

Online Appendix for
“How Worried Should We Be?
The Implications of Fabricated Survey Data for Political Science”

Table of Contents

A. Ethical Considerations	1
B. Coding and Matching Procedures	2
C. Supporting Information for Results in Main Text.....	3
Table A1. Question Code Key	3
Table A2. Differences by Variable	6
Table A3. Question Wording.....	9
Table A4. Regression Models of Support for Democracy.....	11
Table A5. Chow Tests for Support for Democracy Models	12
Table A6. Regression Models of Political Tolerance	13
Table A7. Chow Tests for Political Tolerance Models.....	14
Figure A1: Mean Differences from Fraudulent Responses	15
Figure A2. Mean Differences from Random Responding	16
Figure A3. Mean Differences from Speed Responding.....	17
Figure A4. Mean Differences from Middle Responding	18
Figure A5. Differences in Analyses of Change over Time.....	19
Figure A6. Difference in Proportion Selecting Center of Scale	20
Figure A7. Difference in Proportion Selecting Extremes.....	21
D. Investigating Interviewer Bias in Fraudulent Responses	22
Figure A8. Comparing Modal Interviewer Response to Fraudulent and Clean Responses.....	23
E. Comparing Fieldwork Patterns in Clean and Fraudulent Data.....	24
Figure A9. Distribution of Interviews by Day of Week	24
Figure A10. Distribution of Interviews over Fieldwork Period.....	25
Figure A11. When Enumerators Fabricate First Interview.....	26
F. Replication Using Peru 2017 Survey.....	27
Table A8. Item-Level Effects of Fabricated Data.....	27
References.....	28

A. Ethical Considerations

The data in this paper come from the LAPOP Lab's AmericasBarometer surveys of the public in the Americas. These surveys are approved by Vanderbilt University's Institutional Review Board and adhere to the American Political Science Association's *Principles and Guidance for Human Subjects Research*. See LAPOP's statement on ethics at <https://www.vanderbilt.edu/lapop/ab2018/Ethics-Statement.php>.

B. Coding and Matching Procedures

Identifying fraudulent interviews: Among all cancelled interviews, identify interviews that are flagged by the multi-indicator LAPOP quality control system for being in the incorrect location (or changing locations mid-interview); having an incorrect or missing verification statement by the interviewer; having auditor notes regarding changes in the respondent's or interviewer's voices; appearing to use duplicate photos, audio files, or GPS coordinates of other interviews; having empty photo or audio files; being conducted on devices that were not approved or were modified (e.g., GPS disabled); having background pictures suggesting they did not take place in a residential location (e.g., in a car); and/or having questions that were answered but not read out loud (particularly those for which the audio was missing voices but was not empty). Cross-check these with the auditors' notes and assign the canceled interview a score of "likely fraudulent" if the sum of the evidence supports that designation. Robustness check on our coding: A research assistant independently reviewed the canceled interview and identified fraudulent cases. Our coding and theirs coincided 92% of the time.

Matching fraudulent and clean interviews: We matched the set of interviews coded as fraudulent with corresponding genuine interviews. To do so, we paired fraudulent and clean interviews that were exact matches on sex (dichotomous variable for male, female), age group (specifically the age quota category indicated by the enumerator, with three categories: 18-29, 30-44, 45+), and the primary sampling unit (the final Venezuela survey included 83 primary sampling units averaging 18 interviews per primary sampling unit). For example, a fraudulent interview for a male respondent, in the 18-29 age group, and in primary sampling unit #12 is only matched to a clean interview with a male respondent, in the 18-29 age group, and in primary sampling unit #12. We use these three variables to generate exact matches since sex, age, and geographic cluster are the defining features of sample. We used the `cem` package in Stata (Blackwell et al. 2009) to implement the exact matching.

C. Supporting Information for Results in Main Text

Table A1. *Question Code Key*

Our Variable Code	AmericasBarometer Questionnaire Code
3pt_v1	venprot12
3pt_v2	venprot11
3pt_v3	soct2
3pt_v4	lib4
3pt_v5	lib2c
3pt_v6	lib2b
3pt_v7	lib1
3pt_v8	idio2
4pt_v9	vengrp4
4pt_v10	vengrp3
4pt_v11	sd6new2
4pt_v12	sd3new2
4pt_v13	sd2new2
4pt_v14	poll
4pt_v15	pn5
4pt_v16	pn4
4pt_v17	mil10un2
4pt_v18	mil10un1
4pt_v19	mil10oas2
4pt_v20	mil10oas1
4pt_v21	mil10e2
4pt_v22	mil10e1
4pt_v23	mil10a2
4pt_v24	mil10a1
4pt_v25	ls3
4pt_v26	it1
4pt_v27	fear11
4pt_v28	env2b
4pt_v29	drk1
4pt_v30	cp8
4pt_v31	cp7
4pt_v32	cp6
4pt_v33	cp20
4pt_v34	cp13
4pt_v35	aoj12
4pt_v36	aoj11
5pt_v37	venprot10
5pt_v38	vengrp2
5pt_v39	vengrp1
5pt_v40	venesc3

5pt_v41	venesc2b
5pt_v42	polz1a
5pt_v43	m2
5pt_v44	m1
7pt_v45	venvb20
7pt_v46	venvb19
7pt_v47	venvb18
7pt_v48	venps3
7pt_v49	venps2
7pt_v50	venps1
7pt_v51	vengrp9
7pt_v52	vengrp8
7pt_v53	vengrp7
7pt_v54	vengrp6
7pt_v55	vengrp5
7pt_v56	vengrp16
7pt_v57	vengrp15
7pt_v58	vengrp14
7pt_v59	vengrp13
7pt_v60	vengrp12
7pt_v61	vengrp11
7pt_v62	vengrp10
7pt_v63	venct3
7pt_v64	venct2
7pt_v65	venct1
7pt_v66	venb51
7pt_v67	venb11
7pt_v68	venb10
7pt_v69	ros4
7pt_v70	ros1
7pt_v71	pr3en
7pt_v72	pr3dn
7pt_v73	polz1
7pt_v74	media4b
7pt_v75	media4
7pt_v76	media3
7pt_v77	media2b
7pt_v78	media2
7pt_v79	media1
7pt_v80	ing4
7pt_v81	env1c
7pt_v82	eff2
7pt_v83	eff1
7pt_v84	dst1b
7pt_v85	b6
7pt_v86	b47a

7pt_v87	b43
7pt_v88	b4
7pt_v89	b37
7pt_v90	b32
7pt_v91	b3
7pt_v92	b21a
7pt_v93	b21
7pt_v94	b2
7pt_v95	b18
7pt_v96	b13
7pt_v97	b12
7pt_v98	b11
7pt_v99	b1
7pt_v100	aoj22new
10pt_v101	ven13
10pt_v102	ven12
10pt_v103	l1
10pt_v104	e5
10pt_v105	e3
10pt_v106	e16
10pt_v107	e15
10pt_v108	d6
10pt_v109	d5
10pt_v110	d4
10pt_v111	d3
10pt_v112	d2
10pt_v113	d1

Table A2. Differences by Variable

Question	Standard Deviation		Non-Response Rate		Difference
	Clean	Fraudulent	Clean	Fraudulent	
3pt_v1	36.38	35.08	0.019	0.014	0.005
3pt_v2	37.62	38.14	0.067	0.055	0.012
3pt_v3	16.93	17.98	0.007	0.000	0.007
3pt_v4	28.43	29.02	0.017	0.026	-0.010
3pt_v5	36.04	32.23	0.014	0.029	-0.014
3pt_v6	38.76	34.05	0.019	0.026	-0.007
3pt_v7	37.20	32.85	0.026	0.019	0.007
3pt_v8	26.80	22.70	0.002	0.005	-0.002
4pt_v9	38.89	36.39	0.045	0.064	-0.019
4pt_v10	38.72	37.64	0.055	0.074	-0.019
4pt_v11	25.64	25.13	0.002	0.024	-0.021
4pt_v12	25.99	26.41	0.055	0.060	-0.005
4pt_v13	23.96	25.65	0.007	0.017	-0.010
4pt_v14	36.16	33.42	0.007	0.007	0.000
4pt_v15	36.19	34.60	0.033	0.033	0.000
4pt_v16	27.45	24.23	0.036	0.048	-0.012
4pt_v17	29.40	28.45	0.337	0.340	-0.004
4pt_v18	27.15	26.67	0.382	0.275	0.107
4pt_v19	31.30	30.99	0.346	0.376	-0.029
4pt_v20	28.47	28.59	0.344	0.320	0.025
4pt_v21	31.70	35.62	0.341	0.320	0.022
4pt_v22	33.94	33.62	0.363	0.252	0.111
4pt_v23	33.51	34.04	0.505	0.497	0.007
4pt_v24	29.44	32.84	0.514	0.369	0.145
4pt_v25	32.48	33.53	0.005	0.002	0.002
4pt_v26	32.02	30.37	0.007	0.014	-0.007
4pt_v27*	33.09	27.60	0.005	0.007	-0.002
4pt_v28	25.30	27.10	0.010	0.033	-0.024
4pt_v29	32.52	28.98	0.029	0.041	-0.012
4pt_v30	32.77	28.86	0.007	0.005	0.002
4pt_v31	32.34	28.83	0.012	0.007	0.005
4pt_v32	41.06	36.06	0.007	0.007	0.000
4pt_v33*	25.64	17.53	0.000	0.016	-0.016
4pt_v34*	25.43	19.39	0.012	0.010	0.002
4pt_v35	31.70	30.16	0.017	0.038	-0.022
4pt_v36	32.53	32.39	0.005	0.017	-0.012
5pt_v37	36.38	32.46	0.019	0.024	-0.005
5pt_v38	31.43	29.28	0.067	0.060	0.007

5pt_v39	32.35	30.22	0.069	0.057	0.012
5pt_v40	14.49	15.75	0.007	0.012	-0.005
5pt_v41	21.60	23.29	0.000	0.005	-0.005
5pt_v42	26.36	28.43	0.021	0.014	0.007
5pt_v43	27.21	25.49	0.040	0.033	0.007
5pt_v44	28.89	28.27	0.024	0.017	0.007
7pt_v45	31.05	27.36	0.060	0.064	-0.005
7pt_v46	31.18	27.42	0.064	0.069	-0.005
7pt_v47	34.82	30.78	0.045	0.043	0.002
7pt_v48	37.20	33.58	0.033	0.045	-0.012
7pt_v49	35.62	31.68	0.048	0.069	-0.022
7pt_v50	40.98	37.05	0.055	0.084	-0.029
7pt_v51	34.58	32.86	0.017	0.033	-0.017
7pt_v52	35.54	30.97	0.024	0.031	-0.007
7pt_v53	36.10	32.18	0.026	0.036	-0.010
7pt_v54	39.86	34.47	0.012	0.038	-0.026
7pt_v55	41.76	38.70	0.012	0.033	-0.022
7pt_v56	31.28	30.30	0.036	0.031	0.005
7pt_v57	31.12	28.95	0.045	0.026	0.019
7pt_v58	31.14	26.36	0.026	0.036	-0.010
7pt_v59	39.83	37.68	0.029	0.029	0.000
7pt_v60	40.12	38.05	0.026	0.024	0.002
7pt_v61	30.91	28.15	0.021	0.024	-0.002
7pt_v62	33.17	30.59	0.019	0.021	-0.002
7pt_v63	35.48	34.90	0.074	0.129	-0.055
7pt_v64	30.67	31.79	0.074	0.091	-0.017
7pt_v65	37.22	32.91	0.057	0.093	-0.036
7pt_v66	34.60	31.32	0.062	0.074	-0.012
7pt_v67*	38.60	31.98	0.017	0.029	-0.012
7pt_v68	31.64	31.11	0.033	0.019	0.014
7pt_v69	32.80	31.10	0.057	0.045	0.012
7pt_v70	36.13	31.86	0.043	0.021	0.021
7pt_v71	37.73	33.69	0.069	0.072	-0.003
7pt_v72	36.48	32.40	0.038	0.033	0.005
7pt_v73	26.40	29.90	0.026	0.019	0.007
7pt_v74	31.21	28.90	0.036	0.081	-0.045
7pt_v75	31.65	28.93	0.056	0.063	-0.007
7pt_v76	29.94	27.37	0.038	0.064	-0.026
7pt_v77	31.11	28.01	0.032	0.091	-0.060
7pt_v78	32.79	29.28	0.060	0.058	0.002
7pt_v79	27.94	29.61	0.031	0.043	-0.012
7pt_v80	33.63	32.55	0.040	0.041	0.000
7pt_v81	38.25	35.56	0.010	0.019	-0.010

7pt_v82	30.88	28.96	0.026	0.031	-0.005
7pt_v83	36.73	33.57	0.036	0.036	0.000
7pt_v84	29.45	28.90	0.014	0.026	-0.012
7pt_v85*	38.32	31.12	0.040	0.048	-0.007
7pt_v86	36.50	31.04	0.021	0.029	-0.007
7pt_v87*	14.92	22.69	0.007	0.005	0.002
7pt_v88	36.17	31.13	0.031	0.026	0.005
7pt_v89	30.48	26.26	0.019	0.021	-0.002
7pt_v90	32.50	29.34	0.007	0.019	-0.012
7pt_v91	30.68	27.99	0.033	0.029	0.005
7pt_v92	36.08	30.43	0.019	0.026	-0.007
7pt_v93	31.58	27.97	0.021	0.031	-0.010
7pt_v94	36.18	32.10	0.029	0.029	0.000
7pt_v95	30.78	26.76	0.026	0.024	0.002
7pt_v96	33.72	32.70	0.029	0.029	0.000
7pt_v97	36.85	31.35	0.029	0.033	-0.005
7pt_v98	37.23	31.38	0.019	0.038	-0.019
7pt_v99	30.35	28.18	0.036	0.021	0.014
7pt_v100	29.21	29.32	0.007	0.026	-0.019
10pt_v101	35.81	33.62	0.098	0.090	0.007
10pt_v102	36.89	32.65	0.105	0.107	-0.002
10pt_v103	32.37	30.41	0.107	0.064	0.043
10pt_v104	34.78	32.82	0.010	0.024	-0.014
10pt_v105	32.36	31.88	0.014	0.029	-0.014
10pt_v106	38.48	33.56	0.012	0.031	-0.019
10pt_v107	37.91	33.87	0.014	0.024	-0.010
10pt_v108	40.21	34.97	0.024	0.069	-0.045
10pt_v109	39.30	34.77	0.021	0.036	-0.014
10pt_v110	34.25	31.74	0.021	0.029	-0.007
10pt_v111	34.35	32.19	0.038	0.029	0.009
10pt_v112	30.44	30.89	0.014	0.026	-0.012
10pt_v113	31.39	30.95	0.040	0.048	-0.007

* indicates question with significant difference between variance of clean and fraudulent matched interviews (.05 level with Bonferroni correction).

Table A3. Question Wording

Variable	Question Wording
Support for Democracy	<i>Changing the subject again, 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? Strongly disagree-strongly agree, 7-point scale.</i>
Vote for Maduro 2013	<i>Who did you vote for in the last presidential election of 2013?</i>
Non-Voter in 2013	<i>Did you vote in the last presidential elections of 2013?</i>
Center	<i>On this card there is a 1-10 scale that goes from left to right. The number one means left and 10 means right. Nowadays, when we speak of political leanings, we talk of those on the left and those on the right. In other words, some people sympathize more with the left and others with the right. According to the meaning that the terms “left” and “right” have for you, and thinking of your own political leanings, where would you place yourself on this scale? Tell me the number. Center = 4, 5, 6, or 7.</i>
Right	<i>On this card there is a 1-10 scale that goes from left to right. The number one means left and 10 means right. Nowadays, when we speak of political leanings, we talk of those on the left and those on the right. In other words, some people sympathize more with the left and others with the right. According to the meaning that the terms “left” and “right” have for you, and thinking of your own political leanings, where would you place yourself on this scale? Tell me the number. Right = 8, 9, or 10.</i>
DK/NA Ideology	<i>On this card there is a 1-10 scale that goes from left to right. The number one means left and 10 means right. Nowadays, when we speak of political leanings, we talk of those on the left and those on the right. In other words, some people sympathize more with the left and others with the right. According to the meaning that the terms “left” and “right” have for you, and thinking of your own political leanings, where would you place yourself on this scale? Tell me the number. Don’t know or No answer.</i>
Neighborhood Insecurity	<i>Speaking of the neighborhood where you live and thinking of the possibility of being assaulted or robbed, do you feel very safe, somewhat safe, somewhat unsafe or very unsafe?</i>
Crime Victimization	<i>Now, changing the subject, have you been a victim of any type of crime in the past 12 months? That is, have you been a victim of robbery,</i>

	<i>burglary, assault, fraud, blackmail, extortion, violent threats or any other type of crime in the past 12 months? Yes, No.</i>
Trust in National Police	<i>To what extent do you trust the National Police? Not at all – A lot, 7-point scale.</i>
National Economy Evaluation	<i>Do you think that the country's current economic situation is better than, the same as or worse than it was 12 months ago?</i>
Personal Economy Evaluation	<i>Do you think that your economic situation is better than, the same as, or worse than it was 12 months ago?</i>
Skin Tone	<i>[Interviewer instructions] When the interview is complete, WITHOUT asking, please use the color chart and circle the number that most closely corresponds to the color of the face of the respondent.</i>
Political Tolerance	<i>There are people who only say bad things about the (country) form of government, not just the current (incumbent) government but the system of government. How strongly do you approve or disapprove that such people be allowed to conduct peaceful demonstrations in order to express their views? Please read me the number. 1-10 scale.</i>
Perception of Threat	<i>To what extent do you consider the policies of the opposition threaten the wellbeing of the country? A lot, some, little, or none?</i>
Political Efficacy	Additive index of two items. <i>Now I am going to read some statements. Please tell me to what extent you agree or disagree with the following statements. 1-10 scale.</i> <i>Those who govern this country are interested in what people like you think. How much do you agree or disagree with this statement?</i> <i>You feel that you understand the most important political issues of this country. How much do you agree or disagree with this statement?</i>
Level of Education	<i>How many years of schooling have you completed? Responses sorted into the following three categories: none or primary, secondary, post-secondary.</i>
Age	<i>In what year were you born? Responses are then sorted into the following age cohorts: 18-25, 26-35, 36-45, 46-55, 56-65, and 65+.</i>

Table A4. Regression Models of Support for Democracy

	Clean	Compromised
Respect for Political Institutions	0.0543 (0.187)	0.111 (0.205)
Presidential Approval	0.690** (0.247)	0.611* (0.242)
Vote for Maduro 2013	-0.272+ (0.143)	-0.149 (0.142)
Non-Voter in 2013	-0.194 (0.169)	0.0393 (0.164)
Center	0.333* (0.154)	0.169 (0.149)
Right	0.538** (0.169)	0.654** (0.178)
DK/NA Ideology	0.0392 (0.215)	0.234 (0.251)
Neighborhood Insecurity	0.144 (0.169)	0.143 (0.186)
Crime Victimization	-0.149 (0.147)	-0.198 (0.131)
Trust in National Police	0.370+ (0.221)	0.128 (0.223)
National Economy Evaluation	-0.108 (0.310)	0.0265 (0.320)
Personal Economy Evaluation	0.159 (0.207)	0.269 (0.208)
Woman	-0.231* (0.105)	-0.238* (0.100)
Age	1.635** (0.193)	1.210** (0.183)
Level of Education	0.760** (0.171)	0.607** (0.191)
Skin Tone	-0.00242 (0.332)	0.389 (0.305)
Constant	3.358** (0.313)	3.436** (0.307)
Observations	1,193	1,226

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table A5. *Chow Tests for Support for Democracy Models*

Variable	F-statistic	Prob > F
Respect for political institutions	0.16	0.692
Presidential approval	0.21	0.646
Vote for Maduro 2013	1.42	0.238
Non-voter in 2013	2.97	0.089
Center	1.14	0.289
Right	0.60	0.442
DK/NA Ideology	1.01	0.319
Neighborhood insecurity	0.00	0.991
Crime victimization	0.16	0.690
Trust in national police	2.04	0.158
National economy evaluation	0.57	0.451
Personal economy evaluation	0.51	0.479
Woman	0.01	0.919
Age	6.61	0.012
Level of education	1.02	0.316
Skin tone	1.80	0.184

Note: No coefficient has significant differences between clean and compromised regressions when taking into account the Bonferroni correction for multiple tests.

Table A6. *Regression Models of Political Tolerance*

Variable	Clean	Compromised
Perception of Threat	-0.674* (0.260)	-0.828** (0.238)
Political Efficacy	1.677** (0.397)	2.021** (0.406)
Level of Education	0.680** (0.227)	0.419 (0.267)
Woman	-0.186 (0.175)	-0.0515 (0.160)
Age	-0.0937 (0.261)	0.190 (0.263)
Constant	6.370** (0.341)	6.004** (0.372)
Observations	1337	1330

Standard errors in parentheses

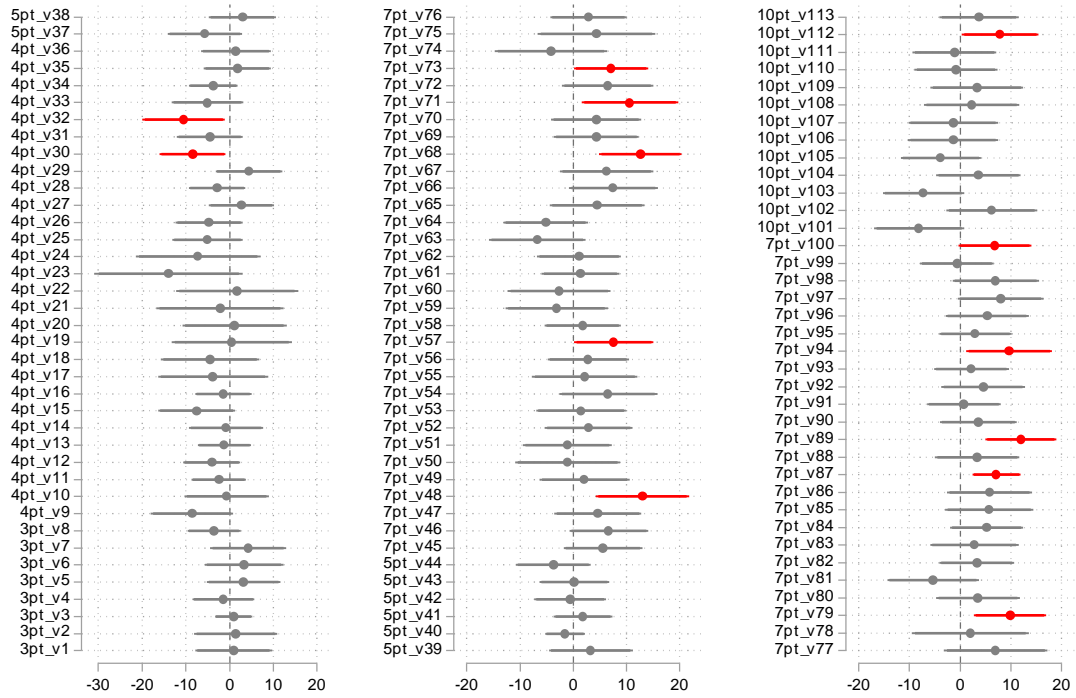
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table A7. *Chow Tests for Political Tolerance Models*

Variable	F-statistic	Prob > F
Perception of threat	0.94	0.334
Political efficacy	1.79	0.185
Level of education	1.80	0.184
Woman	1.28	0.261
Age	2.12	0.149

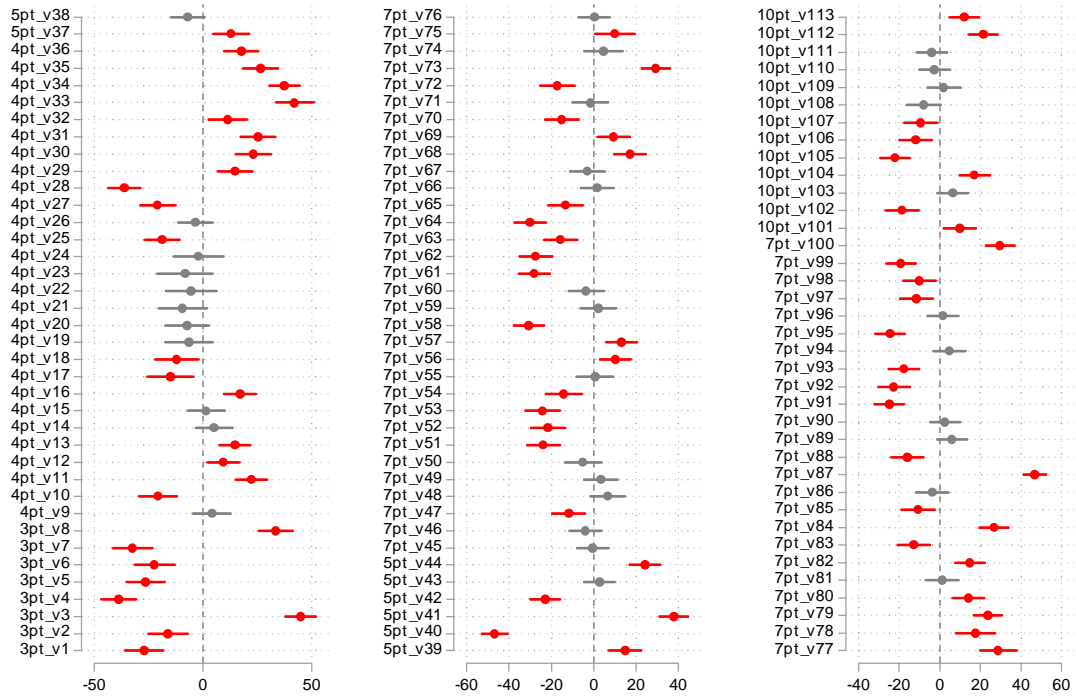
Note: No coefficient has significant differences between clean and compromised regressions when taking into account the Bonferroni correction for multiple tests.

Figure A1. Mean Differences from Fraudulent Responses



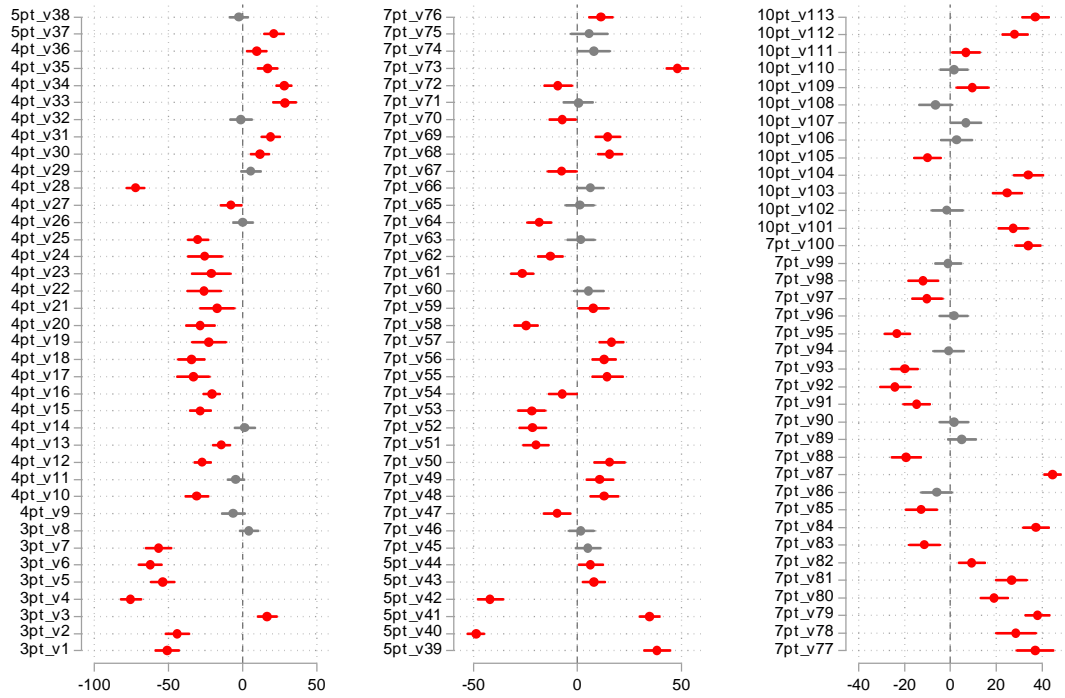
Notes: Values represent the results of difference of means tests for each of the 113 scale items in the survey. Points represent differences (clean data mean minus fraudulent data mean) and bars represent 95% confidence intervals with Bonferroni correction for multiple tests. All responses were rescaled to range from zero to a hundred.

Figure A2. Mean Differences from Random Responding



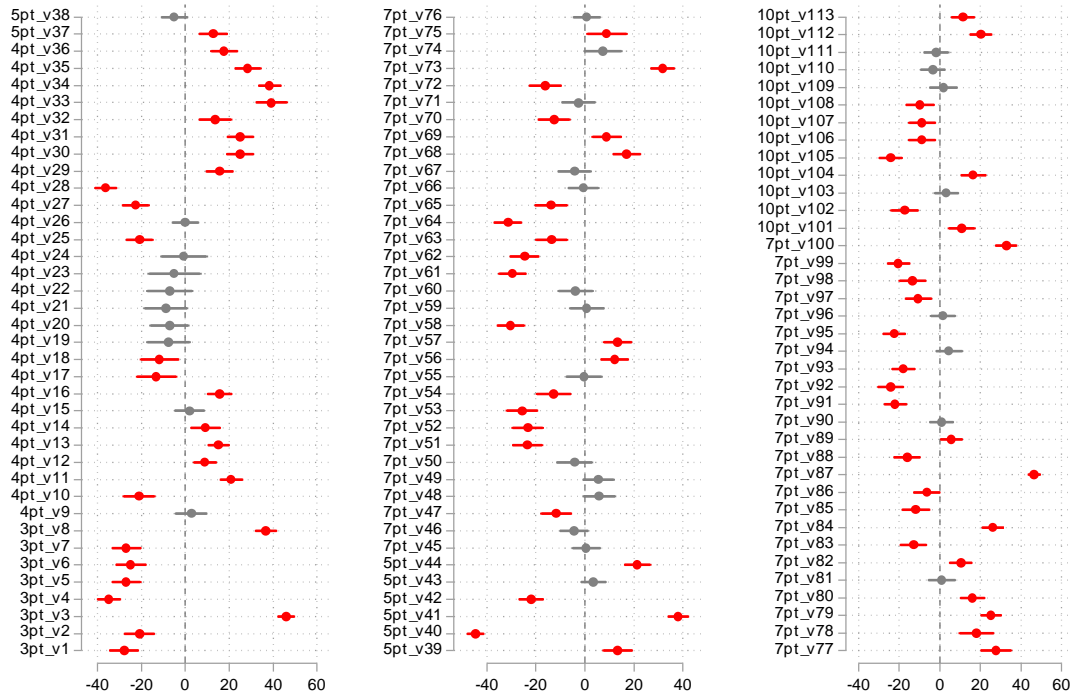
Notes: Values represent the results of difference of means tests for each of the 113 scale items in the survey. Points represent differences (clean data mean minus fraudulent data mean) and bars represent 95% confidence intervals with Bonferroni correction for multiple tests. All responses were rescaled to range from zero to a hundred.

Figure A3. Mean Differences from Speed Responding



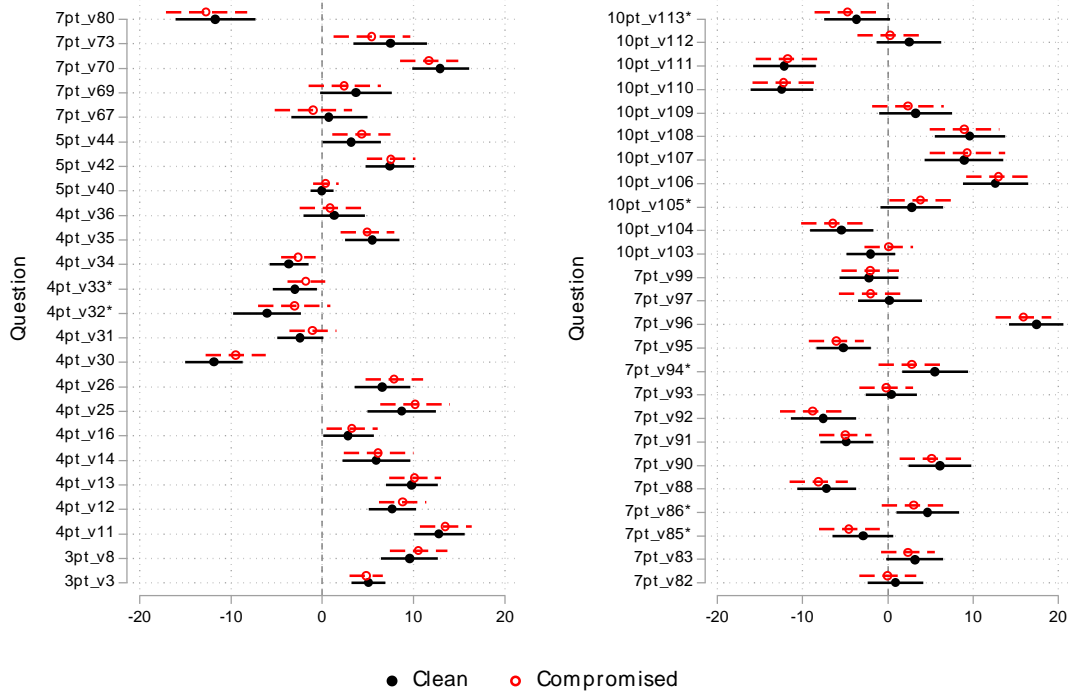
Notes: Values represent the results of difference of means tests for each of the 113 scale items in the survey. Points represent differences (clean data mean minus fraudulent data mean) and bars represent 95% confidence intervals with Bonferroni correction for multiple tests. All responses were rescaled to range from zero to a hundred.

Figure A4. Mean Differences from Middle Responding



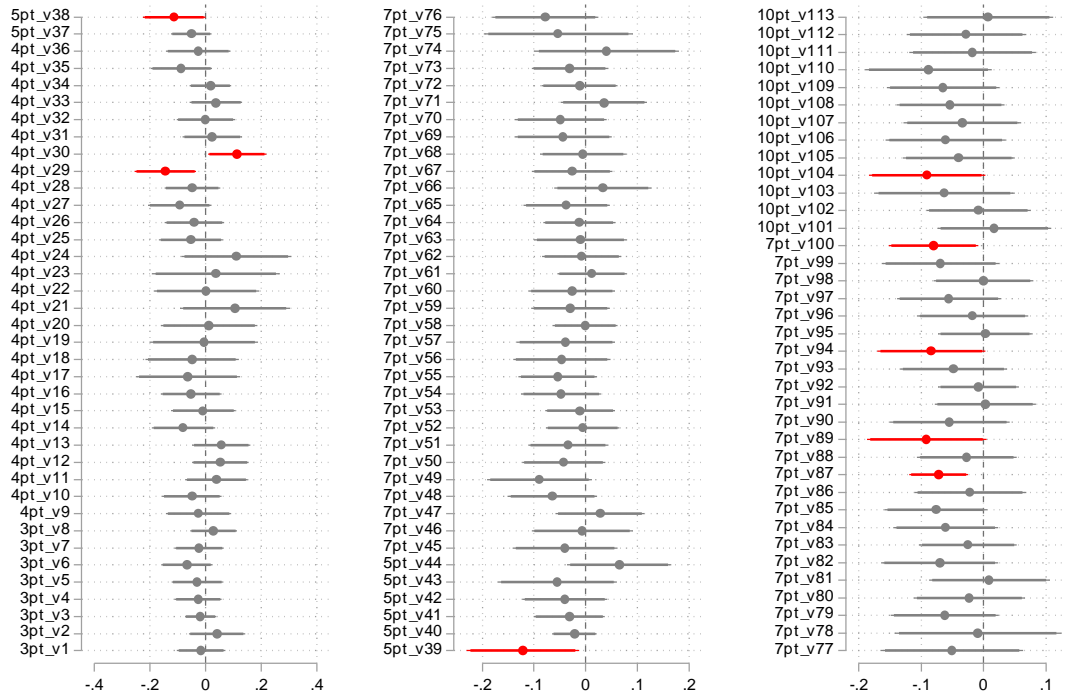
Notes: Values represent the results of difference of means tests for each of the 113 scale items in the survey. Points represent differences (clean data mean minus fraudulent data mean) and bars represent 95% confidence intervals with Bonferroni correction for multiple tests. All responses were rescaled to range from zero to a hundred.

Figure A5. Differences in Analyses of Change over Time



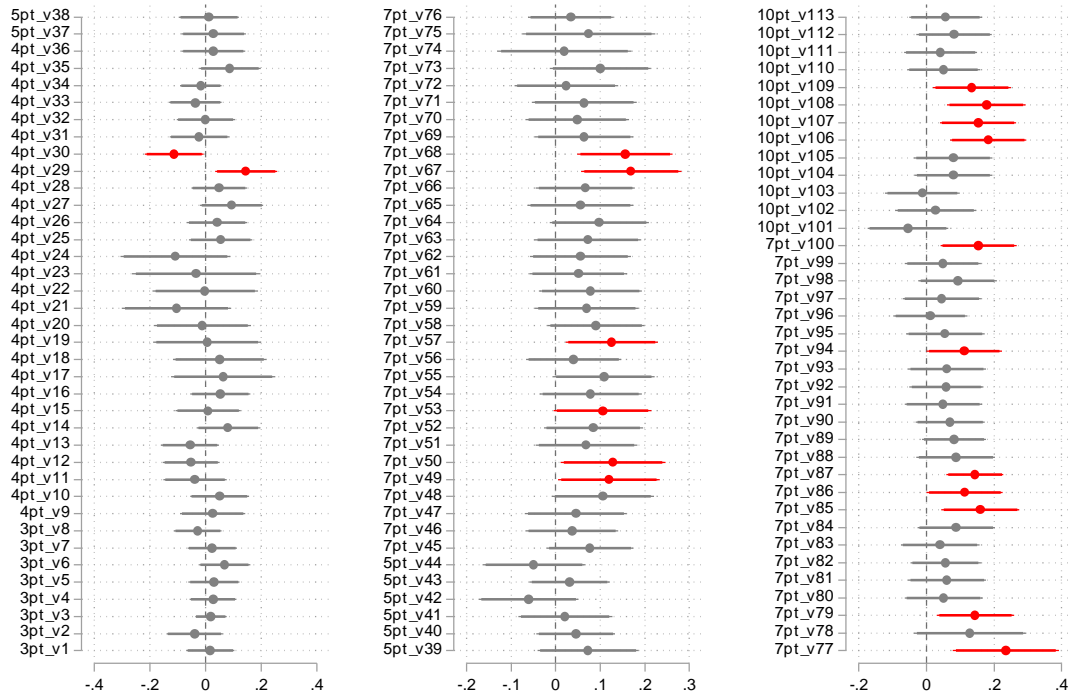
Notes: Each point represents the difference between the 2016/17 mean and the 2014 mean (positive values indicating increases over time and negative values indicating decreases over time) bars represent 95% confidence intervals with Bonferroni correction for multiple tests. All items rescaled to range from 0 to 100 for this analysis. Items with labels followed by an asterisk (*) are cases where conclusion of trend analyses differ when using compromised or clean data.

Figure A6. Difference in Proportion Selecting Center of Scale



Notes: The figure shows the results of difference of proportion test for each of the 113 scale items in the survey. Points represent differences (clean data proportion minus fraudulent data proportion) and bars represent 95% confidence intervals with Bonferroni correction for multiple tests. Of the 113 scale items, 8 (7.1%) were significantly closer to the midpoint in the fabricated data than in the clean data.

Figure A7. Difference in Proportion Selecting Extremes



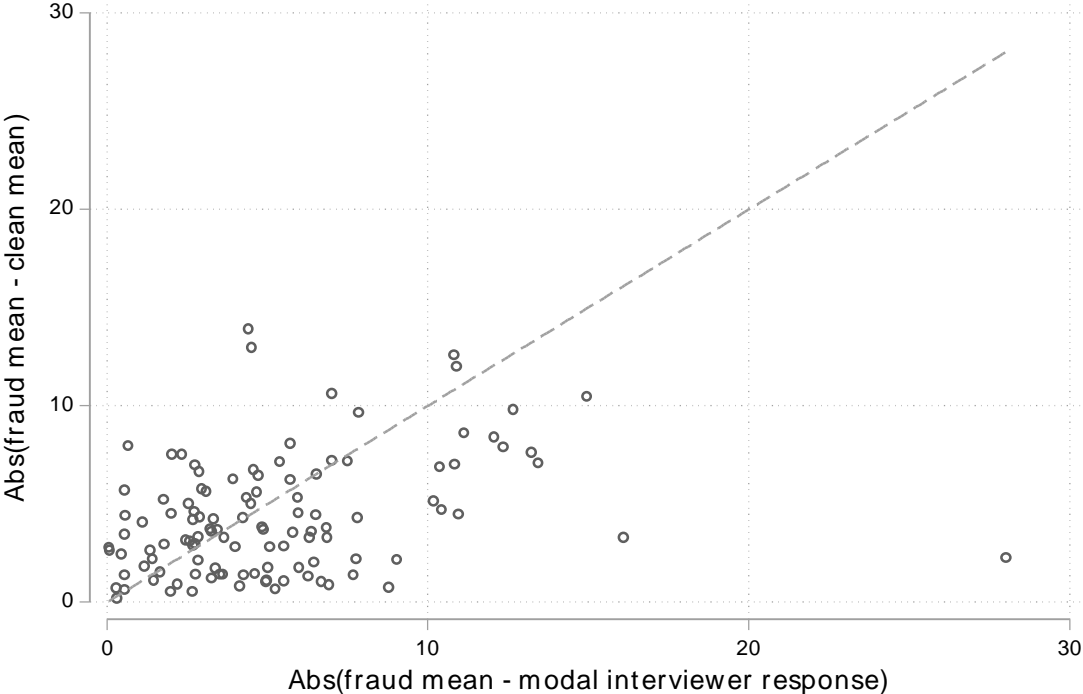
Notes: The figure shows the results of difference of proportion test for each of the 113 scale items in the survey. Points represent differences (clean data proportion minus fraudulent data proportion) and bars represent 95% confidence intervals with Bonferroni correction for multiple tests. In 18 of these 113 items (15.9%), faked responses were significantly less likely to be the minimum or maximum values on the scale.

D. Investigating Interviewer Bias in Fraudulent Responses

Another potential pattern in the generation of fraudulent survey responses is interviewer bias, with fabricators projecting their own preferences or beliefs onto their supposed respondents. To investigate this possibility, we compared the differences between mean responses of fraudulent interviews, matched clean interviews and the modal interviewer response. The measure of modal interviewer response was generated by running a regression of the item of interest on a set of demographic variables using the full clean dataset. From this regression, we generated a predicted value based on the profile of a modal interviewer on the fieldwork team, as described by the local survey firm. In the Venezuela survey, this means a young woman (ages 26 to 35) with higher education from a large city and with a skin tone of 4 on the LAPOP skin color palette.

To the extent that fraudulent responses are a reflection of interviewers' own biases, we expect to see smaller differences between the modal interviewer response and the fraudulent data, compared to the difference between the modal interviewer and clean mean response. Figure A8 plots the absolute difference between the mean of fraudulent responses and the mean of matched genuine responses on the y-axis and the absolute difference between the fraudulent mean and the modal interview response on the x-axis. For this analysis, all items have been rescaled to range from 0 to 100. Values on the dotted 45-degree line would mean that the differences are of equal value. Observations below the 45-degree line denote that the fraudulent data are more similar to the modal interviewer response than the clean data. In a t-test between the absolute differences for all 113 items, we find that the differences with the modal interviewer response (mean=5.2) are significantly greater than those with the clean mean (mean=4.3) at the 0.05 level.

Figure A8. Comparing Modal Interviewer Response to Fraudulent and Clean Responses



E. Comparing Fieldwork Patterns in Clean and Fraudulent Data

The figures below provide descriptive information about fraudulent data. Fraudulent data are more likely to be observed on Sundays, Mondays, and Tuesdays than later in the week (Figure A9). We also observe a spike in fraudulent interviews in mid-December, shortly before the Christmas holiday, when fieldwork was initially slated to end (Figure A10). In 2017, following LAPOP’s redoubled auditing efforts, we observe almost no fraudulent interviews.

Figure A9. *Distribution of Interviews by Day of Week*

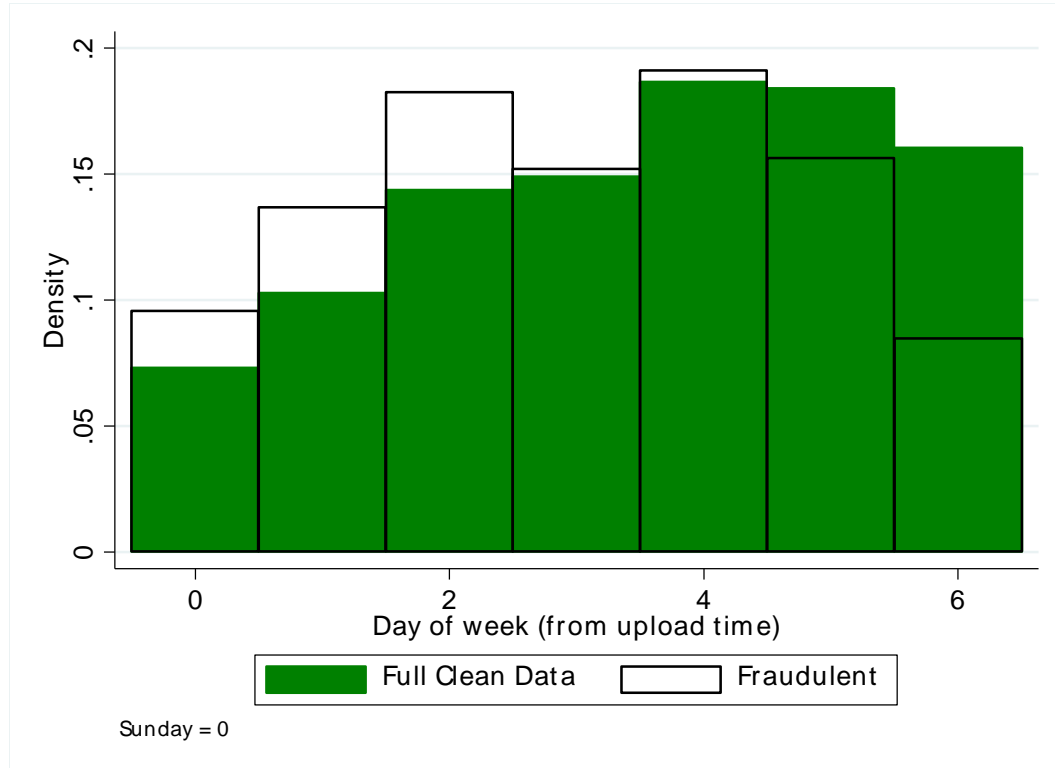


Figure A10. *Distribution of Interviews over Fieldwork Period*

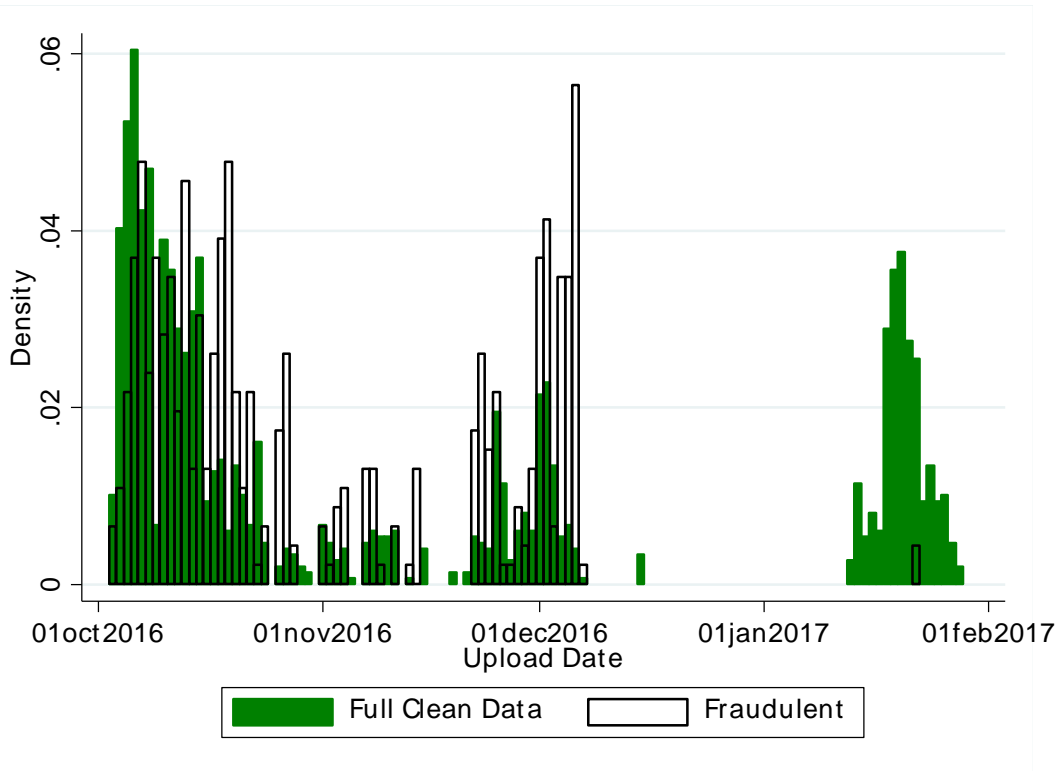
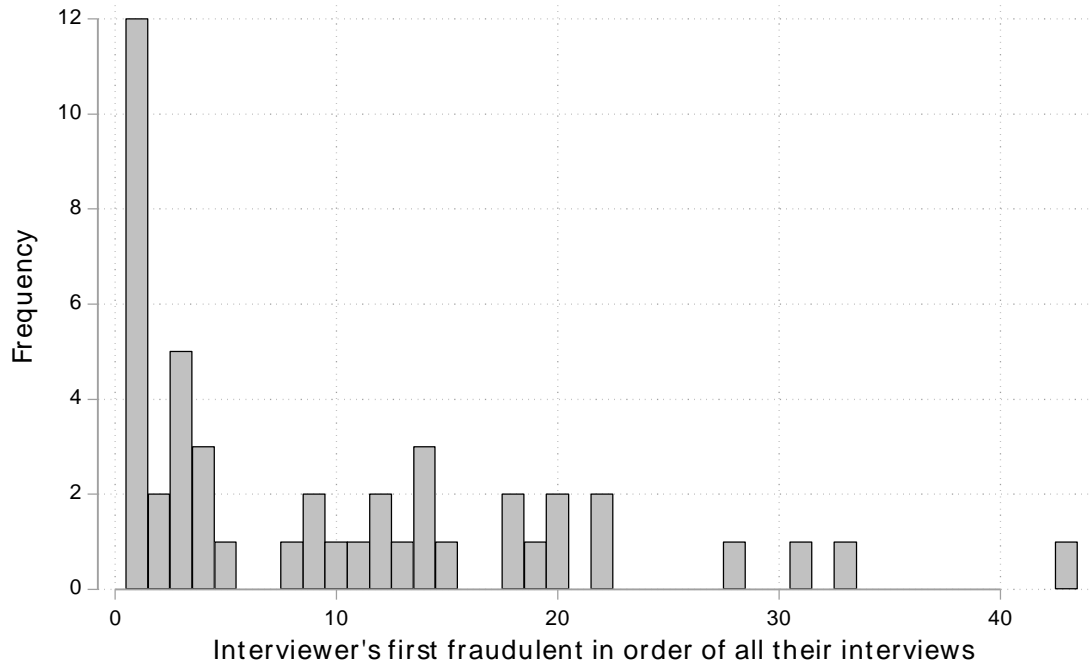


Figure A11 below shows the distribution of fraudulent interviews with respect to enumerators' experience working on the study. Of the 46 enumerators who eventually conducted a fraudulent interview, 12 did so in their first interview on the project, and 23 enumerators conducted a fraudulent interview as one of their first five interviews. Half of all interviewers who eventually recorded a fake interview did so only after conducting eight or more interviews.

Figure A11. *When Enumerators Fabricate First Interview*



Note: 33 out of 79 interviewers had no fraudulent interviews.

F. Replication Using Peru 2017 Survey

We replicated our analysis using data from Peru, the only other survey in the 2016/17 AmericasBarometer with sufficient fraudulent interviews to conduct matching analysis with confidence. As with the Venezuela data, we paired interviews identified as fraudulent (116 interviews in the Peru study) with genuine interviews based on an exact match on sex, age group, and primary sampling unit. This resulted in 116 fraudulent interviews matched with 116 genuine interviews. The table below summarizes the comparison of fraudulent interviews and their matched clean interviews in the Peru data along with the original figures for Venezuela from Table 1 of the paper. Overall, the conclusions are the same: few if any differences. No items show statistically significant differences in terms of means or item non-response between the fraudulent and clean interviews. Only 1.4% of the 72 scale items in the survey register significant differences in terms of the variances. The average magnitudes of the differences (in means) are very similar. It should be noted that the Peru survey has considerably less fraudulent interviews than the Venezuela survey.

Table A8. *Item-Level Effects of Fabricated Data*

Comparison	Venezuela	Peru
Difference in means	11.5%	0.0%
Average magnitude (in SD)	0.13	0.13
Difference in variances	8.9%	1.4%
Item nonresponse	0.0%	0.0%

Notes: Values indicate the results of tests of 113 scale items for Venezuela and 72 scale items for Peru, comparing the fabricated interviews (n=420 for VE, n=116 for PE) and the matched real data (n=420 for VE, n=116 for PE). Tests of significance use Bonferroni correction to account for multiple tests.

References

Blackwell, Matthew, Stefano Iacus, and Gary King. 2009. "cem: Coarsened Exact Matching in Stata." *Stata Journal* 9 (4): 524-546.