Political Analysis (2025), **0**, 1–19
doi:10.1017/pan.2025.10018

ARTICLE

Survey Professionalism: New Evidence from Web Browsing Data

Bernhard Clemm von Hohenberg[1], Tiago Ventura[2], Jonathan Nagler[3], Ericka Menchen-Trevino[4] and Magdalena Wojcieszak[5]

[1]Institute of Political Science, University of the Bundeswehr Munich, Munich, Germany; [2]McCourt School of Public Policy, Georgetown University, Washington, DC, USA; [3]Department of Politics, New York University, New York, NY, USA; [4]Behavioral Research Innovation Center, Chicago, IL, USA; [5]Department of Communication, UC Davis, Davis, CA, USA

Corresponding author: Bernhard Clemm von Hohenberg; Email: bernhard.clemm@gmail.com

(Received 31 October 2024; revised 16 July 2025; accepted 27 July 2025)

Abstract

Online panels have become an important resource for research in political science, but the compensation offered to panelists incentivizes them to become "survey professionals," raising concerns about data quality. We provide evidence on survey professionalism exploring three US samples of subjects who donated their browsing data, recruited via Lucid, YouGov, and Facebook (total n = 3,886). Survey professionalism is common, but varies across samples: by our most conservative estimate, we find 1.7% of respondents on Facebook, 7.6% on YouGov, and 34.7% on Lucid to be professionals (under the assumption that professionals are as likely as non-professionals to donate data after conditioning on observable demographics available from all online survey takers). However, evidence that professionals lower data quality is limited: they do not systematically differ demographically or politically from non-professionals and do not exhibit more response instability. They are, however, somewhat more likely to speed, straightline, and attempt to take questionnaires repeatedly. To address potential selection issues in donating of browsing data, we present sensitivity analyses with lower bounds for survey professionalism. While concerns about professionalism are warranted, we conclude that survey professionals do not, by and large, distort inferences of research based on online panels.

Keywords: survey professionalism; external validity; representativeness; data quality; web browsing data

Edited by: Erin Hartman

1. Introduction

Online panels have become an essential resource for low-cost survey research in political science and other fields. Yet, the recruiting methods of many commercial providers raise urgent questions about data quality and representativeness (Cornesse and Blom 2023; Hopkins and Gorton 2024; Jerit and Barabas 2023; Krupnikov, Hannah Nam, and Style 2021). A complicated web of panel companies, who often recruit via third-parties and even from one another, obscures the origin of respondents (Enns and Rothschild 2022).

From the panelist's perspective, the sheer number of platforms offering payment makes it attractive to become a "survey professional," namely, someone who does many online surveys and spends substantial time answering them. Survey taking is actively advertised as a full-time job (see Section A of the Supplementary Material) and professionals are likely to maximize their revenue by taking as many surveys as possible. Little is known about the extent and the consequences of survey professionalism. Research into the topic has relied entirely on self-reports (e.g., Matthijsse, De Leeuw, and Hox 2015;

Zhang *et al.* 2020), which may not be the best way to measure this phenomenon. Knowing that professionalism is undesirable for researchers, respondents are likely to under-report survey taking.

Our article provides a novel empirical strategy to address the phenomenon, using *behavioral* measures constructed from web browsing data. This allows us to estimate the prevalence of survey professionalism for our population of interest—that is, online survey takers—and put it in perspective to other behaviors such as using social media. We consider three potential downstream consequences for data quality. First, professionals may differ from non-professionals sociodemographically and politically (Krupnikov *et al.* 2021). Second, survey professionalism may entail speeding, satisficing, or insufficient effort in responding (Zhang *et al.* 2020). Third, it is also possible that survey professionals take the same questionnaire multiple times to increase their earnings (Ahler, Roush, and Sood 2025).

We analyze three different US samples with subjects who participated in our surveys and were willing to donate browsing data. We also have basic sociodemographics on participants unwilling to provide browsing data, which allows us to construct target weights based on both groups combined and weight our analyses towards the population of online survey takers. Our identification assumption is that, after conditioning on observable demographics (through weighting), professionals are equally common among those respondents who donate browsing data and those who do not donate. We provide sensitivity analysis results relaxing this assumption, allowing the share of professionals between donors and non-donors to vary. The samples were recruited using different platforms that constitute a range of recruitment approaches common in political science, i.e., through Facebook ads (n = 707), Lucid (n = 2,222), and YouGov (n = 957), with their browsing data totaling over 96 million web visits. We identify survey taking in three ways: based on existing lists of questionnaire websites, by using regular expressions applied to URLs, and by manually coding the most frequent websites in our data.

We first report estimates of the extent of survey professionalism across the three samples, weighted towards the population of online survey takers. Our results indicate that survey professionalism is most prevalent in the Lucid sample (54.6% of all visits, or 43.0% of time spent browsing), followed by YouGov (23.9% of all visits, or 18.2% of time spent browsing), and Facebook (10.2% of all visits, or 8.7% of time spent browsing). Out of four dichotomous measures of survey professionalism we offer, our lowest estimate puts the percentage of survey professionals at 1.7% of Facebook respondents, 7.6% of YouGov respondents, and 34.7% of Lucid respondents. In a sensitivity analysis, in which we drop the assumption that survey professionals are as likely to donate data as non-professionals, the lower bounds for this measure are at 1.1% of Facebook, 4.8% of YouGov, and 19.8% of Lucid respondents.

Second, we compare survey professionals and non-professionals on a range of sociodemographics and political characteristics, finding few consistent differences. Third, we compare the two groups on several indicators of response quality. Speeding through surveys is more common among survey professionals. However, we find only weak evidence for satisficing (measured by straightlining through grid questions), and no evidence that survey professionals show lower response stability over time. Finally, by analyzing visits to well-known questionnaire platforms, we explore how common it is that participants attempt to take the same questionnaire multiple times. We find that 26.7% of Facebook, 71.5% of Lucid, and 15.3% of YouGov respondents attempted to take *at least one* questionnaire more than once. This behavior is much more common among professionals than non-professionals.

In sum, although we uncover the widespread presence of survey professionalism across three types of opt-in online samples, evidence that this affects data quality negatively is limited. Professionals may be more prone to speeding and straightlining, but these behaviors can be measured and contained by excluding respondents. The prevalence of repeated attempted questionnaire taking is arguably our most worrying finding.

2. Extent and Consequences of Survey Professionalism

Previous literature has mostly understood survey professionalism in terms of the number of panels someone is a member of, or the number of surveys done in a certain time period (e.g., Callegaro *et al.* 2014; Gittelman and Trimarchi 2010; Zhang *et al.* 2020). This conceptualization is geared towards measurement via survey self-reports such as "How many Internet surveys have you completed before this one?" (cf. Zhang *et al.* 2020). Such questions do not easily translate into measurements using fine-grained web browsing data, since concepts such as "a survey" cannot be neatly counted in the vast browsing data a subject produces.

We offer a slightly different definition of survey professionalism, which follows the core ideas of previous work, while lending itself to measurement with web browsing data. Our unit of analysis is the visit to a survey site (henceforth "survey visit"), as well as its duration. We use a broad definition of survey visits, including filling out a questionnaire and the steps necessary before and after: signing up on platforms offering survey jobs, selecting surveys on such platforms, and obtaining the rewards. These various activities cannot be cleanly separated from one another, as many online platforms provide several or all of these functionalities. We define and measure a subject's survey professionalism as the number (and duration) of their survey visits. We offer details on measurement in Section 3.2.

We use our measure of survey professionalism in two ways. First, we examine it as a continuous variable, i.e., someone can have more or fewer survey visits or spend more or less time on surveys. We use this continuous variable to measure the overall extent of survey professionalism. Beyond this continuous understanding, the literature has also categorized subjects into "professionals" and "non-professionals." We follow this dichotomy and build three distinct binary measures, defining a professional as someone with (a) more than 100 survey visits a day; (b) more than 50% of their visits to survey sites; (c) more than 50% of their browsing time spent on survey sites. Additionally, we create a fourth measure that counts as professional (d) a subject meeting any of the three criteria above. We present the first approach in the main paper and report the others in the Supplementary Material.

2.1. The Extent of Survey Professionalism

This conceptualization and measurement allows us to quantify the extent of survey professionalism, a goal pursued by researchers for two decades. An early study reported substantial numbers of respondents to be members of twenty or more panels (Ipsos Insight, NPD Group, and TNS 2006). Later studies in the US, the UK (Gittelman and Trimarchi 2010), and the Netherlands (Willems, Van Ossenbruggen, and Vonk 2006) confirmed that membership in multiple panels was common. The issue with these estimates is that they rely on self-reports. As respondents may surmise that heavy survey participation is not desirable from the researcher's perspective, it is plausible that they under-report survey taking. In general, respondents tend to misreport media-related behaviors (Parry *et al.* 2021). Estimating the number of surveys taken in a certain period may be especially challenging to remember. As the internet is now the most important mode for survey recruitment in political science, we specifically target the population of *online* survey takers. Relying on online behavioral data, we ask: *What is the degree of survey professionalism among online survey takers? (RQ1)*

2.2. Sociodemographic Peculiarities of Survey Professionals

A common worry about online survey research is that samples may systematically differ from the general population (Krupnikov *et al.* 2021). Most studies focus on demographic and political characteristics of *convenience samples* compared to the population (Berinsky, Huber, and Lenz 2012; Chang and Krosnick 2009; Huff and Tingley 2015; Malhotra and Krosnick 2007; Valentino *et al.* 2020). There is less evidence on the peculiarities of *survey professionals*, who are conceptually distinct from members of convenience samples. Reviewing a number of U.S. studies, Whitsett (2013) concluded that frequent survey respondents were more likely to have lower education and no full-time job, although no differences emerged in terms of gender and age. More recent studies reached mixed conclusions on variables such as age, education, gender, ethnicity, political interest, and participation (Matthijsse *et al.* 2015; Zhang *et al.* 2020). We combine survey questions on demographics and political variables with our behavioral data and ask: *Do survey professionals differ from non-professionals sociodemographically and politically? (RQ2)*

2.3. Low Response Quality Among Survey Professionals (RQ3)

Concerns that online respondents might deliver lower quality of responses—manifesting in speedy responding, low attention, or a lack of careful thinking to produce meaningful responses (satisficing)—have been raised early on (Cape 2008). Although, the literature evaluates the quality of responses obtained from convenience samples in general (Chang and Krosnick 2009; Ternovski and Orr 2022; Zhang and Gearhart 2020), evidence on professional respondents is scarce. Zhang *et al.* (2020) and Matthijsse *et al.* (2015) find that professionals in the US and the Netherlands were more likely to speed through the questionnaire, but not more likely to engage in straightlining—that is, respondents choosing the same scale position on all items of a battery even if some are reverse-coded, a common way to measure satisficing (Chang and Krosnick 2009). Building on this research, we ask: *Are survey professionals more likely to engage in straightlining (RQ3a) and speeding (RQ3b) than non-professionals?*

A related concern about survey data quality regards insufficient effort responding (IER), which means that survey participants do not pay attention to the questions prior to responding (McGonagle, Huang, and Walsh 2016; Toich, Schutt, and Fisher 2022). As IER results in similar response patterns as random responses (Huang *et al.* 2015), it should increase over-time response instability. Previous studies have explored this issue for convenience samples in general (Holden, Dennie, and Hicks 2013; Shapiro, Chandler, and Mueller 2013), but not for survey professionals. Taking advantage of the multi-wave structure of our three datasets, we ask: *Do survey professionals exhibit higher over-time response instability than non-professionals? (RQ3c)*

2.4. Survey Professionals and Repeated Participation (RQ4)

Finally, we examine a more problematic possibility, namely, that highly active survey professionals may try to complete the same questionnaire[1] multiple times to maximize their revenue. Only a handful of studies have contemplated measuring this type of behavior, partly due to the methodological challenges of ascertaining multiple survey taking attempts by a single individual (Ahler *et al.* 2025; Berinsky *et al.* 2012). Most of these findings are restricted to the "closed system" of MTurk, in which the researcher directly invites participants. Other sample providers often recruit respondents from third parties so that a professional respondent who is member of both panel A and B might be invited to the same survey through both panels (Ternovski and Orr 2022). In contrast to studies that only check whether participation in *one* questionnaire may be repeated, we explore whether our participants may have attempted to repeatedly participate in *any* questionnaire we can identify. We ask: *What is the extent to which participants attempt to take the same questionnaire more than once, and do survey professionals engage in more repeated participation than non-professionals? (RQ4)*

3. Data and Methods

We rely on data collected from three distinct U.S. samples. For our main analyses, we only examine subjects who took the survey and shared their browsing data ("donors"). We use subjects not willing to donate ("non-donors") to construct demographic-based weights to estimate survey professionalism among the entire samples. Our first sample was recruited in 2018 through ads on Facebook for a two-wave survey. All survey participants were paid, and those who uploaded browsing data could win one of five $100 Amazon gift cards. Our second sample was acquired in 2019 for a three-wave survey from Lucid (now Cint), a provider that aggregates respondents from third-party sources. Respondents were paid up to $25 for providing their browsing data. The third sample was recruited in 2018 from the YouGov Pulse panel for a two-wave survey, compensated according to YouGov standards. All three studies received ethical approval.

[1]Throughout the article, we distinguish between "survey sites" / "survey taking," which is the wider concept, including the signing up to surveys and getting paid for them, and "questionnaires" / "questionnaire taking," which refers to the narrower activity of answering questions.

https://doi.org/10.1017/pan.2025.10018 Published online by Cambridge University Press

The three data sets combine survey responses with individual-level records of browsing behavior for two months for the YouGov sample and for 90 days before each wave for the Lucid and Facebook samples. In the Facebook and Lucid samples, after informed consent and some initial survey questions, participants were directed to an open-source tool, Web Historian, where they were asked to submit their browsing history stored in their browsers. In this method, web visits are collected retrospectively and only for desktop devices. In the YouGov sample, browsing data were collected with a tracking tool that participants had consented to install when signing up as panelists. YouGov browsing data were collected on-the-go, i.e., each web visit instantly produces a record, and include both desktop and mobile data. Section B.2.1 of the Supplementary Material provides details on how browsing data are recorded.

Across samples, some donating subjects participated in the surveys but provided little browsing data. As such subjects would distort some of the proportional metrics we calculate—for example, someone who submitted five visits in total, all of which are to a survey site, would be treated as doing surveys 100% of the time—we exclude subjects who submitted data from less than seven days. For details on sample recruitment and data exclusions, see Section B.1 of the Supplementary Material. Our final Facebook sample has $n_{subjects}$ = 707 and $n_{webvisits}$ = 16.4 million in the first wave (up to 90 days), the Lucid sample $n_{subjects}$ = 2,222 and $n_{webvisits}$ = 73.8 million in the first wave (up to 90 days), and the YouGov sample $n_{subjects}$ = 957 and $n_{webvisits}$ = 6.4 million (up to 60 days).[2]

3.1. External Validity

A potential weakness of our research design relates to systematic differences between donors and the population of interest, i.e., all online survey takers (donors plus non-donors). All three studies were—at least in part—open to both donors and non-donors, although the recruitment process differed slightly across the studies. On Facebook and Lucid, participants landing on the survey were first asked a few sociodemographic questions and then invited to donate their browsing data. If they declined, they could not proceed with the survey. On Facebook, out of 2,775 people who participated in our study, there were 820 (29.5%) donors (of which 707 had more than seven days of browsing data and were thus included in the main analysis) and 1,955 non-donors. On Lucid, out of 15,589 initial participants, there were 2,462 (15.8%) donors (of which 2,222 had more than seven days of browsing) and 13,127 non-donors. On YouGov, we analyze two independent sets of respondents. One set consisted of panelists who had installed YouGov's tracking software. The other set consisted of panelists from YouGov's "main" panel of respondents. We refer to the second set as the "non-donor" set (although it may include some respondents who have donated data to other researchers or to YouGov). Both sets of respondents received identical surveys during an identical time frame but we do not know how many panelists YouGov incentivized to provide tracking data to produce the respondents we have in our analysis. On Yougov, we then have 1,179 donors (of which 957 had more than seven days of browsing) and 4,543 non-donors.

As we have basic socio-demographic information about donors and non-donors, we can explore potential differences between these groups. As detailed in Section B.1.2 of the Supplementary Material, we find that donors in the Facebook sample are not significantly different from non-donors in terms of age, gender, education, and ethnicity. In the Lucid sample, donors have a similar gender and ethnicity profile but are somewhat younger, more educated, more liberal, and more likely to identify as Democrats. The starkest differences emerge for the Yougov sample, where donors are younger, less educated, less likely to be white, more positive towards the out-party, more likely to follow politics,

[2]For the continuous measures of professionalism, we focus on wave-1 data in the Lucid and Facebook samples to avoid attrition bias. In the YouGov sample, browsing data was collected for the whole period even if subjects did not return to wave 2. For the binary professionalism measures, we use all waves across all samples. For the analyses based on survey responses from both waves (RQ3c) we only use subjects participating in both. For analyses of repeated participation, we use data from all waves.

more politically knowledgeable and interested, but not significantly different in terms of gender and partisanship.

Although all of these differences are small in size, they raise the question of representativeness. As discussed, we target the population of online survey takers—for which there is no census. However, we assume that donors and non-donors taken together are a fair approximation of the population of online survey takers.[3] To maximize external validity, we construct target weights based on the sets of donors and non-donors combined, and weight the samples used for our main analyses with raking (Deville, Särndal, and Sautory 1993). For the Facebook sample, we base our weights on age, gender, education, and ethnicity. For Lucid and Yougov, we add ideology and partisanship. See Section B.4 of the Supplementary Material for details on the weighting procedure. Section C.1 of the Supplementary Material reports results for *RQ*1 without weights.

Our weighting approach should alleviate concerns of representativeness under the identification assumption that survey professionalism is as common among donor as among non-donors (after adjusting for the weighting variables). However, survey professionalism might be positively related to the willingness to donate browsing data, as both are presumably motivated by monetary rewards. Thus, we may overestimate the extent of professionalism. We run two analyses to address this concern. First, if survey professionals are more common among donors, we should see that donors have lower attrition rates than non-donors. For the sample, where we can test this implication (Yougov), we do *not* find support for differential attrition (see Figure C.14 in the Supplementary Material). Second, we perform a sensitivity analysis for our estimates of the percentage of professionals varying the assumption that the share of professionals is the same among donors as among non-donors, finding that the share of professionals remains substantial (see Section C.6 of the Supplementary Material). This sensitivity analysis provides the necessary caveats when interpreting the prevalence of survey professionalism.

3.2. Measuring Survey Professionalism

In all three browsing data sets, each row represents a web visit by a subject and has three main variables: a subject identifier, the URL, and a timestamp of the visit. From the URL, we first extract the *host*, defined as the part between the *scheme* (e.g., "https://") and the *path* (beginning at the first "/"). The host always includes the *domain*, but sometimes it conveys more information. For example, the URL "https://survey123.panel456.com/r/abcd" would have the host "survey123.panel456.com" and the domain "panel456.com". For details on the complexities of URL parsing, see Clemm von Hohenberg *et al.* (2024).

3.2.1. Identifying Survey Visits

We identify survey visits at the level of URL hosts. As exemplified above, a URL host sometimes provides more information than its domain, allowing us to minimize false negatives. As defined earlier, survey visits comprise both filling out questionnaires and the steps before and after that, such as accepting study invitations and obtaining rewards. To capture the variety of websites linked to survey taking, we follow three different approaches.

- First, we rely on a report of questionnaire software published by Bevec and Vehovar (2021). From their comprehensive list (see their Table 12 on page 7), we filtered out irrelevant types (e.g., "UX tool") and manually verified the web addresses connected to each questionnaire software, such as "qualtrics.com" or "surveymonkey.com".
- Second, assuming that the content of URLs provides clues about their content, we classified all hosts that contained the word "survey" as survey taking.

[3] We recognize that this statement needs to be qualified to the extent that the recruitment processes of the three platforms are a black box, and it is impossible to definitively understand why a given individual was invited to a survey or not by the platforms' mechanisms.

https://doi.org/10.1017/pan.2025.10018 Published online by Cambridge University Press

- Third, an approach based purely on the identification of the word "survey" might miss important sites in our data. Two trained research assistants manually coded the 500 most frequently visited hosts from each of our three datasets for whether people are taking surveys on them, that is either responding to questionnaires or being recruited/rewarded for such activities. Detailed coding instructions can be found in Section B.2.2 of the Supplementary Material. Each of the three lists was coded by two coders with an overlap of at least 10% of cases. Inter-coder reliability for the overlaps was high (Facebook: 92% agreement; Lucid: 88%; YouGov: 94%).

The final list of hosts based on these three approaches can be found in the reproduction materials. We present results broken down by the three approaches in Figure C.8 in the Supplementary Material and highlight the most dominant survey sites in Figure C.9 in the Supplementary Material.

3.2.2. Measuring Survey Professionalism

Based on this list of hosts, we classify each visit by each participant in the three datasets as a survey visit or not. We then compute several variables: the subject's count of survey visits; the subject's time spent on survey sites; the proportion of survey visits out of all visits, the proportion of time spent on survey sites out of the overall browsing time of each participant. To measure time spent, we use timestamps to order all visits for a person chronologically and take the time until the subsequent visit as the duration—unless this time exceeds a threshold of five minutes. In this case, the participant likely paused their browsing activity, and we code the visit duration as missing (cf. Clemm von Hohenberg *et al.* 2024). We use these continuous count and duration variables to measure the degree of survey professionalism (*RQ*1). To address *RQ*2, *RQ*3, and *RQ*4, we dichotomize these continuous metrics to create four alternative classifications of survey professionals. For our primary analyses, we define a survey professional (a) as someone who has on average more than 100 survey visits per active day (i.e., a day on which they were online). In the Supplementary Material, we report results based on three alternative definitions, namely (b) a respondent with more than 50% of all visits to survey sites; (c) a respondent who spends more than 50% of all browsing time on survey sites, and (d) a respondent who meets *any* of the three conditions.

3.2.3. Measuring Repeated Participation

To measure how often individuals attempt to take the same questionnaire multiple times (*RQ*4), we take advantage of the fact that questionnaire platforms assign unique URLs to the same questionnaire. For example, "https://www.surveymonkey.com/r/MT2B22T" will permanently point to the same questionnaire. The methodological challenge here is not to count visits to generic URLs from the same platform, e.g., "https://www.surveymonkey.com/login," as a repeated attempt. We identify patterns in URLs that reliably designate unique questionnaires for eleven common questionnaire platforms (e.g., Qualtrics, SurveyMonkey, QuestionPro). See Section B.2.3 of the Supplementary Material for details.

Whenever a questionnaire URL (e.g., "https://www.surveymonkey.com/r/MT2B22T") appears more than once in a participant's browsing data, the participant potentially took this questionnaire repeatedly. However, some visits to the the same questionnaire can be legitimate, for example, a participant accidentally clicking twice on the questionnaire link when invited, or coming back to a questionnaire to finish it. Hence, we only count visits to the same questionnaire URL as repeated if done at least one hour later (we vary this parameter in the Supplementary Material) and if not happening directly after each other. We emphasize again that we measure *attempts* to take questionnaires repeatedly because we do not know if the participant completed the questionnaire multiple times, as questionnaire owners might have guardrails in place to prevent such behavior.

3.3. Survey-Based Measures

3.3.1. Sociodemographics

To compare professionals and non-professionals in terms of sociodemographics (*RQ*2), we rely on measures of age, gender, education, and ethnicity. We also compare our samples to 2020 U.S. census statistics on these four variables.

https://doi.org/10.1017/pan.2025.10018 Published online by Cambridge University Press

3.3.2. Political Outcomes

To explore whether professionals differ from non-professionals on various political characteristics (*RQ*2), we analyze partisanship, ideology, an out-party feeling thermometer (dropping Independents), political interest, political knowledge, and following politics in the news. All measurements are reported in detail in Table B.3 in the Supplementary Material. For better comparability across data sets, we recode ideology, political interest, political knowledge, and following politics to a scale from 0 to 1. We benchmark our sample statistics against ANES 2020 survey data.

3.3.3. Speeding and Straightlining

Across all samples, we use individual duration (in seconds) of taking the first-wave survey to assess speeding (*RQ*3*b*). To provide comparability with previous work, we also analyze the percentage of participants who were 30%, 40%, or 50% faster than the median duration (Greszki, Meyer, and Schoen 2015). To detect straightlining (*RQ*3*a*), we use multi-item batteries in which, for some items, higher values represent a higher value on the construct, while the reverse is the case for other items (Facebook: five-item battery about gun control attitudes; Lucid: five-item battery measuring attribution of malevolence to the out-party; YouGov: six-item battery on abortion attitudes). We flag subjects if they chose the same scale value on all items except if that value was the midpoint.

3.3.4. Over-Time Response Instability

To identify if survey professionals change their responses to the same item at a higher rate than non-professionals (*RQ*3*c*), we use all variables that were measured across waves with precisely the same question wording and a numeric response scale. We identified 59 such variables in the Facebook data, 151 in the Lucid data, and 31 in the YouGov data. These are primarily standard political science questions, such as issue attitudes, media use, or political preferences.

4. Results

4.1. The Extent of Survey Professionalism (RQ1)

How much do the participants in our three samples engage in survey taking? Figure 1 shows the percent of visits to survey sites out of all visits (bars in dark tones). These are aggregate-level percentages showing how much out of the sample's web visits go to survey sites. For comparison, the figure also plots the percent of visits to *facebook.com*, *google.com*, *amazon.com*, and *youtube.com*. Survey-taking constitutes a substantial part of participants' browsing: 54.6% of all visits in the Lucid sample and 23.9%in the YouGov sample are visits to survey-taking sites. This is much more than visits to, for example, *google.com* (6.0% and 11.3%,respectively). Survey professionalism is least prevalent in the Facebook sample, i.e., 10.2% of all visits are to survey-taking sites, which is less than visits to *google.com* (25.1%).

Figure 2 plots the distribution of survey visits per respondent. Although extreme values exist across the three samples, the shapes of the distributions differ drastically. The Facebook sample has the most skewed distribution—a few people account for most visits—followed by the YouGov sample. The Lucid sample is the least skewed, indicating that it is more common to engage in a lot of survey taking among Lucid participants. The average percentages of survey visits based on individual-level variation, depicted in Figure C.10 in the Supplementary Material, are slightly smaller than the aggregate percentages (Facebook: 8.4%; Lucid: 49.7%; YouGov: 16.3%).

In Section C.5 of the Supplementary Material, we report statistics in terms of visit duration. A substantial share of our samples' browsing time is spent on survey sites, whether measured as an aggregate percentage (Facebook: 8.7%; Lucid: 43%; YouGov: 18.2%) or as an average individual-level percentage (Facebook: 6.8%; Lucid: 41%; YouGov: 15.2%). In Section C.3 of the Supplementary Material, we report which of our three approaches to identify survey sites, and which specific sites, account for most survey visits. Overall, the distribution of the three methods is similar across datasets.

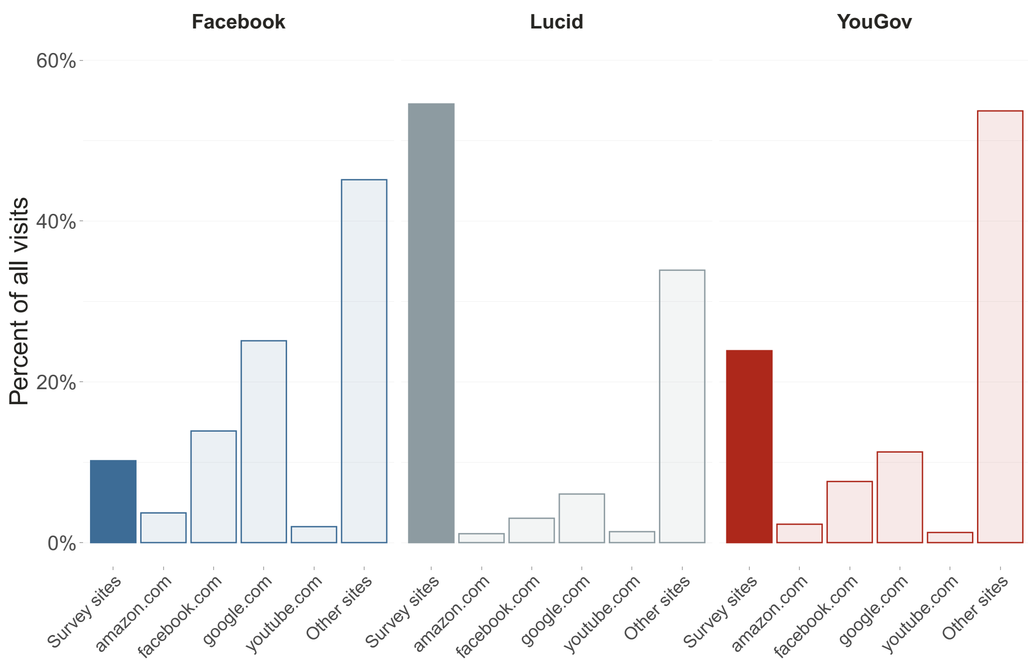

Figure 1. Percent of survey visits (out of all visits) compared to visits to popular web domains.

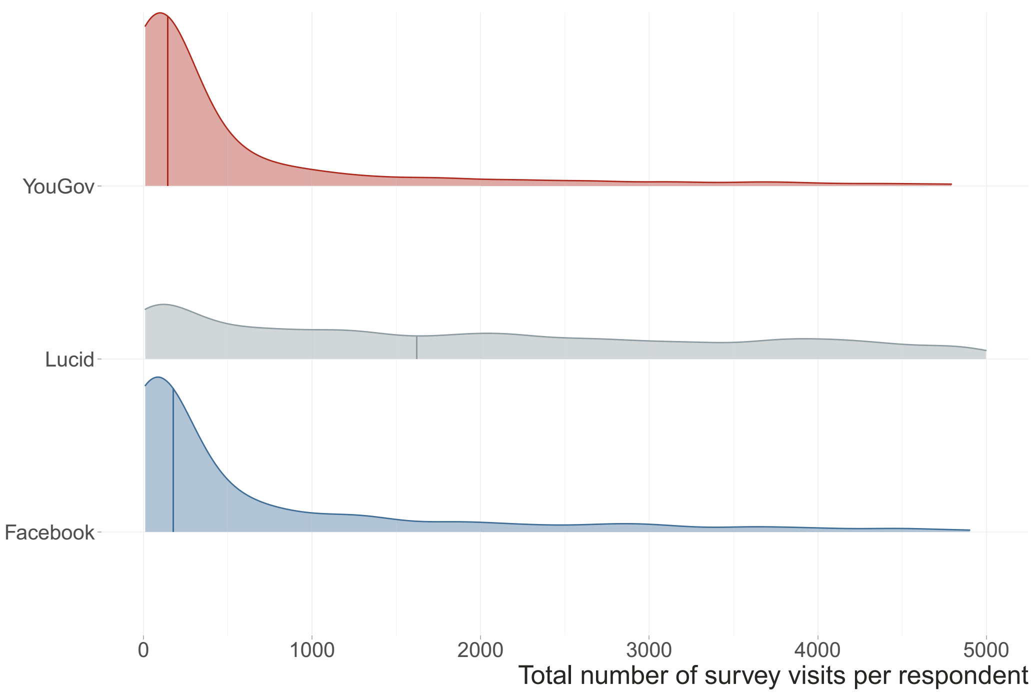

Figure 2. Individual level distribution of survey visits per respondent across the three samples.
Note: Lines represent the median value for the distribution.

https://doi.org/10.1017/pan.2025.10018 Published online by Cambridge University Press

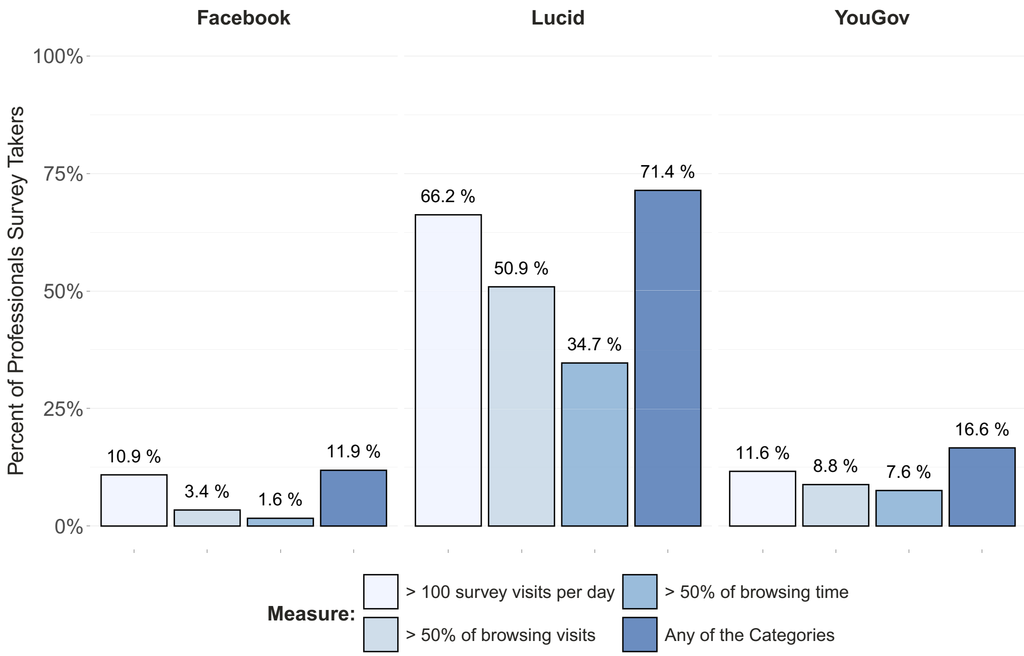

Figure 3. Percent of survey professionals for different definitions of survey professionalism.

Figure C.9 in the Supplementary Material presents the ten most prevalent survey sites across the samples. We see that *swagbucks.com*, *mturk*, *samplicio.us*, and *decipherinc.com* represent some of the largest survey sites. This shows that participants active on one panel, e.g., YouGov, take surveys across different (and somewhat competing) platforms.

To address the next set of research questions (*RQ*2, *RQ*3a,b,c, and *RQ*4), we classify subjects as "professionals" or "non-professionals" using the four different categorizations described earlier. As shown in Figure 3, in the Lucid sample, the estimates of survey professionals vary between 34.7% and 66.2%, and go up to 71.4% when, including subjects categorized as professionals by any of the measures; in the YouGov sample, between 7.6% and 11.6%, with 16.6% when using any of the categories; and in the Facebook sample, between 1.7% and 10.9%, with 11.9% when using any of the categories. For the remainder of the article, we use the first definition of survey professionals (> 100 visits per active day).

As aforementioned, if we assume that survey professionalism is as common among donors as among non-donors (after weighting the sample of donors), these estimates should be unbiased. However, this assumption may be unrealistic and we relax it in a sensitivity analysis (details in Section C.6 of the Supplementary Material). Using the indicator of professionalism yielding the most conservative estimates, the share of professionals in the Facebook and YouGov samples does not change substantially, reducing from 1.7% to 1.1%, and from 7.6% to 4.6%, respectively. Bounds are wider in Lucid, and the estimation for the percentage of professionals would decrease from 34.7% to 20.3%.

In sum, survey taking is prevalent, but also varies substantively across samples: Lucid is the panel with the highest proportion of survey professionals; the Facebook sample has the lowest proportion. The YouGov sample sits between these extremes, indicating that even on a high-quality panel from one of the most reputable survey companies, the presence of survey professionals is not negligible. As the YouGov sample was collected from both desktop and mobile devices, we explore potential differences between these two modes (Section C.2 of the Supplementary Material): professionalism tends to be slightly more prevalent among those who primarily use desktop.

4.2. Sociodemographic and Political Differences (RQ2)

Do survey professionals differ from non-professionals in terms of demographic and political characteristics? Table 1 reports the central tendencies of the two groups, as well as of the U.S. population, on a number of key variables. The first rows, which compare professionals and non-professionals demographically, show the most pronounced differences in the Lucid sample, in which professionals are older, more highly educated, and more ethnically white than non-professionals at conventional levels of statistical significance. We see similar patterns in the Facebook sample, but no statistically significant difference in terms of age in the YouGov sample. Note that in the Facebook and YouGov samples, the number of professional respondents is small, which affects statistical power. Tables D.4, D.5, and D.6 in the Supplementary Material report the same statistics for our alternative categorizations, showing that the differences are not very robust: Only the finding that professionals are older in the Lucid sample remains significant across approaches.

The second part of Table 1 reports political differences. Across the samples, survey professionals tend to be more conservative than non-professionals—although this difference only remains statistically significant across categorizations in the Lucid sample (see Tables D.4, D.5, and D.6 in the Supplementary Material). Professionals also tend to feel more positive towards out-partisans: In the YouGov sample, this difference is statistically significant no matter the categorization of professionals. In the Lucid sample, professionals are further more politically interested. In the YouGov sample, survey professionals are less politically knowledgeable and less likely to follow politics than the non-professionals. Taken together, we do not find consistent sociodemographic or political differences between professionals and non-professionals.

4.3. Response-Quality Differences (RQ3a,b,c)

To examine whether response quality varies between professional and non-professional participants, we again report results using our first categorization of professionals (see Section ES of the Supplementary Material for the other categorizations). The first row in Table 2 reports straightlining behavior on grid questions among survey professionals and non-professionals (*RQ3a*). Overall, across the samples, professionals show a higher incidence of straightlining. For the YouGov sample, this difference is statistically significant at the 95% level—however, it is insignificant for the alternative professionalism measures.

The more pronounced differences appear in terms of speeding (*RQ3b*), reported in rows 2–5. First, consider the absolute measure of duration of taking our surveys. For the Lucid sample, the median professional is 6.9% faster compared to the median non-professional, representing a difference of one minute and twenty-six seconds. This difference is yet more pronounced for the YouGov and Facebook samples: professionals are 18.7% and 13.4% faster, respectively. When we compare speeding rates among professionals and non-professionals using the different criteria for finishing faster than the median (rows 3–5), we again see that professionals speed more than non-professionals (though the difference is not statistically significant throughout).

Next, *RQ3c* asks if professionals show more response instability across waves than non-professionals. Our quantity of interest is the variability of wave-two responses after taking into account responses in wave one. To estimate this parameter, we fit a Bayesian heteroscedastic linear model that allows us to estimate the difference in the standard deviation between professionals and non-professionals explicitly. Formally, the model can be described as:

$$Y_{iw_2} \sim \mathcal{N}(\mu_i, \sigma_i^2)$$
$$\mu_i \sim \alpha + \beta_1 Y_{iw_1} + \beta_2 SP_i + \beta_3 Y_{jw_1} \cdot SP_i + \epsilon_i$$
$$\sigma_i = SP_i \cdot \sigma_{\text{pro}} + (1 - SP_i) \cdot \sigma_{\text{nonpro.}}$$

We regress the wave-two response for outcome (Y_{iw_2}) on the wave-one response for the same outcome (Y_{iw_1}), plus the interaction of the wave-one response with a dummy for survey professionalism (SP_i).

https://doi.org/10.1017/pan.2025.10018 Published online by Cambridge University Press

https://doi.org/10.1017/pan.2025.10018 Published online by Cambridge University Press

Table 1. Survey professionals versus non-professionals versus population (professionals = more than 100 survey visits / day).

		Facebook			Lucid			Yougov		
	U.S. population	Professionals		Non-professionals	Professionals		Non-professionals	Professionals		Non-professionals
Sociodemographics										
Age (median years)	38.2	40–44	⋆⋆	35–39	46	⋆⋆⋆	37	56		55
Gender (% female)	50.8	74.8 (5.3)		74.9 (1.8)	54.4 (1.4)		54.0 (2.0)	58.4 (5.5)		55.2 (2.0)
Education (% Bachelor or more)	30.4	56.0 (6.0)		56.0 (2.0)	47.1 (1.4)	⋆	41.8 (1.9)	43.9 (5.7)		40.7 (2.0)
Ethnicity (% white)	62.6	89.1 (3.7)		84.3 (1.5)	81.6 (1.0)	⋆	77.1 (1.6)	78.5 (4.1)		77.4 (1.5)
Political outcomes										
Partisanship (1–7)	4 (0.059)	3.7 (0.3)	∘	3.1 (0.1)	3.8 (0.1)		3.7 (0.1)	3.6 (0.3)		3.6 (0.1)
Ideology (0–1)	0.54 (0.006)	0.49 (0.03)	⋆⋆	0.40 (0.01)	0.54 (0.01)	⋆⋆	0.50 (0.01)	0.56 (0.03)	∘	0.50 (0.01)
Thermometer out-party (1–100)	17.4 (0.425)	28.9 (3.3)		25.0 (0.9)	28.2 (0.8)		28.8 (1.1)	19.1 (2.9)	⋆	12.6 (0.9)
Political interest (0–1)	0.4 (0.006)	0.65 (0.04)		0.66 (0.01)	0.69 (0.01)	⋆⋆	0.63 (0.02)	0.70 (0.04)	⋆⋆	0.82 (0.01)
Political knowledge (0–1)	0.5 (0.006)				0.65 (0.01)		0.62 (0.02)	0.62 (0.04)		0.68 (0.01)
Following politics (0–1)		0.64 (0.04)		0.61 (0.01)				0.57 (0.04)	⋆⋆	0.69 (0.01)

Note: Standard errors of means in parentheses. Significance of differences between professionals and non-professionals were tested with a Kolgomorov–Smirnoff test for age, chi-squared tests for gender, education and race, and t-tests for all other variables ($\circ p < 0.1$; $\star p < 0.05$; $\star\star p < 0.01$; $\star\star\star p < 0.001$). Sociodemographic population data from the U.S. Census; political variables from ANES 2020. Variables trust, political interest, knowledge, and partisanship were recoded to a scale from 0 to 1 to ensure comparability.

https://doi.org/10.1017/pan.2025.10018 Published online by Cambridge University Press

Table 2. Response quality of survey professionals versus non-professionals (professionals = more than 100 survey visits / day).

	Facebook			Lucid			Yougov		
	Professionals		Non-professionals	Professionals		Non-professionals	Professionals		Non-professionals
Straightliner (%)	2.6 (1.8)		0.9 (0.4)	1.9 (0.4)		1.4 (0.5)	7.8 (2.8)	$\star$	2.9 (0.5)
Duration (median seconds)	733 (77)	$\circ$	846 (25)	1167 (20)		1253 (42)	1466 (139)		1804 (60)
Duration (% 30% faster than median)	33.3 (5.6)	$\star$	20.7 (1.6)	22.9 (1.1)		21.4 (1.5)	33.0 (5.2)	$\star$	21.0 (1.6)
Duration (% 40% faster than median)	26.7 (5.2)	$\star\star$	13.6 (1.4)	15.6 (0.9)		15.2 (1.3)	22.9 (4.6)	$\star\star$	12.7 (1.2)
Duration (% 50% faster than median)	13.1 (3.9)		8.1 (1.1)	10.1 (0.8)		10.7 (1.2)	8.9 (2.7)		6.5 (0.9)

Note: Standard errors of means in parentheses. Significance of differences between professionals and non-professionals were tested with a Kolgomorov–Smirnoff test for survey duration, and chi-squared tests for the proportion of those 30/40/50% faster than the median duration and for proportion of straightliners ($\circ p < 0.1$;$\star p < 0.05$;$\star\star p < 0.01$;$\star\star\star p < 0.001$).

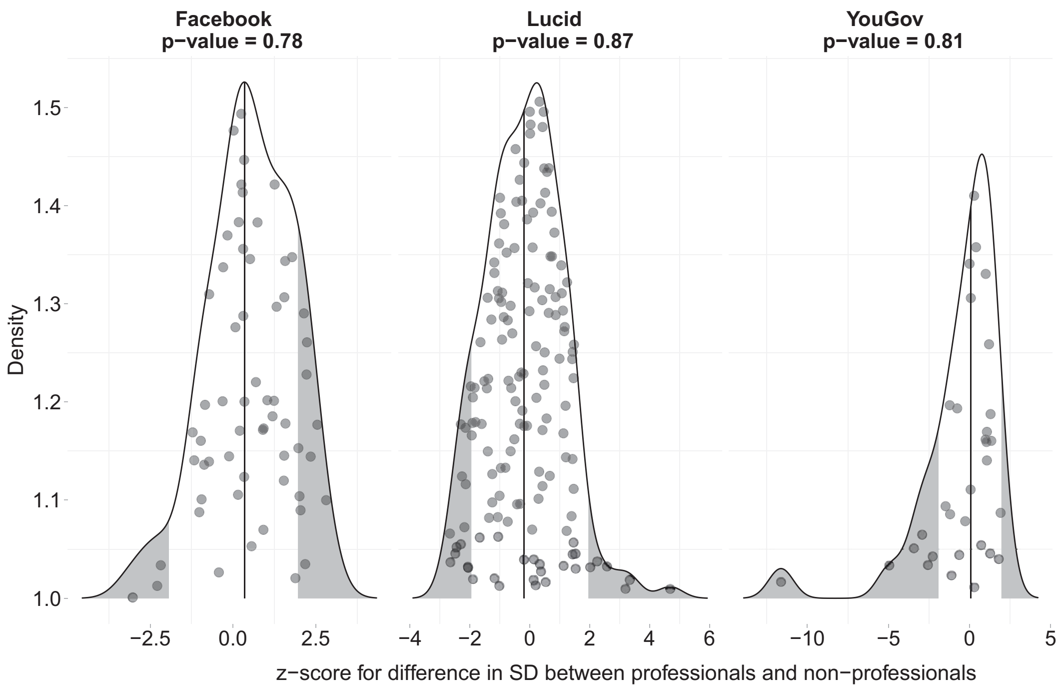

Figure 4. Z-scores for the difference in between-wave standard deviation between professionals and non-professionals.
Note: Gray areas contain z-scores larger than 1.96.

We allow the standard deviation of the residual (Y_{iw_2}) to vary conditional on SP_i. Our quantity of interest is the difference in the posterior distributions between σ_{pro} and σ_{nonpro}, which represents the difference in response stability between professional and non-professionals. The full model specification is discussed in Section E.2 of the Supplementary Material. Similar modeling approaches have been used to investigate the stability of policy attitudes (Alvarez and Brehm 1995; Clark, Nordstrom, and Reed 2008; Garner and Palmer 2011).

Figure 4 presents the distribution of z-scores for differences in the standard deviation between professionals and non-professionals for all survey outcomes measured in both waves. Using this distribution, we calculate exact p-values (Imbens and Rubin 2015, Chapter 5) for the null hypothesis that the average z-score of the interactive terms is equal to zero. Across all three samples, only a few coefficients for the interactive term achieve statistical significance at conventional levels and all three distributions are centered around zero, showing no consistent patterns of lower response stability among professionals. Thirteen outcomes (out of 59) yield statistically significant parameters in the Facebook sample, nineteen (out of 149) in the Lucid sample, and six(out of 31) in the YouGov sample. For some of these outcomes, professionals actually show *lower* standard deviation than non-professionals. The exact p-values are higher than 0.78in all three cases, which does not allow us to reject the null hypothesis. Section E.2.1 of the Supplementary Material reports the point estimates with confidence intervals for all models. To ensure robustness of these findings, we also run an alternative model in which we regress the absolute difference between wave-two and wave-one outcome on the professionalism dummy using a simple OLS model. If professionals showed less response stability, these absolute differences should be higher, which is not the case.

To summarize, despite the high prevalence of survey professionals across the three samples, we do not find substantive differences in the quality of responses between professionals and non-professionals. Professionals do speed through surveys more, but they are not more likely to straightline through grid questions or to show more unstable responses to questions asked across waves. As an additional

Table 3. Repeated questionnaire participation.

	Facebook	Lucid	Yougov
Subjects taking at least one questionnaire repeatedly (%)	26.7 (1.7)	71.5 (1.0)	15.3 (1.2)
Number of repeated questionnaires per participant (mean)	1.1 (0.2)	8.3 (0.4)	0.5 (0.1)
Percent of repeated questionnaires per participants (mean)	5.0 (0.5)	6.8 (0.2)	2.2 (0.3)

Note: Standard errors in parentheses.

piece of evidence about potential quality differences, we also analyze whether the effects of two treatments administered in the Lucid sample vary between professionals and non-professionals, finding no differences (Section E.3 of the Supplementary Material).

4.4. Repeated Questionnaire Participation (RQ4)

To examine if respondents attempt to take the same questionnaires repeatedly, we track if an individual's data show two or more visits to the same questionnaire URL with a time difference of at least one hour and not happening directly after each other, as detailed earlier.[4] For the eleven questionnaire platforms, we analyze, our data sets contain 70,220 unique questionnaires, which have been visited by respondents in our samples 273,840 times. It is not surprising that questionnaires are taken more than once *across individuals*. However, here we ask whether *the same individual* attempts to take the same questionnaire more than once.

Table 3 reports several statistics capturing attempts to take the same questionnaire repeatedly. First, the percent of respondents attempting to take *at least one* questionnaire repeatedly is 26.7% for Facebook, 71.5% for Lucid and 15.3% for YouGov. This statistic does not necessarily imply that participants in this set attempt to take *many* questionnaires repeatedly. The second row speaks to the absolute number of repeated questionnaires per individual, of which we take the average across all participants. For example, if one individual attempted to take five questionnaires repeatedly and the other had fifteen attempts, then their average would be ten (no matter how many questionnaires they did in total). The table shows an average of 1.1 and 0.5 for the Facebook and YouGov samples, but 8.3 for the Lucid sample. In other words, on average, a Lucid participant attempted to take more than eight questionnaires more than once. The third row summarizes, as an average, the share of questionnaires participants take repeatedly out of all questionnaires they take. For example, if one participant took fifty unique questionnaires in total and of those ten multiple times, her percentage would be 20%; if another took one hundred unique questionnaires and ten more than once, his percentage would be 10%; their average would be 15%. The table shows that the average percentages of repeated participation range from 2.2% (YouGov) to 5.0% (Facebook) and 6.9% (Lucid).

We emphasize that our findings do not imply that YouGov participants take *YouGov* questionnaires multiple times, as the company operates its own (closed) questionnaire platform, which we could not include in our identification of unique questionnaire URLs. However, our YouGov panelists *do* attempt to repeat questionnaires on the eleven other platforms we identified, even if they are not sent there by YouGov.

Table 4 shows that among professionals, the share of those who have taken at least one questionnaire repeatedly is much higher than among non-professionals. For example, in the Facebook sample, 74.4% of professionals took at least one questionnaire more than once, whereas this is the case for only 21.1% of

[4] We emphasize that we can only measure *attempts* to take the same questionnaire. We know when the respondent went to the same questionnaire URL more than once, but we do not observe whether the survey provider may have turned away the respondent.

Table 4. Repeated questionnaire participation, professionals versus non-professionals (professionals = more than 50 of browsing time to survey sites).

	Facebook		Lucid		Yougov	
	Non-professionals	Professionals	Non-professionals	Professionals	Non-professionals	Professionals
Subjects taking at least one questionnaire repeatedly (%)	21.1 (1.6)	74.4 (5.0)	44.6 (1.9)	85.2 (1.0)	8.2 (1.0)	69.2 (4.7)
Number of repeated questionnaires per participant (mean)	0.6 (0.1)	5.2 (1.0)	1.6 (0.2)	11.6 (0.6)	0.1 (0.0)	3.5 (0.6)
Percent of repeated questionnaires per participants (mean)	4.8 (0.5)	6.2 (0.9)	5.5 (0.4)	7.4 (0.2)	1.6 (0.3)	6.7 (0.7)

Note: Standard errors in parentheses.

non-professionals. For Lucid, these numbers are 85.1% versus 44.6%; for YouGov, 69.2% versus 8.2%. In all three samples, professionals also show higher absolute and relative numbers of repeated participation.

These results are based on our cutoff of one-hour. We present results with alternative cutoffs (six hours and one day) in Section F.1 of the Supplementary Material. Although the prevalence of repeated participation slightly decreases, it is still high: For example, when using the 24-hour cutoff, the share of subjects taking at least one questionnaire more than once is still 22.5%, 68.2% and 13.1% for Facebook, Lucid, and YouGov, respectively. We also disaggregate results by the questionnaire platforms we identified in Section F.3 of the Supplementary Material.

We also consider an alternative explanation of repeated participation: Several visits to the same questionnaire could be the result of participants taking breaks, which would manifest in long response times for individual questions. However, we do not find that it is very common that respondents take as long as one hour for individual questions (see Section F.5 of the Supplementary Material). Finally, we consider whether our results could be driven by survey providers over-targeting respondents from hard-to-reach groups, who can be reached comparatively well with online surveys (Hopkins and Gorton 2024; Rosenzweig *et al.* 2020). Section F.4 of the Supplementary Material zooms in on respondents over 65, from non-white racial groups, and Republicans. All three groups are somewhat more likely to attempt taking surveys repeatedly. We advise researchers to be mindful of these dynamics when using online surveys to recruit hard-to-reach participants.

5. Discussion

Our article uses a novel measurement approach based on web browsing data to offer three key findings about the extent and effects of survey professionalism in online surveys. First, professional survey-taking represents a substantial portion of the online activity of the samples. Lucid (now Cint) shows the highest prevalence of survey professionalism, followed by YouGov and Facebook. Visits to survey sites are more than half of all visits in the Lucid sample, around a quarter in the YouGov sample, and one tenth in the Facebook sample.

Second, although survey professionals constitute a non-trivial part of our samples, we do not find that they introduce significant inferential problems for online research. Although there are some demographic and political differences between professionals and nonprofessionals, these largely depend on the sample and the categorization of professionals. More importantly, survey professionalism does not seem to have pronounced implications for data quality. Professionals do show a greater tendency to speed through surveys. YouGov professional panelists also engage in more straightlining. But to the extent that these behaviors are observable, data providers, and researchers themselves can decide whether to remove speeders and straightliners from their analyses. Crucially, survey professionals do not show greater instability of responses to the same questions over time. We take this as evidence that they answer questions at least as attentively as non-professionals.

Our third core finding hints at one problematic consequence of survey professionalism. 26.7% of subjects in the Facebook sample, 15.3% in the YouGov sample, and 71.5% in the Lucid sample attempted to take at least one questionnaire multiple times. Across all three samples, repeated participation is substantially higher among survey professionals. These findings suggest that bad actors are present in the survey ecosystem and underscore the importance of building systems to detect repeated participation.

Some limitations of our study are worth highlighting. Our study design, based on behavioral measurement of survey taking, arguably offers higher internal validity than previous studies based on self-reports. Yet, we acknowledge that our behavioral measure requires subjects' willingness to share their web-browsing data. As we show in Table B.1 in the Supplementary Material, donors vary in some key demographics compared to non-donors. Even after adjusting for these differences by weighting, professionalism might be more prevalent among donors. To address this concern, we provide a sensitivity analysis bounding our results. Second, it is unclear to what extent our results generalize to other vendors. We provide a range of recruitment methods commonly used by political scientists. Nevertheless, the prevalence and consequences of survey professionalism may differ among other common vendors such as MTurk or Prolific. Future work should examine our questions across different panels and time periods, potentially based on strategies and measures presented here.

Despite these limitations, we offer previously unavailable evidence on the extent and implications of survey professionalism in U.S. panels. We encourage researchers to continue examining these phenomena and their consequences for what the public knows from research. Decades ago, lab experimentalists were criticized for their reliance on college students (mostly white, more affluent, and better educated than the general population). Online panels were hoped to offer an antidote through access to a more diverse set of individuals partaking in surveys or experiments. Although online samples are also unusual as they attract survey professionals, our finding that these do not, by and large, distort inferences makes us cautiously optimistic.

Acknowledgements. We thank Nejla Asimovic, Michael Bailey, Raphael Callel, Taylor Boas, James Druckman, Ernesto Calvo, and seminar participants at the Summer Institute of Computational Social Science (Rutgers) 2024, at the American Political Science Annual Meeting 2024, at the European Political Science Association 2024 conference, at NYU's Center for Social Media and Politics, and at the Social Data Science and AI Lab at Ludwig-Maximilians-Universitaet Munich for helpful feedback.

Funding Statement. J.N. gratefully acknowledges funding for data collection for this work from Democracy Fund, and that the Center for Social Media and Politics at New York University is supported by funding from the John S. and James L. Knight Foundation, the Charles Koch Foundation, Craig Newmark Philanthropies, the William and Flora Hewlett Foundation, and the Siegel Family Endowment. This research was also supported in part through the NYU IT High Performance Computing resources, services, and staff expertise. M.W. gratefully acknowledges the support of the European Research Council (ERC Starting Grant, EXPO- 756301, PI Magdalena Wojcieszak).

Data Availability Statement. Replication code for this article is available at Clemm von Hohenberg et al. (2025). A preservation copy of the same code and data can also be accessed via Dataverse at https://doi.org/10.7910/DVN/PY8COU.

Author Contributions. B.C.vH., T.V., J.N., and M.W. designed the research. B.C.vH. and T.V. did all the data analysis and wrote the first version of the article. B.Cv.H., T.V., J.N., and M.W. authors contributed to revising the manuscript. J.N., M.W., E.M.T., and B.C.vH. collected the data.

Competing Interests. J.N. was listed as an investigator on a research grant from Facebook for "Misinformation on Social Media: Investigating the Role of Digital Literacy" that was awarded to Princeton University. M.W. was the principal investigator

on two research grants from Facebook and is part of the 2020 Facebook and Instagram Election Study. The remaining authors declare none.

Ethical Standards. The three studies were approved by the Institutional Review Board of the University of Amsterdam / New York University.

Supplementary Material. For supplementary material accompanying this paper, please visit https://doi.org/10.1017/pan.2025.10018.

References

Ahler, D. J., C. E. Roush, and G. Sood. 2025. "The Micro-Task Market for Lemons: Data Quality on Amazon's Mechanical Turk." *Political Science Research and Methods* 13 (1): 1–20. https://doi.org/10.1017/psrm.2021.57

Alvarez, R. M., and J. Brehm. 1995. "American Ambivalence towards Abortion Policy: Development of a Heteroskedastic Probit Model of Competing Values." *American Journal of Political Science* 39 (4): 1055–1082.

Berinsky, A. J., G. A. Huber, and G. S. Lenz. 2012. "Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk." *Political Analysis* 20 (3): 351–368. https://doi.org/10.1093/pan/mpr057

Bevec, D., and V. Vehovar. 2021. *A WebSM Study: Web Survey Software 2021*. Ljubljana: Faculty of Social Sciences, Centre for Social Informatics. https://repozitorij.uni-lj.si/IzpisGradiva.php?lang=eng&id=133467

Callegaro, M., A. Villar, D. S. Yeager, and J. A. Krosnick. 2014. "A Critical Review of Studies Investigating the Quality of Data Obtained with Online Panels Based on Probability and Nonprobability Samples." In: Callegaro M, Baker R, Bethlehem J, Göritz AS, Krosnick JA, Lavrakas PJ, eds. *Online Panel Research: A Data Quality Perspective*, 23–53. Hoboken, NJ: John Wiley & Sons. http://www.wiley.com/WileyCDA/WileyTitle/productCd-1119941776.html

Cape, P. 2008. "ASC Conference—Multiple Panel Members: Saints or Sinners?" *International Journal of Market Research* 50 (5): 702–704. https://doi.org/10.2501/S1470785308200122

Chang, L., and J. A. Krosnick. 2009. "National Surveys via RDD Telephone Interviewing Versus the Internet: Comparing Sample Representativeness and Response Quality." *Public Opinion Quarterly* 73 (4): 641–678. https://doi.org/10.1093/poq/nfp075

Clark, D. H., T. Nordstrom, and W. Reed. 2008. "Substitution Is in the Variance: Resources and Foreign Policy Choice." *American Journal of Political Science* 52 (4): 763–773.

Clemm von Hohenberg, B., S. Bernhard, A. S. Cardenal, A. Guess, E. Menchen-Trevino, and M. Wojcieszak. 2024. "Analysis of Web Browsing Data: A Guide." *Social Science Computer Review* 42 (6): 1479–1504. https://doi.org/10.1177/08944393241227868

Cornesse, C., and A. G. Blom. 2023. "Response Quality in Nonprobability and Probability-Based Online Panels." *Sociological Methods & Research* 52 (2): 879–908. https://doi.org/10.1177/0049124120914940

Deville, J.-C., C.-E. Särndal, and O. Sautory. 1993. "Generalized Raking Procedures in Survey Sampling." *Journal of the American Statistical Association* 88 (423): 1013–1020.

Enns, P. K., and J. Rothschild. 2022. "Do you Know where your Survey Data Come from?" Medium. https://medium.com/3streams/surveys-3ec95995dde2

Garner, A., and H. Palmer. 2011. "Polarization and Issue Consistency over Time." *Political Behavior* 33: 225–246.

Gittelman, S., and E. Trimarchi. 2010. "Online Research . . . and all that Jazz!: The Practical Adaption of Old Tunes to Make New Music." https://api.semanticscholar.org/CorpusID:192628845

Greszki, R., M. Meyer, and H. Schoen. 2015. "Exploring the Effects of Removing "Too Fast" Responses and Respondents from Web Surveys." *Public Opinion Quarterly* 79 (2): 471–503.

Holden, C. J., T. Dennie, and A. D. Hicks. 2013. "Assessing the Reliability of the M5-120 on Amazon's Mechanical Turk." *Computers in Human Behavior* 29 (4): 1749–1754. https://doi.org/10.1016/j.chb.2013.02.020

Hopkins, D. J., and T. Gorton. 2024. "On the Internet, No One Knows You're an Activist: Patterns of Participation and Response in an Online, Opt-in Survey Panel." *Political Research Quarterly* 77 (4): 1397–1414. https://doi.org/10.1177/10659129241268818

Huang, J. L., N. A. Bowling, M. Liu, and Y. Li. 2015. "Detecting Insufficient Effort Responding with an Infrequency Scale: Evaluating Validity and Participant Reactions." *Journal of Business and Psychology* 30: 299–311. https://doi.org/10.1007/s10869-014-9357-6

Huff, C., and D. Tingley. 2015. "'Who Are these People?' Evaluating the Demographic Characteristics and Political Preferences of MTurk Survey Respondents." *Research & Politics* 2 (3). https://doi.org/10.1177/2053168015604648

Imbens, G. W., and D. B. Rubin. 2015. *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge: Cambridge University Press.

Ipsos Insight, NPD Group, and TNS. 2006. "Heavier Responders in Online Survey Research." https://www.ipsos.com/sites/default/files/news_and_polls/2006-10/mr061023-1POV.pdf

Jerit, J., and J. Barabas. 2023. "Are Nonprobability Surveys Fit for Purpose?" *Public Opinion Quarterly* 3: 816–840. https://doi.org/10.1093/poq/nfad037

Krupnikov, Y., H. H. Nam, and H. Style. 2021. "Convenience Samples in Political Science Experiments." In *Advances in Experimental Political Science*, edited by J. N. Druckman, and D. P. Green, 165–183. Cambridge: Cambridge University Press

Malhotra, N., and J. A. Krosnick. 2007. "The Effect of Survey Mode and Sampling on Inferences about Political Attitudes and Behavior: Comparing the 2000 and 2004 ANES to Internet Surveys with Nonprobability Samples." *Political Analysis* 15 (3): 286–323. https://doi.org/10.1093/pan/mpm003

Matthijsse, S. M., E. D. De Leeuw, and J. J. Hox. 2015. "Internet Panels, Professional Respondents, and Data Quality." *Methodology* 11 (3): 81. https://doi.org/10.1027/1614-2241/a000094

McGonagle, A. K., J. L. Huang, and B. M. Walsh. 2016. "Insufficient Effort Survey Responding: An Under-Appreciated Problem in Work and Organisational Health Psychology Research." *Applied Psychology* 65 (2): 287–321. https://doi.org/10.1111/apps.12058

Parry, D. A., B. I. Davidson, C. J. R. Sewall, J. T. Fisher, H. Mieczkowski, and D. S. Quintana. 2021. "A Systematic Review and Meta-Analysis of Discrepancies between Logged and Self-Reported Digital Media Use." *Nature Human Behaviour* 5 (11): 1535–1547. https://doi.org/10.1038/s41562-021-01117-5

Rosenzweig, L., P. Bergquist, K. H. Pham, F. Rampazzo, and M. Mildenberger. 2020. "*Survey sampling in the Global South using Facebook advertisements*." SocArXiv Papers.

Shapiro, D. N., J. Chandler, and P. A. Mueller. 2013. "Using Mechanical Turk to Study Clinical Populations." *Clinical Psychological Science* 1 (2): 213–220. https://doi.org/10.1177/2167702612469015

Ternovski, J., and L. Orr. 2022. "A Note on Increases in Inattentive Online Survey-Takers since 2020." *Journal of Quantitative Description: Digital Media* 2: 1–35. https://doi.org/10.51685/jqd.2022.002

Toich, M. J., E. Schutt, and D. M. Fisher. 2022. "Do you Get What you Pay for? Preventing Insufficient Effort Responding in MTurk and Student Samples." *Applied Psychology* 71 (2): 640–661. https://doi.org/10.1111/apps.12344

Valentino, N. A., D. Kirill Zhirkov, S. Hillygus, and B. Guay. 2020. "The Consequences of Personality Biases in Online Panels for Measuring Public Opinion." *Public Opinion Quarterly* 84 (2): 446–468. https://doi.org/10.1093/poq/nfaa026

Whitsett, H. C. 2013. "Understanding 'Frequent Survey Responders' on Online Panels." Nera Economic Consulting. https://www.nera.com/insights/publications/2013/understanding-frequent-survey-responders-on-online-panels.html

Willems, P., R. Van Ossenbruggen, and T. Vonk. 2006. "The Effects of Panel Recruitment and Management on Research Results: A Study Across 19 Online Panels." Presented at ESOMAR Panel Research. http://www.websm.org/db/12/12228/Web%20Survey%20Bibliography/The_effects_of_panel_recruitment_and_management_on_research_results_A_study_across_19_online_panels/

Zhang, B., and S. Gearhart. 2020. "Collecting Online Survey Data: A Comparison of Data Quality among a Commercial Panel & MTurk." *Survey Practice* 13 (1): 1–10. https://doi.org/10.29115/SP-2020-0015

Zhang, C., C. Antoun, H. Y. Yan, and F. G . Conrad. 2020. "Professional Respondents in Opt-in Online Panels: What Do We Really Know?" *Social Science Computer Review* 38 (6): 703–719. https://doi.org/10.1177/0894439319845102

Cite this article: Clemm von Hohenberg, B., Ventura, T., Nagler, J., Menchen-Trevino, E. and Wojcieszak, M., (2025). Survey Professionalism: New Evidence from Web Browsing Data. *Political Analysis*, 1–19. https://doi.org/10.1017/pan.2025.10018