0} can be negative when only strategic behavior is occurring, but such parameter values can also signal bad acts. If $\rho_{M0} < 0$ or $\rho_{S0} < 0$ while δ_{M0} and δ_{S0} do not differ from zero, probably only strategic behavior occurred. Multimodality in mixture probability π_2 is an indicator that votes were lost from nonleading alternatives in ways that `eforensics` may be adding to F_w , even while the leader’s votes were not actually augmented. The sizes of π_2 and π_3 or of F_t and F_w are not good indicators in themselves for whether `eforensics` estimates are responding to malevolent distortions in votes, to strategic behavior, to administrative failures or what. The parameters are really the guides to interpreting the `eforensics`-frauds.

The round 1 estimates in Table 3 show $\pi_2 = .046$ and $\pi_3 = .00001$, with ρ_{M0} , ρ_{S0} , δ_{M0} and δ_{S0} not differing from zero. $H = 610$ *mesas* are `eforensics`-fraudulent, and

Table 3: Peru 2021 President Election Round 1 **eforensics** Estimates

Type	Parameter	Covariate	Mean	lo ^a	up ^b
mixture probabilities	π_1	No Fraud	.954	.867	.9999
	π_2	Incremental Fraud	.0464	.000132	.133
	π_3	Extreme Fraud	.0000116	6.12e-09	.0000345
turnout	β_0	(Intercept)	.944	.920	.968
vote choice	γ_0	(Intercept)	-2.17	-2.34	-2.08
incremental frauds	ρ_{M0}	(Intercept)	-.378	-1.16	.0217
	ρ_{S0}	(Intercept)	-.199	-.536	.118
extreme frauds	δ_{M0}	(Intercept)	-.0202	-.0726	.0387
	δ_{S0}	(Intercept)	-.0161	-.119	.0675

units fraudulent: 610 fraudulent, 82748 not fraudulent

manufactured votes $F_t = 8770.7$ [178.4, 17394.8]^c

total fraudulent votes $F_w = 32969.8$ [618.2, 58855.3]^c

Note: **eforensics** model parameter estimates (posterior means and credible intervals). $n = 83366$ *mesa* units. ^a 95% HPD lower bound. ^b 95% HPD upper bound. ^c posterior mean [99.5% credible interval].

$F_w = 32969.8$ [618.2, 58855.3]. π_2 exhibits multimodality³ that may indicate that nonleading ballot alternatives lost votes that **eforensics** is counting as part of F_w : perhaps this connects to the problems that produced the 3 *Anulada por Extraviada actas* in Table 2. Because most of the HPD interval for ρ_{M0} is negative, the **eforensics**-frauds may also relate to strategic elector behavior (cf. Bouton and Gratton 2015).

The round 2 estimates in Table 4 show $\pi_1 = .999$, so any **eforensics**-frauds are scant. Indeed, only $H = 70$ *mesas* are **eforensics**-fraudulent. Rare as they may be, extreme frauds are more likely than incremental frauds— $\pi_3 = .001 > \pi_2 = .0001$. Both incremental and extreme frauds parameters are nonzero: $\rho_{M0} < 0$, $\delta_{M0} < 0$ and $\delta_{S0} < 0$. Especially the negative values of the extreme frauds parameters suggest that bad acts occurred. Overall the number of **eforensics**-fraudulent votes is too small to affect the outcome:

$F_w = 6322.8$ [4723.2, 9532.4] is smaller than the difference of $8835970 - 8791730 = 44240$ votes between the parties. The *mesa*-specific estimates of **eforensics**-fraudulent votes

³Across MCMC chains π_2 has estimates .0400 [.0386, .0413], .000292 [.000170, .000428], .132 [.129, .135], .0134 [.0125, .0142].

Table 4: Peru 2021 President Election Round 2 `eforensics` Estimates

Type	Parameter	Covariate	Mean	lo ^a	up ^b
mixture probabilities	π_1	No Fraud	.999	.998	.9996
	π_2	Incremental Fraud	.000115	2.31e-08	.000305
	π_3	Extreme Fraud	.000999	.000332	.00217
turnout	β_0	(Intercept)	1.16	1.14	1.17
vote choice	γ_0	(Intercept)	-.0973	-.112	-.0824
incremental frauds	ρ_{M0}	(Intercept)	-.122	-.259	-.0619
	ρ_{S0}	(Intercept)	-.0301	-.233	.0576
extreme frauds	δ_{M0}	(Intercept)	-.176	-.470	-.00994
	δ_{S0}	(Intercept)	-.0783	-.229	-.00976

units fraudulent: 70 fraudulent, 86167 not fraudulent

manufactured votes $F_t = 2345.2$ [1764.2, 3539.7]^c

total fraudulent votes $F_w = 6322.8$ [4723.2, 9532.4]^c

Note: `eforensics` model parameter estimates (posterior means and credible intervals). $n = 86237$ *mesa* units. ^a 95% HPD lower bound. ^b 95% HPD upper bound. ^c posterior mean [99.5% credible interval].

shown in Tables 5, 6 and 7 reveal that most of the `eforensics`-frauds occur in the Puno *Departamento*: 60 of the 70 `eforensics`-fraudulent *mesas* are in that *Departamento*. `eforensics`-fraudulent votes are typically high proportions of the votes for the leader in these *mesas*. `eforensics`-frauds are scant but geographically concentrated.

Based on the UBIGEO codes, the *Departamentos* that have `eforensics`-fraudulent *mesas* in Tables 5, 6 and 7 are Apurmac (03), Arequipa (04), Ayacucho (05), Cusco (07), Huancavelica (08), La Libertad (12) and Puno (20), with the most in Puno. For a list of UBIGEO codes with *provincia* and *distrito* identifying details, for each *Departamento* click the map at Oficina Nacional de Procesos Electorales (2021).⁴

⁴Note that UBIGEO codes 200905 (*provincia* San Roman) and 030320 (*provincia* Andahuaylas) are not included in the `xls` files available from Oficina Nacional de Procesos Electorales (2021).

Table 5: eforensics-Fraudulent Votes for Peru 2021 Election Round 2 *Mesas*

UBIGEO ^a	<i>Mesa</i>	Observed			eforensics-FraudulentVotes					
		NVoters	NValid	Votes	Turnout	99.5% CI		Votes	99.5% CI	
					Mean	lo	up	Mean	lo	up
030320	4475	274	254	231	53.4	48.6	57.2	144.1	136.4	151.9
040103	5435	71	69	64	14.2	12.0	16.0	38.1	34.1	41.4
050104	900574	167	164	144	33.5	30.0	36.5	88.7	82.0	94.3
050104	900575	168	163	159	33.6	29.7	36.9	90.6	83.7	97.1
050109	900589	262	235	218	50.7	45.1	54.2	138.1	129.5	146.3
071212	16869	300	271	252	58.3	53.3	62.4	158.7	149.6	168.7
071302	16972	300	279	255	59.1	53.5	62.9	159.0	149.4	167.1
080302	901372	213	205	195	42.8	38.2	46.7	114.8	106.3	121.4
120411	901871	280	248	237	54.1	49.0	57.6	147.6	138.4	155.5
200103	902804	249	237	227	49.9	45.1	53.6	134.0	123.5	142.3
200103	902805	250	228	213	48.6	44.2	51.9	131.8	122.7	139.2
200105	902806	254	234	217	49.7	44.6	53.3	134.7	124.9	142.2
200105	902808	255	243	223	51.0	46.9	54.8	136.9	129.2	143.8
200105	902809	265	243	220	51.1	48.0	55.1	140.4	133.0	147.4
200105	902810	266	251	238	53.1	47.9	57.5	142.7	132.5	151.2
200105	902812	265	248	236	52.8	47.5	57.1	142.1	133.1	149.8

Notes: DESCRIP_ESTADO_ACTA is “contabilizada” for all *mesas*. ^a *Código de Ubicación Geografica que denotan “DDppdd” (Departamento, provincia, distrito), fuente INEI. Departamento names: Apurmac (03), Arequipa (04), Ayacucho (05), Cusco (07), Huancavelica (08), La Libertad (12), Puno (20).*

Table 6: eforensics-Fraudulent Votes for Peru 2021 Election Round 2 *Mesas*

UBIGEO ^a	<i>Mesa</i>	Observed			eforensics-FraudulentVotes					
		NVoters	NValid	Votes	Turnout Mean	99.5% CI		Votes Mean	99.5% CI	
						lo	up		lo	up
200201	69893	292	259	243	56.7	51.6	60.7	154.2	144.4	162.7
200201	69896	300	277	261	58.9	53.8	63.3	159.6	149.3	168.2
200201	69899	300	274	254	58.3	52.9	62.3	158.1	148.3	166.8
200201	69901	300	273	265	59.3	53.3	63.9	159.9	149.2	169.2
200201	69902	300	277	269	59.2	54.0	63.3	160.3	149.8	169.5
200201	69903	300	277	257	58.9	54.3	63.1	159.1	148.9	167.4
200201	69904	300	273	256	58.4	53.6	62.4	158.4	149.1	166.3
200201	69905	300	278	262	58.7	53.4	63.1	159.1	149.9	168.0
200201	69906	300	267	259	58.3	53.5	62.2	158.8	150.4	166.4
200201	69942	300	279	265	59.2	54.2	63.1	160.1	150.6	168.8
200201	69944	300	270	248	58.1	53.7	62.4	158.0	149.0	167.3
200201	69945	300	277	251	58.8	53.4	63.4	158.2	148.9	165.9
200201	69947	300	282	260	59.6	53.9	64.4	159.6	149.2	167.9
200201	69948	300	281	252	58.8	53.7	62.8	158.5	149.0	166.9
200204	70007	300	266	252	58.5	52.3	62.6	158.6	148.7	167.2
200204	70018	300	278	259	59.3	54.0	63.4	159.9	149.5	168.5
200217	70198	300	260	253	57.9	53.5	61.7	158.6	149.3	167.6
200403	70436	300	264	253	58.8	54.3	62.5	159.9	151.2	167.3
200410	902861	260	241	217	50.7	46.3	54.5	137.0	128.5	145.2
200501	70585	300	278	258	59.4	54.3	63.3	159.9	150.3	168.5
200501	70586	300	270	265	58.6	53.3	63.0	159.3	149.8	167.3
200601	70737	300	265	252	58.4	53.7	62.2	158.9	149.0	166.9
200701	70853	300	263	245	57.5	52.8	61.0	157.4	147.8	165.6
200701	70866	300	276	246	58.1	53.3	61.9	157.0	148.8	164.8
200705	902873	258	235	225	50.8	45.8	54.7	137.2	127.5	145.0
200807	902879	299	271	245	57.9	52.7	61.9	157.1	147.0	165.7

Notes: DESCRIP_ESTADO_ACTA is “contabilizada” for all *mesas*. ^a Código de Ubicación Geografica que denotan “DDppdd” (Departamento, provincia, distrito), fuente INEI. Departamento names: Puno (20).

Table 7: eforensics-Fraudulent Votes for Peru 2021 Election Round 2 *Mesas*

UBIGEO ^a	<i>Mesa</i>	eforensics-FraudulentVotes								
		Observed			Turnout			99.5% CI		
		NVoters	NValid	Votes	Mean	lo	up	Mean	lo	up
200904	71817	257	235	225	50.7	45.5	54.6	137.0	127.6	145.0
200905	71850	300	275	247	58.3	53.1	62.3	157.8	148.5	166.8
200905	71851	300	281	256	59.3	53.8	63.6	159.1	149.5	167.6
200905	71852	300	276	253	58.6	53.7	62.5	158.7	148.3	166.9
200905	71854	300	286	249	59.4	54.1	64.0	157.8	149.0	165.7
200905	71855	300	293	257	60.2	54.5	64.6	159.2	149.3	166.9
200905	71857	300	290	261	59.6	54.5	63.8	159.8	149.7	167.6
200905	71860	300	294	260	60.6	55.7	64.9	161.0	152.5	168.7
200905	71862	300	285	265	59.8	54.3	64.0	160.6	150.4	168.7
200905	71863	300	282	255	59.4	53.7	63.9	159.2	149.3	167.4
200905	71865	300	281	245	58.5	53.0	62.5	157.0	148.2	165.4
200905	71868	300	282	253	58.9	54.2	62.9	158.6	149.3	166.9
200905	71869	300	284	260	59.6	54.0	63.7	160.0	149.8	167.8
200905	71872	300	284	254	59.5	53.3	64.2	158.6	148.5	166.6
200905	71876	300	282	248	58.5	53.5	62.5	157.2	148.7	164.6
200905	71885	300	286	243	58.7	53.8	62.7	156.4	148.1	164.3
200905	71886	300	290	252	59.9	53.6	64.4	158.5	148.3	166.6
200905	71893	205	197	172	40.5	36.5	43.4	108.0	100.1	114.2
200905	71894	205	198	173	40.9	37.2	44.0	108.9	102.2	115.2
200905	71900	300	283	251	59.2	53.9	63.6	158.3	148.8	166.6
201101	72052	300	265	252	58.4	52.9	62.8	158.7	148.5	167.3
201103	72071	300	271	246	58.3	52.6	62.7	157.6	148.7	166.3
201201	72114	54	54	47	11.0	8.6	12.5	28.5	24.5	31.3
201201	72155	300	275	253	58.4	53.8	62.3	158.0	149.8	166.0
201201	72156	300	271	250	58.0	53.2	62.0	157.5	148.6	166.6
201201	72160	300	271	263	59.2	54.0	63.8	159.8	150.4	168.7
201201	72162	300	267	264	58.7	53.9	62.7	160.0	149.8	168.8
201301	902898	232	218	190	45.6	41.0	49.1	121.6	113.8	129.0

Notes: DESCRIP_ESTADO_ACTA is “contabilizada” for all *mesas*. ^a Código de Ubicación Geografica que denotan “DDppdd” (Departamento, provincia, distrito), fuente INEI. Departamento names: Puno (20).

References

- Aquino, Marco. 2021. “Peru political fog thickens as election count official quits.” *Reuters.com*. June 24, <https://www.reuters.com/world/americas/peru-political-fog-thickens-election-count-official-quits-2021-06-24/>.
- Bouton, Laurent and Gabrielle Gratton. 2015. “Majority runoff elections: Strategic voting and Duverger’s hypothesis.” *Theoretical Economics* 10:283–314.
- 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.
- Gobierno de Datos. 2021a. “Resultados por mesa de las Elecciones Presidenciales 2021 Primera Vuelta - [Oficina Nacional de Procesos Electorales (ONPE)].” Fecha modificada 2021-06-25, Fecha de lanzamiento 2021-06-17, URL: <https://www.datosabiertos.gob.pe/dataset/resultados-por-mesa-de-las-elecciones-presidenciales-2021-primera-vuelta-oficina-naci>
- Gobierno de Datos. 2021b. “Resultados por mesa de las Elecciones Presidenciales 2021 Segunda Vuelta - [Oficina Nacional de Procesos Electorales (ONPE)].” Fecha modificada 2021-06-25, Fecha de lanzamiento 2021-06-18, URL: <https://www.datosabiertos.gob.pe/dataset/resultados-por-mesa-de-las-elecciones-presidenciales-2021-segunda-vuelta-oficina-naci>
- Mebane, Jr., Walter R. 2019a. “eforensics: A Bayesian Implementation of A Positive Empirical Model of Election Frauds.” Class notes for Political Science 485, Oct 21, 2019, URL: <http://www.umich.edu/~wmebane/efslides.pdf>.
- Mebane, Jr., Walter R. 2019b. “Evidence Against Fraudulent Votes Being Decisive in the Bolivia 2019 Election.” Working paper, URL: <http://www.umich.edu/~wmebane/Bolivia2019.pdf>.
- Oficina Nacional de Procesos Electorales. 2021. “NOSOTROS / ONPE en las Regiones.”

URL: <https://www.onpe.gob.pe/nosotros/onpe-regiones/>.