MASS Handout 2022
"DeliverAndo": Combining traditional survey data, ecological momentary assessments, and GPS sensor data through a smartphone-based travel app

Ricardo González, Adolfo Fuentes, Johnjairo Meléndez, Vicente Galaz, and Valentina Salvatierra
LEAS at Universidad Adolfo Ibáñez

1. Introduction

In recent years, Santiago de Chile (like other cities worldwide) has witnessed a surge in digital platform workers (particularly for ride-hailing and delivery companies) moving around the city. Bravo and Castillo (2021) estimate that, in june 2021, 190 thousand workers provided transportation or delivery services over the last 24 months, while 98 thousand have worked for these platforms the last month. Additionally, almost a quarter of households reported having used delivery platforms in the last 24 months. Public discussions have emerged about the precariousness of their working conditions. Nowadays, in Chile, digital platform workers do not have the benefits associated with formal employment, such as labor rights, vacation days and payments to the social security system (health, unemployment insurance, pensions, among others). Nevertheless, in September 2022, a new law will regulate this situation, mainly by giving workers the option to establish a formal employment relationship with digital platforms.

Past research on these workers has relied on the use of traditional surveys to grasp their work routines. For instance, Fielbaum and Tirachini (2020) survey Uber and Cabify drivers in Santiago de Chile about their work routines. They find that these drivers value the flexibility provided by the job but generally have fixed routines, they tend to work long hours and are exposed to dangerous situations. Other researchers have relied on qualitative approaches (e.g. Gandhi et al. 2018, Kasliwal 2020). While these approaches are informative, their findings depend on people's recall capacity and research shows that people tend to overestimate duration, underestimate distance, and that they are not accurate in reporting when a trip starts (Wolf et al. 2003; Forrest and Pearson 2005; Bricka et al. 2009; Stopher and Shen 2011; Kelly et al. 2013; Carrion et al. 2014).

Looking to diminish the measurement error associated with self-reports, our survey research center, LEAS at Universidad Adolfo Ibáñez, began developing a smartphone-based travel app. In order to test this app, LEAS teamed up with Chilean researchers associated with the Fairwork project (based at the Oxford Internet Institute and the WZB Berlin Social Science Centre) to focus the first prototype of the app, "DeliverAndo", on digital platform workers. To the best of our knowledge, a granular characterization of these workers' daily routines is not available.

The structure of this document is as follows. Section 2 presents the substantive research questions that we want to address as they guide the design of the app. Section 3 focuses on the data collection methods, particularly on the inner workings of the DeliverAndo app. Section 4 discusses the pre-test results of both the profiling web survey and the app and the adjustments to the app in order to make it user-friendly. Section 5 discusses next steps.

2. (Substantive) research questions

First of all, our main objective is to characterize the working routines of digital platform workers with precision and granularity. Specifically, we want to learn about when and how many hours they work, and the number, distance, length and speed of trips made. These measurements not only help characterize their workday, but more deeply, they give clues about how digital platform workers plan (or not) their workdays. This is particularly interesting as the relationship between the worker and the digital platform is mediated by an algorithm that is assumed to be programmed to maximize the platform's profit.

In addition, we want to link their working routines, particularly those aspects named above, to their subjective experience over a workday. For instance, we can address the association of the feeling of insecurity with risks associated with traffic. We can also assess the associations of their perceptions on economic stability, their mental health status and labor satisfaction with the number, distance, length and speed of trips made each day (cross-sectionally) and over several days (longitudinally). These are two examples, but there are a handful of additional hypotheses that we can explore with this dataset (the list of all modules is listed below).

3. Data collection methods

We focus on digital platform workers for delivery companies. Due to the absence of a sampling frame, we use snowball sampling. Recruiters contact them where they tend to wait for requests for their services and lead them to a web-based profile survey. This survey collects demographics and attitudes about work and digital platforms. It is conducted externally to the application. Once workers respond to this survey, they are invited to download the DeliverAndo app. To use the app, they must complete a registration process (create an username and password) and agree to two informed consent forms. The first form asks for the authorization to link data from the survey with data collected by the app. The second form asks for agreement with terms and conditions. To complete the registration process, workers are asked to report the platform they work for using the smartphone where the app was downloaded.

3.1 DeliverAndo's user interface

After completing the registration process, the application asks respondents to set the departure and arrival points for each work trip. When they press the "end" button, questions on their subjective experience of the journey are triggered. We discuss the possibility of implementing a stop detection algorithm, following McCool et al. (2020), as respondents might forget to press the "start" button or the "end" button, which would lead to an under- or over-estimation of length, distance and speed. Nevertheless, we decided against it for two reasons. First, because we believe that it is natural for them to look at the app they use to work when the trip starts and ends and we tried to recreate that with DeliverAndo (see pre-testing section below). Second, Santiago de Chile is a massive city –seven times larger than Amsterdam in terms of inhabitants and three times larger in terms of size– and traffic jams are quite common in some areas but in others they happen unpredictably, for example, as a result of demonstrations against the government. These demonstrations have increased in frequency over the last decade, but particularly since the start of the social crisis in October 2019 (see González and Le Foulon 2020). These difficulties, specific to the Chilean context, made us decide against programming a stop detector in DeliverAndo's current (beta) version.

The questions triggered after each work trip are divided into twelve thematic modules, where each module contains about 3 questions on the topic. For example, the block of questions that make up the "risks" module are as follows:

- How worried were you about being involved in a traffic accident during your last trip?
- How worried were you about being robbed during your last trip?
- How worried were you about having your last trip canceled?

The modules are the following: risk, self-sufficiency (autonomy), digital cage (control), work breaks, mental health, labor satisfaction, relationship with the platform, relationship with the customers and discrimination.

The in-app questionnaire includes close-ended questions only. Some items include a list of five response categories, all of them labeled. Other items include a 11-point scale, with labeled endpoints, where users have to use a slider to register their response.

3.2 DeliverAndo's algorithms

The app is only available for Android devices at the moment. While it is possible to make it available on iOS devices, we decided against it because a minority of the general population uses them, especially among people working for digital platforms. The app is not available in the Play Store because the app is still in its beta phase and legal issues that we have not sorted out just yet.

The app collects processed and raw data. On the one hand, the processed data includes location and time. The app updates location and time every 20 seconds, to avoid consuming much battery life from mobile devices. These data are considered sensitive by digital platforms workers (some of them are illegal immigrants). For that reason, the app processes time and location on the device so these data are not stored on our server. On the other hand, the unprocessed data includes the responses to the questions triggered at the end of a work trip. **Figure 1** illustrates the relationship between the different components of the app.

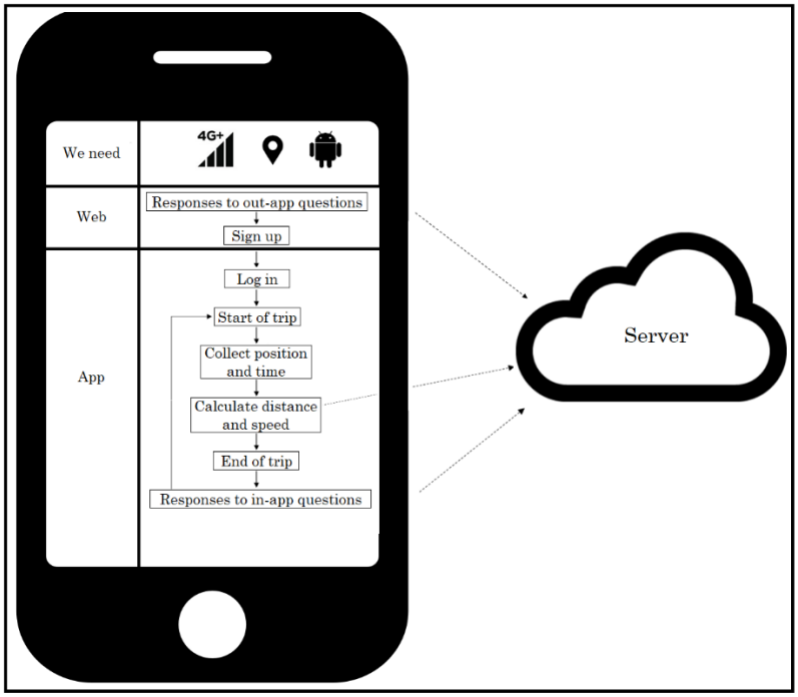
The travel distance is calculated using the Haversine formula. This formula is standard in the literature (e.g., McCool et al. 2021) and is applied to calculate distances on spherical surfaces. The app also calculates speed as the distance traveled in a given time.

Unlike other apps, ours relies only on GPS to estimate the position. Other apps use wifi connections or phone signals to do this. As mentioned above, Santiago de Chile is a large city, with good connectivity overall, so we do not think these settings (which are also invasive) are necessary. We are aware that this may cause, in case of a loss of GPS signal, an increase in the range of the position, which could eventually generate movements that are not happening. However, because GPS works with the average position and Santiago de Chile has a good signal, we believe these problems will be minimal. However, we are going to evaluate this in the testing phase.

4. Pre-collection testing

We interviewed nine people working for delivery companies (UberEats, Rappi and Justo, the largest actors in the market). The testing phase is a mix of a traditional cognitive interviewing and the experience of users handling the app. We are still processing the results of these interviews, so these findings are still preliminary.

Figure 1



The cognitive interview includes questions to inquiry on their work routines and to test whether the items on the web and the in-app questionnaire make sense to them. We use the following questions: Was the question easy to understand? Can you state this question back to me in your own words? Is there a better way to ask the question? What do you think we meant by [TERM] in this item? Was this question clear to you? When you responded to this item, what were you thinking about? Was this question easy or difficult to answer? Was there a better way to ask the question? We receive valuable feedback indicating the rephrasing of some questions. Perhaps the most interesting were the questions about income and sufficiency. The web questionnaire includes an item on the monthly income they earn and the in-app questionnaire includes an item on whether what they made over the day would be sufficient to cover their expenses. The recommendation is to change the timing from monthly to weekly as that is the unit of measurement workers use to measure their income, which fits with the timing of payments from the different apps.

The second component is the user experience. There is a large body of research on ways to test applications before taking them to the field (e.g. Berger and Platzer 2015; Cottrill et al. 2013). An important feature of testing is to measure how familiar people are with digital tools, measured by, for example, how long they have used smartphones for. In our case, the digital skills required to understand and use the application should not be restrictive, given that our target group tend to be digital natives (people under 40), but also and more importantly, they work daily using smartphone applications.

We are also interested in assessing DeliverAndo's usability and stability as part of the daily routines of digital platform workers. Particularly, we are interested in knowing whether they considered the application intuitive to use and found any errors. We assess it with questions such as: "What problems do you think delivery platform workers who participate in the study may have in handling this app?". None of the workers interviewed reported problems using the app. They find it intuitive and easy to use.

However, they did state that the constant use of the application could become tedious, as they usually take between 20 and 25 orders per day. They assume that they must respond to questions at the end of each journey, but that is not the case, when the worker makes more than twelve journeys. That assumption might lead them to under-report journeys and depict a cloudy picture of workers' routines. Nevertheless, if items show up randomly, there is a chance that not all items show up every day, generating item non-response, and thus affecting longitudinal models. We have not figured out a way to solve this issue at the moment of writing.

Also related to the workers' routines is how many days they usually work per week. We want to adapt the project to their working days. However, this line of work gives absolute freedom to choose when to work. They can work in a sporadic or disorganized fashion or adopt routines (fixed working hours) as most people do, which is another reason why it is relevant to study their worker routines using a smartphone-based travel app. The pre-collection testing shed some light on how many days they usually work per week as well. Respondents said that they work between six and seven days a week, depending on whether they meet the level of income they expect to earn net of the costs they have to incur. This is evidence of a significant degree of vulnerability on this line of work and makes the question of working days particularly relevant. Additionally, from these responses, we learned that the study period should be one week, as this is the relevant unit of time for these digital platform workers.

Finally, the testing also involves internal error reporting. Usually, every time the application must be closed due to an error, it must generate a report to the server in order to identify what generated the failure. The beta version we are currently developing does not include this feature. We could only monitor these possible errors during the cognitive interview. The app did not crash while workers were interviewed. Furthermore, workers report that the app was easy to understand and use and did not face any problems during the testing period. Regardless, we reckon that future versions of the app will have to include internal error reporting, since by the nature of back-end, it is impossible for users to be able to reckon what is going wrong in case of errors.

5. Next steps

This document presents our research questions and discusses the app design and the preliminary results of our pre-collection testing. We plan to finish processing the results of the pre-collection testing next week and adjust the design of our project accordingly. After the workshop, we plan to collect all comments, make more adjustments and conduct another round of pre-collection testing focused on the new features and how the app works in a "real-life" situation, i.e. while they are taking orders and traveling across the city to deliver them.

References

Berger, M., and M. Platzer. 2015. "Field Evaluation of the Smartphone-Based Travel Behaviour Data Collection App 'SmartMo'". *Transp. Res. Procedia* 11: 263–79.

Bravo, D. and E. Castillo. 2021. "Estudio Longitudinal Empleo-Covid19: Datos de empleo en tiempo real". Centro de Encuestas y Estudios Longitudinales UC.

Bricka, S.G., J. Zmud, J.Wolf, and J. Freedman. 2009. "Household Travel Surveys with GPS: An Experiment". *Transp Res. Rec.* 2105(1): 51–56.

Carrion, C., F. Pereira, R. Ball, F. Zhao, Y. Kim, K. Nawarathne, N. Zheng, C. Zegras, and M. Ben-Akiva. 2014. "Evaluating FMS: A Preliminary Comparison with a Traditional Travel Survey."

Cottrill, C.D., F. Camara Pereira, F. Zhao, I. Ferreira Dias, H. Beng Lim, M.E. Ben-Akiva, and P.C. Zegras. 2013. "Future Mobility Survey: Experience in Developing a Smartphone-Based Travel Survey in Singapore". *Transp Res. Rec.* 2354(1): 59–67.

Gonzalez, R. and C. Foulon. 2020. "The 2019–2020 Chilean protests: A first look at their causes and participants". *International Journal of Sociology* 50(3): 227-235.

Fielbaum, A. and A. Tirachini. 2020. "The sharing economy and the job market: the case of ride-hailing drivers in Chile". *Transportation* 48: 2235-2261.

Forrest, T.L., and D.F. Pearson. 2005. "Comparison of Trip Determination Methods in Household Travel Surveys Enhanced by a Global Positioning System". *Transp Res. Rec.* 1917(1): 63–71.

Gandhi, A., A. Nizar, Y. Giri, and Y. Ruldeviyani. 2018. "Exploring People's Intention to Become Platform-based Gig Workers : An Empirical Qualitative Study". *International Conference on Information Technology Systems and Innovation (ICITSI)*.

McCool, D., P. Lugtig, O. Mussmann, and B. Schouten. 2021. "An App-Assisted Travel Survey in Official Statistics: Possibilities and Challenges". *Journal of Official Statistics* 37(1): 149-170.

Kaslinwal, R. 2020. "Gender and the Gig Economy: A Qualitative Study of Gig Platforms for Women Workers". ORF Issue Brief *359*.

Kelly, P., P. Krenn, S. Titze, P. Stopher, and C. Foster. 2013. "Quantifying the Difference Between Self-Reported and Global Positioning Systems-Measured Journey Durations: A Systematic Review". *Transp Rev* 33(4): 443–459.

Stopher, P., and L. Shen. 2011. "In-Depth Comparison of Global Positioning System and Diary Records". *Transp Res. Rec.* 2246(1): 32–37.

Wolf, J., M. Oliveira, and M. Thompson. 2003. "The Impact of Trip Underreporting on VMT and Travel Time Estimates: Preliminary Findings from the California Statewide Household Travel Survey GPS Study". *Transp Res. Rec.* 1854: 189–198.