

Current Innovations in Probability-based Household Internet Panel Research (CIPHER)

February 26, 2020

# Comparing Probability Panels to Opt-in Sources in Security and Data Quality

**Courtney Kennedy**

*Director of Survey Research*

# Acknowledgments

- Co-authors: Nick Hatley, Arnold Lau, Andrew Mercer, Scott Keeter, Joshua Ferno, and Dorene Asare-Marfo
- Thank you to Pew Research Center leadership for supporting this project

## Why we did this study

**Over 80% of U.S. public opinion polls are conducted with online opt-in samples**

### How Popular Is Donald Trump?

UPDATED 2 HOURS AGO



Source: FiveThirtyEight.com

## Why we did this study

**What does the influx of opt-in sourcing mean for the quality and security of U.S. polling?**

**While some survey issues are ever present (e.g., satisficing, nonresponse), the ability of people to self-enroll into the process may introduce new potential risks:**

- Fraudulent interviews from outside the U.S. (Ahler et. al 2019; TurkPrime 2018)
- Survey bots (Fors Marsh 2018; Versta Research 2017; Accelerant Research 2019)
- Click farms (Pasternak 2019)
- Trolling (Lopez and Hillygus 2018)

## Why we did this study

**Vendors of online opt-in surveys are aware of these threats and often have measures in place to confront them. These can include:**

- Screening out foreign IP addresses
- Digital fingerprinting to prevent duplicate interviews
- IP proxy testing
- Validation of email addresses
- Machine learning to flag suspicious patterns

## Research questions

- **How effective are efforts to block bogus respondents?**
- **Are there bogus respondents in online surveys?**
- **If so, what is the incidence level?**
- **If so, do they just add noise or are they systematically different?**
- **If so, do common data quality checks identify them?**

## Research design

- We fielded a large study in spring 2019 administering a common questionnaire on six different online survey sources:
  - 3 widely used opt-in survey panels (or “marketplaces”)
  - 1 crowdsourcing platform
  - 2 online probability-based panels (both primarily ABS recruited)
- We obtained at least 10,000 completed interviews with each source

	Crowd sourced	Online Opt-in			ABS Recruited	
		Panel 1	Panel 2	Panel 3	Panel 1	Panel 2
<b>Interviews</b>	10,879	10,002	10,000	11,054	10,178	10,526

## Research design

- Target population: U.S. adults age 18 and older
- Survey administered in English and Spanish
- Opt-in polls fielded “using formatting, style, and respondent sourcing that is normal for political polls” conducted on each panel
- Analysis presented here is all unweighted. Main findings are generally unaffected when weighting (raking on 11 demographics) is applied

## Questionnaire began as a routine opinion poll

(Q1) **Would you say things in this country today are...**

- Generally headed in the right direction
- Off on the wrong track

(Q2) **Do you approve or disapprove of the job Donald Trump is doing as President?**

- Strongly approve
- Somewhat approve
- Somewhat disapprove
- Strongly disapprove

- The first half of the survey was designed to look like a national opinion poll. We modeled off high-profile online opt-in opinion polls.

## Questionnaire asked six open ended questions

How would you say you are feeling today?

When you were growing up, what was the big city nearest where you lived?

When you visit a new city, what kinds of activities do you like to do?

In retirement what skill would you most like to learn?

How do you decide when your computer is too old and it's time to purchase a new one?

What would you like to see elected leaders in Washington get done during the next few years? Please give as much detail as you can.

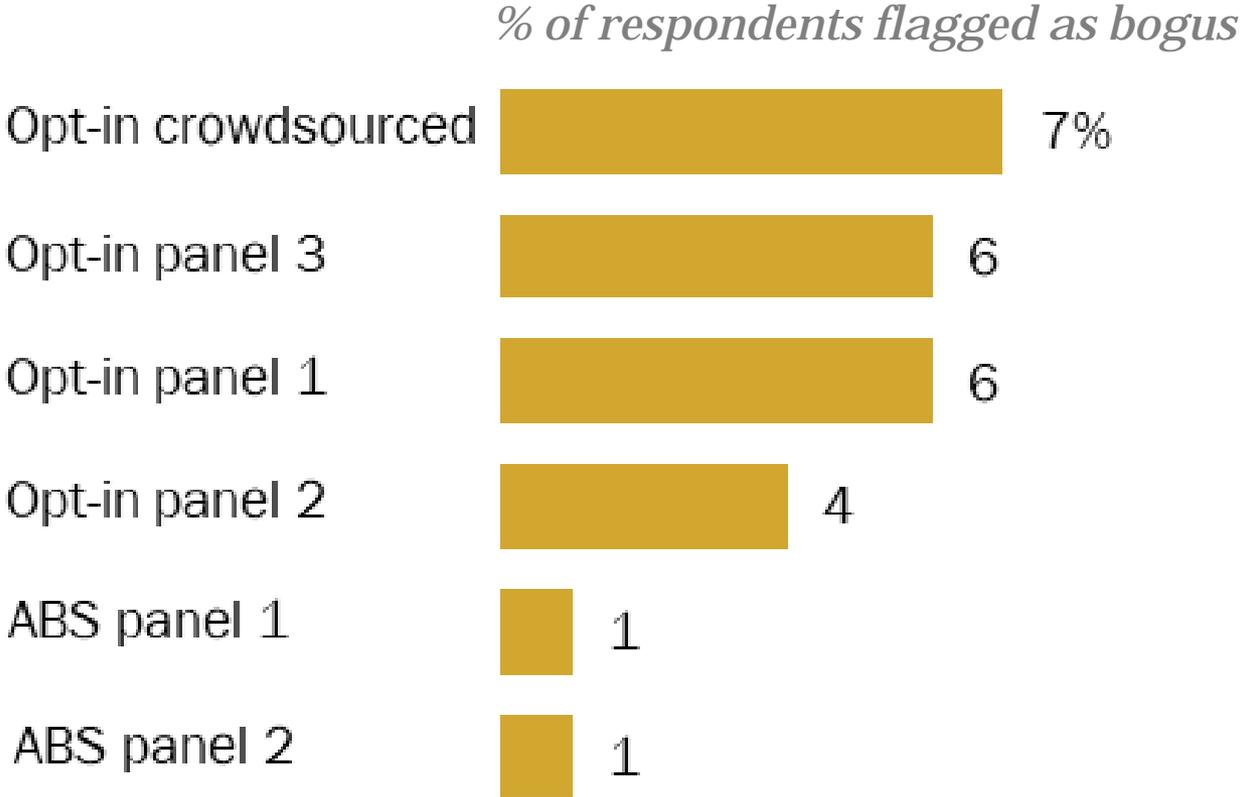
# Coding open-ended answers

- **We manually coded the over 375,000 open-ended answer into one of four categories:**
  - Responsive answer
  - Non sequitur (answer doesn't fit the question)
  - Gibberish (e.g., "Dffvdgggugcfhdggv Jr ffv")
  - Blank/Don't know/Refused (e.g., "Unsure" "idk")
- **Coding principles**
  - i. Disregard spelling and grammatical errors
  - ii. Employ a generous, permissive definition of answers that are responsive to the question

# Research Design

- Created a definition for a “bogus” based on behaviors that are fairly egregious
- Cases were flagged as bogus if they did one or more of the following:
  - Gave multiple non sequitur open-ended answers
  - Took the survey multiple times based on open-ended answers
  - Self-reported living outside the U.S. (the target pop is U.S. adults)
  - Always say approve/favorable regardless of what was asked

# About 1-in-20 opt-in panel cases were flagged as bogus



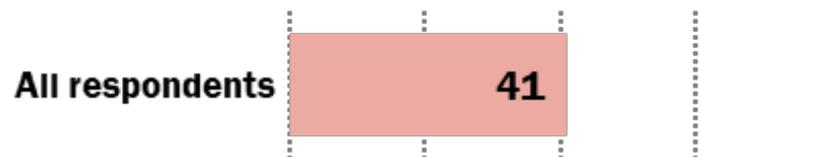
## Flagged behaviors were rare but concentrated in the opt-in sources

	% gave multiple non sequiturs	% approve of everything	Reported living outside the U.S.	Took survey multiple times
Crowdsourced	4	4	1	1
Opt-in panel 3	3	3	1	0
Opt-in panel 1	4	3	2	0
Opt-in panel 2	2	2	1	0
ABS panel 1	0	0	0	0
ABS panel 2	0	0	0	0

# **Do bogus respondents add bias or just noise?**

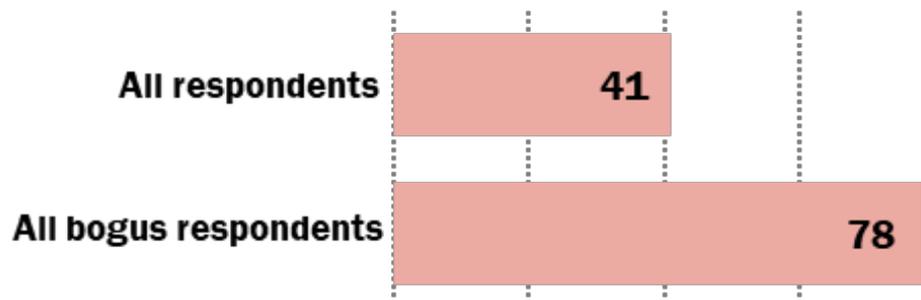
# Bogus respondents are prone giving positive answers

*% Approve of Trump's job performance*



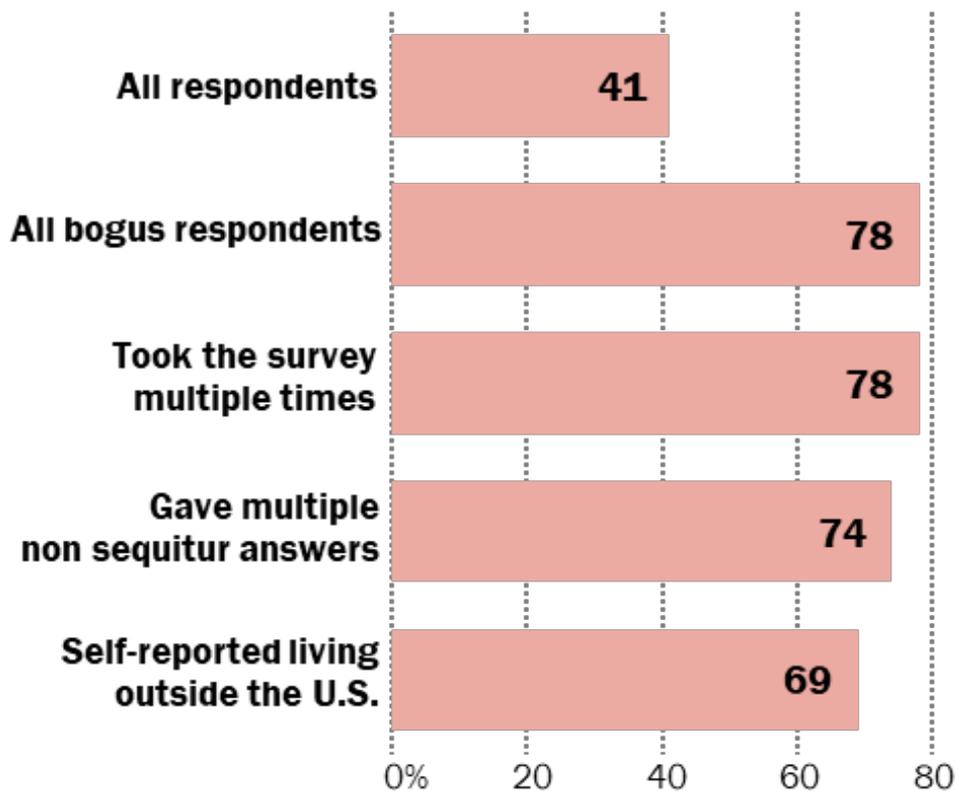
# Bogus respondents are prone giving positive answers

*% Approve of Trump's job performance*



# Bogus respondents are prone giving positive answers

*% Approve of Trump's job performance*



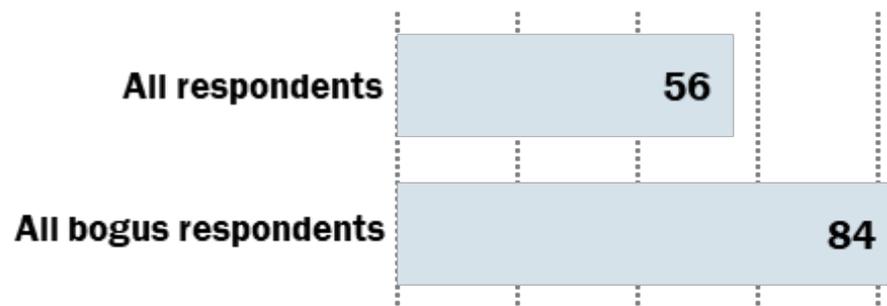
# Bogus respondents are prone giving positive answers

*% Approve of the health care law  
passed by Barack Obama and Congress in 2010*



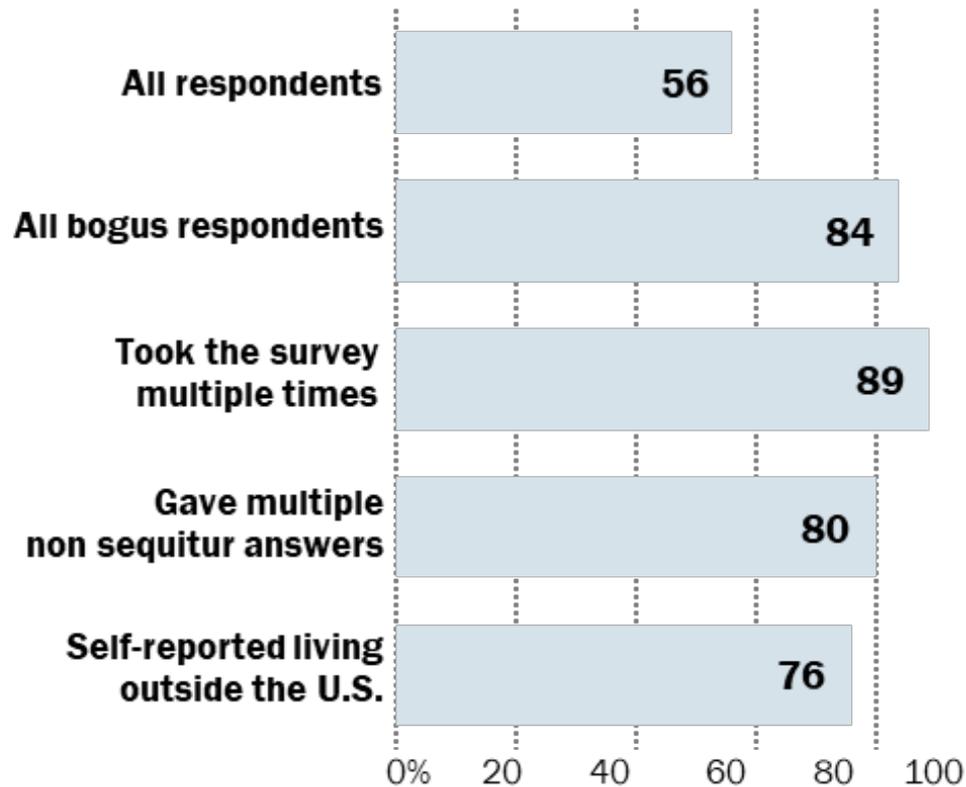
# Bogus respondents are prone giving positive answers

*% Approve of the health care law  
passed by Barack Obama and Congress in 2010*



# Bogus respondents are prone giving positive answers

*% Approve of the health care law passed by Barack Obama and Congress in 2010*

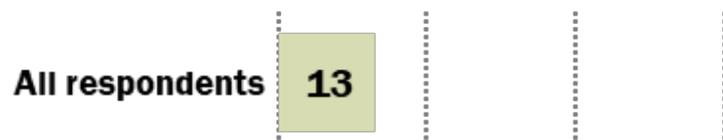


## **Bogus respondents are prone giving positive answers**

*% favorable overall opinion of Russian President Vladimir Putin*

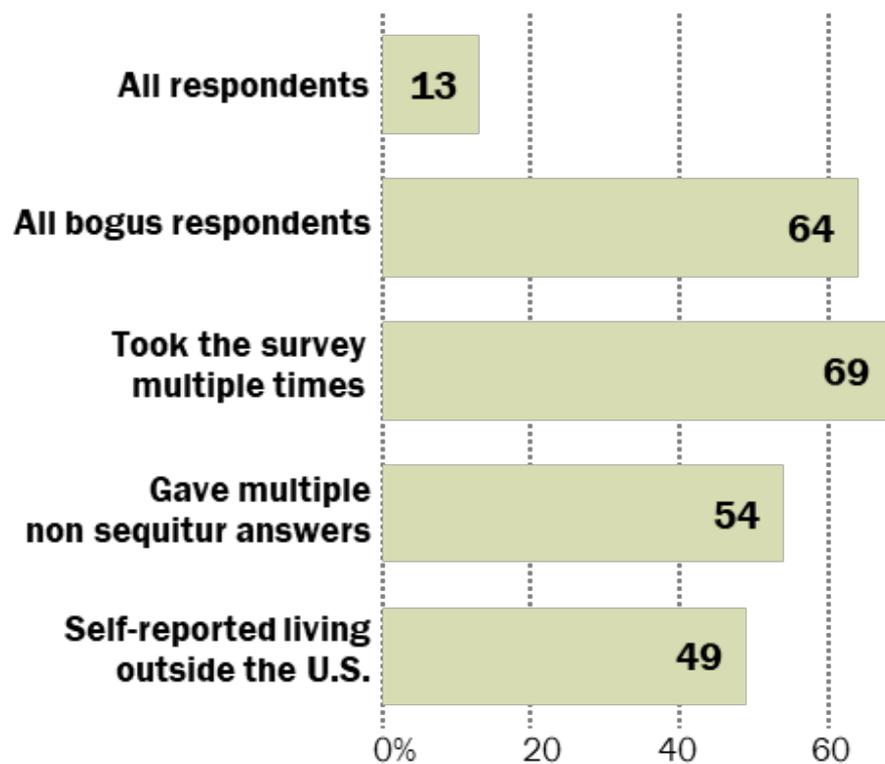
# Bogus respondents are prone giving positive answers

*% favorable overall opinion of Russian President Vladimir Putin*



# Bogus respondents are prone giving positive answers

*% favorable overall opinion of Russian President Vladimir Putin*



**Does this actually affect topline estimates?**

## Bogus cases can slightly inflate approval-type estimates

*% favorable overall opinion of Russian President Vladimir Putin*

	All interviews
Crowdsourced	16
Opt-in panel 3	15
Opt-in panel 2	13
Opt-in panel 1	13
ABS panel 1	10
ABS panel 2	9

## Bogus cases can slightly inflate approval-type estimates

*% favorable overall opinion of Russian President Vladimir Putin*

	All interviews	Bogus cases excluded	<i>Diff</i>
Crowdsourced	16	11	-5
Opt-in panel 3	15	12	-3
Opt-in panel 2	13	10	-3
Opt-in panel 1	13	12	-1
ABS panel 1	10	10	0
ABS panel 2	9	9	0

# **Bogus cases can slightly inflate approval-type estimates**

*% saying things in the country are headed in the right direction*

# Bogus cases can slightly inflate approval-type estimates

*% saying things in the country are headed in the right direction*

	All interviews
Crowdsourced	37
Opt-in panel 3	38
Opt-in panel 2	37
Opt-in panel 1	39
ABS panel 1	30
ABS panel 2	31

## Bogus cases can slightly inflate approval-type estimates

*% saying things in the country are headed in the right direction*

	All interviews	Bogus cases excluded	<i>Diff</i>
Crowdsourced	37	33	-4
Opt-in panel 3	38	36	-2
Opt-in panel 2	37	35	-2
Opt-in panel 1	39	37	-2
ABS panel 1	30	30	0
ABS panel 2	31	31	0

## **Bogus cases can slightly inflate approval-type estimates**

*% favorable overall opinion of German Chancellor Angela Merkel*

## Bogus cases can slightly inflate approval-type estimates

*% favorable overall opinion of German Chancellor Angela Merkel*

	All interviews	Bogus cases excluded	<i>Diff</i>
Crowdsourced	54	52	-2
Opt-in panel 3	43	42	-1
Opt-in panel 2	46	45	-1
Opt-in panel 1	39	38	-1
ABS panel 1	48	48	0
ABS panel 2	52	52	0

## Bogus cases can slightly inflate approval-type estimates

*% approve of Trump's performance as president*

	All interviews	Bogus cases excluded	<i>Diff</i>
Crowdsourced	35	31	-4
Opt-in panel 3	42	40	-2
Opt-in panel 2	42	41	-1
Opt-in panel 1	43	41	-2
ABS panel 1	43	43	0
ABS panel 2	39	39	0

# Other questions appear unaffected by bogus cases

## Other questions appear unaffected by bogus cases

Which do you think should be a higher priority right now?

- Enacting new laws to try to reduce gun violence
- Protecting the right to own guns

## Other questions appear unaffected by bogus cases

*% Higher priority is protecting the right to own guns*

	All interviews	Bogus cases excluded	<i>Diff</i>
Crowdsourced	30	31	+1
Opt-in panel 3	37	37	0
Opt-in panel 2	35	35	0
Opt-in panel 1	40	39	-1
ABS panel 1	36	36	0
ABS panel 2	34	34	0

## Other questions appear unaffected by bogus cases

Who do you trust more to handle each of the following issues?

- a. The economy
  - Republicans in Congress
  - Democrats in Congress

## Other questions appear unaffected by bogus cases

*% Trust Republicans in Congress (over Dems) on the economy*

	All interviews	Bogus cases excluded	<i>Diff</i>
Crowdsourced	45	44	-1
Opt-in panel 3	49	49	0
Opt-in panel 2	51	50	-1
Opt-in panel 1	50	50	0
ABS panel 1	50	50	0
ABS panel 2	47	47	0

# **Do common data quality checks detect bogus cases?**

# This study featured two common data quality checks

## Speeding check

Median survey time was 7 min. We flagged cases answering in under 3 min. (1.5% of all respondents)

## Attention check (or “trap”) question

Paying attention and reading the instructions carefully is critical. If you are paying attention, please choose Silver below. **[RANDOMIZE RESPONSES]**

- Red
- Yellow
- Blue
- Green
- Silver
- Orange

## Two common checks failed to detect most bogus respondents

*% of bogus respondents passing...*

<b>Bogus behavior</b>	<b>...Trap question</b>	<b>...Speeding test</b>	<b>...Both</b>
All bogus respondents	84	87	76

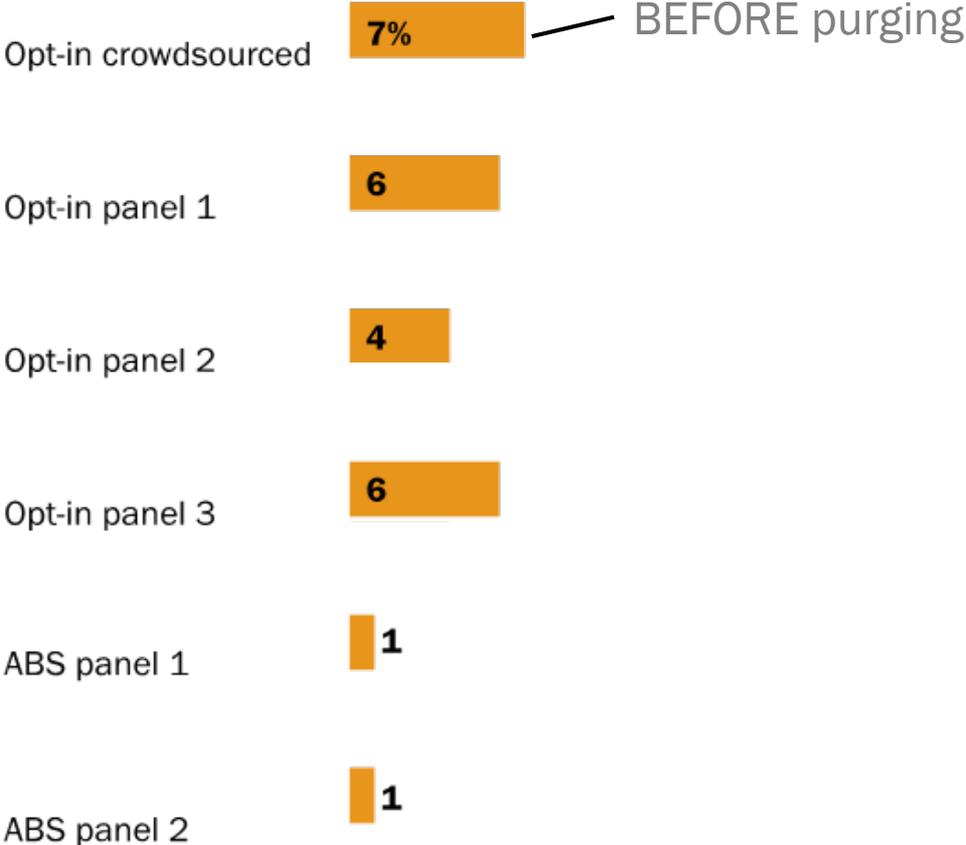
## Two common checks failed to detect most bogus respondents

*% of bogus respondents passing...*

<b>Bogus behavior</b>	<b>...Trap question</b>	<b>...Speeding test</b>	<b>...Both</b>
All bogus respondents	84	87	76
Approve/favor everything	93	93	88
Multiple non sequiturs	81	85	71
Reported living outside U.S.	64	75	52

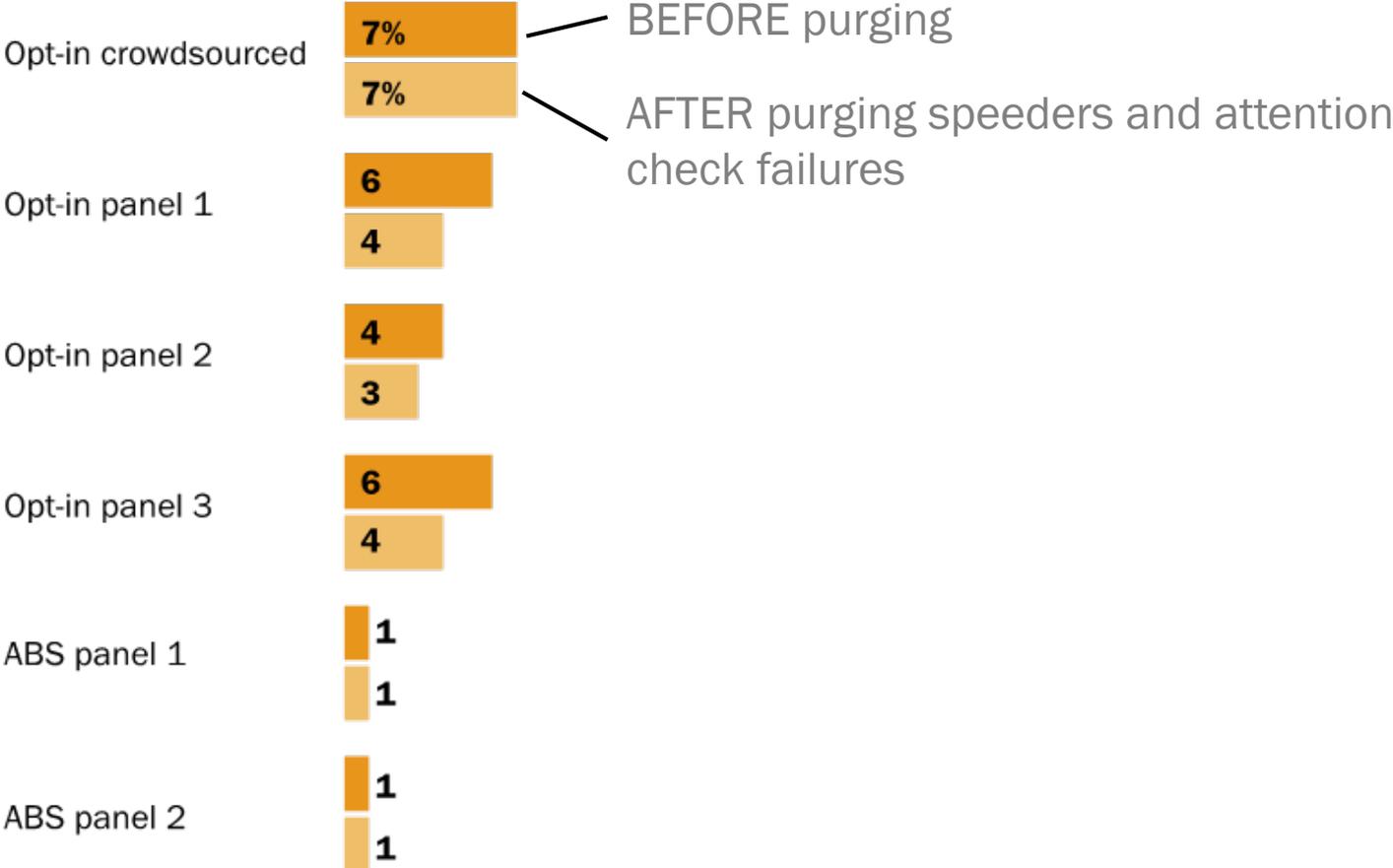
# After removing speeders and attention check failures, most bogus cases remain

*% of interviews identified as bogus*



# After removing speeders and attention check failures, most bogus cases remain

*% of interviews identified as bogus*



# Conclusions

- Crowdsourced and opt-in survey panel respondents were more likely than those recruited via random sampling of addresses to give bogus data
- Bogus respondents are prone to selecting positive answer choices, which can introduce a small upward bias in approval-type questions
- Two common data quality checks fail to detect most bogus respondents
- The crowdsourced sample had a unique set of issues (e.g., plagiarism, foreign IP addresses). After removing foreign IPs, bogus incidence still 4%.
- In absolute terms, the problems uncovered in this study are low incidence but have the potential to get worse in the near future.
- To help people determine whether the researcher performed data quality checks, these should be mentioned in methods statements.

**The report and micro-dataset are available on the Pew Research Center website**

# Thank you

**Courtney Kennedy**

*Director of Survey Research*

[ckennedy@pewresearch.org](mailto:ckennedy@pewresearch.org)



# Research Design

- Created a definition for a “bogus” based on behaviors that are fairly egregious
- Cases were flagged as bogus if they did one or more of the following:
  - Gave multiple non sequitur open-ended answers
  - Took the survey multiple times based on open-ended answers
  - Self-reported living outside the U.S. (the target pop is U.S. adults)
  - Always say approve/favorable regardless of what was asked
- Do you approve or disapprove of the job Donald Trump is doing as President?
- What is your overall opinion of French President Emmanuel Macron?
- What is your overall opinion of British Prime Minister Theresa May?
- What is your overall opinion of Russian President Vladimir Putin?
- What is your overall opinion of German Chancellor Angela Merkel?
- What is your overall opinion of U.S. President Donald Trump?
- Do you approve or disapprove of the health care law passed by Barack Obama and Congress in 2010?

# Research Design

- Created a definition for a “bogus” based on behaviors that are fairly egregious
- Cases were flagged as bogus if they did one or more of the following:
  - Gave multiple non sequitur open-ended answers
  - Took the survey multiple times based on open-ended answers
  - Self-reported living outside the U.S. (the target pop is U.S. adults)
  - Always say approve/favorable regardless of what was asked
- Do you approve or disapprove of the job Donald Trump is doing as President?
- What is your overall opinion of French President Emmanuel Macron?
- What is your overall opinion of British Prime Minister Theresa May?
- What is your overall opinion of Russian President Vladimir Putin?
- What is your overall opinion of German Chancellor Angela Merkel?
- What is your overall opinion of U.S. President Donald Trump?
- Do you approve or disapprove of the health care law passed by Barack Obama and Congress in 2010?

# Research Design

- Created a definition for a “bogus” based on behaviors that are fairly egregious
- Cases were flagged as bogus if they did one or more of the following:
  - Gave multiple non sequitur open-ended answers
  - Took the survey multiple times based on open-ended answers
  - Self-reported living outside the U.S. (the target pop is U.S. adults)
  - Always say approve/favorable regardless of what was asked

In real life, people who genuinely support Putin, Trump, Obamacare, Merkel and Macron are incredibly rare.

This behavior was strongly associated with other flags for bad data quality

## Approving-of-everything was considered bogus

- At first glance, giving uniformly positive answers may not seem like a bogus behavior
- But in this study that meant the respondent expressed a favorable opinion of Vladimir Putin, Emmanuel Macron, Donald Trump, Angela Merkel, Theresa May, and the 2010 Health Care Law (Obamacare)
- In the address-recruited panels, 0% gave that response pattern
- Approving of everything was strongly associated with bad data quality

<b>Data Quality Flag</b>	<b>Respondents who approve of everything</b>	<b>All study respondents</b>
Gave multiple non sequitur answers	40%	2%
Used foreign IP address	15%	1%
Self-reported living outside the U.S.	12%	1%
Took the survey more than once	7%	0%
Unweighted n	1,195	62,639

# Taking the survey multiple times based on open-ends

Question	Respondent 1	Respondent 2	Respondent 3	Respondent 4
<b>How would you say you are feeling today?</b>	Am feeling relaxed and more positive thoughts today, and also a excited day for me.	Am feeling relaxed and perfect minded today. Excited day for me	Am feeling relaxed and clean minded and also a excited day for me	Am feeling relaxed and clean minded today, also an excited day for me.
<b>When you were growing up, what was the big city nearest where you lived?</b>	Los Angeles	Los Angeles	Los Angeles	Los Angeles
<b>When you visit a new city, what kinds of activities do you like to do?</b>	I like to found the famous place of that city for visiting those scenarios .	I like to travel through that city and find the famous place of that city for visiting there.	I like to find out the famous place of the city and like to enjoy those beautiful scenario on that day	I like to find out the famous place of that city for visiting those scenarios.
<b>How do you decide when your computer is too old and it's time to purchase a new one?</b>	when computer starts to works slower and often hanging in the desktop leads to purchase a new one	It becomes slower in working and often hanging in the system leads to purchase a new one.	When computer started to work slower than its actual speed and also a often hanging leads to purchase a new one	When the computer starts to work slower than its actual speed and also a often hanging in the desktop leads to purchase me a new one
<b>In retirement what skill would you most like to learn?</b>	Renovating skills	Renovating skills	Renovating skills	Renovating skills

## Self-reporting living outside the U.S.

- This would not catch foreign respondents concealing their location
- There are several potential explanations for this response
  - They live outside the U.S. and aren't Americans
  - They live outside the U.S. and are Americans (e.g., overseas military)
  - Trolling
  - Answering without reading the question (extreme satisficing)
- 12% of these respondents had a foreign IP address

## More on that definition of “bogus”

Which of the following websites do you use at least once a month?

*[Check all that apply]*

Facebook

Instagram

YouTube

FizzyPress

Doromojo

We could have counted as bogus anyone who reported using a website that doesn't exist

## More on that definition of “bogus”

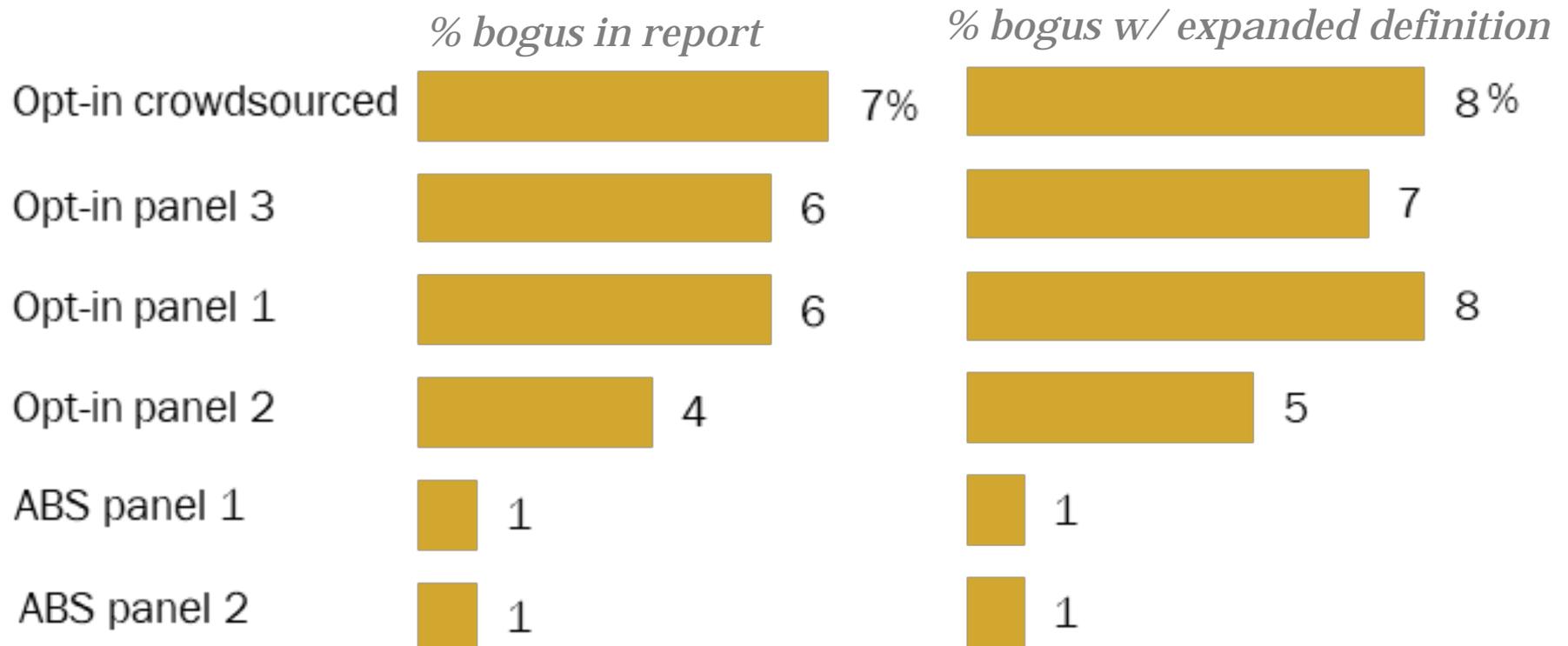
**Q: What would you like to see elected leaders in Washington get done during the next few years?**

Khlnohlhohlhonlholnoylnohljh j

We could have counted as bogus anyone who repeatedly answered with gibberish in the open-ended questions

## More on that definition of “bogus”

Had we counted “repeated gibberish” and “impossible website usage” as bogus, the average bogus incidence in opt-in panels would be 7% instead of 5%. The rate for the ABS is unchanged



## Summing up the open-ended data

- In all six samples most respondents were giving genuine answers
- That said, 2% to 4% of opt-in cases gave multiple non sequitur answers, which compares to 0% of address-recruited cases
- We found five different types of non sequitur answers
  - Plagiarism
  - Product reviews
  - Conversational text
  - Common words
  - Miscellaneous
- Almost all of the plagiarizers (97%) were in the crowdsourced sample
- Most of the conversational and product review type non sequiturs were in the opt-in panel samples