

CHAPTER 2

How Many Gig Workers Are There?

Using Two Methods to Estimate the Online Gig Workforce

2.1 INTRODUCTION

There are no systematic ways to estimate how many people work in the gig economy, despite its place as a new, growing segment of the workforce that has implications for labor market and social protection policies, as well as for regulations governing data privacy, competition, and taxation. Gig workers are seldom measured in labor force and household surveys, in which they may be classified together with day laborers, independent contractors, or self-employed workers. (See chapter 6 for a discussion of labor force surveys.) Tax returns for gig and nongig workers may be similar, as is the case in the United States (Abraham et al. 2018); hence, they do not provide a reliable source of data. In addition, since both gig workers perform tasks from flexible locations and client firms may be located outside the worker's jurisdiction, traditional methods of national data collection and national tax records do not work. Platforms too rarely disclose much detailed data because they are commercially sensitive information.

This chapter uses two methods to estimate the number of gig workers globally.

- 1. The first method uses data science and builds on the mapping database** discussed in chapter 1. It involves collection of data on the number of registered users on each platform through web scraping and manual searches and, where no information on the number of registered workers is available for a platform, uses data on website traffic and unique visitors and other indicators such as the Alexa rank (collected in the database) to estimate the total number of *registered workers* globally. Then this number of registered workers (observed and predicted) and traffic data are used to estimate the share of *active workers* on each platform.⁵⁰
- 2. The second method employs an experimental methodology that uses the random domain intercept technology (RDIT) patented by RIWI⁵¹** to conduct a global survey in 17 low- and middle-income countries, from which it extrapolates the share of gig workers among the working population globally. The RDIT methodology assumes a random distribution of the survey to the internet population in the targeted countries that is accessible on a variety of devices (desktop, mobile, tablet). The survey was conducted in 12 languages in addition to English to reach non-English-speaking populations. Complete responses were collected from 7,015

⁵⁰ Traffic data have the benefit of being widely available, introducing consistency in interpretation across platforms and regions. This type of data also offers insight into how many people use a platform and how intensely they do so, through information on unique monthly users and average time spent on a website.

⁵¹ For information on RIWI, see <https://riwi.com/>.

respondents, including 956 responses from online gig workers; the rest were from respondents who had never done any gig work. The 17 countries include Argentina, Bangladesh, China, the Arab Republic of Egypt, India, Kenya, Lebanon, Mexico, Morocco, Nigeria, the Philippines, Pakistan, República Bolivariana de Venezuela, the Russian Federation, South Africa, Tunisia, and Ukraine) (see appendix D for survey methodology).

2.2 HOW HAVE OTHER STUDIES APPROACHED THIS QUESTION?

This study builds on previous work to estimate the number of gig workers (see Table 2.1 for a summary). Because of differences in scope and methodology, it is difficult to compare the estimates of the studies in the table. While early estimates (Codagnone, Abadie, and Biagi 2016; Kuek et al. 2015) were relatively modest (50 million registered workers), more-recent estimates (Kässi, Lehdonvirta, and Stephany 2021) are much higher (163 million), even though one of the former estimates included both online and location-based gig workers. These later estimates could be higher because of both rapid growth in the gig economy and improvements in methodology. Regarding the latter, larger data sets have been developed, as each study built on the work of the earlier ones, leading to an overall improvement in methods and estimates. Studies using survey-based methods arrive at various estimates because of differences in geographic coverage and the type of gig work estimated (online or location based versus active or registered workers).

TABLE 2.1: Estimates of market size

Reference	Sample	Methodology used	Estimates
Kässi, Lehdonvirta, and Stephany (2021)	Database mapping of 351 online web-based platforms	Machine learning model that includes as predictive features the Alexa rank, estimates for monthly users from siterankdata.com, and Google Trends information	163 million registered users, of which on average only 8.6 percent (14 million) are active
Kuek et al. (2015)	5 large platforms: Upwork, Freelancer, and Zhubajie/Witmart (online freelancing); Amazon Mechanical Turk and Crowdfunder (today FigureEight) (microwork)	Estimation of market size in terms of revenue, using gross revenue figures of top 3 freelance and top 2 microwork platforms, and predicting growth rate/trend with average growth rates of past years	Overall market size of US\$2 billion in 2013, of which US\$1.9 billion is freelance and US\$0.1 billion is microwork; 48 million registered workers, 10 percent of whom are active
Codagnone, Abadie, and Biagi (2016)	39 gig platforms in the US and the UK and other European countries (all types)	Desk research (web searches) on registered gig workers on these platforms plus assumptions	In the US and UK, 1 to 2 percent share of gig workers in the labor force; 52.6 million registered workers ^a globally
Heeks (2017)	Based on sample used by Codagnone, Abadie, and Biagi (2016)	Literature review, combining results from previous studies to calculate market size and workers in the Global South	US\$5 billion, involving around 70 million workers globally, of which 60 million are in the Global South

Source: Study team summary.

Note: UK = United Kingdom; US = United States.

a. On the platforms in the sample. See Codagnone, Abadie, and Biagi (2016).

Many studies have used a small sample consisting of the few large gig platforms for which information is available to estimate the size of the overall online gig economy (Kuek et al. 2015; Pesole and Rani, forthcoming). This approach was spearheaded by Kuek et al. (2015), who estimated the overall market revenue and number of workers on the assumption that the top three gig platforms covered 50 percent of the entire market. Researchers at the Oxford Internet Institute (OII) used the same approach and created the Online Labor Index (OLI), which initially tracked data from the five largest English-speaking platforms (Freelancer.com, Guru.com, Amazon Mechanical Turk [Mturk.com], Peopleperhour.com, and Upwork.com) and was recently expanded to include a

few Spanish- and Russian-language platforms in subsequent rounds,⁵² but overall representation of regional platforms in the OLI remains limited (Stephany et al. 2021).

Some studies use data on revenue and financial transactions to estimate the gig economy market size. Kuek et al. (2015) estimated the total market size to be US\$2 billion in revenue,⁵³ with 48 million workers globally in 2013, based on gross revenue figures and worker data from the five leading gig platforms.⁵⁴ The study then used the prior two years' average market growth rate to predict an overall market size of US\$4.8 billion by 2016. Similarly, Codagnone, Abadie, and Biagi (2016) collected data on registered contractors from a larger sample of 39 large gig platforms from simple web searches. From these numbers, they estimated that in the United States and the United Kingdom, the proportion of gig workers in the labor force was 1 to 2 percent, with a total of 52.6 million registered workers on the sample of platforms that were reviewed. Heeks (2017) expanded those results to include workers in developing nations. Considering survey ratios from other studies and a study of gig platforms in China, Heeks estimated that around 60 million people were involved in gig work in developing nations, of whom 10 percent (Kuek et al. 2015)—6.1 million—would be considered active and up to 3 million of whom would have online labor as their primary income. More recently, Kässi, Lehdonvirta, and Stephany (2021) employed a data-driven approach based on database mapping to estimate 89 million unique registered workers⁵⁵ and 14 million active workers.

Governments and private organizations have conducted surveys to estimate the size of the gig workforce, focusing mostly on developed countries. For example, the United States (Current Population Survey) included the contingent work supplement to the monthly labor force survey. Canada, Denmark, Finland, Sweden, and Switzerland (for an overview, see chapter 6 on social insurance in this report and ILO 2021), too, made efforts to measure gig work through labor force surveys. The EC conducted two COLLEEM surveys, with the later survey across 16 countries finding that 1.4 percent of the working-age population performed gig work as their main form of employment. Other surveys, such as a study of 11,000 workers in 11 countries that focused on low-skill and low-income respondents, found that the share of workers who receive their main income from gig work was much larger in emerging economies (3 to 12 percent in Brazil, China, India, and Indonesia) than in mature markets (1 to 4 percent in Germany, Spain, Sweden, the United Kingdom, and the United States) (BCG Henderson Institute 2019). Another study surveyed 6,000 adults in the United States in 2021 and found that about 36 percent of the US workforce (59 million) performed freelance work⁵⁶ in 2020–21 and that freelancers contribute up to US\$1.3 trillion to the US economy annually (Ozimek 2021). Pew Research Center surveyed 10,348 adults in 2021 to understand Americans' experiences and attitudes about earning money from online gig platforms and found that 16 percent of Americans have earned money from an online gig platform at some point (Anderson et al., 2021).

⁵² The three Spanish-speaking platforms are freelancer.es, twago.es, and workana.es. Three from the Russian-speaking domain are freelance.ru, freelancehunt.ru, and weblancer.ru. See Stephany et al. (2021).

⁵³ While revenue offers valuable insight into individual platforms' business performance, it is not reported very often, making it difficult to use this metric on a broad set of platforms, particularly those that have not been listed on public markets. In addition, gig platforms' business models and associated revenue models differ widely—for example, the working relationship between a platform and the gig workers, pricing and revenue structures, and vetting mechanisms vary across platforms. Those differences cannot be accounted for clearly when interpreting the level of activity on platforms from reported revenue streams.

⁵⁴ With the assumption that the market leaders at the time (Upwork, Freelancer, and Zhubajie/Witmart) held 50 percent of the online freelancing and that Amazon Mechanical Turk and Crowdfunder (today FigureEight) held 80 percent of the market for microwork. See Kuek et al. (2015).

⁵⁵ Calculated from 163 million estimated registered-user accounts divided by 1.83 to account for multihoming. See Kässi, Lehdonvirta, and Stephany (2021).

⁵⁶ In this study, freelancers are defined as "Individuals who have engaged in supplemental, temporary, project- or contract-based work, within the past 12 months (calculated within the US Workers Overall sample)." See Ozimek (2021).

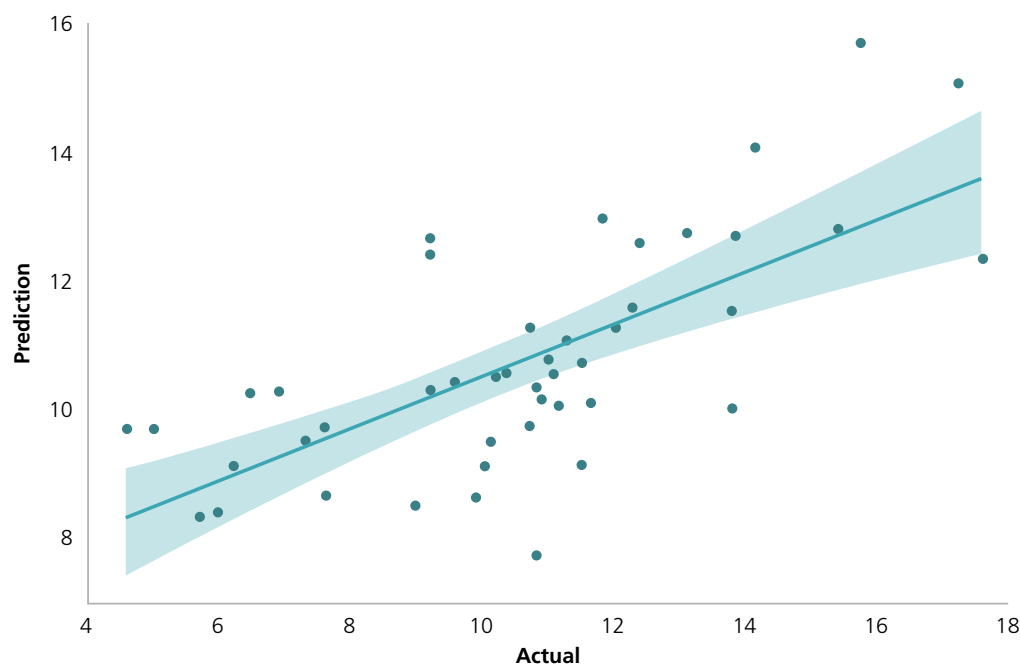
This study contributes to the literature by proposing two alternative methods to estimate the size of the gig labor force and by making additional effort to identify and measure regional/local and non-English-language platforms.

2.3 METHOD 1: WEB SCRAPING AND DATA SCIENCE

The first step was to collect web-scraped data for registered workers with a Python script or retrieved from the platforms' websites, press releases, or third-party reports. Information about the number of registered workers was available online for around 60 percent of the platforms.

The second step was to develop a predictive model for the remaining 40 percent of platforms for which information was not available, by using XGBoost, a tree-based machine learning model (Chen and Guestrin 2016). The model uses parameters such as website traffic (total traffic and number of unique visitors) and Alexa rank as independent predictive features or variables to predict the number of registered workers (dependent variable). These parameters related to website traffic highlight how many people visit a website, how much time they spend on it, and how many pages they visit on average. Traffic and visitors and unique visitors' values were logarithmically transformed, since the data are extremely skewed with few high outliers. This approach to reduce skewness is consistent with that of prior literature (such as Ang, Chia, and Saghafian 2021 and Lütkepohl and Xu 2010). An 80-20 train-test split was used on the 327 observed platforms, and various models including linear and polynomial regressions, random forest, extra trees, and XGBoost (Chen and Guestrin 2016) were experimented with in Python. The hyperparameters of the tree-based regressors were optimized by both grid search and Bayesian optimization. The XGBoost model was found to perform best on the test set, with the lowest mean square error and highest R^2 fit between the actual and predicted values. Figure 2.1 illustrates the plot of the actual versus predicted values for the test set. This fit appears to work well in other studies as well (Kässi, Lehdonvirta, and Stephany 2021, for example).

FIGURE 2.1: Model fit (XGBoost) for the prediction of registered workers on the test set



Source: Elaboration by the study team.

Note: The figure presents the plot for the model predicted values for number of registered workers (log scale) versus the actual data (log scale) for the test set. As observed, apart from outliers, the model performed reasonably well.

The next step was to adjust the estimates for multihoming and multiworking. Multihoming refers to freelancers or gig workers being registered, affiliated, or actively working on more than one online gig work platform. The team surveys conducted for this study (see chapter 4) found that workers are registered on an average of 1.834 platforms.⁵⁷ This means that registered-worker estimates need to be divided by 1.834 to account for multihoming to yield *unique* registered workers. This number is consistent with other studies.⁵⁸ At the same time, multiple workers may be working under a single freelancing account instead (multiworking),⁵⁹ as suggested by interviews with gig workers conducted by the team as well as by other studies in Africa (Melia 2020; Wood et al. 2019b). Reasons to engage in multiworking include lower barriers to entry, for example where subcontractors are not yet able to perform tasks using their own accounts (Melia 2020), and the trust and reputation of more-established accounts (Wood et al. 2019b). To date, there are no systematic studies or surveys of the multiworking phenomenon (Kässi, Lehdonvirta, and Stephany 2021).

This study is among the few that have estimated this phenomenon at a global level. Results suggest that an average of 1.19 workers is performing work from one account.⁶⁰ Therefore, an adjustment factor of 1.19 was added to the estimations of unique registered workers.

So how many online registered gig workers are there?

We estimate that there are 154 million unique registered gig workers worldwide. The total number of registered workers that were found through data collection and predictions using the XGBoost model was divided by 1.83 to account for multihoming, yielding 154 million unique registered gig workers on online gig platforms worldwide. While this is a reasonable estimate and not far from other estimates, the results may still be underestimating the number of gig workers. That is because traffic data were not available for all platforms. Also, some large platforms were excluded because it was not possible to trace traffic on relevant subfolders, further suggesting that these *estimates may be on the lower side*.

And how many of the registered gig workers are active?

Considering the sporadic nature of gig work, the number of registered gig workers may not accurately reflect the size of this group. Gig workers often vary widely in terms of how much time they spend doing gigs and what fraction of their overall income is generated by gig work. A worker may be doing gig work on a full-time or part-time basis, might perform tasks only sporadically (on weekends or some days in a month), or only under certain circumstances (such as loss of a job). The team's global survey in 17 countries found that one in three gig workers does online work as their main occupation, while for two-thirds it is a secondary occupation or is performed only sporadically

⁵⁷ This figure is the weighted average of the responses to the following question: "Which platforms do you work on? Please list all that apply" from the RIWI and Soyfreelancer surveys. Responses were weighted to account for different sample sizes.

⁵⁸ Surveys from the ILO (2021) and Wood et al. (2019a) estimate that on average, workers are active on 1.83 platforms.

⁵⁹ We adopt the term used in Kässi, Lehdonvirta, and Stephany (2021): multiworking. In other literature, this phenomenon has been called "subcontracting" or "re-outsourcing." See Melia (2020) and Wood et al. (2019b).

⁶⁰ Across five surveys with a total of 6,113 responses, workers were asked whether they (a) work on the tasks alone on their own account, (b) hire other people and assign tasks to other gig workers, or (c) sometimes work alone, sometimes hire other people. The responses were coded with 1 for "I work alone always" and 3 for the response "I hire other people and assign tasks to other gig workers" (this is the median of responses in the survey conducted in the Khyber Pakhtunkhwa region in Pakistan survey on how many people a person delegates tasks to); for "sometimes I work alone; sometimes I hire other people," weights are varied between 2 (50 percent alone, 50 percent other people), 2.5 (25 percent alone, 75 percent other people), and 1.5 (75 percent alone, 25 percent other people). The results indicate that between 1.13 and 1.24 would be the factor for multiworking, depending on the weights. Assuming that the 50-50 split for answer (c) is most likely, the resulting factor for multiworking is 1.19.

(see chapter 4). Furthermore, not all users who register end up pursuing gig work. They might have done gig work in the past or might have signed up out of curiosity. This implies that workers registered on gig platforms may not be actively working on them. Therefore, it is important to also estimate the number of *active* workers.

Platform websites do not list how many of their registered workers are active. This is partly because platforms compete with one another for users and funding and because they use various definitions of “active” workers. For example, some platforms may consider workers active if they are submitting bids or proposals (in other words, engaging with the platform), but others may consider workers active only if they are currently working on live projects and generating income or revenue. Existing estimations of active workers have relied largely on small samples and rules of thumb. For example, Kässä, Lehdonvirta, and Stephany (2021) predict that 8.6 percent of registered workers have worked at least once, Kuek et al. (2015) estimate that 10 percent of registered workers could be considered active (with a sample of $n = 5$), and Pesole and Rani (forthcoming) find that, in a sample of given platforms, about one-third of registered workers have completed at least one project successfully, while only 10 percent or fewer have completed 10 projects or earned more than US\$1,000 on the platforms.

In the absence of reliable data on activity levels, we use a proxy indicator for monthly unique website visitors. This study uses a definition of “active” that combines hours worked and percentage of overall income earned through online gig work monthly (see table 4.2 for details). But in the absence of sample-wide data on user behavior, this definition cannot be used for the present approach. Since the traffic data are at the firm level (not the individual level), we use activity on platforms with traffic data, specifically with the time spent on the website by users, as a proxy. The model estimates the share of active workers, defined as the share of registered workers that are likely to be actively using the platform.

The model uses the average number of unique website visitors per month multiplied by the bounce rate to remove one-off or accidental visits.⁶¹ This number is then multiplied by the estimated ratio of workers to clients, to account for workers only, and is subsequently divided by the number of registered users, accounting for multihoming and multiworking. A key input for the formula is the ratio between workers and clients on platforms, which enables an estimate of traffic data generated by workers. However, these data do not exist at the platform level and likely vary across platform types, sizes, and geographies. With the global demand stemming predominantly from high-income countries, there tends to be a larger proportion of clients relative to workers in high-income countries than in low- and middle-income countries. At the same time, there are differences between platform business models as well: smaller platforms and those focusing on high-skill tasks often employ an agency model that has higher barriers to signing up but also greater likelihood of winning a job offer. This suggests that a larger share of registered workers might be active, particularly compared to larger platforms that have low barriers to signing up. Surveys and interviews with 10 platforms conducted for this report⁶² show an average ratio between workers and clients

⁶¹ The bounce rate tells us the percentage of visitors to a website that leave said site without taking an action, such as clicking on a link, filling out a form, or making a purchase. See [https://backlinko.com/hub/seo/bounce-rate#:~:text=Bounce%20Rate%20is%20defined%20as,obviously\)%20didn't%20convert](https://backlinko.com/hub/seo/bounce-rate#:~:text=Bounce%20Rate%20is%20defined%20as,obviously)%20didn't%20convert).

⁶² Al7arefa, Asuqu Elite, BeMyEye, Jolancer, Onesha, SoyFreelancer, Upwork, Workana, Wowzi, Truelancer.

of 75.5 to 24.5.⁶³ While this ratio will not be true for all platforms, it reflects a diverse set of large and small and global and regional platforms.

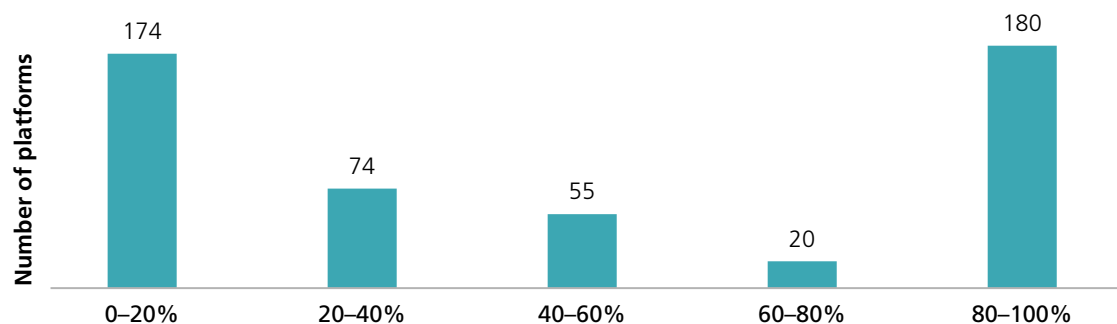
The model is as follows:

$$\text{Estimated share of active workers}^{64} \text{ for each platform } (\text{Percentage}_{\text{active}}) = \left[\frac{V_u * (1 - br) * r}{1.19 * \frac{(W_r)}{1.834}} \right]$$

where V_u is the average number of unique visitors per month; br is the average monthly platform bounce rate; W_r is the number of registered workers (either observed or predicted) for each individual platform; 1.19 is the adjustment factor for multiworking, based on internal surveys conducted by the World Bank; 1.834 is the adjustment factor for multihoming, based on internal surveys conducted by the World Bank; and r is the ratio of workers to client (=0.755).

We find that there are approximately 52 million active gig workers globally. The distribution for share of active workers was found to be generally right (positive) skewed but with a significant share of platforms having high percentages of active workers (see figure 2.2). This indicates that in most cases, only a small fraction (0 to 10 percent) of workers actively engage on the platform, but there is a sizeable percentage (35 percent) of platforms with a large share of active workers (over 81 percent). Large proportions are driven by high traffic figures in relation to the number of registered workers. In some cases, this might be due to a different business model in which platforms curate and keep a pool of vetted workers who are rotated and used across projects. In other cases, there might be overestimation of traffic or underestimation of registered-worker figures.

FIGURE 2.2: Histogram—Percentage of registered workers that are active



Source: Elaboration by the study team.

Note: The share (percentage) of active workers among registered users on a gig platform in the sample of platforms for which traffic data was available (n = 503) is shown. The total numbers of platforms are indicated above the bar graph.

The average proportion of active workers out of registered workers is 37 percent, with a median of 26 percent. This is higher than findings in prior studies that found active-worker shares of 8.6 percent (Kässi, Lehdonvirta, and Stephany 2021), 10 percent (Kuek et al. 2015), and 33 percent

⁶³ This figure was further tested by evaluating common search terms leading to the four top platforms. Using traffic data from the four top platforms, about 100 keywords in terms of traffic that landed on those websites were classified according to whether they likely indicate a buyer/client or a seller/worker. For this purpose, a keyword that includes a verb (for example, translate something) or the term “services” (such as copywriting services) was classified as indicating a buyer/client looking for such a service, while anything that included the term “jobs” (for example, freelance design jobs) was classified as indicating a seller/worker looking for job openings. There are several categories—for example website designer, translation, and others—that could belong to either sellers or buyers and are therefore not marked. We find that the ratio of workers to clients is roughly 70:30, which is close to the ratio we used in our model.

⁶⁴ Some platforms have unusually high numbers of unique visitors observed. Because the share of active workers cannot exceed 100 percent of registered workers, we also apply a 100 percent upper limit to the percentage.

(one-project threshold) or 10 percent (10-project threshold) (Pesole and Rani, forthcoming). On average, global platforms have a slightly higher percentage of active workers than regional platforms (37 percent versus 36 percent).

This estimation model has several limitations. The model relies heavily on traffic data for the estimations. However, other factors besides traffic, which are impossible to capture in this model, likely influence the proportion of active workers significantly. These include the split of demand and supply among website traffic, which was incorporated into the model on the basis of data from a sample of six platforms. Furthermore, the extent to which work requires spending time on the platform and the type of work (especially microwork versus tasks that require more time to complete) are difficult to estimate. Because these data points are not possible to obtain without unique insights into proprietary data owned by the platforms, collaboration with platform providers would be necessary to expand this model in the future.

2.4 METHOD 2: ESTIMATION USING AN RDIT GLOBAL SURVEY

Given the challenges in developing reliable estimates of gig workers, the team also used another experimental approach: an online global survey collected randomly from the internet using populations in selected countries. The survey uses the RDIT, patented by RIWI,⁶⁵ rolled out in 17 low- and middle-income countries to extrapolate the share of gig workers among the working population. The RDIT methodology assumes a random distribution of the survey to the internet population in the targeted countries, accessible on a variety of devices (desktop, mobile, tablet). The survey was conducted in 12 languages in addition to English to reach non-English-speaking populations. Complete surveys were collected from 7,015 respondents, of which 956 responses were from online gig workers and the rest were from respondents who had never done any gig work. The 17 countries were Argentina, Bangladesh, China, Arab Republic of Egypt, India, Kenya, Lebanon, Mexico, Morocco, Nigeria, Pakistan, the Philippines, República Bolivariana de Venezuela, the Russia Federation, South Africa, Tunisia, and Ukraine (see appendix D for survey methodology).

Calculating the global number of online gig workers

We followed a series of steps to calculate the global number of online gig workers *excluding* high-income countries.

- The first step was to **select the 17 countries** while taking into account their market share in the global online gig work industry, geographic representation, and language usage.
- After **piloting the survey** in three countries, we launched it between June and August 2022.
- The collected data were cleaned, and **quality checks** were carried out to remove unreliable responses based on time taken to finish the survey.
- Then **a raking algorithm** based on age, gender, and education was used to assign weights for each response. The weights were constructed in such a way that their sum adds up to the internet-using population of each country during 2021, which is the latest year for which we could

⁶⁵ RIWI implements online surveys using random domain intercept technology. RIWI allows internet users to opt in to anonymous surveys on any web-enabled device. While using the web or apps, internet users may randomly come across an RIWI survey via dormant domains (websites that are no longer in use), incorrect URLs, and links within apps and websites. Instead of encountering a “page does not exist” notification or an advertisement, a RIWI survey or message test is rendered full site on the page. Web users then decide whether they would like to anonymously participate in the research and do so without incentivization. See <https://riwi.com>.

get internet penetration data for the sampled countries from World Bank’s World Development Indicators (WDI).

- Next, we **calculated the proportion of online gig workers at the country level** by applying the weights constructed from the raking procedure.
- After that, we **multiplied the result by the internet-using population** of the country to arrive at the total number of online gig workers in the sampled countries. To arrive at a regional-level estimate, we used Semrush data to calculate each sampled country’s share of internet traffic to online gig platforms. For instance, Kenya, Nigeria, and South Africa account for 80.6 percent of the internet traffic flow to online gig platforms from Sub-Saharan African countries. Using our global survey-based estimation, we determined that the number of online gig workers in these three countries is 17.5 million (the share of online gig workers from the survey multiplied by their internet-using population).
- We then **used this information to estimate the number of online gig workers for the remaining countries**, which account for 19.35 percent of the traffic flow, giving us roughly 21.7 million gig workers in Sub-Saharan Africa.
- We **replicated the process for the rest of the regions and added the results to arrive at the global number of online gig workers**.⁶⁶ This calculation provides us with a more reasonable estimate of the online gig worker population in each region.

To obtain the global number of online gig workers, **we totaled the regional estimates, excluding North America from the calculation** because no country from the region was sampled. We then **incorporated estimates from previous studies on online gig workers based in North America to arrive at the final global estimate**. (See appendix C for details.)

The primary question used to identify online gig workers reads as follows.

“Does this describe ANY work you did in the last 12 Months? Yes/NO”

“Some people find short, ONLINE tasks or jobs through a website or an app. These tasks (also called gigs) are done entirely online and digital platforms coordinate payment for the work done”

Defining “active” gig workers

To assess activity levels, we divided gig workers into three groups—main, secondary, and marginal gig workers—based on the study by Urzì Brancati, Pesole, and Fernández-Macías (2020) in the EU. This classification uses the number of hours worked on online gigs and the percentage of personal income earned from the online gig economy to determine whether a gig worker is main, secondary, or marginal (table 2.2).

⁶⁶ Because China was underrepresented in the Semrush data, we used the traffic share for the Philippines to estimate the figure for the East Asia and Pacific region, excluding China. We then added the number of online gig workers in China estimated from our global survey.

TABLE 2.2: Classification of gig workers based on earnings and working hours

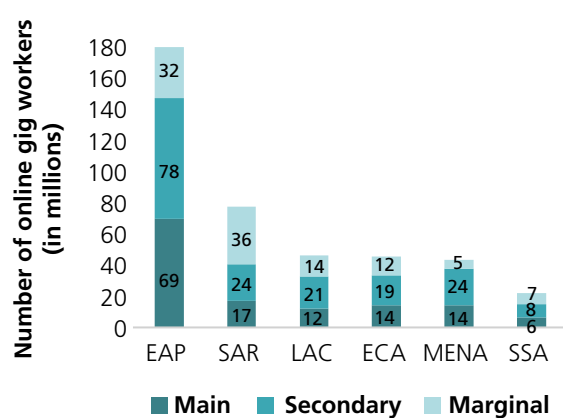
	Less than 10 hours a week	Between 10 and 20 hours a week	More than 20 hours a week
Less than 25 percent of personal income	Marginal	Secondary	Secondary
25 to 50 percent of personal income	Secondary	Secondary	Main
More than 50 percent of personal income	Secondary	Main	Main

Source: Adapted from Urzi Brancati, Pesole, and Fernández-Macias 2020.

So how many online gig workers are there? How many are “active?”

We estimate that there are 132.5 million main, 173.7 million secondary, and 106.2 million marginal gig workers. The total number of online gig workers, excluding North America, is 412.5 million. Adding in estimates of the online gig worker populations from other studies suggests that the number of online gig workers globally could be around 435 million.⁶⁷ In other words, we estimate that the share of online gig workers in the global labor force ranges from **4.4 to 12.5 percent**.⁶⁸

The East Asia and Pacific region accounts for 51 percent of online gig workers, followed by the South Asia region and the Sub-Saharan region (see Figure 2.3). Secondary and marginal online gig workers account for 42 and 26 percent of the online gig workers, respectively.

FIGURE 2.3: Estimated number of online gig workers based on the global online gig work survey

Source: Elaboration by the study team.

Note: Non-high-income countries in these regions are not included. EAP = East Asia and Pacific; ECA= Europe and Central Asia; LAC = Latin America and the Caribbean; MENA = Middle East and North Africa; SAR = Southeast Asia region; SSA = Sub-Saharan Africa.

These estimates are substantially higher than previous estimates. The difference could be due to the following reasons. First, the team’s global gig work survey was conducted in multiple languages, including Bangla, Mandarin, Arabic, Hindi, Swahili, Spanish, Hausa, Tagalog, Urdu, Russian, and English, to try to reach people from non-English-speaking populations. This approach led to a higher response rate in non-English-speaking countries, picking up respondents who would

⁶⁷ For example, Codagnone, Abadie, and Biagi (2016) estimated 52 million gig workers in the United States and the United Kingdom and other European countries. Among these, 44 million are registered users on online gig platforms. Assuming that the United States accounts for 50 percent of these gives an estimate of 22 million online gig workers. Adding those to the 412.5 million online gives an estimate of 435 million online gig workers globally.

⁶⁸ WDI data show that the global number of laborers was 3.46 billion in 2021.

have been missed in English-only surveys. Second, our survey was more recent and captured the current trend toward increasing gig work due to COVID-19. Third, the survey made a special effort to reach online gig workers on regional/local platforms who often get overlooked in studies that use platform data or survey data for only the large global platforms. As such, our study could reflect more comprehensive coverage of the online gig work market and may have identified gig workers who are often missed.

However, these estimates also have their limitations. Despite the assumption that the RDIT leads to a random selection of respondents, a recent study by Soundararajan et al. (2022) found that this may not always be the case. They discovered that the method overrepresents male, younger, and more educated members of the country's population. However, it should be noted that their study relied on an online survey to draw conclusions about the broader labor force, including offline workers. In contrast, our study focused solely on internet users, using an online survey to collect data and making it a better fit for our purpose. Also, not everyone who starts filling in the questionnaire completes it, as there is no incentive to do so. We found that individuals who identified as online gig workers and high school graduates were most likely to drop out before finishing the survey. Furthermore, although the raking procedure relies on good-quality nationally representative survey data on internet usage for seven countries,⁶⁹ for the remainder of the countries we had to rely on regional averages, an approach which may affect the quality of our results. The raking procedure by itself may not eliminate all biases, either.⁷⁰ Last, in the absence of accurate data for all countries, the estimation is built on the assumption that the relationship between traffic flow to online gig platforms and the number of online gig workers is proportional across all countries.

We conducted a robustness check for a few countries to understand whether and how these estimates could be biased. The Vietnam 2021 Labor Force Survey (LFS) asks if respondents use the internet to carry out their work on a regular basis, which is very helpful in estimating the number of online gig workers. We limited the analysis to self-employed individuals who use the internet and work in occupations and industries that are very similar to online gig work.⁷¹ This gives us the share of online gig workers in occupation-industry cells. We applied these figures to the Philippines, which is another East Asia and Pacific country, assuming a share of online gig workers in occupation-industry cells similar to that in Vietnam. For the main gig workers, our estimates are 6 percent lower than the LFS-based estimation for the Philippines and close to 20 percent higher than that for Vietnam. If we focus on the main, secondary, and tertiary gig workers, our estimates are more than four times higher than what the LFS-based results suggest.

2.5 CONCLUSION

This chapter describes the use of two models to estimate the size of the gig workforce. While not directly comparable, the two estimations show a possible range of the size of the gig economy. Our first approach used data science methods and estimated that there are a **total 154 million unique registered** and **60 million active gig workers**. Our second estimation model used a global survey and estimated that there are **435 million gig workers**.

The two methods complement each other and should be read in tandem. The first method (using web traffic data) traces the number of workers from a relatively comprehensive list of platforms,

⁶⁹ For details of the data sources, see appendix C.

⁷⁰ Soundararajan et al. (2022) used propensity score reweighting to address bias, but the resulting sample was not representative and yielded estimates that were at odds with nationally representative surveys.

⁷¹ See the mapping in appendix I.

thereby allowing a reasonable inference of the total market size. While this is a good base, the first method is an underestimate, since the total figure is missing data for the Chinese market.⁷² Traffic from mainland China is likely not captured fully in the present data, because of difficulties in accessing information on traffic on Chinese websites. For example, traffic predictions in our sample are higher for Hong Kong SAR, China, than they are for mainland China, which is unlikely to be true. This would imply that the total figures of registered and active workers on Chinese websites are underestimated. Also, the ratio for the split between workers and clients used to estimate active workers is based on assumptions and a very small sample of data, which is hard to confirm. Therefore, our first method gives us a *lower bound*.

The second approach is based on a global survey of workers and relies on information on the share of online gig workers in the sampled countries, which had larger proportions of gig workers than other countries within their region. These estimates are used to calculate the number of online gig workers in the regions they are drawn from, which could introduce an *upward bias*.

Although the two approaches used to calculate the figures yield different results, both methods confirm that online gig workers constitute a non-negligible portion of the overall labor force. According to the data science–based approach, the number of unique registered online gig workers is 154 million globally, which can be considered a lower bound for the reasons previously discussed. Meanwhile, the survey-based approach suggests that there are 132.5 million main gig workers, but when we include those who engage in gig work as secondary or marginal workers, the estimate could be as high as 435 million, providing an upper bound estimate.

In other words, we estimate that there are between 154 million and 435 million gig workers globally, which means that the share of online gig workers in the global labor force ranges between 4.4 and 12.5 percent.

⁷² The team's survey-based estimate after excluding China was 283 million, which is closer to the data science estimate, especially for main gig workers (74 million) versus the data science range (58 million to 91 million), making the two estimates comparable. However, another reason for the higher survey-based estimate is that it was conducted in several languages and was hence more successful in identifying gig workers who do not speak English and gig workers who work on regional/local platforms who may have been missed by the data science method.

References

- Abraham, Katharine G., John C. Haltiwanger, Kristin Sandusky, and James R. Spletzer. 2018. "Measuring the Gig Economy: Current Knowledge and Open Issues." NBER Working Paper 24950, National Bureau of Economic Research, Cambridge, MA.
- Anderson, M., McClain, C., Faverio, M., & Gelles-Watnick, R. 2021. "The state of gig work in 2021." https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2021/12/PI_2021.12.08_Gig-Work_FINAL.pdf.
- Ang, Yu Qian, Andrew Chia, and Soroush Saghafian. 2021. "Using Machine Learning to Demystify Startups' Funding, Post-Money Valuation, and Success." In *Innovative Technology at the Interface of Finance and Operations*, edited by V. Babich, J. R. Birge, and G. Hilary. Springer Series in Supply Chain Management, vol 11. Edinburgh, Scotland: Springer, Cham. https://doi.org/10.1007/978-3-030-75729-8_10.
- BCG Henderson Institute. 2019. "The New Freelancers: Tapping Talent in the Gig Economy."
- Chen, Tianqi, and Carlos Guestrin. 2016. "XGBoost: A Scalable Tree Boosting System." In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–94. New York: ACM. <https://doi.org/10.1145/2939672.2939785>.
- Codagnone, Cristiano, Fabienne Abadie, and Federico Biagi. 2016. *The Future of Work in the 'Sharing Economy': Market Efficiency and Equitable Opportunities or Unfair Precarisation?* Joint Research Centre, Institute for Prospective Technological Studies, Publications Office of the European Union, Luxembourg. <https://data.europa.eu/doi/10.2791/431485>.
- Heeks, Richard. 2017. "Decent Work and the Digital Gig Economy: A Developing Country Perspective on Employment Impacts and Standards in Online Outsourcing, Crowdwork, Etc." Development Informatics Working Paper 71, University of Manchester, UK. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3431033.
- ILO (International Labour Organization). 2021. *World Employment and Social Outlook: The Role of Digital Labour Platforms in Transforming the World of Work*. Geneva: ILO.
- Kässi, Otto, Vili Lehdonvirta, and Fabian Stephany. 2021. "How Many Online Workers Are There in the World?" Open Research Europe, 1–53.
- Kuek, Siou Chew, Cecilia Paradi-Guilford, Toks Fayomi, Saori Imaizumi, Panos Ipeirotis, Patricia Pina, and Manpreet Singh. 2015. "The Global Opportunity Online Outsourcing." World Bank, Washington, DC. <http://hdl.handle.net/10986/22284>.
- Lütkepohl, Helmut, and Fang Xu. 2010. "The Role of the Log Transformation in Forecasting Economic Variables." *Empirical Economics* 42: 619–38. <https://doi.org/10.1007/s00181-010-0440-1>.
- Melia, Elvis. 2020. "African Jobs in the Digital Era: Export Options with a Focus on Online Labour." Discussion Paper 3/2020, German Development Institute, Bonn.
- Ozimek, Adam. 2021. "Freelance Forward Economist Report 2021." Upwork. <https://www.upwork.com/research/freelance-forward-2021>.
- Pesole, A., and U. Rani. Forthcoming. "How many online gig workers?: Estimates based on selected online web-based and location-based platforms." European Commission Working Paper, European Commission, Brussels.

-
- Soundararajan, V., S. Soubeiga, D. Newhouse, A. Palacios-Lopez, U. J. Pape, and M. Weber. 2022. "How Well Do Internet-Based Surveys Track Labor Market Indicators in Middle-Income Countries?" Policy Research Working Paper 10359, World Bank, Washington, DC.
- Stephany, Fabian, Otto Kässi, Uma Rani, and Vili Lehdonvirta. 2021. "Online Labour Index 2020: New Ways to Measure the World's Remote Freelancing Market." *Big Data & Society* 8 (2): 20539517211043240.
- Urzi Brancati, M.C., A. Pesole, and E. Fernández-Macías. 2020. "New Evidence on Platform Workers in Europe: Results from the Second COLLEEM Survey." Joint Research Centre Paper JRC 118570, Publications Office of the European Union, Luxembourg.
- Wood, Alex J., Mark Graham, Vili Lehdonvirta, and Isis Hjorth. 2019a. "Good Gig, Bad Gig: Autonomy and Algorithmic Control in the Global Gig Economy." *Work Employment and Society* 33 (1): 56–75.
- Wood Alex J., Mark Graham, Vili Lehdonvirta, and Isis Hjorth .2019b. "Networked but Commodified: The (Dis)Embeddedness of Digital Labour in the Gig Economy." *Sociology* 53 (5): 931–50.