Survey Modes and Survey Experiments: A Comparative Analysis between RDD Cell Phone Survey and Opt-In Online Survey*

Takeru Eto† Ikuma Ogura‡ Masahiro Zenkyo§

07/02/2024

Abstract

This paper examines the impact of survey mode on response distribution and experimental outcomes in Japan, emphasizing the comparison between an opt-in online survey and a random digit dialing (RDD) cell phone survey. Results from online surveys are often criticized for their lack of generalizability. Likewise, results of online survey experiments are often questioned with regard to their external validity. Notably, online survey respondents are not randomly selected from the population. Moreover, it is not quite clear how experimental and other results differ between those who participate in opt-in online surveys and those who are randomly selected from the population. In particular, the effects of survey mode—online surveys and RDD phone surveys—have not been sufficiently studied in Japan compared to other countries. In this study, we analyze the effects of different survey modes on response distribution and experimental results. For this purpose, we have developed an original mobile phone survey system using RDD and interactive voice response (IVR). We compare the results of an online survey with those of an RDD cell phone surveys to examine the differences in response distribution and experimental results between the two methods.

*Paper prepared for the WAPOR 77th and WAPOR Asia Pacific 7th Joint Annual Conference (July 31, 2024; Seoul, South Korea). This work was supported by JSPS Grant-in-Aid for Scientific Research (Grant ID: 23H05426). ***Please do not cite or circulate without permission of the authors.***

†JX Press, Email: t.eto@jxpress.net

‡Hitotsubashi University, Email: ikuma.ogura@r.hit-u.ac.jp

§Kwansei Gakuin University, Email: masahirozenkyo@kwansei.ac.jp

1 Introduction

In contemporary developed countries, extensive studies are underway to develop alternative approaches to replace traditional survey methods for comprehending public attitudes and behaviors, as well as assessing the validity of these methods. This is because conducting face-to-face interviews with randomly selected individuals is costly, and recent shifts in the information environment due to the spread of the Internet and the transition of personal devices from landlines to mobile phones have made it possible to use alternative and new survey methods (Olson et al., 2021). The COVID-19 pandemic has accelerated this transition (Coffey et al., 2024; Sakshaug et al., 2020). While some comparative survey projects (e.g., the Comparative Study of Electoral Systems project) continue to conduct face-to-face interviews with randomly selected individuals, many scholars have shifted their focus from single-frame to methods combining multiple survey modes with new online survey systems using the text messaging functions of mobile phones (Asimov and Blohm, 2024; Conrad et al., 2017; Kim and Couper, 2021; Kocar, 2022; Olson et al., 2021; Seol et al., 2023).

Public opinion polls by the mass media and academic surveys in Japan are also in a transitional phase. Until the early 2000s, academic surveys in the field of Japanese social sciences generally adopted face-to-face interviews with individuals randomly selected from electoral rolls or basic resident registers. However, since the 2010s, empirical studies using online surveys targeting monitors registered with research firms have increased rapidly, owing to their cost-effectiveness. The development of new survey systems is a pressing concern for public polling in Japan. Most Japanese media have used a random digit dialing (RDD) system targeting both landlines and mobile phones for public polling; however, they are compelled to seek new approaches to collect large amounts of data effectively because the costs associated with their current survey methods have escalated to a level that seriously affects the financial situation (Eguchi, 2022).

These suggest that the Japanese mass media and academic scholars face a trade-off between representativeness and cost-effectiveness. They have explored the possibility of

using online surveys based on non-probability panels, because of the increasing difficulty of conducting surveys based on traditional methods. However, representativeness of opt-in online-based survey is low and its coverage error is large (Couper et al., 2007; Couper and Miller, 2008). One way to address these problems is to make a probability-based panel and conduct surveys with registered panelists.[1] Although there are some trial studies that have built such panels for academic web-based surveys in Japan (e.g., Hirasawa et al., 2023), the cost makes it challenging for many scholars. Additionally, repeated surveys on the same panel may result in the problem of survey fatigue (Jeong et al., 2023; Porter et al., 2004).

In this study, we propose a new RDD system targeting mobile phones that can also be used for academic survey research. Our RDD system also targets randomly generated phone numbers, similar to general RDD. However, the target was limited to mobile phones. According to the 2023 White Paper Information and Communications in Japan by the Ministry of Internal Affairs and Communications, the installation rate of landlines in Japanese households has been declining since approximately 2010, falling below 70% in 2018. On the other hand, the White Paper also reports that a) the penetration rate of mobile phones exceeded 90% in 2009, b) it has remained high until this day, and c) the ownership rate of smartphones has increased sharply since 2010 (87% as of 2020). These suggest that there were almost no coverage errors when the target was limited to mobile phones. While Peytchev and Neely (2013) pointed out that the estimated bias from removing landline phones was not negligible, a recent study encouraged a switch from RDD to a single-frame cell phone design (Soullier et al., 2022).

The two main features of our RDD are as follows. First, the survey is conducted by machine voice automatically ("Robocalls") rather than by human operators. This significantly reduces personnel costs compared with the general RDD. Second, our RDD is able to pay respondents incentives without obtaining personal information. As outlined in Section 2, we developed a system that pays respondents points as an incentive through a dedicated

[1]There is no research firm that has a probability-based sample for online surveys in Japan. The efforts of Japanese academic scholars to create open probabilistic panels have not progressed sufficiently.

site. This makes it possible to use our RDD for academic surveys where incentive payments are often required by Institutional Review Boards. Because most Japanese RDD do not adopt an incentive payment system, operators need to collect personal information, such as respondents' addresses, when incentive payments are required; However, in our RDD, there is no need to collect any personal information.

There have been no previous studies on automated robocall surveys for cell phones with payment incentives in Japan. Thus, the quality of the data obtained from our RDD study is unclear. To examine our data quality, we compared the results of our RDD with an opt-in web survey. Specifically, we compared the distributions of respondents' gender, age, and education obtained from the web survey, in which we used quota sampling to match the marginal distributions of demographics with the census with those obtained from our RDD. Furthermore, we demonstrated the reproducibility of the survey experimental results in the previous studies.

The remainder of this paper is organized as follows. Section 2 describes the details of the RDD process. Section 3 explains the research design used to compare the data quality collected by our RDD with that of the web survey, and Section 4 summarizes the results. In Section 5, we discuss the possibilities of our RDD and issues to be addressed to improve the system.

2 New RDD System for Mobile Phones in Japan

To conduct surveys of randomly selected individuals, we developed an original RDD system for cell phones. By integrating an incentive payment system with a randomly generated cell phone number system, we can not only conduct probability-based sample surveys but also pay incentives to those who provide complete answers. Additionally, we have adopted an automated system using machine voice in our RDD system rather than using a human operator to make the calls. This method has the advantage of collecting large amounts of

data quickly and inexpensively. These two systems have been developed and operated by JX Press,[2] a Japanese news agency founded in 2008. JX Press has provided several services using these systems since 2020, making it possible for us to design an original RDD system.

An outline of our RDD system is shown in Figure 1. The our RDD system consists of the following four steps. First, random cell phone numbers are generated. Second, we conduct an automated survey using the data. Third, we send an SMS with incentive payment information to the respondents who have completed the survey. Fourth, incentive points are paid through a dedicated incentive payment website.

Generating Phone Numbers

To conduct the survey, we first generate the phone numbers for the calls. The number-generation mechanism used was developed by JX Press prior to this system. Cell phone numbers are generated using the Japanese format +81-A0-BCDE-FGHI, ensuring the format validity of the number combinations by adhering to specific rules (e.g., A ranges from 7 to 9, B is not 0, and some BCDE combinations are currently not in use). Because the active status of phone numbers is not verified before dialing, there is a possibility of calling numbers that are not assigned to the users.

Conducting Survey

Our RDD survey is conducted using an Interactive Voice Response (IVR) system. Using the Communication Platform as a Service (CPaaS) provided by Vonage, calls are made automatically to randomly generate numbers through a Web Application Programming Interface (API).[3] Respondents hear a voice message explaining the survey's purpose, incentives to be

[2]The company is a startup primarily targeting the media sector and has secured investments from major Japanese media companies. In 2019, the company joined the Japan Association for Public Opinion Research (JAPOR), which consists of major media companies conducting public opinion polls in Japan.

[3]CPaas does not call all requested numbers; some numbers may be rejected by the system or calls cannot be made because of errors. Additionally, if the owner of a mobile phone rejects calls from the CPaas, calls cannot be made through this system.

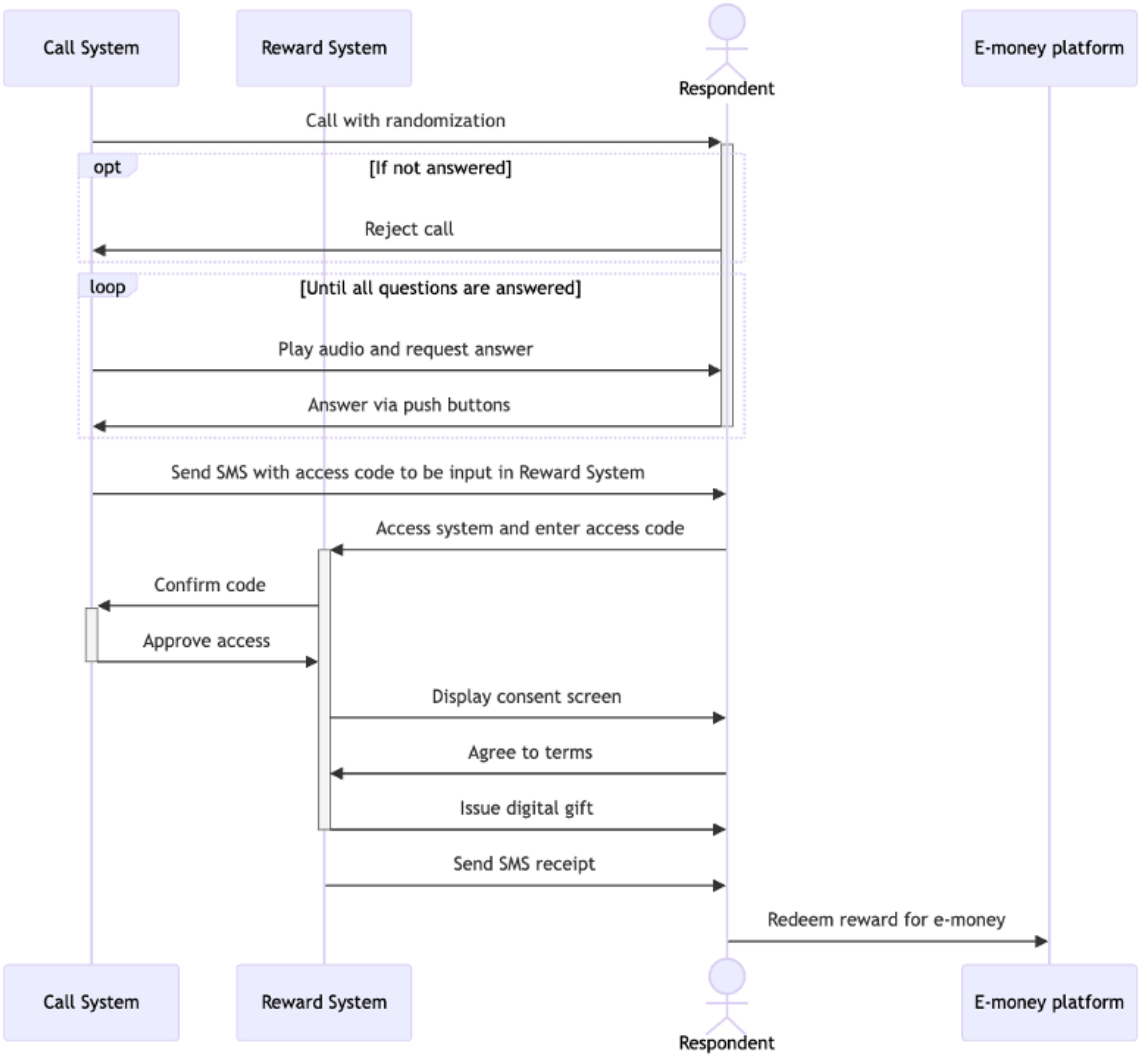

Figure 1: Process of Our RDD System

paid, and the number of questions before participating in the survey. They answered the questions by pushing button[4] under the guidance of a prerecorded machine voice created using a commercial API. Those who completed all questions automatically received an SMS as an incentive payment.

The other details of the survey procedure are as follows. 1) The numbers used to conduct the survey are the same as those used in the public polling conducted by JX Press in collaboration with other mass media. 2) The caller ID is not hidden; thus, respondents

[4]When the number of options for each question exceeds 10, respondents are required to answer using button #. For example, they would press '8#' for option number 8, and '11#' for option number 11.

are notified of caller's phone number. 3) The survey is conducted from 9am to 10pm on weekends and holidays and from 6pm to 10pm on weekdays. 4) Each phone number is called only once.

Payment Incentives

Respondents who completed all the questions received an access code to obtain incentives via SMS, which was sent via Vonage. An example of this message is as follows.

> *This message is from JX Press. Thank you for participating in our survey. As a token of our appreciation, we would like to offer you a digital gift worth *** JPY. Please search for "JX Press Academic Research Survey" within the next three days and enter the code ********* on the dedicated website. Due to the prevalence of SMS-based fraud, we are not providing a URL link directly.*

The access codes are ten alphanumeric characters long, excluding easily confusable characters such as 0, O, o, I, i, 1, L, l, 2, Z, 5, S, 8, and B. To mitigate the risk of the message being mistaken for SMS-based fraud, we do not include a URL link in the SMS. Instead, we instruct them to use a search engine (e.g., Google) to find our website.

Respondents can receive digital gifts by accessing our dedicated website[5] and entering their access codes and mobile phone numbers. When the entered information matches the database records, e-money that can be exchanged for one of the six options widely used in Japan (e.g., Amazon gift cards, PayPay points, or Google Pay) is issued to the respondent through an external platform. An SMS receipt is also sent to the respondent if the e-money is issued.

To convey persuasively that this website is for incentive payments, we display the phone number used to conduct the survey on the website. The privacy policy of the survey and the contact form for respondents are also provided on this website.

[5]https://academic-research.jxpress.net/.

3 Research Design

To evaluate the quality of data obtained through our RDD system, we compare the results of our RDD survey with those from an opt-in web survey conducted around the same time as the RDD survey. Specifically, we analyze the differences in the distribution of the three demographic variables and the outcomes of the survey experiment between our RDD and web surveys.

Data

Between March and April 2024, we conducted two surveys: an RDD survey based on our original system and an opt-in web survey. The RDD survey for cell phones was conducted between March 2 and March 5, 2024. For this survey, approximately 350,000 mobile phone numbers were randomly generated, 255,927 of which were called. Prior to answering the survey questions, participants listened to a brief recorded message explaining the survey's purpose, number of questions, and incentives offered. Following this message, 1,494 individuals started the survey. Upon completion, each participant was asked to provide permission to use their responses in academic research. Finally, we obtained complete data from 978 participants after excluding 54 who refused to provide permission to use their responses. The valid response rate is approximately 0.38%.

The opt-in web survey was conducted on April 20, 2024. We recruited 1,045 participants for our web survey via Purespectrum, using quota sampling to ensure that the marginal distributions of gender (male or female), age (18–24, 25–34, 35–44, 45–54, 55–64, or 65+ years), and region (Chubu, Chugoku, Hokkaido, Kansai, Kanto, Kyusyu, Shikoku, or Tohoku) were balanced with the target population based on the 2020 census. The design of the web survey, including the number of questions, order, wording, and response options, was identical to that of the RDD survey. However, there was a significant difference in the incentives between the RDD and web surveys, as the cost per incident (CPI) in the web survey was set at 0.5

USD per person (approximately 80 JPY at the time of this survey).

Comparison of the Distributions of Demographic Variables

First, we examined how the distributions of gender, age, and education level, which are the key demographic variables correlated with various social attitudes, differed between the RDD and web surveys. As mentioned above, the distributions of gender and age in the web survey were adjusted to align with the census data, suggesting minimal bias in these variables. Conversely, no such adjustment has been made for education level; thus, we anticipate some bias for this variable, as opt-in web surveys in Japan often exhibit a higher proportion of respondents with advanced educational backgrounds in practice.

By comparing the distribution of these three variables between the RDD and web surveys, we evaluated the quality of the data obtained from the RDD survey in terms of distributional bias. The response rate (approximately 0.4%) for the RDD survey was not high; hence, some bias is expected among respondents. However, if there are no significant biases in the gender and age distributions of the data obtained from the RDD, despite not employing allocation recovery as in the web survey, we can conclude that the data from the RDD are highly valid. Additionally, if the distribution of education does not skew toward higher education levels as it does in web surveys, the data obtained from the RDD can be considered more representative than those from opt-in web surveys.

Comparison of the Results of Survey Experiments

Second, we examined whether the experimental results obtained from web surveys in prior studies could be replicated in our RDD. Telephone surveys generally convey experimental stimuli through auditory rather than visual information. Auditory information may enhance respondents' survey satisficing tendencies (Krosnick, 1991) to the same extent, or even more than visual information. Although the results of survey experiments are not significantly affected by coverage errors (Moniz et al., 2024), survey satisficing affects experimental out-

comes (Alvarez et al., 2019). Whether or not our RDD yields results comparable to those of previous studies is a crucial benchmark for assessing the validity of our RDD.

Specifically, we included an experiment on preferences for different types of immigrants in both the RDD and web surveys. Previous studies have consistently demonstrated that immigrants with high skills are generally preferred to those with low skills, both in Japan and abroad (e.g., Hainmueller and Hiscox, 2010; Igarashi et al., 2022; Valentino et al., 2019). Thus, we deemed this immigration experiment suitable for testing experimental reproducibility. The question for the experiment (vignette) was as follows: "Do you agree or disagree with the opinion that the Japanese government should allow [low-/high-skilled] immigrants to come and live in this country more than ever before?" The response options were provided on a 7-point Likert scale ranging from 1 (strongly agree) to 7 (strongly disagree). Either "low-skilled" or "highly skilled" was randomly presented to the respondents as text (web) or voice (RDD).

4 Results

Distribution of Demographic Variables

Figure 2 summarizes the distribution of the respondents' gender (top left panel), age (top right), and education (bottom left) from our RDD and web surveys. First, a comparison of the distribution of respondents' gender shows that participants in our RDD survey were heavily skewed toward men. Although slightly more than half of the web survey respondents, for which we used census-based quotas to recruit participants, were male, over 70% were male. Although not as stark as that of sex, the distributions of age and education also differed between the RDD and web surveys. Compared with the quota-sampled web survey, the RDD survey respondents were skewed toward younger generations. In addition, the RDD respondents were more likely to be educated than those in the web survey. Given that opt-in web survey respondents in Japan tend to be more educated than the country's

population, this pattern suggests that the participants in our RDD survey were even more skewed in terms of education level. We also conducted chi-squared tests to examine whether the distributions of these three variables were statistically different between the RDD and web surveys and found that the null hypotheses of distributional equivalence were rejected ($p < 0.01$).

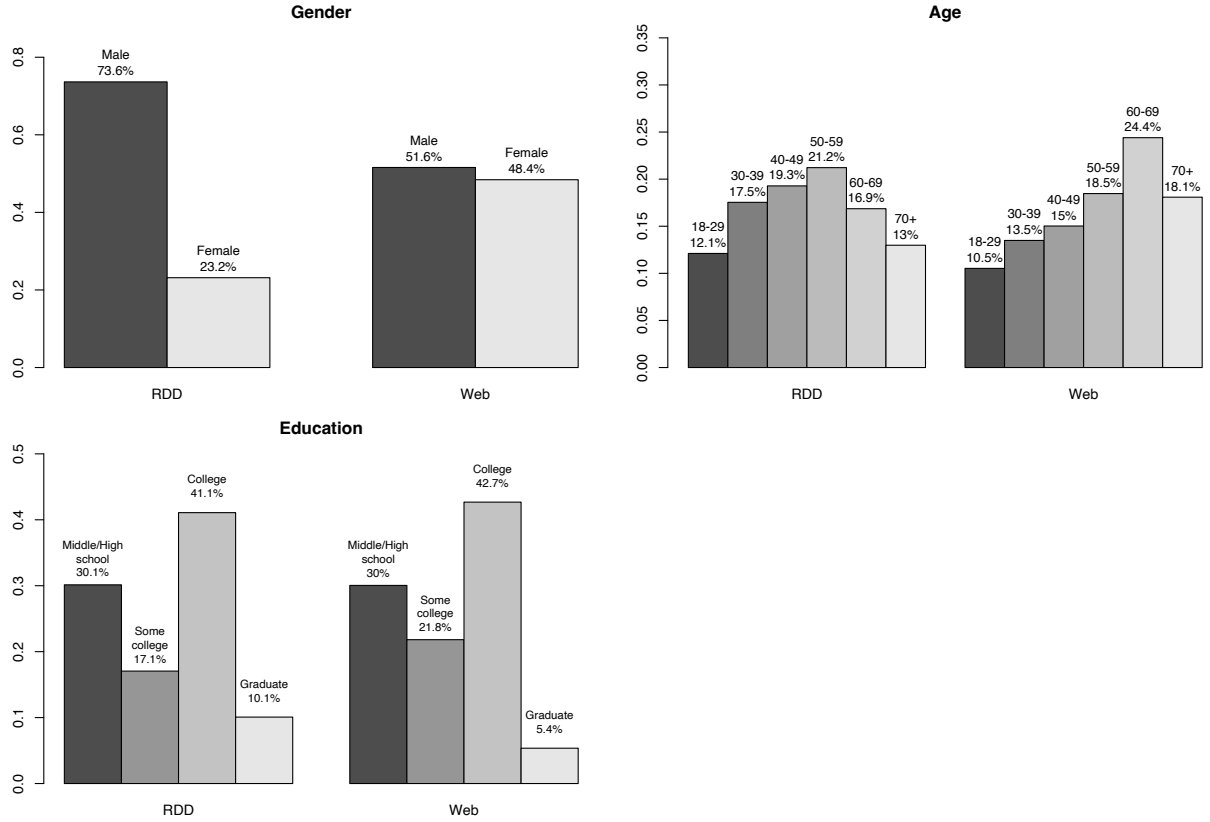

Figure 2: Distributions of Demographic Variables from the RDD and Web Surveys

Survey Experiment

Figure 3 shows the average differences in respondents' disagreements between high- and low-skilled immigrants coming to Japan. The solid circles and triangles represent the point estimates of the average treatment effects with and without the demographic control variables (i.e., gender, age, and education), and the vertical segments are the corresponding 95% confidence intervals. The figure clearly shows that the estimated treatment effects were

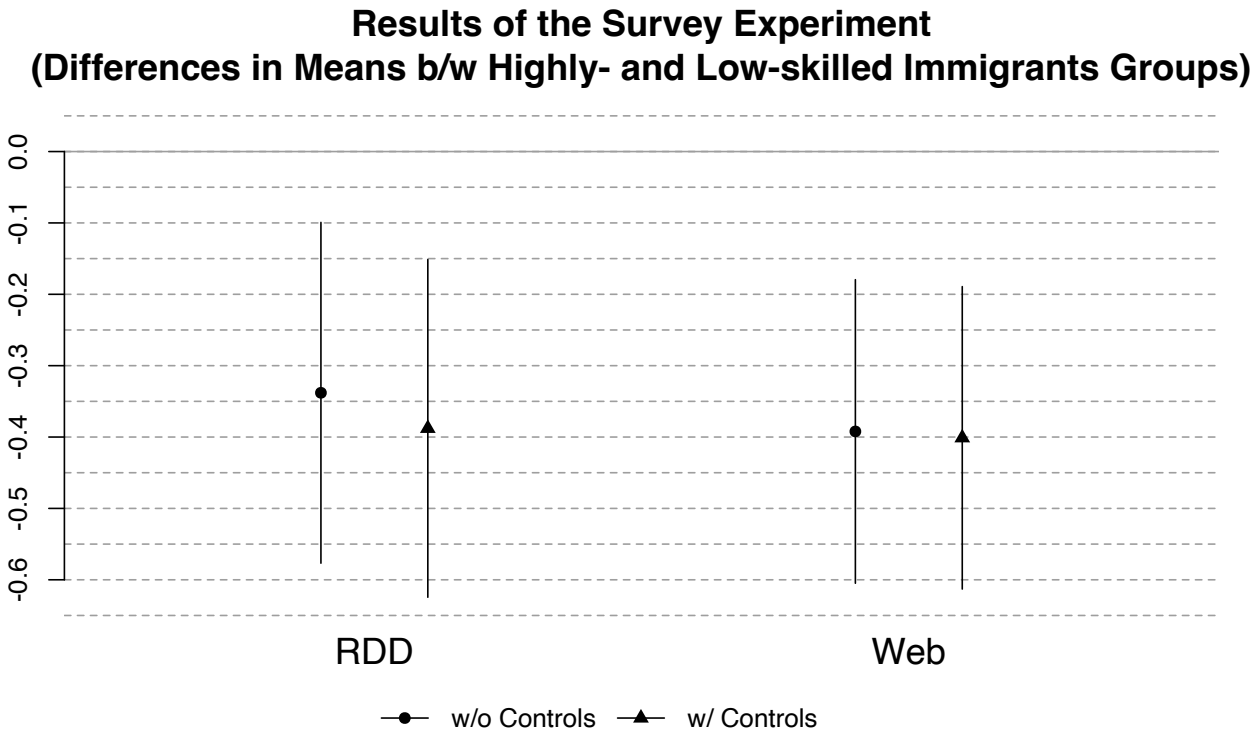

Figure 3: Results of the Survey Experiment from the RDD and Web Surveys

similar between the RDD and the web surveys. The negative and statistically significant ($p < 0.05$) treatment effects indicated that, on average, both our RDD and web survey respondents preferred high-skilled immigrants to low-skilled immigrants to come to Japan, a pattern also found in previous studies conducted in Japan (Igarashi et al., 2022). Controlling for demographic variables, whose distributions differed significantly between the surveys, did not affect the findings of the survey experiment. Thus, we can conclude that the results of the immigration experiment can be replicated regardless of the survey.

5 Discussions

This study introduces a newly developed RDD system for mobile phones in Japan, which incorporates an incentive payment system that does not require respondents to provide personal information. We then compared the data obtained through our RDD system and

an opt-in web survey and reported the findings.

Our RDD system has significant room for improvement, especially in terms of response rate. Because of the extremely low response rate, respondents to our current system are not representative of the Japanese population; as described in the previous section, RDD respondents are skewed toward male, young, and educated participants. Because individuals' decisions about whether to participate in the survey are not random, the low response rate makes the survey respondents unrepresentative of the population, even if the system randomly generates cell phone numbers to call. To overcome this problem, we need to develop a way to nudge individuals who would not participate in the current RDD survey to take the call and complete the survey. We are currently working on a design to increase the response rate and hope to present a modified system in the near future.

At the same time, our RDD system has potential, as we have shown that the results of survey experiments consistently found in previous studies can also be replicated in the RDD survey. Although the RDD survey conveys the experimental treatment through audio rather than visually, as in previous studies, which were mostly based on online surveys, we were able to obtain similar results. Thus, by solving the problem of low response rates, our RDD system will be able provide scholars with a platform for conducting survey experiments on a probability-based sample at a reasonable cost. However, we examined only the reproducibility of the results using the immigration experiment, and we still do not know whether the results reported in previous studies can be replicated when the topics under study or the experimental designs are different. Therefore, a series of replication studies is necessary to examine the types of survey experiments in which our RDD system can produce high-quality results.

References

Alvarez, R. M., Atkeson, L. R., Levin, I., & Li, Y. (2019). Paying attention to inattentive survey respondents. *Political Analysis*, *27*(2), 145–162. https://doi.org/10.1017/pan.2018.57

Asimov, A., & Blohm, M. (2024). Sequential and concurrent mixed-mode designs: A tailored approach. *Journal of Survey Statistics and Methodology*, *smae016*. https://doi.org/10.1093/jssam/smae016

Coffey, S., Maslovskaya, O., & McPhee, C. (2024). Recent innovations and advances in mixed-mode surveys. *Journal of Survey Statistics and Methodology*, *smae025*. https://doi.org/10.1093/jssam/smae025

Conrad, F. G., Schober, M. F., Antoun, C., Yan, H. Y., Hupp, A. L., Johnston, M., & Zhang, C. (2017). Respondent mode choice in a smartphone survey. *Public Opinion Quarterly*, *81*(S1), 307–337. https://doi.org/10.1093/poq/nfw097

Couper, M. P., Kapteyn, A., Schonlau, M., & Winter, J. (2007). Noncoverage and nonresponse in an internet survey. *Social Science Research*, *36*(1), 131–148. https://doi.org/10.1016/j.ssresearch.2005.10.002

Couper, M. P., & Miller, P. V. (2008). Web survey methods: Introduction. *Public Opinion Quarterly*, *72*(5), 831–835. https://doi.org/10.1093/poq/nfn066

Eguchi, T. (2022). Designing the 2021 house of representatives election survey : Reasons for adopting online surveys. *Policy and Survey*, *22*, 29–34. https://doi.org/https://sucra.repo.nii.ac.jp/records/2000012

Hainmueller, J., & Hiscox, M. J. (2010). Attitudes toward highly skilled and low-skilled immigration: Evidence from a survey experiment. *American Political Science Review*, *104*(1), 61–84. https://doi.org/10.1017/S0003055409990372

Hirasawa, K., Sugino, I., & Kaeriyama, A. (2023). The advantages of mixed-mode surveys for randomly selected people: Focusing on comparison of web and mail methods.

Contemporary Sociological Studies, *36*, 77–92. https://doi.org/https://doi.org/10.7129/hokkaidoshakai.36.77

Igarashi, A., Miwa, H., & Ono, Y. (2022). Why do citizens prefer high-skilled immigrants to low-skilled immigrants? identifying causal mechanisms of immigration preferences with a survey experiment. *Journal of Marketing*, *9*(2), 116–131. https://doi.org/10.1509/jm.15.0007

Jeong, D., Aggarwal, S., Robinson, J., Kumar, N., Spearot, A., & Park, D. S. (2023). Exhaustive or exhausting? evidence on respondent fatigue in long surveys. *Journal of Development Economics*, *161*, 102992. https://doi.org/10.1016/j.jdeveco.2023.102992

Kim, S., & Couper, M. P. (2021). Feasibility and quality of a national rdd smartphone web survey: Comparison with a cell phone cati survey. *Social Science Computer Review*, *39*(6), 1218–1236. https://doi.org/10.1177/0894439320964135

Kocar, S. (2022). Survey response in rdd-sampling sms-invitation web-push study. *Survey Research Methods*, *16*(3), 283–299. https://doi.org/10.18148/srm/2022.v16i3.7846

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied cognitive psychology*, *5*(3), 213–236. https://doi.org/10.1002/acp.2350050305

Moniz, P., Ramirez-Perez, R., Hartman, E., & Jessee, S. (2024). Generalizing toward nonrespondents: Effect estimates in survey experiments are broadly similar for eager and reluctant participants. *Political Analysis*, 1–14. https://doi.org/10.1017/pan.2024.8

Olson, K., Smyth, J. D., Horwitz, R., Keeter, S., Lesser, V., Marken, S., & Wagner, J. (2021). Transitions from telephone surveys to self-administered and mixed-mode surveys: Aapor task force report. *Journal of Survey Statistics and Methodology*, *9*(3), 381–411. https://doi.org/10.1093/jssam/smz062

Peytchev, A., & Neely, B. (2013). Rdd telephone surveys: Toward a single-frame cell-phone design. *Public Opinion Quarterly*, *77*(1), 283–304. https://doi.org/10.1093/poq/nft003

Porter, S. R., Whitcomb, M. E., & Weitzer, W. H. (2004). Multiple surveys of students and survey fatigue. *New Directions for Institutional Research*, *2004*(121), 63–73. https://doi.org/10.1002/ir.101

Sakshaug, J. W., Beste, J., Coban, M., Fendel, T., Haas, G. C., Hülle, S., & Zins, S. (2020). Impacts of the covid-19 pandemic on labor market surveys at the german institute for employment research. *14*(2), 229–233. https://doi.org/10.18148/srm/2020.v14i2.7743

Seol, D.-H., Jang, D.-H., & LoCascio, S. P. (2023). Rdd with follow-up texting: A new attempt to build a probability-based online panel in south korea. *Asian Journal for Public Opinion Research*, *11*(3), 257–273. https://doi.org/10.15206/ajpor.2023.11.3.257

Soullier, N., Legleye, S., & Richard, J.-B. (2022). Moving towards a single-frame cell phone design in random digit dialing surveys: Considerations from a french general population health survey. *BMC Medical Research Methodology*, *22*(1), 94. https://doi.org/10.1186/s12874-022-01573-1

Valentino, N. A., Soroka, S. N., Iyengar, S., Aalberg, T., Duch, R., Fraile, M., & Kobayashi, T. (2019). Economic and cultural drivers of immigrant support worldwide. *British Journal of Political Science*, *49*(4), 1201–1226. https://doi.org/10.1017/S000712341700031X