Online Appendix for

Corruption and Political Support: The Case of Peru's *Vacuna-gate* Scandal

Oscar Castorena
Noam Lupu
Adam Wolsky
Elizabeth J. Zechmeister

A. Question Wording

Corruption perceptions: Thinking of the politicians of Peru, how many of them do you believe are involved in corruption? (1) None; (2) Less than half of them; (3) Half of them; (4) More than half of them; (5) All
Pensando en los políticos de Perú, ¿cuántos de ellos cree usted que están involucrados en corrupción? (1) Ninguno; (2) Menos de la mitad; (3) La mitad de los políticos; (4) Más de la mitad; (5) Todos

System support (index): I will ask you to answer the following questions using a number on a scale ranging from 1 to 7, where 1 means NOT AT ALL and 7 means A LOT. If your opinion is between not at all and a lot, you would choose an intermediate score.

- To what extent do you respect the political institutions of Peru?
- To what extent do you think that citizens' basic rights are well protected by the political system of Peru?
- Still using the scale from 1, "not at all" to 7, "a lot"… To what extent do you feel proud of living under the political system of Peru?
- To what extent do you think that one should support the political system of Peru?

Le pediré que responda las siguientes preguntas utilizando un número de una escala que va de 1 a 7, en la que 1 significa NADA y 7 significa MUCHO. Si su opinión está entre nada y mucho, elegiría un puntaje intermedio.

- *¿Hasta qué punto tiene usted respeto por las instituciones políticas de Perú?*
- *¿Hasta qué punto cree usted que los derechos básicos del ciudadano están bien protegidos por el sistema político peruano?*
- *Siempre usando la escala de 1, "nada" a 7, "mucho"… ¿Hasta qué punto se siente usted orgulloso de vivir bajo el sistema político del Perú?*
- *¿Hasta qué punto piensa usted que se debe apoyar al sistema político del Perú?*

Support for democracy: I'm going to read a statement. Please tell me your opinion using a scale from 1, which means "strongly disagree" to 7, which means "strongly agree". You can use any number between 1 and 7: Democracy may have problems, but it is better than any other form of government. To what extent do you agree or disagree with this statement?
Le voy a leer una frase. Por favor dígame su opinión usando una escala que va de 1, que significa "muy en desacuerdo" a 7, que significa "muy de acuerdo". Puede usar cualquier número entre 1 y 7: Puede que la democracia tenga problemas, pero es mejor que cualquier otra forma de gobierno. ¿Hasta qué punto está de acuerdo o en desacuerdo con esta frase?

Tolerance for military coup: Some people say that under some circumstances it would be justified for the military of this country to take power by a coup d'état (military coup). In your opinion, would a military coup be justified when there is a lot of corruption? (1) It would be justified (2) It would not be justified
Alguna gente dice que en ciertas circunstancias se justificaría que los militares de este país tomen el poder por un golpe de Estado. En su opinión se justificaría que hubiera un golpe de estado por los militares frente a mucha corrupción? (1) Se justificaría (2) No se justificaría

Tolerance for executive coup: Do you believe that when the country is facing very difficult times it is justifiable for the president of the country to close the Congress and govern without Congress? (1) Yes, it is justified (2) No, it is not justified
¿Cree usted que cuando el país enfrenta momentos muy difíciles, se justifica que el presidente del país cierre el Congreso y gobierne sin Congreso? (1) Sí se justificaría (2) No se justificaría

Region: In what department of Peru do you live?
¿En qué departamento de Perú vive usted?

Urban/rural: Do you live in … (1) a city, (2) the periphery or outskirts of a city, (3) a town in a rural area, (4) a rural area?
¿Usted vive en (1) una ciudad, (2) en la periferia o alrededores de una ciudad/asentamientos humanos, (3) en un pueblo/poblado cercano a una zona/un área rural, (4) en un área/una zona rural?

Gender: For statistical purposes, could you please tell me what your gender is? (1) Man/masculine (2) Woman/feminine (3) Other
Para fines estadísticos, ¿me podría por favor decir cuál es su género? (1) Hombre/masculino (2) Mujer/femenino (3) Otro

Age: How old are you?
¿Cuál es su edad en años cumplidos?

Education: What is the highest level of education you have reached?
(0) None, (1) Primary (incomplete or complete), (2) Secondary (incomplete or complete), (3) Tertiary or university or higher (incomplete or complete)
¿Cuál es el máximo nivel educativo que usted alcanzó? (0) Ninguna, (1) Primaria (incompleta o completa), (2) Secundaria (incompleta o completa), (3) Superior, técnica o universitaria (incompleta o completa)

B. Sample Design

The 2021 AmericasBarometer survey in Peru was carried out between January 22 and March 26, 2021, as part of LAPOP's 2021 AmericasBarometer. Survey fieldwork was carried out by Instituto de Estudios Peruanos on behalf of LAPOP. Key funding for the AmericasBarometer came from USAID, Vanderbilt University, the National Science Foundation, and the Inter-American Development Bank.

Questionnaire pretesting took place on December 12 and 15, 2020 and interviewer training took place from January 8 to 11, 2021. Pilot surveys were conducted between January 13 and January 15. A full copy of the 2021 AmericasBarometer Peru questionnaire and additional information can be found at LAPOP's website at https://www.vanderbilt.edu/lapop/.

The survey collects information from a nationally representative sample of voting-age respondents, who are 18 years of age or older, are citizens or permanent residents of Peru and have access to a functioning mobile phone. The study excludes individuals with no access to mobile phones or with only landline phones. Participation in the AmericasBarometer survey is voluntary. Eligible respondents who agree to participate in the survey are administered the questionnaire after giving their consent to interviewers.

The sampling frame corresponds to all possible mobile phone numbers available in the country, drawn from the National Telephone Numbering System. A sample with 500,000 mobile phone numbers was generated by the firm. The survey firm used an automatic dialing system to call the mobile phone numbers. LAPOP Lab approved a final dataset of 3,038 complete interviews. All calls were conducted in Spanish and data was collected with STG.

With a confidence level for the national unweighted sample at 95 percent, the estimated margin of error is 1.8 percent, assuming a 50/50 response distribution on dichotomous variables.

The mobile phone number is the final unit of selection. In other words, the survey is conducted with any eligible individual who answers the call. Interviewers calling mobile phone numbers screen informants who answer the call to determine their eligibility. The study excludes business-only mobile phones.

LAPOP Lab instructed interviewers to call during business hours and on weekends, except when the potential respondent requested an appointment outside that timeframe. Callbacks after unsuccessful attempts were staggered over times of day and days of the week to maximize the chance of contacting informants and to minimize nonresponse. If no one was available to answer our call, interviewers were instructed to call back at least 4 times before a final disposition was made for that number.

Mobile phone coverage and response rates vary across socio-demographic groups, introducing survey errors that can potentially lead to biased estimations. To mitigate these sources of error, LAPOP Lab used a "responsive design" strategy. The LAPOP team continuously monitored both collected data and para-data with the goal of reducing bias in survey estimates without significantly increasing the costs of the survey. More specifically, they pre-identified elements

that can affect costs and errors of survey estimates, monitored those elements during the initial data collection stages, and adjusted those elements while data collection is in progress. With the purpose of balancing the sample to mirror population distributions, the fieldwork team screened out individuals when necessary from overrepresented population group(s) during the final stages of data collection.

The variables monitored were region, age, gender, urbanization, and education. When imbalances were found between the sample and target population, the LAPOP team also examined correlations between the relevant variables and region, cellphone provider, the time and day of interview, and interviewer effects. Where necessary, LAPOP instituted a filter to screen our individuals from overrepresented groups until the weighting effect fell below 1.5.

The LAPOP team's analysis of the Peru survey found that the effective sample was imbalanced with regard to region at the end of February. In particular, respondents from the Lima and Callao provinces were overrepresented. Beginning February 26, interviews were assigned at random to a version of the questionnaire that included the filter or one that did not; when that mixed approach did not fully balance the sample, beginning on March 9 all interviews filtered out Lima area residents (those living in the city as well as the Callao and Lima provinces) until the fieldwork ended on March 26 (previously rescheduled interviews went forward). The regional distribution of the sample is summarized below.

Table A1: *Regional Distribution of Interviews*

	Before filter		After filter		Total		Census	
Region	**No.**	**%**	**No.**	**%**	**No.**	**%**	**No.**	**%**
Amazonas	33	1%	6	1%	39	1%	235,657	1%
Ancash	103	4%	26	6%	129	4%	736,289	4%
Apurimac	35	1%	9	2%	44	1%	266,194	1%
Arequipa	125	5%	46	10%	171	6%	991,969	5%
Ayacucho	42	2%	10	2%	52	2%	403,809	2%
Cajamarca	113	4%	34	7%	147	5%	872,747	4%
Cusco	100	4%	21	5%	121	4%	809,150	4%
Huancavelica	18	1%	3	1%	21	1%	219,406	1%
Huanuco	57	2%	9	2%	66	2%	463,043	2%
Ica	99	4%	30	6%	129	4%	580,269	3%
Junin	113	4%	36	8%	149	5%	828,244	4%
La Libertad	147	6%	42	9%	189	6%	1,195,702	6%
Lambayeque	109	4%	33	7%	142	5%	811,573	4%
Lima	1,002	39%	32	7%	1,034	34%	7,638,447	37%
Loreto	64	2%	11	2%	75	2%	510,979	3%
Madre de Dios	14	1%	4	1%	18	1%	91,491	0%
Moquegua	13	1%	3	1%	16	1%	126,342	1%
Pasco	17	1%	3	1%	20	1%	169,985	1%
Piura	142	6%	47	10%	189	6%	1,206,440	6%
Puno	63	2%	14	3%	77	3%	817,428	4%
San Martin	73	3%	15	3%	88	3%	516,331	3%

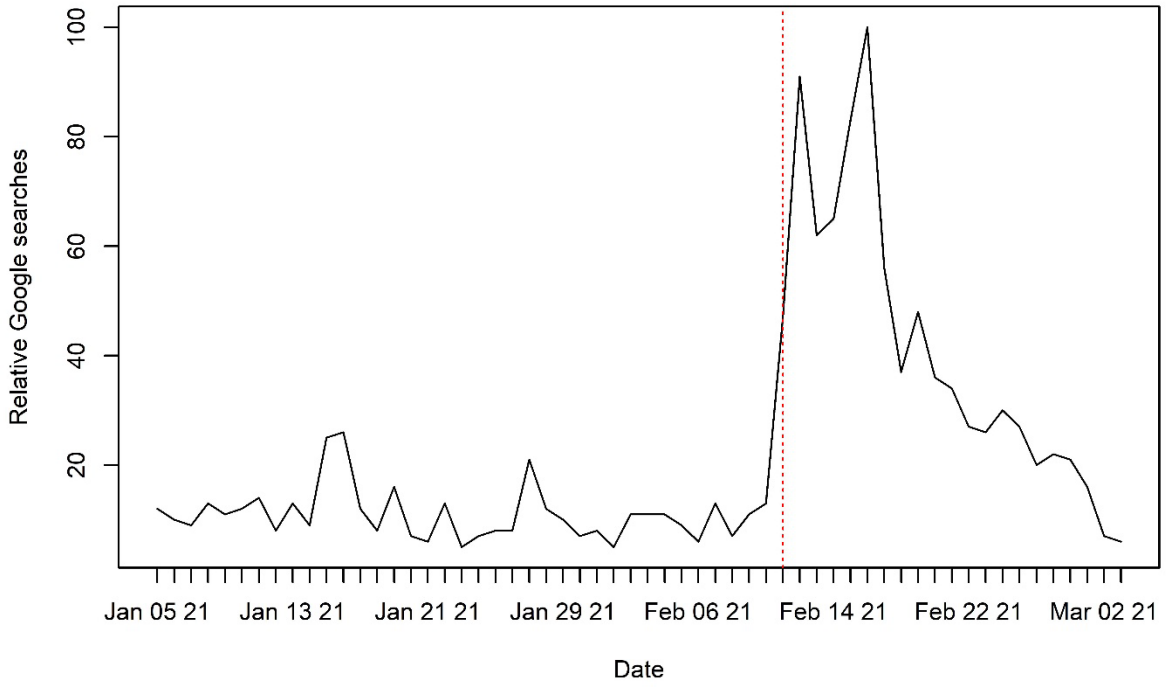
Tacna	31	1%	13	3%	44	1%	237,769	1%
Tumbes	30	1%	8	2%	38	1%	148,193	1%
Ucayali	28	1%	11	2%	39	1%	300,098	1%
Total	2,571	100%	466	100%	3,037	100%	20,177,555	100%

Interviewers obtained voluntary and informed consent from participants using an IRB-approved consent protocol. LAPOP research is overseen by the Institutional Review Board at Vanderbilt University. The survey did not use deception. Since the survey was conducted over the phone, informed consent was obtained verbally from all respondents prior to beginning the survey. The consent script was programmed into the computer software used by interviewers; they had to click that the respondent had consented in order to proceed to the questionnaire. Respondents also consented to have the interview recorded for quality control purposes. Respondents were not compensated.

The response rate (AAPOR R3) for the survey was 2.8%.

C. Additional Evidence

Figure A1: *Google Search Patterns*



Note: Figure shows the relative number of Google searches in Peru for the term "vizcarra" between January 5 and March 3, 2021. Red line indicates February 11.
Source: Google Trends

Figure A2: *National Newspaper Covers from February 12, 2021*

Source: https://www.rcrperu.com/portadas/portadas-de-principales-diarios-a-nivel-nacional-y-regional-viernes-12-de-febrero-del-2021/

Table A2: *Balance Tests*

	Prior to Scandal	Post-Scandal	Difference	P-value
Female	0.451	0.479	-0.028	0.129
18-25	0.239	0.230	0.008	0.589
26-35	0.242	0.267	-0.024	0.129
36-45	0.208	0.215	-0.007	0.660
46-55	0.167	0.150	0.017	0.193
56-65	0.091	0.094	-0.004	0.739
66+	0.053	0.044	0.009	0.270
None or Primary Ed.	0.069	0.074	-0.005	0.570
Secondary Ed.	0.337	0.335	0.002	0.890
Tertiary/Higher Ed.	0.594	0.591	0.003	0.868
Live in a city	0.448	0.474	-0.026	0.155
Live on outskirts of city	0.269	0.260	0.009	0.565
Live in a town/village	0.102	0.105	-0.003	0.814
Live in a rural area	0.181	0.162	0.019	0.161
Lima area	0.400	0.294	0.106	0.000
Observations	1,332	1,706		

Table A3: *Treatment and Control Groups After Coarsened Exact Matching*

	Control	**Treatment**
Female	0.481	0.481
	(0.014)	(0.012)
18-25	0.249	0.231
	(0.012)	(0.010)
26-35	0.252	0.270
	(0.012)	(0.011)
36-45	0.219	0.219
	(0.012)	(0.010)
46-55	0.152	0.152
	(0.010)	(0.009)
56-65	0.088	0.088
	(0.008)	(0.007)
66+	0.041	0.041
	(0.006)	(0.005)
None or Primary Ed.	0.073	0.069
	(0.007)	(0.006)
Secondary Ed.	0.331	0.335
	(0.013)	(0.012)
Tertiary/Higher Ed.	0.596	0.596
	(0.014)	(0.012)
Lima area	0.292	0.292
	(0.013)	(0.011)
Live in a city	0.480	0.480
	(0.014)	(0.012)
Live on outskirts of city	0.261	0.261
	(0.012)	(0.011)
Live in a town/village	0.098	0.098
	(0.008)	(0.007)
Live in a rural area	0.161	0.161
	(0.010)	(0.009)
Observations	1,285	1,672

Standard errors in parentheses

Table A4: *Item Nonresponse with CEM Weights*

	Prior to Scandal	Post Scandal	Difference	P-value
Perceptions of Corruption	0.008	0.013	-0.005	0.34
System Support	0.032	0.015	0.017	0.05
Support for Democracy	0.007	0.004	0.003	0.25

Figure A3: *Comparison of Pre- and Post-Scandal Variable Distributions*

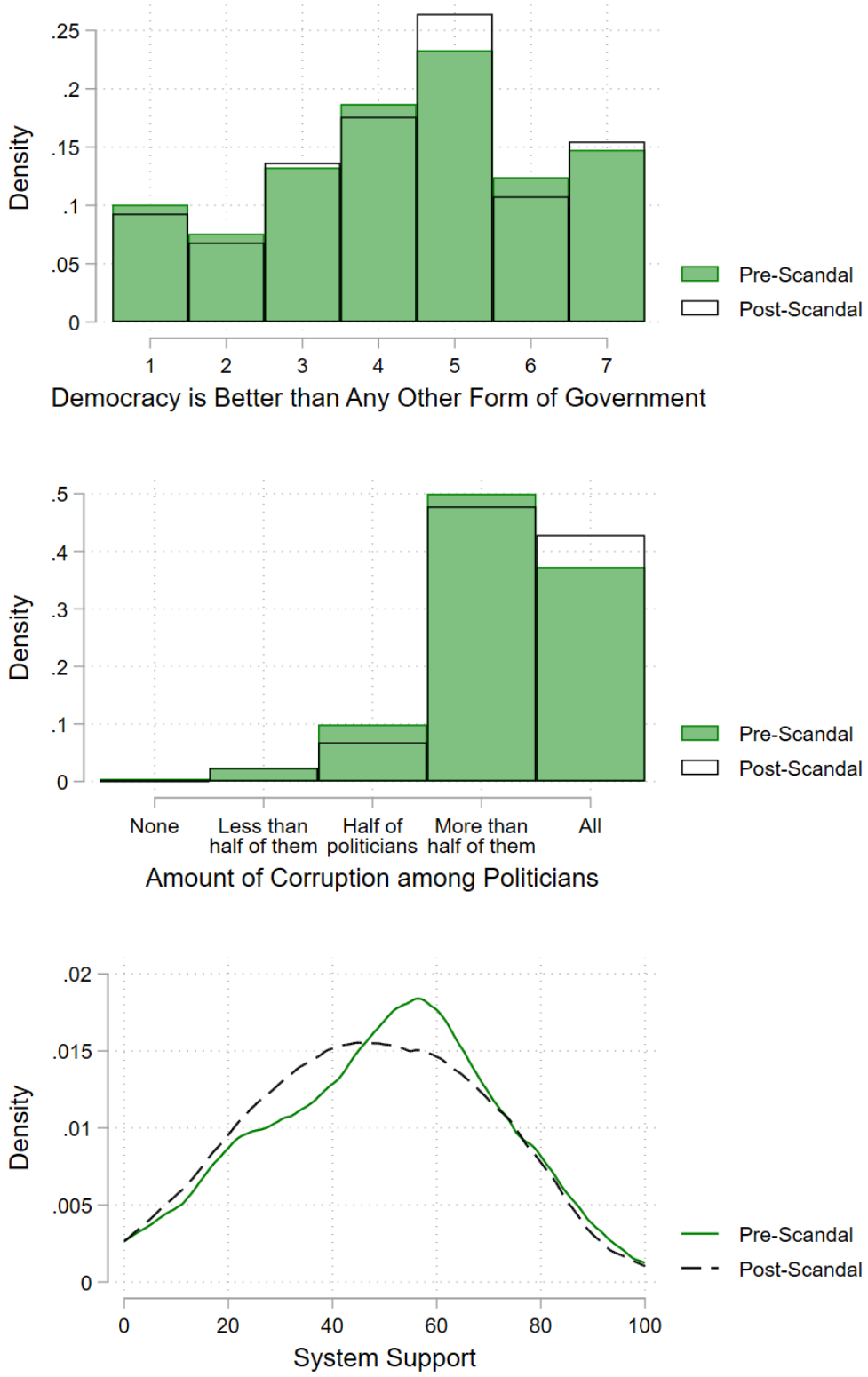


Table A5: *Effect of Scandal on Corruption Perceptions*

	CEM weights	**No weights**		**Survey weights**	
	(1)	(2)	(3)	(4)	(5)
Post-Scandal	2.484**	2.401*	2.432*	2.917*	3.005*
	(0.934)	(0.941)	(0.953)	(1.211)	(1.227)
Female			2.506**		2.846*
			(0.953)		(1.177)
26-35			4.356**		3.549*
			(1.331)		(1.641)
36-45			1.332		0.750
			(1.416)		(1.810)
46-55			2.361		2.210
			(1.555)		(1.950)
56-65			1.655		-0.804
			(1.857)		(2.541)
66+			2.228		0.898
			(2.344)		(2.956)
Secondary			5.988**		4.543+
			(1.925)		(2.620)
Tertiary or higher			4.376*		3.840
			(1.909)		(2.622)
City outskirts			0.175		0.850
			(1.195)		(1.478)
Town in rural area			0.922		2.383
			(1.627)		(1.963)
Rural area			-0.907		0.412
			(1.410)		(1.897)
Lima area			-0.403		0.169
			(1.043)		(1.257)
Constant	80.24**	80.35**	72.70**	79.98**	72.77**
	(0.694)	(0.700)	(2.358)	(0.939)	(3.309)
Observations	1,489	1,517	1,492	1,517	1,492

Standard errors in parentheses
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table A6: *Effect of Scandal on System Support*

	CEM weights		**No weights**		**Survey weights**
	(1)	(2)	(3)	(4)	(5)
Post-Scandal	-2.031^{+}	-2.044^{+}	-2.387*	-2.196	-2.760^{+}
	(1.207)	(1.180)	(1.202)	(1.476)	(1.509)
Female			3.507**		3.743**
			(1.193)		(1.441)
26-35			-5.660**		-5.877**
			(1.685)		(2.062)
36-45			-6.630**		-7.537**
			(1.767)		(2.063)
46-55			-5.233**		-7.809**
			(1.936)		(2.309)
56-65			-2.433		-4.328
			(2.355)		(2.952)
66+			-4.620		-5.110
			(3.097)		(3.349)
Secondary			-0.819		-2.163
			(2.680)		(2.957)
Tertiary or higher			-0.829		-3.564
			(2.647)		(2.946)
City outskirts			-0.625		-1.922
			(1.453)		(1.771)
Town in rural area			0.252		-0.829
			(2.142)		(2.748)
Rural area			1.886		1.095
			(1.730)		(2.112)
Lima area			-0.914		-1.038
			(1.276)		(1.583)
Constant	49.40**	49.33**	53.05**	49.41**	56.50**
	(0.921)	(0.893)	(3.152)	(1.136)	(3.505)
Observations	1,417	1,465	1,437	1,465	1,437

Standard errors in parentheses
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table A7: *Effect of Scandal on Support for Democracy*

	CEM weights		**No weights**		**Survey weights**
	(1)	(2)	(3)	(4)	(5)
Post-Scandal	1.573	0.894	1.533	-1.278	-0.719
	(1.115)	(1.098)	(1.099)	(1.377)	(1.371)
Female			1.571		0.615
			(1.094)		(1.352)
26-35			-3.874*		-5.205**
			(1.543)		(1.912)
36-45			-4.491**		-6.331**
			(1.627)		(1.950)
46-55			-2.090		-2.704
			(1.783)		(2.217)
56-65			-1.225		-2.082
			(2.145)		(2.743)
66+			1.001		1.992
			(2.756)		(3.457)
Secondary			2.765		2.520
			(2.297)		(2.780)
Tertiary or higher			7.830**		7.177**
			(2.276)		(2.784)
City outskirts			-5.630**		-5.339**
			(1.356)		(1.679)
Town in rural area			-7.364**		-11.86**
			(1.912)		(2.484)
Rural area			-2.954+		-1.639
			(1.603)		(1.952)
Lima area			5.468**		5.510**
			(1.184)		(1.443)
Constant	55.05**	55.64**	52.16**	55.36**	54.95**
	(0.839)	(0.823)	(2.771)	(1.020)	(3.452)
Observations	2,942	3,022	2,967	3,022	2,967

Standard errors in parentheses
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table A8: *Effect of Scandal on Individual System Support Items*

	Respect for Political Institutions	**Basic Rights Are Protected**	**Pride in Political System**	**People Should Support Political System**
	(1)	(2)	(3)	(4)
Post-Scandal	-1.590	0.338	-3.836*	-2.986+
	(1.684)	(1.456)	(1.683)	(1.710)
Constant	61.670**	35.285**	45.682**	59.015**
	(1.281)	(1.110)	(1.282)	(1.301)
Observations	1,442	1,438	1,439	1,437

Notes: Variables are rescaled to [0,100] and all models include CEM weights. Standard errors in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table A9: *Effect of Scandal on Tolerance for Military Coup*

	CEM weights	No weights		Survey weights	
	(1)	(2)	(3)	(4)	(5)
Post-Scandal	-0.027	0.070	0.052	-0.157	-0.165
	(0.159)	(0.156)	(0.166)	(0.189)	(0.199)
Female			0.029		0.129
			(0.164)		(0.197)
26-35			0.174		0.193
			(0.234)		(0.283)
36-45			0.007		0.037
			(0.241)		(0.293)
46-55			-0.471^{+}		-0.217
			(0.262)		(0.316)
56-65			-0.237		-0.331
			(0.316)		(0.376)
66+			-0.619		-1.235*
			(0.458)		(0.565)
Secondary			0.319		0.364
			(0.359)		(0.422)
Tertiary or higher			-0.068		0.128
			(0.357)		(0.422)
City outskirts			0.517**		0.515*
			(0.197)		(0.236)
Town in rural area			0.880**		1.013*
			(0.333)		(0.402)
Rural area			0.401^{+}		0.580^{+}
			(0.239)		(0.296)
Lima area			-0.568**		-0.558**
			(0.177)		(0.210)
Constant	0.113	0.020	-0.057	0.183	-0.178
	(0.121)	(0.117)	(0.430)	(0.139)	(0.509)
Observations	649	670	655	670	655

Notes: The variable is recoded so that higher values indicate more tolerance of a military coup. Logistic regression results. Note that due to space constraints this item was administered only to one quarter of the sample. Standard errors in parentheses.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table A10: *Effect of Scandal on Tolerance for Executive Coup*

	CEM weights	**No weights**		**Survey weights**	
	(1)	(2)	(3)	(4)	(5)
Post-Scandal	0.100	0.069	0.061	0.110	0.097
	(0.109)	(0.107)	(0.110)	(0.129)	(0.133)
Female			0.145		-0.107
			(0.109)		(0.131)
26-35			-0.029		-0.050
			(0.154)		(0.185)
36-45			-0.134		-0.279
			(0.162)		(0.195)
46-55			-0.095		-0.100
			(0.177)		(0.218)
56-65			-0.315		-0.062
			(0.215)		(0.257)
66+			-0.319		-0.585^{+}
			(0.282)		(0.348)
Secondary			0.215		0.238
			(0.241)		(0.280)
Tertiary or higher			-0.173		-0.162
			(0.239)		(0.282)
City outskirts			-0.053		-0.031
			(0.133)		(0.161)
Town in rural area			-0.460*		-0.678**
			(0.202)		(0.254)
Rural area			-0.082		-0.147
			(0.157)		(0.188)
Lima area			0.120		0.080
			(0.117)		(0.137)
Constant	-0.309**	-0.289**	-0.193	-0.276**	-0.060
	(0.083)	(0.081)	(0.284)	(0.097)	(0.338)
Observations	1,421	1,461	1,433	1,461	1,433

Notes: The variable is recoded so that higher values indicate more tolerance of a coup. Logistic regression results. Standard errors in parentheses.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table A11: *Heterogeneous Effects of Scandal on Corruption Perceptions*

	(1)	(2)	(3)	(4)
Post-Scandal	1.270	-2.998	3.983	4.118
	(1.295)	(1.994)	(3.866)	(4.500)
Female	0.932			
	(1.386)			
Post-Scandal × Female	2.550			
	(1.866)			
Age		-1.365**		
		(0.484)		
Post-Scandal × Age		2.036**		
		(0.655)		
Education			1.211	
			(1.132)	
Post-Scandal × Education			-0.590	
			(1.490)	
Political knowledge				-0.355
				(1.033)
Post-Scandal × Political knowledge				-0.498
				(1.386)
Constant	79.790**	83.913**	77.184**	81.357**
	(0.964)	(1.475)	(2.942)	(3.314)
Observations	1,489	1,489	1,489	1,489

Notes: Models include CEM weights. Standard errors in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table A12: *Heterogeneous Effects of Scandal on System Support*

	(1)	(2)	(3)	(4)
Post-Scandal	-1.986	-6.085*	-9.760^{+}	-3.891
	(1.670)	(2.614)	(5.115)	(5.707)
Female	3.603^{+}			
	(1.841)			
Post-Scandal × Female	-0.183			
	(2.411)			
Age		-1.971**		
		(0.651)		
Post-Scandal × Age		1.510^{+}		
		(0.857)		
Education			-2.628^{+}	
			(1.454)	
Post-Scandal × Education			3.040	
			(1.954)	
Political knowledge				-0.265
				(1.283)
Post-Scandal × Political knowledge				0.569
				(1.710)
Constant	47.687**	54.708**	56.075**	50.250**
	(1.268)	(1.981)	(3.810)	(4.240)
Observations	1,417	1,417	1,417	1,417

Notes: Models include CEM weights. Standard errors in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table A13: *Heterogeneous Effects of Scandal on Support for Democracy*

	(1)	(2)	(3)	(4)
Post-Scandal	0.196	2.842	-7.253	3.650
	(1.547)	(2.405)	(4.618)	(5.249)
Female	-1.265			
	(1.679)			
Post-Scandal × Female	2.865			
	(2.232)			
Age		-0.044		
		(0.590)		
Post-Scandal × Age		-0.466		
		(0.788)		
Education			3.604**	
			(1.331)	
Post-Scandal × Education			3.489*	
			(1.775)	
Political knowledge				4.609**
				(1.191)
Post-Scandal × Political knowledge				-0.765
				(1.597)
Constant	55.656**	55.167**	45.948**	40.433**
	(1.163)	(1.798)	(3.464)	(3.869)
Observations	2,942	2,942	2,942	2,942

Notes: Models include CEM weights. Standard errors in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$