Policy Research Working Paper 10359

How Well Do Internet-Based Surveys Track Labor Market Indicators in Middle-Income Countries?

Vidhya Soundararajan
Sidiki Soubeiga
David Newhouse
Amparo Palacios-Lopez
Utz Johann Pape
Michael Weber

WORLD BANK GROUP
Poverty and Equity Global Practice
March 2023

Policy Research Working Paper 10359

Abstract

Online surveys are convenient, cost-effective, speedy, and increasingly popular instruments for data collection. This study investigates whether online surveys that used Random Domain Intercept Technology to recruit respondents were accurately measured labor market outcomes in six middle-income countries in the aftermath of the COVID-19 pandemic. Compared with the national average, online surveys oversampled males, youth, those with higher levels of education, and those in smaller households. Reweighting using propensity score estimates fails to equalize the means of variables excluded from the model. When comparing the employment-to-population ratio from the internet surveys to the most recent relevant nationally representative surveys, the average deviation is 30 percent. Reweighting using propensity scores in that case worsened the bias. Internet survey estimates of informal and self-employment rates also tend to be inconsistent with benchmark data, although the latter are available for fewer countries. Overall, the results suggest that despite the advantages and convenience of recruiting internet survey participants through Random Domain Intercept Technology, the resulting sample is not representative and even after propensity score reweighting, it can yield estimates that are at odds with nationally representative surveys.

This paper is a product of the Poverty and Equity Global Practice. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at dnewhouse@worldbank.org.

How Well Do Internet-Based Surveys Track Labor Market Indicators in Middle-Income Countries?[¥]

Vidhya Soundararajan, Sidiki Soubeiga, David Newhouse, Amparo Palacios-Lopez, Utz Johann Pape, Michael Weber

[¥] We thank the RIWI corporation for sharing the data from their online surveys for this paper. The data were collected through the Random Domain Intercept Technology (RDIT), patented by RIWI corporation.

1. Introduction

Internet-based surveys are a convenient option for policy makers and research organizations to collect data in a cost-effective way. Online surveys present many advantages. First, it is possible to achieve large sample sizes in a short amount of time, enabling the tracking of frequently changing events. This would enable policy makers to collect survey data more frequently on key policy indicators. Second, with the increased spread of the Internet, online surveys can be substantially more cost-effective than regular physical surveys in reaching out to more respondents. Third, the anonymity of respondents makes it possible to ask extremely sensitive questions, garner honest responses, and maintain the respondent's safety, which could otherwise be difficult in face-to face or phone surveys (Braunsberger et al. 2007; Granello and Wheaton 2004).

The COVID-19 pandemic curtailed economic activities worldwide and adversely affected labor market participation. These changes were abrupt and ever-evolving due to the unpredictable nature of the pandemic, which also in many countries prevented the implementation of face-to-face surveys. In such a setting, internet surveys are a potentially attractive option for tracking the frequently changing economic situation.

The main disadvantage of online surveys is that they are typically based on non-probability-based sampling techniques. Respondents are selected based on convenience and internet access rather than at random from a census-based sample frame.[1] Online surveys are not nationally representative because some individuals in the target population may not have access to the internet, those with internet access may be more or less likely to be approached for a survey, and not all of those who are approached may respond to the survey. Since the probability of inclusion in the sample depends on respondent characteristics, the resulting sample is prone to selection bias. A large literature corroborates this evidence on the lack of representativeness of online surveys (e.g., Evans and Mathur 2005; Granello and Wheaton 2004; Steinmetz and Tijdens 2009; Steinmetz et al. 2014; Steinmetz, Tijdens, and de Pedraza 2009; Valliant and Dever 2011). Despite these issues, surprisingly, many surveys continue to be conducted online, for example, through river sampling (American Association for Public Opinion Research 2013), and many opinion polls continue to be conducted using nonprobability online panels (Callegaro, Villar, Yeager, and Krosnick 2014a; Callegaro, Baker, Bethlehem, Göritz, Krosnick, Lavrakas, 2014b).

In comparison with Internet surveys, other modes of data collection have different strengths and weaknesses. Random Digit Dialing (RDD) phone surveys and traditional face-to-face surveys are common and more traditional alternatives. RDD is a method of probability sampling that provides a sample of households, families, or persons through a random selection of their telephone numbers. Since RDDs rely on the frame of those who own a phone, and not on the whole population, they may not be nationally representative. Biases can also arise based on the person who answers the phone call, because interviewers and respondents sometimes rush through conversations, and due to non-responses arising from network non-availability or other reasons (Holbrook et al. 2003).[2] From the perspective of those requesting the survey, online

[1] Increasingly, many companies are trying to conduct Internet surveys with a frame from which probability samples can be drawn (e.g., Kapteyn, Smith, and van Soest 2007; Malhotra and Kuo 2008; Moskalenko and McCauley 2009; Skitka and Bauman 2008), but these are rare because it is hard to develop a frame without knowledge of the universe of Internet users.
[2] For example, Kohut et al. (2012) shows that RDD response rates decreased from 36% in 1997 to 9% in 2012.

surveys are cheaper to conduct compared to phone surveys.[3] Traditional face-to-face surveys typically employ carefully designed probabilistic samples and are therefore representative. Another major advantage of in-person interviews is the ability to obtain high-quality and deeply insightful answers from these surveys, while Internet surveys are poorly suited for collecting answers to complex questions. However, face-to-face surveys are costly and are therefore difficult to implement frequently. They can also be difficult to field in conflict-affected environments due to security concerns.

How important is the lack of representativeness of online surveys and to what extent can it be corrected using statistical reweighting techniques? To investigate these questions, we utilized data from online surveys to track labor market indicators during the COVID-19 pandemic. The surveys were administered over three data collection waves in Brazil, the Arab Republic of Egypt, Indonesia, Kenya, Sri Lanka, and Türkiye, between December 2020 and August 2021.

Participants were recruited using Random Domain Intercept Technology (RDIT), a method that randomly selects internet users that mistype a web address or click on a broken link. The RDIT was developed by the RIWI corporation, and it works in the following way. RIWI takes ownership or temporary control of unused/abandoned websites or expired domains. Usually, users receive a "this page does not exist" message while they click on such websites while browsing the internet. However, because of RIWI's ownership, users would now be led to a link to take up an RIWI survey. RIWI, by autodetecting the IP address of the survey participant, offers the survey in the local language. Participants are anonymized and do not get any incentives to take up the survey. The survey can reach countries where there could be potentially government restrictions because no entity has a list of RIWI-controlled domains, and RIWI often changes the list of domains it owns.[4]

RDIT has advantages over other methods of selecting Internet survey respondents including Internet panels or email-based solicitation. In panels, respondents might be conditioned to taking the survey and may even change their behavior if they have been repeatedly answering survey questions, or they may make more errors than fresh respondents, speeding through a survey or answering strategically to avoid follow-up questions (Toepoel et al. 2009). RDIT methodology, by the virtue of randomly targeting respondents and employing a repeated cross-section survey, avoids these issues. Email solicitation may also be problematic if the emails reach the respondents' spam box, and because respondents are not anonymous (Schonlau and Couper, 2017; Roder-DeWan et al. 2019). Further, RDIT surveys are also consistent and reliable over time, as we see in our results in section 5. Another set of studies showed that an RDIT survey conducted every month for 21 months in India produced reliable estimates with low standard errors (Seeman et al. 2010; Seeman et al. 2016; Rizo et al, 2011). This is unlike online panel surveys where stability can be compromised due to high attrition among panel members (Toepoel et al. 2009).

[3] However, responding to an online survey may not be cheap for responders living particularly in a Lower-Middle-Income (LMI) country, and having to pay a non-negligible amount of money to buy internet data. But data costs are increasingly cheaper in many developing countries.

[4] RDIT has been used in several previous studies for examining attitudes related to mental illness, health care, and vaccination. For example, Sargent et al. (2022) tracked COVID-19 vaccination rates in real time, Seeman et al. (2016) studied attitudes on mental illness, and Roder-DeWan et al. (2019) studied the expectations of health care quality among low-income countries.

Despite these advantages, Internet surveys conducted with RDIT could be biased because of non-probability-based sampling.[5] To better understand these biases, we take two approaches to assess the quality of internet survey estimates. The first approach compares time-invariant measures, particularly demographic variables, calculated from the online surveys to nationally representative surveys collected prior to the crisis, contained in the World Bank's Global Monitoring Database (GMD). The results indicate that online surveys represent a substantially different demographic group compared to the national picture, overrepresenting male, younger, and more educated members of the country's population. Therefore, using the raw online survey results to track labor market indicators would be misleading.

As a first quality check, we attempt to address the bias described above using a propensity score adjustment, based on the estimated probability of participation in the Internet survey. In particular, we estimate a logit model of selection into the online survey, as opposed to the GMD survey, in each country. The model is used to construct weights equal to the inverse of the predicted propensity scores. The results show that, for variables included in the propensity score calculation, the reweighting process produces a smaller deviation of the means of the internet survey from the GMD survey. The deviations between variables excluded from the propensity score model, however, remain relatively high. For example, the deviation for model variables was 32.0 percent in Brazil, but 56.8 percent for the non-model variables. In Indonesia, while the deviation for model variables was 15.3 percent, the deviation was 582.0 percent for non-model variables. This result highlights that it is problematic in this context to assume that selection bias becomes ignorable after conditioning on demographic characteristics such as age, gender, and education.

The second quality check compares reweighted estimates of labor market indicators from the internet survey with other nationally representative data available in the same time period. This comparative data came from PNAD in Brazil, labor force surveys from Sri Lanka, Türkiye, and Indonesia, and the Continuous Household Survey from Kenya. Looking at the employment-to-population ratio, the reweighted online survey captures the national picture reasonably well in Türkiye, but considerably overestimates the employment-to-population ratio in the other countries. On average, the internet surveys give employment-to-population ratios that are 30 percent above the rates from the benchmark data. We also examined the relationship between internet coverage in a country and how well the online survey in that country captures the employment-to-population ratio accurately. Notably, Türkiye has the lowest deviation and the highest Internet coverage, but other than that there is no systematic relationship between Internet coverage and prediction accuracy.

We also examine two other labor outcomes, namely, the formal employment-to-total employment ratio, and the self-employment-to-total employment ratio. The sample size for these indicators is much smaller than that for employment as these questions were only administered for a subsample of those employed.[6] The performance of the online survey in tracking formality

[5] Besides non-probability-based sampling, another issue with standard Internet surveys is the potential duplication of responses. The same respondent can potentially answer the survey several times. However, the RDIT can identify the device of the survey respondent using a combination of the respondent's IP address, device details such as device type, internet browser, operating system, and other device details, and thus prevent the same device owner from participating in the survey multiple times. While the same person could potentially take the survey from multiple devices, the opportunity cost of time could prevent them from doing so.

[6] The employment related questions were administered to all those taking up the survey. We define as employed those who mention that they were "employed at work" and "employed but not at work". The formality and self-employment related questions were administered only to those who are "employed at work".

is poorer for these indicators, with the average relative absolute deviation of 39.1 percent, and for those in self-employment, it is even higher with a deviation of 166.9 percent.

Previous studies from developed countries that attempt to reduce bias in online surveys find that the propensity score-based reweighting methods do not help in bringing the web survey[7] estimates close to the probability-based reference survey/sample (Smyk et al. 2020; Schonlau et al. 2009; Yeager et al. 2011). We add to this literature by showing that these findings also tend to hold in developing countries. We compare the reduction in bias after adjusting the estimates based on propensity scores to estimates when we do not adjust at all. For the employment-to-population ratio, the propensity score adjustment increased the average bias ratio across countries/waves by 18 percent, indicating that reweighting using propensity scores worsened the accuracy of the estimates. The employment measures that we are interested in are not included in propensity score matching models. To the best of our knowledge, other studies have only examined the accuracy of variables used in the propensity score model or in other calibration methods, and not the non-model variables. We are the first to attempt to study whether non-model variables can be estimated accurately using such adjustment techniques. We also contribute to the literature by testing our method across six countries, which gives a coarse sense of whether Internet surveys may be less biased in some settings over others.

Overall, the results suggest considerable caution when interpreting labor market indicators collected through random intercept domain technology. Although the internet survey results do match benchmark data in some cases, the sample of internet responders is highly skewed, and in most cases, there are large discrepancies in outcomes even after reweighting. While there may be more room to experiment with the collection of additional demographic data to reweight the sample more effectively, achieving credibly representative estimates with internet surveys continues to be a major ongoing challenge.

The next sections are laid out in the following format. Section 2 presents the previous literature on the methods and issues in utilizing online surveys to be representative of a larger target population. Section 3 presents the data sources. Section 4 describes the methodology. Section 5 presents the results, and section 6 concludes the study.

2. Literature and Methods Review

2.1. Overview of Issues in Online Surveys

There are two broad types of web surveys. The first one is a probability-based sample where there is a sampling frame with e-mail addresses available for all eligible persons and a random sample is asked to visit a website and complete a survey. The second type, in contrast, is a *non-probability-based* web survey, like entertainment surveys, self-selected web surveys, and surveys made up of volunteer panels of Internet users. In the latter type of surveys, the probability of selection differs for different members of the population, making the surveys less representative. For instance, in volunteer web surveys, open invitations on websites are used to select respondents.[8] The probability of receiving such an invitation and the probability of accepting it are unknown. Due to the lack of a sampling frame and volunteer-based data collection, it is difficult to ascertain the extent to which the results from these surveys can be generalized for the whole population (Valliant and Dever, 2011).

[7] The terms web surveys and online surveys will be used interchangeably.

[8] Many market research and survey companies use convenience sampling to track sentiment, behavior, consumer preferences, etc.

One can distinguish between three distinct populations. The first is the target population or the set of persons for which inferences are to be made. The second is the potential covered population, which in this case includes persons having access to the Internet and who use it. The third is the realized sample, which includes persons who are exposed to the offer to participate in the survey and accept the offer. The potential covered population may not cover the target population well because of those without Internet access or because of those who have access but do not use the Internet and are hence unreachable. The realized sample may not cover the target population or the potentially covered population. Due to self-selection, the reachable respondents among the potentially covered population may be a biased sample. Thus, the realized sample of volunteers will be expected to overrepresent or underrepresent some demographic groups of the target population, making inferences based on this sample biased.[9]

2.2.Methods for Correcting Bias

Post-stratification, raking, propensity score adjustment (PSA), and matching are popular methods to address biases in online surveys. The efficacy of these different adjustment procedures in improving the representativeness of the web surveys is still debated (Bethlehem and Stoop 2007, Vehovar et al. 1999). Nonetheless, to apply these methods, we need a probability-based reference survey (or a subsample of the survey) that represents the distribution in the true population.

Poststratification indicates a process of stratification after the sample has been selected and is often appropriate when the sample is not properly balanced by the representation. After the data is obtained, typically, weights are determined for each stratum by dividing the proportion of population in that stratum (from the reference-survey) by the proportion of the sample in that stratum.

$$w_i = {}^{p_p}/_{p_s}$$

The most prevalent method for weighting is iterative proportional fitting, more commonly referred to as raking. With raking, a researcher chooses a set of variables where the population distribution is known, and the procedure iteratively adjusts the weight for each case in the internet sample until the sample distribution aligns with the population for those variables. Raking is popular because it is relatively simple to implement, and it only requires knowing the marginal proportions for each variable used in weighting.

Propensity score adjustment aims to correct differences caused by the varying inclinations of individuals to participate in web surveys (Duffy et al. 2005). That is, it adjusts for selection-bias due to observed covariates. More technically, a propensity score (ps_i) is the conditional probability that a person will be in one condition rather than in another (e.g., 'being in the web or the reference survey') given a set of observed covariates used to predict the person's condition. The web surveys are then weighted by the inverse of the propensity score, that is $1/ps_i$ (see Lee 2006, Schonlau et al. 2004, 2007). Since the propensity score refers to both the web and reference survey respondents, the propensity score weight for the reference survey is the inverse of propensity of being in the reference survey ($1/(1- ps_i)$).

[9] This general criticism is not only applicable to web surveys. In phone surveys, too, for example, coverage may be incomplete due to non-response. However, it is possible to randomly sample respondents from the potential covered population because telephone users list is usually known. This contrasts with a Web survey where there is currently no way to identify all Internet users (Valliant and Dever, 2011).

$$w_i = \begin{cases} 1/ps_i \\ 1/(1-ps_i) \end{cases}$$

While poststratification typically adjusts based on demographic variables, propensity score adjustments often use demographic and webographic (lifestyle/attitudinal) questions. Webographic variables capture general attitudes or behaviors that are hypothesized to differ between the web sample and the general population.[10]

An alternative to propensity score adjustment is matching. This technique utilizes a sample of cases representative of the population. This serves as the "target" sample. The aim of the matching exercise is to pair each case in the target sample with the most similar case from the online opt-in sample. When the closest match has been found for all of the cases in the target sample, any unmatched cases from the online opt-in Internet survey sample are discarded. At the end of the procedure, the matched cases from the Internet survey should be a set that closely resembles the target population.

There are several ways to conduct the matching procedure. To perform the matching, the target sample and the online survey data can be combined and a selection model for the Internet survey can be implemented as before. A propensity score can be constructed as in the above section, and matching can be based on those. If there are many variables for selection, machine learning algorithms, such as random forest, can be employed. Random forest algorithms can measure the probability that each case belongs to either the target sample or the survey and produce a measure of the similarity between each case and every other case. The similarity measure can also be used for matching. Another approach is to match based on Mahalanobis distance (Diamond and Sekhon, 2005).

2.3. Review of Studies Applying the above Methods

The selection of covariates for constructing propensity scores has been widely studied.[11] Covariate selection is usually done in a variety of ways, such as stepwise regression that drops variables that are not significant (Berk and Newton 1985, Lieberman et al. 1996), or conducting one-step covariate selection based on theoretical and logical significance. Some studies argue that there are no serious errors if the model for propensity score adjustment is mis-specified (Drake, 1993), other studies note that the choice of covariates is critical for propensity score weighting (Isaksson and Forsman, 2003). Covariates may be country specific (Stenbjerre, 2002). In countries with homogenous populations and high Internet penetration, there could be fewer covariates that discriminate web survey participation, unlike in countries that have low Internet penetration. There is also a general consensus in the literature that webographic questions, in addition to regular demographic variables, are needed to calculate the propensity scores more accurately.

Several studies examine non-economic averages and how adjusting can potentially reduce bias between volunteer Internet surveys and non-Internet probability-based surveys. Most studies indicate that various weighting schemes do not adequately reduce the difference between weighted Internet survey data and the reference survey. Varedian and Forsman (2003) investigate

[10] Many Internet survey companies, such as, Harris Interactive, increasingly use webographic questions that are thought to best capture the differences between the general population and those willing to answer a web survey.
[11] There is a tendency to use only those covariates that are statistically significantly different between treatment and comparison groups. Rosenbaum (2002) offered three cautions against this approach because the relationship between the outcome variable and one covariate is not important, statistical significance is not relevant always especially because it depends on the sample size.

the efficacy of propensity score-based weights in a web marketing survey that asks questions on hygiene products and attitude towards local banks in a northern European country. They find that estimates from the web survey and the phone surveys (which they use as a comparison) are quite different, and various weighting schemes did not make a difference regarding this. Some studies find that the type of measure matters. Schonlau et al (2004) found among propensity score adjusted web-survey estimates, only 8 of 37 estimates were not significantly different from phone survey estimates. Web survey estimates were significantly more likely to agree with RDD phone survey estimates for found factual questions, and those regarding personal health, and when the questions involved two as opposed to multiple categories of responses. Lee and Valliant (2009) show that weighting adjustment that combines propensity score adjustment and calibration adjustment has the potential to reduce bias for volunteer panel web survey estimates that are contaminated by sample selection bias. However, their results are for model variables used for propensity score adjustment or in the calibration exercise, and not for non-model variables.

Malhotra and Krosnick (2007) compare the 2000 and 2004 American National Election Study (ANES) data collected via using probability sampling with simultaneous Internet surveys of volunteer samples. The non-probabilistic samples were weighted using the raking method using the US CPS so as to match sample proportions with the population proportions on age and education. A comparison of the results yielded many differences in the distributions of variables and in the associations between variables. Applying the weights did almost nothing to reduce the differences between modes/sampling methods in the distributions of the political variables. Chang and Krosnick (2009) simultaneously fielded a probabilistic telephone survey, a probabilistic Internet survey, and a non-probabilistic internet survey. The non-probabilistic Internet survey used a weighting procedure to adjust for variable propensity of individuals to have regular access to email and the Internet. The comparisons showed that the probability samples were more representative than the nonprobability sample in terms of demographics and electoral participation, even after re-weighting.

Brüggen et al. (2016) compare the accuracy of results obtained from 18 opt-in online "panels" participating in the Dutch Online Panel Comparison Study (NOPVO) with data from a probability sample collected from the Internet (the LISS panel; Scherpenzeel 2009), and two probability samples collected via computer assisted personal interviewing, and compared all of them with the Dutch government's registry of all 16 million residents of the country, the Municipal Basic Administration (MBA). The non-probability sample surveys were weighted using General Regression (GREG) and Horvitz-Thompson Estimators. Results indicate that the nonprobability samples yielded less accurate estimates of proportions than the probability samples, and that weighting does not reduce selection bias in the level estimates. Dutwin and Buskirk (2017) compare RDD surveys and nonprobability Internet panel surveys to a high-quality in-person survey. The surveys were weighted using propensity weighting, raking, and sample matching. Their results showed that nonprobability samples attained the greatest estimated bias, and the in-person sample, the lowest. The weighting techniques were not able to improve the measures. Maccinnis et al. (2018) compare data across a variety of probability and nonprobability sampling methods in the United States, using a set of 50 measures of 40 benchmark variables. The probability samples interviewed by telephone or the Internet were the most accurate. Internet surveys of a probability sample combined with an opt-in sample were less accurate; least accurate were internet surveys of opt-in panel samples. These results were not altered by implementing poststratification weights provided by survey companies.[12]

[12] Cornesse et al. 2020 provides an exhaustive literature review of studies comparing non-probability and probability samples.

Two studies have shown promising results in using non-probability-based surveys, and using statistical methods to make them representative. Goel et al. (2015) employed an online opt-in survey of social and political attitudes. The survey utilized 14 demographic questions and 49 attitudinal questions that were drawn from the 2012 General Social Survey (GSS) and recent Pew Research Center studies. To correct for the biases in the non-representative online data, the study generates population-level and subgroup-level estimates via model-based poststratification [Gelman and Little, 1997, Wang et al., 2015]. After statistical correction, the median absolute difference between non-probability and probability surveys is 7.4 percentage points. The authors find this to be comparable to the difference between various probability surveys themselves, and hence conclude that "non-representative surveys are a promising tool for fast, cheap, and (mostly) accurate measurement of public opinion." Wang et al. (2015) conduct a series of daily voter intention polls in the Xbox gaming platform. After adjustments using regression and poststratification, the corrected estimates are comparable to estimates from a leading poll analyst that was based on aggregating traditional polls during the election cycle. From these results, the authors argue non-representative polling shows promise not only for election forecasting, but also for measuring public opinion on a broad range of social, economic and cultural issues.

Few studies look at economic variables. Smyk et al. (2020) study Wage Indicators (WI), an online survey in 17 countries and check if they provide wage distributions that match with benchmark data in each of these countries. They use covariate balancing propensity score to reweight the data. While these weights could match demography and human capital endowment of the representative samples, they are unable to match wage distributions from WI to the representative reference samples.

Propensity score adjustment seems to work better than post-stratification in settings where they have been compared directly. Isaksson and Forsman (2003) study political polls in Sweden and find that propensity score adjustment gives more accurate estimates than poststratification weighting. Yeager et al (2011) compared Internet surveys using non-probability samples of American adults with aggregate benchmark surveys. They tested the results for demographic and some non-demographic variables. Results indicate that post-stratification of non-probability samples did not consistently improve accuracy.

Szolnoki and Hoffmann (2013) compared a representative face-to-face survey with 2,000 respondents and a telephone survey with 1,000 respondents were compared with two online surveys, one based on quota sampling (2,000) and the other on snowball sampling (3,000) using identical questions. The objective was to assess consumer data for research in the wine business. The study did not conduct any reweighting for their online survey based on quota sampling. For the online survey with snowball sampling, they did not undertake reweighting because a weighting factor greater than five is seen as very problematic (Bandilla et al. 2003). Face-to-face data delivered the best results, followed by the telephone interviews and finally the online quota survey.

Some studies do not directly study web surveys, but use data collected during regular survey interviews from households with and without Internet access. This helps avoid the possibility that collecting data via the Internet as opposed to face to face or over the phone may directly affect the answer to survey questions. Dever et al (2008) compare estimates collected during telephone interviews from households with and without Internet access using data from the 2003 Michigan Behavioral Risk Factor Surveillance System in the United States. Their analysis results suggest that statistical adjustments can reduce, if not eliminate, coverage bias in the situation we study. Schonlau et al. (2009) study a survey of respondents that have Internet access in the Health and Retirement Study. They analyze whether propensity score and Genetic Matching algorithm can correct for imbalances between Internet and non-Internet samples for variables used for adjustment and for variables not used in the adjustment. Using these methods, they find that the estimated bias is almost always reduced, but significant differences in many cases remain.

Mercer et al (2018) study public opinion polls and find that the choice of adjustment variables is more important than any reweighting method for the accuracy of estimates. Raking performs nearly as well as more elaborate techniques based on matching. In their application, both the demographic and political variables resulted in lower average bias than adjusting on demographics alone. Very large sample sizes do not fix the shortcomings of online opt-in samples. While an online opt-in survey with 8,000 interviews may sound more impressive than one with 2,000, the study finds virtually no difference in accuracy. Ferri-Garcia (2020) uses Machine Learning methods as an alternative to propensity score adjustment (PSA) for reducing or eliminating selection bias in online surveys. The results indicate that the ML algorithms are more effective in removing selection bias than logistic regression when used for PSA, but this difference is not important when the dimensionality of the data is low, and the covariates are not very discriminant.

In sum, the existing literature broadly suggests that various weighting schemes do not adequately reduce the difference between weighted Internet survey data and the reference survey. Propensity score adjustment seems to work better than post-stratification in settings where they have been compared directly, but other studies show that the choice of adjustment variables is more important than any reweighting method for the accuracy of estimates. Only one study examines economic indicators (wages) from web surveys. While this study was able to match demography and human capital variables with reference samples, they were unable to match wage distributions. Clearly, all studies are based on high income countries, with little research on developing countries. In this study, we try to address these gaps by studying the accuracy of Internet surveys in a set of six developing countries using the propensity score adjustment method.

3. Data and Descriptive Statistics

3.1. Data Sources

The study uses several sources of data. The primary data source is the web-survey data collected through the Random Domain Intercept Technology. The survey is offered to the online population at random, provided they incur a browsing data input error. Specifically, if an individual stumbles across a broken link, they will be invited to participate in the survey. After excluding bots, users are invited to participate in a survey in their country's language. The targeting is based on the IP-address that also reveals information about different regions across the country. Respondents are identified by creating a unique non-personally identifiable respondent code which is a combination of the respondent's IP address, device details such as device type, internet browser, operating system, and other device details. Using this code, respondents are filtered out from taking the survey more than once, as the various survey waves are designed as repeated cross-sections. While the same person could potentially take the survey from multiple devices, the opportunity cost of time could prevent them from doing so.

The survey collected information on basic demographic details, such as age, gender, education, and household size. Beyond these basic details, the survey obtained information on respondents' employment status and type, and how they coped with the COVID-19 crisis. Personally identifiable information, such as name, residential address, or identity card details, were not collected. Enticements were not offered to participate in the survey, and respondents could easily exit the survey if/when they wanted to.

Table 3.1: Survey timelines and sample sizes

		Brazil	Egypt, Arab Rep.	Indonesia	Kenya	Sri Lanka	Türkiye
Panel A: Internet survey							
Wave-1	Dec 10th to 29th, 2020	1,098	1,071	1,198	1,024	1016	1,180
Wave-2	April 28th to May 9th, 2021	1,106	1,031	1,114	1,016	1,049	1,168
Wave-3	August 5th to 23rd, 2021	1,146	1,032	1,316	1,020	1,009	1,007
Panel B: Global Monitoring Data base							
Global Monitoring Database		Overall: 350,749 Internet: 272,598 Year: 2019	Overall: 34,733 Internet: 11,043 Year: 2015	Overall: 804,698 Internet: 80,840 Year: 2018	Overall: 52,841 Internet: 12,355 Year: 2015	Overall: 61,768 Internet: Not available Year: 2016	Overall: 30,731 Internet: 20,256 Year: 2018
Panel C: Comparative data from country labor force surveys							
Sample size		193,000 HH/month		793,542 HHs and 203,464 individuals	11,997 individuals	25,750 HH	58,560 HH
Survey name and timeframe		PNAD: Dec 2020; May and, June 2021		National Labor Force Survey – Indonesia; February 2021.	KCBS-Kenya; 28th Sep 2020 to 30th Nov 2020 (4th quarter of the survey)	LFS – Sri Lanka; Oct - Dec 2021	Household labor force survey; Jan - March 2021 Apr - June 2021

Notes: Wave-1 data were collected between December 10th and 29th, 2020, wave-2 between April 28th and May 9th, 2021, and wave-3 between August 5th and 23rd, 2021. Each wave had over 1000 "completed" surveys per country. A "completed" survey is one where respondents answered until the eighteenth question, which included all questions related to employment status and characteristics. For the analysis, we use only completed surveys, and ignored partial-response surveys. The sample size for these surveys in each wave is presented in panel A in table 3.1.

Second, we use the Global Monitoring Database, a probability-based survey, to compare demographic variables and individual attributes with our survey. The GMD is the repository of multitopic income and expenditure household surveys used by the World Bank to monitor global poverty and shared prosperity. The household survey data are typically collected by national statistical offices in each country, and then compiled, processed, and harmonized.[13] The data are nationally representative, and hence represent the demographics of the entire country. Since our survey targeted individuals and not households, we use the individual surveys from GMD for comparison. While the GMD survey years do not match our Internet survey years in all countries, they are not older than the online surveys by more than a few years. Since demographic characteristics, such as age and gender are unlikely to vary within a few years, the GMD is a valid comparison.[14] Specifically, we used the GMD survey from Brazil in 2019, Egypt in 2015, Indonesia in 2018, Kenya in 2015, Sri Lanka in 2016, and Türkiye in 2018.[15] The details are presented in panel B in table 3.1.

Third, we compare labor market indicators from the online survey with probabilistic sample-based surveys through the pandemic depending on their availability. In Brazil, we use the National Household Sample Survey – PNAD, a nationally representative survey of 193,000 households per month, conducted during the pandemic in December 2020 and May/June 2021. In Indonesia, we use the labor force surveys conducted in August 2020 and February 2021, each comprising 793,542 and 203,464 households respectively. In Kenya, we use the Kenya Continuous Household Survey comprising 11,997 households conducted in the last quarter of 2020 (October to December 2020). We use labor force indicators from the fourth quarter of 2020's Quarterly Report of the Sri Lanka Labor Force Survey covering up to 25,750 households, and from the Türkiye Household labor force survey covering 58,560 households for each quarter.[16] The comparative data are drawn based on whether data are available in the same month as an online survey wave, or within one month before or after. The exception is Türkiye for which the comparative household labor force surveys are conducted two months prior to wave-3 of the online survey. The details are presented in panel C in table 3.1.[17]

3.2. Data Description and Patterns

Are the demographics represented in the online data nationally representative, or do they over-sample or under-sample specific demographic groups? The answer to this question will indicate whether and how we can use the survey data to make broader conclusions, or if other adjustments will be required in order for them to be nationally representative. We compare the online survey's wave-2 data with the GMD, and with the Internet-using population from the GMD if Internet usage data are available in that country.[18]

The comparison with GMD relies on the assumption that the slightly dated GMD reflects the demographics at the time of the online survey accurately. Indeed, the pandemic itself could

[13] The process of compiling the GMD is coordinated by the Data for Goals (D4G) team and supported by the six regional statistics teams in the Poverty and Equity Global Practice.

[14] Admittedly, educational attainment might change drastically within a few years, and there have been instances of this in the past (for example, literacy rate changed from 53% in 2014 to 89% in 2019 in Côte d'Ivoire), but these types of drastic changes are minimal in the countries and period we study. For example, Brazil's illiteracy rate among adults remained somewhat stable at 7.95% in 2015 and 6.8% in 2018; Indonesia's adult illiteracy rate changed from 7.9% in 2015 to 8.55% in 2020; Sri Lanka's illiteracy rate changed from 10.5% in 2016 to 7.9% in 2019 (UNESCO Institute of Statistics, 2020).

[15] The GMD was also available in Egypt in 2019, but the variable pertaining to Internet usage is only available in the 2015 survey.

[16] For Sri Lanka, see http://www.statistics.gov.lk/LabourForce/StaticalInformation/QuarterlyReports/4thQuarter2020 (last accessed December 4, 2021); For Türkiye, see https://www.tuik.gov.tr/ (last accessed December 4, 2021).

[17] Egypt is the only country for which we are unable to find labor force survey data in the same time frame as our online surveys were conducted.

[18] We present all the demographic variables from the online data across waves in appendix figures A.4.1 to A.4.6. The figures indicate that the demographics captured by the online survey data are stable across different waves. Therefore, to make comparisons with GMD, we could potentially use any of the online survey waves. We use the median-wave, wave-2, for comparisons.

swiftly influence the distribution of demographic variables in countries. For example, the urban share of residents could have been heavily influenced by the internal migration of citizens returning to their home villages or to other rural areas or to take up remote work, or by people migrating from rural to urban areas in order to search for better livelihoods during the pandemic. These changes happened swiftly, in the aftermath of the onset of the pandemic in March 2020 in low- and middle-income countries (see, for example, evidence for Brazil (Garcia, 2021), Sub-Saharan Africa (FAO, 2020), and Peru (Dupraz-Dobias, 2020)). We take up such swift changes as a caveat while interpreting the data.

Data from figure 3.1 indicate that in Kenya and Indonesia, the urban share is well captured reasonably accurately in the online survey. For Brazil, the survey underrepresents urban areas. In Egypt and Sri Lanka, the survey overrepresents urban areas. Türkiye has a very high share of urban data points in the online survey, but location data are not available for Türkiye in the GMD. Overall, this mixed picture is consistent with the expectation that urbanization itself might be changing as a result of the COVID-19 pandemic, as noted above. Comparison with the Internet using population in GMD (figure 3.2) presents a different picture and shows that urban residents are underrepresented in the online survey. Female respondents are underrepresented in the online survey in all countries, which is consistent with many studies, particularly, Valliant and Dever (2011) for the United States. Younger people are overrepresented in online surveys in all countries and older people are underrepresented in the online survey, which is consistent with findings in Steinmetz et al. (2009) for the cases of Germany and the Netherlands, and by Valliant and Dever (2011) in the United States. These results are robust when we compare the online survey with the Internet-using population in GMD in figure 3.2.

Primary and secondary education are underrepresented in the online survey across countries, except Sri Lanka for primary education and Indonesia for secondary education. This trend broadly holds true even if we compare the online survey with the Internet using population from GMD. Tertiary education is overrepresented in the online surveys in all countries. This holds true if we compare internet using population from GMD, except in the case of Indonesia. Consistent with this, the literature shows that those with higher education are overrepresented in online surveys in most countries, with one exception. For example, as per Steinmetz et al. (2009), in the Netherlands, low and medium educated persons are underrepresented, and in Germany highly educated persons are underrepresented. In the United States, Valliant and Dever (2011) find that individuals in web survey volunteer samples have better education and higher income.[19] Households with one member are overrepresented in the online surveys, but households with three or more members are underrepresented in online surveys, and this is consistent in comparison with the Internet using population.

Web surveys could be targeted to specific kinds of individuals, which adds to another layer of bias. Steinmetz et al. (2009) shows that low and medium education workers are more likely to fill the survey in Germany because the links are provided in the trade union website. In our case, the survey was posed to those who mistype web addresses or those who stumble on a broken link. While we are unable to quantify this bias, its direction is unclear. While those who access the Internet regularly are more likely to click on broken links, those who access the Internet sparsely are also likely to mistype addresses because of their unfamiliarity.

[19] A possible reason why respondents in the Internet surveys have higher access to education than others is that those with access to Internet itself are more educated than those without access (Schonlau et al., 2009).

Overall, Internet surveys are biased in the demographic coverage of the population, in comparison to nationally representative data from the GMD. The sample is young, more educated, and more male compared to the national averages. The sample is not necessarily more urban than GMD, and the pattern of the bias depends on the country. This lack of a pattern in the urbanization bias could be because, as explained above, COVID-19 itself could have impacted urbanization that is not being reflected in the slightly dated GMD data.

Figure 3.1. Absolute deviation of Online Surveys from GMD

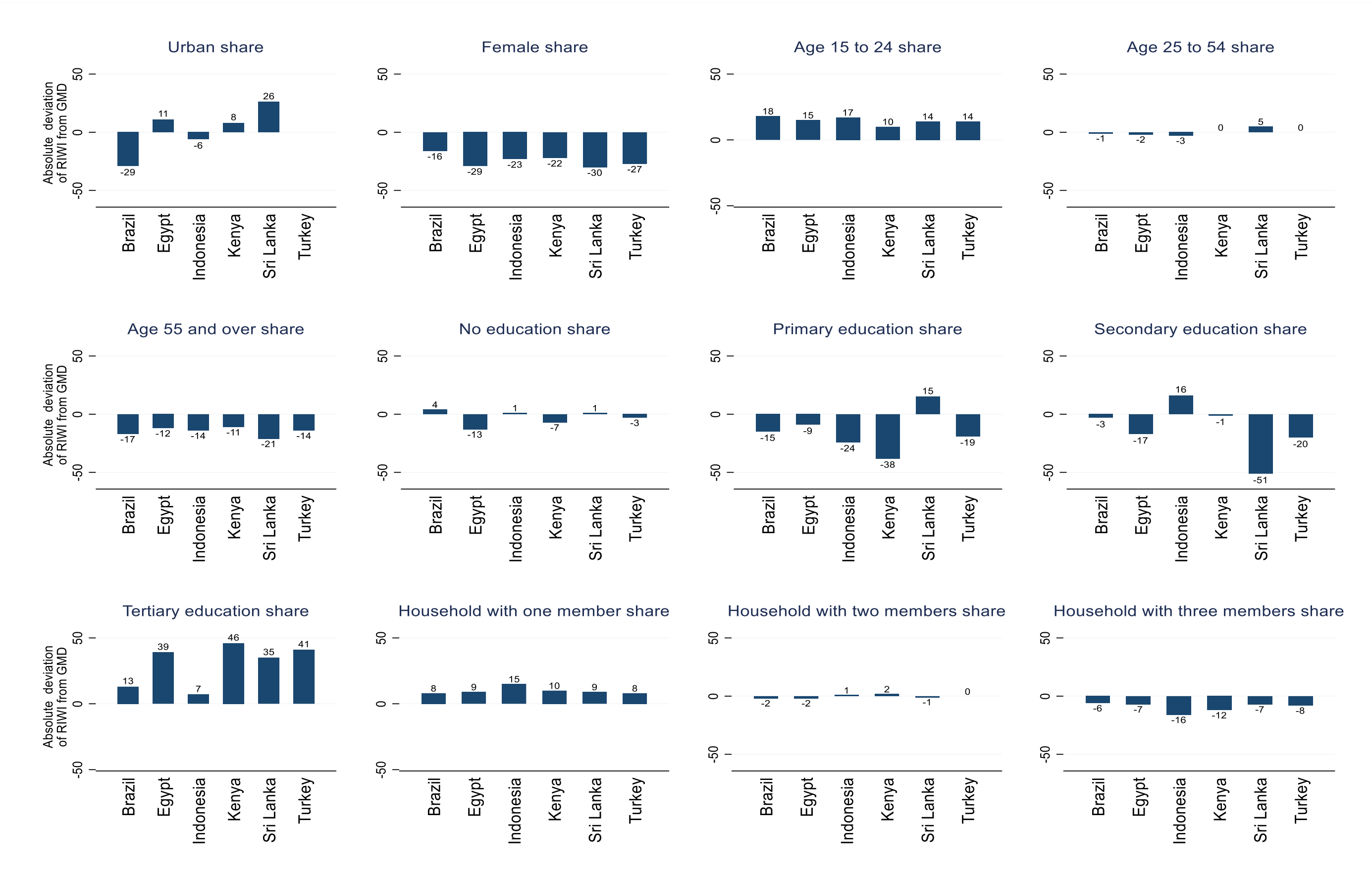
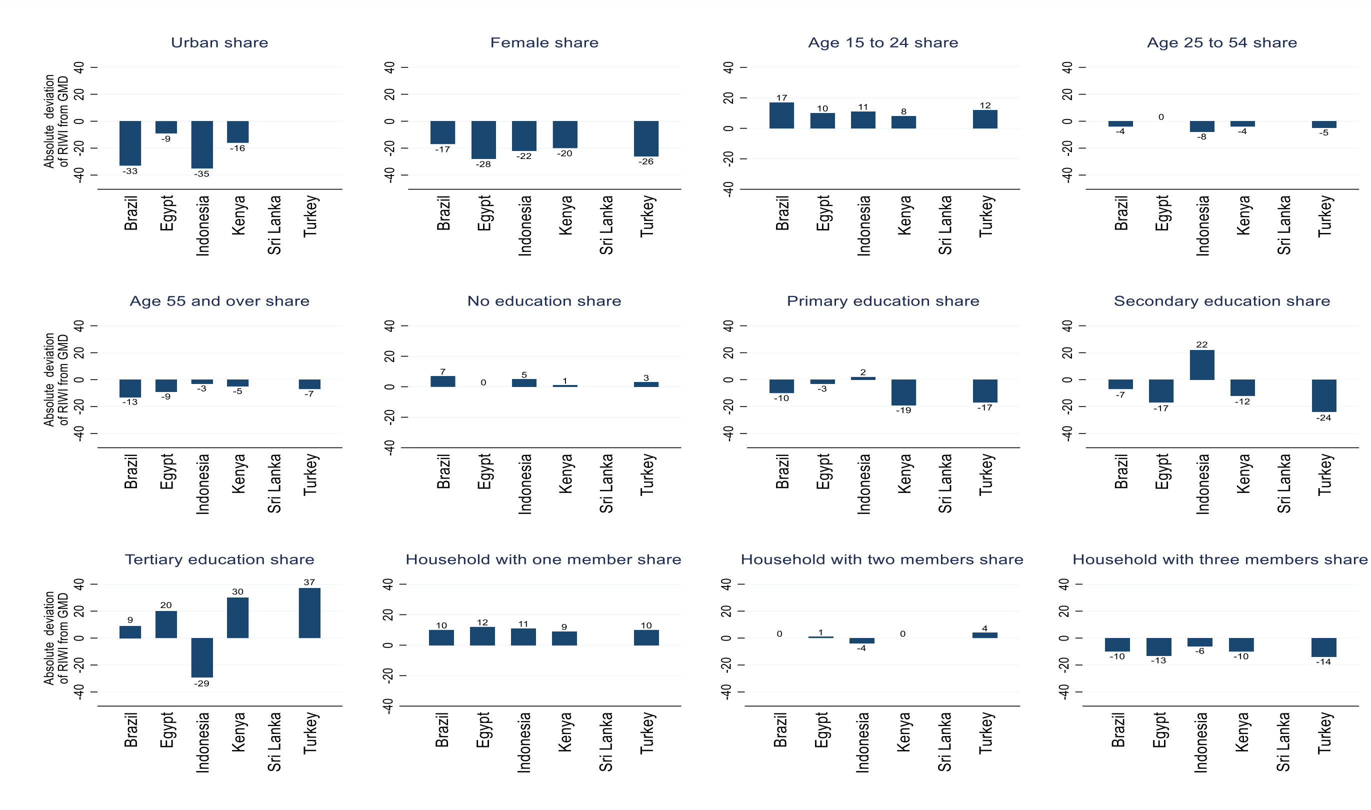


Figure 3.2. Absolute deviation of Online Surveys from the Internet using sample in GMD



4. Methodology

Next, in an attempt to overcome the biases in the online survey, as observed in section 3.2, we construct weights based on the propensity scores to examine whether applying such weights would make online surveys nationally representative. For this, we estimated a logit model of the following type:

$$y_i = \beta_0 + \beta_1 \boldsymbol{X}_i + \varepsilon_i \qquad \text{--- (1)}$$

Here, the dependent variable, y_i, takes the value 1 for the online survey data (for a specific wave) and 0 for the Internet using population of the GMD. The exception is Sri Lanka where Internet usage is unavailable in the GMD, and where the dependent variable takes the value 0 for all GMD observations. $\boldsymbol{X}_i$ is a vector that denotes the set of control variables that are associated with an individual's participation in the online survey. We estimate this model for each cross-section of the online survey wave with the GMD. Equation (1) describes the possible factors that influence a person who has access to the internet and was exposed to the Internet survey, as to whether they take up the survey or not.

Variables in $\boldsymbol{X}_i$ include gender, age categories, education, rural versus urban residence, sector of work pre-COVID, household size categories, a variable set of assets specific to each country, and region fixed effects.[20] We also include many interactions of these demographic variables between themselves, and with region.[21] These demographic factors are highly correlated with an

[20] The regional aggregations are different across countries, based on data availability. In some cases, the lowest administrative unit level where the data is available differs across Internet Surveys and GMD. In such cases, we harmonize at a higher level when running models or even in simply comparing indicators from the two datasets. Specifically, we define region at the state level in Brazil, province level in Indonesia and Sri Lanka, region level in Egypt, and county level in Kenya. For Türkiye, location data are missing in GMD, so region fixed effects are not used.

[21] The specific interaction variables used in each country/wave model are available upon request.

individual's inclination to take up the survey, as illustrated in section 3.2 (also observed in Dever (2011), and Valliant and Dever (2011)). Specifically, those who take up the Internet surveys are more likely to be younger individuals, more educated individuals, and males. Additionally, economic variables are important because higher income often correlates with whether an individual has the luxury of time to participate in the survey. Due to the unavailability of some economic variables in specific survey waves, models between GMD and wave-2 data do not include the asset variables, and models between GMD and wave-1 data do not include variables on assets and the pre-COVID sector of work. Moreover, in the education variable, the "no education" category was not collected in wave 1, and hence we use fewer categories on education in wave-1 regressions.

Asset data in the online surveys varies across countries. In Brazil, the online survey collected information about the availability of land phones, cell phones, computer, and piped water. In Egypt, the online survey collected data on the availability of land phone, computer, and piped water. In Indonesia, the asset variables include the availability of land phone, computer, car, air conditioner, and fridge. In Türkiye, the data on assets include the availability of computer, car, air conditioner, cable connection, and stove. In Kenya, the data on assets include the availability of computer, piped water, modernized floor, and modernized wall. In Sri Lanka, the assets include toilet and piped water.

In all regressions, standard errors were clustered at the region level, except in Türkiye where the regional indicators were not available.[22] From the model in equation (1), we obtained the predicted probability of participation in the online survey (p) and calculate weights for the online survey observations as $(1-p)/p$. The GMD is weighted by its own survey weights. Theoretically, the weighted means of variables used in the model should match between the online surveys and the GMD (Li et al, 2018).

We also test whether non-model variables match across the datasets. For this, we re-estimate the models after dropping the original model variables one by one. Using these restricted models, we predict the probability of participating in the online surveys (p), as before. Applying the weights as before, we compare the means of the dropped variable, and calculate the relative mean deviation of the Internet survey mean (for each wave) from the GMD mean.

5. Results

5.1. Results from Reweighting

This section presents the weighted averages of model variables based on propensity score weights obtained from the logit models in each country. Appendix tables A.4.1 to A.4.6 present the weighted averages from GMD (column 1), online survey wave-1 (column 2), online survey wave-2 (column 3), and online survey wave-3 (column 4), along with the confidence intervals. Results indicate that we do not reject the null that the GMD means are the same as the online

[22] Clustering usually provides more conservative standard error estimates compared to robust standard errors. Still, we face the problem of the small number of clusters biasing our standard error estimates downwards because the number of regions in each country is low and this could lead to over-rejection of the null hypothesis (Cameron and Miller 2015). The ideal way to proceed is to make corrections by estimating the Wild cluster bootstrap-*t* statistics which would provide more conservative standard errors. This downward bias will however not hurt us because, ideally, we are hoping that the null hypothesis of GMD and online survey mean equivalence is not rejected. But if it is not rejected with clustered standard errors, it will not be rejected with downward adjustments using the Wild method either.

surveys means in most instances across countries.[23] Appendix tables A.4.7 to A.4.12 present the weighted averages along with confidence intervals for non-model variables. Notably, we reject the null hypothesis that the GMD means are same as the online survey means in more instances for the non-model variables compared to the model-variables. These details are summarized in table A.4.13.

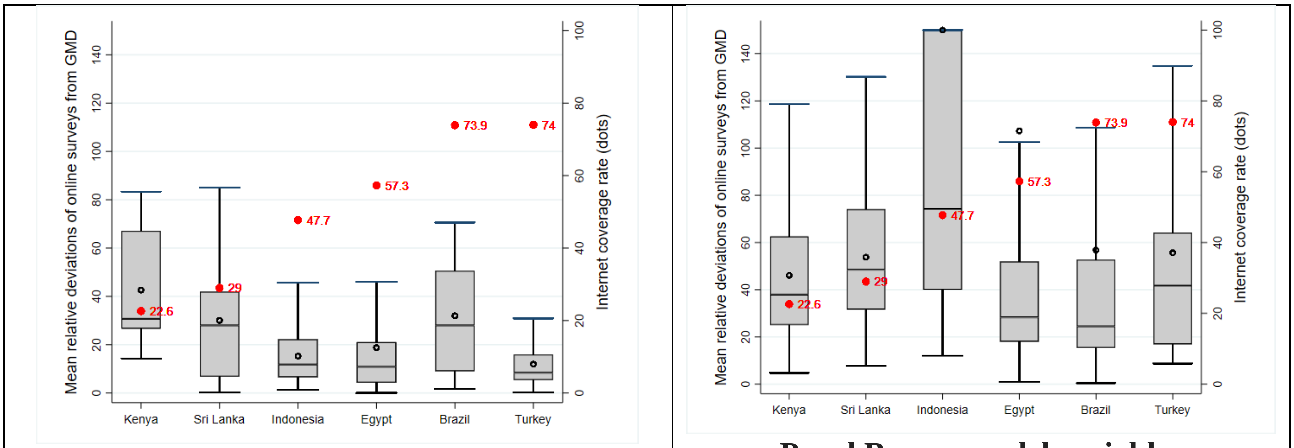
Compiling data from tables A.4.1 to A.4.12, tables A.4.14 and A.4.15 present the average of the relative deviations of online survey means with GMD means for model-variables and non-model variables, respectively. In both these tables, column (1) specifies the online survey waves that were considered in obtaining the deviation. For example, the variable age is available in all waves, and hence age was dropped to estimate models between wave-1 with GMD, wave-2 with GMD, and wave-3 with GMD. The average presented for age is the relative mean from all the three waves, and this is represented as '1-2-3' in column 1. Since asset variables were only available in wave-3, the models that drop the asset variables were only estimated between wave-3 and GMD. This is represented as '3' in column 1 for all asset variables. Columns (3)-(8) present the deviation for each country.

Results indicate that the average relative deviations for model variables between the Internet surveys and GMD range from 0 precent for cell phone usage in Egypt, to 87.1% percent for the "no education" category in Egypt. For non-model variables, the average varies widely, ranging from as low as 0.4 percent for computer usage in Brazil, to as high as 10,722.2 percent for working in "other sector" in Indonesia. Figure 5.1 presents a boxplot of the average deviations across all countries for model variables (panel A) and non-model variables (panel B). The spread of the deviation and the mean values of deviation are both lower for model variables compared to non-model variables. The mean values of these deviations for each country are reported in table 5.1.

It is useful to check if the models perform better in countries that have higher Internet coverage data. For this, we present data on Internet coverage across the countries from various official sources (red dot markers in the figure) in the second y-axis and arrange the countries in the x-axis in ascending order of internet coverage. Internet coverage in 2019 is highest in Türkiye and Brazil, followed by Egypt, Indonesia, Sri Lanka, and Kenya in that order. For model variables, countries with higher Internet coverage seem to have lower median values of deviation on average. Türkiye, which has the lowest mean deviation for the model variables (as per table 5.1), has the highest Internet coverage. For non-model variables, evidently, there is no clear indication that the best performing models are in countries where Internet coverage is the highest. The median deviation for Kenya, the country that has the least Internet coverage, seems close to the median deviation for Türkiye, the country that has the highest Internet coverage. Collectively, these results indicate that even after reweighting using inverse probability weights, it is challenging to match the means of non-model variables of the online surveys with nationally representative surveys.

[23] Online survey means statistically significantly different from GMD are indicated by yellow shaded cells in tables A.4.1 to A.4.6.

Figure 5.1. Average deviations and Internet coverage



Panel A: model variables

Panel B: non-model variables

Note: The Y-axis presents the relative deviation of internet survey means from GMD means. The panel A is for model variables, and panel B is for non-model variables. The boxplot shows the distribution of these relative deviations. The mean deviations are presented as a black circular marker. The deviations are winsorized at 150 percent. The second Y-axis presents Internet coverage ratio, represented by a red marker. The data source for Internet coverage are the following: Brazil: NIC.br; Egypt: Ministry of Communications and Information Technology; Indonesia: BPS-Statistics; Kenya: National Bureau of Statistics; Sri Lanka: ITU based on Department of Census and Statistics; Türkiye: Turkish Statistical Institute.

Table 5.1: Mean of average deviation of model and non-model Online survey variables from GMD

	Model variables	Non-model variables
Brazil	31.98%	56.8%
Egypt, Arab Rep.	18.79%	107.27%
Indonesia	15.26%	581.95%
Kenya	42.61%	46.12%
Sri Lanka	30.03%	53.76%
Türkiye	12.18%	54.76%

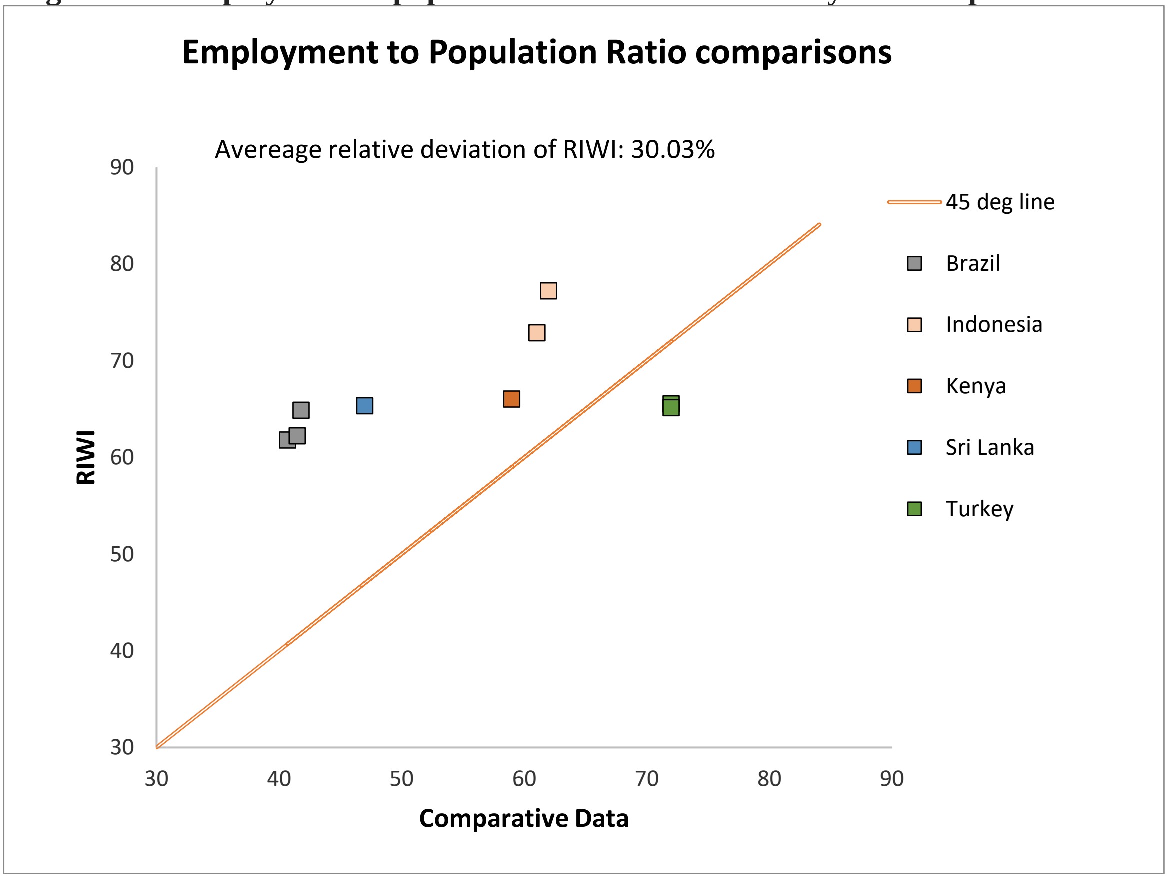
5.2. Comparing Labor Market Indicators

The main labor market indicator we track from the online surveys is the employment to population ratio in different countries. Specifically, we check if our measure tracks indicators from other nationally representative sources (if they exist) in the same/similar time frame.

We define someone to be employed if they worked for pay or for any kind of business, farming or other activity to generate income or future income, even if only for one hour in the previous week. We also define someone to be employed even if they did not work in the previous week, but if they have a job, business or family farm to which they would return to once the restrictions ease or once the work season begins. This definition encompasses those who are employed at work and employed but currently not at work. We calculate the employment to population ratio for each country by dividing the number of people who are employed by the total number of "complete" respondents and applying the model-based weights as described in section 4. For the comparative data, the definition of “employed” is the same as used in the online survey. That is, someone is defined as employed during the reference week if they performed some work for at least one hour or were temporarily not at work but had a job to go back to. Figure 5.2 presents a scatter plot of the employment to population ratio in online surveys and the comparative data.[24] The 45-degree line is presented for reference. Evidently, the online survey estimates are mostly above the 45-degree line, indicating that the online survey perhaps overrepresents the employment to population ratio.

We calculate the relative bias as the absolute value of the difference between the online estimates and the comparative data-based estimates divided by comparative data estimates. For the propensity score adjusted online estimates, the bias ranges from 9 percent to 55 percent, with an average of 30.0 percent. Without weight adjustments, the deviations ranged from 3 percent to 62 percent, with around the same average of 29.8 percent. Notably, the weight adjustment sometimes improved the accuracy of Internet sample surveys and sometimes reduced their accuracy. We also calculate bias reduction due to propensity score adjustment as (bias of unadjusted online survey estimates - bias of adjusted online survey estimates)/bias of unadjusted online survey estimates. We find the average of the bias reductions across countries/waves to be -18 percent, indicating that the weight adjustments overall worsened the accuracy of Internet sample surveys. [25]

Figure 5.2: Employment to population ratio: Internet survey and comparative data



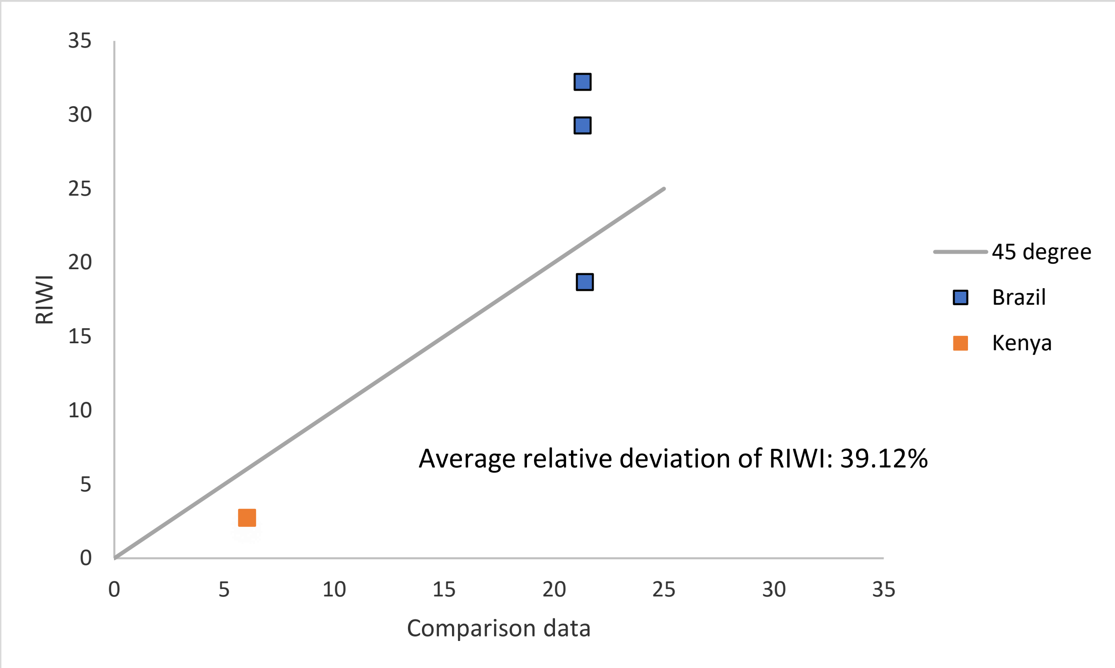
[24] The results for each country are presented as bar graphs in appendix figures A.4.6 to A.4.11. These figures also include comparative summary data from other nationally representative sources as line graphs whenever available.
[25] We do not have enough information from the comparative data to test the statistical significance of employment measures between our online surveys and the comparative data.

Note: Each point represents the combination of employment to population ratio as (comparative data, Internet survey data). Average relative deviation is the average of relative values of absolute deviation of the Internet survey from the comparative data. Specifically, we calculate the average of abs(Online-comparative)/comparative.

Data on other labor market indicators were collected by the survey, including the formal versus informal nature of employment, and type of employment. Respondents are classified to be working formally if they have a written contract, if they have any form of social insurance, and if they get annual paid or sick leave. Those who do not satisfy any of these three conditions are deemed to be working informally. The respondents are classified as self-employed if they own a business or business operated by the household or a family farm. They are instead classified as an employee if they work as an employee for another house, or another company, or as an apprentice, trainee, etc. The questions related to formality and self-employment are asked only to those who responded that they are *currently* employed at work, and not to those employed but not *currently* at work. Therefore, the survey sample size for these questions is lower than for the previous graph.

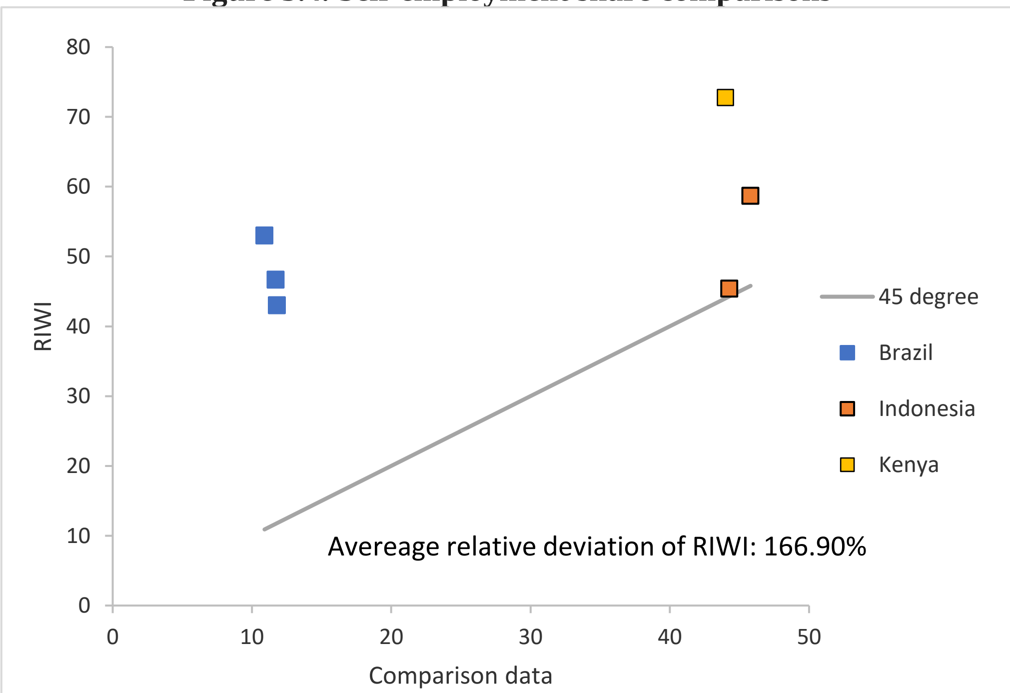
The comparative data for formal and self-employment are available for fewer countries and fewer time periods. Figure 5.3 and figure 5.4 present a scatter plot of the formal share between online surveys and the comparative data, and a scatter plot of the self-employment share between online surveys and the comparative data, respectively. For the propensity score adjusted online estimates of formal employment, the deviations from the comparative data range from 12.7 percent to 55.04 precent, with an average of 39.1 percent. Without weight adjustments, the deviations ranged from 14 percent to 29 percent, with a lower average of 21 percent. The average of bias reductions due to propensity score weight-based adjustments across countries/waves is -1.20 percent. Here again, the weight adjustment worsened the accuracy of Internet sample surveys. For the propensity score adjusted online estimates of self-employment, the deviations from comparative data range from 6 percent to 377 percent, with an average of 166.9 percent. Without weight adjustments, the deviations ranged from 38 percent to 402 percent, with a higher average of 210 percent. The average of bias reductions due to propensity score weight-based adjustments across countries/waves is 26 percent. In this case, the weight adjustment improved the accuracy of Internet sample surveys.

Figure 5.3. Strictly Formal Employment Share comparisons



Note: Each point represents the combination of formal employment share (comparative data, Internet survey data). Average relative absolute deviation is the average of relative values of absolute deviation of the Internet survey from the comparative data. Specifically, we calculate the average of abs(Internet-Comparative)/comparative.

Figure 5.4. Self-employment share comparisons



Note: Each point represents the combination of self-employment share (comparative data, Internet survey data). Average relative absolute deviation is the average of relative values of absolute deviation of the Internet survey from the comparative data. Specifically, we calculate the average of abs(Internet-Comparative)/comparative.

Since we do not have comparative literature focusing on employment status or related indicators such as self-employment and formality rate, we are unable to compare the bias itself directly with the other estimates in the literature. However, we can compare our bias-reduction from propensity score weights-based adjustments with others'. For example, using propensity score weights, and a combination of propensity score weighting and calibration, Lee and Valliant (2009) find a bias reduction in the range of 1 percent to 94 percent for health status, and 18 percent to 98 percent for health insurance coverage. However, both these variables themselves were used in the propensity score weight estimation, and health status was used in the calibration of weights. Specifically, the study finds that matching after adjustments is better in cases where the variable in question itself is used for the adjustment. In other words, the matching for model variables performs better than non-model variables, a result we also observe.

6. Conclusion

Traditional surveys that track economic and social indicators use probability sampling techniques based on a complete sampling frame. While these surveys accurately represent the national population, they are expensive and difficult to conduct in national crisis situations, such as the COVID-19 pandemic or fragile or conflict affected environments. Quicker, less expensive, and more convenient methods of data collection have emerged, including phone surveys and online surveys. While Random Digit Dialing technique-based phone surveys are based on a representative sampling frame, they may be biased, since the population of phone owners who would answer the phone and respond to the survey may differ from the general population. Internet surveys are even more convenient and inexpensive to administer, but their scope in developing countries could be curtailed by lower levels of internet access and educational attainment. These surveys are based on non-probabilistic sampling frames because it is difficult to ascertain the universe of Internet users.

In this study, we utilized three waves of online surveys conducted in six countries during the COVID-19 pandemic between December 2020 and August 2021. The objectives were to capture key economic variables during the pandemic that would have otherwise been harder to collect through physical door-to-door surveys. We compared economic indicators from these surveys with regular labor force surveys conducted in these countries around the same time in order to assess whether these surveys accurately captured and tracked indicators. We tracked three key indicators, namely, the employment-to-population ratio, the formal share of employment, and the share of self-employment.

Examining the demographic representation of the online survey data, we find that the raw or unweighted online-survey data are not nationally representative. They overrepresent men, younger individuals, and the highly educated. To attempt to overcome these biases, we estimated a propensity score model to estimate weights to make the internet survey data more representative. For this, we estimated a model to predict the selection of observations in the online survey compared to a nationally representative survey from the Global Monitoring Database (GMD) using a variety of independent variables. From this model, we further predicted inverse probability weights for adjustments. The means of variables included in the propensity score model seem to be relatively close to the GMD means, in comparison to the non-model variables whose online survey means deviated widely from the GMD mean. For example, the deviation for model variables was 32.0 percent in Brazil, but 56.8 percent for the non-model variables. In Indonesia, while the deviation for model variables was 15.3 percent, the deviation was 582.0 percent for non-model variables.

Applying these estimated weights, we compared the economic indicators from the online survey with those derived from nationally representative comparative labor force surveys available during the pandemic. The results indicate that there is substantial deviation of the online survey indicators from the labor surveys. The average across countries of the relative values of the absolute deviation of the employment-to-population ratio from the online survey to that of the labor force survey data is 30.0 percent. The same figure for formal share of employment is 39.1 percent, and for self-employment it is 166.9 percent. The reweighting procedure worsens the accuracy of online estimates of the employment-to-population ratio by 18 percent on average, and the formal employment share by 1.2 percent. In the case of self-employment rate estimates, model-based re-weighting reduced the bias by 26.4 percent.

We conclude that the bias from non-representative Internet survey data can be substantial. Corrections using propensity score-based weights based on a small set of demographic variables can actually worsen this bias, and therefore, the reliance on non-representative internet surveys could be misleading. However, Internet surveys may be useful to identify trends in times of emergency among the population that they represent (mostly men, younger individuals, and the highly educated) for specific purposes. On the other hand, phone surveys are much preferable to Internet surveys because they can be drawn from a nationally representative sample or utilize RDD when mobile phone coverage is high.

REFERENCES

American Association for Public Opinion Research (2013). "Report of the AAPOR Task Force on Nonprobability Sampling," available at https://www.aapor.org/AAPOR_Main/media/MainSiteFiles/NPS_TF_Report_Final_7_revised_FNL_6_22_13.pdf (accessed September 30, 2019).

Bethlehem, J., & Stoop, I. (2007). Online panels. A paradigm theft? *The challenges of a changing world*, 113-131.

Braunsberger, K., Wybenga, H., & Gates, R. (2007). A comparison of reliability between telephone and web-based surveys. *Journal of Business Research*, *60*(7), 758-764.

Bruggen, E., J. A. van den Brakel, and J. Krosnick (2016), "Establishing the Accuracy of Online Panels for Survey Research," Statistics Netherlands. Available at https://www.cbs.nl/en-gb/background/ 2016/15/establishing-the-accuracy-of-online-panels-for-survey-research (accessed September 30, 2019).

Callegaro, M., Villar, A., Yeager, D. S., & Krosnick, J. A. (2014a). A critical review of studies investigating the quality of data obtained with online panels based on probability and nonprobability samples. Online Panel Research: A Data Quality Perspective, Wiley (2014), pp. 23-53

Callegaro, M., Baker, R. P., Bethlehem, J., Göritz, A. S., Krosnick, J. A., & Lavrakas, P. J. (Eds.). (2014b). Online panel research: A data quality perspective. John Wiley & Sons.

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. Journal of human resources, 50(2), 317-372.

Chang, L., and J. A. Krosnick (2009), "National Surveys via RDD Telephone Interviewing versus the Internet," Public Opinion Quarterly, 73, 641–678.

Cornesse, C., Blom, A. G., Dutwin, D., Krosnick, J. A., De Leeuw, E. D., Legleye, S., ... & Wenz, A. (2020). A review of conceptual approaches and empirical evidence on probability and nonprobability sample survey research. Journal of Survey Statistics and Methodology, 8(1), 4-36.

Dever, J. A., Rafferty, A., & Valliant, R. (2008, June). Internet surveys: Can statistical adjustments eliminate coverage bias? In Survey Research Methods (Vol. 2, No. 2, pp. 47-60).

Diamond, A., & Sekhon, J. S. (2013). Genetic matching for estimating causal effects: A general multivariate matching method for achieving balance in observational studies. *Review of Economics and Statistics*, *95*(3), 932-945.

Duffy, B., Smith, K., Terhanian, G., & Bremer, J. (2005). Comparing data from online and face-to-face surveys. *International Journal of Market Research*, *47*(6), 615-639.

Dupraz-Dopias (2020). COVID-19 shifts Peru's internal migration into reverse. The New Humanitarian. 8 July 2020.

Dutwin, D., and T. D. Buskirk (2017), "Apples to Oranges or Gala versus Golden Delicious?," Public Opinion Quarterly, 81, 213–239.

Evans, J. R., & Mathur, A. (2005). The value of online surveys. *Internet research*.

Fan Li, Kari Lock Morgan & Alan M. Zaslavsky (2018) Balancing Covariates via Propensity Score Weighting, Journal of the American Statistical Association, 113:521, 390-400, DOI: 10.1080/01621459.2016.1260466

FAO (2020). Migrant workers and remittances in the context of COVID-19 in sub-Saharan Africa. Rome. Last accessed on March 29th, 2022. Link: https://doi.org/10.4060/cb0184en

Ferri-García, R., & Rueda, M. D. M. (2020). Propensity score adjustment using machine learning classification algorithms to control selection bias in online surveys. PloS one, 15(4), e0231500.

Garcia, Raphael (2021). Can Brazil's remote workers thrive outside big cities? BBC, January 2021. Last access on March 29th, 2022. Link: https://www.bbc.com/worklife/article/20210121-the-challenge-of-remote-work-in-developing-nations

Goel, S., Obeng, A., & Rothschild, D. (2015). Non-representative surveys: Fast, cheap, and mostly accurate. *Work Pap*.

Granello, D. H., & Wheaton, J. E. (2004). Online data collection: Strategies for research. *Journal of Counseling & Development*, *82*(4), 387-393.

Holbrook, A. L., Green, M. C., & Krosnick, J. A. (2003). Telephone versus face-to-face interviewing of national probability samples with long questionnaires: Comparisons of respondent satisficing and social desirability response bias. *Public opinion quarterly*, 67(1), 79-125.

Andrew Kohut, Scott Keeter, Carroll Doherty, Michael Dimock, and Leah Christian. 2012. Assessing the representativeness of public opinion surveys. Pew Research Center, Washington, DC, 2012.

Kapteyn, A., Smith, J. P., & Van Soest, A. (2007). Vignettes and self-reports of work disability in the United States and the Netherlands. *American Economic Review*, *97*(1), 461-473.

Lee, S. (2006). Propensity score adjustment as a weighting scheme for volunteer panel web surveys. *Journal of official statistics*, *22*(2), 329.

Macinnis, Bo, Jon A. Krosnick, Annabell S. Ho, and Mu-Jung Cho. 2018. The accuracy of measurements with probability and nonprobability survey samples. Public Opinion Quarterly, Vol. 82, No. 4, Winter 2018, pp. 707–744.

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, 286–323.

Malhotra, N., & Kuo, A. G. (2008). Attributing blame: The public's response to Hurricane Katrina. *The Journal of Politics,* 70(1), 120-135.

Mercer, A., Lau, A., & Kennedy, C. (2018). For weighting online opt-in samples, what matters most? Pew research center. January 2018.

Moskalenko, S., & McCauley, C. (2009). Measuring political mobilization: The distinction between activism and radicalism. *Terrorism and political violence*, *21*(2), 239-260.

Pasek, J. (2016), "When Will Nonprobability Surveys Mirror Probability Surveys? Considering Types of Inference and Weighting Strategies as Criteria for Correspondence," International Journal of Public Opinion Research, 28, 269–291

Rizo C, Deshpande A, Ing A, Seeman N. A rapid, Web-based method for obtaining patient views on effects and side-effects of antidepressants. Journal of Affective Disorders. 2011;130(1-2):290-3. doi: 10.1016/j.jad.2010.07.027.

Roder-DeWan, S., Gage, A. D., Hirschhorn, L. R., Twum-Danso, N. A., Liljestrand, J., Asante-Shongwe, K., Rodríguez, V., Yahya, T., & Kruk, M. E. (2019). Expectations of healthcare quality: A cross-sectional study of internet users in 12 low-and middle-income countries. PLoS Medicine, 16(8), e1002879. https://doi.org/10.1371/journal.pmed.1002879

Rosenbaum, Paul R. 1987. ''Model-Based Direct Adjustment.'' Journal of the American Statistical Association 82:387-94.

Sargent RH, Laurie S, Weakland LF, Lavery JV, Salmon DA, Orenstein WA, Breiman RF (2022). Use of Random Domain Intercept Technology to Track COVID-19 Vaccination Rates in Real Time Across the United States: Survey Study. J Med Internet Res 2022;24(7):e37920. https://www.jmir.org/2022/7/e37920/

Scherpenzeel, A. C., and J. G. Bethlehem (2011), "How Representative Are Online Panels? Problems of Coverage and Selection and Possible Solutions," in Social and Behavioral Research and the Internet: Advances in Applied Methods and Research Strategies, eds.M. Das, P. Ester, and L. Kaczmirek, pp. 105–132, New York: Routledge.

Seeman, N., Tang, S., Brown, A. D., & Ing, A. (2016). World survey of mental illness stigma. Journal of affective disorders, 190, 115-121.

Schonlau M, Couper MP. Options for Conducting Web Surveys. Statist Sci. 2017;32(2):279-92. doi: 10.1214/16-STS597.

Schonlau, Matthias, Kinga Zapert, Lisa Payne Simon, Katherine Sanstad, Sue Marcus, John Adams, Mark Spranca, Hongjun Kan, Rachel Turner, and Sandra Berry. 2004. ''A Comparison Between a Propensity Weighted Web Survey and an Identical RDD Survey.'' Social Science Computer Review 22:128-38.

Schonlau, M., Van Soest, A., Kapteyn, A., & Couper, M. (2009). Selection bias in web surveys and the use of propensity scores. *Sociological Methods & Research*, *37*(3), 291-318.

Skitka, L. J., & Bauman, C. W. (2008). Moral conviction and political engagement. *Political Psychology*, *29*(1), 29-54.

Smyk, M., Tyrowicz, J., & Van der Velde, L. (2021). A cautionary note on the reliability of the online survey data: the case of wage indicator. *Sociological Methods & Research*, *50*(1), 429-464.

Steinmetz, S., & Tijdens, K. (2009). Can weighting improve the representativeness of volunteer online panels? Insights from the German Wage indicator data. *C&M Newsletter*, *5*.

Steinmetz, S., Bianchi, A., Tijdens, K., & Biffignandi, S. Improving web survey quality–Potentials and constraints of propensity score weighting. *Online Panel Research: A Data Quality Perspective. Wiley*.

Szolnoki, G., and D. Hoffmann (2013), "Online, Face-to-Face and Telephone Surveys—Comparing Different Sampling Methods in Wine Consumer Research," Wine Economics and Policy, 2, 57–66.

UNESCO Institute of Statistics, 2022. Other Policy Relevant Indicators (OPRI), 2022. Link: https://apiportal.uis.unesco.org/bdds. Last accessed on April 5, 2022.

Valliant, R., & Dever, J. A. (2011). Estimating propensity adjustments for volunteer web surveys. *Sociological Methods & Research*, *40*(1), 105-137.

Vehovar, V., Manfreda, K. L., & Batagelj, Z. (1999, May). Web surveys: Can the weighting solve the problem. In *Proceedings of the Section on Survey Research Methods, American Statistical Association* (pp. 962-967).

APPENDIX

Figure A.4.1: Comparison of female share across the three Internet survey waves

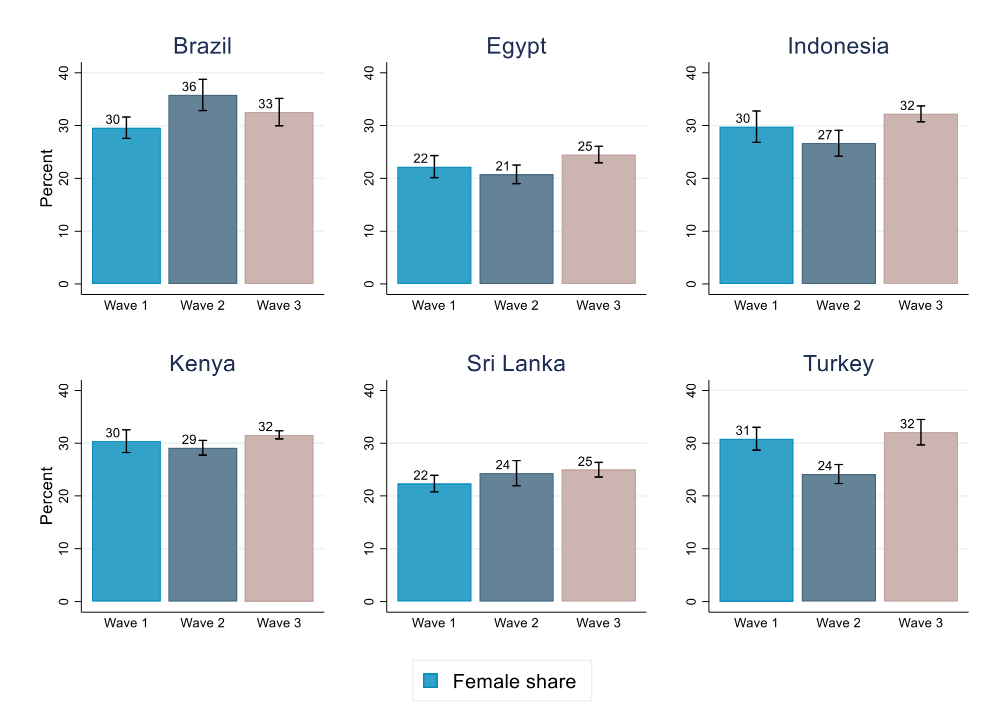


Figure A.4.2: Comparison of urban share across the three Internet Survey waves

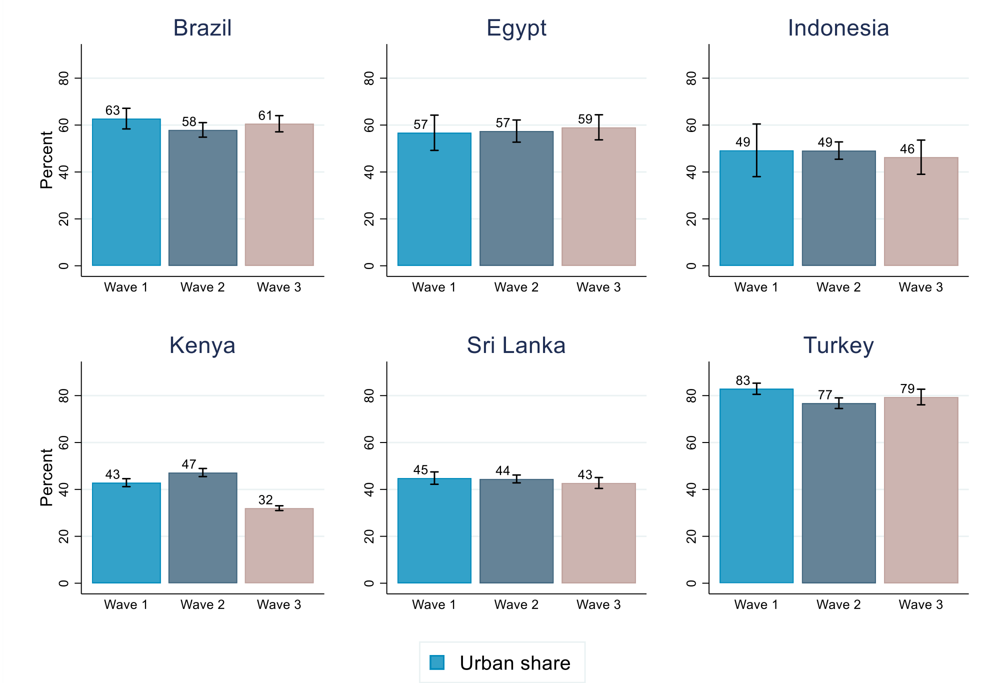


Figure A.4.3: Comparison of age distribution across three Internet Survey waves

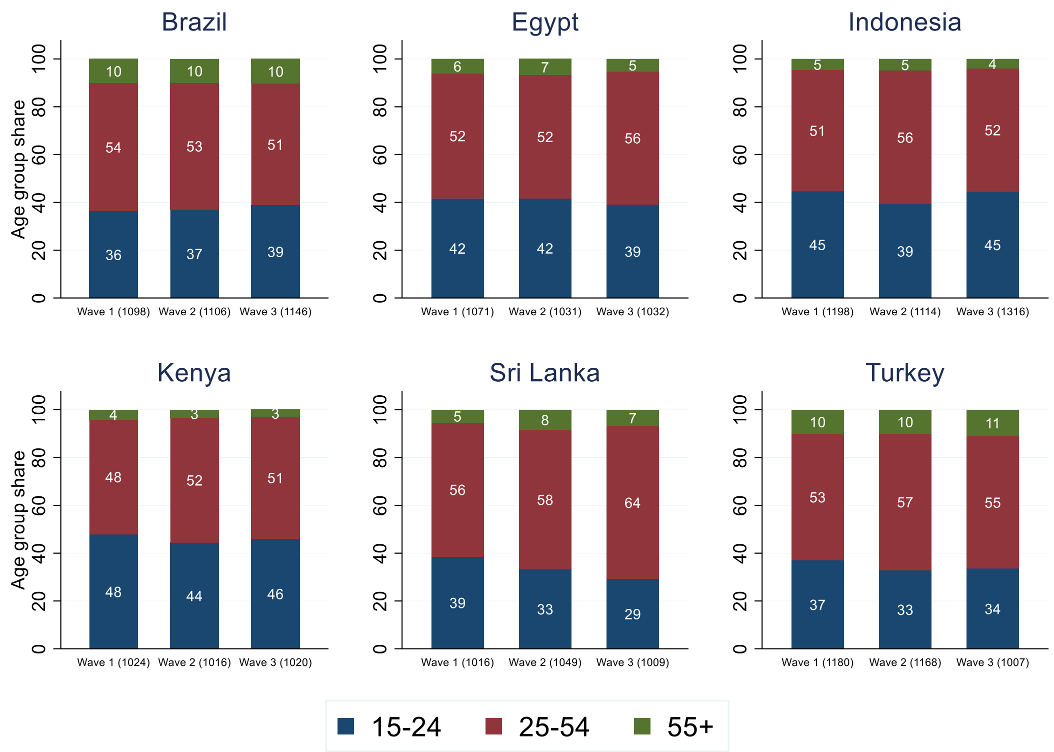


Figure A.4.4: Comparison of education categories across three Internet Survey waves

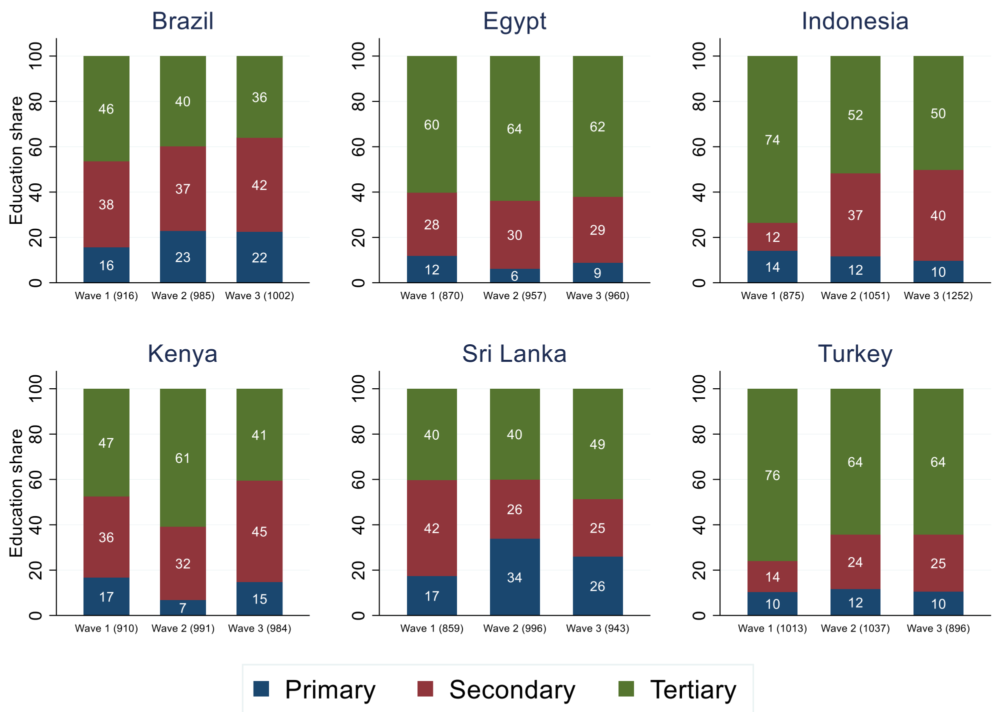


Figure A.4.5: Comparison of household size distribution across three Internet Survey waves

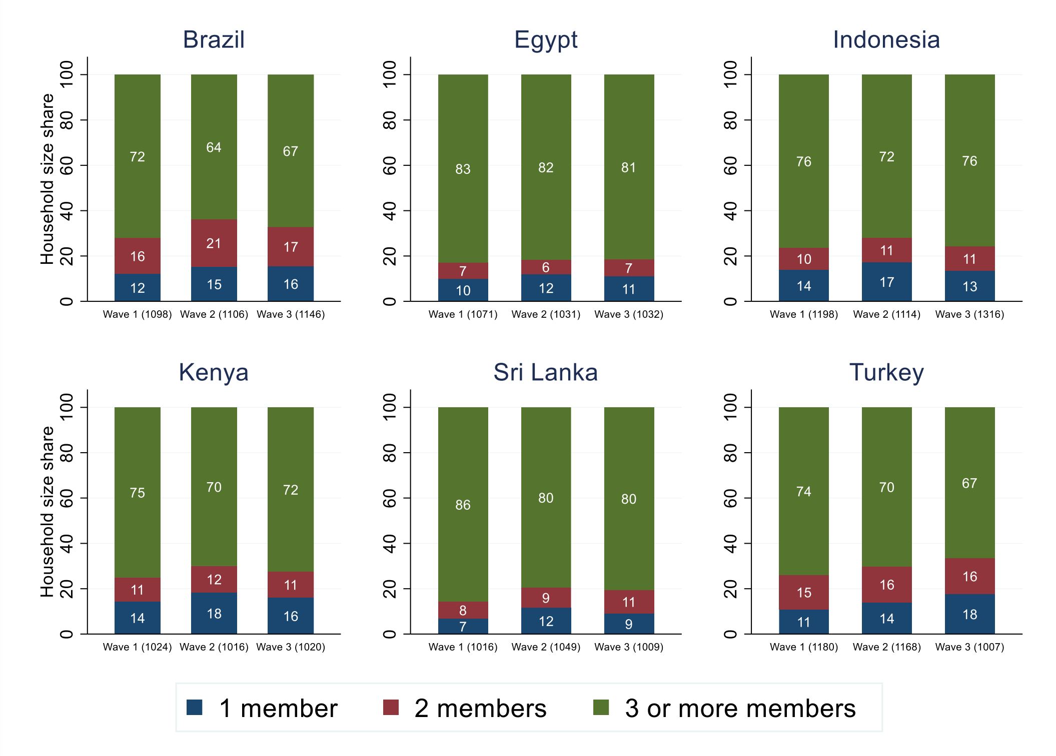


Table A.4.1: Comparing weighted Internet Survey data with GMD for the model variables, Brazil

	GMD	RIWI-wave1	RIWI-wave2	RIWI-wave3
Age between 15 and 24	20.05%	20.80%	22.31%	22.51%
	(18.91,21.18)	(18.57,23.04)	(19.06,25.55)	(16.86,28.15)
Age between 25 and 54	57.24%	60.45%	62.45%	65.32%
	(56.54,57.94)	(56.32,64.57)	(56.80,68.09)	(57.04,73.60)
Age between 55 and 100	22.71%	18.75%	15.24%	12.17%
	(21.14,24.27)	(13.24,24.26)	(9.75,20.74)	(5.91,18.43)
Female	52.85%	53.16%	48.61%	45.95%
	(52.48,53.23)	(48.78,57.53)	(44.91,52.30)	(41.09,50.81)
Urban	91.30%	90.89%	90.23%	94.24%
	(87.70,94.89)	(87.41,94.38)	(85.46,94.99)	(92.01,96.46)
No education	3.85%		3.20%	0.86%
	(2.65,5.06)		(-0.41,6.81)	(0.16,1.55)
Primary education	30.38%	25.89%	23.31%	21.79%
	(27.89,32.87)	(21.97,29.80)	(17.93,28.69)	(19.22,24.36)
Secondary education	39.85%	41.15%	43.74%	45.77%
	(38.52,41.17)	(38.08,44.22)	(38.52,48.96)	(41.87,49.68)
Availability of Landline phone	29.57%			29.68%
	(20.23,38.90)			(22.15,37.22)
Cellphone access	99.84%			99.76%

	(99.81,99.87)			(99.63,99.90)
Availability of Computer	52.64%			58.69%
	(46.82,58.46)			(51.68,65.71)
Piped water access	99.38%			99.19%
	(98.91,99.85)			(98.56,99.83)
Pre-covid sector: Not working	39.72%		37.91%	33.85%
	(37.19,42.25)		(31.47,44.34)	(28.67,39.02)
Pre-covid sector: Agriculture	3.74%		3.00%	2.54%
	(2.36,5.12)		(1.53,4.47)	(0.47,4.60)
Pre-covid sector: Industry	7.45%		7.58%	6.28%
	(5.90,9.00)		(5.17,9.99)	(2.58,9.99)
Pre-covid sector: Commerce	15.46%		16.58%	17.00%
	(14.93,15.99)		(14.74,18.42)	(11.60,22.41)
Pre-covid sector: Services	29.47%		30.55%	33.34%
	(27.47,31.47)		(23.51,37.59)	(27.21,39.46)
Pre-covid sector: Other sectors	4.14%		4.38%	6.99%
	(3.93,4.36)		(3.72,5.04)	(2.94,11.05)
One member household	4.85%	6.24%	4.47%	10.11%
	(4.33,5.36)	(4.24,8.23)	(2.96,5.97)	(4.90,15.32)
Two member household	20.90%	24.25%	20.14%	22.12%
	(19.55,22.25)	(19.92,28.58)	(16.01,24.27)	(17.17,27.07)
Three member household	29.15%	25.95%	29.00%	30.47%
	(28.38,29.91)	(20.49,31.41)	(24.09,33.91)	(19.55,41.40)
Four member household	25.55%	24.84%	28.12%	22.69%
	(24.77,26.33)	(18.58,31.10)	(21.41,34.82)	(16.37,29.02)
Five member household	11.64%	11.19%	11.63%	9.73%
	(10.69,12.59)	(7.85,14.53)	(7.41,15.86)	(5.43,14.03)
Six member household	4.45%	4.04%	3.27%	3.45%
	(3.88,5.01)	(2.00,6.08)	(1.00,5.55)	(1.12,5.78)
Seven member household	1.92%	1.90%	1.80%	0.69%
	(1.49,2.36)	(0.94,2.87)	(0.80,2.79)	(0.05,1.33)
Eight member household	1.55%	1.59%	1.57%	0.74%
	(1.07,2.02)	(0.76,2.43)	(0.85,2.29)	(-0.10,1.58)

Note: Cells in yellow indicate that the corresponding RIWI mean is statistically from GMD means.

Table A.4.2: Comparing weighted Internet Survey data with GMD for the model variables, Arab Republic of Egypt

	GMD	RIWI-wave1	RIWI-wave2	RIWI-wave3
Age between 15 and 24	31.71%	30.00%	35.13%	43.59%
	(29.91,33.51)	(25.40,34.60)	(27.64,42.63)	(33.30,53.88)
Age between 25 and 54	51.93%	52.14%	56.92%	52.95%
	(47.53,56.32)	(35.02,69.27)	(38.75,75.09)	(44.16,61.74)
Age between 55 and 100	16.36%	17.86%	7.95%	3.46%
	(10.45,22.27)	(0.25,35.47)	(-3.89,19.78)	(-4.77,11.69)
Female	49.25%	51.31%	43.23%	44.87%
	(47.95,50.55)	(40.75,61.87)	(34.20,52.26)	(23.39,66.35)
Urban	66.15%	72.91%	72.10%	72.82%
	(12.17,120.14)	(49.62,96.20)	(52.01,92.18)	(20.75,124.88)
No education	6.80%		1.35%	0.40%
	(1.58,12.02)		(-0.08,2.77)	(0.21,0.59)
Primary education	8.68%	6.70%	5.02%	1.77%
	(7.27,10.10)	(1.96,11.44)	(1.92,8.12)	(1.67,1.88)
Secondary education	45.13%	43.37%	47.13%	39.11%
	(41.10,49.15)	(33.82,52.91)	(38.07,56.19)	(30.56,47.66)
Availability of Landline phone	49.74%			51.69%
	(30.88,68.59)			(45.28,58.10)
Cellphone access	99.86%			99.87%
	(99.60,100.11)			(99.72,100.02)
Availability of Computer	72.69%			74.85%
	(59.88,85.50)			(53.83,95.87)
Piped water access	94.87%			94.48%
	(84.94,104.79)			(93.52,95.43)
Pre-covid sector: Not working	56.40%		56.32%	50.92%
	(51.37,61.43)		(43.91,68.73)	(28.74,73.11)
Pre-covid sector: Agriculture	4.03%		1.47%	0.86%
	(-1.55,9.60)		(-0.16,3.10)	(0.10,1.61)
Pre-covid sector: Industry	5.57%		5.67%	4.56%
	(3.47,7.67)		(3.09,8.25)	(-6.29,15.41)
Pre-covid sector: Commerce	6.44%		6.10%	8.81%
	(5.26,7.62)		(4.16,8.04)	(4.95,12.66)
Pre-covid sector: Services	24.44%		27.09%	31.71%
	(20.75,28.13)		(17.64,36.54)	(9.87,53.54)
Pre-covid sector: Other sectors	3.12%		3.36%	3.15%
	(2.60,3.64)		(2.57,4.16)	(0.37,5.92)
One member household	0.57%	0.42%	0.61%	0.58%
	(-0.14,1.29)	(-0.12,0.96)	(-0.07,1.29)	(-0.18,1.34)
Two member household	4.90%	4.56%	2.57%	4.34%
	(3.93,5.87)	(2.23,6.90)	(-1.59,6.73)	(-0.71,9.39)
Three member household	13.65%	16.16%	19.02%	14.86%

	(10.64,16.66)	(-5.08,37.40)	(3.79,34.25)	(7.76,21.95)
Four member household	25.74%	24.15%	24.51%	26.13%
	(19.90,31.58)	(13.05,35.24)	(17.51,31.51)	(2.36,49.91)
Five member household	28.48%	30.30%	25.97%	27.17%
	(25.34,31.62)	(10.94,49.65)	(15.83,36.11)	(21.46,32.87)
Six member household	14.42%	14.21%	17.05%	17.90%
	(9.56,19.28)	(8.16,20.25)	(8.90,25.19)	(-1.08,36.88)
Seven member household	7.75%	6.41%	6.70%	5.60%
	(4.16,11.34)	(3.93,8.89)	(6.00,7.41)	(1.23,9.98)
Eight member household	4.48%	3.80%	3.56%	3.42%
	(-0.92,9.88)	(0.40,7.20)	(-0.97,8.10)	(2.30,4.54)

Note: Cells in yellow indicate that the corresponding RIWI mean is statistically from GMD means.

Table A.4.3: Comparing weighted Internet Survey data with GMD for the model variables, Indonesia

	GMD	RIWI-wave1	RIWI-wave2	RIWI-wave3
Age between 15 and 24	28.15%	27.63%	28.80%	32.59%
	(26.42,29.88)	(24.86,30.41)	(25.07,32.53)	(27.58,37.60)
Age between 25 and 54	63.93%	67.54%	66.33%	64.83%
	(62.82,65.04)	(65.34,69.74)	(60.97,71.69)	(59.42,70.25)
Age between 55 and 100	7.92%	4.82%	4.87%	2.57%
	(6.94,8.89)	(2.77,6.88)	(2.59,7.15)	(1.00,4.15)
Female	49.12%	45.67%	48.11%	51.48%
	(48.62,49.63)	(41.24,50.09)	(40.79,55.43)	(47.07,55.88)
Urban	84.36%	85.59%	83.34%	81.03%
	(78.74,89.97)	(81.26,89.92)	(78.28,88.41)	(75.19,86.86)
No education	0.17%		0.18%	0.20%
	(0.05,0.29)		(0.01,0.36)	(0.08,0.33)
Primary education	9.42%	10.05%	7.64%	7.96%
	(7.83,11.01)	(6.02,14.08)	(5.19,10.08)	(1.79,14.12)
Secondary education	12.60%	8.56%	13.12%	17.52%
	(11.28,13.91)	(6.23,10.89)	(7.84,18.39)	(12.13,22.91)
Availability of Landline phone	10.93%			10.79%
	(8.26,13.60)			(7.07,14.52)
Availability of Computer	55.92%			46.49%
	(54.57,57.26)			(40.52,52.47)
Availability of Car	33.53%			24.42%
	(31.84,35.23)			(17.22,31.61)
Availability of Airconditioning	29.46%			22.75%
	(23.34,35.59)			(13.55,31.95)
Availability of Fridge	80.66%			73.16%
	(78.39,82.93)			(67.02,79.31)
sectorprecovid2==Not working	34.54%		37.41%	38.35%
	(33.20,35.87)		(29.82,44.99)	(29.02,47.68)

sectorprecovid2==Agriculture	1.06%		1.27%	1.48%
	(0.60,1.52)		(0.62,1.93)	(0.77,2.18)
sectorprecovid2==Industry	1.47%		0.86%	1.98%
	(0.51,2.43)		(0.25,1.47)	(0.69,3.27)
sectorprecovid2==Commerce	1.08%		0.89%	1.36%
	(0.65,1.51)		(0.67,1.11)	(0.62,2.10)
sectorprecovid2==Services	53.77%		59.48%	56.72%
	(52.67,54.86)		(52.64,66.32)	(46.53,66.91)
sectorprecovid2==Other sectors	0.09%		0.08%	0.12%
	(0.04,0.15)		(0.02,0.14)	(0.05,0.19)
One member household	6.23%	6.09%	6.28%	6.96%
	(5.09,7.38)	(4.15,8.03)	(3.93,8.63)	(3.57,10.35)
Two member household	15.37%	13.82%	14.92%	12.75%
	(14.73,16.02)	(11.40,16.24)	(9.67,20.17)	(5.61,19.89)
Three member household	30.47%	35.16%	33.44%	36.22%
	(29.31,31.63)	(30.62,39.70)	(19.72,47.16)	(21.26,51.19)
Four member household	26.40%	26.76%	25.07%	24.55%
	(25.36,27.44)	(22.03,31.50)	(19.18,30.95)	(16.77,32.33)
Five member household	12.69%	11.55%	10.45%	11.89%
	(11.92,13.47)	(9.74,13.37)	(6.15,14.74)	(9.19,14.60)
Six member household	5.38%	4.03%	6.12%	4.15%
	(4.47,6.30)	(3.09,4.98)	(3.21,9.02)	(1.33,6.97)
Seven member household	1.71%	1.27%	1.83%	1.77%
	(1.36,2.05)	(0.88,1.66)	(0.38,3.28)	(1.05,2.48)
Eight member household	1.74%	1.32%	1.89%	1.71%
	(1.17,2.31)	(0.61,2.02)	(0.50,3.29)	(0.95,2.46)

Note: Cells in yellow indicate that the corresponding RIWI mean is statistically from GMD means.

Table A.4.4: Comparing weighted Internet Survey data with GMD for the model variables, Kenya

	GMD	RIWI-wave1	RIWI-wave2	RIWI-wave3
Age between 15 and 24	36.49%	26.28%	28.93%	21.18%
	(33.96,39.03)	(15.83,36.73)	(22.56,35.30)	(-0.34,42.71)
Age between 25 and 54	55.92%	72.70%	68.86%	77.71%
	(52.07,59.76)	(61.61,83.80)	(61.61,76.11)	(54.81,100.62)
Age between 55 and 100	7.59%	1.02%	2.21%	1.11%
	(6.12,9.06)	(-0.35,2.38)	(0.87,3.56)	(-1.02,3.23)
Female	48.98%	46.41%	38.73%	21.22%
	(48.05,49.91)	(29.73,63.08)	(28.39,49.08)	(-1.28,43.71)
Urban	62.73%	91.17%	81.78%	92.65%
	(41.75,83.70)	(87.69,94.65)	(69.10,94.45)	(85.77,99.53)
No education	2.04%		0.52%	0.16%
	(1.00,3.07)		(0.01,1.03)	(-0.13,0.44)

Primary education	26.46%	22.48%	16.46%	44.90%
	(20.57,32.35)	(10.15,34.81)	(6.38,26.54)	(-6.94,96.74)
Secondary education	42.57%	35.68%	38.69%	16.19%
	(40.83,44.32)	(18.94,52.42)	(29.03,48.35)	(-5.03,37.42)
Own Computer	19.84%			25.27%
	(15.21,24.47)			(-3.92,54.46)
Access to piped water	62.29%			90.25%
	(44.08,80.50)			(79.78,100.73)
Has a modern floor	78.36%			94.57%
	(67.19,89.52)			(89.25,99.89)
Has a modern wall	71.26%			83.90%
	(65.15,77.38)			(66.44,101.35)
Worked pre-COVID	74.91%		81.22%	89.84%
	(73.28,76.54)		(74.72,87.73)	(76.98,102.70)
One member household	8.68%	16.22%	15.28%	12.00%
	(6.98,10.38)	(1.13,31.32)	(4.25,26.30)	(-3.58,27.57)
Two member household	11.71%	9.98%	13.96%	16.08%
	(7.37,16.04)	(3.26,16.70)	(6.53,21.39)	(-3.20,35.37)
Three member household	15.65%	14.39%	17.14%	51.20%
	(12.57,18.74)	(2.69,26.09)	(8.67,25.61)	(4.08,98.33)
Four member household	15.47%	16.45%	15.83%	4.81%
	(14.04,16.90)	(2.88,30.02)	(6.29,25.36)	(-3.15,12.77)
Five member household	15.51%	24.80%	16.18%	6.50%
	(13.56,17.47)	(4.67,44.92)	(8.74,23.61)	(-2.46,15.46)
Six member household	12.12%	12.85%	13.52%	4.55%
	(9.31,14.93)	(0.76,24.94)	(-1.71,28.76)	(-0.89,9.99)
Seven member household	7.61%	2.08%	3.80%	1.47%
	(4.97,10.25)	(0.19,3.97)	(0.67,6.93)	(-0.55,3.48)
Eight member household	13.25%	3.23%	4.29%	3.39%
	(7.60,18.90)	(1.07,5.39)	(1.07,7.52)	(-1.21,7.99)

Note: Cells in yellow indicate that the corresponding RIWI mean is statistically from GMD means.

Table A.4.5: Comparing weighted Internet Survey data with GMD for the model variables, Sri Lanka

	GMD	RIWI-wave1	RIWI-wave2	RIWI-wave3
Age between 15 and 24	18.55%	17.64%	27.85%	29.40%
	(17.19,19.90)	(13.84,21.45)	(4.57,51.13)	(17.93,40.86)
Age between 25 and 54	52.69%	53.56%	55.28%	58.01%
	(51.48,53.90)	(34.79,72.34)	(39.69,70.88)	(49.86,66.15)
Age between 55 and 100	28.76%	53.56%	16.86%	12.60%
	(26.92,30.60)	(34.79,72.34)	(7.37,26.35)	(9.12,16.08)
Female	54.34%	54.99%	50.12%	52.65%
	(53.39,55.30)	(40.11,69.86)	(39.16,61.09)	(48.14,57.15)
Urban	17.76%	20.41%	27.18%	Not included

	(1.45,34.07)	(-2.25,43.07)	(9.68,44.68)	
No education	3.01%		0.69%	0.76%
	(1.63,4.38)		(0.47,0.91)	(0.45,1.07)
Primary education	17.35%	21.68%	13.01%	10.07%
	(12.41,22.29)	(-4.00,47.36)	(9.76,16.25)	(6.96,13.18)
Secondary education	76.32%	71.64%	80.47%	83.08%
	(71.33,81.30)	(48.46,94.81)	(76.66,84.28)	(79.16,87.00)
Toilet availability	99.21%			98.98%
	(98.80,99.63)			(96.61,101.35)
Piped water facility availability	34.90%			53.05%
	(20.71,49.09)			(34.89,71.20)
Pre-covid sector: Not working	50.02%		53.32%	
	(47.70,52.33)		(41.32,65.32)	
Pre-covid sector: Agriculture	12.39%		1.85%	
	(4.95,19.83)		(0.79,2.90)	
Pre-covid sector: Industry	9.43%		8.42%	
	(7.09,11.78)		(1.78,15.07)	
Pre-covid sector: Commerce	6.89%		8.26%	
	(5.42,8.36)		(4.85,11.66)	
Pre-covid sector: Services	17.33%		24.48%	
	(13.81,20.84)		(13.24,35.72)	
Pre-covid sector: Other sectors	3.94%		3.68%	
	(3.54,4.33)		(1.95,5.40)	
Worked pre-COVID	49.98%			50.16%
	(47.67,52.30)			(43.45,56.87)
One member household	2.19%	1.03%	2.39%	0.80%
	(1.93,2.44)	(-0.34,2.40)	(1.87,2.92)	(0.51,1.10)
Two member household	10.37%	21.55%	6.89%	6.73%
	(9.22,11.52)	(-0.75,43.85)	(4.41,9.38)	(3.37,10.08)
Three member household	18.65%	23.98%	24.64%	18.18%
	(17.55,19.75)	(7.38,40.57)	(3.62,45.67)	(9.28,27.07)
Four member household	26.81%	19.85%	26.99%	30.60%
	(25.55,28.06)	(10.79,28.91)	(17.37,36.60)	(16.80,44.40)
Five member household	21.69%	20.55%	21.15%	25.95%
	(20.89,22.49)	(8.17,32.94)	(8.33,33.97)	(24.15,27.75)
Six member household	11.78%	6.10%	10.43%	7.02%
	(10.79,12.78)	(0.11,12.09)	(7.32,13.53)	(4.76,9.27)
Seven member household	4.96%	4.02%	5.40%	2.18%
	(4.03,5.88)	(-0.89,8.92)	(0.93,9.86)	(1.21,3.15)
Eight member household	3.55%	2.93%	2.11%	8.55%
	(2.61,4.50)	(0.82,5.04)	(0.55,3.67)	(7.76,9.34)

Note: Cells in yellow indicate that the corresponding RIWI mean is statistically from GMD means.

Table A.4.6: Comparing weighted Internet Survey data with GMD for the model variables, Türkiye

	GMD	RIWI-wave1	RIWI-wave2	RIWI-wave3
Age between 15 and 24	20.51%	20.89%	23.11%	20.86%
	(19.96,21.07)	(18.38,23.39)	(20.69,25.53)	(18.35,23.38)
Age between 25 and 54	62.22%	60.24%	60.70%	64.32%
	(61.56,62.89)	(57.22,63.26)	(57.89,63.50)	(61.36,67.29)
Age between 55 and 100	17.26%	18.88%	16.19%	14.81%
	(16.74,17.78)	(16.46,21.29)	(14.07,18.30)	(12.61,17.01)
Female	49.85%	50.10%	50.23%	43.03%
	(49.16,50.54)	(47.02,53.19)	(47.36,53.11)	(39.97,46.09)
No education	7.61%		7.72%	5.12%
	(7.24,7.97)		(6.19,9.26)	(3.76,6.49)
Primary education	26.88%	26.39%	25.14%	23.62%
	(26.27,27.50)	(23.67,29.11)	(22.65,27.63)	(20.99,26.25)
Secondary education	45.37%	44.55%	45.01%	44.46%
	(44.68,46.06)	(41.48,47.61)	(42.15,47.87)	(41.39,47.53)
Availability of Computer	61.64%			56.11%
	(60.97,62.31)			(53.04,59.18)
Availability of Car	54.43%			43.94%
	(53.74,55.11)			(40.87,47.01)
Availability of Airconditioning	23.72%			21.88%
	(23.13,24.30)			(19.32,24.44)
Availability of Cable TV	18.79%			25.04%
	(18.25,19.33)			(22.36,27.73)
Ownership of a stove	29.15%			29.09%
	(28.53,29.78)			(26.28,31.90)
Pre-covid sector: Not working	48.64%		50.05%	42.98%
	(47.95,49.33)		(47.18,52.92)	(39.91,46.04)
Pre-covid sector: Agriculture	5.20%		2.97%	4.38%
	(4.89,5.50)		(1.99,3.94)	(3.11,5.64)
Pre-covid sector: Industry	11.16%		11.24%	8.97%
	(10.73,11.60)		(9.43,13.06)	(7.21,10.74)
Pre-covid sector: Commerce	7.84%		7.62%	7.53%
	(7.47,8.21)		(6.10,9.15)	(5.90,9.17)
Pre-covid sector: Services	23.77%		24.92%	32.53%
	(23.19,24.36)		(22.43,27.40)	(29.63,35.43)
Pre-covid sector: Other sectors	3.39%		3.20%	3.61%
	(3.14,3.64)		(2.19,4.21)	(2.45,4.76)
One member household	4.05%	6.11%	4.11%	6.72%
	(3.78,4.32)	(4.63,7.58)	(2.97,5.24)	(5.17,8.27)
Two member household	12.37%	13.94%	14.73%	14.51%
	(11.92,12.82)	(11.80,16.07)	(12.69,16.77)	(12.33,16.69)
Three member household	22.46%	22.76%	23.09%	25.03%
	(21.89,23.04)	(20.17,25.35)	(20.67,25.51)	(22.35,27.71)

Four member household	26.85%	25.29%	25.70%	25.62%
	(26.24,27.46)	(22.61,27.97)	(23.19,28.21)	(22.92,28.32)
Five member household	15.50%	12.97%	16.12%	14.46%
	(15.00,16.00)	(10.89,15.04)	(14.01,18.23)	(12.28,16.63)
Six member household	10.28%	10.23%	9.75%	5.87%
	(9.86,10.70)	(8.36,12.10)	(8.05,11.45)	(4.42,7.33)
Seven member household	4.98%	5.08%	3.27%	4.64%
	(4.68,5.28)	(3.73,6.43)	(2.25,4.30)	(3.34,5.94)
Eight member household	3.50%	3.63%	3.24%	3.16%
	(3.25,3.76)	(2.48,4.79)	(2.22,4.25)	(2.07,4.24)

Note: Cells in yellow indicate that the corresponding RIWI mean is statistically from GMD means.

Table A.4.7: Summary comparison in models after dropping a variable, Brazil

	GMD	RIWI-wave1	RIWI-wave2	RIWI-wave3
Age between 15 and 24	20.05%	36.20%	38.03%	31.32%
	(18.91,21.18)	(30.88,41.51)	(31.60,44.46)	(19.53,43.12)
Age between 25 and 54	57.24%	56.07%	54.36%	58.27%
	(56.54,57.94)	(51.52,60.63)	(48.50,60.21)	(43.83,72.71)
Age between 55 and 100	22.71%	7.73%	7.61%	10.41%
	(21.14,24.27)	(5.90,9.55)	(5.69,9.53)	(5.10,15.72)
Female	52.85%	34.21%	34.01%	30.27%
	(52.48,53.23)	(29.04,39.38)	(28.21,39.81)	(24.85,35.69)
Urban	91.30%	61.16%	53.11%	92.67%
	(87.70,94.89)	(52.51,69.82)	(46.18,60.05)	(89.21,96.13)
No education	3.85%		6.33%	3.02%
	(2.65,5.06)		(2.12,10.53)	(0.87,5.18)
Primary education	30.38%	15.00%	14.88%	13.25%
	(27.89,32.87)	(11.42,18.57)	(10.40,19.36)	(7.18,19.32)
Secondary education	39.85%	33.53%	31.58%	36.07%
	(38.52,41.17)	(29.87,37.19)	(27.99,35.16)	(29.54,42.59)
Availability of Landline phone	29.57%			26.12%
	(20.23,38.90)			(19.41,32.83)
Cellphone access	99.84%			75.90%
	(99.81,99.87)			(69.79,82.02)
Availability of Computer	52.64%			52.41%
	(46.82,58.46)			(45.61,59.22)
Piped water access	99.38%			79.39%
	(98.91,99.85)			(69.33,89.46)
Pre-covid sector: Not working	39.72%		35.20%	26.99%
	(37.19,42.25)		(29.91,40.50)	(22.37,31.61)
Pre-covid sector: Agriculture	3.74%		3.18%	4.44%
	(2.36,5.12)		(1.90,4.45)	(1.85,7.02)
Pre-covid sector: Industry	7.45%		7.69%	4.30%
	(5.90,9.00)		(2.49,12.89)	(2.15,6.45)

Pre-covid sector: Commerce	15.46%		15.21%	18.97%
	(14.93,15.99)		(11.98,18.45)	(11.93,26.00)
Pre-covid sector: Services	29.47%		20.82%	21.06%
	(27.47,31.47)		(15.28,26.37)	(15.97,26.14)
Pre-covid sector: Other sectors	4.14%		17.90%	24.26%
	(3.93,4.36)		(12.54,23.25)	(16.74,31.78)
One member household	4.85%	15.81%	11.15%	18.94%
	(4.33,5.36)	(12.18,19.43)	(8.15,14.15)	(10.82,27.06)
Two member household	20.90%	22.68%	24.45%	22.83%
	(19.55,22.25)	(18.74,26.62)	(20.39,28.52)	(17.41,28.24)
Three member household	29.15%	19.00%	19.51%	23.06%
	(28.38,29.91)	(14.56,23.44)	(15.68,23.35)	(14.85,31.27)
Four member household	25.55%	16.60%	22.71%	18.54%
	(24.77,26.33)	(11.70,21.49)	(16.05,29.37)	(12.98,24.10)
Five member household	11.64%	9.91%	11.53%	8.96%
	(10.69,12.59)	(6.71,13.10)	(7.12,15.94)	(4.69,13.23)
Six member household	4.45%	4.78%	2.41%	5.34%
	(3.88,5.01)	(3.07,6.49)	(0.69,4.14)	(2.02,8.65)
Seven member household	1.92%	4.08%	3.06%	0.91%
	(1.49,2.36)	(2.07,6.08)	(1.36,4.75)	(-0.05,1.88)
Eight member household	1.55%	7.15%	5.17%	1.42%
	(1.07,2.02)	(5.22,9.07)	(2.39,7.96)	(0.26,2.57)

Note: Cells in yellow indicate that the corresponding RIWI mean is statistically from GMD means.

Table A.4.8: Summary comparison in models after dropping a variable, Arab Republic of Egypt

	GMD	RIWI-wave1	RIWI-wave2	RIWI-wave3
Age between 15 and 24	31.71%	40.33%	44.76%	44.70%
	(29.91,33.51)	(34.80,45.86)	(31.50,58.01)	(36.71,52.68)
Age between 25 and 54	51.93%	54.76%	51.94%	50.89%
	(47.53,56.32)	(46.63,62.89)	(34.99,68.89)	(41.76,60.02)
Age between 55 and 100	16.36%	4.91%	3.30%	4.41%
	(10.45,22.27)	(2.12,7.69)	(-2.16,8.77)	(-4.12,12.94)
Female	49.25%	24.60%	not converging	30.53%
	(47.95,50.55)	(17.82,31.38)		(-4.00,65.06)
Urban	66.15%	49.59%	51.26%	57.81%
	(12.17,120.14)	(17.84,81.35)	(26.42,76.11)	(-6.72,122.34)
No education	6.80%		2.64%	1.27%
	(1.58,12.02)		(1.78,3.51)	(-0.58,3.11)
Primary education	8.68%	8.02%	3.24%	2.89%
	(7.27,10.10)	(1.27,14.76)	(0.32,6.16)	(-0.49,6.28)
Secondary education	45.13%	22.36%	23.44%	21.16%
	(41.10,49.15)	(16.48,28.24)	(20.07,26.81)	(2.81,39.50)
Availability of Landline phone	49.74%			50.19%
	(30.88,68.59)			(44.16,56.22)

Cellphone access	99.86%			74.43%
	(99.60,100.11)			(44.72,104.13)
Availability of Computer	72.69%			60.70%
	(59.88,85.50)			(36.73,84.66)
Piped water access	94.87%			56.58%
	(84.94,104.79)			(39.66,73.50)
Pre-covid sector: Not working	56.40%		44.52%	35.37%
	(51.37,61.43)		(38.21,50.82)	(18.95,51.79)
Pre-covid sector: Agriculture	4.03%		3.76%	5.10%
	(-1.55,9.60)		(2.51,5.01)	(-13.76,23.97)
Pre-covid sector: Industry	5.57%		5.34%	5.33%
	(3.47,7.67)		(1.05,9.63)	(-9.31,19.97)
Pre-covid sector: Commerce	6.44%		5.95%	15.49%
	(5.26,7.62)		(5.00,6.91)	(3.10,27.88)
Pre-covid sector: Services	24.44%		26.65%	26.15%
	(20.75,28.13)		(21.27,32.03)	(12.04,40.25)
Pre-covid sector: Other sectors	3.12%		13.78%	12.56%
	(2.60,3.64)		(10.99,16.57)	(3.86,21.26)
One member household	0.57%	6.58%	12.24%	10.04%
	(-0.14,1.29)	(5.14,8.03)	(0.79,23.69)	(4.55,15.53)
Two member household	4.90%	9.66%	2.64%	5.51%
	(3.93,5.87)	(0.36,18.95)	(-3.59,8.88)	(-2.21,13.22)
Three member household	13.65%	10.74%	21.05%	14.46%
	(10.64,16.66)	(-0.60,22.08)	(1.86,40.25)	(5.89,23.03)
Four member household	25.74%	19.89%	17.95%	18.10%
	(19.90,31.58)	(16.86,22.92)	(14.05,21.84)	(13.75,22.45)
Five member household	28.48%	29.07%	17.86%	16.07%
	(25.34,31.62)	(6.34,51.80)	(7.46,28.27)	(1.50,30.64)
Six member household	14.42%	11.79%	10.89%	16.08%
	(9.56,19.28)	(6.82,16.75)	(2.73,19.05)	(-5.31,37.48)
Seven member household	7.75%	4.21%	6.96%	5.91%
	(4.16,11.34)	(0.11,8.30)	(4.76,9.16)	(-2.13,13.95)
Eight member household	4.48%	8.07%	10.41%	13.82%
	(-0.92,9.88)	(7.70,8.44)	(1.66,19.16)	(-12.82,40.47)

Note: Cells in yellow indicate that the corresponding RIWI mean is statistically from GMD means.

Table A.4.9: Summary comparison in models after dropping a variable, Indonesia

	GMD	RIWI-wave1	RIWI-wave2	RIWI-wave3
Age between 15 and 24	28.15%	38.83%	36.81%	42.62%
	(26.42,29.88)	(34.39,43.28)	(28.81,44.81)	(31.72,53.52)
Age between 25 and 54	63.93%	57.54%	57.10%	53.85%
	(62.82,65.04)	(52.00,63.09)	(47.29,66.91)	(42.41,65.29)
Age between 55 and 100	7.92%	3.62%	6.09%	3.53%
	(6.94,8.89)	(1.78,5.47)	(3.23,8.95)	(1.02,6.03)

Female	49.12%	29.93%	30.38%	40.03%
	(48.62,49.63)	(26.33,33.53)	(24.08,36.67)	(32.81,47.25)
Urban	84.36%	49.78%	48.90%	50.96%
	(78.74,89.97)	(36.93,62.63)	(42.71,55.09)	(40.98,60.93)
No education	0.17%		4.14%	4.08%
	(0.05,0.29)		(1.16,7.12)	(1.80,6.37)
Primary education	9.42%	13.00%	7.74%	6.13%
	(7.83,11.01)	(7.63,18.38)	(4.35,11.14)	(2.67,9.59)
Secondary education	12.60%	6.75%	29.56%	35.09%
	(11.28,13.91)	(3.14,10.36)	(22.51,36.61)	(26.79,43.38)
Availability of Landline phone	10.93%			21.51%
	(8.26,13.60)			(15.53,27.49)
Availability of Computer	55.92%			28.34%
	(54.57,57.26)			(22.79,33.90)
Availability of Car	33.53%			24.41%
	(31.84,35.23)			(17.22,31.61)
Availability of Airconditioning	29.46%			22.72%
	(23.34,35.59)			(13.53,31.91)
Availability of Fridge	80.66%			49.67%
	(78.39,82.93)			(40.21,59.13)
sectorprecovid2==Not working	34.54%		22.36%	22.83%
	(33.20,35.87)		(15.03,29.68)	(18.41,27.25)
sectorprecovid2==Agriculture	1.06%		12.20%	8.15%
	(0.60,1.52)		(9.12,15.27)	(5.34,10.96)
sectorprecovid2==Industry	1.47%		8.26%	7.33%
	(0.51,2.43)		(3.82,12.71)	(1.40,13.26)
sectorprecovid2==Commerce	1.08%		15.43%	10.88%
	(0.65,1.51)		(9.02,21.85)	(7.67,14.09)
sectorprecovid2==Services	53.77%		32.46%	31.06%
	(52.67,54.86)		(26.36,38.56)	(18.27,43.84)
sectorprecovid2==Other sectors	0.09%		9.29%	19.75%
	(0.04,0.15)		(6.06,12.53)	(10.86,28.63)
One member household	6.23%	12.16%	16.22%	9.18%
	(5.09,7.38)	(8.03,16.30)	(9.21,23.23)	(4.80,13.56)
Two member household	15.37%	9.78%	10.99%	8.00%
	(14.73,16.02)	(7.59,11.97)	(6.80,15.17)	(3.38,12.61)
Three member household	30.47%	17.83%	15.07%	15.67%
	(29.31,31.63)	(13.76,21.90)	(9.16,20.99)	(8.80,22.54)
Four member household	26.40%	22.14%	21.38%	25.14%
	(25.36,27.44)	(18.17,26.12)	(17.53,25.22)	(19.34,30.95)
Five member household	12.69%	17.71%	12.23%	20.23%
	(11.92,13.47)	(12.22,23.20)	(8.45,16.01)	(16.34,24.13)
Six member household	5.38%	8.35%	12.65%	6.06%
	(4.47,6.30)	(4.14,12.55)	(6.62,18.68)	(1.93,10.19)
Seven member household	1.71%	6.32%	4.77%	9.69%

	(1.36,2.05)	(3.41,9.24)	(0.69,8.85)	(4.97,14.41)
Eight member household	1.74%	5.70%	6.69%	6.03%
	(1.17,2.31)	(2.72,8.69)	(4.29,9.09)	(4.54,7.52)

Note: Cells in yellow indicate that the corresponding RIWI mean is statistically from GMD means.

Table A.4.10: Summary comparison in models after dropping a variable, Kenya

	GMD	RIWI-wave1	RIWI-wave2	RIWI-wave3
Age between 15 and 24	36.49%	36.09%	43.44%	50.35%
	(33.96,39.03)	(18.04,54.14)	(29.94,56.94)	(19.32,81.38)
Age between 25 and 54	55.92%	63.03%	53.12%	49.20%
	(52.07,59.76)	(44.49,81.57)	(37.03,69.20)	(18.19,80.20)
Age between 55 and 100	7.59%	0.88%	3.44%	0.45%
	(6.12,9.06)	(-0.20,1.96)	(-0.86,7.74)	(-0.40,1.31)
Female	48.98%	43.54%	28.18%	11.87%
	(48.05,49.91)	(19.28,67.80)	(12.55,43.81)	(-2.09,25.84)
Urban	62.73%	43.65%	40.17%	49.77%
	(41.75,83.70)	(35.47,51.82)	(25.20,55.14)	(29.59,69.95)
No education	2.04%		1.01%	0.68%
	(1.00,3.07)		(0.26,1.75)	(-0.42,1.77)
Primary education	26.46%	13.64%	6.01%	15.09%
	(20.57,32.35)	(4.84,22.45)	(1.62,10.40)	(-8.71,38.88)
Secondary education	42.57%	22.49%	21.48%	14.68%
	(40.83,44.32)	(7.89,37.09)	(14.30,28.66)	(0.52,28.85)
Own Computer	19.84%			34.43%
	(15.21,24.47)			(-0.74,69.60)
Access to piped water	62.29%			77.90%
	(44.08,80.50)			(52.42,103.38)
Has a modern floor	78.36%			91.98%
	(67.19,89.52)			(83.12,100.85)
Has a modern wall	71.26%			77.01%
	(65.15,77.38)			(54.66,99.36)
Worked pre-COVID	74.91%		75.57%	81.37%
	(73.28,76.54)		(65.10,86.04)	(63.44,99.30)
One member household	8.68%	24.64%	26.12%	19.69%
	(6.98,10.38)	(5.08,44.20)	(4.84,47.41)	(-0.21,39.58)
Two member household	11.71%	7.38%	8.29%	10.02%
	(7.37,16.04)	(2.76,11.99)	(3.19,13.38)	(1.14,18.89)
Three member household	15.65%	9.47%	14.61%	41.70%
	(12.57,18.74)	(1.77,17.17)	(5.17,24.05)	(-2.12,85.52)
Four member household	15.47%	19.49%	15.18%	6.18%
	(14.04,16.90)	(3.53,35.45)	(5.13,25.23)	(-3.07,15.43)
Five member household	15.51%	16.29%	12.31%	7.72%
	(13.56,17.47)	(2.18,30.41)	(5.53,19.08)	(-1.19,16.63)
Six member household	12.12%	15.94%	16.26%	6.28%
	(9.31,14.93)	(-0.51,32.40)	(-0.87,33.39)	(-1.14,13.71)

Seven member household	7.61%	2.95%	3.36%	1.23%
	(4.97,10.25)	(0.19,5.70)	(0.10,6.62)	(-0.35,2.81)
Eight member household	13.25%	3.85%	3.87%	7.18%
	(7.60,18.90)	(0.79,6.90)	(-0.52,8.26)	(-3.48,17.84)

Note: Cells in yellow indicate that the corresponding RIWI mean is statistically from GMD means.

Table A.4.11: Summary comparison in models after dropping a variable, Sri Lanka

	GMD	RIWI-wave1	RIWI-wave2	RIWI-wave3
Age between 15 and 24	18.55%	42.84%	46.54%	38.73%
	(17.19,19.90)	(30.12,55.56)	(22.33,70.75)	(26.83,50.63)
Age between 25 and 54	52.69%	50.09%	48.63%	58.16%
	(51.48,53.90)	(33.76,66.41)	(28.17,69.09)	(47.24,69.08)
Age between 55 and 100	28.76%	7.07%	4.83%	3.11%
	(26.92,30.60)	(-0.61,14.75)	(0.75,8.91)	(1.80,4.42)
Female	54.34%	24.09%	19.36%	19.53%
	(53.39,55.30)	(10.55,37.62)	(9.79,28.94)	(7.61,31.45)
Urban	17.76%	36.71%	32.73%	
	(1.45,34.07)	(23.66,49.75)	(23.36,42.10)	
No education	3.01%		2.71%	1.19%
	(1.63,4.38)		(1.54,3.87)	(0.70,1.69)
Primary education	17.35%	31.07%	30.67%	9.79%
	(12.41,22.29)	(-6.02,68.16)	(19.58,41.76)	(6.46,13.13)
Secondary education	76.32%	27.59%	21.03%	82.14%
	(71.33,81.30)	(11.71,43.46)	(16.62,25.45)	(74.01,90.28)
Toilet availability	99.27%			51.06%
	(99.17,99.36)			(50.13,51.98)
Piped water facility availability	34.90%			76.71%
	(20.71,49.09)			(65.06,88.36)
Pre-covid sector: Not working	50.02%		57.79%	
	(47.70,52.33)		(51.04,64.53)	
Pre-covid sector: Agriculture	12.39%		3.21%	
	(4.95,19.83)		(0.92,5.50)	
Pre-covid sector: Industry	9.43%		7.36%	
	(7.09,11.78)		(-0.06,14.78)	
Pre-covid sector: Commerce	6.89%		9.98%	
	(5.42,8.36)		(-2.28,22.24)	
Pre-covid sector: Services	17.33%		15.81%	
	(13.81,20.84)		(1.73,29.90)	
Pre-covid sector: Other sectors	3.94%		5.85%	
	(3.54,4.33)		(1.48,10.22)	
Worked pre-COVID	65.43%			42.71%
	(60.36,70.50)			(40.88,44.54)
One member household	2.19%	2.43%	Not converging	5.77%
	(1.93,2.44)	(-0.52,5.39)		(2.71,8.84)
Two member household	10.37%	15.42%	Not converging	8.84%

	(9.22,11.52)	(-2.85,33.69)		(2.20,15.47)
Three member household	18.65%	34.53%	Not converging	14.11%
	(17.55,19.75)	(-2.14,71.20)		(4.59,23.63)
Four member household	26.81%	15.18%	Not converging	27.36%
	(25.55,28.06)	(4.81,25.55)		(15.65,39.07)
Five member household	21.69%	17.89%	Not converging	31.70%
	(20.89,22.49)	(-2.56,38.33)		(23.90,39.50)
Six member household	11.78%	3.40%	Not converging	3.43%
	(10.79,12.78)	(-0.57,7.37)		(1.76,5.10)
Seven member household	4.96%	4.79%	Not converging	2.65%
	(4.03,5.88)	(-3.12,12.71)		(1.14,4.15)
Eight member household	3.55%	6.36%	Not converging	6.14%
	(2.61,4.50)	(-5.98,18.70)		(1.54,10.75)

Note: Cells in yellow indicate that the corresponding RIWI mean is statistically from GMD means.

Table A.4.12: Summary comparison in models after dropping a variable, Türkiye

	GMD	RIWI-wave1	RIWI-wave2	RIWI-wave3
Age between 15 and 24	20.51%	38.55%	36.75%	34.38%
	(19.96,21.07)	(35.55,41.55)	(33.98,39.52)	(31.45,37.32)
Age between 25 and 54	62.22%	52.80%	54.17%	56.04%
	(61.56,62.89)	(49.72,55.88)	(51.31,57.03)	(52.97,59.12)
Age between 55 and 100	17.26%	8.65%	9.07%	9.57%
	(16.74,17.78)	(6.92,10.39)	(7.42,10.72)	(7.75,11.39)
Female	49.85%	28.43%	23.29%	24.65%
	(49.16,50.54)	(25.65,31.22)	(20.87,25.72)	(21.98,27.32)
No education	7.61%		8.32%	6.80%
	(7.24,7.97)		(6.74,9.91)	(5.25,8.36)
Primary education	26.88%	10.07%	10.26%	8.64%
	(26.27,27.50)	(8.21,11.92)	(8.52,12.01)	(6.91,10.38)
Secondary education	45.37%	11.65%	22.00%	19.51%
	(44.68,46.06)	(9.67,13.63)	(19.62,24.38)	(17.06,21.96)
Availability of Computer	61.64%			34.00%
	(60.97,62.31)			(31.07,36.93)
Availability of Car	54.43%			25.20%
	(53.74,55.11)			(22.51,27.89)
Availability of Airconditioning	23.72%			19.69%
	(23.13,24.30)			(17.23,22.15)
Availability of Cable TV	18.79%			46.85%
	(18.25,19.33)			(43.76,49.94)
Ownership of a stove	29.15%			39.19%
	(28.53,29.78)			(36.17,42.21)
Pre-covid sector: Not working	48.64%		47.98%	40.84%
	(47.95,49.33)		(45.11,50.85)	(37.80,43.88)

Pre-covid sector: Agriculture	5.20%		4.56%	8.91%
	(4.89,5.50)		(3.37,5.76)	(7.14,10.67)
Pre-covid sector: Industry	11.16%		4.81%	3.67%
	(10.73,11.60)		(3.58,6.03)	(2.51,4.84)
Pre-covid sector: Commerce	7.84%		6.98%	5.89%
	(7.47,8.21)		(5.51,8.44)	(4.44,7.35)
Pre-covid sector: Services	23.77%		25.66%	29.10%
	(23.19,24.36)		(23.15,28.17)	(26.29,31.91)
Pre-covid sector: Other sectors	3.39%		10.01%	12.25%
	(3.14,3.64)		(8.28,11.73)	(10.22,14.28)
One member household	4.05%	16.49%	9.80%	5.44%
	(3.78,4.32)	(14.20,18.77)	(8.09,11.50)	(4.04,6.84)
Two member household	12.37%	18.00%	14.33%	16.95%
	(11.92,12.82)	(15.63,20.37)	(12.32,16.34)	(14.63,19.27)
Three member household	22.46%	16.40%	18.55%	17.55%
	(21.89,23.04)	(14.12,18.69)	(16.32,20.78)	(15.20,19.90)
Four member household	26.85%	18.91%	19.56%	18.83%
	(26.24,27.46)	(16.50,21.33)	(17.28,21.84)	(16.41,21.25)
Five member household	15.50%	11.25%	15.14%	17.53%
	(15.00,16.00)	(9.30,13.20)	(13.08,17.20)	(15.18,19.88)
Six member household	10.28%	6.71%	9.25%	4.06%
	(9.86,10.70)	(5.17,8.25)	(7.59,10.91)	(2.84,5.28)
Seven member household	4.98%	5.44%	4.80%	4.18%
	(4.68,5.28)	(4.04,6.84)	(3.58,6.03)	(2.94,5.41)
Eight member household	3.50%	6.79%	8.57%	8.20%
	(3.25,3.76)	(5.24,8.35)	(6.96,10.18)	(6.50,9.90)

Note: Cells in yellow indicate that the corresponding RIWI mean is statistically from GMD means.

Table A.4.13.: Number of model variables and non-model variables from the Internet survey with means statistically same as GMD

A. Model variables (share whose RIWI means are statistically same as GMD)						
	Brazil	Egypt, Arab Rep.	Indonesia	Kenya	Sri Lanka	Türkiye
Weighted wave-1	80%	93%	47%	60%	80%	100%
Weighted wave-2	82%	91%	86%	65%	86%	100%
Weighted wave-3	88%	96%	81%	67%	83%	100%
B. Non-model variables						
	Brazil	Egypt, Arab Rep.	Indonesia	Kenya	Sri Lanka	Türkiye
Weighted wave-1	27%	33%	13%	47%	53%	100%
Weighted wave-2	50%	57%	23%	59%	50%	100%
Weighted wave-3	58%	62%	15%	67%	28%	100%

Note: These summarized are based on tables A.4.1 to A.4.12

Table A.4.14: Mean relative deviation of Internet survey from GMD for model variables

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Internet survey waves	Variable	Brazil	Egypt, Arab Rep.	Indonesia	Kenya	Sri Lanka	Türkiye
1-2-3	Age between 15 and 24	9.1%	17.9%	6.6%	30.2%	37.8%	5.4%
1-2-3	Age between 25 and 54	9.6%	4.0%	3.6%	30.7%	5.6%	3.0%
1-2-3	Age between 55 and 100	32.2%	46.5%	48.4%	80.9%	61.3%	10.7%
1-2-3	Gender	7.2%	8.4%	4.6%	27.6%	4.0%	5.7%
1-2	Urban	1.6%	9.8%	2.2%	41.1%	34.0%	25.0%
2-3	No education	64.8%	87.1%	11.8%	83.3%	75.9%	7.4%
1-2-3	Primary education	22.1%	48.2%	13.7%	40.8%	30.6%	1.5%
1-2-3	Secondary education	9.3%	7.2%	25.1%	29.1%	6.8%	5.4%
3	Availability of Landline phone	66.8%	3.9%	1.3%			
3	Cellphone access	66.7%	0.0%				
3	Availability of Computer	70.5%	3.0%	16.9%	27.4%		9.9%
3	Piped water access	66.7%	0.4%		44.9%	52.0%	
3	Availability of Car			27.2%			23.9%
3	Availability of Airconditioning			22.8%			8.4%
3	Has a modern floor				20.7%		
3	Has a modern wall				17.7%		
3	Availability of Fridge			9.3%			
3	Toilet availability					0.2%	
3	Availability of Cable TV						25.0%
3	Ownership of a stove						0.2%
2-3	Non-employed	39.8%	4.9%	9.7%		6.6%	8.0%
2-3	Agriculture	50.6%	71.1%	29.7%		85.1%	30.8%
2-3	Industry	39.1%	10.0%	38.1%		10.7%	12.6%
2-3	Commerce	39.1%	21.0%	21.8%		19.9%	3.5%
2-3	Services	38.9%	20.3%	8.1%		41.3%	15.9%

2-3	Other sector	58.2%	4.3%	22.2%		6.6%	5.8%
2-3	Worked pre-COVID				14.2%	0.4%	
1-2-3	One member household	48.3%	11.7%	4.9%	67.1%	41.9%	30.7%
1-2-3	Two members household	8.5%	22.0%	10.0%	23.8%	58.8%	15.5%
1-2-3	Three members household	5.3%	22.2%	14.7%	81.6%	21.1%	4.8%
1-2-3	Four members household	8.0%	4.2%	4.5%	25.9%	13.6%	5.0%
1-2-3	Five members household	6.8%	6.6%	11.0%	40.8%	9.1%	9.2%
1-2-3	Six members household	19.4%	14.6%	20.6%	26.7%	33.4%	26.9%
1-2-3	Seven members household	23.8%	19.5%	12.1%	67.8%	28.0%	14.6%
1-2-3	Eight members household	18.7%	19.8%	11.5%	72.6%	66.3%	7.3%

Note: The percentages are the absolute difference between the means of Internet survey and GMD divided by GMD mean

Table A.4.15: Mean relative deviation of Internet survey from GMD for the non-model variables

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Internet survey waves	Variable	Brazil	Egypt, Arab Rep.	Indonesia	Kenya	Sri Lanka	Türkiye
1-2-3	Age between 15 and 24	75.5%	36.4%	40.0%	19.4%	130.2%	78.3%
1-2-3	Age between 25 and 54	3.0%	2.5%	12.1%	9.9%	7.7%	12.7%
1-2-3	Age between 55 and 100	62.2%	74.3%	44.3%	79.1%	82.6%	47.3%
1-2-3	Gender	37.9%	44.0%	31.9%	43.1%	61.4%	48.9%
1-2-3	Urban	25.4%	20.1%	40.9%	29.0%	95.5%	10.0%
2-3	No education	43.0%	71.3%	1578.4%	58.6%	35.2%	64.1%
1-2-3	Primary education	52.7%	45.7%	30.3%	56.2%	66.5%	60.9%
1-2-3	Secondary education	15.4%	50.5%	119.8%	54.1%	48.0%	78.3%
3	Land phone	11.7%	0.9%	98.9%			
3	Cell phone	24.0%	25.5%				
3	Computer	0.4%	16.5%	83.1%	73.5%		44.8%
3	Piped	20.1%	40.4%		25.1%	119.8%	
3	car			75.7%			53.7%
3	ac			74.3%			17.0%
3	Floor modern				17.4%		
3	Wall modern				8.1%		
3	Fridge			79.5%			
3	Toilet accommodation					48.6%	

3	TV cable						149.3%
3	Stove						34.4%
2-3	Non-employed	21.7%	29.2%	56.4%		15.5%	8.7%
2-3	Agriculture	16.8%	16.6%	606.6%		74.1%	41.8%
2-3	Industry	22.8%	4.2%	320.2%		22.0%	62.0%
2-3	Commerce	12.2%	74.1%	778.7%		44.8%	17.9%
2-3	Services	28.9%	8.0%	60.6%		8.8%	15.2%
2-3	Other sector	409.2 %	322.1%	10722.2 %		48.5%	228.3%
2-3	Worked pre-COVID				4.7%	34.7%	
1-2-3	One-member household	215.5 %	1587.7%	101.0%	170.5 %	87.2%	161.2%
1-2-3	Two members household	11.6%	51.9%	37.6%	26.9%	31.7%	32.8%
1-2-3	Three members household	29.6%	27.2%	46.9%	70.9%	54.7%	22.1%
1-2-3	Four members household	24.5%	27.6%	13.3%	29.3%	22.7%	28.9%
1-2-3	Five members household	12.9%	27.6%	34.2%	25.3%	31.8%	14.3%
1-2-3	Six members household	24.4%	18.1%	67.7%	37.9%	71.0%	35.1%
1-2-3	Seven members household	74.8%	26.5%	305.1%	67%	25.0%	9.6%
1-2-3	Eight members household	201.1 %	140.3%	252.9%	62.5%	76.1%	124.4%

Note: The percentages are the absolute difference between the means of Internet survey and GMD divided by GMD mean

Figure A.4.6. Employment to population ratio, Brazil

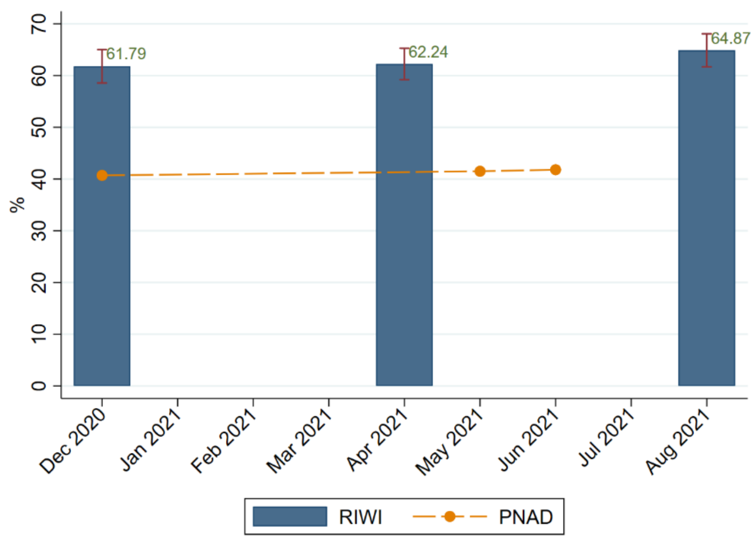


Figure A.4.7. Employment to population ratio, Arab Republic of Egypt

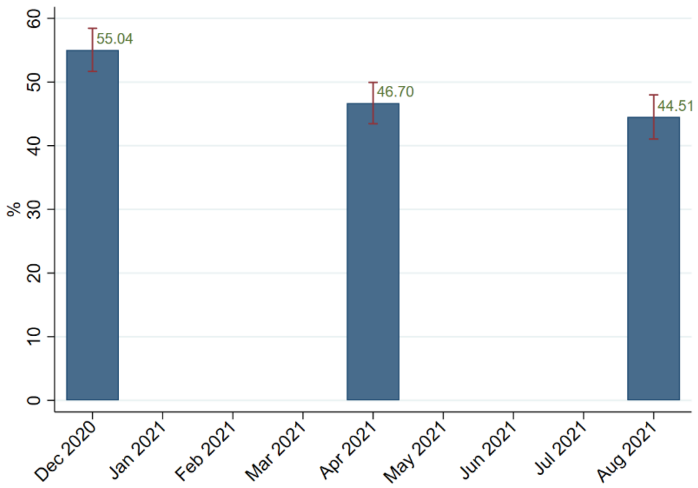


Figure A.4.8. Employment to population ratio, Indonesia

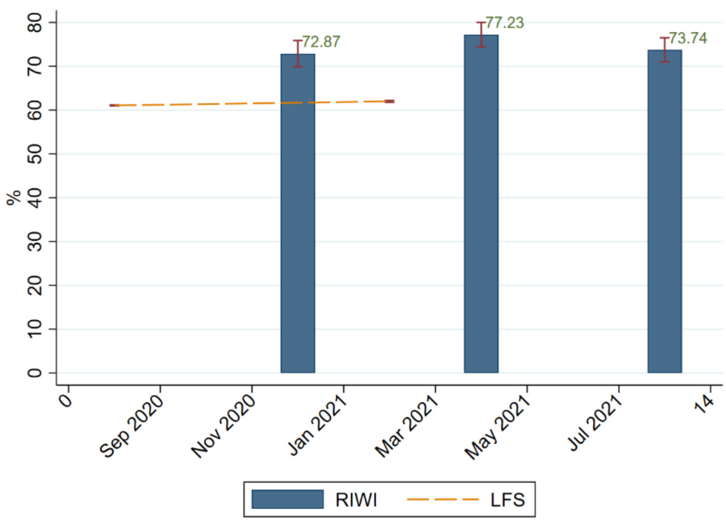


Figure A.4.9. Employment to population ratio, Kenya

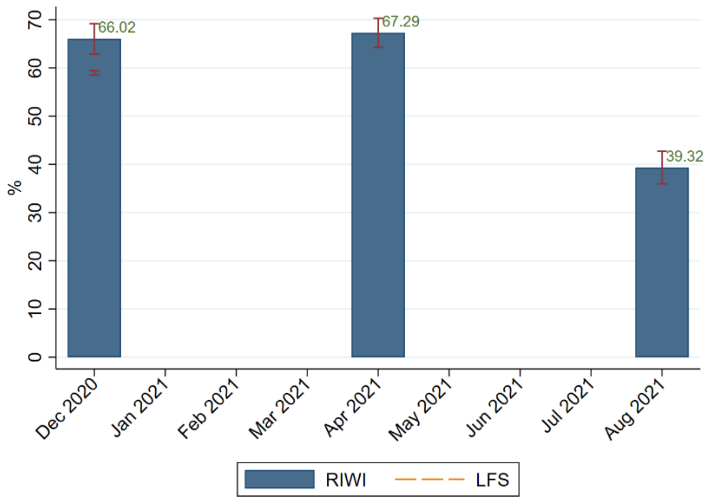


Figure A.4.10. Employment to population ratio, Sri Lanka

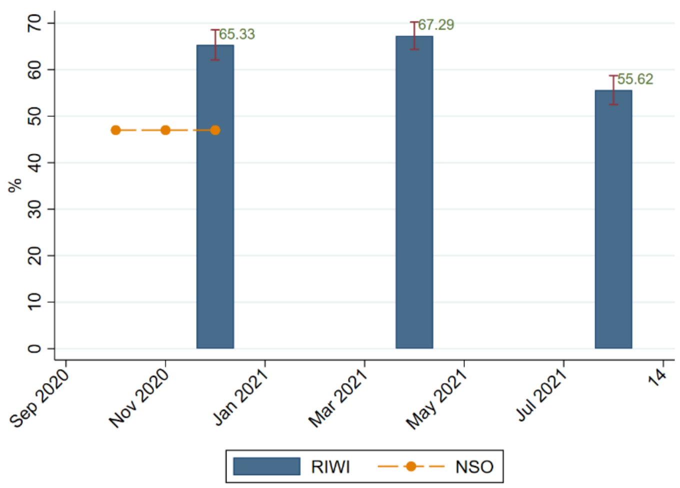


Figure A.4.11. Employment to population ratio, Türkiye

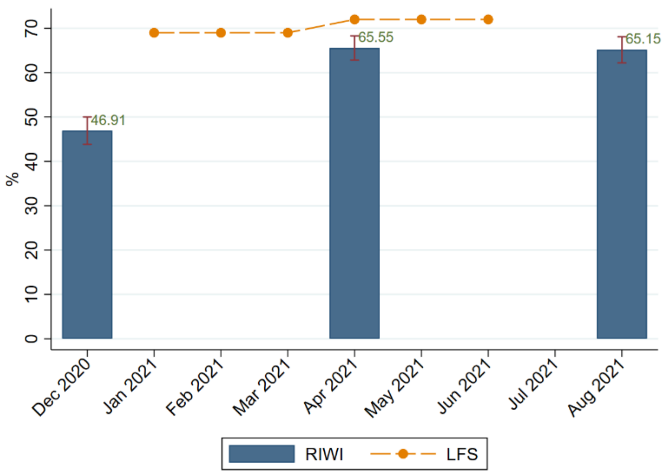


Figure A.4.12: Brazil, strictly formal share

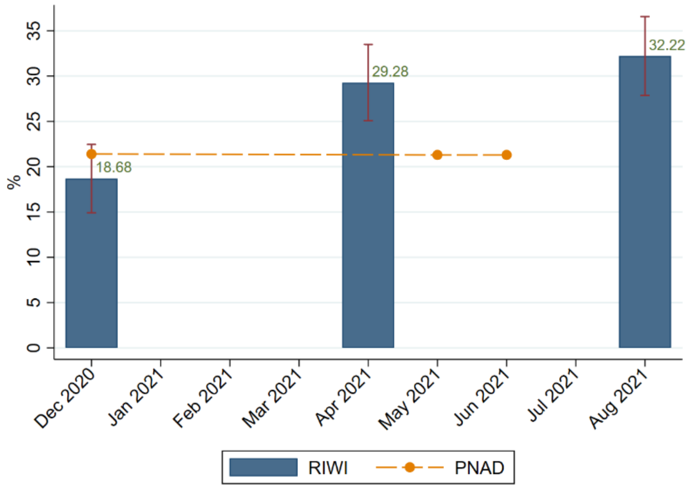


Figure A.4.13: Arab Republic of Egypt, strictly formal share

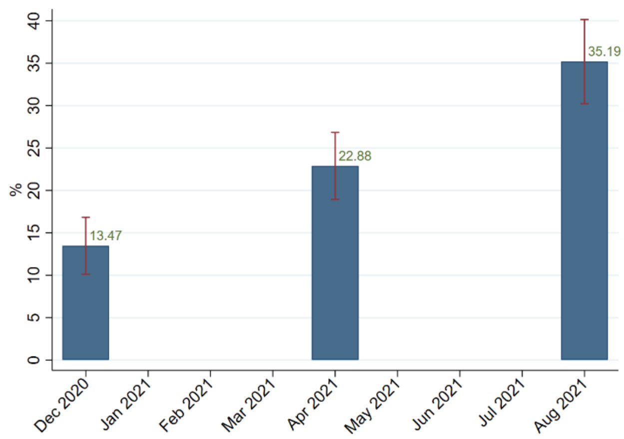


Figure A.4.14: Indonesia, strictly formal share

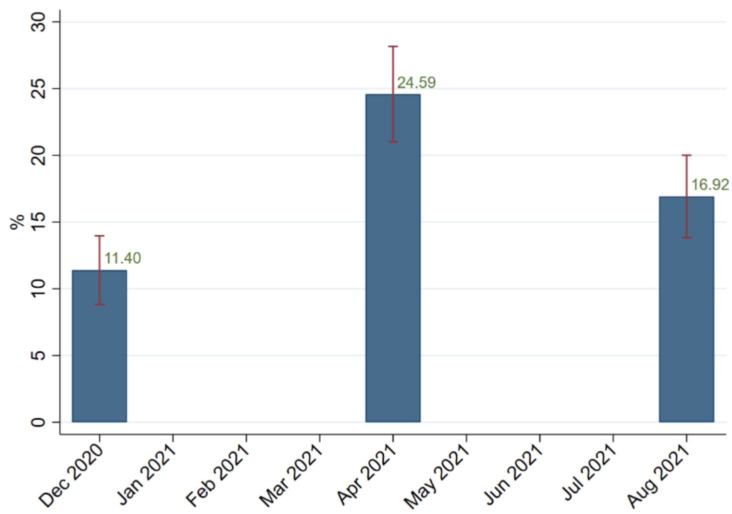


Figure A.4.15. Kenya, strictly formal share

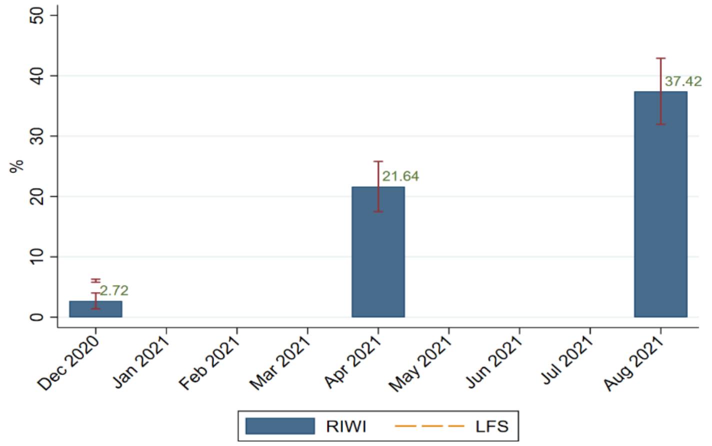


Figure A.4.16: Sri Lanka, strictly formal share

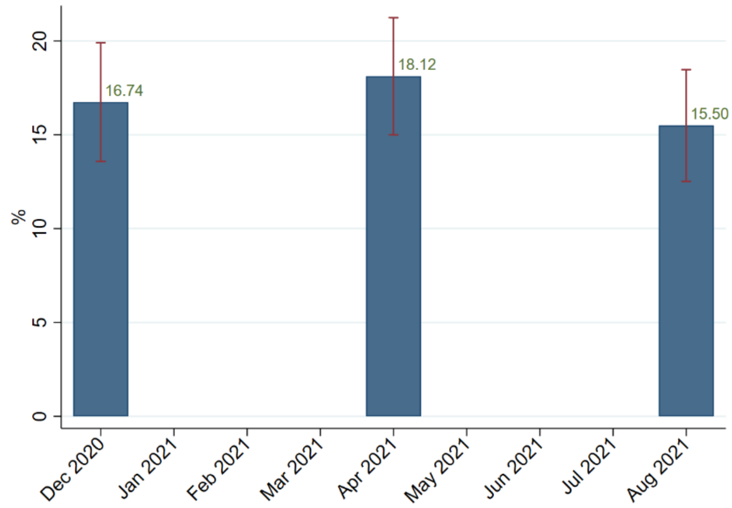


Figure A.4.17: Türkiye, strictly formal share

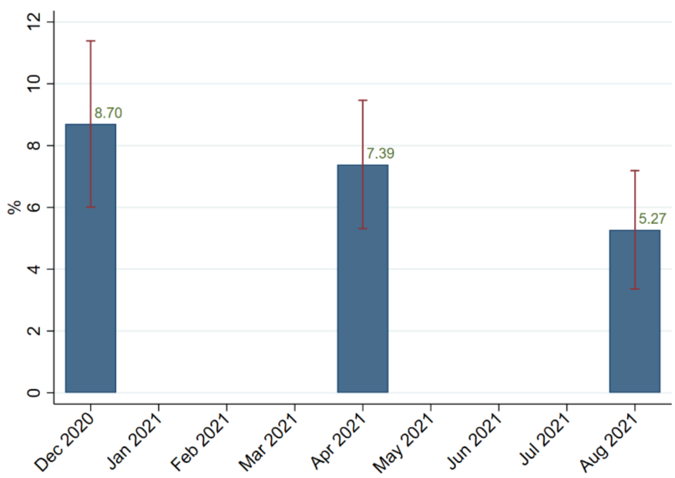


Figure A.4.18: Self-employment rates, Brazil

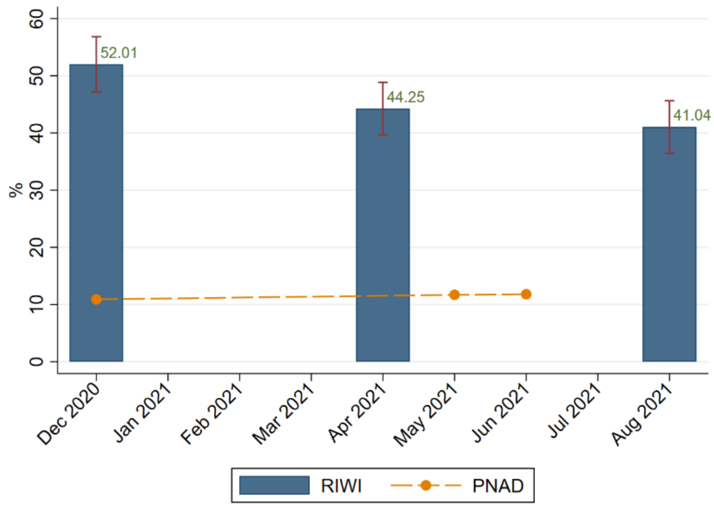


Figure A.4.19: Self-employment rates, Arab Republic of Egypt

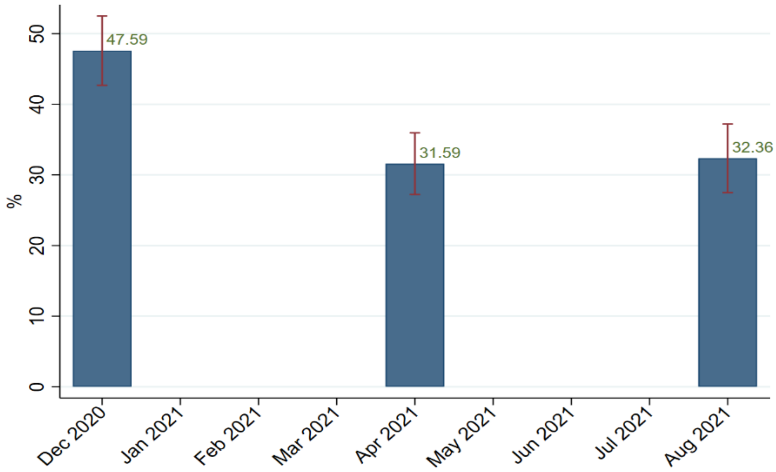


Figure A.4.20: Self-employment rates, Indonesia

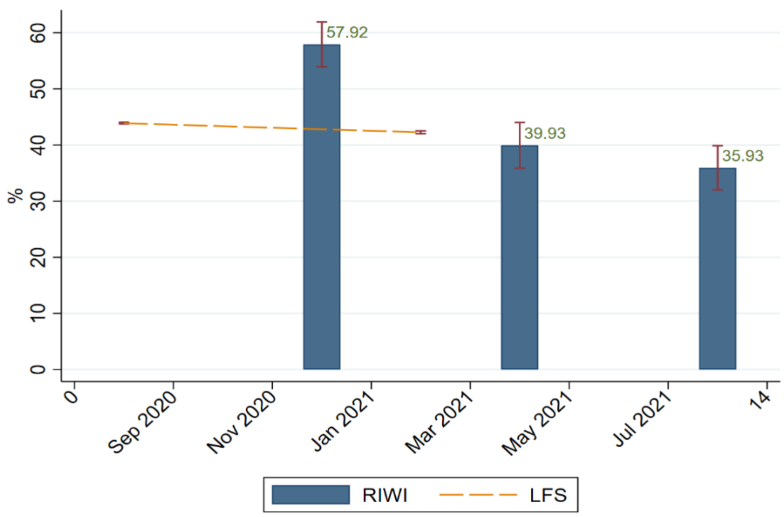


Figure A.4.21: Self-employment rates, Kenya

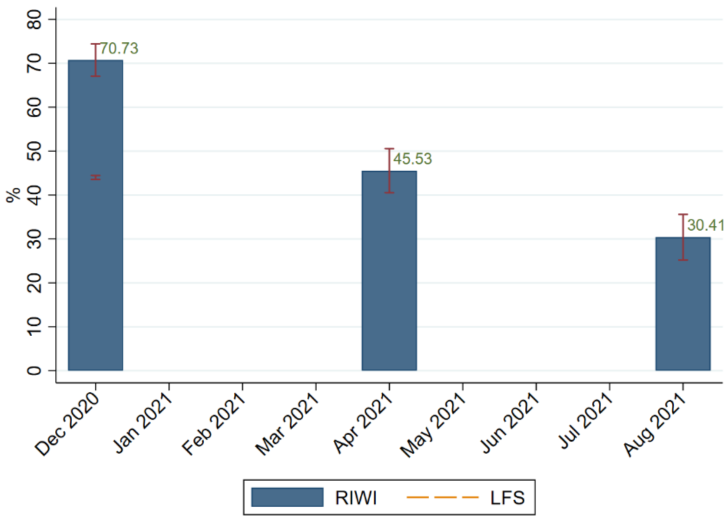


Figure A.4.22: Self-employment rates, Sri Lanka

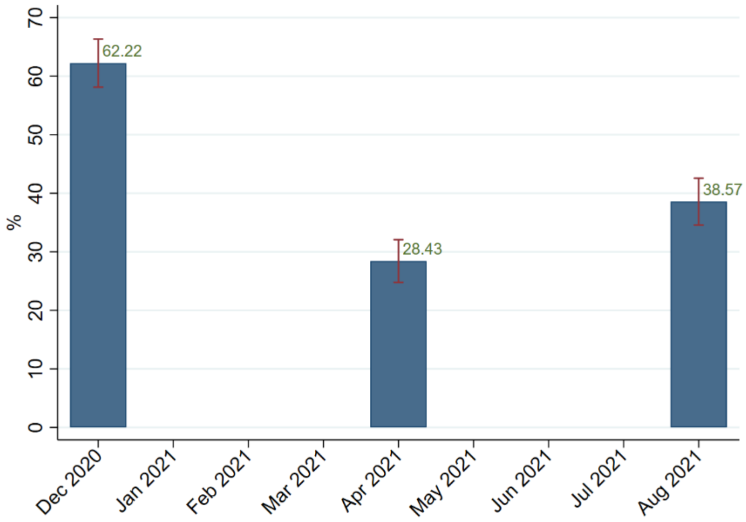


Figure A.4.23: Self-employment rates, Türkiye