TAXI PROJECT
ASOCIACION DE PRESION

eticas

Observatorio de Trabajo, Algoritmo y Sociedad

MARCH 2023

Adversarial audit of ride-hailing platforms

Algorithmic compliance with competition, labor and consumer law in Spain

CONTENT

INTRODUCTION

As artificial intelligence and algorithmic systems proliferate, the need to understand how they work and impact communities becomes more and more urgent. However, "black box" technologies that lack transparency are not always easy to audit. For this reason, Eticas has been conducting **adversarial algorithmic audits as a tool to independently examine the impact and, to the extent possible, the functioning of algorithmic systems** in order to detect potential anomalies or practices that could be unfair or harmful towards protected groups or society as a whole (Eticas, 2021). Due to restricted access to the algorithms and the databases used to design, develop and validate them, adversarial algorithmic audits rely on analyses of the populations affected, secondary sources, and data scraped via different collection mechanisms.

This adversarial algorithmic audit examines **ride-hailing platforms** which have revolutionized cities' transportation industry worldwide by making it easier and more convenient for people to get around. In Spain, these apps act as mediators between passengers and license holders for private hire vehicles (PHV), but they do not directly exploit such licenses, originally designed for chauffeurs, limousines, official transportation, or pre-booked trips. However, **the growth of three ride-hailing platforms, Uber, Cabify and Bolt, has stretched the boundaries of the PHV regulatory framework** in Spain to accommodate their business models.

The audit was developed by Eticas, the Taxi Project 2.0, an organization aiming to improve the conditions for workers in the taxi industry, and Observatorio TAS, who defends the interests of workers in the platform economy. Organizations who aim to identify when and where algorithmic systems used by ride-hailing platforms in Spain can cause harm with a focus on three main concerns:

1. The **competition** implications of using similar algorithms to set up ride fares, as these algorithms could be harming consumer choice even in the absence of an established cartel to set up prices.
2. The **labor** compliance of ride-hailing apps, and in particular the extent to which app processes incorporate existing labor protections, specifically in relation to absence from work and payment transparency.
3. Potential geographic discrimination in **consumer** prices emerging from the logic of the algorithms used to set prices, which could disproportionately harm less affluent neighborhoods in ways that traditional taxis do not.

In examining these three concerns, we seek to identify how algorithms and AI systems challenge traditional notions of compliance with competition, labor and consumer law, and the extent to which current legal frameworks sufficiently address these new challenges. With this, we aim to uncover the **potentially harmful effects of using algorithms in the platform economy on competitors, workers and users**.

COMPETITION LAW

Competition law encourages companies to ensure that consumers have true choice, as Spain's National Commission for Markets and Competition points out in its guide, "The benefits of competition for consumers" (CNMC). However, competition is not an end in itself, but an instrument at the service of society, as consumers benefit from more affordable, better quality products that are better suited to their needs. It also indirectly benefits businesses and the public sector by supporting economic growth, employment and innovation. On the flipside, restrictions on competition benefit few powerful actors and harm all others (Eticas, 2022). With this in mind, we examined whether and to what extent the new business model pioneered by ride-hailing platforms underscored by the use of pricing algorithms complies with competition law in Spain.

Law 15/2007 for the Defense of Competition (LDC), or the Competition Act, governs competition in Spain. Article 1 of the LDC prohibits collusive conduct, defined as "all agreements, collective decisions or recommendations, or concerted or consciously parallel practices [...], which have as their object, produce or may produce the effect of prevention, restriction or distortion of competition in all or part of the national market". The law prohibits the **direct or indirect fixing of prices** or any other trading or service conditions, and **the application of dissimilar conditions to equivalent transactions**.

Algorithmic price-fixing

Contrary to established, traditional taxis that use a combination of time and distance to calculate prices transparently, ride-hailing apps use "surge pricing" to set fares for similar or equal services. Surge pricing uses complex, **opaque algorithms to adjust fares** based on supply and demand by applying a "surge multiplier" to standard rates. Dynamic or surge pricing is calculated in real time and it is specific to different areas within a city. Uber, Cabify and Bolt all report using surge or dynamic pricing to set their fares.

Recent studies have shown that pricing algorithms tend to systematically collude with one another (Calvano et al., 2020). The use of similar algorithms or self-learning algorithms, which become more precise as they accumulate information, to set fares can lead to price convergence or **price-fixing**. Price-fixing usually refers to an agreement between companies to set the prices for their goods or services at a specific level and it can prevent, restrict or distort competition. With this, we set out to explore whether **price-fixing can occur due to algorithmic processes** even in absence of direct coordination between companies.

In Spain, two allegations of price fixing by PHV platforms have been brought forward to the CNMC in 2018 and 2019 respectively. After reviewing the complaints together, the CNMC determined that Uber and Cabify set their prices independently and differently and found no indication of anti-competitive conduct (CNMC, 2020). However, due to the insufficient attention dedicated to the case according to a dissenting opinion by a CNMC Board member (Vila and Rivas, 2020) and recent allegations of price-fixing abroad

(Towards Justice, 2022), the issue of possible price collusion by ride-hailing platforms merits further investigation.

Data collection

The first part of this report, developed by the Taxi Project 2.0, explores how the algorithms of ride-hailing platforms affect competition. In particular, we investigate whether the algorithms of the three main ride-hailing platforms in Spain are fixing prices.

In order to examine this, we collected price data from Uber, Cabify and Bolt using a combination of sock-puppet and scraping methods. First, we selected 8 routes in Madrid and 7 routes in Andalusia and sent automated trip requests for each route every 10 minutes between 11 October 2021 and 11 January 2022. This process is referred to as sock-puppet method or programmatically-constructed traffic to a platform which allows auditors to collect a large amount of system outputs (in this case, trip fares). We then scraped the fares for each route in the three ride-hailing platforms. The resulting average prices are reported below (Table 1).

			Uber		Cabify		Bolt	
	km	Min.	avg. price	std. dev.	avg. price	std. dev.	avg. price	std. dev.
Paseo de las Acacias, Madrid - Hospital Quirón Salud, Pozuelo de Alarcón	10,9	20	12,75	2,62	11,31	1,07	16,49	6,81
Atocha - Paseo de la Castellana, 259, Madrid	13,8	19	14,00	3,96	13,96	1,28	21,00	8,89
Atocha - Calle Orense, 6, Madrid	7,7	20	9,29	2,51	8,34	1,09	12,20	5,45
Atocha - Calle Serrano, Madrid	5,3	14	6,11	1,90	5,20	0,92	7,46	3,46
Calle Velázquez - Paseo de la Castellana, 81, Madrid	3,6	12	5,69	1,55	4,47	0,81	6,55	2,66
Aeropuerto de Barajas T4 - Avenida Bruselas, Madrid	8,9	10	14,80	1,79	15,01	0,24	10,95	1,60
Aeropuerto de Barajas T4 - Calle María de Molina, Madrid	13,8	20	19,09	3,81	16,01	1,45	17,84	2,75
Aeropuerto de Barajas T4 - Plaza Castilla, Madrid	13,3	14	17,44	3,80	15,34	0,93	15,47	2,25
Aeropuerto de Málaga a Puerto Banús	57,8	44	84,49	8,87	65,29	4,59	55,90	18,61
Aeropuerto de Málaga - Málaga	8,4	10	15,21	1,70	10,15	1,04	8,69	2,91
Aeropuerto de Málaga - Marbella	53,5	37	78,98	8,29	60,22	4,31	52,01	17,31
Aeropuerto de Málaga - Nerja	68,3	48	92,81	10,38	77,09	5,62	67,80	2,16
Estación de Autobús de Marbella - Puerto Banús	8,5	10	10,37	1,49	11,20	0,87	9,59	1,54
Hotel Marriott Marbella Palacio - Hipercor Puerto Banús	20,8	20	25,72	3,40	25,45	1,08	19,51	6,53
Bulevar San Pedro de Alcántara - Hotel Puente Romano, Marbella	7	14	8,68	1,20	9,48	0,87	7,84	0,84

Table 1. Trip length (km and minutes), average price and standard deviation for each service provider and trip

Following this, we set out to explore whether there were any instances of algorithmic price-fixing in these routes. We did this by studying the correlation between the fares of the three service providers through linear regression analysis for each route. In statistics, linear regression is a mathematical model used to approximate the dependency relationship between a dependent variable and one or more independent variables, also adding a random term. This method is applicable in many situations where the relationship between two or more variables is examined. The analysis allows us to obtain the Pearson correlation coefficient, which is a measure between -1 and +1 that indicates the positive or negative linear dependence between two quantitative random variables (Table 2), and the R squared (R^2) or coefficient of determination, which tells us what percentage of the variance of a dependent variable is explained by the movements of the independent variable, usually expressed with a number from 0 to 1.

The Pearson correlation coefficient is interpreted as follows:

Strength of association	**Positive**	**Negative**
Strong	0.5 to 1.0	-0.5 to -1.0
Moderate	0.3 to 0.5	-0.3 to -0.5
Weak	0.1 to 0.3	-0.1 to -0.3
None	0	0

Table 2. Interpretation of the Pearson correlation coefficient (r)

Findings

Based on these observations, we found a moderate positive and statistically significant correlation of prices for all monitored routes in Andalusia between Uber and Cabify, and a strong positive and statistically significant correlation of prices for 5 out of 8 trips in Madrid between Uber and Bolt (Table 3). The instances of strong and moderate correlation of prices are highlighted in green in the table below.

	Uber-Cabify			**Uber-Bolt**			**Cabify-Bolt**		
	r	**R^2**	**ø**	**r**	**R^2**	**ø**	**r**	**R^2**	**ø**
Paseo de las Acacias, Madrid, - Hospital Quirón Salud, Pozuelo de Alarcón	0,26	0,07	2,53	**0,59**	0,34	5,51	0,26	0,07	6,57
Atocha - Paseo de la Castellana, 259, Madrid	0,19	0,04	1,26	**0,57**	0,33	7,30	0,26	0,07	8,60
Atocha - Calle Orense, 6, Madrid	0,36	0,13	2,34	**0,65**	0,42	1,91	0,36	0,13	5,09
Atocha - Calle Serrano, Madrid	**0,44**	0,19	1,70	**0,66**	0,44	1,42	**0,44**	0,20	3,10
Calle Velázquez - Paseo de la Castellana, 81, Madrid	0,25	0,06	1,50	**0,56**	0,31	1,29	0,28	0,08	2,56

Aeropuerto de Barajas T4 - Avenida Bruselas, Madrid	0,11	0,01	0,24	0,26	0,07	1,54	0,04	0,00	1,60
Aeropuerto de Barajas T4 - Calle María de Molina, Madrid	0,14	0,02	3,77	0,30	0,09	2,62	0,15	0,02	2,72
Aeropuerto de Barajas T4 - Plaza Castilla, Madrid	0,12	0,01	0,93	0,28	0,08	2,16	0,15	0,02	0,92
Aeropuerto de Málaga - Puerto Banús	**0,41**	0,17	4,18	0,03	0,00	8,96	0,16	0,03	4,70
Aeropuerto de Málaga - Málaga	**0,46**	0,21	0,92	0,08	0,01	1,72	0,12	0,01	1,07
Aeropuerto de Málaga - Marbella	**0,42**	0,18	7,51	0,07	0,01	8,27	0,14	0,02	4,43
Aeropuerto de Málaga - Nerja	**0,42**	0,17	9,44	0,02	0,00	10,38	0,15	0,02	2,13
Estación de Autobús de Marbella - Puerto Banús	**0,42**	0,17	0,79	0,16	0,03	1,52	0,14	0,02	0,86
Hotel Marriott Marbella Palacio - Hipercor Puerto Banús	**0,40**	0,16	3,11	0,03	0,00	3,25	0,01	0,00	1,09
Bulevar San Pedro de Alcántara - Hotel Puente Romano, Marbella	**0,41**	0,17	1,09	0,11	0,01	1,19	0,12	0,01	0,86

Table 3. Pearson correlation coefficient, coefficient of determination and standard deviation for the binomials Uber-Cabify, Uber-Bolt, and Cabify-Bolt.

Our findings suggest that **the pricing algorithms of Uber, Cabify and Bolt are colluding in some of the most important routes in Andalusia and Madrid**. In particular, we observe a strong and statistically significant positive correlation between the fares of Uber and Cabify on the one hand, and Uber and Bolt on the other, particularly for routes in Madrid. Conversely, we observe weaker correlations between Cabify and Bolt prices. This leads us to think that the high correlations between the prices of Uber and the other two ride-hailing apps are not a spurious correlation, but are probably due to coordination of the algorithms of Bolt and Cabify with the price system of Uber.

This, in turn, is a possible violation of Law 15/2007 for the Defense of Competition (LDC) in Spain which prohibits direct and indirect price collusion. Even if there was no explicit agreement between ride-hailing companies, there is scope to suggest **indirect price-fixing by algorithmic means**. The indirect price coordination via pricing algorithms creates an uneven playing field and **harms other actors in the market**, such as traditional taxis and potential new entrants.

LABOR LAW

Labor rights are the mark of modern society according to the Spanish government (Law 12/2021). However, the surge of the platform economy and ride-hailing apps pose a challenge for workers' rights. Algorithms in particular have revolutionized the way services are provided, leading to gains in efficiency and productivity. However, algorithms increasingly mediate the relationships between employers, workers and users in opaque ways which can disadvantage the contractually weaker party and potentially harm workers.

Legal precedents outside of Spain have highlighted the ways in which algorithms can be harmful for platform workers. In 2021, a case in the Court of Bologna found that Deliveroo's algorithm penalized delivery workers for absences from work shifts for legally protected reasons, thereby limiting workers' opportunities to secure job assignments in the future (Lomas, 2021). The Labour Court of Bologna ruled that Deliveroo engaged in discrimination against workers in cases of absences due to legally protected reasons such as illness, the need to care for a minor or a disabled person and the right to strike (Tribunale Ordinario di Bologna, 2019). Importantly, the decision also confirmed that algorithms are subject to judicial review in cases of non-compliance with existing labor protections.

Another area of concern for labor rights in the algorithm-driven platform economy is the lack of payment transparency. This issue became subject to debate in court as early as 2014 when Uber riders launched a class action lawsuit in San Francisco. The plaintiffs alleged that the platform misled customers to believe that drivers receive 100% of tips, while in reality Uber kept approximately 40% of gratuity payments (Christophi, 2017). Despite the settlement, the lack of payment transparency persists in the platform economy, as delivery apps such as Instacart in the United States continue to attract negative attention for their opaque payment structures with inconsistent and unreliable commission rates for job assignments (Kerr, 2021). More recently, Washington D.C.'s Attorney General filed a lawsuit against Amazon for keeping a portion of the tips paid to FLEX drivers while customers were misinformed that the delivery workers would receive 100% of the gratuity (DeGeurin, 2022).

The cases in other countries above demonstrate that **opaque algorithms in the platform economy may not be compliant with existing legislation on labor rights**. With these concerns in mind, we set out to explore how the use of algorithms by ride-hailing apps affects compliance with existing labor laws in Spain and what new challenges it poses to workers' rights.

Algorithmic transparency

In Spain, the legal framework which governs ride-hailing platforms offers **little protection to workers**. Uber, Cabify and Bolt operate as intermediaries through PHV licenses owned by fleet companies and exploited by hired drivers. The operation of PHV on urban routes is not regulated on the national level, while autonomous communities (comunidades autónomas, autonomous political and administrative regions in Spain) provide little meaningful labor safeguards to drivers.[1] The drivers, who are classified as employees, earn a fixed income and incentives according to their performance, often determined by opaque algorithms. On the other hand, self-employed drivers that work directly for the platforms are a minority in Spain, and they are not classified as employees but as independent contractors, despite suffering similar control from the app as fleet employees. These drivers receive all their pay as a proportion of their performance.

In recognition of the increasing influence of algorithms on labor relations, Spain introduced an amendment to the Workers' Statute Law (Royal Legislative Decree 2/2015) known as **Ley Rider (Rider's Law) guaranteeing the right of workers to algorithmic transparency** i.e. "to be informed by the company of the parameters, rules and instructions on which the algorithms of artificial intelligence systems are based, that affect decisions that may affect working conditions, access to and maintenance of employment, including profiling" (Law 12/2021).[2]

The right to algorithmic transparency, hailed as a ground-breaking accomplishment, was designed to enable **"the neutralization of algorithmic punishments, penalties for performance and bias"** according to Yolanda Díaz, Minister for Employment and the Social Economy (Aranguiz, 2021). This section of the report examines the extent to which this goal was accomplished with view to algorithmic punishments for absences from work and payment transparency.

Data collection

To examine the challenges to labor rights posed by algorithmic decision-making in the platform economy, we conducted in-depth interviews with PHV drivers for Uber and Cabify, and a PHV fleet manager for a PHV fleet company. Our interviews revealed that the operation of ride-hailing platforms as intermediaries between PHV fleet and passengers in Spain creates **an opaque decision-making structure** involving both algorithms and human agents. This structure determines, among other issues, the allocation of shifts and trips for drivers and their payment, and as such, it has important implications for labor rights.

Based on our conversations with PHV drivers, we identified at least two levels to the decision-making structure with partial overlap and without a clear focal point of accountability:

- At the first level, ride-hailing apps' algorithms connect vehicles with passengers, process payments, calculate drivers' scores, determine whether a driver is allowed

[1] In the Community of Madrid, for example, the collective agreement between Aseval and Unauto, the employers' associations in the PHV sector, and the main PHV driver unions, UGT, CC OO and SLT (Aseval, 2022) allows two consecutive days of rest, but workers can voluntarily waive this right. Similarly, the agreement denotes four to six unjustified service cancellations in a month as a "serious offense" for drivers, but it fails to specify what constitutes justified and unjustified rejection (Consejeria de Economia, Hacienda y Empleo, 2022).

[2] The second provision of Ley Rider also recognized workers in the platform economy as employees rather than contractors under certain conditions, but this change did not apply to PHV drivers.

to log into the platform, and send warnings to human agents from the PHV license fleet companies among others.

- At the second level, PHV license fleet companies set the general rules and frameworks for operation, such as when and how long drivers work and whether and how they can receive tips, while fleet managers enforce the rules and carry out the organization of the work, including assigning work slots and determining punishments.

Findings

Absence from work

Our interviews revealed that **ride-hailing platforms in Spain do not adequately accommodate lawful reasons for absence from work** and do not provide sufficient transparency regarding their processes. PHV drivers consistently reported feeling pressured to work more and longer shifts despite legal provisions for rest during work hours and days off. For example, one driver expressed concern that workers cannot decline “a single minute of the assigned working hours” as this can result in sanctions, pay cuts and even dismissal. A PHV fleet manager similarly shared that if workers fail to meet the minimum requirements for earnings from completed trips, they may be assigned to worse cars or less lucrative areas. This can create a “vicious cycle” and further constraining the ability to achieve targets and secure profitable job assignments.

The drivers explained that these punishments can be determined and enforced by the managers in PHV license fleet companies. Screenshots of the Cabify application for PHV fleet managers obtained by Eticas corroborate this, as the app allows managers to deactivate drivers’ profiles from the platform.

However, we found that **algorithms can sanction drivers** too. Excessive or unjustified ride cancellations during an ongoing shift can result in severe penalties. One driver noted that, in cases of excessive cancellations, passenger complaints and low customer satisfaction scores, an algorithm can lock out workers from connecting to the app for a period of time (for example, a day) or indefinitely, thereby directly limiting drivers’ opportunity to work. This phenomenon is known as ‘robo-firing’, where algorithms can produce automated decisions to dismiss workers from employment, often without transparency regarding the reasons (Worker Info Exchange, 2021).

Despite the potentially severe sanctions, decision-making algorithms in ride-hailing apps do not specify what constitutes excessive and unjustified ride cancellation. In the Uber app, for example, drivers have a list of options to select from as a reason for declining the ride (Figure 1). The app does not specify whether all options, such as “I have accepted the trip by mistake” or “I have taken the wrong way”, are considered justified reasons.

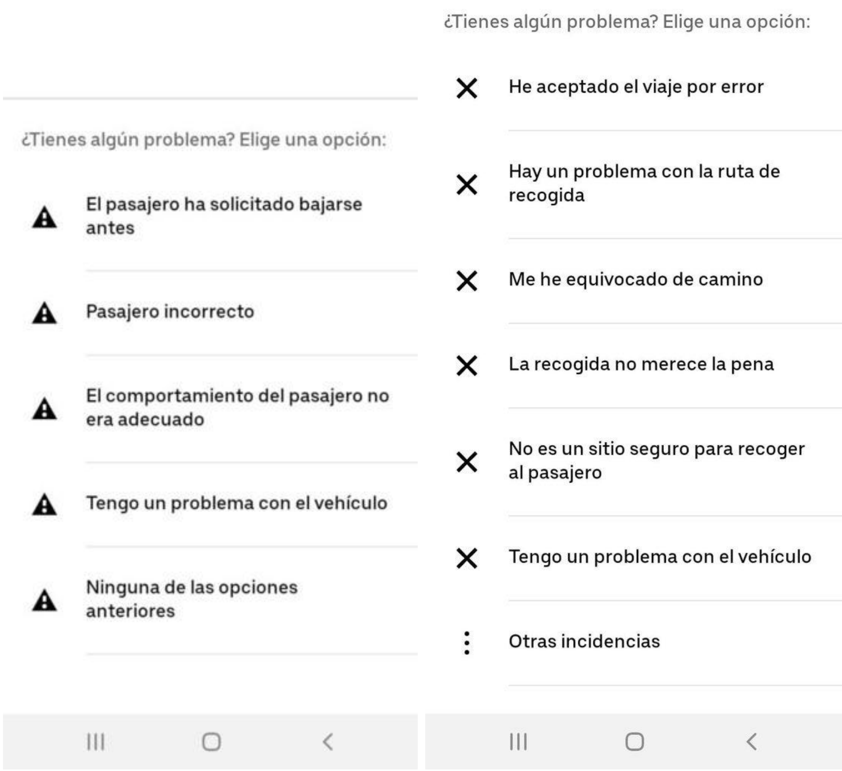

Figure 1. Canceling a trip in the Uber app
Screenshots provided by PHV drivers

The issue of unjustified ride cancellations as grounds for punishment raises further doubts in the case of the Cabify app, which does not collect information about the reason for declining a trip (Figure 2). To exacerbate this problem, Cabify only allows a maximum of two cancellations in 24 hours, as shared by a PHV driver.

Figure 2. Canceling a trip in the Cabify app
Screenshot provided by PHV drivers

It remains unclear whether and how ride cancellations or absences from work affect drivers' ability to work, even for legally protected reasons. Our fieldwork study found evidence that an internal ranking score for drivers appears to be connected to a monthly reward system, but it is not specified whether the same ranking is also used to determine sanctions.

Eticas received screenshots of Cabify's application for fleet management that indicate an internal ranking score for employees ("DO") that differs from the drivers' public score based on customer feedback that is displayed on the user app. The internal ranking system, which appears to be algorithmically-determined, assigns a score of 0 to drivers who are currently offline and have not earned any profits for the day. Drivers who have been online for about the same period of time and earned similar amounts, on the other hand, are granted different ranking scores for unspecified reasons. This suggests that the ranking algorithm takes into account factors other than the number of hours worked, trips made, and amount of money earned. The monthly frequency of performance evaluations, on the other hand, may indicate that, beyond ride cancellations, absences from work shifts for both justified and unjustified reasons also factor in this score.

Overall, we found that ride-hailing apps and PHV companies do not provide sufficient transparency on procedures related to absence from work. The extent to which workers are protected in cases of absences for lawful remains unclear, while there are strong disincentives for any absences or ride cancellations. This not only limits drivers' future opportunities for assignments and earning potential, but it also raises **concerns about the platforms' compliance with labor rights**.

Payment transparency

Our interviews revealed that **ride-hailing apps in Spain lack transparency in their payment structures**, especially in the case of performance incentives and tips. While each platform and PHV operator has different rules regarding tips, Uber and Cabify drivers generally report difficulties with receiving gratuity for several reasons. Some PHV operators and ride-hailing platforms, such as Cabify, forbid drivers from receiving tips in cash. This rule not only limits the opportunities to receive gratuity, but breaking it can also result in sanctions for drivers. In cases when users tip through the app, platforms provide little information regarding when and what proportion of tips is paid out to workers.

During our interviews, PHV drivers were generally **skeptical about receiving tips from ride-hailing platforms**. One driver remarked that they are still waiting to receive their tips from Uber, while another noted that they have yet to receive any tips from Cabify. This perception may be due to a lack of transparency in the way payments are processed. A PHV fleet manager explained in an interview that tips made through ride-hailing apps are usually added directly to payroll along with the fixed wage and other monetary rewards for the workers, making the exact amount received in tips unclear. However, the fleet manager also shared that some PHV companies do not distribute tips to drivers at all.

The screenshots of Cabify's application for fleet managers (received separately by Eticas) reveal that the tip amount for each ride received through the user app is clearly visible. However, the payment slips issued to workers only contain a single category without a breakdown of different income streams, such as wages earned, performance bonuses, and tips. At best, this could point to the lack of transparency in the payment structures of ride-hailing platforms and PHV companies. At worst, however, it can imply that workers are not fairly compensated for their work.

In the cases of both absence from work and payment transparency, we reached concerning conclusions about ride-hailing platforms' protection of labor rights. On the one hand, our findings raise **doubts about Uber's and Cabify's compliance with existing labor safeguards** such as lawful absences from work. On the other hand, our interviews clearly demonstrate that **algorithmic punishments and opaque penalties for performance persist in the mobility sector**, despite the promise of Ley Rider's guarantee for greater transparency. Beyond legal implications, our findings have significant social repercussions, as the PHV sector employs over 20.000 in the Community of Madrid alone, usually members of vulnerable groups with little bargaining power, such as older unemployed people who have difficulties returning to the job market and migrants (Cortés, 2022).

CONSUMER LAW

Ride-hailing apps use surge pricing algorithms to determine ride fares based on supply and demand in a given area and time. For example, higher demand for transportation services during rush hour may cause trip prices to go up in busy areas. Conversely, the low supply of cars in remote and less busy areas may drive fares up. This means that geographic price discrimination occurs, where platforms charge different rates for the same service in different locations.

With regards to consumer protection, the General Consumer and User Protection Act in Spain prohibits any form of discrimination based on place of residence. While the Act allows for "differences in access conditions directly justified by objective criteria" (Article 49), the opaque nature of the pricing algorithms used by ride-hailing platforms precludes any formal assessment of objectivity. The final part of this report attempts to reverse-engineer how ride-hailing apps factor geographic location in their pricing algorithms through an exploratory study of geographic price discrimination.

Geographic price discrimination

While geographic price discrimination has important legal implications, differences in prices based on location can also signal socio-economic discrimination. Previous studies have revealed that surge pricing algorithms can discriminate neighborhoods based not only on geographic location, but also on demographic makeup due to variance in supply and demand in areas with different population characteristics. Pandey and Caliskan found that neighborhoods with large non-white populations, higher poverty levels, younger residents and high education levels are associated with higher fares on ride-hailing apps (2021). Similarly, Chang et al. show that Uber charges higher prices for trips to more expensive hotels (2021). To probe this issue further, we set out to investigate whether ride-hailing apps discriminate based on the socio-economic characteristics of neighborhoods by algorithmic means.

Data collection

Initially, we planned to conduct this part of the audit through crowdsourcing data directly from users and test whether the fares for completed trips (comparable in terms of length of trip in minutes and kilometers, time of day, etc.) correlate with the socio-economic profiles of different neighborhoods in which the trips took place. We launched our crowdsourcing campaign in June 2022 to collect trip data from Uber users in Madrid, in which we asked riders to obtain information about their trips from the platform[3] and send it to Eticas anonymously for analysis. This required users to request to download their data from Uber and wait for "a few days", as per Uber's policies, to receive access. Due to

[3] The data requested include product type (UberX, Comfort, etc.), trip origin and end point, time, distance, and fare amount.

the difficulties with this process, however, we did not gather sufficient data for our analysis.

For this reason, we opted to conduct a limited, exploratory study of socio-economic price discrimination by collecting trip fares for 20 routes using the Uber, Cabify and Bolt apps. We selected routes in neighborhoods with different median incomes as an indicator of the socio-economic makeup of the area based on 2018 data from the Spanish Statistical Office (INE) reported in El País (Andrino et al., 2021).[4]

We identified four low-income neighborhoods, two medium-income neighborhoods, and four high-income neighborhoods in Madrid and Malaga (Table 4). The low-income neighborhoods in both Madrid and Malaga represent the bottom 1% in their respective autonomous communities, the Community of Madrid and Andalusia. The selected medium-income neighborhoods sit at the top 24-41%, while high-income neighborhoods are in the top 1-6% of their communities. We also considered relative proximity to one another and relative distance from the city center in our selection of neighborhoods in order to partially control for the higher supply and demand of transportation services in busy areas.

	Low-income Neighborhoods	**Medium-income neighborhoods**	**High-income neighborhoods**
Madrid	5.250-6.650 EUR	19.950 EUR	33.600-36.750 EUR
Málaga	5.250-6.650 EUR	14.350 EUR	25.550-28.350 EUR

Table 4. Median income in EUR for selected neighborhoods

We then selected two routes in each neighborhood (a total of 20 routes), with approximate lengths of 2 km and 4 km respectively (Table 5). We collected the fares for each route from the standard services of the three main ride-hailing platforms in Spain, Uber, Cabify and Bolt.

City	**District**	**Median income (EUR)**	**Start point**	**End point**	**Length (km)**	**UberX fare (EUR)**	**Cabify fare (EUR)**	**Bolt fare (EUR)**
Madrid	Villa-Vallecas	5250	Carr. Cañada Real, 90	Cam. Leña, 10	1.9	5.09	5.5	-
Madrid	Villa-Vallecas	5250	Av. Mediterráneo, 127	Ctra. Vertedero Municipal Valdemingómez, 155	3.8	5.77	5.5	5
Madrid	Vicálvaro	6650	C. de Boyer, 2	C. Dehesa Vieja, 8	2	7.15	5.54	5.75
Madrid	Vicálvaro	6650	Blvr. de José Prat, 29	Carr. de Vicálvaro a la Estación de O'donnell, 19	3.9	7.21	5.8	5.85

[4] Income refers to the net income per person, calculated by taking the entire household income and dividing it by the number of people living in it, where the first adult counts as 1, the second as 0.5 and the minors as 0.3.

Madrid	Vicálvaro	19950	Av. de Miguel Delibes, 30	C. Vereda de la Cebolla	2	5.25	5.5	5
Madrid	Vicálvaro	19950	C. de Cerceda, 20	C. Pilar Bellosillo, 12	3.9	7.25	5.5	5.5
Madrid	Fuencarral-Pardo	33950	C. de Frómista, 1	C. de Cebreiro, 2	2.1	5.27	5.5	5.3
Madrid	Fuencarral-Pardo	33600	Distrito Telefónica Edificio Norte 1	C. de Navarrete, 9	3.9	7.26	5.5	5.2
Madrid	Barajas	36750	Av. de Logroño, 179	Parque Juan Carlos I	2.1	5.69	4.5	5
Madrid	Barajas	36750	Av. de Logroño, 179	Feria de Madrid	3.9	7.23	5.5	5.8
Málaga	Ronda I.-Campan	5250	CEIP María de la O	Lugar Cuidad de los Niños 1	1.9	5.59	5.5	6
Málaga	Ronda I.-Campan	5250	Calle pedagoga María Montessori N, 8	Arquitecto Francisco Peñalosa, 18	4.1	5.08	5.72	6
Málaga	Malaga-Norte	6650	C. Alcalde José Luis Estrada	Cam. de Los Alcabuceros, 6	2	7.26	12.48	6
Málaga	Malaga-Norte	6650	Finca La Pola	C. Ana Sólo de Zaldívar	3.9	7.18	12.48	18.6
Málaga	Ronda I.-Campan	14350	Av. de las Malagueñas	C. la Orotava, 38	1.9	5	5.72	6
Málaga	Ronda I.-Campan	14350	C. José María Jacquard, 18	Av. de José Ortega y Gasset, 201	4.2	7.12	5.72	6.6
Málaga	Málaga-Norte	25550	C. Bogor, 4	C. Julio Verne, 6	2	5	5.72	6
Málaga	Málaga-Norte	25550	C. Julio Verne, 6	C. Trombón, 22	4.1	5	5.5	6
Málaga	Málaga-Este	28350	C. de la Minilla, 3	C. Monte Miramar, 38	2.1	7.59	5.77	6
Málaga	Málaga-Este	28350	Camino de los Almendrales	C. las Espuelas, 12	4	6.41	5.57	6

Table 5. Price and length of selected trips for four Uber services

Findings

To explore correlations between trip fares and median income for indications of socio-economic price discrimination, we calculated the price per kilometer for each trip and conducted a linear regression analysis. Below, the Pearson correlation coefficient (r) indicates the linear dependency (either positive or negative) between two variables (Table 6), the coefficient of determination (R^2) indicates the proportion of the variation in the dependent variable that is predictable from the independent variable, and the P-value denotes the level of significance.

Strength of association	Positive	Negative
Strong	0.5 to 1.0	-0.5 to -1.0
Moderate	0.3 to 0.5	-0.3 to -0.5
Weak	0.1 to 0.3	-0.1 to -0.3
None	0	0

Table 6. Interpretation of the Pearson correlation coefficient (r)

Our findings demonstrate that surge pricing changes according to median neighborhood income for all three ride-hailing platforms (Table 7). In particular, we observe a weak to moderate, statistically significant negative correlation between price per kilometer and median income for Uber, Cabify and Bolt. In other words, there are indications that **prices in ride-hailing apps tend to be lower in more affluent neighborhoods**. While this is the case for all three ride-hailing platforms, the correlation between trip fares and median neighborhood income is stronger in the case of Cabify compared to Uber and Bolt.

	Uber			Cabify			Bolt		
	r	R^2	P	r	R^2	P	r	R^2	P
All routes (n=20)	-0.11	0.01	0.009	-0.37	0.14	0.001	-0.27	0.08	0.002

Table 7. Price per kilometer and median income: Pearson correlation coefficient (r), coefficient of determination (R^2) and P-value

At the same time, the low R^2 values in Table 7 suggest that median neighborhood income explains little of the variance in the trip fares. In other words, surge pricing takes into account multiple variables, which include the supply of cars, user demand and the length of travel according to the terms of ride-hailing apps. Nonetheless, there is a scope to infer a relationship between the socio-economic characteristics of neighborhoods and trip fares.

While this exploratory analysis relies on a small sample, it nevertheless raises **concerns regarding algorithmic price discrimination on the basis of the geographic location and socio-economic makeup of neighborhoods**. In particular, we have observed that under certain circumstances, customers in lower income neighborhoods tend to pay higher trip fares in ride-hailing apps, making mobility services less accessible to disadvantaged groups.

This, in turn, calls into question ride-hailing platforms' compliance with the General Consumer and User Protection Act in Spain regarding discrimination based on place of residence. The opaque nature of the pricing algorithms used by Uber, Cabify and Bolt make it difficult to determine whether they rely on objective criteria for the differences in access conditions, but our preliminary analysis suggests that this issue merits further investigation from a both legal and social perspective.

CONCLUSION

This report outlined the findings of Eticas', the Taxi Project's and Observatorio TAS' adversarial audit of ride-hailing platforms in Spain regarding their compliance with competition, labor, and consumer law. The main conclusions of the audit are as follows:

- The pricing algorithms of Uber, Cabify and Bolt appear to collude in some of the most important routes in Andalusia and Madrid, which suggests indirect price-fixing by algorithmic means in breach of the Law for the Defense of Competition.
- The use of algorithms in ride-hailing platforms to mediate labor relations lacks transparency in payment and profiling, and it can lead to discrimination against platform workers for absences due to legally protected reasons.
- Uber's pricing algorithm can discriminate based on the socio-economic characteristics of neighborhoods, making mobility services less accessible in low-income neighborhoods, which may constitute an infringement of the General Consumer and User Protection Act.

While our findings raise **doubts about ride-hailing platforms' compliance** with applicable legislation in the areas of competition, labor and consumer law**, the lack of transparency in the algorithms used by mobility service providers is a persistent issue across these areas** despite recent legal advances in favor of algorithmic transparency.

Given the concerning results of this audit, we make the following **recommendations** on a national level in Spain:

- The **CNMC must investigate the issue of indirect price-fixing** by ride-hailing platforms based on the evidence of this report and further inquiry.
- Self-employed PHV drivers must be included in the employment provisions of Ley Rider and **recognized as employees**, rather than contractor workers.
- More **robust mechanisms for enforcement of Ley Rider's provision for algorithmic transparency** must be put in place, including disclosure of information on algorithmic processes and transparent communication about worker profiling, performance assessment and payment structures.
- Authorities must explore the differences in access to services of ride-hailing platforms based on geographical location and socio-economic characteristics presented in this study.

On a European Union level, we endorse **the Workers' Recommendations on the Draft EU Platform Work Directive** on fair and transparent algorithmic management, developed by the Worker Info Exchange and Observatorio del Trabajo, Algoritmo y Sociedad, including:

- **Disclosure of information on automated decision-making** and explanations of the impact of these systems on workers;
- Disclosure of performance factors, and **prohibition of robo-firing of workers**;

- **Prohibition of predictive behavioral profiling technologies** that affect working conditions;
- **Prohibition of dynamic pay**;
- **Transparency and due process** in termination of work and performance management;
- **Transparent communication and social dialogue** about the impact of automated decision-making systems, including for risk and safety management;
- **Right to access** all personal data and an explanation of how platforms have processed workers' data at work.

Annex: Uber's Data Management[5]

For any data controller to be able to carry out Article 6 of the European General Data Protection Regulation (**GDPR**), there are **six provisions** listed under which **data can be lawfully processed.** That is, if:

(a) the data subject has **given consent** to the processing of his or her personal data for one or more specific purposes;
(b) processing is necessary for the **performance of a contract** to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract;
(c) processing is necessary for **compliance with a legal obligation** to which the controller is subject;
(d) processing is necessary in order to **protect the vital interests of the data subject** or of another natural person;
(e) processing is necessary for the **performance of a task carried out in the public interest** or in the exercise of official authority vested in the controller;
(f) processing is necessary for the purposes of the **legitimate interests** pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child.

For any data controller to be able to carry out the processing of personal data, it is essential that such processing is duly justified by one or more of the legal bases provided for in the article above. In this sense, the basis of legitimate interest is one of the most commonly used but is subject to much interpretation by the supervisory authority – being also one of the most controversial aspects regulated in GDPR.

In this regard, legitimate interest is one of the grounds on which the processing of personal data does not require the express consent of the data subject. However, neither the GDPR nor the Spanish regulation (i.e., the Organic Law 3/2018 on Data Protection and Guarantee of Digital Rights, hereinafter, the **"LOPDGDD"**), establishes a clear definition of the concept of legitimate interest.

Nevertheless, the LOPDGDD does include several cases in which the processing of personal data may be carried out based on legitimate interest; specifically, these refer to the processing of **contact data**, **individual entrepreneurs**, and **professionals**. In the same vein, the GDPR indicates that **it is necessary to weigh the legitimate interests of those who will process the data against the interests and fundamental rights of the data subject** based on three standards: (*i*) appropriateness of the processing, (*ii*) necessity; and (*iii*) proportionality. Therefore, in the analysis to be carried out by the controller regarding the weighting of the legitimate interest as a legal basis for the processing, the following factors must be considered:

[5] This analysis will be included in the second report.

(*i*) the **nature and source of the legitimate interest**

(*ii*) the **impact on the data subjects**, such as the nature of the data (whether or not it affects data considered as sensitive), the manner in which the data is processed, the reasonable expectations of the data subject, the position of the controller and the data subject (i.e. the balance of power between both)

(*iii*) additional safeguards such as **data minimization and extensive use of anonymization techniques**, data aggregation, privacy enhancing technologies, privacy by design and data protection impact assessments

In order to be able to base a processing operation on legitimate interest, the test of weighing of rights must be carried out based on the criteria of suitability, necessity and proportionality in order to determine whether there are less intrusive measures for the rights and freedoms of data subjects. Additionally, it is important to mention that the processing of data by a data controller may be covered by different legal bases depending on the purposes of such processing.

Moreover, in the Opinion 06/2014 on the concept of legitimate interest of the data controller issued by the Article 29 Working Group, it is stated how the use of legitimate interest as a justification for data processing "*should also not be perceived as a preferential option nor should its use be unduly extended because it is considered less restrictive than the other legal grounds. The legitimate interest must also be expressed with sufficient clarity and must be sufficiently specific to allow the balancing test to be conducted against the interests and fundamental rights of the data subject. It must also represent a real and present interest, i.e., it must not be speculative*."

Last, according to this same guideline, we must bear in mind that the interest at stake must also be "pursued by the data controller, requiring a real and current interest that corresponds to a current interest or one that is expected in the not too near future. In other words, and as quoted in the same text, "*interests that are too vague or speculative will not suffice*." Therefore, a "legitimate interest" will be understood as one that is **lawful under applicable EU and national law, that is specific in its articulation and that represents a real and present (i.e., not speculative) interest.**

Minimization

Article 5 in GDPR addresses the general principles relating to the processing of personal data. The third of the listed principles refers to the concept of data minimisation, indicating how any data processed needs to be “adequate, relevant, and limited to what is necessary in relation to the purposes for which [the data] are processed”.

With this, it is possible to infer that data minimisation encompasses several requirements, such as:

- Only data that will be processed can be collected, that is, those that are strictly necessary for the stated purposes

· Data can only be collected for an intended purpose, thus, data cannot be collected and stored for the purpose of being used later

· Personal information can only be used for the purpose for which data were collected, but no other objective than that

Data Flows in Uber

In order to see to what extent Uber is sharing the data they collect from their users, during the course of this audit a Man In The Middle Attack was conducted to read the communication between the users' application and Uber's servers. In short, a MITM attack is a strategy to intercept the communication between two parties, relaying the information exchanged between such parties. In what follows, the results of this MITM attack are discussed in detail.

Dynamic and Static Analysis

Uber collects personal data from three different sources:[6]

- **Provided by users** to Uber, such as during account creation
- **Created during use of their services**, such as location, app usage and device data
- From **other sources**, such as other users or **account** owners, **business partners**, **vendors**, insurance and financial solution providers, and governmental authorities.

This analysis, however, focuses on the **data created during the use of their services**. To do so, the analysis was carried out with a OnePlus6T device updated with Android version 9.0, corresponding to the last update to date. Moreover, and to know which data is collected from the user, the registered permissions inside the application were analyzed. In this regard, there are two types of permissions: **normal** and **sensitive**.

Sensitive Permissions

Depending on the version of the android system, sensitive permissions can be required during the execution of the application. However, previous versions ask the user to grant permission when the Uber app is first installed. In Android version 9.0, **sensitive permissions** consist of:

- **Camera**: access to the camera of the device.
- **Phone**: allows the application to read the state of the device, including the number of the device, the information of the network, the status of calls in progress, a list of all phone accounts registered on the device and making calls.
- **SMS**: the application can send and receive SMS.
- **Microphone**: uses the device's microphone to record audio.
- **Contacts**: uses contacts information.
- **Storage**: uses external SD to read and write data.

[6]https://www.uber.com/legal/es-es/document/?country=united-states&lang=en&name=privacy-notice

- **Location**: uses the location data of the device. Uber collects precise or approximate location data.

During the execution of the application, if the user authorizes one single permission, she is **automatically giving the permission to a number of single permissions that belong to the same group**. In other words, if the user gives permission to READ_PHONE_STATE, she is not only giving information about the network to which the device is connected to, but also to the number of the device, and all the other permissions that belong to the group "Phone".

In the analysis, the Camera permission was required when the user was scanning the code for Scooter service, and location permission was only required when the application was launched. In terms of **geolocation**, the application allows a **manual selection** of the **pick up point** when location permission is denied.

Uber uses SMS verification instead of asking for permissions to know the phone number. In SMS verification, the user must manually enter the phone number so there is no need to obtain it through system permissions. During the verification process, the application itself uses an API provided by Android called SMS Retriever API to **read the message without the need to ask the user for any SMS permission**. However, and according to the official documentation, it should not be possible to access any third-party SMS.

The application also allows the user to contact the driver, for example, to recover any belongings lost during the journey. The application **hides the real number at both ends** of the communication. However, this does not necessarily forbid **Uber from recording the conversation without the user's prior consent.**

Normal Permissions

In addition to sensitive permissions, which the user must approve, there exist **permissions that are granted by default to the application**. These permissions are known as **normal permissions.** Normal permissions are automatically gained by the system, once the application is installed in the device, only if the application is signed by reliable sources such as the Play Store. The Uber application installed in Android version 9.0 (API28) has the following remarkable normal permissions:

- ACESS_NETWORK_STATE**:** uses network information to know whether the device is connected to a WiFi network, signal level and the quality of the network.
- ACCESS_WIFI_STATE**:** uses WiFi information state.
- RECEIVE_BOOT_COMPLETED and ACTION_BOOT_COMPLETED**:** allows the application to start as soon as the system has finished booting.
- VIBRATE**:** uses the vibration of the device.
- GET_ACCOUNTS**:** Uber has access to the accounts known by the device, such as Google accounts or Facebook among others.
- INTERNET: allows Uber application to open internet connections and connect to the Internet.

- BLUETOOTH and BLUETOOTH_ADMIN**:** allows the Uber application to connect to paired Bluetooth devices, or find and pair Bluetooth devices.
- WAKE_LOCK**:** prevent the processor from going to sleep or the screen from dimming when the application is in use.
- SYSTEM_ALERT_WINDOW**:** allows an application to create overlay windows that allow it to be displayed on top of all other applications.
- MODIFY_AUDIO_SETTINGS**:** allows an application to modify global audio settings
- CHANGE_NETWORK_STATE**:** allows applications to change the state of network connectivity.

Normal permissions are supposed to not pose any threat to the user's privacy. However, the inappropriate use of some of these could be harmful. For example, through ACCESS_WIFI_STATE permission, Uber could know whether the device is connected to a WiFi network and know if the user is on the street or in a closed space. Moreover, this permission allows for **making a network topology**.

A network is a collection of devices connected to each other to allow the sharing of data. Hence, network topology is the description of how a network is arranged and how all the components are interconnected to each other, including physical and logical connectivity. By knowing the network topology, Uber could infer the behavior of different types of users, sharpening their prediction tools.

It is not clear what the purpose of the BLUETOOTH and BLUETOOTH_ADMIN permissions are in relation to Uber's activity, these allow the application to **manage the devices that are close to a given User's device**. For example, Uber could control the Bluetooth of the driver and connect to other nearby devices with open Bluetooth.

On a different note, the SYSTEM_ALERT_WINDOW permission should be restricted, since these windows are intended for system level interaction with the user. For example, to follow a route on Google maps or similar, parallel to the use of other applications. A problematic aspect of this permission could come from fraudulent malware, normally tied to payment systems, that uses this technique to position itself above other applications and see where the user presses the device, being able to **identify patterns such as PIN numbers**.

Privacy Note

According to Uber's Privacy Note,[7] the company collects:

- **Location data**: precise or approximate location data, without specifying if data is only collected when the Uber app is running in the foreground (app open and on-screen) or also in the background (app open but not on-screen). Uber states that this data is collected to enable and enhance use of apps, including to improve pick-ups, facilitate deliveries, enable safety features, and prevent and detect fraud.

[7]https://www.uber.com/legal/es-es/document/?country=united-states&lang=en&name=privacy-notice

- **Transaction information**: related to the use of their services, including the type of services requested or provided, order details, payment transaction information (such as a restaurant's or merchant's name and location and amount of the transaction), delivery information, date and time the service was provided, amount charged, distance traveled, and payment method. Additionally, if someone uses another user's promotion code, they may associate both of the users' names.

- **Usage data**: data about how users interact with their services, including access dates and times, app features or pages viewed, app crashes and other system activity, and type of browser. They also collect data regarding the third-party sites or services used before interacting with their services, which they use for marketing. In some cases, they collect this data through cookies, pixels, tags, and similar tracking technologies that create and maintain unique identifiers.

- **Device data**: data about the devices used to access their services, including the hardware models, device IP address or other unique device identifiers, operating systems and versions, software, preferred languages, advertising identifiers, device motion data, and mobile network data.

- **Communications data**: Uber enables users to communicate with each other and Uber through their mobile apps and websites. For example, they allow drivers and riders, and delivery persons and delivery recipients, to call, text, or send files to each other (generally without disclosing their telephone numbers to each other). To provide this service, Uber receives some data regarding the calls, texts, or other communications, including the date and time of the communications and the content of the communications. Uber may also use this data for customer support services (including to resolve disputes between users), for safety and security purposes, to improve our services and features, and for analytics.

- **Safety recordings**: In certain jurisdictions, and where permitted by law, users can record the audio and/or video of their trips through an in-app feature or using a dashcam. In app recordings are encrypted and stored on users' devices and are only shared with Uber if submitted to customer support by the users in connection with safety incidents.

The reasons why these data are collected are further developed in Annex 1, but it remains unclear whether the collection and processing of this data is justified under GDPR. On the one hand, Uber states that it does not sell information to third parties for their direct marketing except with the users' consent. However, it does share personal data with other users, affiliates, partners and subsidiaries such as Google and social media companies including Facebook and TikTok, marketing platforms, or research partners, among others. In this regard, it is not clear whether the data shared undergoes an anonymization process, which is paramount to safeguarding the users' privacy.

Analysis of Network Communications

Some of the data obtained and sent to the servers includes:

- Android identifier of the device, known as androidID.
- Operating system and its version, for example, Android 9.0 Pie.
- Api version, for example, API 28 (Pie).
- Device model and manufacturer, for example, OnePlus 6T.
- Country code and exact GPS location, ES, 41.390205, 2.154007.
- Mobile number associated with the SIM and the corresponding network, for example, +34666777888, Vodafone ES.

It should be noted that all the data that the application sends is associated with the personal account of the initial registration. The main data used for registration includes name, surname, email and mobile number.

During the use of the application, **Uber contacts some third-party servers and services.** Third-party servers and services are created by companies or developers that aren't Uber. This means that, **when they contact a third-party server, they are giving access to personal data without the user's permission**. An application that makes a large number of third party connections may steal user's private information, which results in a privacy breach. The third-party connections generally do not have a secure mechanism for data transmission, so it may be possible for the user's personal information to be breached in the process. Moreover, if an application contacts a particular third-party server many times, it may also be transmitting the user's private data (Kumar & Singh, 2015).

The servers and domains with which the Uber application communicates the most are detailed below:

➢ https://api.braintreegateway.com/

This service is used by the **payment system** to allow users to make payments by credit card and through Paypal. Once the application is running, it sends a request to this service to identify the country in which the user is located and the supported card types in the country. The application does not contact this server again unless the user wants to add a new payment system. It is important to note the fact that **the application can identify if the Paypal app is installed on the device and ready to use without the need for user interaction.**

➢ https://*.cfe.uber.com/

This domain belongs to Uber and the application contacts it several times. During the registration, once the mobile number is provided, the device contacts this server and **sends the collected data without asking for any permission**, except in the case of location permissions. The collected data includes:

- Application's version and the identifier
- Telephone operator's name associated with the device, for example Vodafone ES.
- Level of the battery (value from 0 to 1), and the charging status of the device.
- IP direction

- Whether the device is rooted, in other words, if it contains unauthorized modifications.
- Whether the app is running in an emulated environment, i.e. not on a mobile device
- The device model, including the architecture, OS version and IMEI identifier.
- The geolocation of the device. If it is in movement, this information will include the height, angle, and speed.

This data is sent every 4 seconds as a package when the user is not taking any actions and the application is running. Moreover, **information with geolocation data, precise location data, as well as the destination, vehicle's location and identifier is also sent every 4 seconds.**

This periodic data traffic is called **polling**. **The information is saved in an internal database of the device** and read whenever the application is reopened. Then, the server answers with the following information: drivers who are in the range of the last location coordinates, vehicle's orientation and the estimated time to arrive at the selected pick-up point.

On the other hand, when the user performs an action in the application, such as pressing a button or entering in a new window, the application sends an 'event' to the server. This **'event'** called **'analytics'** describes where the user is inside the application and what they are doing ,t and includes the following information:

- The application's version and identifier
- Telephone operator's name associated with the device. For example, Vodafone ES.
- Available memory left on the device
- Level of the battery (value from 0 to 1), and the charging status of the device.
- IP direction
- Whether the device is rooted. In other words, if it contains unauthorized modifications.
- Language of the device
- The device model, including the architecture, OS version, screen size in pixel and serial number.
- Amount of time the device has been on.
- The geolocation of the device. If it is in movement, this information will include the height, angle and speed.
- Time that Uber application has been running, including the last time it was opened.

All this data is updated and sent to the server each time the user performs an action in the application.

If the application is open, but not on-screen (also known as **background mode**), a package of information is sent to the server the next time the application is reopened, including the time it has been inactive, as well as the new and old location. The time during which the application is totally closed is recorded in the phone's internal memory, and an event similar to the background event is sent to the server when the application is reopened. Furthermore, information of the points of interest (hotspots determined by Uber for picking up new passengers) near the device are also sent through this domain. Most of these points are extracted from the Google maps API.

	Registration	**Application is running in foreground and user is making an action**	**Application is running in foreground and user is not making any action**	**Open app in background**	**Closed app**
Data package	• App's version and identifier • Phone operator • Level of the battery and charging status of the device • IP direction • Whether the device is rooted • Whether the app is running in an emulated environment • The device model, including the architecture, OS version and IMEI identifier • Geolocation of the device (if there is movement, this includes the height, angle and speed)	• App's version and identifier • Phone operator • Available memory on the device • Level of the battery and charging status of the device • IP direction • Whether the device is rooted • Language of the device • The device model, including the architecture, OS version, screen size in pixel and serial number. • Time the device has been on • Geolocation of the device (if there is movement, this includes the height, angle and speed) • Time that Uber application has been running, including the last time it was opened	• App's version and identifier • Phone operator • Level of the battery and charging status of the device • IP direction • Whether the device is rooted • Whether the app is running in an emulated environment • The device model, including the architecture, OS version and IMEI identifier • Geolocation of the device (if there is movement, this includes the height, angle and speed)	• Time it has been inactive • New location • Old location	• Time when it was closed • App's version and identifier • Phone operator • Level of the battery and charging status of the device • IP direction • Whether the device is rooted • Whether the app is running in an emulated environment • The device model, including the architecture, OS version and IMEI identifier • Geolocation of the device (if there is movement, this includes the height, angle and speed)
Location	No	Yes	Yes	N/A	No
Analytics	No	Yes	No	No	No
Points of interest	No	Yes	Yes	N/A	N/A
Frequency	In the registration	When user makes an action in the application	Every 4 seconds	When the app is reopened	When the app is reopened

Table 6: Data collected throughout the different status of the app (from registration to closed app)

➢ https://cn-geo1.uber.com

This server sends data related to app crashes due to code errors that cause unexpected device shutdowns or wrong actions. **Uber identifies the user** with device's data such as its telephone operator, version, model, used memory and location, among others. Moreover, along with this data, it sends a list of all the pieces of code that have been executed until reaching the failure.

Discussion

This analysis highlights many privacy concerns. From a data minimization, GDPR-compliant point of view, a **high percentage of Uber's required permissions are unnecessary** for the application to run and **could be harmful for a user's security** such as ACCESS_WIFI_STATE or SYSTEM_ALERT_WINDOW. The number of permissions requested of the user could be used as a parameter to evaluate the trustworthiness of an application.[8] To run the application it was only necessary to approve Camera permissions, therefore, the request of multiple other permissions makes one think that they could be used for information gathering purposes for third party servers.

Moreover, while **normal permissions are granted to the application by default** (ACESS_NETWORK_STATE, GET_ACCOUNTS, etc.) because they are not supposed to pose any threat to the user's privacy, they are considered by android developers as a dangerous category in some cases.[9] **Dangerous permissions put the privacy of users in jeopardy**. For example, allowing the ACCESS_WIFI_STATE permission can allow an app to send sensitive data to the app developer, since this permission is not needed to verify connectivity, because it allows apps to access private information about the Wi-Fi network that the user is connected to. Permissions like these can be considered dangerous because if a user approves one of the permissions in one of the categories, the app will have access to the permissions of the entire group of permissions. On the other hand, the SYSTEM_ALERT_WINDOW permission allows an app to overlay any other on the phone, blocking out the rest. This is used by some applications to display the most aggressive advertising possible. in fact, Google recommends that Android developers use this permission as little as possible.

For example, allowing the ACCESS_WIFI_STATE permission can allow an app to send sensitive data to the app developer, as this permission is not required to verify connectivity, because it allows apps to access private information about the Wi-Fi network to which the user is connected. Permissions like these can be considered dangerous because if a user approves one of the permissions in one of the categories, the app will have access to the permissions of the entire group of functionalities.

Another example is when an app can **access the internal memory of a user's phone** is dangerous in the sense that it could allow Uber to upload any number of personal images and files from the phone memory to a remote server without the knowledge of the user. Additionally, in the case of data collection, **Uber should provide the user with**

[8] Kumar, P.,Singh.,M.,(2015) Mobile applications: analyzing private data leakage using third party connections. *IEEE.* 28 September 2015. Available on: https://ieeexplore.ieee.org/document/7275584
[9] Khatoon, A.,Corcoran,P.,(2017) Android permission system and user privacy-A review of concept and approaches. *IEEE.* 18 December 2017. Available on: https://ieeexplore.ieee.org/document/8210616

information about which data they collect and its exact purpose prior to the installation of the application. Uber currently **collects data** like phone numbers and battery levels **without permissions** or even **without the user's knowledge**.

Beyond these specific examples, the excessive request for permissions can lead to vulnerabilities in terms of security and personal data that are stored in the applications and that they take advantage of to sell or share the user's personal data with third parties, to know the status of the phone, the status of the mobile network, calls, access to the phone memory, contact list or access to the mobile camera or microphone, even when the ability to use these functionalities goes beyond what is strictly necessary to offer the service it manages.

The analysis identified multiple cases of sensitive data being sent to Uber servers. **Much of this information is not necessary for the proper functioning of the application,** including:

- Available memory on the device.
- Level of the battery (value from 0 to 1), and the charging status of the device.
- IP direction of the device and whether it is connected to WiFi.

As it is mentioned in Uber Privacy Notice, Uber **collects the content of all communications**, meaning that calls made between riders and users are recorded by Uber **without the user's prior consent**. They also send the **user's precise geolocation position every 4 seconds** to its server despite the fact that, back in 2014, the Federal Trade Commission, or FTC, reported that Uber stopped using a tool informally named 'God View map'. This tool allowed riders to follow passengers' position in a map, and although this tool is no longer in use, the Uber application is still compromising its users' privacy in a similar way as the 'God View map' did. However, and with regards to privacy, it is worth noting that Uber **does not anonymize data related to app crashes**.

With the help of a third-party server, **the application is able to identify whether Paypal is installed on the device** and ready to use without the need for user interaction.

In the company's Privacy Notice it is not whether location data and other types of data are collected when the app is running in the foreground (app open and on-screen) or background (app open but not on-screen). This could be a potentially interesting point for further research, analyzing which data collects Uber when the app is running in the background.

With this, it is obvious that Uber makes an **excessive use of analytics**. And while it is true that the use of all the data Uber collects and sends to its servers makes it easier for developers and designers to adjust and perfect their application, **Uber should always allow the user to choose which data is shared.**

ACKNOWLEDGEMENTS

This project is the result of a collaboration between Eticas, Taxi Project and *Observatorio TAS*. The team would like to thank the interviewed PHV drivers and fleet manager whose invaluable responses contributed to the findings and discussions presented in this report. We also want to acknowledge UATAE's collaboration.

Research Leads:

- Dr. Gemma Galdon-Clavell, Founder of Eticas
- Iliyana Nalbantova, Junior Ethics and Technology Researcher at Eticas
- Sergi Cutillas, the Taxi Project and Observatorio TAS

Contributing Researchers:

- Ferran Celades, the Taxi Project
- Miguel Azores, Junior Ethics and Technology Researcher at Eticas
- Matteo Mastracci, Ethics and Technology Researcher at Eticas
- Virginia Bertelli, Ethics and Technology Researcher at Eticas

Other Contributors:

- Tito Álvarez, the Taxi Project and Observatorio TAS
- Marta Vergara, Lead Storyteller at Eticas
- Isabela Miranda, Project manager at Eticas

Recommended citation: "Eticas, the Taxi Project, & Observatorio TAS. (2023). Adversarial audit of ride-hailing platforms: Algorithmic compliance with competition, labor and consumer law in Spain. Association Eticas Research and Innovation. Taxi Project 2.0. Observatorio TAS"

BIBLIOGRAPHY

ADCU (2020). Uber drivers take unprecedented international legal action to demand their data. https://www.adcu.org.uk/news-posts/uber-drivers-take-unprecedented-international-legal-action-to-demand-their-data.

Andrino, B., Llaneras, K., & Grasso, D. (2021). El mapa de la renta de los españoles, calle a calle. https://elpais.com/economia/2021-04-29/el-mapa-de-la-renta-de-los-espanoles-calle-a-calle.html.

Aseval (2022). Sindicatos y patronales de las PHV firman una subida salarial del 12% en el primer convenio laboral estatutario. https://aseval-madrid.com/sindicatos-y-patronales-de-las-PHV-firman-una-subida-salarial-del-12-en-el-primer-convenio-laboral-estatutario/.

Ballestrero M. V. (n.d.). Ancora sui rider. La cecità discriminatoria della piattaforma. https://www.rivistalabor.it/ancora-sui-rider-la-cecita-discriminatoria-della-piattaforma/.

Calo, R., & Rosenblat, A. (2017). The taking economy: Uber, information, and power. Colum. L. Rev., 117, 1623.

Calvano, E., Calzolari, G., Denicolo, V., & Pastorello, S. (2020). Artificial intelligence, algorithmic pricing, and collusion. American Economic Review, 110(10), 3267-97.

Chang, Y., Winston, C., Yan, J. (2022). Does uber benefit travelers by price discrimination? In Journal of Law and Economics.

Chen, L., Mislove, A., & Wilson, C. (2015, October). Peeking beneath the hood of uber. In Proceedings of the 2015 internet measurement conference (pp. 495-508).

Christophi, H. (2017). Uber Pays Customers $344,000 for Tip Trickery. https://www.courthousenews.com/uber-pays-customers-344000-for-tip-trickery/.

CNMC (n.d.). Los beneficios de la competencia para los consumidores: Preguntas y respuestas.

Consejeria de Economia, Hacienda y Empleo (2022). Resolución de 26 de enero de 2022, de la Dirección General de Trabajo de la Consejería de Economía, Hacienda y Empleo, sobre registro, depósito y publicación del Convenio Colectivo del Sector de Transporte de Pasajeros de la Comunidad de Madrid en Vehículo de Turismo Mediante Arrendamiento con Licencia PHV, suscrito por Unauto PHV Madrid y

Aseval y por la representación sindical CCOO, SLT y UGT. https://noticias.juridicas.com/base_datos/CCAA/719289-convenio-colectivo-del-sector-de-transporte-de-pasajeros-de-la-comunidad-de.html#c5.

Cortés, I. (2022). Semanas de 60 horas en Uber y Cabify: "Si libras dos días, vamos a tener problemas". https://www.elconfidencial.com/juridico/2022-06-14/jornadas-interminables-uber-cabify_3439310/.

DeGeurin, M. (2022). Amazon Sued for Allegedly Stealing Driver's Tips, Launches New Gamified Tipping System. https://gizmodo.com/amazon-amazon-prime-gig-work-alexa-1849865607.

Drahokoupil, J., & Piasna, A. (2017). Work in the platform economy: beyond lower transaction costs. Intereconomics, 52(6), 335-340.

Eticas (2021). Lanzamos un proyecto de auditorías externas de algoritmos. https://eticasfoundation.org/auditorias-externas-algoritmos/.

Eticas (2022). Guia de Recomendaciones.

Eticas. (2022). The External Audit of the VioGén System.

Eyert, F., Irgmaier, F., & Ulbricht, L. (2022). Extending the framework of algorithmic regulation. The Uber case. Regulation & Governance, 16(1), 23-44.

Farrar, J., Cutillas, S. (2022). Workers' Recommendations on the Draft EU Platform Work Directive.

Graham, M., & Woodcock, J. (2018). Towards a fairer platform economy: introducing the Fairwork Foundation. Alternate Routes, 29.

Guillén Navarro, N. A. (2018). El arrendamiento de vehículos con conductor (PHV) y su entramado jurídico: el avance de Uber, Cabify y la economía colaborativa. Revista De Estudios De La Administración Local Y Autonómica (9), 128–147. https://doi.org/10.24965/reala.v0i9.10470.

Jiao, J. (2018). Investigating Uber price surges during a special event in Austin, TX. Research in Transportation Business & Management, 29 , 101-107. https://doi.org/10.1016/j.rtbm.2018.02.008.

Kerr, D. (2021). Why Are Some Instacart Workers Calling for an App Boycott? https://themarkup.org/2021/10/12/why-are-some-instacart-workers-calling-for-an-app-boycott.

Kumar, P., Singh ,M. (2015). Mobile applications: analyzing private data leakage using third party connections. IEEE. 28 September 2015. Available on: https://ieeexplore.ieee.org/document/7275584.

Lomas, N. (2021). Italian court rules against 'discriminatory' Deliveroo rider-ranking algorithm. https://tcrn.ch/38aQeoX.

Reuters (2020). Spain's Supreme Court rules food delivery riders are employees, not freelancers. https://www.reuters.com/article/uk-spain-glovo-ruling-idUKKCN26E2NR.

Rosenblat, A. (2018). Uberland: How algorithms are rewriting the rules of work. Univ of California Press.

Susser, D. (2019, January). Invisible influence: Artificial intelligence and the ethics of adaptive choice architectures. In Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society (pp. 403-408).

Todolí Signes, A. (2015) El impacto de la UBER ECONOMY en las relaciones laborales: los efectos de las plataformas virtuales en el contrato de trabajo, IUSLABOR (3) pp. 1-26.

Worker Info Exchange (2021). Dutch & UK courts order Uber to reinstate 'robo-fired' drivers. https://www.workerinfoexchange.org/post/dutch-uk-courts-order-uber-to-reinstate-robo-fired-drivers.

Yeung, K. (2018). Algorithmic regulation: A critical interrogation. Regulation & Governance, 12(4), 505-523.

eticas

TAXI PROJECT
ASOCIACION DE PRESION

CONTACT
info@taxiproject.eu
info@eticas.tech