**Digital Advertising During the 2015 Census Test: Impact on Hard-to-Survey Populations**

Matthew Virgile and Nancy Bates

U.S Census Bureau

***Introduction***

As response rates to probability surveys continue to decline, methodologists search for ways to encourage response. This paper reports results from the 2015 Census Test in the designated market area (DMA) of Savannah, Georgia, with particular emphasis on hard-to-survey populations. The test area contained 439,918 housing units identified on the Census Bureau’s master address file (MAF). The main goal of the test was to evaluate the success of driving Internet response through manipulation of a relatively new contact strategy namely, digital advertising.

**The 2015 Savannah Census Test**

The 2015 Census Test exposed unique households to different communications campaign elements including partnership activities, an earned media campaign, a traditional paid media campaign (i.e. broadcast and cable television, radio, print, and out-of-home) and an extensive digital advertising effort.

In the 2015 Census Test, there were three different digital ad types: Pre-registration, Awareness, and Completion. Pre-registration advertisements encouraged users to pre-register for the test census in advance of the questionnaire becoming available. Awareness ads consisted of 30- and 15-second video ads. The primary objective of awareness ads was to promote awareness of the test, though participants also could navigate to the test’s landing page through each video ad. Completion ads were designed to drive users to the landing page to begin the survey.

To distinguish between different sources of Internet traffic, the team used unique links for the digital advertisements, partnership events, email contacts, social media, postal mailings, text messages, and other media. This allowed us to analyze the source of entry for online submissions. The 2015 Census Test overlaid three experimental designs to test the effects and interactions between digital advertising and mailings. The following section describes these designs.

**Digital Advertising Panels**

The 2015 Census Test included a digital advertising experiment with two fully crossed factors. As shown in Table 1 (tables and figures found in Appendix), one manipulation was per-household spending level while the second manipulation compared general versus targeted digital advertising. A fifth panel served as a control and did not receive digital advertising. The 106 non-P.O. box ZIP codes in the Savannah DMA were divided into five digital advertising panels and all households in the DMA were assigned to one of the five. Each digital advertising panel received different targeted advertising strategies and per-household spending levels.

*For purposes of this paper, we focus more broadly on the impact of digital advertising on hard-to-survey populations rather than the spend level and targeted advertising experiments.*

**Mailout Panels**

The 2015 Census Test also incorporated mailings to a subset of households in the Savannah media market. The agency sampled 90,000 households to be mailed one of three mail strategies. Selected addresses were assigned to either a pre-registration mail strategy, mailings that contained a unique geocoded identifier, or mailings that contained no such identifier (non-ID cases). In all three of the mailings, households were first sent instructions for completing the test form online, without a paper form. If households did not respond online by the fourth contact, a paper questionnaire was mailed (see Table 2). For purposes of this paper, we combine all 90,000 households sampled into *any* of the three mailing strategies. To understand the interactions between mailings and advertising, we separate the majority of findings in our paper into households that received some type of mailing versus households that did not. The former represents a closer strategy to the one currently planned for the 2020 Census, while the latter allows us to study response behavior under a social marketing campaign devoid of any direct mail contacts.

***Operationalizing “Hard-to-Survey” in the Savannah Test***

Recently, the Census Bureau developed a new hard-to-survey metric known as the Low Response Score, or LRS (Erdman and Bates, in press). The score is simply a geographical area’s fitted value from a regression model predicting the area’s mail return rate from the 2010 Census. In all, 25 predictor variables were used as input to develop the LRS[[1]](#footnote-1) . To operationalize hard-to-survey areas for the Savannah 2015 Census Test, we used the LRS as a guide. To define hard-to-survey tracts, we performed an exploratory factor analysis. We input the 25 LRS variables contained on the 2014 Census Planning Database, a tract-level public database containing selected variables from the 2010 Census and 2009-2013 estimates from American Community Survey. We then retained only those LRS predictor variables that had *positive* coefficients in the model and sufficiently high factor loadings. This left 11 variables for input to the final factor analysis.

We classified tracts as hard-to-survey if the factor score was in the top 10 percent. We labeled households located in tracts associated with high factor scores on Factor 1 as “Young, Mobile, Renters” and households aligned with high factor scores on Factor 2 as “Female-Headed, Low Income/Education” (see Table 3). While the LRS served as a useful guidepost to classify hard-to-survey areas, the score predicts only one mode of self-response -- mail. One of the main Savannah Test objectives was to encourage online response. Consequently, we defined a third geography reflecting households with lower prevalence of home Internet connectivity. We leveraged data from the 2013 Federal Communication Commission (FCC) indicating the number of households per 1,000 connected to residential high speed Internet. We chose to define tracts containing between zero and 400 connected households per 1,000 as exhibiting “low Internet connectivity”.

This classification method isolated three distinct groups of tracts containing a high proportion of households correlated with lower self-response. It also provided a denominator of hard-to-survey addresses we could determine to be respondents or nonrespondents in the test. Figure 1 illustrates the location and distribution of the hard-to-survey tracts.

***Results***

First, we examine the impact of mailings on response rates, among households in each hard-to-survey tract type and among all households. The response rates are displayed in Table 4. We find that households in mailout panels were many times more likely to respond than households not in mailout panels, regardless of whether they were in a hard-to-survey tract or not (overall, 49.0 percent versus only 7.5 percent). Thus, mailing materials were highly effective for all households regardless of geography. For households in mailout panels, those in any type of hard-to-survey tract had lower response rates than the overall rate. Households in “Young/Mobile/Renter” areas and in “Female-Headed, Low Income/Education” areas had similar response rates (37.2 percent versus 37.0 percent), those in “Low Internet Connectivity” areas had the lowest response rate (32.9 percent), and all of these are lower than the overall mailout response rate (49.0 percent).

For households not in mailout panels, those in two of the three hard-to-survey tract types - “Young/Mobile/Renter” and “Low Internet Connectivity” - had similar response rates (5.0 percent and 4.7 percent, respectively) that were lower than the overall response rate (7.5 percent). One unexpected finding is that households in the “Female-Headed, Low Income/Education” tracts had a response rate (8.3 percent) close to the overall rate (7.5 percent).

For those households that responded, Table 5 provides the mode distributions – the percent responding by Mail, Internet, or telephone questionnaire assistance (TQA) in each of the hard-to-survey tract types and overall. For households in mailout panels, Internet was by far the most popular mode, regardless of whether households were in any hard-to-survey tracts or not. Internet was especially popular among households in “Young/Mobile/Renter” tracts (74.3 percent), a few points higher than the overall proportion of Internet respondents (69.0 percent). This hard-to-survey tract type was the most similar in mode distribution to the overall mode distribution. The percent reporting by Internet was lower for households in “Female-Headed, Low Income/Education” and “Low Internet Connectivity” tracts but still the preferred option (55.2 percent and 53.7 percent, respectively), and the remainder reported by mail or TQA. Of interest to note, of those in a mail panel, two of the three hard-to-survey groups (Female-Headed, Low Income/Education and Low Internet Connectivity) had above average responses via telephone questionnaire assistance[[2]](#footnote-2) (17.3 percent compared to 11.8 percent overall). The latter group stands to reason as many households lack internet as a response option.

Next, we turn attention to the subset of households that responded online. Here, we are primarily interested in the source of online entry and who responded due to advertising versus a direct-mail piece. Table 6 lists the *source* households used to find the online form. For those in mailout panels, a URL to access the survey was part of the mailed materials. In our dataset for analysis, a blank entry for the source URL variable usually represented responses using the URL from the mailed materials. As expected, blank entries were the predominant source overall (83.7 percent), as well as for households in each of the “hard-to-survey” tract types. The rate is similar for “Young/Mobile/Renter” areas (86.8 percent) and “Low Internet Connectivity” areas (84.2 percent) but lower for “Female-Headed/Modest Income” areas (69.1 percent). The remainder entered through digital ads or a URL found on traditional ads (e.g. radio, TV, billboards), and very few entered due to other sources. About one-quarter of responses from mailout household located in the “Female-Headed, Low Income/Education” areas came in by a traditional ad. This is far higher than either of the other two hard-to-survey groups.

For households not in mailout panels, the URL found on traditional ads was the dominant source of entry overall (67.1 percent). Again, this was most prevalent among households in “Female-Headed, Low Income/Education” areas (72.0 percent). The second most common source of entry overall was digital ads (23.8 percent). This was most popular with those in “Young/Mobile/Renter” areas (26.2 percent). Overall, 7 percent of online submissions from non-mailout households responded using a URL found in a mail piece. The majority of these were cases that responded by pre-registration – a household member saw a pre-registration advertisement, signed up to receive a text or email message, and then were sent a message containing the URL used in some of the mailings.

Digital ads can be served by various platforms –as banner ads on websites, as displays on social media sites such as facebook, or as ads displayed as a result of search engine inquiries (e.g. Google or Bing). The distribution of ad platforms resulting in web survey responses across the three hard- to-survey areas indicated a bit of variation, but the same rank ordering (table 7). That is, for all three groups Social – Facebook Posts (ads that appear in the Facebook Newsfeed on desktop and mobile devices) were most successful followed by Google search ads; followed by Facebook Display ads (ads that appear in the right rail of Facebook on desktop devices); followed by Bing search ads. This confirms that ads placed on social media and as a result of self-initiated Census-related searches are much more effective than ads targeted to sites based on user demographics and search histories. However, all four platforms seem to work equally well among the different hard-to-survey groups.

***Conclusions***

For the 2020 Decennial Census, the Census Bureau plans to expand the use of digital advertising as a means to raise awareness and encourage online response. Recent studies indicate that smartphone adoption by racial and ethnic minorities is approaching or even surpassing that of whites and that younger age groups have the highest rate of smartphone ownership (Rainie, 2016). Since these subpopulations have historically lower rates of Census self-response, digital ads present a potentially powerful new communication medium.

The purpose of our paper was to explore digital advertising’s impact among three different hard-to-survey populations: (1) Young, renter households that move frequently, (2) Female-headed households with low education and modest income and (3) areas with below-average internet connection at home.

As expected, the three areas designated as hard-to-survey had lower response rates compared to the overall Savannah DMA. Overall, the mailout panel response rate was 49.0 percent compared to 37.2 percent, 37.0 percent and 32.9 percent for the Young/mobile/renters, female-headed/low income, and low internet areas, respectively.

Overall, 69 percent of the mail panel households that responded to the test did so using the internet mode. This rate was higher among households located in the young/mobile/renter tracts where close to three-quarters (74 percent) responded by internet. This is perhaps not surprising given the usage of technology and smartphones by younger cohorts. However, given this group has historically low self-response it’s encouraging to see high online uptake. Alternatively, among responding households in the other two hard-to-survey areas, just over half selected the internet option (55 percent in female-headed areas and 54 percent in areas with low internet connectivity). Of note, however is that both of these areas had above average selection of the telephone questionnaire assistance (TQA) option. For both groups, close to 2 in 10 (17 percent) opted to call the toll free number and provide their response over the phone. This is higher than the overall DMA TQA response rate of ~12 percent. For these populations, the phone still serves as an important mode, presumably in households that do not have internet connection and do not want to wait for the paper form under an internet –push implementation.

Of the overall online submissions received, nearly half (49.2 percent) were directly attributable to the advertising campaign – 35.5 percent to traditional advertising and 13.7 percent to digital ads. Among the mail panel households, 10.3 percent of internet responses were directly attributable to traditional ads and 5.8 to digital ads. For our three hard-to-survey areas in the mail panels, results were similar to overall patterns with the exception of the female-headed, low education/income tracts. For these areas, one-quarter (24.9 percent) of the internet responses were attributable to traditional advertising (e.g. television ads, radio and billboards) – a lower proportion of these area’s responses were attributable to the URL provided in the mail materials. This suggests that *traditional advertising* plays an important part to encourage response among this particular group.

Digital ads were delivered across various platforms including ads on social media and search engine ads. For all three hard-to-survey groups, social media post ads had the largest click-resulting-in-submission rate followed by Google search ads; followed by Facebook Display ads (ads that appear in the right rail of Facebook on desktop devices); followed by Bing search ads.

***References***

Erdman, C. and Bates, N. (in press). The Low Response Score: A Metric to Locate, Predict, and Manage Hard-to-Survey Populations. *Public Opinion Quarterly.*

Rainie, L. (2016). Digital Divides 2016. Pew Research Center. <http://www.pewinternet.org/2016/07/14/digital-divides-2016/>

Appendix

**Table 1. Description of Digital Advertising Panels and Treatments**

|  |  |  |
| --- | --- | --- |
|  | **Targeted** | **Nontargeted** |
| **High spend level** (per household) | **Panel A**  General digital advertising +  Targeted advertising | **Panel B**  General digital advertising |
| **Low spend level**  (per household) | **Panel C**  General digital advertising +  Targeted advertising | **Panel D**  General digital advertising |
| **No spend** | **Panel E**  No digital (control condition for the other communication activities) | |

**Table 2. Mail Contact Strategy Panels**

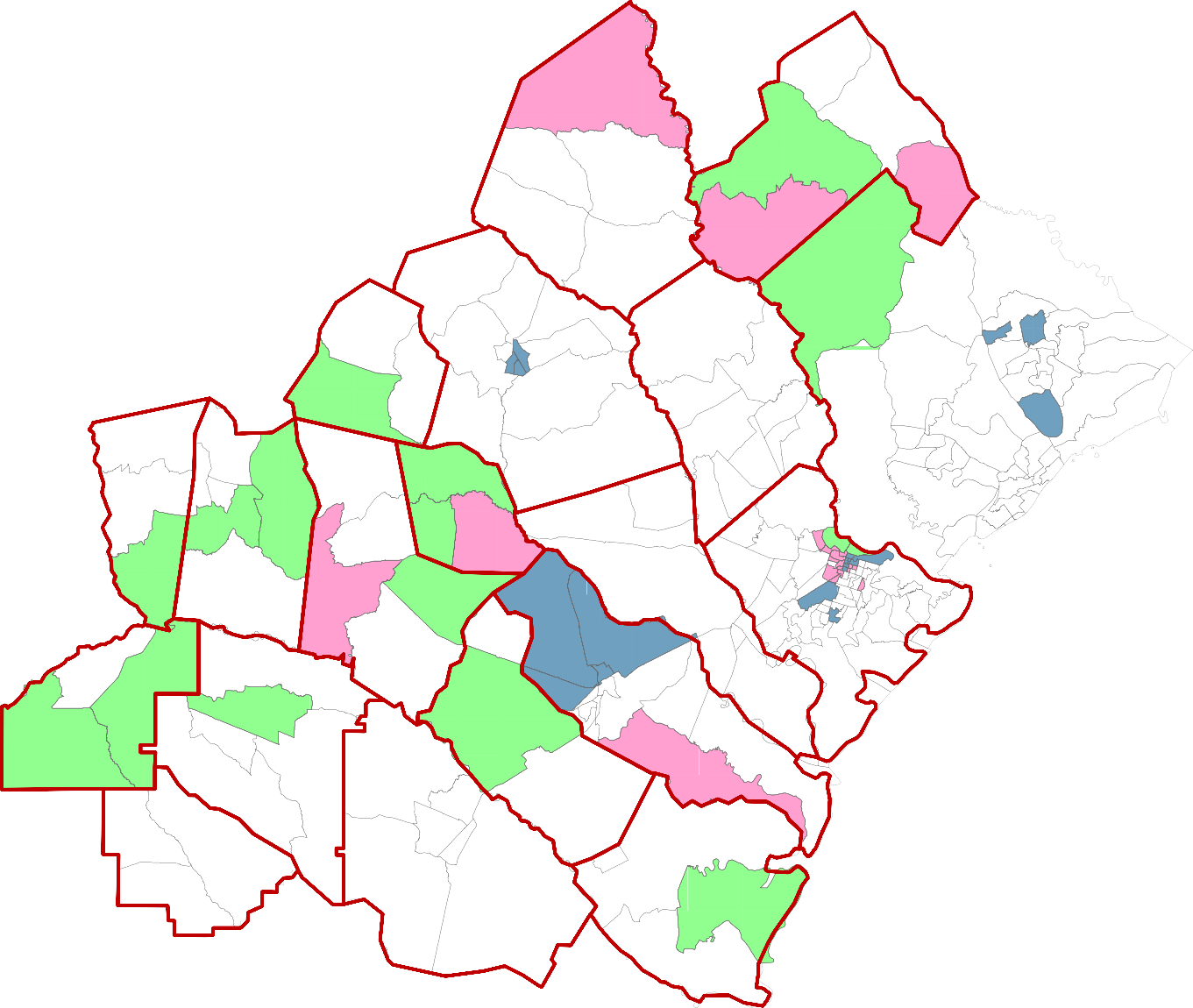
| Mail Panel | Sample  Size | Pre-Notice (Feb 23) | Contact 1  (March 23) | Contact 2  (April 1) | Contact 3  (April 8) | Contact 4  (April 15/16) |
| --- | --- | --- | --- | --- | --- | --- |
| “Notify Me” Pre-registration mailing panel | 30,000 | Postcard invite to register | Choice of email/text\* | Choice of email/ text\* | Choice of email/text\* | Mail  questionnaire |
| Non-ID mailing panel | 30,000 | N/A | Letter  (no ID) | Postcard  (no ID) | Postcard  (no ID) | Mail  questionnaire |
| ID mailing panel | 30,000 | N/A | Letter | Postcard | Postcard | Mail  questionnaire |

*\*Choice of email or text sent to households that pre-registered. Otherwise, addresses were sent mail materials identical to the ID panel.*

**Table 3. Factor Loadings from Selected LRS Predictor Variables**

|  |  |  |
| --- | --- | --- |
| Census Tract-level variable | Factor 1: Young/Mobile/Renters | Factor 2: Female-Headed, Low Income/Education |
| % age 5-17 | -44 | 25 |
| % age 18-24 | 84 | 9 |
| % < high school graduate | -23 | 80 |
| % persons below poverty | 44 | 74 |
| % female-headed, no husband | 16 | 80 |
| Mean number persons per hhd | 8 | 17 |
| % hhds with related child <6 years old present | 48 | 23 |
| % who moved in last year | 85 | -4 |
| % renters households | 86 | 31 |
| % vacant households | -23 | 23 |
| % Non-hispanic, Black | 15 | 78 |
| *Eigenvalue* | 3.85 | 2.26 |

**Figure 1. Location of Tracts Scoring High on Hard-to-Survey Factors**



Young mobile renters

Low internet

Female-headed low educ/income

**Table 4. Response Rates for Households in Mailout Panels Versus Non-Mailout, by Hard-to-Survey Tract Type and Overall**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Mailout Panels (ID, non-ID, Pre-Reg)** | | | | **Non-mailout** | | | |
| **Young/ Mobile/ Renter** | **Female-Headed, Low Inc/Ed** | **Low Internet Connectivity** | **Overall** | **Young/ Mobile/ Renter** | **Female-Headed, Low Inc/Ed** | **Low Internet Connectivity** | **Overall** |
| Response Rate | 37.2 | 37.0 | 32.9 | 49.0 | 5.0 | 8.3 | 4.7 | 7.5 |
| Total | 4,818 | 3,844 | 3,621 | 90,000 | 19,746 | 18,272 | 25,497 | 319,918 |

**Table 5. Mode of Response Distributions for Households in Mailout Panels Versus Non-Mailout, by Hard-to-Survey Tract Type and Overall**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Mode** | **Mailout Panels (ID, non-ID, Pre-Reg)** | | | | **Non-mailout** | | | |
| **Young/ Mobile/ Renter** | **Female-Headed, Low Inc/Educ** | **Low Internet Connectivity** | **Overall** | **Young/ Mobile/ Renter** | **Female-Headed, Low Inc/Educ** | **Low Internet Connectivity** | **Overall** |
| Mail | 18.7 | 27.5 | 29.1 | 19.2 | 0 | 0 | 0 | 0 |
| Internet | 74.3 | 55.2 | 53.7 | 69.0 | 98.8 | 98.9 | 99.2 | 99.2 |
| TQA | 7.0 | 17.3 | 17.3 | 11.8 | 1.2 | 1.1 | 1.0 | 0.8 |
| Total N | 1,792 | 1,424 | 1,191 | 44,064 | 991 | 1,508 | 1,190 | 23,849 |

**Table 6. Source of Entry for Online Responses in Mailout Panels Versus Non-Mailout, by Hard-to-Survey Tract Type and Overall**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Mailout Panels (ID, non-ID, Pre-Reg)** | | | | **Non-mailout** | | | |
| **Young/ Mobile/ Renter** | **Female-Headed, Low Inc/Educ** | **Low Internet Connectivity** | **Overall** | **Young/ Mobile/ Renter** | **Female-Headed/ Low Inc/Educ** | **Low Internet Connectivity** | **Overall** |
| URL from mail pieces & pre-registration | 86.8 | 69.1 | 84.2 | 83.7 | 8.3 | 6.2 | 6.9 | 7.0 |
| URL from traditional ads (e.g. TV, radio, print) | 8.5 | 24.9 | 10.6 | 10.3 | 62.4 | 72.0 | 67.1 | 67.1 |
| Digital ads | 4.4 | 5.6 | 4.4 | **5.8** | 26.2 | 19.2 | 23.5 | **23.8** |
| Partners/events | 0.1 | 0.3 | 0.3 | 0.1 | 1.3 | 1.3 | 0.7 | 1.0 |
| Email (pre-reg) | 0.2 | 0.1 | 0.5 | 0.1 | 1.3 | 0.5 | 1.4 | 0.8 |
| Other | 0 | 0 | 0 | 0 | 0.3 | 0.3 | 0.1 | 0.1 |
| Total | 1,331 | 786 | 639 | 30,391 | 979 | 1,491 | 1,181 | 23,649 |

**Table 7. Distribution of Ad Platforms Resulting in Completed Survey among Households Responding via Digital Ads – Hard-to-Survey Tract Type and Overall**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Ad Platform** | **Young/Mobile/ Renter** | **Female-Headed, Low Inc/Ed** | **Low Internet Connectivity** | **Overall** |
| Social – Facebook Post | 59.5 | 43.6 | 52.4 | 42.6 |
| Search – Google | 24.3 | 23.8 | 18.8 | 28.7 |
| Social – Facebook Display | 8.1 | 12.9 | 14.7 | 13.0 |
| Search – Bing | 5.4 | 12.9 | 8.8 | 10.6 |
| Display – Banners | 2.7 | 4.0 | 3.5 | 4.2 |
| Rich Media – Video Pre-roll | 0 | 0 | 1.2 | 0.7 |
| Content Marketing | 0 | 3.0 | 0.6 | 0.3 |
| Display – Facebook Exchange | 0 | 0 | 0 | 0 |
| Total | 74 | 101 | 170 | 2,962 |

1. The 25 variables included % age 5-17; % age 18-24; % not HS graduate; % below poverty; % female headed hhd, no husband; Mean number per hhd; % related child <6; % moved in last year; % renter occupied; % vacant houses; % non-Hispanic, Black. % married family hhds; median household income; median house value; % Hispanic, population density; % non-Hispanic White; % age 65+; % males, % college graduates; % moved in last 5 years; % single family units; % single person hhds; [↑](#footnote-ref-1)
2. Telephone questionnaire assistance represents responses captured from calls made to a 1-800 found in the mail materials. [↑](#footnote-ref-2)