The gig economy in Poland – evidence based on mobile big data

Beręsewicz Maciej [*], Nikulin Dagmara[†], Szymkowiak Marcin[‡], Wilak Kamil[§¶]

Abstract

Gig economy creates a growing extent of job relations on the labour market. In this paper we focused on measuring the size and characteristics of the platform economy in Poland. We propose a different approach to measure gig economy based on mobile smartphones data that is collected through programmatic algorithms in online marketing. Our focus is on two types of task: food delivery (using apps like Bolt Courier, Takeaway, Glover, Wolt and transport services (Bolt Driver, Free Now, iTaxi and Uber). Using such data it was possible to provide upper bound for the number of drivers and couriers at very low aggregation levels. Our results show an increasing extent of platform economy in Poland. Focusing on the delivery and transportation activities performed with an use of applications, we report growing trend between January 2018 and December 2020. In particular, the most sharp increase was reported for Takeaway and Glover among the delivery apps, while for the transport apps, the increases were at a similar level. Taking into account demographic structure off apps user, we confirm the evidence existing in the past studies: young men are the dominated group of platform workers. Putting together the number of platform workers and the working populations our estimates show that the share of active app users accounts for about 0.5-2% of working population in the largest 9 Polish cities.

[*]Corresponding author: maciej.beresewicz@ue.poznan.pl. Poznań University of Economics and Business, Poland; Statistical Office in Poznań, Poland. **Acknowledgements**: The study was conducted within the research project *Economics in the face of the New Economy* financed within the Regional Initiative for Excellence programme of the Minister of Science and Higher Education of Poland, years 2019-2022, grant no. 004/RID/2018/19, financing 3,000,000 PLN (for Maciej Beręsewicz, Marcin Szymkowiak and Kamil Wilak).

[†]Gdańsk University of Technology, Poland.

[‡]Poznań University of Economics and Business, Poland; Statistical Office in Poznań, Poland.

[§]Poznań University of Economics and Business, Poland; Statistical Office in Poznań, Poland.

[¶]The views expressed in this paper are those of the authors and not the Statistics Poland nor Statistical Office in Poznań.

The gig economy in Poland – evidence based on mobile big data

1 Introduction

The changing world of work is indicating a growing tendency towards labour relationships based on non-standard employment contract. A typical employee-employer indefinite contract meaning full-time job on the regularly basis and performed in the employer's premises in not longer a predominant view of contemporary labour market. The driving forces related to digitalisation and flexibilization contribute to the emergence of new forms of employment. One of the recently observed expanding phenomena is the platform-mediated labour. According to the newest report issued by the International Labour Organization (ILO), the number of online web-based and location-based (taxi and delivery) platforms increased from 142 in 2010 to over 777 in 2020 (ILO, 2021).

Platform work may be defined as "non-standard work facilitated by online platforms which use digital technologies to 'intermediate' between individual suppliers (platform workers) and buyers of labour" (Hauben et al., 2020, p. 98).

Originally, *the gig economy* was identified with passenger transport services, mainly with the company *Uber*, that started to arrange work via the online platforms. Nowadays, the range of platform-based activities is rapidly growing and covers several types of job like: odd jobs (e.g. *TaskRabbit*), cleaning (e.g. *Helpling*), care (e.g. *care.com*), food delivery (e.g. *Deliveroo*) or programming and translating (e.g. *Upwork*) (Koutsimpogiorgos et al., 2020, cf.). That is why the new phenomenon of the platform work is getting more and more attention, both across academics ass well as international institutions (see among others the recent reports published by Eurofound (2019) or ILO (2021)).

Research related to these populations is becoming important for several reasons. Primarily, the "platformisation" of labour relations may increase the competitiveness and create new opportunities on the labour market. Moreover, the de-standardisation of employment may results in the worse working conditions in terms of job security and protections as platform workers are predominantly self-employed who are not covered by institutional frames like collective pay agreement scheme. On the other hand, platform work may create opportunities for some groups in the labour market, by increasing labour market participation (Eurofound, 2019) and enabling greater flexibility (Lehdonvirta, 2018).

However, the existing studies on the platform economy is rapidly expanding, the provided evidence is mostly anecdotal and based on interviews and personal first-hand experiences (De Stefano and Aloisi, 2018). This is due to the fact that workers in the gig economy, despite its size, constitute a hard-to-reach and identify population (see Bohning et al., 2017, Chapter 1). There is no sampling frame, nor a register that fully covers it and it may happen that members of the population are indistinguishable from the rest of the population (e.g. they may be only identified by the content of their smartphone or laptop). Information is fragmented, and thus there is a need for modern data sources such as big data, that may be helpful to assess the size and its characteristics. As these populations are often restricted to urban areas, general population sample surveys do not provide accurate estimate of reference population (e.g. economically active population) and administrative data often suffers from over-coverage (e.g. outdated information about the place of residence, delays in reporting).

In this study, we propose an alternative way to measure the size and characteristics of working age (18–65) population involved in providing their delivery or transportation services, using mobile big data obtained via advertisements systems on smartphones. For this study, we purchased historical data (period 2018-2020) about users of the following apps: *Uber*, *Bolt Driver*, *Glover*, *Wolt Courier*, *TakeAway Courier* and *Bolt Courier*. For comparison, we used *iTaxi* and *FREE NOW* apps that are prepared for the licensed taxi drivers. Background characteristics, being a result of classification algorithms applied by the data provider, include information on the location, country of origin, age and gender as well as other characteristics such as being a student or having children. In addition, we report descriptive statistics about the average length of app activity within working (Mondays-Thursdays) and weekends (Friday-Sunday) for day (8-18) and night (18-8).

Our contribution to the literature is twofold. First, we use big data collected via smartphones to passively measure the gig economy in Poland. This is in line with the recent trends in official statistics – to use all available information rather than create new ones (e.g. surveys). We underline problems regarding these sources (e.g. measurement error) and strengths (e.g. exact location) that may open a new avenue to measure these populations. This approach allows not only to cover hard-to-reach and identify populations as foreigners, but also plan targeted surveys to obtained more detailed information concerning motivations, working conditions or quality of life. Second, we provide a detailed information regarding the size and background of workers which was not previously available in Poland at any level of spatial aggregation. We believe that that similar data sources may be identified in other countries and thus our experiences and results may be used by other researchers and policy makers.

The paper has the following structure. Section 2 contains literature review and state the main issues in measuring gig economy by classical (e.g. sample surveys) and modern data sources (e.g. administrative data, the Internet, big data). Sections 3 and 4 present data sources about gig economy in Poland, selection of mobile apps and provide assessment of coverage the main data source used in the study. Section 5 presents main results and compares it with reference population from the Labour Force Survey (LFS). Finally, Section 6 provides discussion and future steps in analysis of the gig economy by mobile big data.

2 Issues in measuring gig economy

The "platformisation" of labour relations is described by several terms. Among the most common labels are the "platform economy" and "gig economy", but also "on-demand economy", "collaborative economy" which are used interchangeably. Importantly, the gig economy may be defined from the broader and narrower perspective. The broad framework cover precarious and casual work with the use of technological intermediation (Aleksynska et al., 2019), while in narrower perspective some of the researchers focus on the specific types of platforms, or identifies gig economy with the digital labour markets, without definition details (for an overview see Koutsimpogiorgos et al. (2020)). Based on the definitions proposed by Eurofound, the platform work means "a form of employment that uses an online platform to enable organisations or individuals to access other organisations or individuals to solve problems or to provide services in exchange for payment" (Eurofound, 2018, p. 9). Moreover, they mention the main characteristics of the platform work like paid work organised through an online platform; involvement of three parties: the online platform, the client and the worker; job based on providing specific tasks or solve specific problems; the work is outsourced or contracted out; jobs are broken down into tasks; services are provided on demand (Eurofound, 2018). In short, following the review made by Koutsimpogiorgos et al. (2020), there is a lack of common definition on the gig economy both across academics, policymakers, and practitioners.

The literature review shows an increasing number of studies describing the relatively new phenomenon of the gig economy. The main message which may be derived from, is that the hitherto term describing typical job as those based on indefinite contract, performing on the regular basis and fixed hours, is becoming to outdated and replaced by new forms of employment. Besides the problems with understanding and defining the platform economy, a related concern is about their measurement. In this section, we provide the past evidence of the studies measuring the extent of platform economy. Please note, that we focus solely on the European labour market and do not provide past evidence for remaining countries which is to find among others in Friedman (2014), Mitea (2018), Riggs et al. (2019) for New Zealand, Koustas (2019), Abraham et al. (2018) for US.

Looking at the previous evidence compiled in Table 1 it turns out, that the majority of past estimates is based on the survey studies. It is important to underline the numerous limitations related to estimates derived from surveys postulated in the literature. Predominant doubt is on the appropriate understanding of the definition of platform work by respondents. As Bonin and Rinne (2017) argue the respondents may misclassified other online activities like using websites for job search or search engines and professional networks as a forms of platform work. Moreover, in case of the internet surveys, many inconsistencies may occur, caused by the poor reliability of the pooling technique which may result in over-representation of people who are more likely to being engage in platform work (Piasna and Drahokoupil, 2019). Another part of estimates is general population based, see among others Aleksynska et al. (2019) present the survey results for Ukraine; Piasna and Drahokoupil (2019) for Bulgaria, Hungary, Latvia, Poland and Slovakia; Berg (2016) for US and India. Remaining evidence is case study specific and describes selected types of platforms with the use of qualitative methods, see among others Meszmann (2018) for Hungary, Lenaerts (2018) for Belgium, Sedláková (2018) for Slovakia, Lee et al.

(2018) for Hong Kong. However, these anecdotal evidence bring important knowledge on the characteristics on platform workers rather then evidence on the extent of exiting phenomena. Importantly, when it comes to US the use of transaction (financial) data is observed. Koustas (2019) derived data from transaction-level dataset from a large financial aggregator and bill-paying application. Similarly, Farrell et al. (2019) show the results based on the de-identified Chase checking accounts. However, to the best of our knowledge, there is a lack of Europe related studies using bank big data.

Table 1: Past estimates of platform economy in Europe

Author / Institution	Data source	Coverage	Scope of definition of platform work	Results
ETUI Internet and Platform Work Survey Piasna and Drahokoupil (2019)	Face to face survey	Bulgaria, Hungary, Latvia, Poland and Slovakia in 2018–2019; working age adults (aged 18–64); in total, 4,731 respondents	work using online platforms at least weakly. Online platforms are internet websites or apps through which workers can find short jobs or tasks, such as IT work, data entry, delivery, driving, personal services, etc.	0,4% in Poland and Slovakia; 0,5% in Latvia; 0,8% in Bulgaria; 1,9% in Hungary
University of Hertfordshire, Foundation for European Progressive Studies, UNI Europa Huws et al. (2019)	large-scale online surveys	11 European countries in 2016–2019	platform work conducted at least once a week	between 4.7 per cent in the UK, 6.2 per cent in Germany and 9.5 per cent in Austria (2016) 28.5 percent of citizens in Czech Republic (2019)
Collaborative Economy (COLLEEM) survey implemented by the Joint Research Centre of the European Commission Brancati et al. (2020)	large-scale online surveys	2017 and 2018 in 14 and 16 EU member states; 38,878 responses (internet users aged 16–74 years)	workers who have ever gained income from providing services via online platforms, where the match between provider and client is made digitally, payment is conducted digitally via the platform, and work is performed either (location-independent) web-based or on-location	4.1 per cent of the adult population in Finland and 9.9 per cent in the United Kingdom
ILO 2015 Berg (2016)	"1,167 responses of which 814 were from AMT and 353 were from Crowdflower"	2015, US and India	Amazon Mechanical Turk and Crowdflower platforms	Only characteristics of crowd workers
Drahokoupil and Piasna (2019)	administrative data provided by Smart and a survey of workers	Belgium 2016–2018	food delivery platforms, Deliveroo	3,828 member-riders in October 2017
Aleksynska et al. (2019)	"Internet survey; 1000 respondents of the age 18 and older"	Ukraine 2017	Workers of online platforms are	Only characteristics of online workers
Lee et al. (2018)	"self-reported online survey was conducted among Uber, 295 respondents"	Hong Kong 2016	Uber users	analysis of users' intention to participate in Uber

3 Data sources about the gig economy in Poland

3.1 Probability and non-probability sample surveys

The main source of the online population in Poland is the Information and Communication Technology (ICT) survey. It has a standardized methodology across European Union members and provides estimates at country or regional levels based on a probability sample of enterprises and households and its members. The latter provide information about the online population by activity (last 3 months, last 12 months) or devices used to connect to the Internet. For detailed description of the design and methodology of the ICT survey see Statistics Poland (2020b).

Despite being a general population survey, between 2017 and 2020 it covered demand side of the platform economy. The ICT survey included the question about transportation services (*Have you used any website or app to arrange transportation (e.g. by car) from another person in the last 12 months?*) with the following options *Yes, from dedicated websites or applications (e.g. UBER, BlaBlaCar)*, *yes, from other websites or applications (including social networking sites)* and *no*. Proposed answers changed overtime and are presented in Appendix A.1. In addition, in 2018 and 2019 it included two questions directly devoted to the gig economy supply side, i.e. *Have you found a paid job through intermediary websites or applications (eg. TakeTask, Sir Local, Upwork, TaskRabbit, Freelancer, Amazon Mechanical Turk) in the last 12 months? (excluding temporary employment agency / employment agency websites)* and *Does the revenue obtained from an order found via intermediary websites or applications constitute?*. Results for both questions are not reported by Statistics Poland which may be due to no or small number of positive answers to these questions. In particular, lack of reporting of about the supply side suggests that microtasks are of negligible size.

As the ICT survey covers online population, the LFS focuses on determining the size of active, inactive and unemployed population according to the ILO definitions. This survey is a probability sample survey with a rotating panel conducted on quarterly basis and the methodology behind it may be found in Statistics Poland (2020c).

At the time of writing the paper, the LFS did not include questions or special sections directly connected to the gig economy. However, there are plans to include a special module in the LFS as a coordinated action suggested by the Group of Experts on Measuring Quality of Employment within United Nations Economic Commission for Europe[1].

In the paper, we used LFS to obtain working 18–65 population at municipality levels in Poland as a reference for the gig economy supply side. Direct estimates and its standard errors were obtained using the methodology used in LFS (results may be found in Appendix D).

[1] See for instance https://unece.org/statistics/events/group-experts-measuring-quality-employment and the Handbook on Measuring Quality of Employment (https://unece.org/statistics/publications/handbook-measuring-quality-employment).

Sample surveys conducted by Statistics Poland are not the only source of information on the online or target population. For example, The Office of Electronic Communications[2] presents yearly report entitled "Public opinion survey on the functioning of the telecommunications services market and consumer preferences" which is conducted by market research companies. This survey covers variety of topics such as mobile phones customers, the Internet access or digital media. This survey provides more detailed information on the online population but sample size is significantly smaller than in surveys conducted by Statistics Poland (about 1600-2000) and the detailed description of the methodology is not published. We used this survey to indicate the level of smartphone coverage in Poland.

Another source that we considered in this study is the non-probability sample of the Internet users conducted by the Polish Internet Research company (PBI; pol. *Polskie Badania Internetu*). It focuses on measuring online audience and their characteristics. PBI between 2016 and 2020 conducted the Megapanel PBI/Gemius survey focused solely on Internet users. Panelists were were asked to install a special computer application that tracked all movement on the Internet. This survey was based on about 150k panelists recruited by pop-ups on main Polish websites.

In 2020 PBI with the Gemius company[3] and the Radio Research Committee[4] launched a new survey – The Mediapanel. It is a combination of internet research standard, the radio audience standard (Radio Track) and data from the single-source cross-media Gemius measurement. The Mediapanel includes passive measurement of Internet, TV and radio consumption. As of 2021 the panel of over 280k users.

We contacted PBI with inquire to provide the estimate about the number of users between 2018 and 2020 using at least one of 25 websites devoted to freelancing or testing such as www.mturk.com, www.clickworker.com, www.fiverr.com or www.upwork.com[5]. Simiarly as in previous cases, almost none of the above-mentioned websites met the minimum number of active users requirement. PBI set up this threshold to 50k for audited (verified by the PBI staff) and 105k for non-audited media channels (not-verified by the PBI staff).

Thus, existing survey based data sources in Poland do not allow to accurately measure the gig economy. But this is not the only limitation of these sources. As they are based on limited sample sizes these surveys may provide reliable estimates at the national level, while the gig economy is connected with using of the Internet and smartphones at municipal level. There are no sources that provide reliable estimates at low level of spatial aggregation or cross-sections. To overcome this, one

[2]The President of the Office of Electronic Communications (UKE) is a regulatory authority responsible for telecommunications and postal activities and frequency resources management. It's also a supervisory authority responsible for controlling compliance of products emitting or vulnerable to emission of electromagnetic field, including radio equipment placed on the market in Poland. For more details please see https://www.uke.gov.pl/en/about-us/.

[3]Gemius is an international research and technology company providing data, as well as advanced tools for digital and traditional marketing activities such as web analytics, online campaigns' management and ad serving. For more see: https://www.gemius.com/about-us.html.

[4]Details may be found here https://badaniaradiowe.pl/.

[5]For the whole list please consult Appendix B.3.

may consider using administrative data or modern data sources such as big data.

3.2 Administrative data

The legal basis regulating passenger transport in Poland is the road transport act (Ustawa z dnia 6 września 2001 r. o transporcie drogowym, Dz.U. 2001 nr 125 poz. 1371). It imposes on entrepreneurs providing transport services, including taxi drivers, obligations, including having the appropriate license. Until 2019, however, this act did not apply to private persons providing this type of services. This legal loophole allowed for applications such as Uber to function in Poland, through which any person meeting the requirements imposed by such an application, but not necessarily meeting the requirements imposed on taxi drivers, could provide transport services. Nevertheless, the problem in the case of Uber applications was the form of payment for the service, which took place without the cash register. However, such a practice was not in line with the regulation on cash registers (Rozporządzenie Ministra Finansów z dnia 20 grudnia 2001 r. w sprawie kas rejestrujących, Dz.U. 2001 nr 151 poz. 1711), which, since 2002, imposes the obligation to register turnover by cash registers on all entities providing services in transport, including private individuals. Such a requirement was related to the installation of a cash register in the vehicle, which in practice was not met by Uber drivers.

The lack of coherent legislation in the field of road transport has had negative consequences. The law that was not adapted to the new forms of service was a source of conflicts between taxi and Uber drivers. Due to the greater burden on taxi drivers, resulting in higher operating costs and less competition, this group of service providers felt a growing sense of injustice. Uber drivers, in turn, were exposed to checks and financial penalties due to the lack of a cash register.

The remedy for the changing situation in the field of road transport and the growing popularity of applications such as Uber was to be an amendment to the road transport act (Ustawa z dnia 16 maja 2019 r. o zmianie ustawy o transporcie drogowym oraz niektórych innych ustaw, Dz.U. 2019 poz. 1180), commonly known as "Uber lex". All the introduced changes included both entrepreneurs and private individuals providing transport services, and their aim was to align the obligations imposed on taxi and Uber drivers. ”Uber Lex” extended the obligation to have a license also to Uber users, but its obtaining was also greatly facilitated. The problem of recording turnover has been solved by enabling the use of the virtual cash registers. The ”Uber lex” came into force in 2020 with a transition period until the end of March, which was then extended to the end of September and the end of December due to pandemics. This period served as a time for Uber drivers to adapt to the new regulations.

A taxi license is valid within the municipality in which it was issued, the body competent for granting taxi license is the head of the municipality. Databases on licensed drivers are created independently by administrators, there is no common IT system with which a central register would be created. Moreover, in practice, data is often collected on paper. Among the information collected as part of the records of people with a taxi license, specified by the act, there is no method of accounting for turnover by the driver (cash register / virtual cash register). The lack of this information makes it impossible to separate people who settle services through the application.

Taking into account the above-mentioned limitations in the register of persons holding a license, and bearing in mind that the obligation to have a license by Uber users has been imposed from 2020, with a transition period lasting until the end of the year, data from municipal registers are not suited for the purpose of measuring the gig economy.

4 Big data from smartphones

4.1 Programmatic advertisement system

Big data sources can be generally described as: "high volume, velocity and variety of data that demand cost-effective, innovative forms of processing for enhanced insight and decision making". In addition to generating new commercial opportunities in the private sector, Big data are potentially a very interesting data source for official statistics, either for use on their own, or in combination with more traditional data sources such as sample surveys and administrative registers (Daas et al., 2015).

Big data sources are used to study gig economy from different angles. For instance, change in income was measured at the national level via transaction-level datasets such as bill-paying applications (Koustas, 2019, cf.) or detailed behaviour and matching between customers and workers based on detailed data generated by Uber[6] or Lyft[7]. However, access to these data is restricted and often limited to researchers from financial institutions or providers of the apps.

In this study, we also used big data but its source was different. In particular, we purchased aggregated data from the Selectivv company[8] that uses *programmatic* advertisement platform to monitor about 350,000 apps and over 17 mln mobile websites of over 20 mln smartphone users in Poland.

Programmatic is a IT system that automates displaying advertisements to the Internet users based on their individual characteristics such as historic behavior on web sites or location. It is based on real-time biding systems that in milliseconds decide whether to show given ad to given user (Busch, 2016).

In practice, this system works as follows. When a user of a mobile device opens a website or an application that contains ad space with an advertisement displayed by some company (e.g. Selectivv), this company receives basic information from the mobile device along with the individual user number. This number is assigned through two systems: *The Google advertising ID* (GAID) and *An identifier for Advertisers from Apple* (IFDA). The first works on smartphones with Android, the second with iOS.

[6]See: https://eng.uber.com/research/.

[7]See https://eng.lyft.com/tagged/data-science.

[8]For more information see https://selectivv.com/en/.

The main difference between data gathered through *programmatic* and mobile phone providers systems is the level of available information. The latter may provide only information their customers, often without distinction between business and private users, the location based on signalling or the type of smartphone. Most of these data are taken from call details records (CDR) that by the Polish law are separated from the consumer relationship management systems. Thus, there is no information about what apps are installed, nor background details about given user.

On the other hand, companies working with *programmatic* systems, cover multiple providers as well as native and foreign population that are often missed by surveys or registers. As almost all apps contain advertisements the level of information collected is rich. It may include: location (GPS / WiFi), time of use, web sites visited, application installed, device type, operating system and its settings (such as language) and mobile operator. *Programmatic* systems do not have access to activities within mobile applications and provide only information about the app name, time spend and when and where it was used (depending on settings on the smartphone).

Based on the passively collected data, companies profile consumers of mobile devices using machine learning algorithms or heuristic rules. In our case, Selectivv claims that their algorithms predict with high accuracy up to 360 variables about a single user. Despite a rich information, the company does not know personality such as name, surname or personal ID, only his/her characteristics. Selectivv also use algorithms that allow to identify users of multiple devices (private and business) based on GPS location and WiFi signals.

In the next subsections we will describe selection of variables used in this study, assessment the coverage of Selectivv data by comparing with existing official statistics and administrative data and mobile apps selected to measure the size of a gig economy.

4.2 Variables derived from smartphones

We obtained aggregated data for various socio-demographic characteristics about smartphones users that were specified during discussion with Selectivv's staff. The type and levels of all variables were limited by the project budget and thus this is the only data that we have. All variables are based on classification or heuristic algorithms but as this constitute Selectivv know-how the level of errors associated with using these methods is unknown. For the study, we selected the following variables:

- `Sex` (Male, Female) – variable is derived based on user's activity (e.g. visited web-pages, installed applications, information provided for the apps) and sample surveys conducted via advertisements systems conducted by Selectivv.

- `Age group` (18–30, 31–50, 51–64) – variable is derived based on user's activity (e.g. visited web-pages, installed applications, information provided for the apps) and sample surveys conducted via advertisements systems conducted by Selectivv.

- `Country` (Polish, Ukrainian, Other) – Ukrainians are defined as people who have SIM card of Polish provider, have set Russian or Ukrainian language at smartphone and at least one time during last year were in Ukraine where may change SIM card to Ukrainian operator.
- `Residence` (Cities, Functional Urban Areas and Province (called also Voivodeships); for more see Appendix C) – this information is derived based on location metadata (e.g. GPS, WiFi) as the most frequent night location (18:00–8:00) for a given period. We provided a shapefile with borders of these areas.
- `Student` – whether given person is a student; derived based on location (e.g. points of interests; POI) and browsing history.
- `Parent of child 0-4` – derived based on POIs such as kinder gardens, offline and online shops visits.
- `Parent of child 5-8` – derived based on POIs such as primary schools, offline and online shops visits.
- `Time spend in the app` – average and standard deviation of time spend in the app (in seconds).

For the study, we defined active user as *a person who used a given app for at least one minute* within a given period (month, half-year between 2018 and 2020). We decided not to use threshold such as one, two or more hours because we did not actually know prior the analysis how long given users use these apps. Moreover, as these apps are dedicated to drivers and couriers they may be more motivated to use them frequently. Finally, the pragmatic reason was a limited budget of our project which allowed to obtain data for long periods such as months or half-years.

We obtained four datasets for the period of 2018 and 2020 for the population 18–64 years:

1. number of active users by app within a given month,
2. number of active users by sex, group age and country within a given half-year; for each cross-section we got share of students, parents with children in 0–4 age and parents with children in 0–4 age,
3. number of active users by city, functional urban areas and provinces by within a given half-year,
4. average and standard deviation of usage time for the following periods: Mondays-Thursdays 8:00-18:00, Mondays-Thursdays 18:00–8:00; Fridays-Sundays 8:00-18:00 and Fridays-Sundays 18:00–8:00 within a given half-year.

All datasets used in the study are made public and available as supplementary materials for this paper.

4.3 Assessment of coverage error

According to the Selectivv website they collected data about 21 mln unique smartphone users in Poland in 2019[9]. This number refers to all Polish and non-Polish citizens. According to Consumer Satisfaction surveys conducted by the Office of Electronic Communications (OEC) there were 22,5 mln[10] Polish citizens with smartphones. This suggest that the coverage of Selectivv data is high. Unfortunately, this company does not report any data regarding background or spatial distribution which limits comparison.

One way to assess the coverage is to verify share and dynamics of foreigners within their databases. This population is often missed by official statistics due to lack of up-to-date sampling frames. In the recent report, Selectivv (2021) reported 1,274 mln Ukrainians as of January 2020. According to the recent experimental studies Statistics Poland (2020a,d) reports 1,351 mln Ukrainians as of 31.12.2019. The latter number was obtained based on the integration of 9 administrative sources that refer to the registered population, in particular economically active. Once again comparison with Selectivv data suggests that they cover population Ukrainian population quite well.

Figure 1 presents dynamics of the Selectivv users by country of origin. It shows that between 2018 and 2020 the number of smartphone users in their database increased by 8% however the change differs between sub-population. For Polish users the increase was around 6% which in absolute numbers was about 2 mln users between 2018 and 2020.

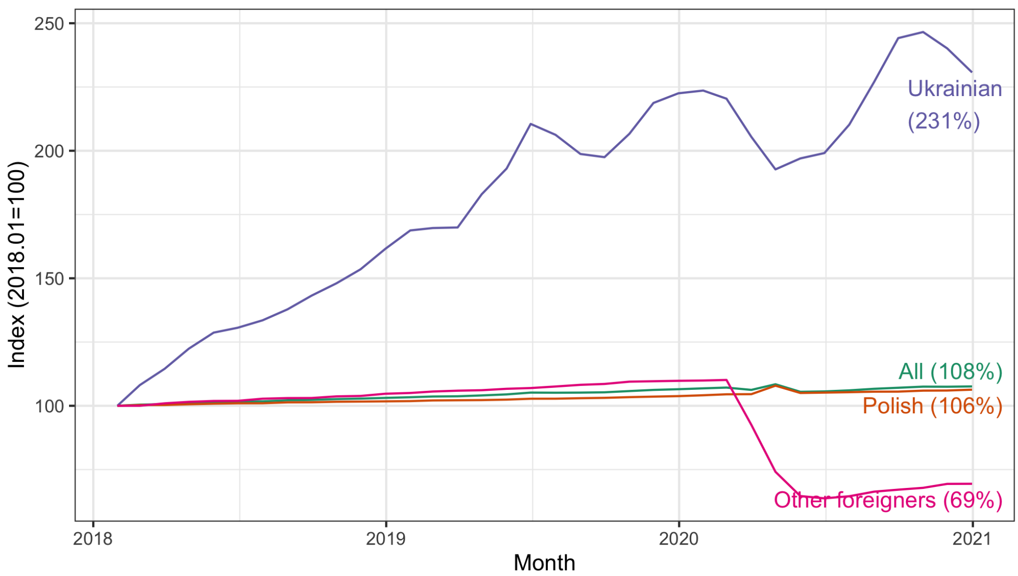

Figure 1: Monthly change in the number of smartphone users covered by Selectivv between 2018 and 2020 (2018.01 = 100)

[9] See https://index.selectivv.com/.

[10] See https://uke.gov.pl/akt/badania-konsumenckie-2019,286.html as of November 2019 70% of people in 16+ had smartphone (96% had one private and 6% had business device). Note that this survey was based on 1600 sample and thus the uncertainty may be high.

The highest changes are visible for Ukrainians and other foreigners. For the latter, we observe large decrease after the first COVID lockdown in Poland in March 2020 but on the contrary to the Ukrainians this number did not change significantly in the rest of 2020. In case of Ukrainians we observe high dynamics with over 200% increase between 2018 and 2020. This increase is explained by the migration to Poland which is also observed in official sources.

To verify further the coverage of foreigners population we compare dynamics with the index of insured foreigners according to the The Social Insurance Institution in Poland (ZUS). If a foreigner is legally employed in Poland then it should be present in the ZUS register. Therefore, this register covers economically active foreigners, which is in line with the structure of migration to Poland from 2018.

Figure 2 presents comparison between monthly and quarterly data from ZUS with Selectivv data by two groups of foreigners: Ukrainians and non-Ukrainians (other). Trend lines for Ukrainians in ZUS, in particular sharp drop and increase in 2020, is similar to those observed in Selectivv data. The main difference for this subgroup is the level of Selectivv and ZUS indices (note we use January 2019 as a baseline). This may be due to the fact that not all Ukrainians may apply for the legal job as well as Selectivv covers also children.

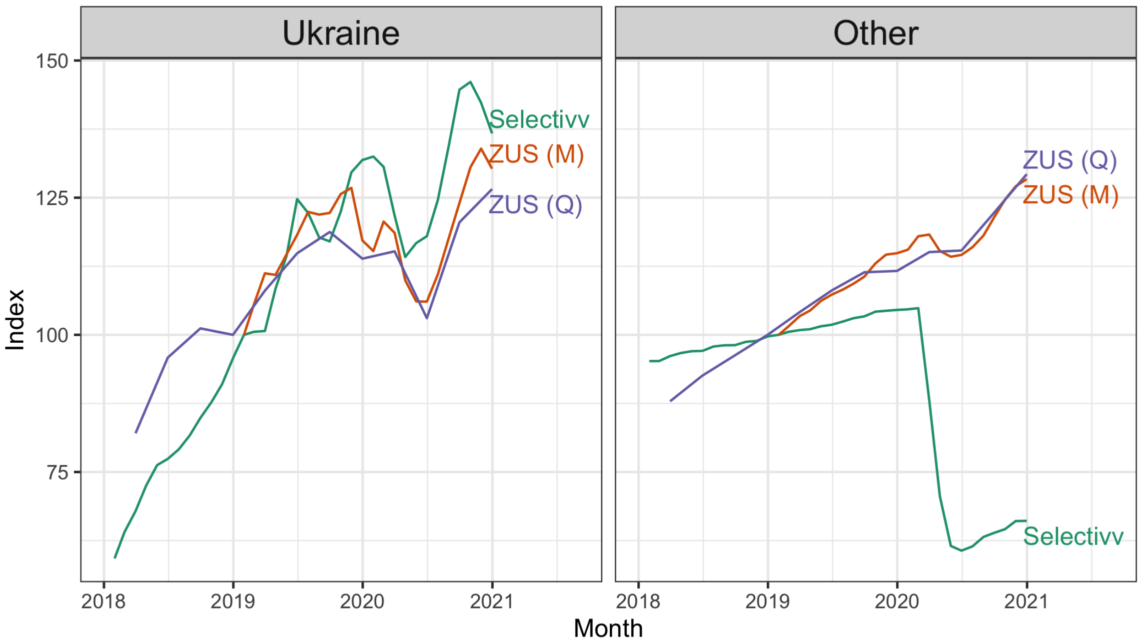

Figure 2: Comparison of change of number of users in Selectivv and insured in ZUS (2019.01 = 100). Note: ZUS (M) – monthly statistics; ZUS (Q) – quarterly statistics

On the contrary, drop in the index for non-Ukrainian foreigners is significantly higher than observed in ZUS registers. There may be several reasons. First, we do not know what is the number of other foreigners in Selectivv databases and thus the decrease may be due to low initial levels. In ZUS we observe about 200k non-Ukrainan foreigners. Second, ZUS covers all foreigners including the ones who may live in Poland for several years and are assimilated (e.g. use Polish language), while Selectivv may cover the short-term working migration. Finally, another possible explanation that the drop refers to the illegally working sub-population, that during the COVID pandemic, lost their source of income.

According to the overall comparison of Selectivv data with existing official or administrative sources we conclude that this source may cover majority of smartphone users in Poland. Certainly, access to more detailed information may improve our justification but to our knowledge there is no other available data source in Poland with this level of coverage. Thus, we believe that this data may provide accurate information regarding the size of the gig economy populations in Poland.

4.4 Applications used in the study

Next, we focused on identification of mobile apps that may be used to identify members of gig workers population. Table 2 contains list of such applications categorized by: transportation, delivery, services, crowdwork and microtasks. We identified provider i.e. company that develops given app for workers and clients as well as specific app devoted only to workers.

This is a crucial part of our analysis. We decided to focus on apps that are prepared only for workers in order to distinguish between demand and supply side. Based on our analysis we found that majority of transportation and deliverable platforms creates apps for workers, while platforms for services, crowdwork or microtasks have only one app for both side of the market. One remark should be done in reference to Uber as it provides one app for both of their services: Uber Eats and Uber. Programmatic services do not have access to what is done within the app so we may have only information about the time use.

In the column *Comments* we specified whether given app was possible to use in the study. For instance, there were three providers Pyszne.pl, lieferando and takeaway which held app takeaway.com courier. Additionally, information about the number of downloads presented in the Google Play store or the App store was misleading. For example, for task rabbit there was over 100k downloads or for Fiverr were over 10 mln app downloads but data provided by Selectiv indicated that none of users in their databases had these apps. This suggests that Google Play may present information about the global number of downloads, not only limited to Poland.

Finally, based on the limitations and our main assumptions, we used 8 apps for further analysis: *Uber*, *Bolt Driver*, *Glover*, *Wolt Courier*, *TakeAway Courier* and *Bolt Courier*. For comparison, we used *iTaxi* and *FREE NOW* as these apps are designed for licensed taxi drivers. Detailed description regarding the apps may be found in Appendix B.1.

Table 2: Mobile applications by categories and whether there is a separate app for workers

Category	Provider	App for workers	Comments
Transportation	Uber	Uber Driver	one app for Uber;✓
	Bolt	Bolt Driver	✓
	Lyft	Lyft Driver	no or small # users
	FREENOW (Mytaxi)	FREE NOW for drivers	✓
	iTaxi	iTaxi Kierowca K3	✓
	Optitaxi	–	no or small # users
Delivery	Deliveroo	Deliveroo Rider	no or small # users
	Glovo	Glover	✓
	UberEats	Uber Driver	one app for Uber; ✓
	Coopcycle	–	GP 5k+
	Take Eat Easy	–	no or small # users
	Pyszne.pl	Takeaway.com Courier	Same as pyszne.pl; ✓
	Lieferando	Takeaway.com Courier	Same as pyszne.pl; ✓
	Takeaway	Takeaway.com Courier	Same as pyszne.pl; ✓
	Bolt Food	Bolt Courier	✓
	Wolt	Wolt Courier Partner	✓
Services	TaskRabbit	Tasker by TaskRabbit	GP 100k+
	Helpling	Helpling Partner	no or small # users
	Fiverr	–	GP 10 mln+
	Upwork	Upwork for Freelancers	no or small # users
	Freelancer	–	GP 5 mln+
	PeoplePerHour	–	GP 100k+
	Toptal	–	GP 10k+
	Guru	–	GP 5k+
	FlexJobs	–	GP 1k+
	Truelancer	–	GP 500k+
	ClearVoice	–	–
Crowdwork	Amazon Mechanical Turk	Turkdroid	GP 50k+
	Clickworker	–	GP 500k+
	Microworkers	–	GP 5k+
	CrowdFlower	–	–
	Spare5	–	–
Microtasks	App Jobber	–	GP 100k+
	ShopScout	–	GP 50k+
	Streetspotr	–	GP 500k+

Note: GP – google play downloads. ✓– app was selected for analysis.

5 Results

Figure 3 presents change in the number of active users (in thousands) between January 2018 and December 2020. Only Glover and Bolt Courier are observed for shorter period as these services started operating in Poland in mid 2019 and beginning of 2020. In particular, growth in the number of users of Glover is significantly higher than other apps. The main reason for that is the scope of activity that is not only delivering food, including fast-foods like McDonald's but also groceries from the biggest chain in Poland – Biedronka.

In case of other delivery apps the Takeaway has the highest absolute increase from 7.5k to close to 14k users at December 2020. On the other hand, Wolt and Bolt Couriers had around 6k and 2k users at the end of 2020. The reason for such discrepancies are the regions where these companies operate. For instance, in Takeaway only 14% were users from Warsaw (Poland's capital city) but for other apps the share was between 35% and 40%. For more details see Appendix 5. While interpretation of these results please note that users of these apps may overlap, but we do not know to what extent.

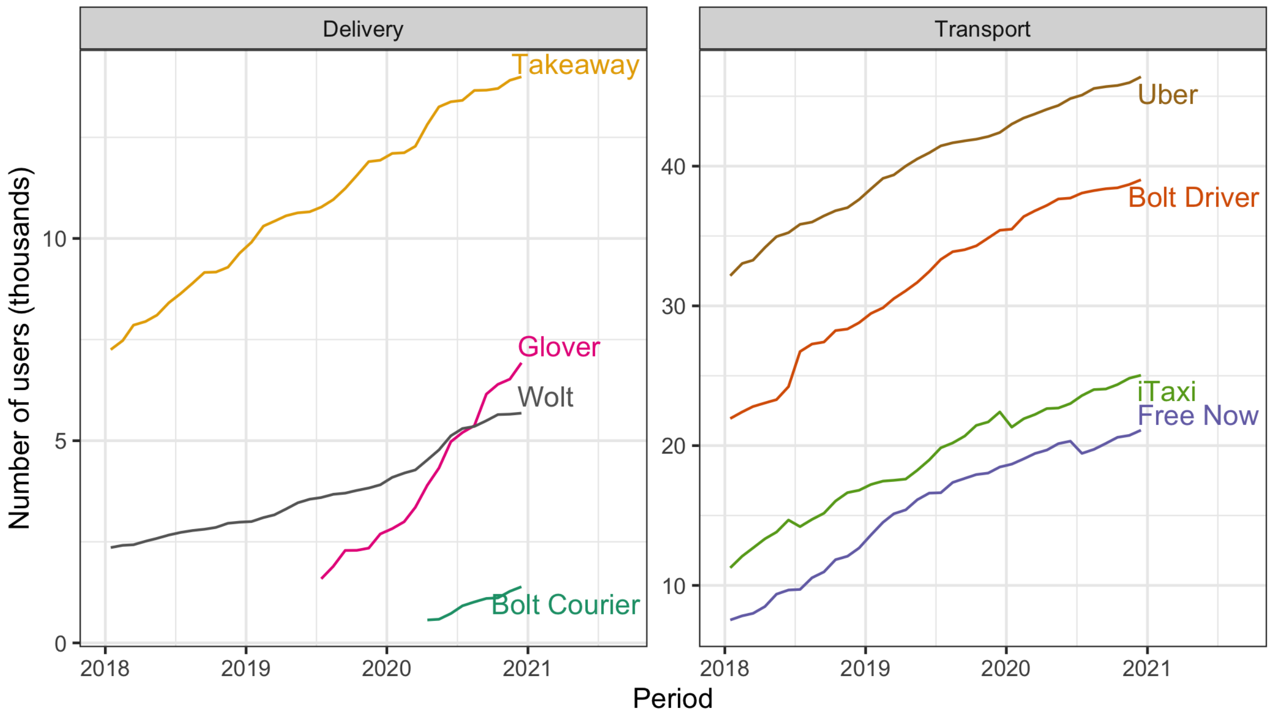

Figure 3: Number of monthly active users of selected apps in Poland by type and between 2018 and 2020

Table 3 contains demographic characteristics of active apps users for the last period in our analysis (2020HY2) that may indicate the selection process. The main difference between transportation and delivery is the age structure. Majority (95%) of couriers are between 18 and 30 which is mainly due to the mode of transportation – bikes and scooters. This mode may also suggest the difference in terms of gender because for drivers females account for about 10-12% while for couriers it is around 5-7% of active users. This findings are in line with those from survey results showing that on average younger people are involved in platform work (Piasna and Drahokoupil, 2019; Brancati et al., 2020).

Age structure may also explain the difference in the share of foreigners with other nationalities, students and parents with 0–4 children. For the first group, majority of recent migrants in Poland are young people from Ukraine, Belarus, Russia, India, Bangladesh or Nepal.

Table 3: Demographic characteristics of gig economy app users at 2020HY2

	Gender		Age group			Country			Student	Parent	
App	Men	Women	18–30	31–50	51–64	PL	UA	other		0–4	5–10
Transportation											
Uber	88.0	12.0	49.2	46.4	4.4	66.0	24.1	9.8	2.7	3.5	3.1
Bolt	86.3	13.7	56.2	37.8	6.0	65.5	26.1	8.5	3.1	4.4	1.9
FREE NOW	88.6	11.4	40.1	52.4	7.5	76.8	21.7	1.5	0.8	2.0	3.2
iTaxi	88.7	11.3	28.2	58.1	13.7	77.9	20.2	2.0	4.4	0.6	1.1
Delivery											
Takeaway	89.7	10.3	94.6	4.5	0.9	62.1	31.1	6.7	8.1	8.1	1.6
Glover	93.8	6.2	94.1	5.7	0.2	61.6	27.5	10.8	7.8	7.8	1.4
Wolt	92.3	7.7	95.7	2.9	1.4	54.2	28.4	17.5	7.9	8.0	1.5
Bolt Courier	94.7	5.3	100.0	–	–	62.3	27.5	10.2	0.0	0.0	0.0

Note: PL – Poland, UA – Ukraine, other – other foreigners.

In the next figure 4 we present change in the number of active users by country (top), age (middle) and sex (bottom) to show dynamics for different apps. The number of Polish users is significantly higher and the curve is steeper than for Ukraine and other countries. For each app we observe an increase of active users that is higher than overall increase presented in figure 1. This suggest that existing users of smartphones may install apps to start providing their services. We also do not observe any COVID-19 effect, in particular for the other category.

For the age group the relationship differ between apps. For delivery, majority of users are people between 18-30 years. The highest increase is for Glover and Takeaway as these apps have the highest coverage and popularity. In case transport, we observe different structure for apps devoted to licensed taxi drivers and Uber and Bolt. For iTaxi majority are drivers between 31-50, while for FREE NOW there is a slight but increasing difference between 18-30 and 31-50. This may suggest that younger taxi drivers select FREE NOW and iTaxi is an app for the taxi corporations. Interesting results are for Bolt and Uber. Variations between age groups may suggest that users of these apps are different users as the there is similar trend and a slight difference between 18-30 and 31-50 age groups for Uber. For Bolt increase in the number of users in 18-30 is higher than other groups but from 2019HY2 growth slows down.

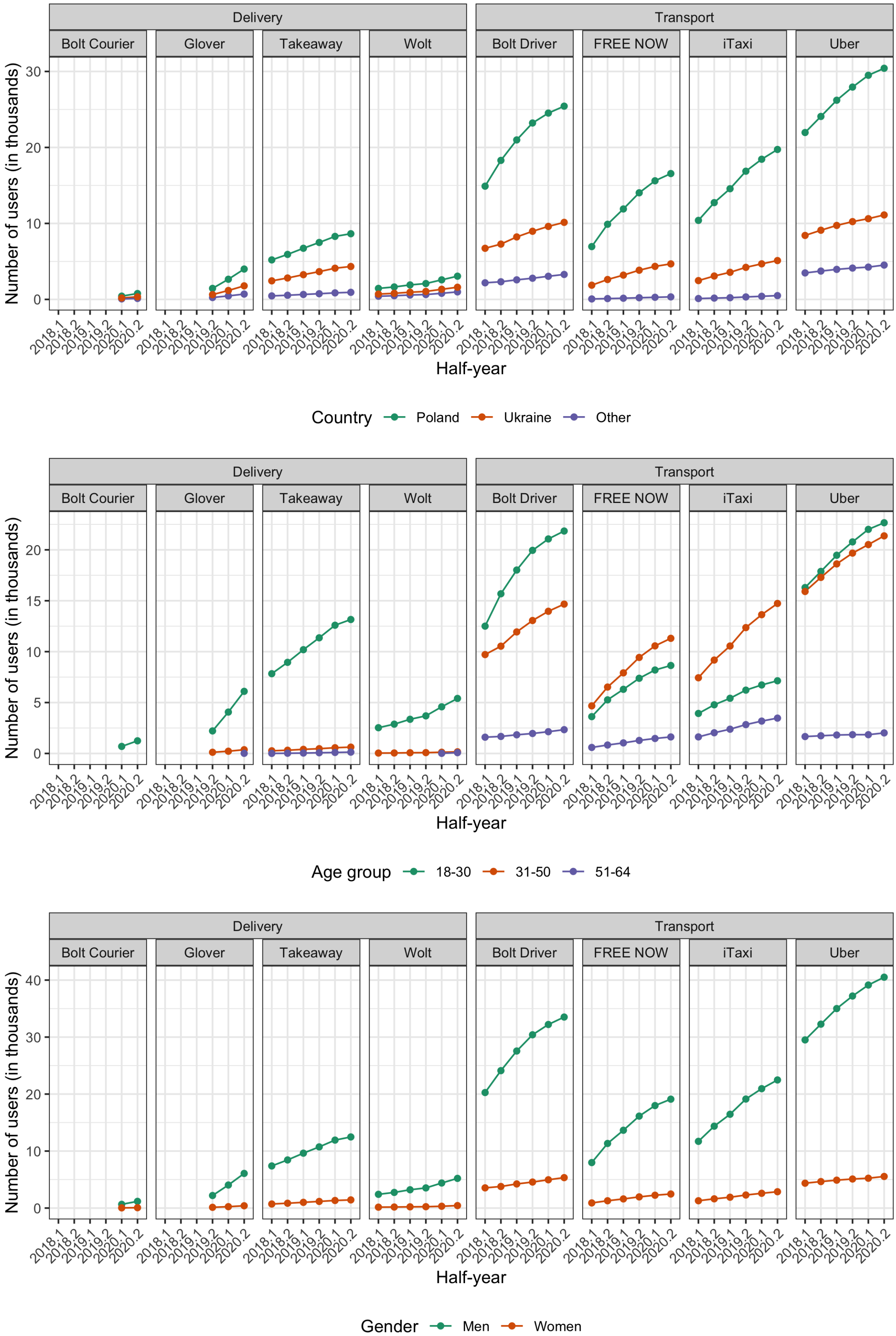

Figure 4: Number of monthly active users of selected apps in Poland by country, age and gender between 2018 and 2020

Finally, as expected majority of users involved in the transportation and delivery are males. Previous evidence for other European countries, show that women account for ca. 13% of riders, see e.g. Drahokoupil and Piasna (2019) for Belgium. Our results therefore confirm this proportion. Interestingly, the share of women involved in platform work is growing, especially if the transportation sector is considered. This trend is recently observed and reported in Europe. Following the survey evidence for European countries provided by Brancati et al. (2020), platform work is becoming a source of income for increasing proportion of women. Importantly, the gender differences are strongly related to the type of task performed as platform work. Women are more likely to be over-represented in feminised task, like the translation or interactive services, while the transportation and delivery services are more male-dominated types of work Brancati et al. (2020).

Furthermore, we investigate what is the share of people providing their services via apps to the working population in a given city based on the LFS estimates. Due to very small sample sizes we do not present results for the Sopot city, thus the comparison is done for 19 Polish cities. As the LFS estimates at the city levels have relatively high variance we decided to also use confidence intervals for comparison. For details about the LFS please refer to Table 7 and for the number of active users to Table 6 in Appendix.

Figure 5 presents 95% confidence intervals (CI) and point of the relation between number of active users to the 16–65 working population for the last period, i.e. 2020HY2. Comparison of transportation apps is not possible for all cities as Uber is available in 9 cities (including Sopot), Bolt in 10 cities (including Sopot) and FREE NOW in 8 cities (including Sopot) while iTaxi and all Delivery apps are available in all study cities.

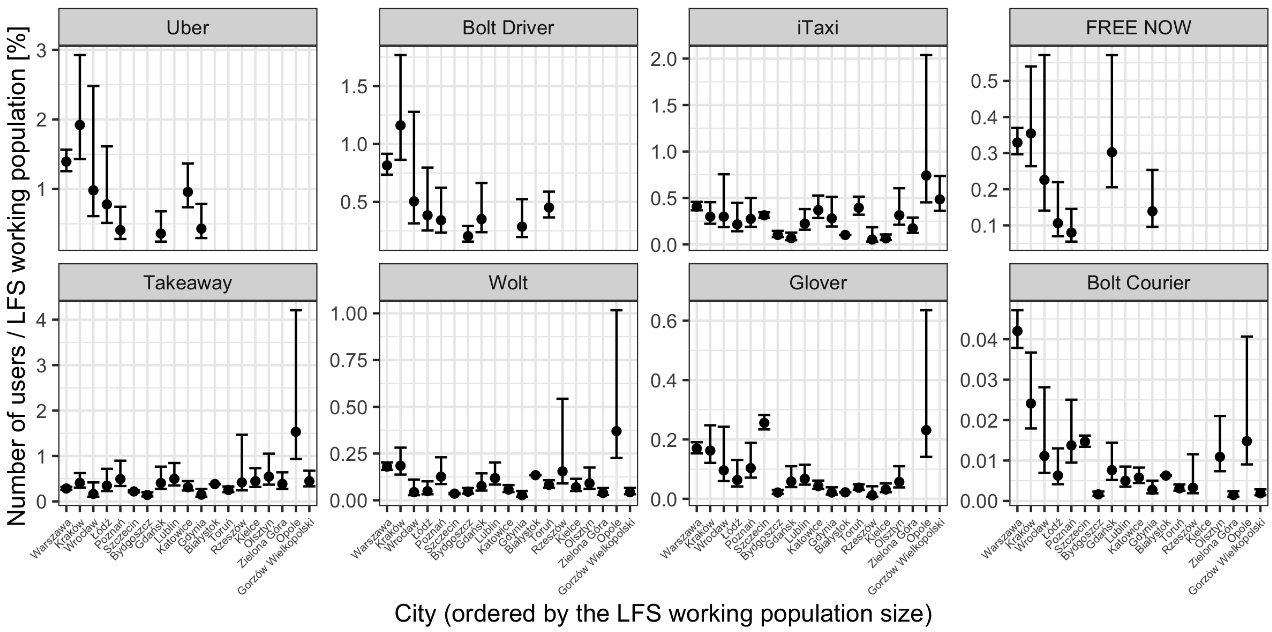

Figure 5: Point and 95% interval of ratio of the number of active users and the LFS working 18–65 population by app for 2020HY1. Note that Y axis differs between the apps.

In general, share of active users of transportation apps is 0.5%-2% of the working population in these cities. Popularity of FREE NOW is rather small as the share is about 0.1% to 0.5%. In general, the number of users is similar between cities, with the highest share Kraków and Warsaw. The reason for that may be due to the their character. Kraków is the city with the large number of foreign tourists and Warsaw is a capital city with the large number of foreign visitors and students. Share of iTaxi users is almost equality distributed between the cities. The reason for that may be policy of the taxi corporations that may optimize the number of taxi drivers to the given market, while Uber or Bolt control the matching between driver and customer.

For the delivery, the share of users is smaller than for transportation apps. Notable example is Takeaway, as the most popular app for ordering food in Poland. The highest variability is observed for Bolt as majority of users operate in Warsaw and Kraków. For other apps, the pattern is similar and about 0.5% of working age population for Takeaway and 0.1-0.2% for Wolt and Glover. Note that, these numbers are upper bound as the working population may be underestimated due to non-response among foreign-born population.

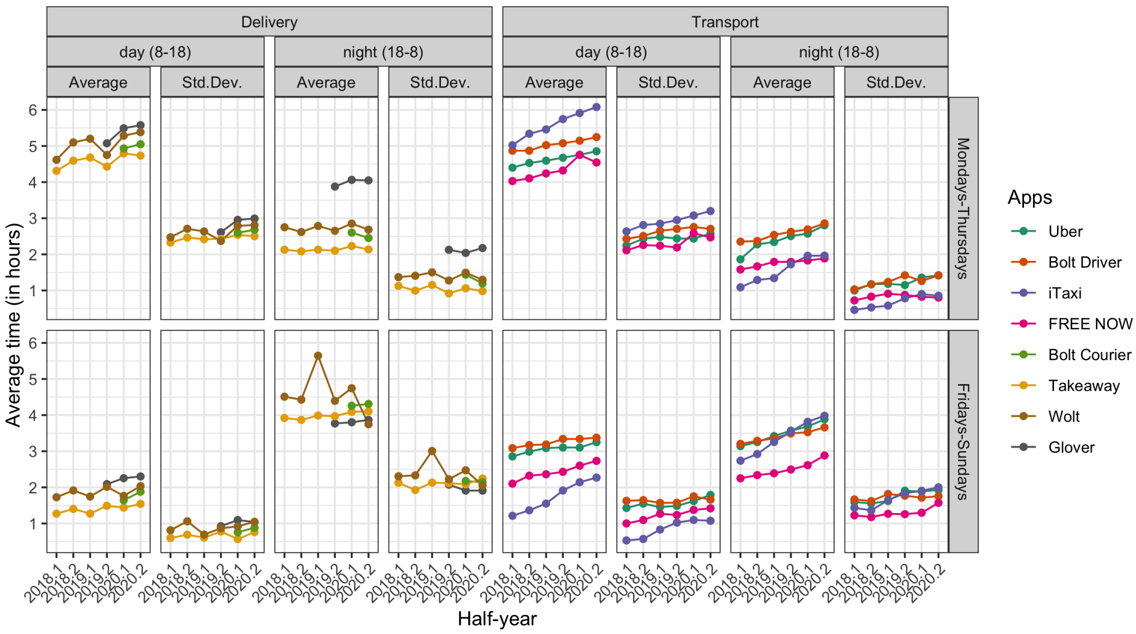

Figure 6: Average and standard deviation (Std. Dev.) of activity time by app type, app, weekdays and hours

Finally, we investigate average and standard deviation of activity time in working days (Mondays-Thursdays) and weekends (Fridays-Sundays) in day time (8:00–18:00) and at night time (18:00-8:00). Figure 6 presents changes in trends in both descriptive statistics.For delivery riders the greater activity is observed during the working days and during the day. However, an opposite trend is reported for weekends, where the level of activity is greater at night compared to the during the day. It is consistent with the intuition that during the weekends, interest in transport services is greatest at night, when people's entertainment activity is greater. However, working at night can be a serious concern when it comes to balancing leisure time with work. For transportation services, in turn, the average activity during the week is greater during the day compared to night hours. At weekends, the activity during the day is lower as during the weekend, but at night is on a similar level.

6 Conclusions

Platform work poses a rapidly increasing working relations worldwide. In the broad sense it covers all job activities conducted with the use of digital platforms and applications. Along with the technology spill over, the prevalence of web-mediated activities is increasing in all economies, both the developing as well developed ones. Against this background, the literature concerning the platform work is growing, delivering new knowledge about this phenomenon. However, due to data scarcity, the evidence on the size of this segment on the labour market, is rather limited.

In this paper we focused on measuring the size and characteristics of the platform economy in Poland. For this purpose we used big data based on advertisements system about 22 mln smartphone users. Using such data it was possible to provide upper bound for the number of drivers and couriers at very low aggregation levels.

Our results show an increasing extent of platform economy in Poland. Focusing on the delivery and transportation activities performed with an use of applications, we report growing trend between January 2018 and December 2020. In particular, the most sharp increase was reported for Takeaway and Glover among the delivery apps, while for the transport apps, the increases were at a similar level. Taking into account demographic structure off apps user, we confirm the evidence existing in the past studies: young men are the dominated group of platform workers. Analysing the gender proportion, a growing trend in favour of the women is observed, especially as far as the transport apps are considered. Putting together the number of platform workers and the working populations our estimates show that the share of active app users accounts for about 0.5-2% of working population in the largest 9 Polish cities.

Despite, undoubted advantages of these source, such as passive collection of data and accurate measure of time use, there are some limitations. While the under-coverage may not be a serious problem of our data, mis-classification is. Not all platforms provide separate apps for workers and potential customers. This affects identification of the members of the gig economy supply side. We cannot determine the threshold which should be applied to distinguish those groups. One may consider using ILO definition of a working person that is based on the question whether given person worked at least one hour in the last week. But applying this rule for real-time data may be a challenge as there is no interview date. We may consider using a reference day, say middle of the month, but this may certainly introduce another errors and limit possibility of comparison with existing official statistics.

Second, mis-classification error introduce additional bias into characteristics such as gender, age or country of origin. All these variables are result of algorithms are not publicly available and based on the partial information disclosed by Selectivv we may have an idea whether their approach is correct or not. One may consider using audit samples to verify the level of errors but this introduces sampling errors as well as additional costs.

Finally, the costs of using big data is non-negligible. Such companies collects gigabytes and terabytes of data in short period and use cloud services to store and process these data. Costs of these services are decreasing but access and computation time makes it very costly to compile required data aggregations and statistics. Prior the analysis, researchers need to specify target population and quantities and negotiate the costs with these companies. This limit possibilities of data exploration as access to individual level data is either forbidden by the company or limited due to how the data is stored.

Another limitation is access to reliable information about the reference, working age, population at low levels of spatial aggregation. Due to small sample sizes in general population surveys or coverage errors in administrative data (i.e. out-dated information, under-coverage of foreign-born population) estimation of the prevalence of the gig economy in these areas is characterized with high mean square error. To overcome this, official statistics and researchers may use small area estimation that use multiple sources to provide reliable estimates for areas of interest (see, for example Van den Brakel and Krieg, 2009; Rao and Molina, 2015).

References

Abraham, K. G., Haltiwanger, J. C., Sandusky, K., and Spletzer, J. R. (2018). Measuring the Gig Economy: Current Knowledge and Open Issues.

Aleksynska, M., Bastrakova, A., and Kharchenko, N. (2019). Working Conditions on Digital Labour Platforms: Evidence from a Leading Labour Supply Economy.

Berg, J. (2016). Income security in the on-demand economy: Findings and policy lessons from a survey of crowdworkers. Technical report, International Labour Office, Geneva.

Bohning, D., Van der Heijden, P. G., and Bunge, J. (2017). *Capture-recapture methods for the social and medical sciences*. CRC Press.

Bonin, H. and Rinne, U. (2017). Omnibusbefragung zur Verbesserung der Datenlage neuer Beschäftigungsformen. Technical report, IZA.

Brancati, U., C., Pesole, A., and Fernández-Macías, E. (2020). New evidence on platform workers in

Europe. Results from the second COLLEEM survey. Technical report, Publications Office of the European Union, Luxembourg.

Busch, O. (2016). *Programmatic advertising.* Springer, Berlin.

Daas, P. J., Puts, M. J., Buelens, B., and van den Hurk, P. A. (2015). Big data as a source for official statistics. *Journal of Official Statistics*, 31(2):249.

De Stefano, V. and Aloisi, A. (2018). *European Legal Framework for "Digital Labour Platforms".* European Commission, Luxembourg.

Drahokoupil, J. and Piasna, A. (2019). Work in the Platform Economy: Deliveroo Riders in Belgium and the SMart Arrangement.

Eurofound (2018). Employment and working conditions of selected types of platform work. Technical report, Publications Office of the European Union, Luxembourg.

Eurofound (2019). Platform work: Maximising the potential while safeguarding standards ? Technical report, Publications Office of the European Union, Luxembourg.

Farrell, B. D., Greig, F., and Hamoudi, A. (2019). The Evolution of the Online Platform Economy: Evidence from Five Years of Banking Data. *AEA Papers and Proceedings*, 109:362–366.

Friedman, G. (2014). Workers without employers : shadow corporations and the rise of the gig economy. *Review of Keynesian Economics*, 2(2):171–188.

Hauben, H., Lenaerts, K., and Wayaert, W. (2020). The platform economy and precarious work. Publication for the committee on Employment and Social Affairs. Technical report, Policy Department for Economic, Scientific and Quality of Life Policies, European Parliament, Luxembourg.

Huws, U., Spencer, N. H., Coates, M., and Holts, K. (2019). The platformisation of work in Europe: results from research in 13 European countries. Technical report, Foundation for European Progressive Studies, UNI Europa and University of Hertfordshire, Brussels.

ILO (2021). *World Employment and Social Outlook. The role of digital labour platforms in transforming the world of work.* International Labour Office, Geneva.

Koustas, B. D. K. (2019). What Do Big Data Tell Us about Why People Take Gig Economy Jobs? *AEA Papers and Proceedings*, 109:367–371.

Koutsimpogiorgos, N., Slageren, J. V., Herrmann, A. M., and Frenken, K. (2020). Conceptualizing the Gig Economy and Its Regulatory Problems. *Policy & Internet*, pages 1–21.

Lee, Z. W. Y., Chan, T. K. H., Balaji, M., and Chong, A. Y.-L. (2018). Why people participate in the sharing economy: an empirical investigation of Uber. *Internet Research*, 28(3):829–850.

Lehdonvirta, V. (2018). Flexibility in the gig economy : managing time on three online piecework platforms. *New Technology, Work and Employment*, 33(1):13–30.

Lenaerts, K. (2018). Industrial Relations and Social Dialogue in the Age of Collaborative Economy. Technical report, CEPS.

Meszmann, T. T. (2018). Industrial Relations and Social Dialogue in the Age of Collaborative Economy. National Report Hungary. Technical Report 27, CELSI.

Mitea, R. D. E. (2018). The Expansion of Digitally Mediated Labor: Platform-Based Economy, Technology-Driven Shifts in Employment, and the Novel Modes of Service Work. *Journal of Self-Governance and Management Economics*, 6(4):7–12.

Piasna, A. and Drahokoupil, J. (2019). Digital labour in central and eastern Europe: evidence from the ETUI Internet and Platform Work Survey.

Rao, J. and Molina, I. (2015). *Small Area Estimation*. Wiley Series in Survey Methodology. Wiley, 2 edition.

Riggs, L., Sin, I., and Hyslop, D. (2019). Measuring the "gig" economy: Challenges and options.

Särndal, C.-E., Swensson, B., and Wretman, J. (2003). *Model assisted survey sampling*. Springer Science & Business Media.

Sedláková, M. (2018). Industrial Relations and Social Dialogue in the Age of Collaborative Economy. National report: Slovakia. Technical Report 28, CELSI.

Selectivv (2021). Ukrainians in Poland 2020 – did they leave, did they come or stayed? https://selectivv.com/en/ukrainians-in-poland-2020-did-they-leave-did-they-come-or-stayed/.

Shao, J. and Tu, D. (2012). *The jackknife and bootstrap*. Springer Science & Business Media.

Statistics Poland (2020a). *Appendix – The foreign population in Poland during the COVID-19 pandemic*.

Statistics Poland (2020b). *Information society in Poland in 2020*.

Statistics Poland (2020c). *Labour force survey in Poland IV quarter 2020*.

Statistics Poland (2020d). *The foreign population in Poland during the COVID-19 pandemic*.

Van den Brakel, J. A. and Krieg, S. (2009). Estimation of the monthly unemployment rate through structural time series modelling in a rotating panel design. *Survey Methodology*, 35(2):177–190.

Appendix

A Sample surveys – details

A.1 The ICT survey

The question *Have you used any website or app to arrange transportation (e.g. by car) from another person in the last 12 months?* is asked only for people who used the Internet in the last year. Possible answers changed overtime which may be found below:

- 2017
  - yes, from dedicated websites or applications (e.g. UBER, BlaBlaCar)
  - yes, from other websites or applications (including social networking sites)
- 2018
  - yes, from intermediary websites or applications dedicated specifically to organizing transport (e.g. BlaBlaCar, yanosiktls.pl, jedziemyrazem.pl)
  - yes, from other websites or applications (including social networking sites)
- 2019
  - yes, from websites or applications that specialize in organizing trips (e.g. BlaBlaCar, jedziemyrazem.pl)
  - yes, from other websites or applications (including social networking sites)
- 2020
  - yes, offered by the company (e.g. public transport, plane, taxi, Uber, Bolt, carsharing, electric scooters),
  - yes, offered by a private person (e.g. BlaBlaCar, jedziemyrazem.pl)
- 2021
  - yes, offered by the company (e.g. public transport, plane, taxi, Uber, Bolt, carsharing, electric scooters),
  - yes, offered by a private person (e.g. BlaBlaCar, jedziemyrazem.pl)

B Mobile apps and websites identified for the study

In this section we describe apps used in the study focusing on: 1) regional availability; 2) application process and who may apply; 3) how given person specify working ours; 4) how given person is informed about the work.

B.1 Gig economy apps

B.1.1 Takeaway.com courier

B.1.2 Glover

B.1.3 Wolt courier

B.1.4 Bolt courier

B.1.5 Uber driver

B.1.6 Bolt driver

B.2 Non-Gig economy apps

B.2.1 FREE NOW for drivers

B.2.2 iTaxi Kierowca K3

B.3 Websites

For the study we identified the following websites:

- http://testuj.pl
- http://testarmy.com/
- http://www.utest.com/
- http://www.applause.com/
- http://whatusersdo.com/
- http://trymyui.com/
- http://www.testbirds.com/
- http://www.usertesting.com/
- https://crowdsourcedtesting.com/
- https://mycrowd.com/

- https://test.io/
- https://usabilityhub.com/
- https://globalapptesting.com/
- https://www.bugfinders.com/
- https://www.fiverr.com/
- https://www.upwork.com/
- https://www.freelancer.pl/
- https://www.toptal.com/
- https://www.guru.com/
- https://www.flexjobs.com/
- https://www.truelancer.com/
- https://www.clearvoice.com/
- https://www.mturk.com/
- https://www.clickworker.com/
- https://www.microworkers.com/

C Functional urban areas

At the time of joining Urban Audit project Polish public statistics did not have data on commuting. Thus in 2003, 2006, 2009 and 2011 editions of the project LUZ was defined as a one or more rings of LAU1 and LAU 2 units surrounding the city. For the 2012 edition of the project broader spatial coverage of Functional urban areas for Polish cities participating in the project has been verified (and the newly proposed by Eurostat cities was defined) on the basis of the results of "survey of employment-related population movements in 2006". The Functional urban area is an area in which 15% or more of the population commutes to work to the city center.

In the project we focused on province city cores, their functional urban areas and the rest of the province. List of all spatial aggregation units are presented in Table 4.

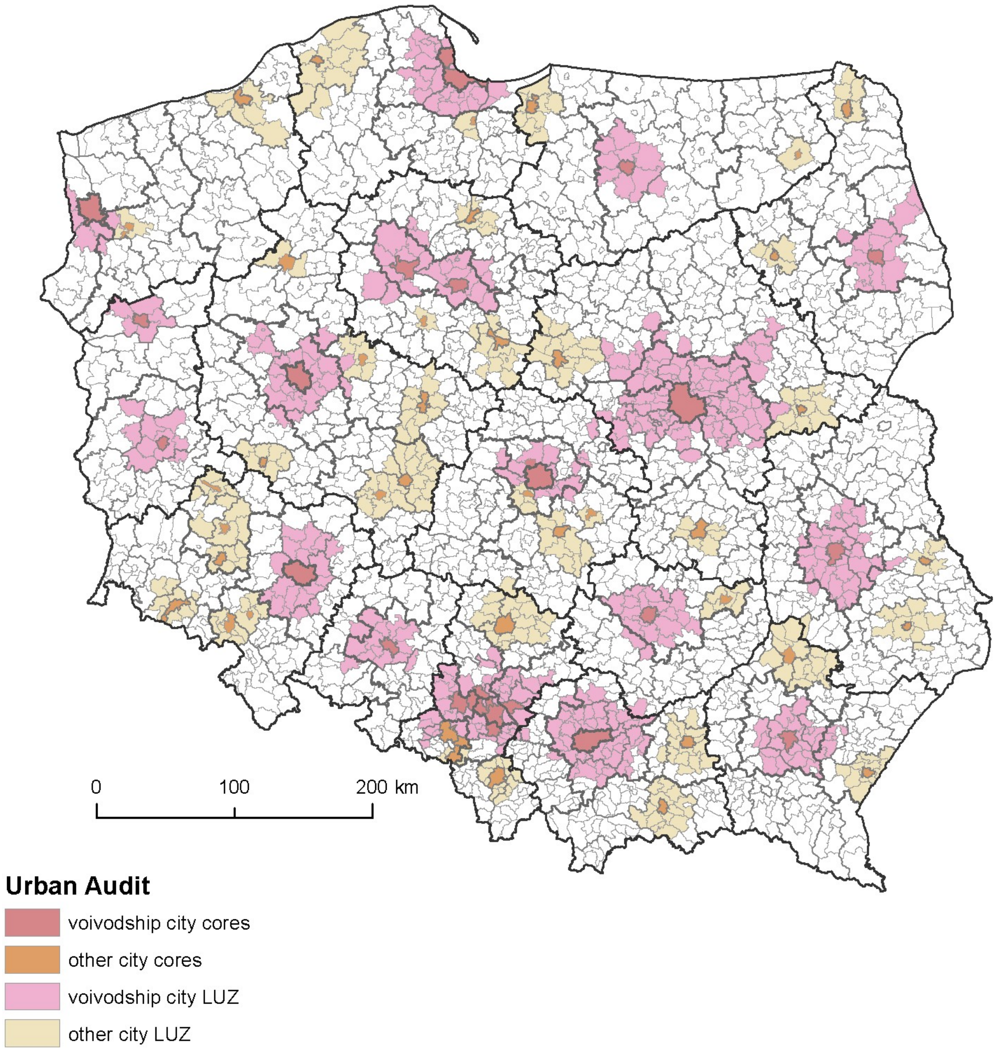

Figure 7: Functional urban areas (FUAs) in Poland in 2017. Source: https://stat.gov.pl/en/regional-statistics/regional-surveys/urban-audit/larger-urban-zones-luz/

Table 4: Spatial aggregation defined in the study based on functional urban areas

Province	**Functional area**	**City**
Dolnośląskie	Wrocław	Wrocław
Kujawsko-pomorskie	Bydgoszcz	Bydgoszcz
	Toruń	Toruń
Łódzkie	Łódź	Łódź
Lubelskie	Lublin	Lublin
Lubuskie	Gorzów Wielkopolski	Gorzów Wielkopolski
	Zielona Góra	Zielona Góra
Małopolskie	Kraków	Kraków
Mazowieckie	Warszawa	Warszawa
Opolskie	Opole	Opole
Podkarpackie	Rzeszów	Rzeszów
Podlaskie	Białystok	Białystok
Pomorskie	Trójmiasto	Gdańsk
		Gdynia
		Sopot
Śląskie	GZM	Katowice
		GZM2 (13 cities)
Świętokrzyskie	Kielce	Kielce
Warmińsko-mazurskie	Olsztyn	Olsztyn
Wielkopolskie	Poznań	Poznań
Zachodniopomorskie	Szczecin	Szczecin

Note: GMZ2 cities are: Bytom, Chorzów, Dąbrowa Górnicza, Gliwice, Jaworzno, Mysłowice, Piekary Śląskie, Ruda Śląska, Siemianowice Śląskie, Sosnowiec, Świętochłowice, Tychy and Zabrze.

Table 5: Number of active users in Warsaw, all cities and in total for 2020HY2

App	Warsaw	All Cities	Total	Warsaw % of total	Cities % of total
Transportation					
Uber	12,637	27,868	46,054	27.4	60.5
Bolt Driver	7,799	16,878	38,855	20.1	43.4
iTaxi	3,692	10,024	25,345	14.6	39.6
FREE NOW	2,954	5,692	21,573	13.7	26.4
Delivery					
Takeaway	1,878	13,563	13,910	13.5	97.5
Wolt	1,956	5,060	5,645	34.7	89.6
Glover	2,358	5,906	6,482	36.4	91.1
Bolt Courier	533	1,168	1,243	42.9	94.0

Note: Warsaw % of total is calculated as Warsaw to Total, Cities % of total is calculated as All Cities to Total. All Cities include cities listed in Table 4 except for GMZ2 (13 cities).

Table 6: Number of active users in cities for 2020HY2

	Transport				Delivery			
City	Uber	Bolt	iTaxi	FREENOW	Takeaway	Wolt	Glover	Bolt
Białystok	–	–	96	–	630	205	56	24
Bydgoszcz	–	274	178	–	297	101	71	12
Gdańsk	472	542	104	450	778	156	170	32
Gdynia	479	401	373	120	245	48	53	10
Gorzów Wielkopolski	–	–	178	–	311	27	15	5
Katowice	1578	–	482	–	558	106	121	21
Kielce	–	–	12	–	442	75	65	18
Kraków	6808	4166	994	1219	1619	792	893	170
Łódź	1762	793	405	230	908	140	261	33
Lublin	–	–	344	–	958	226	187	18
Olsztyn	–	–	190	–	533	90	122	20
Opole	–	–	310	–	912	238	204	30
Poznań	835	701	506	64	1288	350	410	62
Rzeszów	–	–	41	–	459	169	36	22
Sopot	80	249	179	35	196	38	42	8
Szczecin	–	–	461	–	387	52	254	55
Toruń	–	336	370	–	276	94	66	11
Warszawa	12637	7799	3692	2954	1878	1956	2358	533
Wrocław	3217	1617	931	620	540	165	505	78
Zielona Góra	–	–	178	–	348	32	17	6

D Labour Force Survey

The estimation process in Polish LFS is based on survey weights, They are derived by correction of primary weights determined as the reciprocals of selection probabilities for dwellings, which compensate the disproportionate construction of the sample. In the first stem of correction, secondary weights are calculated by dividing primary weights by interview rates $R = (K - N)/K$, where K is the number of interviewed dwellings and N is the estimate of the number of dwellings that were qualified for the survey but were not interviewed. The interview rates are calculated across NUTS 2 regions and categories of place of residence. In the second step of correction, Final weights for the results concerning population are calculated in the third step. The purpose of this step is the adjustment of the LFS results to the current demographic estimates. It is given by calculation of the so-called modifiers for each of 48 categories defined by the place of residence (urban/rural), sex and 12 age groups separately for all 17 NUTS 2 regions. The modifiers are calculated by dividing the number of people in each group according to the demographic estimates by the number of people in these categories calculated from the LFS results with the appliance of the secondary weights from the second step. The final weights result from multiplication of the secondary weights by adequate modifiers. The final weights for generalizations concerning households are calculated as mean values of final wages attributed to household members. In case when only some persons in a household submitted the questionnaire for the respondents, other members of such household are given zero final weight. The final weight for a household comprises the mean value of the weights of the household members including persons with zero final weight. The final weight is attributed to the person who is the household head, while the estimate of the number of households in a given quarter is calculated by adding such weights for all household heads.

The above process results finally in complex ("multi-storey") estimators of the ratio type. The variances of such complex estimators cannot be estimated with the ordinary textbook methods and some special, approximate procedures must be employed. Since 2003, it has been decided to use for this purpose one of the most popular approximate methods, based on the resampling and bootstrap rule (see for instance Särndal et al. (2003)). Resampling based methods make it possible treating different parameter estimators uniformly and they eliminate the necessity for derivation of complicated analytical formulae. The application of bootstrap rule to the complex sample designs most frequently used in sample surveys requires introducing adequate modifications. Detailed review of such aspects can be found in the monograph of Shao and Tu (2012). In the case of a complex two-stage sampling design used in the LFS, variance estimation is performed with the use of data from quarterly samples, drawn with the appliance of stratification at the level of primary sampling units. The chosen variant of bootstrap method is applied separately in each stratum to obtain the corresponding variance component estimate.

In each stratum, from which n_h primary sampling units were drawn initially, a random sub-sample of $n_h - 1$ elements is drawn independently by simple random sampling with replacement. After drawing a sub-sample of primary sampling units the respective dwellings (together with respondents), sampled from them for the LFS, are collected to form a bootstrap sub-sample of dwellings. Such procedure is repeated B times, and for the b-th sub-sample of respondents ($b = 1, 2, \ldots, B$) modified weights are determined which are then used to calculate the estimate of the surveyed parameter $\hat{t}_b^*$.

Weight modification comprises their calibration which assures consistency with demographic data for a surveyed quarter. After B iterations we compute the bootstrap variance estimate for the surveyed parameter estimator $\hat{t}$ (in each stratum) according to the formula:

$$\hat{V}(\hat{t}) = \frac{1}{1 - B} \sum_{b=1}^{B} (\hat{t}_b^* = \hat{t})^2. \tag{1}$$

In order to obtain variance estimate for the considered subpopulation all the variance component estimates obtained independently for subpopulation strata must be summed up. Number of bootstrap iterations B recommended in the relevant literature should be of the order of several hundreds. In the computations presented in this publication we assumed $B = 500$.

Table 7: Direct estimates, its standard errors (in thousands) and sample sizes of the working population size (18–64) based on the Labour Force Survey between 2018HY1 and 2020HY

	2018HY1			2018HY2			2019HY1			2019HY2			2020HY1		
City	$\hat{N}$	$SE(\hat{N})$	n	$\hat{N}$	$SE(\hat{N})$	n	$\hat{N}$	$SE(\hat{N})$	n	$\hat{N}$	$SE(\hat{N})$	n	$\hat{N}$	$SE(\hat{N})$	n
Białystok	132.2	2.9	988	132.4	11.5	934	128.2	0.6	892	138.9	1.6	879	127.4	0.1	845
Bydgoszcz	175.9	20.2	1359	170.4	24.0	1250	165.5	13.4	1179	163.3	20.2	1046	183.6	27.9	1052
Gdańsk	180.5	33.0	847	190.5	42.3	842	207.0	16.4	852	206.6	36.4	807	170.3	40.9	660
Gdynia	142.7	3.9	363	145.1	14.4	351	134.8	8.1	326	132.4	6.9	292	145.3	33.5	319
Gorzów Wielkopolski	67.0	9.8	1002	59.0	13.8	809	63.4	5.9	848	53.9	11.7	722	52.8	9.2	713
Katowice	139.1	29.5	819	142.3	20.4	725	150.3	16.9	705	167.4	37.0	690	155.5	23.7	661
Kielce	84.7	16.5	869	88.9	22.6	870	86.4	19.0	884	86.8	29.6	867	89.0	17.9	807
Kraków	408.0	15.3	904	418.3	20.3	793	415.6	63.3	776	430.5	53.5	756	360.8	63.2	640
Lublin	176.0	20.7	1371	167.6	19.2	1195	162.8	26.5	1156	153.0	30.4	1033	160.0	33.7	1022
Łódź	328.5	16.1	637	334.3	16.9	510	291.5	1.8	489	298.3	18.2	448	238.7	63.0	442
Olsztyn	79.5	18.7	959	83.3	27.2	967	76.4	25.9	891	83.6	22.9	878	82.5	20.2	953
Opole	55.5	16.5	786	59.8	11.8	800	60.1	16.3	841	56.4	15.5	782	54.1	17.6	870
Poznań	268.5	41.9	893	269.0	64.9	841	259.0	28.7	774	278.1	65.1	719	232.6	53.5	785
Rzeszów	101.8	22.9	1050	105.3	31.8	1022	107.5	30.7	1095	104.7	27.0	991	91.1	33.2	1052
Sopot	6.0	17.4	16	4.2	23.8	10	7.0	12.8	17	3.4	10.4	8	3.5	10.7	8
Szczecin	200.5	37.1	1147	196.2	20.5	1055	205.0	14.0	951	189.3	25.8	759	191.1	9.2	671
Toruń	91.5	12.8	1176	98.8	10.0	1211	92.3	14.7	1055	95.2	14.4	1039	95.8	11.4	916
Warszawa	968.2	26.8	1827	978.5	63.5	1677	899.7	28.9	1660	944.8	19.4	1445	909.2	50.7	1454
Wrocław	320.9	61.2	858	350.5	49.1	1006	343.2	67.8	977	342.2	92.4	833	314.6	97.0	805
Zielona Góra	63.9	11.0	856	65.9	10.9	810	65.2	16.0	784	65.9	12.3	790	68.5	13.9	811