The Economic Effects of Mobile Internet Access – Evidence from Roam-Like-at-Home*

Martin Quinn[1], Miguel Godinho de Matos[1], and Christian Peukert[2]

[1]Católica-Lisbon School of Business and Economics
[2]HEC Lausanne

April 13, 2021

Abstract

We evaluate the Roam-Like-At-Home regulation's welfare effects, which drastically reduced the price of accessing the mobile internet for EU residents when traveling abroad in the EU. Estimates from individual-level usage data suggest that consumer surplus increased at least by 2.73 EUR/user/day. A decomposition shows the heterogeneous impact of the regulation on different user segments. We estimate that around 45% of the gains stem from a reduction in deadweight loss, i.e., new users accessing the mobile internet. We then discuss how the regulation affected content provider surplus and its implications for various policy areas.

Keywords: Mobile Data, Roaming, Consumer Surplus, Regulation, Telecom, Difference-in-differences, Online Experiment.

JEL Codes: L96, L51, O33, D62

*Quinn (corresponding author): martinquinn@ucp.pt, Godinho de Matos: miguel.godinhomatos@ucp.pt, Peukert: christian.peukert@unil.ch. We are grateful to Stefan Bechtold, Chiara Farronato, Mark Trembley, Leonardo Madio, Robert Somogyi, Hannes Ullrich and conference and seminar participants at the Media Economics Workshop Braga 2019, Media and Digitization Meeting Zurich 2019, WISE 2019, SCECR 2020, VIDE Brown Bag 2020, ETH Center for Law and Economics Brownbag 2020, Louvain Economics of Digitization Seminar 2020, CESifo Economics of Digitization Conference 2020, ifo Digital Transformation and Innovation Seminar 2021.

1 Introduction

It is now widely known that digitization, and the internet in specific, are associated with firm productivity, innovation, employment, and, more broadly, economic growth (Czernich et al., 2011; Cardona et al., 2013; Hjort and Poulsen, 2019). The literature has started to develop a deeper understanding of the channels through which digitization generates value, mostly highlighting lower costs on the demand- and supply-side (Goldfarb and Tucker, 2019), translating into greater product variety and more efficient entry, which creates substantial welfare effects (Brynjolfsson et al., 2003; Aguiar and Waldfogel, 2018).

However, the bottleneck to participate in the value gains caused by digitization is unrestricted access to the internet. Evidence from within and across countries documents a *digital divide* in terms of income, education and geography (Schleife, 2010; Viard and Economides, 2015; Silva et al., 2018). Most research in this literature is focused on fixed-line technology, while device usage and applications are increasingly shifting towards mobile (Einav et al., 2014). According to data from Statcounter, the majority of web traffic in 2019 came from mobile devices.[1] Mobile broadband can be complementary to fixed-line broadband, not only in closing the gap in internet access between rural and urban areas (Prieger, 2013), but also in individual-level adoption decisions (Xu et al., 2019). Not having access to the mobile internet can have meaningful economic implications. Evidence shows that mobile devices are linked to different internet consumption and e-commerce behavior, and that consumers interact more with geographically local content on mobile devices (Ghose et al., 2013; Fong et al., 2015; Mang et al., 2016; Xu et al., 2017).

Policy measures to increase internet penetration include the forced unbundling of infrastructure (Nardotto et al., 2015), subsidies for infrastructure investments (Briglauer et al., 2019), minimum coverage requirements (Fabrizi and Wertlen, 2008), and (wholesale) price regulation (Vogelsang, 2003; Spruytte et al., 2017). An important aspect of price regulation concerns *roaming*.

Within the limits of regulatory frameworks, network operators may charge different prices for using voice or data services on networks owned by other operators (Zucchini et al., 2013), especially across national borders. Depending on a consumer's physical location, the price to access the mobile internet can be prohibitively high. For example, the average price per MB used abroad for consumers from OECD countries was at least 2.5 USD in 2010. Such prices were many times more expensive than comparable domestic use (OECD, 2011). There are several examples

[1]See https://gs.statcounter.com/platform-market-share/desktop-mobile-tablet/worldwide/.

of consumers being billed enormous sums for casual mobile internet usage while on vacation.[2] Starting in 2007, European regulators imposed rules that gradually lowered the price-based gap in mobile telecommunication between locals and visitors within the European Economic Area (EEA), i.e., in all EU member states, Iceland, Liechtenstein and Norway. However, survey evidence suggests that most consumers still chose not to use the mobile internet while traveling abroad in 2012 and 2013 (Mang et al., 2016). In the last phase of the regulation, effective June 15th, 2017, telecommunication providers in the EU were no longer allowed to charge for mobile voice and data services when consumers use networks in foreign countries within the EEA. Coined *Roam-Like-At-Home* (RLAH), the regulation promised EU residents that the price for mobile telecommunication services is the same as at home when traveling in the EEA. Travel outside the EEA was not affected.[3]

Imposing a price cap on network operators can have intricate welfare effects (Spruytte et al., 2017), not only by transferring surplus from network operators to consumers, but also by redistributing surplus across consumer types (e.g., Chillemi et al., 2019; Maillé and Tuffin, 2017), and by increasing surplus of content providers through increased demand (e.g., Greenstein et al., 2016). In this paper, we aim to empirically evaluate some of these welfare effects in the context of the RLAH regulation.

We have access to an anonymous dataset from a network operator in a EU country we denote by MOBILE. We can track the mobile data usage of all clients who used their mobile phone at least once while traveling abroad between September 2016 and December 2017. We can compare mobile data usage in the home country to mobile data usage when traveling in the EEA and outside of the EEA. Using a difference-in-differences specification, we show that the regulation increased mobile data usage of European travelling within EEA by an average of 177%. We estimate that this corresponds to an increase in consumer surplus of at least EUR 2.73 per user per day. Decomposing the surplus gains by consumer types based on pre-regulation mobile data usage, we highlight important heterogeneity. Users with higher pre-regulation usage experience higher absolute gains, mainly driven by the revaluation of existing mobile data usage, and less by additional usage. Conversely, users with lower pre-regulation usage experience higher relative gains, mostly driven by additional usage. These results suggest that the regulation was partially effective in closing a digital divide between locals and visitors.

[2] See, for example, https://www.theguardian.com/money/2012/may/25/data-roaming-smartphone-abroad.

[3] See https://europa.eu/youreurope/citizens/consumers/internet-telecoms/mobile-roaming-costs/index_en.htm

To understand which types of consumption drive consumer surplus and to investigate content provider surplus, we conduct an online experiment on mobile data usage of Europeans conditional on different data allowances. We show that while music/video consumption and the use of review platforms are the services that experience the highest growth in usage, communication, social media and search explain more than half of the consumer surplus gains.

Estimates from MOBILE allow us to calibrate a back-of-the-envelope calculation of the overall welfare effects of the regulation. Assuming that the average user of our network operator is representative of all EU citizens, we calculate the total consumer surplus increase between June 2017 and December 2017 to be around EUR 2B.

We also discuss the surplus effects of content providers. We show that the regulation stimulated the adoption of online services. To do so, we use both a range of publicly available data to study the effects of RLAH on consumers' adoption of content providers and data from our online experiment. Our results consistently suggest that RLAH rules' implementation increased travelers' usage of various content providers while roaming. Our experiment's results suggest that consumer surplus gains originate from adopting new content services such as music/services and review platforms and increased use in content categories such as communication, social media, and search.

Finally, we perform a rough estimation of the average change in network operator surplus in the EEA. We show that losses to network operators are smaller than our estimates of consumer surplus gains, and depend on tourism flows and come from from the decrease in the marginal revenue from incoming tourists.

By showing that RLAH led to a transfer of surplus from network operators to consumers and content providers, we make several contributions. We add to the (mostly theoretical) literature on telecommunication regulation (Genakos and Valletti, 2011; Chillemi et al., 2019), including the distribution of rents between internet service providers and content providers (Easley et al., 2018). More broadly, we contribute to the related literature on the welfare effects of digitization and the mobile internet (Brynjolfsson et al., 2003; Aguiar and Waldfogel, 2018; Ghose and Han, 2011; Ghose et al., 2013; Xu et al., 2019).

2 Regulation of roaming charges in Europe

The mobile internet dates back to the introduction of second-generation mobile networks in the early 1990s and has since diffused widely. According to data from the International Telecommunications Union (ITU), global mobile broadband penetration increased from 4% in 2007 to 70% in 2018. More than 90% of the world population is covered by at least a third-generation mobile network in 2018. Access to the mobile internet is extraordinarily convenient, as it grants the ability to consume or provide information outside the reach of a fixed-line internet connection at home or work. By the early 2000s, many countries had privatized the telecommunications sector (Waverman and Sirel, 1997), and mobile telecommunication networks are now typically operated by several firms that compete for consumers (Li and Xu, 2004). Interconnection between networks, which enables termination of voice calls across networks and internet access through a competitor's infrastructure, is governed by regulation, mostly concerning network access fees (Vogelsang, 2003; Jullien et al., 2013). Within the regulatory frameworks, network operators may pass on some of those fees (wholesale roaming fee) to consumers for using off-network telecommunication services in the same country and when crossing national borders (retail roaming fee). While national roaming charges often only translate into differentiated price structures for on-net and off-net voice calls (Zucchini et al., 2013), international roaming charges often also translate into a differentiated price structure for national and international data services. Historically, retail roaming fees have been above cost, despite efforts of wholesale regulation (Infante and Vallejo, 2012). High roaming charges are among the most significant deterrents of mobile data usage for international travelers. In a Eurobarometer survey from 2014, 52% of respondents who travel to other EU countries said that they switch off their phone and never use it or switch off the data roaming capabilities of their phone when traveling.[4]

Starting in 2007, roaming regulation in the EU had four major regulatory rounds that gradually introduced wholesale and retail price caps for voice, short message services (SMS), and mobile data services (see Infante and Vallejo, 2012; Spruytte et al., 2017 for a detailed discussion).

Figure 1 shows the evolution of the regulated wholesale and retail price caps for mobile data services in the EU over time. The wholesale price cap for data services reduced from €c100/MB in July 2009 to €c5/MB before the implementation of RLAH rules. Finally, in June 15, 2017, RLAH rules stipulated the end of retail roaming charges for EU residents that travel to countries

[4]See https://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_414_en.pdf.

within the EEA. Wholesale prices were reduced to €c0.77/MB for 2017 while retail price for in-plan roaming were capped to a minimum of half the wholesale price (hence €c0.385/MB for 2017) and to the wholesale price for out-of-plan roaming within Europe. In 2018, 2019 and 2021, price caps for wholesale and retail prices continued to decline.

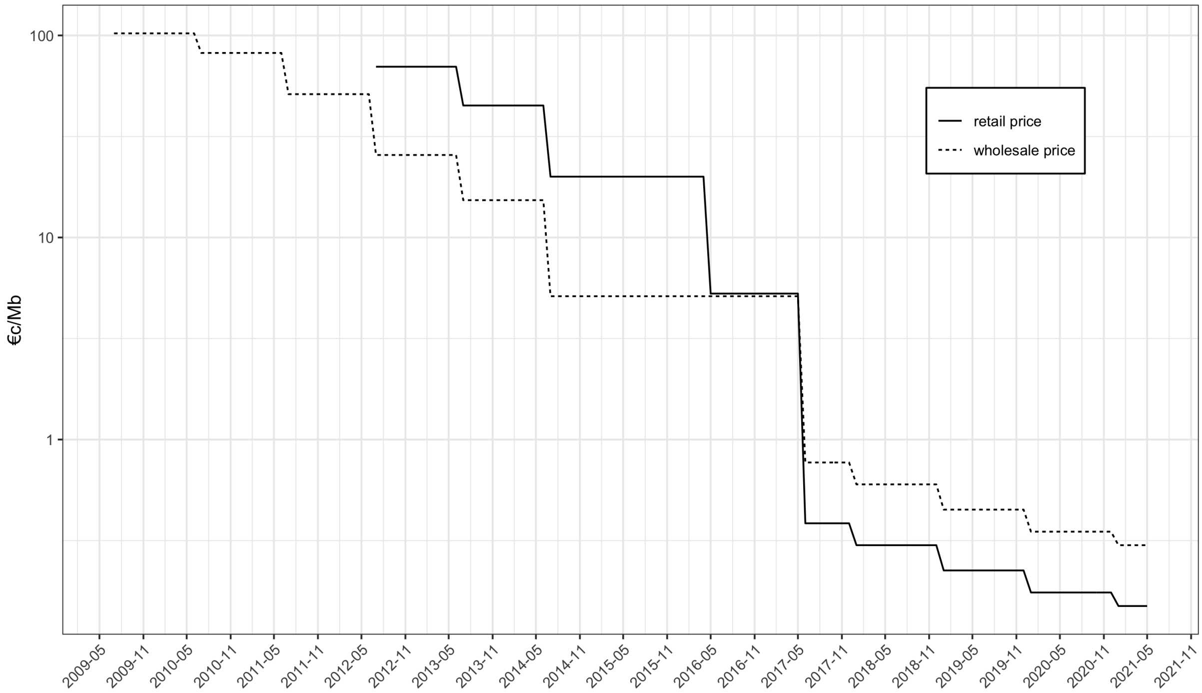

Figure 1: Evolution of retail and wholesale caps.

So far, empirical studies on the effects of the RLAH regulation are scarce. In a report to the European Parliament and the Council, the EC claims mobile operators' complied with the new roaming rules.[5] Using data from the Body of European Regulators for Electronic Communications (BEREC), the EC reported an increase in total roaming data traffic in the EEA from 15.78 million GB in Q3 2016 to 84.37 million GB in Q3 2017. The average data roaming consumption per month and subscriber increased from 59MB in Q3 2016 to 243MB in Q3 2017 and average prices for domestic mobile internet decreased. The counterfactual in these statistics is not clear (domestic prices might have decreased even more in the absence of RLAH). Therefore, it remains undetermined whether the RLAH regulation had waterbed effects and, if so, whether they were strong enough to be overall welfare reducing (Baake and Wagner, 2018).

[5] See https://ec.europa.eu/transparency/regdoc/rep/1/2018/EN/COM-2018-822-F1-EN-MAIN-PART-1.PDF.

3 Methodology

We follow Brynjolfsson et al. (2003), Hausman (1981) and Hausman and Leonard (2002) to estimate the effect of RLAH on consumer surplus.

We start with a simple model where consumers draw utility from accessing the internet at a price p_m, and from consuming and providing content at a price p_c. We model RLAH as a change in prices between periods $t \in \{0, 1\}$. The consumer surplus effect of RLAH can be expressed as the compensating variation CV. In our case, CV measures the payment that a consumer would require to remain in their (lower) initial level of utility after the price reduction.

Considering the utility level of period $t = 1$, we can write the consumer surplus change as

$$CV = e(p_{m0}, p_{c0}, u_1) - e(p_{m1}, p_{c1}, u_1) \tag{1}$$

where p_{mt} captures the price per MB purchased from the network operator, and p_{ct} captures the price per MB of content, where t is equal to 0 before RLAH and 1 after RLAH.

We assume that demand follows a standard log-linear form:

$$q(p_m, p_c, y) = A(p_m + p_c)^{\epsilon} y^{\delta} \tag{2}$$

where A is a shift parameter, and y denotes income, δ the income elasticity and ϵ the total price elasticity. Our demand specification captures the fact that consumers purchase a data allowance from the network operator and spend MBs on content. That is, consumers are sensitive to the total price per MB consumed, i.e. $p_m + p_c$.

Using Roy's identity, the indirect utility function and the expenditure function, we follow Hausman (1981) to rewrite the compensating variation such that:

$$CV = [\frac{1-\delta}{1+\epsilon} y^{-\delta}\big((p_{m0} + p_{c0})q_0 - (p_{m1} + p_{c1})q_1\big) + y^{1-\delta}]^{\frac{1}{1-\delta}} - y \tag{3}$$

Prior studies on the welfare effects of telecommunication technology such as Brynjolfsson (1996) and Hausman (1997) suggest that CV measures are not very sensitive to the estimated income elasticity. This is in line with the idea that we can ignore income effects when purchase amounts are a small fraction of the consumer's annual income (Willig, 1976), which we assume to be the case with roaming charges. Hence we assume $\delta = 0$, which allows us to arrive at a Marshallian expression of consumer surplus change:

$$CV = \frac{(p_{m0} + p_{c0})q_0 - (p_{m1} + p_{c1})q_1}{1+\epsilon} = \frac{p_{m0}q_0 - p_{m1}q_1}{1+\epsilon} + \frac{p_{c0}q_0 - p_{c1}q_1}{1+\epsilon} \tag{4}$$

We make the additional assumption that prices of online content do not change across time $p_{c0} = p_{c1} = p_c$. This assumption is reasonable because a large part of online content, such as internet search, is free. Furthermore, travelers are a relatively small share of overall internet users and are unlikely to change content providers' pricing choices.

$$CV = \frac{p_{m0}q_0 - p_{m1}q_1}{1+\epsilon} + p_c\frac{q_0 - q_1}{1+\epsilon} \tag{5}$$

To compute expression (5), we need estimates for p_{m0}, p_{m1}, q_0, q_1, p_c and ϵ. Estimating ϵ is challenging because we do not observe a continuous exogenous variation in prices. However, we exploit the discrete exogenous change in prices induced by RLAH. That is, we approximate ϵ as the mid point arc-elasticity of demand (Allen and Lerner, 1934), such that

$$\epsilon = \frac{(q_1 - q_0)}{(p_{m1} + p_c) - (p_{m0} + p_c)} \times \frac{(p_{m1} + p_c) + (p_{m0} + p_c)}{(q_1 + q_0)}. \tag{6}$$

This simplifies (5) to

$$CV = (p_{m0} - p_{m1}) \times \frac{q_0 + q_1}{2}. \tag{7}$$

Figure 2 illustrates that the approximation of ϵ as the arc elasticity can lead to an overestimation of the consumer surplus gains. The latter is small in the presence of large differences in prices, as is the case with RLAH.

As we are specifically interested in the effect of the regulation we will focus on estimating the share of q_1 that can be attributed to RLAH. However, we do not observe exact roaming retail

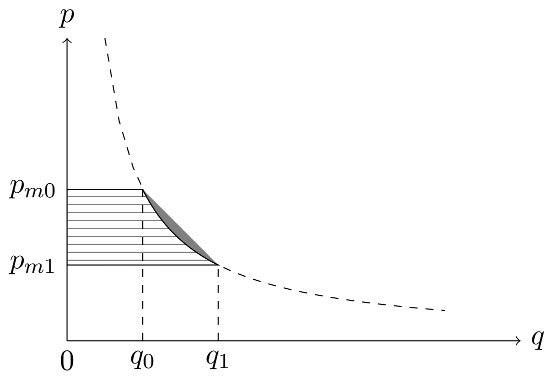

Figure 2: Consumer surplus calculation
Note: The hatch grey area captures the change in consumer surplus while the filled grey area represent overestimation from use of the arc-elasticity.

prices p_{m0} and p_{m1} at the individual-level. We know the maximum price that network operators can charge before and after RLAH, and we approximate average values of p_{m0} and p_{m1}. The maximum price per MB that operators could charge for roaming was established at the domestic price p_h + €c5 before the regulation and decreased to €c0.385 after.[6] In practice, this means that if a consumer pays a domestic price that is lower than the price cap for roaming – which, by definition, is the case before RLAH – the operator can reduce the roaming data allowance to match the regulated roaming price per MB. However, if a consumer pays a higher price at home than the regulated roaming price after RLAH, they will enjoy their entire domestic data allowance while roaming.

Hence, the price per MB p_{mt} paid by users in period t can be defined as the maximum between price per MB at home p_h and the regulated price p_{rt}. Hence $p_{m0} = \max(p_{r0}, p_h) = \max(0.05+p_h, p_h) = 0.05 + p_h$ and $p_{m1} = \max(p_{r1}, p_h) = \max(0.00385, p_h)$.

Plugging these expression in (7), the consumer surplus change after RLAH can be calculated as:

$$CV = (0.05 + p_h - \max(0.00385, p_h)) \times \frac{q_0 + q_1}{2} \tag{8}$$

The following section describes the data sources and empirical strategies that we use to calibrate the consumer surplus calculations laid out above.

[6]See section 2 and https://ec.europa.eu/digital-single-market/en/roaming-tariffs.

4 Data and Empirical Strategy

4.1 Observational Data

We have access to an anonymized panel dataset from an European network operator, from now on called MOBILE. Our sample includes anonymized information from all clients who used their mobile phone at least once while traveling abroad between September 2016 and December 2017. These clients represent 5% of domestic clients which is a similar number to the one reported in the official Eurobarometer statistics on roaming usage during that time period (European Commission, 2017). We observe the average daily number of MBs downloaded and uploaded – while abroad – for each mobile phone number and week. We also observe from which country the mobile internet traffic originated.

Table B.6 provides the descriptive statistics of mobile internet usage of MOBILE's clients. There is an overall increasing trend in mobile internet usage. Traffic while travelling within EU increased by 177 percent from an average of 31 MB/day in the period before RLAH (average from week 39 of 2016 to week 23 of 2017) to 85 MB/day in the period after RLAH (average from week 24 of 2017 to week 52 of 2017). In contrast, the average mobile data usage while being abroad in a country outside EEA remained stable around 7 MB/day.

We use the dataset described above to estimate the quantity effect of RLAH via difference-in-differences (DiD). We compare an individual i's demanded quantity of mobile data when traveling in the EEA and outside of the EEA, before and after RLAH:

$$MobileData_{ijt} = \delta After_t \times RLAH_{ijt} + a_i + \kappa_j + \theta_t + v_{ijt}, \tag{9}$$

where $MobileData_{ijt}$ is the sum of download and upload traffic in MBs initiated by individual i, abroad in country j, in week t. $RLAH_{ijt}$ indicates whether individual i is consuming mobile data while abroad in a country of the EEA where RLAH rules apply. $After_t$ is a dummy variable equal to one throughout the period for which the RLAH policy is in effect. We also include individual-fixed effects a_i, country-fixed effects κ_j, and week-fixed effects θ_t. The error term v_{ijt} follows the usual assumptions. In our estimations, we cluster standard errors at the individual level to allow for arbitrary serial correlation.

The identifying assumption of DiD requires that without RLAH, the mobile data consumed by the treatment group (travelers subject to RLAH) and those in the control group (travelers not subject to RLAH) follow similar trends. While this is not testable, we provide evidence that at least the necessary condition – similar trends before RLAH – holds. We provide details of this robustness analysis in appendix B.3. Finally, we estimate equation (9) separately for different deciles of pre-RLAH mobile data usage.

In the specification of equation 9, the estimation of δ that we denote by $\hat{\delta}$ measures the change in megabyte consumption due to the regulation. Our consumer surplus estimation requires estimating q_0 and q_1. We recover q_0 from the sample mean of RLAH travel activity before the regulation $\hat{q_0}$, and compute $\hat{q_1} = \hat{q_0} + \hat{\delta}$. We then plug our quantity estimates in the consumer surplus formula:

$$CV = (0.05 + p_h - \max(0.00385, p_h)) \times \frac{2\hat{q_0} + \hat{\delta}}{2} \tag{10}$$

Finally, we calibrate the missing price data with information from the European Commission. We use the average price per MB for domestic mobile data in countries similar to those where MOBILE operates - between €c0.15 and EUR €c0.383 in 2017 - and we consider two scenarios:[7]

- A *low domestic price scenario* where $p_h = 0.0015$, which sets roaming retail prices such that $p_{m0} = 0.05 + 0.0015 = 0.0515$ and $p_{m1} = \max(0.00385, 0.0015) = 0.00385$.
- A *high domestic price scenario* with $p_h = 0.00383$, which sets roaming retail prices such that $p_{m0} = 0.05 + 0.00383 = 0.05383$ and $p_{m1} = \max(0.00385, 0.00383) = 0.00385$.

4.2 Online experiment

The observational dataset obtained from MOBILE does not allow us to determine how clients allocated their data allowances across content providers. To study how consumers distributed the additional megabytes they consume after RLAH, we ran an online experiment among 2000 European consumers through the crowd-sourcing platform Clickworker.

We provide the questionnaires' details and descriptive statistics in appendix A.[8]

[7] See https://ec.europa.eu/digital-single-market/en/news/mobile-broadband-prices-europe-2017

[8] While we don't know whether our data is representative of the population of European consumers that are affected by the RLAH regulation, we can compare some key demographics of our participants to consumers in a Eurobarometer survey from August 2017 (see European Commission, 2017). We select the subsample of consumers in Eurobarometer that state that they have traveled at least once in the EU in the previous 12 months. Compared

In the online experiment we ask participants about their monthly data allowance and present them with two choice scenarios. In each scenario, we ask participants to envision that they are traveling abroad without access to Wifi or a desktop computer. We place our first scenario in the context before RLAH and give participants a data allowance of 140 MB.[9] Participants are then asked to allocate the MBs they have across six different content provider types: Communication, Search, Social Media, News, Music/Video, Review platforms, Transportation and none. For each category, we give examples of popular platforms (e.g., WhatsApp, Instagram, Spotify, etc.) and instructions on how units of consumption (messages, minutes of scrolling and posting, number of songs, etc.) translate into MBs. We place the second scenario after RLAH, and we tell participants that they have a mobile data allowance while traveling equal to the one they enjoy at home. We also ask them to allocate their allowance across content. In both scenarios, participants can leave unused MBs to a residual category. Table A.2 provides the descriptive statistics on respondents' mobile data choices.

For each respondent, we use their declared MB allowance and price paid at home to compute a hypothetical MB allowance when traveling abroad before and after the regulation.[10] Suppose a user pays more than the regulated price per megabyte (domestic price + €c5 before the regulation and €c0.385 after the regulation). In that case, she will benefit from her entire home allowance when traveling within the EU. However, if she pays less, she will have her MB allowance reduced to where her price per MB matches the actual regulated price.

After computing a users' roaming data allowance before and after RLAH, we determine consumption for each content provider by excluding the category that captures unused MBs. This category allows us to compute the average amount of MBs consumed per user per content-type:

$$MobileData_{ict} = \sum Content_{ict}(\gamma_c + \delta_c After_t) + a_i + v_{ict}, \tag{11}$$

where $Content_{ict}$ indicates respondent i's allocation of MBs to content of type c in scenario t, and the δ_c coefficients reflect the difference between scenarios. In this case $After_t$ is a dummy

to this sample, our participants are more likely to be male (56% vs. 50%) and less likely to have tertiary education (54% vs. 87%), but our participants have a similar age profile (86% vs 87% are younger than 35). Our sample consists of participants living in all EU countries except Luxembourg and Romania, but is skewed towards Germany (43%), Spain (14%), Italy (13%) and France (7%).

[9] This is the average allowance we observe in pre-study data (N=400).

[10] To better match our MOBILE analysis, we keep users who have spent one or more MB in both scenarios. This decision trims our dataset to 1233 unique users and does not significantly affect our results.

variable equal to 1 in the scenario that captures the RLAH rules. Since we have two observations per respondent, we also control for individual-fixed effects a_i.

Using the estimation of each δ_c we denote by $\hat{\delta}_c$ we compute q_{0c} and q_{1c}. We set q_{0c} as the sample mean of data consumption before the regulation $\hat{q_{0c}}$ for content provider c, and $\hat{q_{1c}} = \hat{q_{0c}} + \hat{\delta}_c$. Then, we recover how much each content provider weights in each individual's mobile data consumption (see equation (12)) and we weight equation (10) by each content type share:

$$share_{0c} = \frac{\hat{q_{0c}}}{\sum_c \hat{q_{0k}}}, \qquad share_{\delta c} = \frac{\hat{\delta}_c}{\sum_c \hat{\delta}_c} \tag{12}$$

$$CV_c = (0.05 + p_h - \max(0.00385, p_h)) \times \frac{2\ share_{0c} \times \hat{q_0} + share_{\delta_c} \times \hat{\delta}}{2} \tag{13}$$

4.3 Observational data from exemplary content providers

We complement our main analysis with estimates of the impact of RLAH on demand for content using publicly available data from a few exemplary content providers. We access data from Google Trends, the joint industry committee of German publishers and advertisers (IVW), TripAdvisor, and Twitter. We provide details on these datasets and analyses in appendix C.

5 Results

This section provides descriptive evidence and uses the methods and data described above to estimate mobile data consumption changes caused by RLAH. We then use our quantity change estimates to calculate consumer surplus changes on average by consumption intensity and content types.

5.1 Changes in quantities

Before implementing the regression specification presented in section 4.1, we provide model-free evidence of the impact of RLAH rules on the use of mobile data by travellers using our MOBILE database.

The left panel of figure 3 shows the average weekly mobile data consumption (sum of uploads and downloads) while abroad. We distinguish between data consumption abroad in EEA countries, where RLAH rules apply, and non-EEA countries, where the regulation does not apply. The figure highlights that travelers' traffic patterns were stable before introducing the RLAH rules and that trends were similar across country groups. The right panel of figure 3, plots the difference in mobile data consumption between RLAH-countries and non-RLAH countries after controlling for country group effects. There is no significant difference in the trends before RLAH comes into effect, but after RLAH, travelers' consumption patterns in countries affected by the regulation more than triples, while mobile data consumption for those traveling to countries not affected by RLAH remains unchanged.

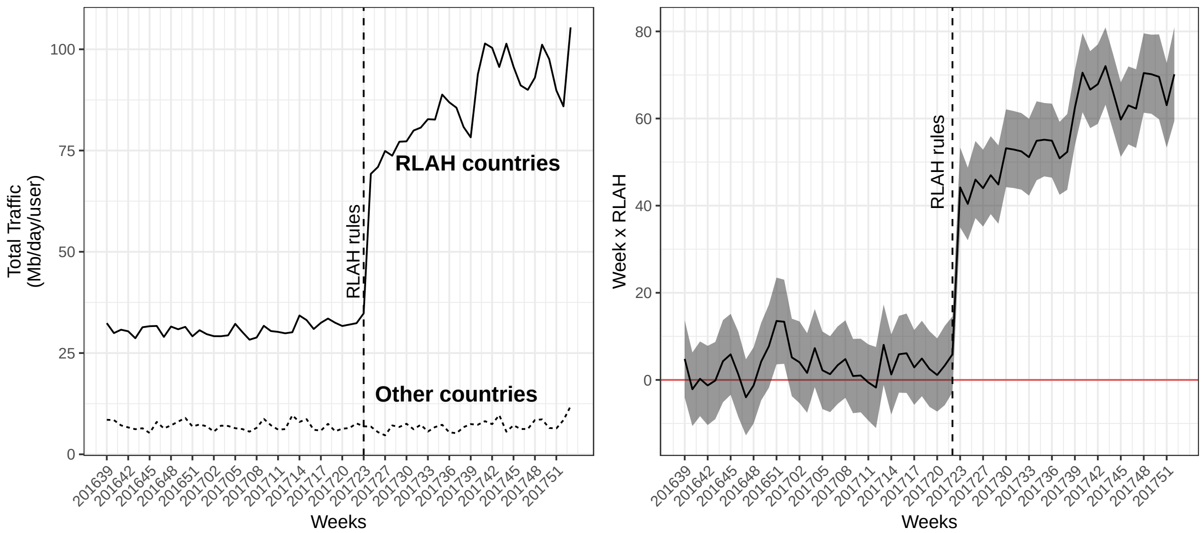

Figure 3: Impact of RLAH rules on mobile data usage
Note: 95% confidence interval for the right plot are computed using results from equation (9)

Table 1 implements our DiD specification model for data from MOBILE. Column (1) provides estimates for a DiD model without fixed-effects. Column (2) includes week and individual fixed effects, and column (3) provides the fully specified estimation of equation (9).[11]

Using columns (1)-(3), we estimate that the RLAH increased daily mobile internet usage by at least 54.17MB ($CI_{95\%}[52.33, 56.02]$), depending on the model specification. Considering the effect of the regulation on European travellers, our estimate suggest an increase of about 177% from a baseline of 24.2MB+7.03MB.

In columns (4)-(6), we repeat this exercise with a log-transformed dependent variable. Results indicate the relative increase in MB consumption for RLAH countries to be around 3.25 ($CI_{95\%}[3.06, 3.42]$) times larger than the one in other countries.[12]

[11] We report estimates for downloads and uploads separately in appendix B.2.
[12] As ($exp(1.178) = 3.25$)

	Dependent variable:					
	Total MB/user/day			log (Total MB/user/day)		
	(1)	(2)	(3)	(4)	(5)	(6)
After x RLAH	55.575***	54.379***	54.176***	1.229***	1.178***	1.178***
	(0.692)	(0.938)	(0.941)	(0.031)	(0.028)	(0.028)
RLAH	24.163***	37.833***		1.914***	2.204***	
	(0.366)	(0.932)		(0.032)	(0.025)	
After	−0.106			0.085**		
	(0.325)			(0.029)		
Constant	7.030***			−0.278***		
	(0.245)			(0.030)		
Week FE	No	Yes	Yes	No	Yes	Yes
User FE	No	Yes	Yes	No	Yes	Yes
Country FE	No	No	Yes	No	No	Yes
Observations	787,870	787,870	787,870	787,870	787,870	787,870
R^2	0.053	0.384	0.385	0.107	0.569	0.571
Adjusted R^2	0.053	0.304	0.305	0.107	0.514	0.515
Residual Std. Error	127.189	109.011	108.922	2.654	1.958	1.954
Note:	+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001 Cluster robust standard errors in () Errors clustered by Mobile phone number					

Table 1: Regression results for the impact of RLAH rules on mobile data usage

To provide additional insight into how RLAH affected the demand for mobile internet, we estimate quantity effects separately for deciles of pre-RLAH consumption intensity. We define deciles according to users' mobile data consumption when traveling within the EEA before RLAH. We present the results of this estimation in table 2 and in figure 4. The analyses show that the relative increase in data usage from the regulation was higher for the lower deciles than the upper deciles. In the first decile, 99.7% of post-RLAH consumption is new, while for the tenth decile, only 25.3% of megabytes consumed would not have been used absent RLAH. This suggests that the regulation had very different effects for different types of users.

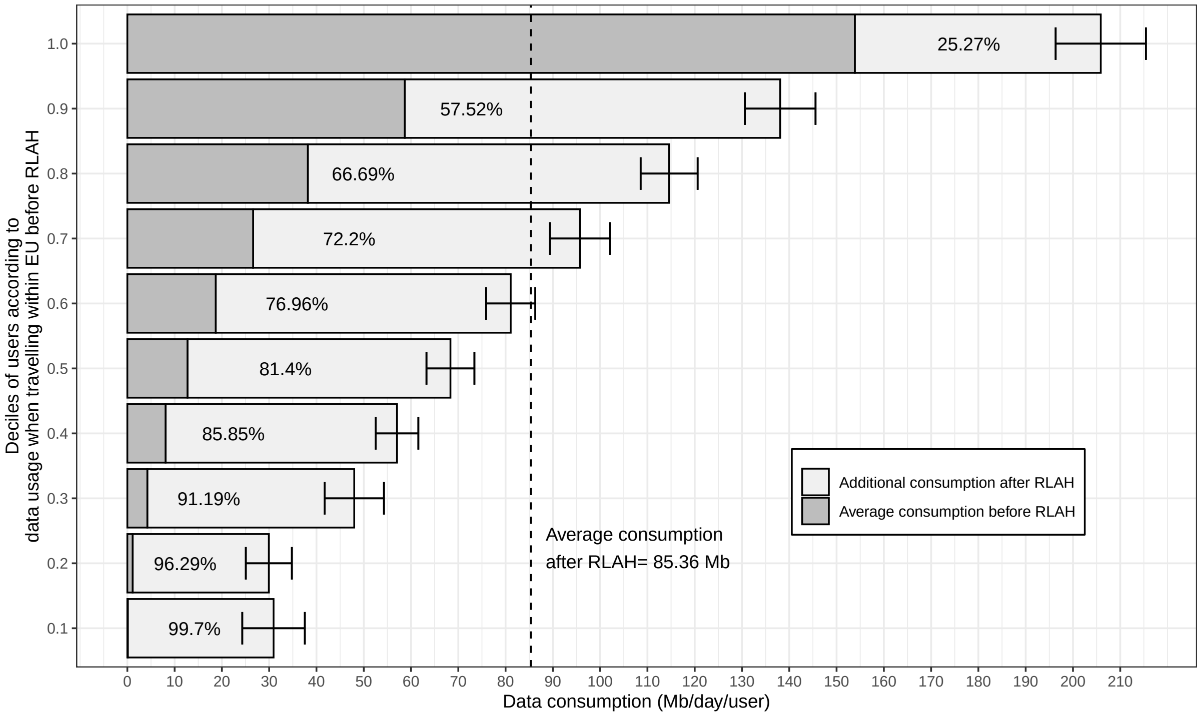

Figure 4: Estimation of data consumption heterogeneity between deciles using MOBILE
Note: 95% confidence interval for the additional effects are computed using results from equation (9) for each deciles.

Table 2: Impact of roaming like at home on mobile data consumption while traveling by decile

	Dependent variable:									
	Total (Mb) by decile									
	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
After × RLAH	30.831***	28.814***	43.768***	48.961***	55.624***	62.398***	69.100***	76.421***	79.427***	52.024***
	(3.599)	(2.648)	(3.399)	(2.424)	(2.714)	(2.791)	(3.411)	(3.253)	(4.054)	(5.200)
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
User FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	64,836	65,876	70,398	83,268	85,899	82,915	77,028	73,785	67,961	63,969
R^2	0.416	0.413	0.369	0.344	0.337	0.322	0.302	0.327	0.322	0.305
Adjusted R^2	0.335	0.334	0.289	0.276	0.269	0.251	0.222	0.247	0.233	0.207
Residual Std. Error	73.193	69.791	75.369	80.395	81.526	94.720	105.116	112.334	129.198	218.472
Ave. before RLAH (Mb)	0.093	1.111	4.226	8.071	12.713	18.684	26.605	38.1689	58.656	153.858

Note: + $p<0.1$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$
Cluster robust standard errors in (); Errors clustered by mobile phone number

5.2 Changes in content consumed

We complement our main results with data from our experiment that contrasts the data consumption profile in a scenario before RLAH to a scenario after RLAH. In particular, we decompose the quantity effect by content types. In figure 5 we show average consumption in MB per day in the scenario that resonates the situation before RLAH and the scenario where RLAH rules apply. More precisely, we use our experiment to compute the average weight of each content providers in the total data consumption before and after the regulation. We show those weight in the percentages of figure 5 and wield them to decompose average consumption in Mb per day computed with the MOBILE database, before and after RLAH. The category of communication sees the most considerable increase, perhaps because users expect to be sharing more data-heavy content such as photos and videos through messaging applications. The smallest increase is in the usage of review platforms, perhaps because there are decreasing returns to this type of information when traveling.

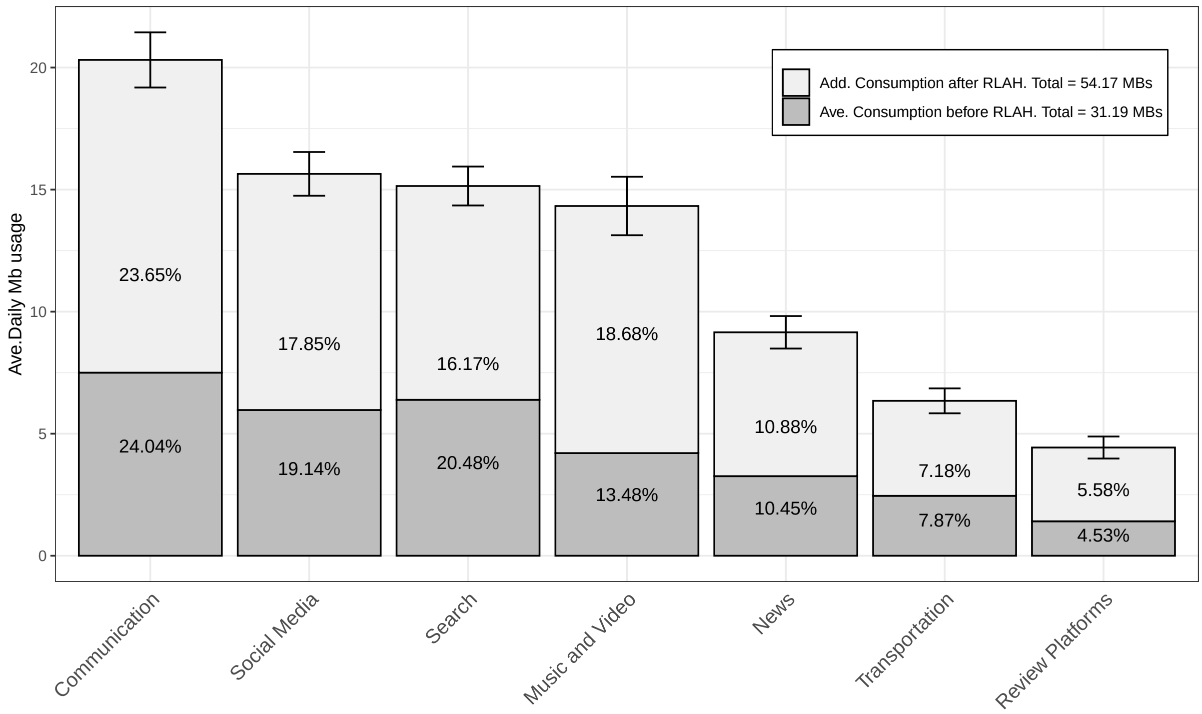

Figure 5: Average daily consumption by content types
Note: Percentages capture (12), i.e. the share for each content type depending on the period. We multiply these shares by the quantities estimated by equation (9) in table 1 - column (3) to find the average daily MB usage for each content type. 95% confidence interval for the additional effects are computed using results from equation (11).

Table 3 shows the increase in the number of experimental subjects that would consume different content categories using mobile data because of RLAH. In this table we also include estimates

Content providers	Online survey	Database
Communication	+39% CI[+35% ,+43%]	–
Search	+43% CI[+38% , +48%]	+13% CI[+1% , +25%]
Social Media	+54% CI[+48% , +60%]	+51% CI[+20% , +81%]
News	+56% CI[+49% , +62%]	+87% CI[+12% , +161%]
Music & Video	+90% CI[+82% , +98%]	–
Transportation	+42% CI[+35% , +48%]	–
Review Platforms	+62% CI[+53% , +72%]	+45% CI[+21% , +68%]

Table 3: Impact of RLAH rules on activity of Content Providers
Note: we compute a ratio and its 95% confidence intervals of the additional unique users after RLAH divided by the number of users in a category before the regulation. Results for the online survey can be found in appendix A.2, while the content providers analysis can be found in appendix C.

for similar statistic calculated from different observational databases that we collected: *Google Trends* for Search, a range of *German publishers* for News, *TripAdvisor* for Review Platforms and *Twitter* for Social Media. The detailed analyses and identification strategies for the numbers reported in the table are described in appendix C.

Together, the online experiment and the analyses of our multiple observational datasets that originate from distinct content providers suggest large increases in consumption of online content while traveling.

5.3 Changes in consumer surplus

We use our quantity estimates and price calibration to establish how RLAH changed consumer surplus that we report in terms of average daily gains.

First, we use the quantity results in column (3) of table 1, and the price scenarios described in section 4.1. Setting the *low domestic price scenario* (with $p_h = 0.0015$), we find that the average daily consumer surplus gain per user increased between €2.73 and EUR €2.82. Using the *high price scenario* (with $p_h = 0.00383$), we find the consumer surplus increased between €2.86 and €2.95.[13]

[13]To compute the range for both cases, we use the mean of MB consumption before the regulation for Europeans travelling within the EEA as an estimate of $\hat{q}_0 = 31.193$. We use the causal estimates and standard errors from column (3) of table 1 as an estimate of $\hat{\delta}$, which gives us $\hat{q}_0 = 31.193$ and $\hat{q}_1 = 31.193 + 54.176 = 85.36$. Overall we have $\hat{q}_0 + \hat{q}_1 = 116.56$

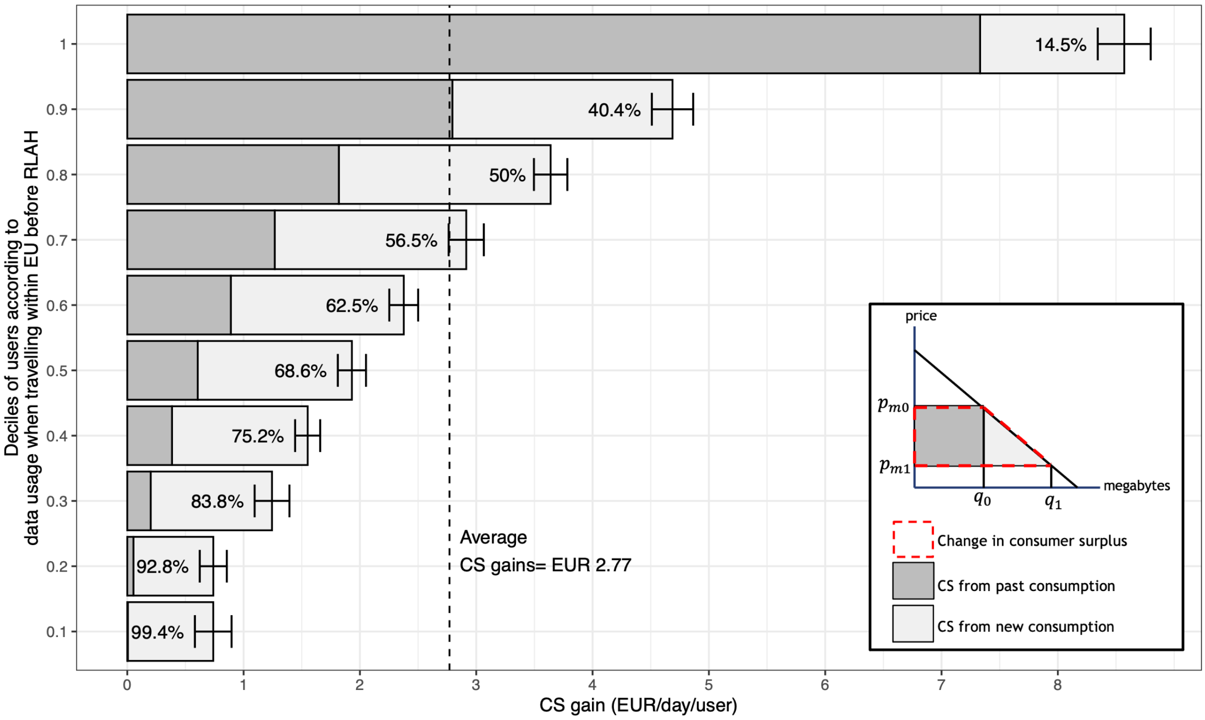

Figure 6: Estimation of Consumer Surplus heterogeneity between deciles using MOBILE
Note: 95% confidence interval for the additional effects are computed using results from equation (9) for each deciles.

Second, we decompose daily consumer surplus gains by user type using our estimates from table 2 and setting prices to the *low domestic price scenario* (with $p_h = 0.0015$). We depict our estimates in figure 6. We find that upper deciles (i.e., heavier mobile data consumers before RLAH) enjoyed higher absolute gains from RLAH. These results originate mostly from the revaluation of mobile data consumption that would occur even without RLAH. Conversely, lower deciles (i.e., users with low or no mobile data consumption in the EEA before the regulation) experience higher relative gains in consumer surplus that originate from usage that would not occur without RLAH. Third, we use the data from our online experiment to provide a perspective of how consumer surplus distributes across different types of content consumption decisions. We set prices to the *low domestic price scenario* (with $p_h = 0.0015$) and provide these results in figure 7. Our analysis suggests that the largest increase in consumer surplus originates from an increase in the consumption of communication services (€0.77 per day/user) and social media (€0.60 per day/user). However, the largest relative increases stem from music and video (88.1%) and review platforms (86.8%).

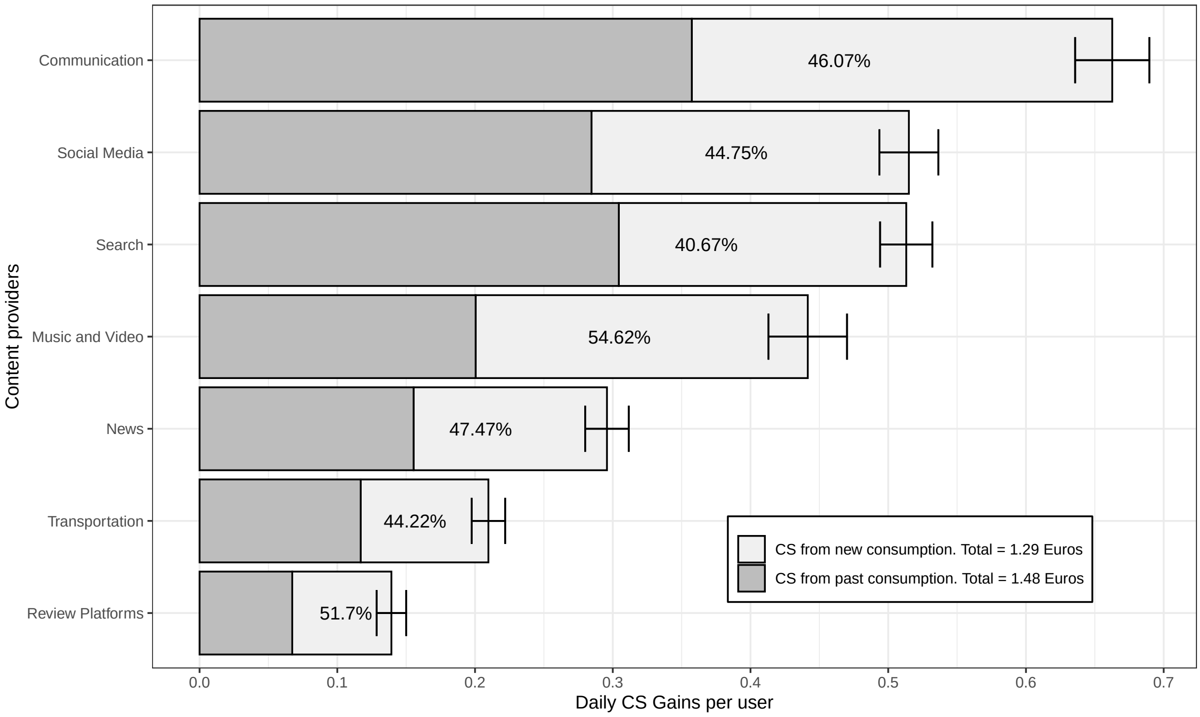

Figure 7: Estimation of Consumer Surplus heterogeneity between content provider types using our online experiment.
Note: 95% confidence interval for the additional effects are computed using results from equation (11).

5.4 Robustness and limitations

Our analysis's main limitation is that before RLAH, consumers may have primarily used other means to access the internet to circumvent high roaming charges when traveling before RLAH. If that is the case, our estimates of increased mobile internet usage and consumer surplus are misleading. We cannot rule out this possibility entirely, but we find some evidence suggesting that such substitution appears to be marginal.

An important substitution scenario is a switch from public Wifi networks towards roaming access. To test this hypothesis, we obtained data from the number of connections to the Open Wifi Hotspot in Milano. Milano is one of Italy's most popular tourist destinations in Europe and collects data on tourists access to the city's public hotspot.[14]

To access Milano's Wifi hotspot, users need to register using their mobile phone number, verified via an SMS code. Milano's Open Wifi Hotspot dataset records the tourists' country of origin based on the phone number, which allows us to separate EU visitors from non-EU visitors. We plot access data to the hotspot from June and July 2017 in figure B.2. Suppose that EU tourists substituted between Wifi usage and mobile data after RLAH rules. In that case, Milano's hotspot

[14] See Statista report.

should have fewer users from EU countries relative to users from non-EU countries after RLAH. However, there seems to be no decrease in the usage of the Wifi hotspot of EU-users.

It may also be that consumers substitute cellular roaming on their smartphones with internet access via Wifi networks on their laptop devices. The Milano hotspot dataset does not have information for this robustness check. However, as mentioned earlier, we also collected data from TripAdvisor and IVW. These datasets allow us to distinguish between internet traffic originating from mobile devices (smartphones and tablets) and computer devices (laptops and desktop computers). Figures C.9 and C.11 show no changes in desktop usage when RLAH comes into effect, suggesting the absence of strong substitution patterns between mobile and desktop internet usage.

Another possibility is that before RLAH, consumers used cheap local prepaid SIM cards to access the mobile internet while traveling. However, data from the International Telecommunication Union (ITU) shows a consistent downward trend in the number of prepaid subscriptions in the EEA countries since 2011 (see figure B.3). Suppose there was a strong substitution between using the mobile internet within the domestic plan's roaming tariff and local prepaid SIM cards before RLAH. In that situation, we would expect to see a drop in the number of local prepaid SIM cards after RLAH. This does not seem to hold. Additionally, figure B.4 shows that the average domestic mobile broadband traffic in the EEA countries follows a similar trend as in non-EEA countries. If the effect of RLAH on tourists' use of domestic SIM cards was substantial, we expected to see a slower growth of domestic traffic in EEA than in non-EEA countries.

Taken together, the evidence that we collected suggests that it is unlikely that the results that we report are just a consequence of consumers switching between free and cheap means to access the internet while traveling towards roaming networks after the RLAH.

6 Discussion of overall welfare effects

So far, we have focused on per individual RLAH induced consumer surplus changes. In this section, we provide an estimate for the total consumer surplus effect of RLAH and turn to a discussion of how network operator surplus and content provider surplus may have changed because of RLAH.

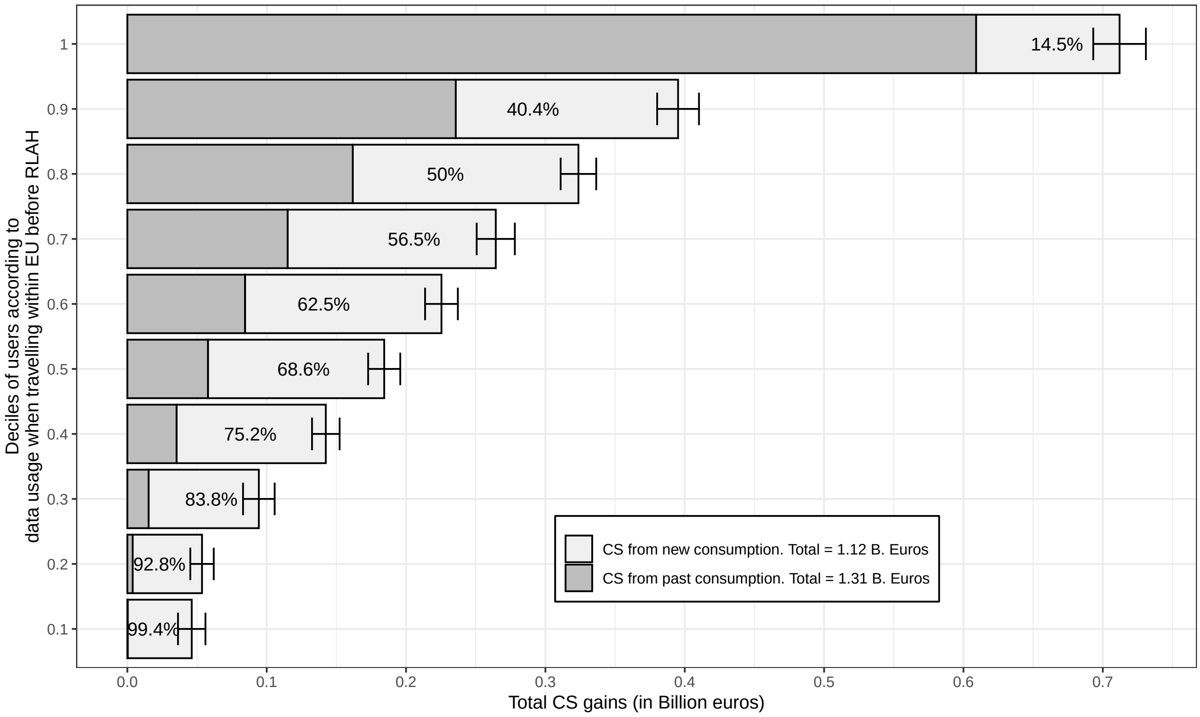

Figure 8: Estimation of Total CS gains for 2017
Note: 95% confidence interval for the additional effects are computed using results from equation (9) for each deciles.

6.1 How much did European consumers gain?

Suppose MOBILE's clients are representative of clients of all network operators in the EU, in mobile internet consumption and travel behavior. In that case, we can use our estimates of consumer surplus changes to calculate aggregate consumer surplus in the EU.

First, we use our estimate of the average consumer surplus increase and tourism statistics from 2017 to provide an EU-wide estimate of the policy's impact.

According to Eurostat, EU countries' residents spent 846.32m nights traveling in the EU between June and December 2017.[15] Assuming that the quantity effects we estimate in section 5.1 are representative across the EU, the total consumer surplus gain of RLAH would be around EUR 2.352B from June 2017 to December 2017. The estimate would more than double for subsequent years because of a longer time horizon and a change in the wholesale price cap - which determines the de-facto data allowance while roaming - that reduced from €c0.385/MB in 2017 to €c0.3/MB in 2018, €c0.225/MB in 2019 and EUR €c0.175/MB in 2020.

To provide additional insights to this analysis, we decompose the surplus gain across deciles of users as in section 5.3. First, for each decile of mobile data use before RLAH, we recover

[15] See https://ec.europa.eu/eurostat/databrowser/view/TOUR_OCC_NIM__custom_61795/default/table?lang=en.

the number of days traveled within Europe from our MOBILE data. Second, we compute each decile's weight by dividing the number of days traveled in the decile by the total number of days of travel. Finally, we use the decile weights to distribute travel nights from tourism data by the corresponding category of mobile data intensity.

We present the results from this computation in figure 8. The figure highlights that consumer surplus gains from the intensive margin represent 54% of consumer surplus gains. Consumer surplus gains from new consumption (extensive margin) represent 46% of the overall consumer surplus gains. Like before, there are higher relative gains for deciles with lower consumption than for deciles with higher consumption of mobile data prior RLAH.

6.2 How were European network operators affected?

In this section, we approximate the change in surplus for network operators after RLAH. Network operators draw two types of revenue from roamers. Roaming charges they levy on their customers during travels (national roamers) and wholesale costs they charge to other network operators because of visitors traveling to the operator's country (roaming visitors). Both these dimensions were affected by the new regulatory environment.

We focus on travelers to/from countries that were affected by the roaming like at home rules because we assume that these were the only group of individuals who affected the network operator's surplus. In that case, the total roaming profit of a network operator is determined by:

$$
\begin{aligned}
\Pi_t^{nop} =& N_t^{nat.r}\pi_t^{nat.r} + N_t^{visitors}\pi_t^{visitors} \\
\pi_t^{nat.r} =& q_t^{nat.r} \times (p_{mt} - w_t) \\
\pi_t^{visitors} =& q_t^{visitors} \times (w_t - c_{ope})
\end{aligned}
$$

π_t^{nop} denotes the average daily roaming profit of a network operator in moment $t \in \{0, 1\}$ with 0 marking the period before RLAH, and 1 marking the period after the regulation. $N_t^{nat.r}$ denotes the number of outgoing roamers that travel abroad and use roaming services. $N_t^{visitors}$ denotes the number of incoming visitors that use the network operator's infrastructure to access mobile data. $\pi_t^{nat.r}$ denotes the average daily roaming profit per national roaming customer, and $\pi_t^{visitors}$ denotes the average daily roaming profit per roaming visitor. $q_t^{nat.r}$ is the average amount of MBs that national roamers use per day while traveling and $q_t^{visitors}$ is the average amount of MBs that roaming visitors consume per day. p_{mt} is the retail price per megabyte for roaming nationals. w_t is the wholesale price per megabyte for both national roamers and roaming visitors. In our approximation, we assume that wholesale prices are the same for all network operators in all countries affected by RLAH. Finally, c_{ope} are operating costs per megabyte in €c.

We set w_t as the wholesale price cap before/after RLAH, the retail price using the same assumptions introduced in section 4.1, and quantity estimates based on table 1.

With these parameters, we calculate the changes in marginal revenue from national roamers and roaming visitors.

For national roamers, the change in the marginal revenue ($\pi_0^{nat.r} - \pi_1^{nat.r}$) depends on the domestic price per megabyte. For domestic prices lower than €c0.385, the marginal revenue of national roamers decreased by $31.19p_h + 32.86$. When domestic prices were higher than €c0.385, the marginal revenue from national roamers decreased by $65.72 - 54.17p_h$. For roaming visitors, the difference in marginal revenue ($\pi_0^{visitors} - \pi_1^{visitors}$) changed by $€c90.22 + 54.17 \times c_{ope}$.

Suppose RLAH change mobile data usage but not the likelihood of travel (which we think is a reasonable assumption). In that case, and with the parameterization we laid out above, the profit change for a network operator because of RLAH will be:

$$\Delta\Pi = \begin{cases} N_1^{nat.r} \times [31.19p_h + €c32.86] + N_1^{visitors}[€c90.22 + 54.17 \times c_{ope}] & \text{if } p_h \leq 0.385 \\ N_1^{nat.r} \times [€c65.72 - 54.17p_h] + N_1^{visitors}[€c90.22 + 54.17 \times c_{ope}] & \text{if } p_h > 0.385 \end{cases} \tag{14}$$

We can use equation 14 to map the 2017 post-regulation gains for any European network operator depending on their network operating cost (c_{ope}), their domestic price per megabyte p_h, the number of roaming visitors $N_1^{visitors}$ and number of national roamers $N_1^{nat.r}$.

According to the distribution of domestic prices and operating costs across European countries (see European Commission (2016)), operators' marginal profit decreased visitors more than national roamers.

For the average operating cost c_{ope} and domestic price p_h across European countries in 2017 (c_{ope} =€c0.4/MB and p_h =€c0.24/MB),[16] the daily marginal revenue loss because of RLAH amounted to €c40.3 per national roamer and €c111.9 per visitor due to RLAH.

Operators in countries with a large number of incoming visitors lost more from the regulation.

[16] See reports from the European Commission (2016).

Additionally, network operators with higher operating costs were also more penalized by RLAH. These heterogeneous impacts are likely to have been substantial given the heterogeneity in roaming operating cost across countries in Europe (European Commission, 2016).

6.3 Who else was affected by RLAH?

As we have shown in section 5.2, consumers split their additional mobile internet allowance by different content providers online. According to our survey's results, search engines, social networks, and music and video services benefited the most from the added data allowances induced by RLAH.

In Europe, a few large firms dominate these content categories. The immediate consequence is that it is likely that the RLAH regulation transferred surplus from European network operators towards multinational content providers such as Google, Facebook, Instagram, and other tech giants, that operate in the markets that travelers seem to value the most.

For example, 90% of Europeans with access to the internet use Google as a search engine,[17] making it likely that travelers' online search activity will benefit Google directly. If RLAH increases users' search behavior, Google certainly captures part of the enlarged pie.

Using Google's financial report and our estimates for how much search activity increased, we can "guesstimate" that Google may have benefited in the order of €M2.5 in 2017 because of roaming like at home.[18] Similar arguments allow us to identify other net benefit receivers from RLAH. Based on our survey estimates, a firm such as Youtube could have profited €M1, and firms like Spotify or Facebook could have generated additional revenues of €M5 from EU consumers only for 2017.

Comparing our estimates of consumer surplus gains and operator losses, it seems clear that RLAH was welfare enhancing. However, it was also redistributive, and it is not evident that redistribution occurred in a way that the regulators anticipated.

[17] See https://gs.statcounter.com/search-engine-market-share/all/europe

[18] We recover that 76% of the European population has access to Internet in 2017, among which 90% are using Google as their search engine. We establish Google revenue from European travels between June and December 2017 equal to $846.32M \times €0.041 \times 0.76 \times 0.9 = €M23.73$. As we have established an increase in Google search activity of European travellers whn they travel within Europe about 13% due to the regulation (see table 3), this suggests that Google gained around €M2.5 for 2017 from its search engine due to the RLAH regulation.

7 Implications for policy

The regulation's initial goal was to increase consumer surplus, reduce the digital divide for travelers, and make the first step toward a unified digital European market. Our results suggest that the regulation succeeded in increasing the total consumer surplus and allowing more travelers to access the Internet. However, our analyses also underline several other nuanced results.

First, we show that a price reduction of internet access for consumers generates an increase in surplus for content providers at the expense of revenue from internet service providers that also bear the cost of an increased volume of internet traffic. Furthermore, our online experiments show that consumers do not value content homogeneously, and different content providers benefited differently from RLAH. These effects and surplus transfers are at the core of the net neutrality debate (Greenstein et al., 2016).

With a net neutrality regime, policies like RLAH transfer ISP's revenue directly to other market agents, which triggers the possibility that the regulation could have an ambiguous mid-term effect on the digital divide. On the one hand, the RLAH might have reduced the digital divide between EU travelers, allowing more users to access the Internet. On the other hand, weakening ISP's revenue to the benefit of content producers may reduce ISP's incentives to invest in bandwidth quality in the future (Pil Choi and Kim, 2010).

Another nuanced aspect of our results is that even if the RLAH resulted in increased mobile data usage by EU citizens and decreased travelers' digital divide, the increased use seems to benefit large and non-EU content providers the most. While the regulation increased consumer surplus, it also reinforced the non-EU platform position in the EU market. This may strike with the current anti-trust investigation the European Commission is conducting against big US firms that benefited from RLAH rules.

Finally, a large span of the literature has been studying the fact that access to information significantly shapes individual preferences (Gentzkow, 2006; Gavazza et al., 2019; Durante and Knight, 2012) and our results show that one of the main use of cheaper internet access while traveling is media content. EU travelers can now access local or home country-related news more easily while traveling. These subtle effects highlight that regulation can affect trade patterns of digital goods and services even within a trade union like the EU and highlight the connection between otherwise distinct regulatory aspects of telecommunications and media policy.

8 Conclusion

Since the introduction of RLAH rules on the 15h June 2017, European citizens can enjoy free data roaming when traveling in the EEA. These rules are the outcome of several telecom regulation waves to provide users with more robust mobile data capabilities to reinforce the European digital single market. More precisely, the rules sharply reduced prices for mobile internet access and allowed a significant number of European citizens to access mobile data while traveling.

In this paper, we analyze the consumer surplus effects of such price decreases. We partner with a major European mobile carrier which gives us access to a dataset of mobile internet usage statistics for 90,000 anonymized users from September 2016 to December 2017. We show that the daily consumption of mobile data per user increased by 177% because of roaming like at-home rules. This quantity increase translates into a consumer surplus gain of at least €2.73 per user per day. We then decompose this surplus gain by different consumer segments. We show that users with high data consumption before the regulation experienced higher absolute gains than other users. These gains originated mainly from the revaluation of past data consumption levels and not from additional data consumption. Conversely, users with very low data consumption before the roaming like at-home rules, obtained most of their surplus gains from increased mobile data use.

We also discuss the effect of RLAH rules with respect to different content providers. We do so using an online experiment designed to study how consumers allocate the additional data allowance they obtained from roaming like at-home rules across different online content types. Our results suggest that consumer surplus gains from RLAH derive from a combination of new data consumption content types such as Music & video, and Review Platforms, and revaluation of previous consumption levels for other categories such as Communication, Social Media and Search. We complement our online experiment results with an evaluation of how RLAH affected user activity in Google search, the IVW, TripAdvisor, and Twitter.

Our results suggest that RLAH increased total consumer surplus and allowed more travelers to access the Internet. At the same time, content providers also benefited from RLAH. These gains outweigh the losses of internet service providers, facing lower revenues and higher volumes of mobile internet use.

Our results have broad implications. Namely, we discuss how RLAH links to net neutrality and the market position of non-EU services within the EU, as well as media policy aspects.

Our study has several limitations. First, we cannot rule out completely that EU citizens substituted other free means of accessing the Internet with roaming. Second, we cannot observe individual-level prices and we approximate the arc-elasticity of demand in our estimations. Finally, we do not observe actual internet use of clients of MOBILE and used online experiments to study how consumers spend the additional data allowance provided by RLAH.

Finally, we think our results may stimulate more research on the unintended effect of data access. For example, cities' shape and organization may have been impacted by users having easier access to both information platforms and transportation means. As places of interest can be more easily detected and reached when one has access to the Internet, firms may change their investment in quality and location choice accordingly. More broadly, if better access to information changed users' optimal choices, firms may also undertake different strategic decisions to adapt, which provides interesting avenues for future research.

References

Aguiar, L., and Waldfogel, J. (2018). "Quality Predictability and the Welfare Benefits from New Products: Evidence from the Digitization of Recorded Music." *Journal of Political Economy*, *126*(2), 492–524.

Allen, R. G. D., and Lerner, A. P. (1934). "The concept of arc elasticity of demand." *The Review of Economic Studies*, *1*(3), 226–230.

Baake, P., and Wagner, L. (2018). "Eu roaming regulation: theoretical model suggests a positive assessment." *DIW Weekly Report*, *8*(6), 59–64.

Baker, S. R., and Fradkin, A. (2017). "The impact of unemployment insurance on job search: Evidence from google search data." *Review of Economics and Statistics*, *99*(5), 756–768.

Briglauer, W., Dürr, N. S., Falck, O., and Hüschelrath, K. (2019). "Does state aid for broadband deployment in rural areas close the digital and economic divide?" *Information Economics and Policy*, *46*, 68–85.

Brynjolfsson, E. (1996). "The contribution of information technology to consumer welfare." *Information Systems Research*, *7*(3), 281–300.

Brynjolfsson, E., Hu, Y. J., and Smith, M. D. (2003). "Consumer surplus in the digital economy:

Estimating the value of increased product variety at online booksellers." *Manage. Sci.*, *49*(11), 1580–1596.

Cardona, M., Kretschmer, T., and Strobel, T. (2013). "Ict and productivity: conclusions from the empirical literature." *Information Economics and Policy*, *25*(3), 109 – 125.

Chillemi, O., Galavotti, S., and Gui, B. (2019). "The impact of data caps on mobile broadband internet access: A welfare analysis." *Information Economics and Policy*, 100843.

Cleveland, R. B., Cleveland, W. S., McRae, J. E., and Terpenning, I. (1990). "Stl: A seasonal-trend decomposition." *Journal of official statistics*, *6*(1), 3–73.

Czernich, N., Falck, O., Kretschmer, T., and Woessmann, L. (2011). "Broadband infrastructure and economic growth." *The Economic Journal*, *121*(552), 505–532.

Durante, R., and Knight, B. (2012). "Partisan control, media bias, and viewer responses: Evidence from berlusconi's italy." *Journal of the European Economic Association*, *10*(3), 451–481.

Easley, R. F., Guo, H., and Krämer, J. (2018). "From net neutrality to data neutrality: A techno-economic framework and research agenda." *Information Systems Research*, *29*(2), 253–272.

Einav, L., Levin, J., Popov, I., and Sundaresan, N. (2014). "Growth, adoption, and use of mobile e-commerce." *American Economic Review*, *104*(5), 489–94.

European Commission (2016). "Impact Assessment – Proposal for a Regulation of the European Parliament and of the Council amending Regulation (EU) No 531/2012 as regards rules for wholesale roaming markets."

European Commission (2017). "The end of roaming charges within the eu – flash eurobarometer 454 report."

Fabrizi, S., and Wertlen, B. (2008). "Roaming in the mobile internet." *Telecommunications Policy*, *32*(1), 50–61.

Fong, N. M., Fang, Z., and Luo, X. (2015). "Geo-conquesting: Competitive locational targeting of mobile promotions." *Journal of Marketing Research*, *52*(5), 726–735.

Gavazza, A., Nardotto, M., and Valletti, T. (2019). "Internet and Politics: Evidence from U.K. Local Elections and Local Government Policies." *The Review of Economic Studies*, *86*(5), 2092–2135.

Genakos, C., and Valletti, T. (2011). "Testing the "waterbed" effect in mobile telephony." *Journal of the European Economic Association*, *9*(6), 1114–1142.

Gentzkow, M. (2006). "Television and voter turnout." *The Quarterly Journal of Economics*, *121*(3), 931–972.

Ghose, A., Goldfarb, A., and Han, S. P. (2013). "How is the mobile internet different? search costs and local activities." *Information Systems Research*, *24*(3), 613–631.

Ghose, A., and Han, S. P. (2011). "An empirical analysis of user content generation and usage behavior on the mobile internet." *Management Science*, *57*(9), 1671–1691.

Goldfarb, A., and Tucker, C. (2019). "Digital economics." *Journal of Economic Literature*, *57*(1), 3–43.

Greenstein, S., Peitz, M., and Valletti, T. (2016). "Net neutrality: A fast lane to understanding the trade-offs." *Journal of Economic Perspectives*, *30*(2), 127–50.

Hausman, J. (1981). "Exact consumer's surplus and deadweight loss." *American Economic Review*, *71*(4), 662–76.

Hausman, J. A. (1997). "Valuing the effect of regulation on new services in telecommunications." *Brookings papers on economic activity. Microeconomics*, *1997*, 1–54.

Hausman, J. A., and Leonard, G. K. (2002). "The competitive effects of a new product introduction: A case study." *The Journal of Industrial Economics*, *50*(3), 237–263.

Hjort, J., and Poulsen, J. (2019). "The arrival of fast internet and employment in africa." *American Economic Review*, *109*(3), 1032–79.

Infante, J., and Vallejo, I. (2012). "Regulation of international roaming in the european union—lessons learned." *Telecommunications Policy*, *36*(9), 736–748.

Jullien, B., Rey, P., and Sand-Zantman, W. (2013). "Termination fees revisited." *International Journal of Industrial Organization*, *31*(6), 738–750.

Li, W., and Xu, L. C. (2004). "The impact of privatization and competition in the telecommunications sector around the world." *The Journal of Law and Economics*, *47*(2), 395–430.

Maillé, P., and Tuffin, B. (2017). "Enforcing free roaming among eu countries: An economic analysis." In *2017 13th International Conference on Network and Service Management (CNSM)*, 1–4, IEEE.

Mang, C. F., Piper, L. A., and Brown, N. R. (2016). "The incidence of smartphone usage among tourists." *International Journal of Tourism Research*, *18*(6), 591–601.

Nardotto, M., Valletti, T., and Verboven, F. (2015). "Unbundling the incumbent: Evidence from uk broadband." *Journal of the European Economic Association*, *13*(2), 330–362.

OECD (2011). "International mobile data roaming." *OECD Digital Economy Papers*, *180.*

Pil Choi, J., and Kim, B.-C. (2010). "Net neutrality and investment incentives." *The RAND Journal of Economics*, *41*(3), 446–471.

Preis, T., Moat, H. S., and Stanley, H. E. (2013). "Quantifying trading behavior in financial markets using google trends." *Scientific reports*, *3*, 1684.

Prieger, J. E. (2013). "The broadband digital divide and the economic benefits of mobile broadband for rural areas." *Telecommunications Policy*, *37*(6-7), 483–502.

Schleife, K. (2010). "What really matters: Regional versus individual determinants of the digital divide in germany." *Research Policy*, *39*(1), 173–185.

Silva, S., Badasyan, N., and Busby, M. (2018). "Diversity and digital divide: Using the national broadband map to identify the non-adopters of broadband." *Telecommunications Policy*, *42*(5), 361–373.

Spruytte, J., Van der Wee, M., de Regt, M., Verbrugge, S., and Colle, D. (2017). "International roaming in the eu: Current overview, challenges, opportunities and solutions." *Telecommunications Policy*, *41*(9), 717–730.

Viard, V. B., and Economides, N. (2015). "The effect of content on global internet adoption and the global "digital divide"." *Management Science*, *61*(3), 665–687.

Vogelsang, I. (2003). "Price regulation of access to telecommunications networks." *Journal of Economic Literature*, *41*(3), 830–862.

Waverman, L., and Sirel, E. (1997). “European telecommunications markets on the verge of full liberalization.” *Journal of Economic Perspectives*, *11*(4), 113–126.

Willig, R. D. (1976). “Consumer’s surplus without apology.” *The American Economic Review*, *66*(4), 589–597.

Xu, J., Forman, C., and Hu, Y. J. (2019). “Battle of the internet channels: How do mobile and fixed-line quality drive internet use?” *Information Systems Research*, *30*(1), 65–80.

Xu, K., Chan, J., Ghose, A., and Han, S. P. (2017). “Battle of the channels: The impact of tablets on digital commerce.” *Management Science*, *63*(5), 1469–1492.

Zucchini, L., Claussen, J., and Trüg, M. (2013). “Tariff-mediated network effects versus strategic discounting: Evidence from german mobile telecommunications.” *International Journal of Industrial Organization*, *31*(6), 751 – 759.

A Appendix: Online experiment

A.1 Descriptive statistics

Table A.1: Descriptive statistics for the online experiment dataset. The table exhibit data only for the 1233 users that use both a positive amount of megabyte before and after.

	Scenario	Variable	Avg. (mb)	Std. (mb)
1	Before RLAH	Search	2.234	98.384
2	Before RLAH	Social Media	2.088	110.878
3	Before RLAH	News	1.140	63.060
4	Before RLAH	Transportation	0.858	55.906
5	Before RLAH	Music and Video	1.471	97.270
6	Before RLAH	Communication	2.623	99.220
7	Before RLAH	Review Platforms	0.494	46.319
8	Before RLAH	Unused Megabytes	1.061	98.667
9	After RLAH	Search	21.058	970.488
10	After RLAH	Social Media	22.876	1,081.179
11	After RLAH	News	13.807	787.541
12	After RLAH	Transportation	9.222	612.732
13	After RLAH	Music and Video	23.225	1,434.399
14	After RLAH	Communication	30.159	1,358.691
15	After RLAH	Review Platforms	6.989	524.343
16	After RLAH	Unused Megabytes	9.390	909.418

Table A.2: Descriptive statistics for the online experiment dataset. The table exhibit data for all the 1856 users.

	Scenario	Variable	Avg. (mb)	Std. (mb)
1	Before RLAH	Search	1.522	86.421
2	Before RLAH	Social Media	1.447	96.355
3	Before RLAH	News	0.781	53.982
4	Before RLAH	Transportation	0.583	47.215
5	Before RLAH	Music and Video	1.011	82.695
6	Before RLAH	Communication	1.784	89.037
7	Before RLAH	Review Platforms	0.347	39.865
8	Before RLAH	Unused Megabytes	4.312	217.014
9	After RLAH	Search	21.191	945.122
10	After RLAH	Social Media	21.671	1,036.447
11	After RLAH	News	12.518	723.958
12	After RLAH	Transportation	8.884	579.335
13	After RLAH	Music and Video	20.200	1,266.852
14	After RLAH	Communication	28.048	1,268.359
15	After RLAH	Review Platforms	6.739	545.821
16	After RLAH	Unused Megabytes	15.102	1,347.423

Table A.3: Demographic statistics for the online experiment dataset. The table exhibit data for all the 1856 users.

variable	category	perc.
gender	Female	0.438
gender	Male	0.557
gender	I don't want to answer	0.003
gender	Other	0.003
age	Under 18	0.002
age	18 - 24	0.234
age	35 - 44	0.239
age	25 - 34	0.384
age	45 - 54	0.095
age	55 - 64	0.039
age	65 - 74	0.004
age	I don't want to answer	0.004
degree	Less than high school degree	0.034
degree	Bachelor's degree	0.289
degree	High school graduate	0.236
degree	Master's degree	0.246
degree	Professional degree	0.027
degree	Some college but no degree	0.129
degree	Doctoral degree	0.022
degree	I don't want to answer	0.016
occupation	Unemployed	0.089
occupation	Student	0.221
occupation	Managers	0.073
occupation	Professionals	0.220
occupation	Clerical Support Worker	0.055
occupation	Other	0.111
occupation	I don't want to answer	0.056
occupation	Craft and Related Trade Workers	0.029
occupation	Service and Sales Workers	0.081
occupation	Armed Forces Occupations	0.009
occupation	Elementary Occupations	0.020
occupation	Agricultural, Forestry and Fishery Workers	0.009
occupation	Plant and Machine Operators	0.016
occupation	Retired	0.012
occupation	None	0.001

A.2 Results

Table A.4: Estimation results of daily data consumption per user for the 1233 users that use both a positive amount of megabyte before and after. Column (1) and (2) estimate the global effect of being after te regulation while column (3) estimates equation (11).

	Dependent variable:		
	Data consumption (Mb)		
	(1)	(2)	(3)
After	116.430***	116.430***	
	(3.249)	(9.504)	
After × Communication			27.537***
			(1.238)
After × Search			18.824***
			(0.873)
After × Social Media			20.788***
			(0.982)
After × News			12.667***
			(0.729)
After × Music and Video			21.755***
			(1.311)
After × Transportation			8.364***
			(0.559)
After × Review Platforms			6.495***
			(0.494)
Constant	10.907***		
	(2.297)		
Individual FE	No	Yes	Yes
Cont. provider FE	–	–	Yes
Observations	2,466	2,466	17,252
R^2	0.343	0.371	0.257
Adjusted R^2	0.342	0.365	0.199
Residual Std. Error	80.667	79.286	23.358
Note:	*p<0.1; **p<0.05; ***p<0.01 Cluster robust standard errors in () Errors clustered by individual		

We also modify equation (11) to capture how the adoption of content typed changes with the regulation

$$Adoption_{ict} = \sum Content_{ict}(\gamma_c + \delta_c After_t) + a_i + v_{ict}, \tag{15}$$

where $Adoption_{ict}$ is a dummy variable that is equal to 1 when user i spend positive amount in content service c at time t. This allow us to compare our survey results with observational data from exemplary content provider we describe below. In this case, we only clean the database from outliers, which trims our dataset to 1856 unique users. Results are displayed in table A.5, and are used to compute percentage increases of the first column of table 3 related to the online survey.

Table A.5: Estimation results of online service adoption on all 1856 users in the online experiment database. It estimates equation (15) with Adoption of the service (0 or 1) as a dependent variable.

	Dependent variable:
	Adoption
After × Communication	0.235***
	(0.013)
After × Search	0.238***
	(0.014)
After × Social Media	0.246***
	(0.013)
After × News	0.224***
	(0.014)
After × Music and Video	0.280***
	(0.013)
After × Transportation	0.154***
	(0.013)
After × Review Platforms	0.158***
	(0.012)
Individual FE	Yes
Cont. provider FE	Yes
Observations	25,984
R^2	0.349
Adjusted R^2	0.299
Residual Std. Error	0.418
Note:	*p<0.1; **p<0.05; ***p<0.01
	Cluster robust standard errors in ()
	Errors clustered by individual

B Appendix: MOBILE dataset

B.1 Descriptive statistics

Table B.6: Descriptive statistics for mobile internet use for RLAH mobile Dataset

	Location	Period	Variable	#Countries	#Clients	Avg.	Std.	Q05	MED	Q95
1	Home	Before	Downloads (Mb/day)	1	90,345	77.493	129.993	0.942	34.129	300.998
2	Home	Before	Uploads (Mb/day)	1	90,345	6.723	8.605	0.287	3.933	22.826
3	Home	Before	Tot traf (Mb/day)	1	90,345	84.216	136.359	1.317	38.986	320.696
4	Home	Before	Number of days	1	90,345	5.416	2.029	1	7	7
5	Abroad within EU	Before	Downloads (Mb/day)	29	84,006	27.435	71.437	0.005	8.243	106.823
6	Abroad within EU	Before	Uploads (Mb/day)	29	84,006	4.082	8.489	0.005	1.436	16.475
7	Abroad within EU	Before	Tot traf (Mb/day)	29	84,006	31.517	76.525	0.020	10.297	122.341
8	Abroad within EU	Before	Number of days	29	84,006	2.212	1.652	1	2	6
9	Abroad outside EU	Before	Downloads (Mb/day)	104	15,227	5.786	23.741	0.0004	1.240	19.806
10	Abroad outside EU	Before	Uploads (Mb/day)	104	15,227	1.287	3.819	0.0003	0.358	5.102
11	Abroad outside EU	Before	Tot traf (Mb/day)	104	15,227	7.073	26.123	0.001	1.714	25.232
12	Abroad outside EU	Before	Number of days	104	15,227	1.679	1.221	1	1	4
13	Home	After	Downloads (Mb/day)	1	90,345	92.233	145.130	1.259	42.691	351.924
14	Home	After	Uploads (Mb/day)	1	90,345	8.049	9.963	0.358	4.756	27.113
15	Home	After	Tot traf (Mb/day)	1	90,345	100.281	152.354	1.726	48.626	375.113
16	Home	After	Number of days	1	90,345	5.384	2.076	1	7	7
17	Abroad within EU	After	Downloads (Mb/day)	29	84,260	77.489	155.677	0.043	27.210	312.542
18	Abroad within EU	After	Uploads (Mb/day)	29	84,260	8.800	13.709	0.061	3.849	35.096
19	Abroad within EU	After	Tot traf (Mb/day)	29	84,260	86.289	164.373	0.133	32.347	341.554
20	Abroad within EU	After	Number of days	29	84,260	2.753	1.996	1	2	7
21	Abroad outside EU	After	Downloads (Mb/day)	106	15,253	5.571	24.556	0.0005	1.270	17.799
22	Abroad outside EU	After	Uploads (Mb/day)	106	15,253	1.373	4.523	0.001	0.365	5.296
23	Abroad outside EU	After	Tot traf (Mb/day)	106	15,253	6.944	27.356	0.001	1.776	22.491
24	Abroad outside EU	After	Number of days	106	15,253	1.584	1.138	1	1	4

B.2 Results

In table B.7, we report results from estimating equation (9) for each of the outcomes of interest (total traffic, download traffic, and upload traffic).

Table B.7: Impact of roaming like at home on mobile data consumption while traveling

	Dependent variable:								
	Total (Mb)			Downloads (Mb)			Uploads (Mb)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
After × RLAH	55.575***	54.379***	54.176***	50.877***	49.898***	49.733***	4.698***	4.480***	4.443***
	(0.692)	(0.938)	(0.941)	(0.650)	(0.880)	(0.884)	(0.063)	(0.087)	(0.087)
RLAH	24.163***	37.833***		21.403***	32.560***		2.760***	5.272***	
	(0.366)	(0.932)		(0.329)	(0.866)		(0.048)	(0.094)	
After	−0.106			−0.185			0.079^{+}		
	(0.325)			(0.297)			(0.043)		
Constant	7.030***			5.740***			1.290***		
	(0.245)			(0.214)			(0.039)		
Week FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
User FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Country FE	No	No	Yes	No	No	Yes	No	No	Yes
Observations	787,870	787,870	787,870	787,870	787,870	787,870	787,870	787,870	787,870
R^2	0.053	0.384	0.385	0.049	0.372	0.373	0.053	0.424	0.425
Adjusted R^2	0.053	0.304	0.305	0.049	0.291	0.292	0.053	0.350	0.351
Residual Std. Error	127.189	109.011	108.922	120.193	103.787	103.707	11.239	9.317	9.309

Note: + p<0.1; * p<0.05; ** p<0.01; *** p<0.001

Cluster robust standard errors in (); Errors clustered by Mobile phone number

B.3 Parallel trend assumption for the MOBILE's dataset

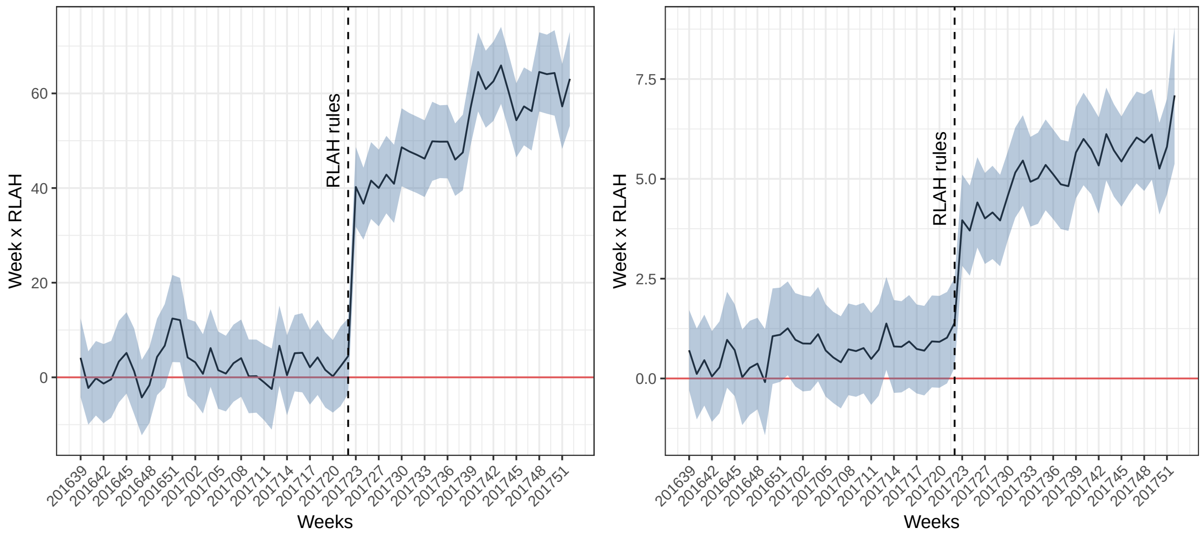

Figure B.1: Regression results for parallel trend assumption of downloads (left) and uploads (right). 95% confidence intervals are computed in using a variation of column (6) and (9) in table B.7.

B.4 Robustness checks

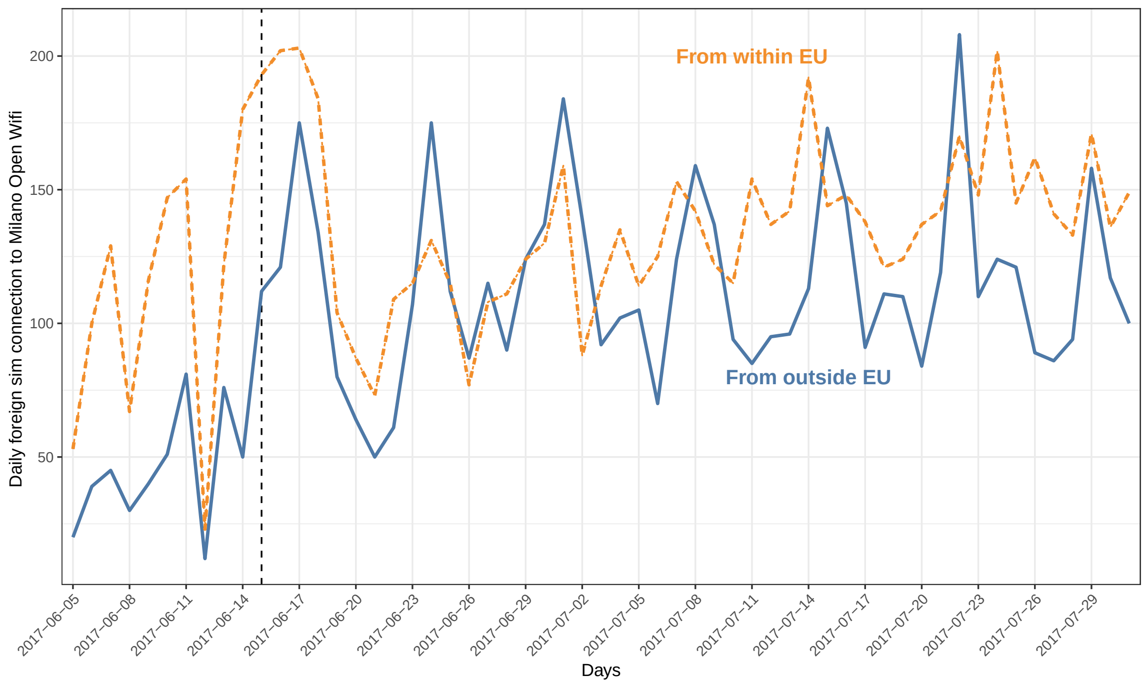

Figure B.2: Number of Milano Wifi Hotspot connection accross time and sim origin
Source: Open Data - Comune de Milano.

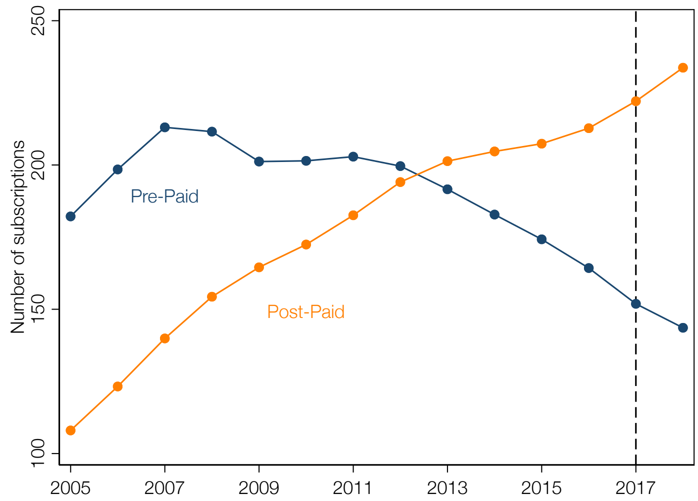

Figure B.3: Number of cellular telephone subscriptions, pre-paid and post-paid, EEA
Source: ITU World Telecommunication/ICT Indicators Database, 2019.

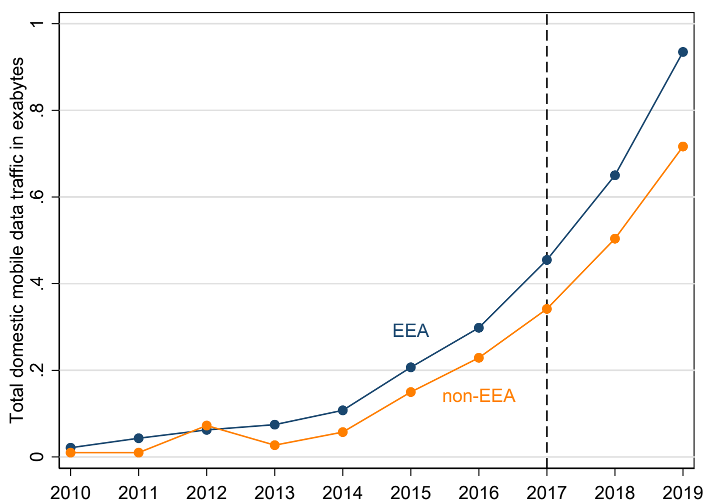

Figure B.4: Domestic mobile broadband traffic in exabytes, average of EEA and US/CA/BR.
Source: ITU World Telecommunication/ICT Indicators Database, 2019.

C Appendix: Detailed content provider analysis

C.1 Google Trends

C.1.1 Data

To study how RLAH changes information consumption behavior, we use data from Google Search Trends Tool (GSTT) to track indexes of the intensity of searching for a set of specific travel-related keywords, before and after the RLAH rules, in and out of the RLAH territories.

GSTT provides a broad keyword search match between the search query and the actual search. It bundles together several expressions that contain the keywords of interest. A search activity index obtained from GSTT is conditional on the date range selected and the geography of interest filtered in the tool. Google calculates the index, dividing the number of searches in the keyword of interest to all searches in the filtered geography and normalizing it 100. [19]

Google compiles the search indexes from random samples of the search history for each time range and geography. Multiple queries to the GSTT yield different search index results for the same keywords, geographies, and period. The variability depends on the search volume of the region and the search volumes of the keywords of interest. Keywords and geographies with more search will have more stable query results.

For each keyword in our dataset, we queried Google Trends 12 times in three consecutive weeks, and we used the average of these 12-time series as our final time series of interest. A similar data collection procedure was used by Baker and Fradkin (2017) and Preis et al. (2013).

We focused on travel-related keywords to link the use of mobile internet while traveling with Google search activity. GSTT does not separate keyword searches from mobile devices from keyword searches originating in desktop or laptop devices. Likewise, GSTT does allow separating search queries of travellers from those of residents. To achieve this, we focused on the Italian language.

Our GSTT dataset contains 71,808 observations at the country-keyword-month level. It covers the period from January 2004 to December 2019. It contains GSTT indexes for 19 different keywords across nine countries within the EU and 11 countries outside the EU.

We present summary statistics of the GSTT dataset in table C.8

[19] The reference constant for the normalization is the highest search index value for the period requested in the GSTT

Table C.8: Descriptive statistics for Google Trend dataset.

			Outside EU				Within EU			
	period	keyword	#Countries	#Months	Avg.	Std.	#Countries	#Months	Avg.	Std.
1	Before	vicino	11	161	8.070	14.860	9	161	9.600	12.730
2	Before	previsioni	11	161	5.550	13.820	9	161	19.690	19.460
3	Before	pioggia	9	161	5.760	14.240	9	161	10.780	16.630
4	Before	ristorante	11	161	19.970	22.780	9	161	23.230	18.090
5	Before	benzina	11	161	7.520	15.570	9	161	20.750	18.540
6	Before	biglietto	11	161	4.220	12.890	9	161	11.120	15.860
7	Before	città	11	161	6.940	14.540	9	161	18.790	18.370
8	Before	ritorno	10	161	7.160	14.730	9	161	11.960	16.050
9	Before	viaggio	11	161	11.730	16.920	9	161	19.460	19.090
10	Before	prezzi	11	161	12.630	19.410	9	161	26.100	21.530
11	Before	storia	11	161	13.960	17.280	9	161	24.170	18.870
12	Before	bellissima	11	161	12.150	19.150	9	161	15.240	16.320
13	Before	antico	11	161	11.720	18.020	9	161	18.330	17.960
14	Before	scultura	10	161	6.110	15.560	9	161	11.120	18.090
15	Before	piazza	11	161	18.410	20.400	9	161	34.570	21.340
16	Before	strada	11	161	23.130	21.840	9	161	36.590	23.570
17	Before	viale	11	161	9.260	16.520	9	161	20.920	20.430
18	Before	tendenza	9	161	4.130	13.180	9	161	7.510	15.480
19	Before	chiesa	11	161	8.990	14.350	9	161	17.860	16.440
20	After	vicino	11	31	14.920	15.830	9	31	29.040	21.090
21	After	previsioni	11	31	6.110	7.890	9	31	21.560	17.460
22	After	pioggia	9	31	6.150	9.690	9	31	14.820	13.500
23	After	ristorante	11	31	21.940	19.920	9	31	41.610	20.100
24	After	benzina	11	31	7.680	9.180	9	31	38.570	26.380
25	After	biglietto	11	31	7.360	11.240	9	31	25.970	20
26	After	città	11	31	9.290	10.240	9	31	24.820	17.740
27	After	ritorno	10	31	6.210	9.020	9	31	10.700	10.130
28	After	viaggio	11	31	12.170	17.860	9	31	22.180	17.640
29	After	prezzi	11	31	23.100	23.310	9	31	33.590	20.160
30	After	storia	11	31	9.640	8.990	9	31	28.480	16.250
31	After	bellissima	11	31	15.700	19	9	31	24.350	19.360
32	After	antico	11	31	16.320	21.310	9	31	19.480	13.480
33	After	scultura	10	31	9.690	12.290	9	31	21.840	19.600
34	After	piazza	11	31	21.390	23.610	9	31	47	25.810
35	After	strada	11	31	24.460	22.390	9	31	45.640	25.570
36	After	viale	11	31	8.950	10.990	9	31	22.290	17.280
37	After	tendenza	9	31	6.810	12.340	9	31	11.510	13.970
38	After	chiesa	11	31	10.140	11.420	9	31	28.820	19.860

C.1.2 Identification Strategy

Using the Google Search Trends Tool (GSTT) data that we collected, we test if low-cost access to mobile data increases the use of online search after RLAH.

GSTT provides geographic- and time-specific search activity but does not separate desktop from mobile search. Furthermore, GSTT does not differentiate the search traffic of travellers' from the search traffic of each country's residents.

Since our identification strategy relies on the ability to segregate search behavior of those affected RLAH, from those who are not, we needed to adapt our GSTT data collection to achieve that goal. In particular, we focused on studying search traffic of Italian keywords related to travel. Italian is the first language only in Italy and San Marino (which is in Italy).[20] The Italian population is among the leader in smartphone adoption worldwide.[21] Italians are also heavy travellers .[22] Furthermore, focusing on a single language ensures that we do not have to worry with different travel patterns across countries.

In our strategy, we assume that within the EU, changes in Italian search patterns in countries other than Italy will most likely originate from persons traveling and affected by RLAH.

We then use our dataset with monthly search activity of keyword k in date t in country i to replicate our primary analysis and to contrast the search activity for Italian keywords in EU countries and countries outside the EU before and after RLAH.

Equation (16) provides the econometric specification that we use in this analysis, where δ_m captures monthly seasonality. D^m_{itk} are month of the year dummies. θ_t are month-year fixed effects, a_i are country fixed effects, λ_k are keywords fixed effects and v_{ikt} the idiosyncratic error term.

Equation (17) provides a variation that decomposes the effect of RLAH by the month of the year. In section C.1.4 and C.1.4 we present additional robustness analysis for the GSTT data. In particular, we use Bayesian Structural Time series to re-estimate the effects of our main specification. The results are consistent.

[20]See https://web.archive.org/web/20120210212620/http://www.ethnologue.com/show_language.asp?code=ita. A first issue may be that Italian is also recognised as an official language in part of Switzerland, which was not subject to RLAH regulation. A second issue might be that Italian as a large diaspora across the world (see for example https://www.repubblica.it/static/speciale/2016/referendum/costituzionale/estero.html that sum-up vote participation in 2016 from Italian living abroad). We assume that such diaspora does not change during our analysis and can be tackled by controlling for country fixed effect.

[21]See https://newzoo.com/insights/trend-reports/newzoo-global-mobile-market-report-2019-light-version/

[22]See https://www.istat.it/it/files//2019/02/Viaggi-e-vacanze-Anno-2018_rev.pdf

$$Search_{ikt} = \beta_1 after_t \times EUcountry_i + \sum_{m=1}^{12} \delta_m D^m_{itk} + \theta_t + a_i + \lambda_k + v_{ikt} \quad (16)$$

$$Search_{ikt} = \sum_{m=1}^{12} \beta_m after_t \times EUcountry_i \times D^m_{itk} + \theta_t + a_i + \lambda_k + v_{ikt} \quad (17)$$

C.1.3 Results

Figure C.5 shows average Google trend activity of the same list of Italian keyword for both EU and non-EU countries from January 2004 to December 2019. First red dashed line corresponds to the beginning of the EU roaming regulation, while the second red dashed line pictures the introduction of RLAH rules. Black dashed lines underline all roaming regulation waves laid out by the European Commission. The graph shows successive increases in Google trend activity after roaming regulations for EU countries compared to non-EU countries. In the subsequent analysis we focus only on the introduction of RLAH rules. A more detailed analysis including all roaming regulation can be found in C.1.4.

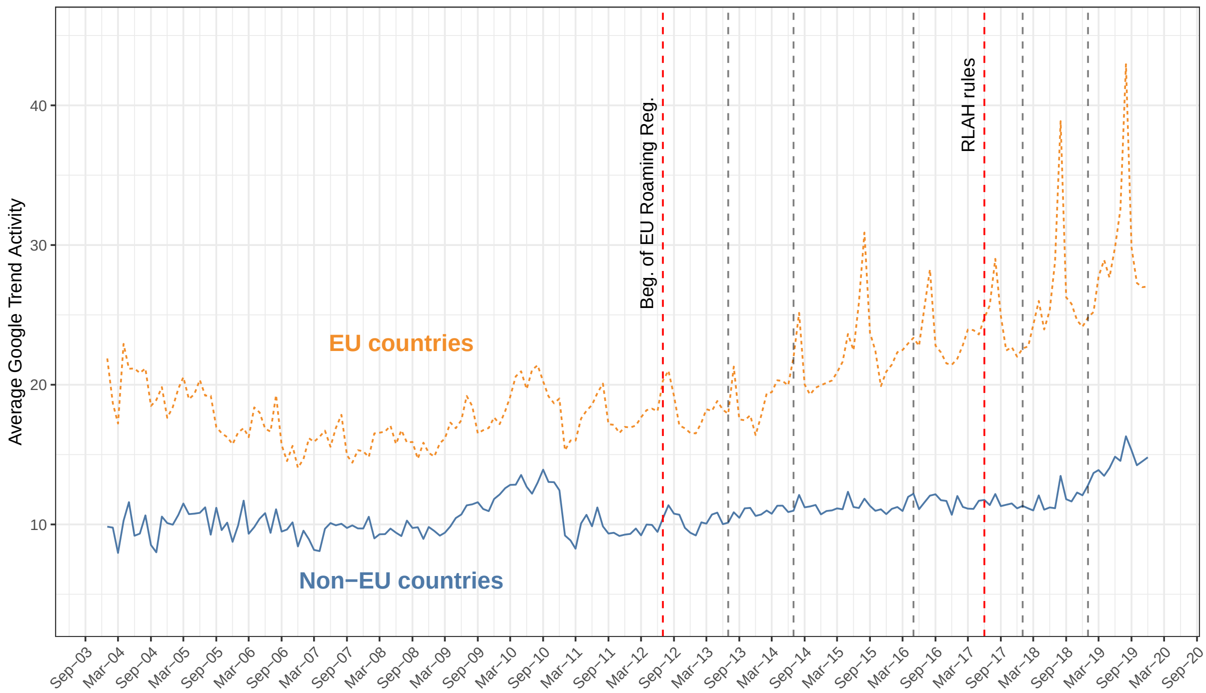

Figure C.5: Model free evidence of the impact of RLAH rules on Google trend activity

We complement this graph with the strategy depicted in C.1.2 that quantify the impact of roaming regulations on Google search activity. Because Google trend data exhibit seasonality, especially around Italian holidays, we adjust our time series by Loess, using the methodology developed in Cleveland et al. (1990).

Firstly, table C.9 summarizes the strategy laid out in 16. Columns (1) and (2) correspond to 16 with data ranging from January 2004 to December 2019. Columns (3) and (4) also relates to 16 with data from July 2014 to December 2019. It captures the incremental effect of RLAH rules implementation relative to the previous waves of roaming regulation. All models are adjusted for seasonality and include time, country and keyword fixed effects. We cluster standard errors at the country and keyword level. They depict a positive and statistically significant increase in Google trend activity within EU countries after roaming regulations implementation compared to non-EU countries.

Table C.9: Impact of roaming like at home on search volume

	Dependent variable:			
	Google trend activity			
	Jan 2004 to Dec 2019		Jul 2014 to Dec 2019	
	(1)	(2)	(3)	(4)
EU	8.333***		11.308***	
	(0.126)		(0.287)	
After	2.163***		1.318***	
	(0.215)		(0.288)	
EU x After	5.942***	5.942**	2.967***	2.967*
	(0.318)	(2.256)	(0.426)	(1.264)
Constant	10.544***		11.390***	
	(0.085)		(0.194)	
Country	No	Yes	No	Yes
Month	No	Yes	No	Yes
Keyword	No	Yes	No	Yes
Month of the year	No	Yes	No	Yes
Observations	71,808	71,808	24,684	24,684
R^2	0.097	0.476	0.132	0.563
Adjusted R^2	0.097	0.474	0.132	0.561
Residual Std. Error	15.396	11.752	16.617	11.817

Note: + p<0.1; * p<0.05; ** p<0.01; *** p<0.001
Cluster robust standard errors in ()
Errors clustered by country and keyword

Appendix C.1.4 decomposes the effect of the regulation by month of the year laid out in 17. The results underline that Google trend activity within EU countries increases significantly after both regulation for almost every month. The increase is lower for Christmas holidays and particularly higher around summers.

As a robustness check, we provide an alternative analysis using Bayesian structural time series equations. Details about this analysis can be found in appendix C.1.4.

All identification strategies highlight that RLAH changed the amount of Google search performed by EU residents while traveling. Considering only July 2014 to December 2019 to measure the effect of RLAH since the last regulation (column (3) and (4) in table C.9), our results shows an increase of 12% in Google activities due to RLAH. Considering a longer time span (column (1) and (2)) increases the effect to 30% as roaming regulations started in 2012 within the EU.

C.1.4 Supplementary analysis on Google trend data

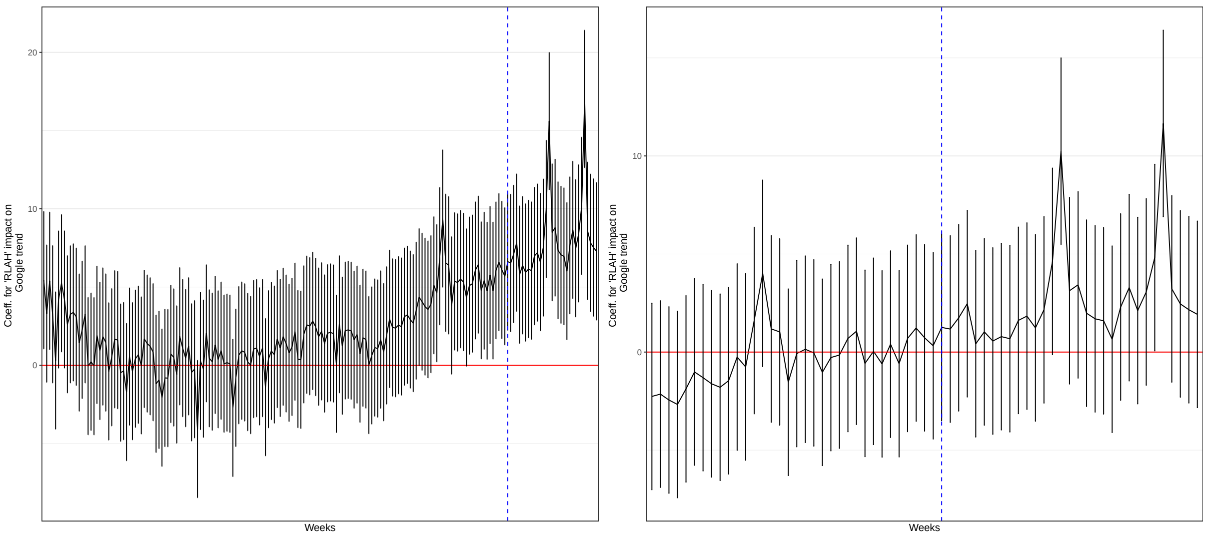

Figure C.6: Regression results for parallel trend assumption for Google Trend's dataset

Parallel trend assumption for the Google Trend dataset

Impact of past roaming regulation

Table C.10: Impact of roaming like at home on Google search volumes with all roaming regulation

	Dependent variable:							
	Google trend activity							
	Jan 2004 to Jun 2013	Jul 2012 to Jun 2017	Jul 2013 to Apr 2016	Jul 2014 to Jun 2017	May 2016 to Dec 2017	Jul 2017 to Dec 2018	Jan 2018 to Dec 2019	all
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$EU \times After_{2012}$	0.518							0.518
	(1.424)							(1.424)
$EU \times After_{2013}$		0.025						0.025
		(0.498)						(0.522)
$EU \times After_{2014}$			3.010***					3.010***
			(0.836)					(0.876)
$EU \times After_{2016}$				0.960				0.960
				(0.787)				(0.697)
$EU \times After_{2017}$					0.894^{+}			0.894^{+}
					(0.507)			(0.459)
$EU \times After_{2018}$						1.638^{+}		1.638*
						(0.848)		(0.673)
$EU \times After_{2019}$							0.439	0.439
							(1.300)	(1.173)
Country	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Keyword	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month of the year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	42,636	8,976	12,716	13,464	7,480	6,732	8,976	71,808
R^2	0.454	0.563	0.565	0.569	0.569	0.565	0.576	0.478
Adjusted R^2	0.452	0.559	0.562	0.566	0.565	0.561	0.572	0.476
Residual Std. Error	11.170	9.812	10.698	11.154	11.454	11.781	12.468	11.729

Note: + p<0.1; * p<0.05; ** p<0.01; *** p<0.001
Cluster robust standard errors in ()
Errors clustered by country and keyword

Table C.11: Impact of roaming like at home on Google search volumes per month

	Dependent variable:											
	Google trend activity											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$EUxAfter_{2014}$	2.962^{+}	3.383*	3.514*	3.062^{+}	3.578*	4.006*	4.832**	7.318**	3.825*	3.585*	2.411	3.661*
	(1.583)	(1.524)	(1.455)	(1.679)	(1.668)	(1.563)	(1.724)	(2.242)	(1.754)	(1.536)	(1.571)	(1.633)
$EUxAfter_{2017}$	1.863*	0.902	2.202^{+}	3.237*	1.817	2.570^{+}	3.449^{+}	6.862**	3.593*	2.728*	2.974**	1.651
	(0.924)	(1.224)	(1.322)	(1.343)	(1.131)	(1.505)	(1.760)	(2.455)	(1.573)	(1.337)	(1.070)	(1.194)
Country	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Keyword	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,984	5,984	5,984	5,984	5,984	5,984	5,984	5,984	5,984	5,984	5,984	5,984
R^2	0.482	0.485	0.489	0.463	0.470	0.474	0.470	0.472	0.485	0.493	0.476	0.492
Adjusted R^2	0.477	0.481	0.484	0.458	0.465	0.469	0.465	0.467	0.480	0.489	0.471	0.487
Residual Std. Error	11.418	11.346	11.397	11.999	11.797	11.786	12.043	12.731	11.691	11.432	11.805	11.513

Note: + p<0.1; * p<0.05; ** p<0.01; *** p<0.001
Cluster robust standard errors in ()
Errors clustered by country and keyword

Table C.12: Impact of roaming like at home on Google search volumes by month

	Dependent variable:	
	Google trend activity	
	Jan 2004 to Dec 2019	Jul 2014 to Dec 2019
	(1)	(2)
January	4.488*	1.513
	(1.850)	(0.938)
February	3.965*	0.991
	(2.008)	(0.989)
March	5.223*	2.248*
	(2.082)	(1.074)
April	5.833**	2.858*
	(2.167)	(1.240)
May	4.961**	1.986*
	(1.900)	(0.983)
June	5.919**	2.945*
	(2.256)	(1.280)
July	6.745**	3.770**
	(2.147)	(1.301)
August	11.087***	8.113***
	(2.725)	(2.051)
September	6.264**	3.289**
	(2.184)	(1.254)
October	5.441**	2.466*
	(1.963)	(1.030)
November	4.937*	1.962*
	(2.015)	(0.978)
December	4.683*	1.708
	(2.142)	(1.149)
Country	Yes	Yes
Month	Yes	Yes
Keyword	Yes	Yes
Observations	71,808	24,684
R^2	0.476	0.564
Adjusted R^2	0.475	0.562
Residual Std. Error	11.747	11.803
Note:	+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001	
	Cluster robust standard errors in ()	
	Errors clustered by country and keyword	

Impact of RLAH regulation by month

Alternative Strategy for Google trend analysis Bayesian Structural Time Series

As a robustness check, we use a Bayesian structural time series analysis as an alternative strategy to create counterfactual time series.

For keyword k in country i, we have:

$$y_t = \mu_t + \tau_t + \beta^T \mathbf{x}_t + \epsilon_t \tag{18}$$

$$\mu_t = \mu_{t-1} + \delta_{t-1} + \eta_t \tag{19}$$

$$\delta_t = \delta_{t-1} + \omega_t \tag{20}$$

$$\tau_t = -\sum_{s=1}^{S-1} \tau_{t-s} + \gamma_t \tag{21}$$

with ϵ_t, η_t and ω_t being centered normal random errors with potentially different constant variance across time. μ_t captures the current trend, τ_t depicts the seasonality while $\mathbf{x}_t$ is a vector of covariates that are correlated with y_t but are not impacted by RLAH rules. β^T captures the difference of covariates across time, relative to y_t.

Using this strategy, we are able to use different specification. Firstly, we use the time series average across EU countries and keywords as dependent variable and times series at from countries outside EU as covariates. As the latter should be correlated with the former before roaming regulations implementation, we create a counterfactual estimate of times series in EU country for our keywords in the case of no regulation. We customize 18 such that y_t depicts the average search index at period t across keywords and countries within the EU, and $\mathbf{x}_t$ is a vector of time series at the keyword-country level for countries outside the EU.

Secondly, we focus only on EU countries and isolate a first period including the regulation implementation date, with the same number of months before and after the regulation. Similarly, we isolate a matching period of the same length before the regulation and not overlapping our initial selection. Time series for the period including the regulation are used as dependent variable while times series in the period before are used as covariates. Our specification creates a counterfactual for months in the period after the regulation is implemented in EU countries based on the previous period. We specify 18 such that y_t depicts average search index at period month t across keywords and countries for the period after the regulation and $\mathbf{x}_t$ is a vector of time series at the keyword-country-month level that take place in the period before the regulation.

Results

figure C.7 corresponds to the first strategy related to (18). Top sub-figure uses Google trend activity outside EU to predict activity within EU countries. As a results of RLAH rules implementation, Google trend monthly activity increased by 36% (p<0.001) on average within EU countries compared to the prediction performed using countries outside EU. The lower panel decomposes such effect. Considering data from January 2004 to December May 2017, left lower panel sub-figure underlines that Google trend monthly activity increased by 26% (p<0.001) after July 2014 regulation. Finally, right lower panel sub-figure exhibits a corresponding increase of 23% (p<0.001), capturing the incremental effect of RLAH rules passed in June 2017 relative to the roaming regulation passed in July 2014.

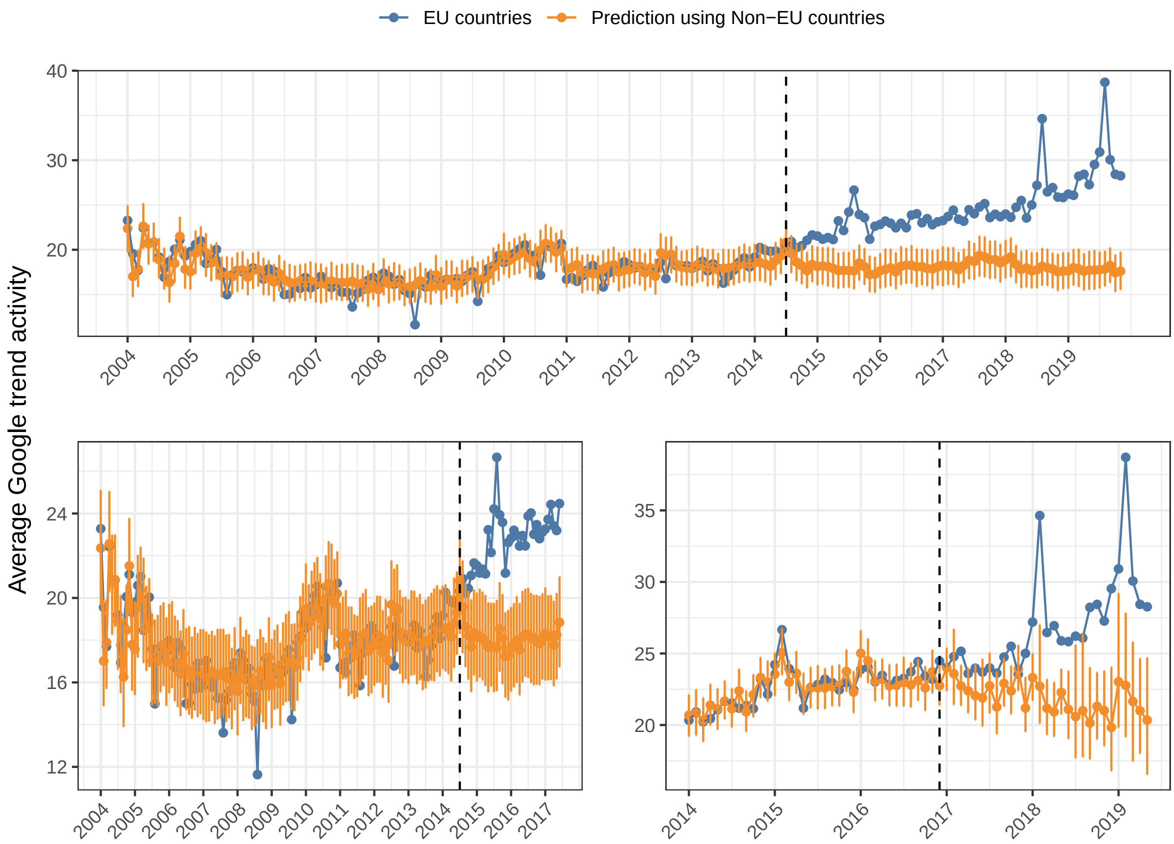

Figure C.7: Impact of roaming regulation on Google trend activity in EU countries, using activity outside EU countries as counterfactual

Finally, figure C.8 corresponds to the second strategy used with 18 and focuses on EU countries. Top sub-figure shows how Google trend monthly activity from 2007-2011 is used to predict monthly average activity between 2012 and 2016 in absence of regulation. According to our strategy, the introduction of July 2014 roaming regulation increased Google trend monthly activity by 17% (p<0.001) on average, with a large spike around August. Sub-figure on the lower part shows the use of activity from 2013-2015 to predict activity for 2016-2018, hence comparing the impact of the two roaming regulations. Google trend monthly activity increased by 9% ($p<0.001$) on average after June 2017 regulation, compared to after July 2014 regulation. However, such effect appears largely tied to summer 2018.

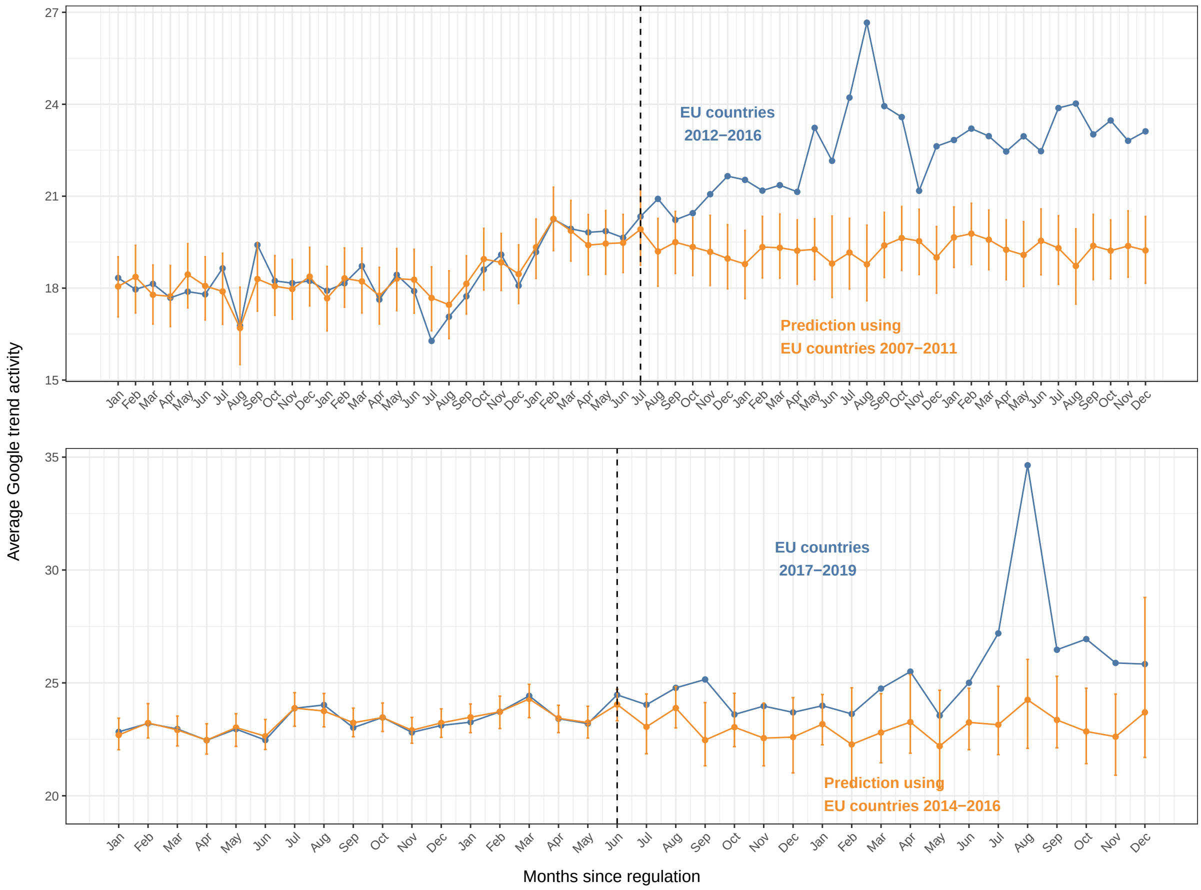

Figure C.8: Impact of roaming regulation on Google trend activity in EU countries, using past activity within EU countries as counterfactual

Our approach shows some limitations. Firstly, we are unable to disentangle search that originates from mobile from those that are on desktop. This means that our data only capture additional search activity due to a better mobile data access, but are not able to quantify how much of the desktop search is shifting to mobile. Secondly, Google trends doesn't distinguish between intensive and extensive search margin, as it return global search activity. This means that we are unable to identify whether a higher search activity for a keyword is due to more user searching or more search per user, which should be both affected by a better access to mobile data.

C.2 German publishers

C.2.1 Data

We recover a panel dataset of audiences for websites of different German publishers for the period between January 2015 and December 2019. The data originates from "The German Audit Bureau of Circulation" (IVW) open database.[23] The IVW is an association of publishers, advertisers, and advertising agencies that has been neutrally recording audience data on the German market of online websites.

For each website and each month, we obtained the number of unique users that accessed the website and the total number of visits. The dataset allows us to separate mobile and desktop audiences and to distinguish audiences that originate from within or from outside Germany. We keep in our sample websites with one observation on both mobile and desktop devices, before and after the regulation. Our final dataset contains 41,377 observations, with 396 unique publishers over 59 months.

We display summary statistics of the IVW dataset in table C.13

[23] IVW stands for "Informationsgemeinschaft zur Feststellung der Verbreitung von Werbeträgern". See https://www.ivw.eu/

Table C.13: Descriptive statistics for the IVW dataset on German website audience.

	device	period	variable_name	n_website	n_month	avg	stdev	q05	med	q95
1	Desktop	Before RLAH	Users	396	29	949, 360.100	3, 397, 830	4, 159.2	160, 157	4, 283, 445
2	Mobile	Before RLAH	Users	396	29	1, 259, 301	10, 848, 370	1, 026.850	63, 390	3, 724, 626
3	Desktop	Before RLAH	Visits	396	29	5, 089, 910	24, 783, 618	8, 420	527, 185	19, 306, 617
4	Mobile	Before RLAH	Visits	396	29	7, 105, 653	55, 088, 764	1, 915.850	203, 542.500	16, 876, 275
5	Desktop	Before RLAH	Visits per user	396	29	4.139	3.430	1.587	3.225	9.257
6	Mobile	Before RLAH	Visits per user	396	29	4.053	4.488	1.374	2.840	10.487
7	Desktop	After RLAH	Users	396	30	959, 412.500	4, 243, 936	4, 055.400	138, 396	4, 141, 079
8	Mobile	After RLAH	Users	396	30	2, 200, 959	18, 918, 759	1, 972	100, 546	5, 435, 500
9	Desktop	After RLAH	Visits	396	30	4, 085, 792	17, 689, 243	7, 785.400	388, 220	16, 118, 026
10	Mobile	After RLAH	Visits	396	30	12, 025, 137	108, 215, 495	3, 436.200	288, 491	22, 681, 125
11	Desktop	After RLAH	Visits per user	396	30	3.741	3.777	1.432	2.644	9.172
12	Mobile	After RLAH	Visits per user	396	30	3.649	5.443	1.283	2.442	9.450

C.2.2 Identification Strategy

We analyse how a better access to mobile data changed consumer visits to news website when travelling. To do so, we use a panel dataset of German websites that breakdown audience metrics according to the origin (inside and outside Germany) and the device (mobile or desktop).

Our specification modifies the initial analysis of mobile data. We focus only on audience outside Germany (travellers) and contrast website visits between mobile and desktop across time. We expect the amount of visit on mobile to increase after the regulation, relative to the traffic from desktop.

Our identification strategy requires two assumptions on the variation in the number of website visits outside Germany. Firstly, we assume that such variation is from German travellers. Secondly, we assume that this traffic mainly originates from Europe, as we cannot distinguish the exact countries where the traffic originates from.[24] Both assumptions allow us to measure the effect of the regulation on website visit, as we contrast the change in mobile traffic due to the regulation relative to desktop traffic.

The empirical approach follows 22, where $Visits_{idt}$ is the traffic for website i, in month t for device type d. $MobileDevice_d$ is a dummy variable equal to one when audience originates from mobile devices. As before, δ_m capture seasonality of month m while θ_t are month-year fixed effects, a_i are website fixed effects and v_{idt} the idiosyncratic error term.

$$Visits_{idt} = \alpha_1 after_t \times MobileDevice_d + \sum_{m=1}^{12} \delta_m D^m_{itk} + \theta_t + a_i + v_{idt} \tag{22}$$

C.2.3 Results

We analyze both how extensive and intensive margin change after the regulation, depending on the device. figure C.9 show the number of unique user (left panel) and the average number of visit per user (right panel) contrasted across 60 months and across devices. RLAH rules resulted in an increase in the number of unique mobile users visiting German publisher websites, relative to desktop users.

table C.14 breakdowns the impact of RLAH rules on website visit from travellers, as laid out in equation (22). The average number of unique visitor per website per month increase by 31.5%

[24]Studies shows that German mostly travel within europe https://ec.europa.eu/eurostat/web/tourism/publications

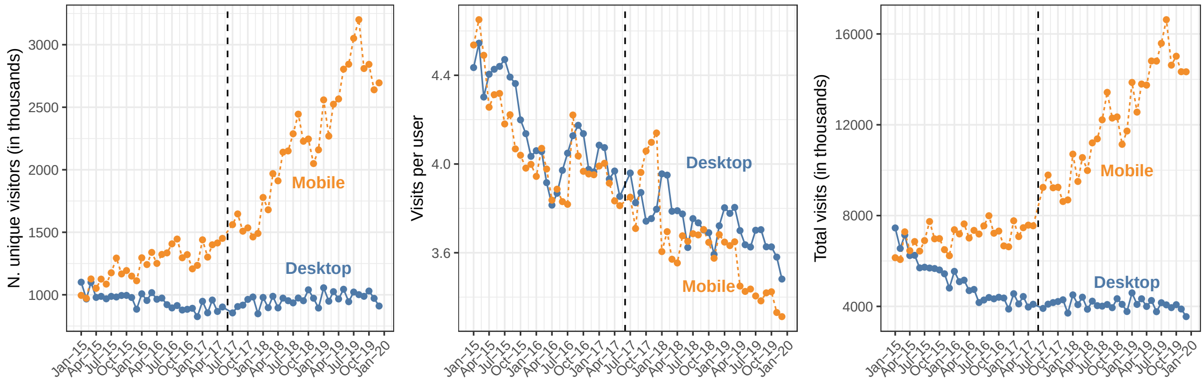

Figure C.9: Model free evidence of the impact of RLAH rules on website visits

after the regulation, while the average number of visit per user does not change. This indicates that RLAH rules had an impact on the extensive margin of website audience, attracting new users, while had no effect on the intensive margin as the number of visit per user follows does not change significantly.

Table C.14: Impact of RLAH on website traffic

	Dependent variable:		
	log(User)	log(Visits)	log(Visits/User)
	(1)	(2)	(3)
MobileDevice x After	0.315**	0.335**	0.021
	(0.114)	(0.114)	(0.056)
Trends	Yes	Yes	Yes
Publisher FE	Yes	Yes	Yes
Month seasonality	Yes	Yes	Yes
Observations	41,377	41,377	41,377
R^2	0.878	0.899	0.695
Adjusted R^2	0.898	0.877	0.692
Residual Std. Error	0.798	0.799	0.337
Note:	+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001		
	Cluster robust standard errors in ()		
	Errors clustered by publisher		

C.2.4 Parallel trend assumption for the Website visits' dataset

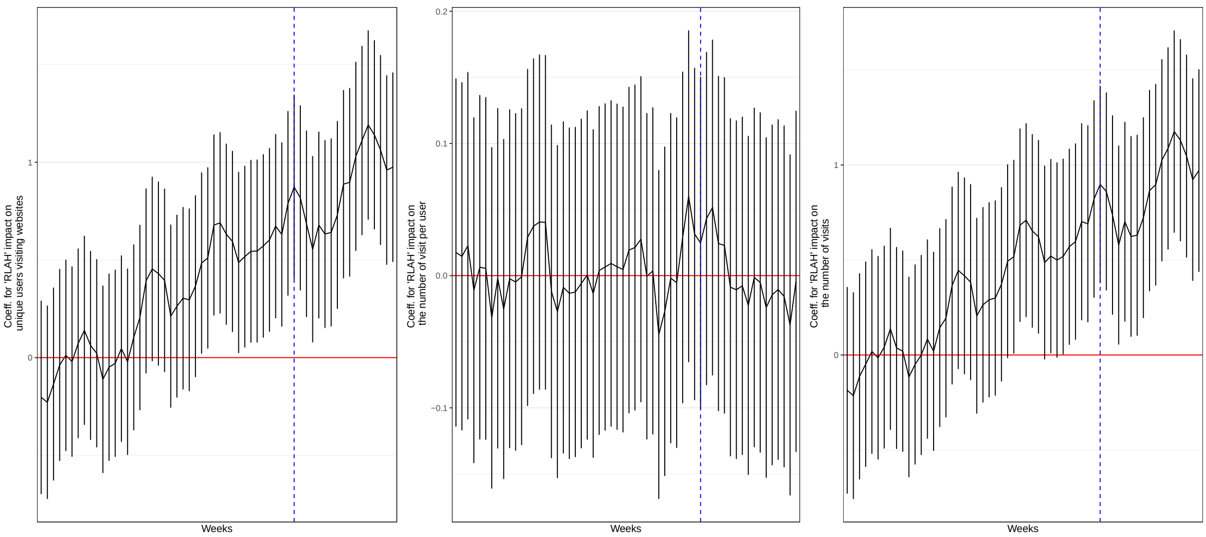

Figure C.10: Regression results for parallel trend assumption for IVW dataset

C.3 Tripadvisor

C.3.1 Data

To provide additional insights into how access to mobile data can generate externalities to economic agents, we obtained data from user-generated content websites.

First, we collected an aggregated anonymised dataset of restaurant reviews from TripAdvisor. We obtained data before and after the RLAH rules for users traveling in geographies affected by RLAH and geographies not affected by the regulation.

Our TripAdvisor dataset contains reviews for the top five hundred restaurants in the largest five cities of the top twenty countries visited by Portuguese consumers. We selected reviews written in Portuguese.

Ten cities in our sample are in geographies affected by RLAH. To ensure that our dataset is anonymous, we aggregate all review data at the restaurant's level, and we did not obtain usernames or the actual text of the restaurant's reviews. Instead, our dataset registers only aggregate user statistics for each country and each restaurant that allows measuring user contributions over time. For each review, we registered the date and the country of residence of the posting user. The country of residence is self-declared and is available in some user profiles on TripAdvisor's website. We use this information to separate review contributions that originate from the home location of the posting user, from review contributions posted while traveling.

For each review, we also obtained the device type used to post it online. TripAdvisor reviews allow separating contributions that originate from smartphones from reviews that originate from desktop/laptop computers, and this information allows us to generate a device-based identification strategy.

The TripAdvisor dataset covers the period from June 2016 to July 2018. It contains 2554 reviews posted by 1987 users for 1358 restaurants. We display summary statistics of the TripAdvisor dataset in table C.15.

Table C.15: Descriptive Statistics for the TripAdvisor dataset.

	location	period	device	#months	#review	#users	#restaurants	#country
1	Not EU	Bef. RLAH	Not mobile	13	212	179	164	8
2	Not EU	Bef. RLAH	Mobile	13	269	214	190	9
3	Not EU	Aft. RLAH	Not mobile	13	150	132	120	8
4	Not EU	Aft. RLAH	Mobile	13	279	230	216	9
5	EU	Bef. RLAH	Not mobile	13	388	319	303	10
6	EU	Bef. RLAH	Mobile	13	449	392	326	10
7	EU	Aft. RLAH	Not mobile	13	312	269	253	10
8	EU	Aft. RLAH	Mobile	13	634	548	440	10

C.3.2 Identification Strategy

We study how low-cost access to mobile data changes user-generated content contributed over the internet. For that purpose, we analyze restaurant review posts in TripAdvisor before and after the RLAH regulation.

TripAdvisor allow the geolocalization of user-generated content. We use posts with geolocalization information to determine the location of each platform's users generating content online. TripAdvisor and Twitter also store information on the type of device used to post content to their platforms. These platforms record whether posts or reviews originate from mobile devices, desktop or laptop computers. Together, the geolocalization of posts and the type of device posting allow us to replicate the difference-in-difference strategy that we presented earlier.

Our initial specification replicates the analysis of mobile data: we contrast the posting behavior of users traveling in the RLAH geographies against the posting behavior of users traveling outside the RLAH zones. As an alternative econometric specification, we contrast the posting behavior of users employing mobile devices to tweet or share restaurant reviews to the posting behavior of users that utilize computers to post content on these platforms. Contrasting behavior of users posting in their mobile device versus users posting using their computers relies on two identifying assumptions. First, users posting content via mobile devices while traveling in the EU will be affected by the RLAH rules after the regulation enters into effect, and users posting content through desktop devices while traveling in the EU will not. Second, the trends in user-generated content for mobile and desktop devices would be similar in the absence of RLAH rules.

The first assumption originates from statistics that suggest that users that access the internet via their computers when traveling are more likely to be connected to a WIFI hot-spot than to be using mobile data plans.[25] The second assumption is the standard parallel trend requirement for differences-in-differences estimation. Like before, we test for the existence of parallel trends across groups before the implementation of the RLAH regulation.

The first empirical approach estimates 23 and the second empirical approach estimates 24. In both these equations REV_{idt} is the number of reviews for country i, in device type d in month t. $MobileDevice_d$ is a dummy variable equal to one for posts originating from mobile devices, and $EUtravel_i$ is a dummy variable equal to one for countries affected by the RLAH rules. θ_t are month fixed effects, a_i are country fixed effects and v_{idt} the idiosyncratic error term.

$$REV_{idt} = \beta_1 after_t \times EUtravel_i + \theta_t + a_i + v_{idt} \tag{23}$$

$$REV_{idt} = \alpha_1 after_t \times MobileDevice_d + \theta_t + a_i + v_{idt} \tag{24}$$

C.3.3 Results

Figure C.11 provides model-free evidence of the impact of RLAH on user-generated contributions on TripAdvisor using both strategies laid out in section C.3.2. The picture tracks the 24 months from June 2016 to June 2018. It highlights that after RLAH, there was an increase in the average number of reviews, on the number of travellers reviewing and on the number of review per traveller. We complement this figure with regression in table C.16.

Table C.16 quantifies the impact of RLAH on user-generated content in Trip Advisor for travellers originating from the country where MOBILE operates. Columns (1), (3), and (5) correspond to equation (23) from section C.3.2 and columns (2), (4), and (6) correspond to equation (24) from the same section. We estimate both models with time and country fixed effects, and we cluster standard errors at the country level. Both identification strategies translates in positive and statistically significant increases in TripAdvisor review activity. The total number of reviews posted by travellers while visiting RLAH geographies increased between 49% (column (3)) and

[25] see https://www.ipass.com/wp-content/uploads/2016/11/iPass-Mobile-Professional-Report-2016.pdf

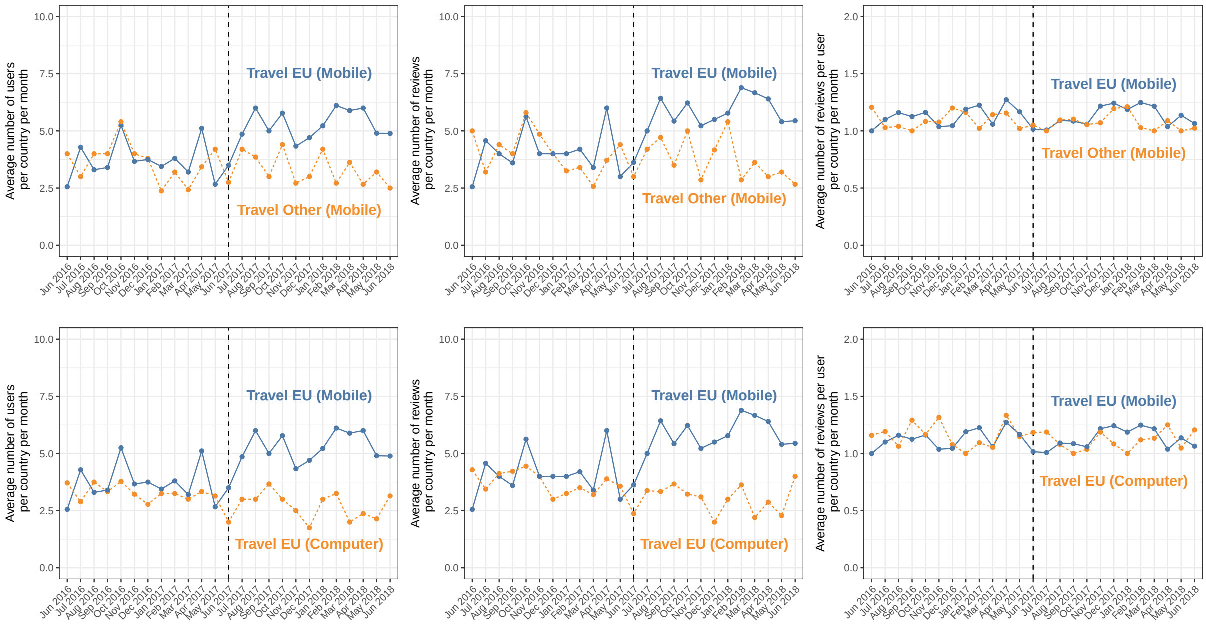

Figure C.11: Model free evidence of the impact of RLAH rules on user generated content on TripAdvisor

44% (column (4)) while the number of unique reviewers increaed between 50% (column (1)) and 46% (column (2)). However, RLAH rules had no impact on the number of review per user. This underlines how the regulation attracted new reviewers, but did not generate more activity by user. We conduct an analog analysis considering Restaurant reviewed in C.3.5 that shows parallel results. The increased participation of travellers in disseminating information about restaurants is evidence of the information spillovers that the RLAH enabled within the geographies affected by the new roaming rules.

Table C.16: Impact of roaming like at home on trip advisor contributions

	Dependent variable:					
	Users		Reviews		Reviews/User	
	(1)	(2)	(3)	(4)	(5)	(6)
EUTravel x After	1.657***		1.773***		0.023	
	(0.438)		(0.453)		(0.049)	
MobileDevice x After		2.122**		2.402**		0.037
		(0.422)		(0.486)		(0.063)
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	365	430	365	430	365	430
R^2	0.759	0.628	0.749	0.608	0.124	0.079
Adjusted R^2	0.727	0.595	0.715	0.57	0.007	−0.003
Residual Std. Error	1.959	2.111	2.226	2.35	0.246	0.280
Note:	+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001 Cluster robust standard errors in () Errors clustered by country					

C.3.4 Parallel trend assumption for the TripAdvisor's dataset

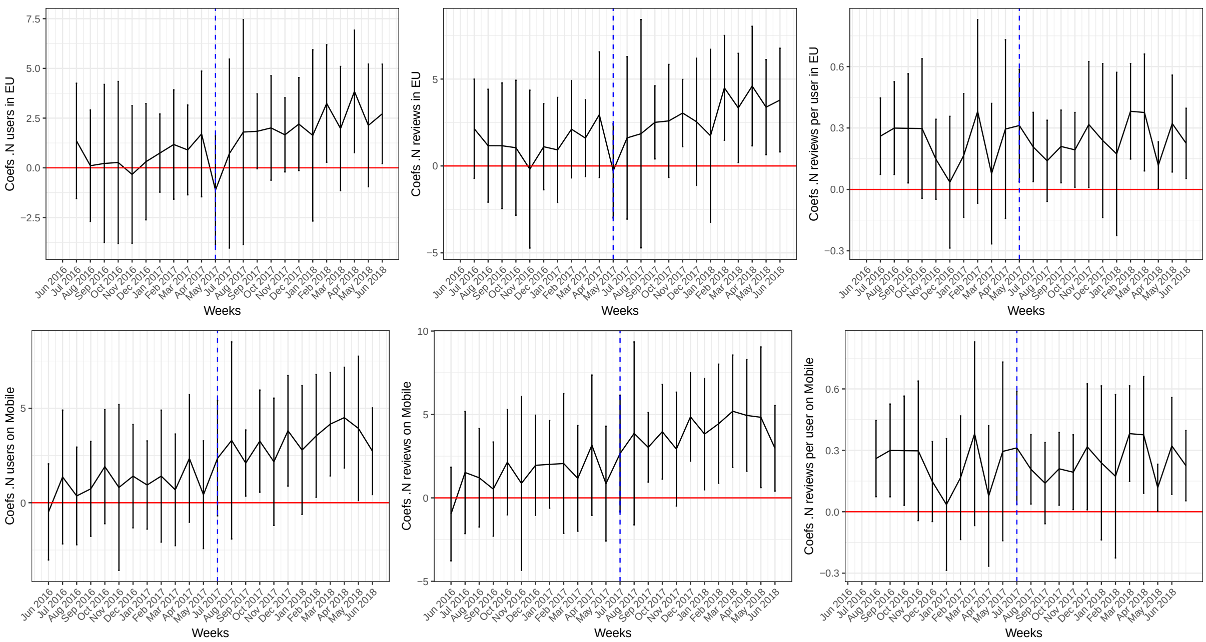

Figure C.12: Regression results for parallel trend assumption for TripAvisor dataset

C.3.5 Tripadvisor analysis at the Restaurant level

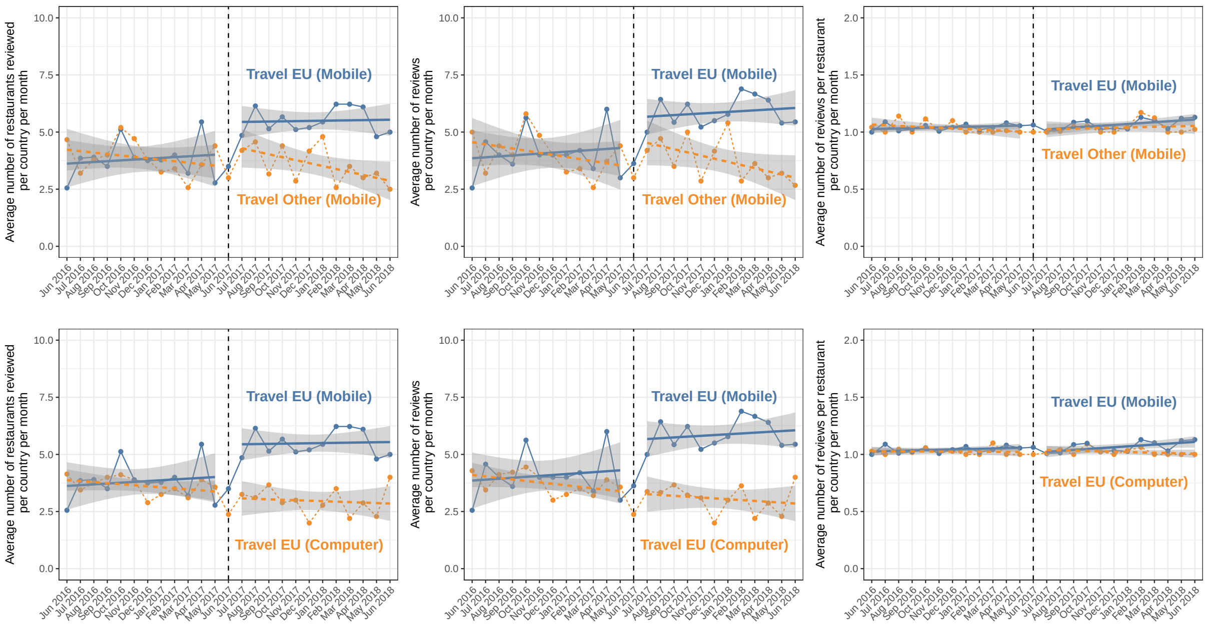

Figure C.13: Model free evidence of the impact of RLAH rules on user generated content on TripAdvisor at the restaurant level

Table C.17: Impact of RLAH rules on TripAdvisor contributions by restaurants

	Dependent variable:					
	Restaurants		Reviews		Reviews per restaurant	
	(1)	(2)	(3)	(4)	(5)	(6)
RLAH x After	1.657***		1.773***		0.014	
	(0.438)		(0.453)		(0.039)	
Mobile x After		2.673**		2.894**		0.046*
		(0.892)		(0.931)		(0.021)
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	365	430	365	430	365	430
R^2	0.759	0.625	0.749	0.605	0.188	0.150
Adjusted R^2	0.727	0.592	0.715	0.571	0.079	0.077
Residual Std. Error	1.959	2.118	2.226	2.354	0.134	0.117

Note: + p<0.1; * p<0.05; ** p<0.01; *** p<0.001
Cluster robust standard errors in (); Errors clustered by country

C.4 Twitter

C.4.1 Data

Second, we used Twitter's Sample Stream API (TSSA) to obtain geo-localised tweets before and after the RLAH rules.[26]

We focused on geo-localised tweets written in the European Union's most spoken languages.[27] We collected all tweets written in German, French, and Italian. We also obtained Tweets posted in English and Spanish and Portuguese by selecting tweets originating from Great Britain, Spain, and Portugal.[28]

For each user posting geo-localised tweets in the languages mentioned above, we obtained her historical tweets. We did so following Twitter's development policy, which limits the number of Tweets we can collect. Twitter allows access to 3200 of a user's most recent Tweets, re-Tweets, and status updates. The rate at which we can collect such information is limited to 900 requests per 15 min window or 100k per 24 hours.[29] Given these restrictions in data collection, in our dataset, user timelines are shorter for users who are the most active and longer for those with less activity. Our data collection is also biased towards households who remained twitter users until 2020.

To ensure that we have complete twitter histories, we discarded all users with 3200 tweets. We also required users in our sample to have used twitter before May 2017. Our final dataset contained 13,369 tweets posted by 911 users across 113 different countries during the year 2017. We acknowledge that our twitter data collection has several limitations, but it still allows us to study how some users reacted to RLAH rules in 2017.

We display summary statistics of the TripAdvisor dataset in table C.18.

C.4.2 Identification Strategy

Twitter identification strategy mimics the one of TripAdvisor. We contrast the number of tweets from EU according to two strategies. The first empirical approach estimates 25 and the second

[26]TSSA provides a free endpoint with a real-time stream of 1% of all public tweets selected randomly. Details about the API are available here: https://developer.twitter.com/en/docs/tweets/sample-realtime/overview/GET_statuse_sample

[27]Due to infrastructure capacity constraints, we could not process all tweets from Twitter's Sample Stream and needed to bound our data collection efforts

[28]We did not filter on the English, Spanish and Portuguese languages directly to avoid collecting data from North and Latin America

[29]A request can contain a page with up to 200 tweets

Table C.18: Descriptive Statistics for the Twitter dataset.

	period	location	ave. #user	ave. #tweets	ave. #tweets per user	ave. #country
1	2017-01	Outside EU	0.795	7.542	1.789	83
2	2017-01	Within EU	2.233	10.067	3.400	30
3	2017-02	Outside EU	0.819	3.831	1.268	83
4	2017-02	Within EU	2.733	14.233	3.564	30
5	2017-03	Outside EU	0.952	6.193	2.166	83
6	2017-03	Within EU	3.333	15.067	3.420	30
7	2017-04	Outside EU	1	9.169	3.620	83
8	2017-04	Within EU	3.500	15.400	3.994	30
9	2017-05	Outside EU	1.048	9.928	2.233	83
10	2017-05	Within EU	4.267	16.300	2.351	30
11	2017-06	Outside EU	1.143	7.714	1.948	91
12	2017-06	Within EU	3.362	8.979	2.452	47
13	2017-07	Outside EU	1.024	8.193	2.242	83
14	2017-07	Within EU	5.167	20.300	3.148	30
15	2017-08	Outside EU	1.157	10.386	2.342	83
16	2017-08	Within EU	6.867	30.800	3.207	30
17	2017-09	Outside EU	0.880	6.253	1.489	83
18	2017-09	Within EU	4.533	13.067	2.626	30
19	2017-10	Outside EU	1.060	5.506	1.470	83
20	2017-10	Within EU	4.733	16.967	2.708	30
21	2017-11	Outside EU	0.867	4.482	1.129	83
22	2017-11	Within EU	3.700	15.567	2.372	30
23	2017-12	Outside EU	1.145	7.687	1.962	83
24	2017-12	Within EU	4	16.200	2.339	30

empirical approach estimates 26. In both these equations $TWEETS_{idt}$ is the number of tweets for country i, in device type d in month t. $MobileDevice_d$ is a dummy variable equal to one for posts originating from mobile devices, and $EUtravel_i$ is a dummy variable equal to one for countries affected by the RLAH rules. θ_t are month fixed effects, a_i are country fixed effects and v_{idt} the idiosyncratic error term.

$$TWEETS_{idt} = \beta_1 after_t \times EUtravel_i + \theta_t + a_i + v_{idt} \tag{25}$$

$$TWEETS_{idt} = \alpha_1 after_t \times MobileDevice_d + \theta_t + a_i + v_{idt} \tag{26}$$

C.4.3 Results

Figure C.14 and table C.19 repeat the TripAdvisor analysis for the case of twitter. The results show that after RLAH, there are more users and more geo-localized tweets for Twitter users residing in the EU and traveling in the EU than for Twitter users residing in the EU and traveling to other regions. The change in behavior in our sample manifests in the extensive rather than the intensive margin.

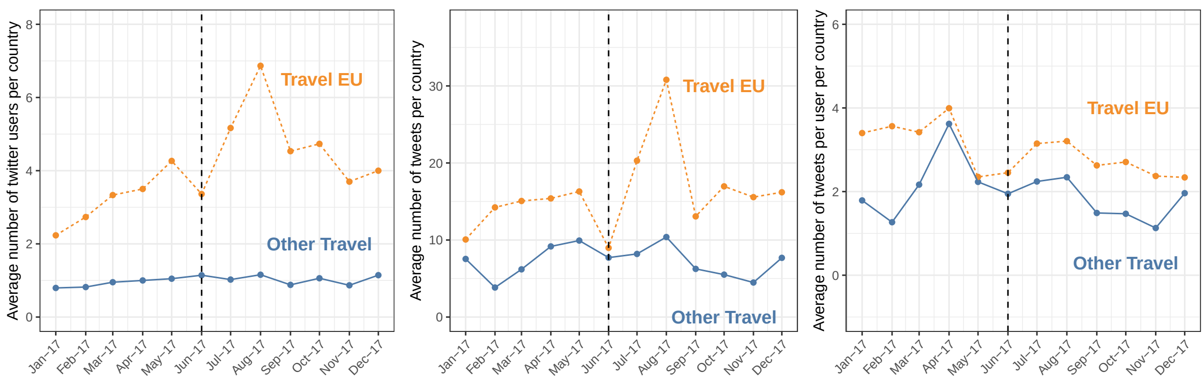

Figure C.14: Model free evidence of the impact of RLAH rules on user generated content on Twitter

Table C.19: Impact of roaming like at home on Twitter contributions

	Dependent variable:		
	Users	Tweets	Tweets/User
	(1)	(2)	(3)
EUTravel x After	1.648**	6.020*	−0.022
	(0.521)	(2.907)	(0.598)
Month FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
Observations	1,381	1,381	1,381
R^2	0.826	0.783	0.257
Adjusted R^2	0.808	0.762	0.183
Residual Std. Error	1.769	15.724	5.092

Note: + p<0.1; * p<0.05; ** p<0.01; *** p<0.001
Tweets in mobile devices while traveling
Cluster robust standard errors in ()
Errors clustered by country

C.4.4 Parallel trend assumption for the Twitter's dataset

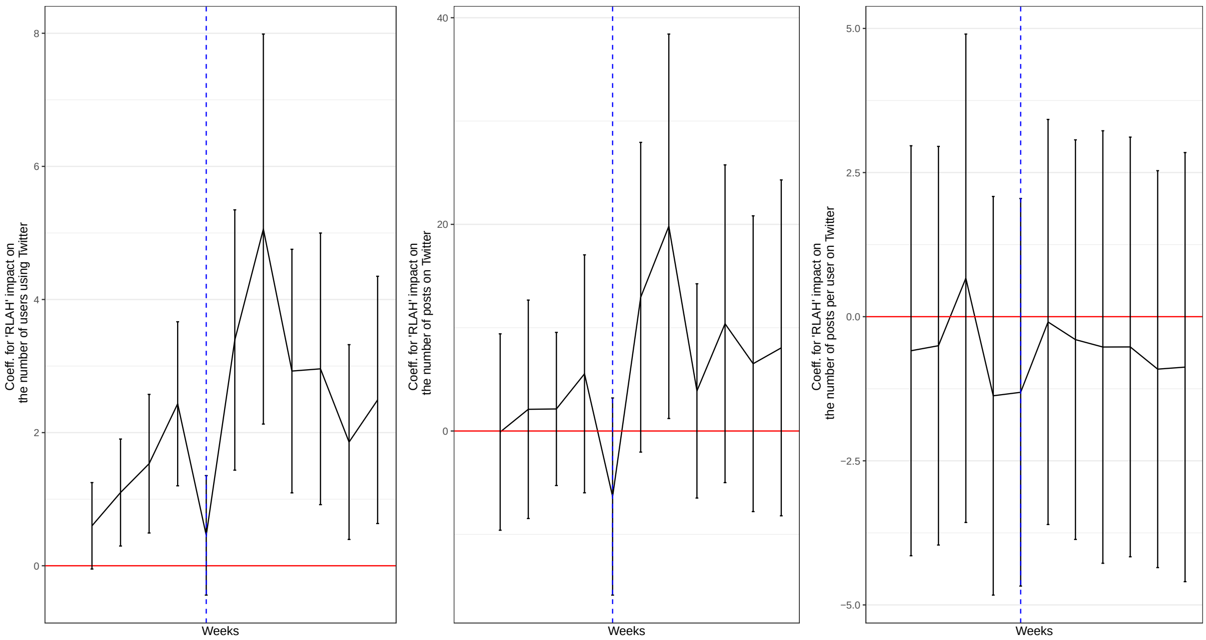

Figure C.15: Regression results for parallel trend assumption for Twitter's dataset

D Appendix: Online survey extract

Disclaimer

Thank you for contributing to our research project. We need your help to determine how people use their smartphones. Please read the following information carefully before deciding whether to participate in this research survey. Purpose of the research: Your information will be used in an academic study.

What you will do in this research:

You will answer a web survey. The survey contains questions about your use of mobile in different scenarios. The survey also asks demographic questions. You may exit the survey at any time.

Time required:

This survey should only take you about 5 minutes.

Risks:

The risk and discomfort associated with participation in this study is no greater than that experienced in everyday life. This means that you will not be taking any additional risks by choosing to participate in this study.

Benefits:

You will be entitled to a payment. At the end of the survey you will receive a personalized code. Your payment will be processed after you have entered your code on the Clickworker platform*. Information about the time and the method of payment will be reiterated at the end of the survey. In addition to the task completion fee, the knowledge that we may generate from your participation could be of value to society.*

Anonymity:

Your responses will be anonymous and stored in encrypted form on a secure server. Only members of the research team participating in this study will have access to your answers and will use it to inform their research.

Contact:

If you have any questions, concerns, or suggestions related to this study, the researchers can be reached at: < deleted for anonymity >. The research project is lead by < deleted for anonymity > and < deleted for anonymity > from < deleted for anonymity > and < deleted for anonymity > from < deleted for anonymity > .

Consent:

By selecting to continue, you indicate that you are at least 18 years old, and you agree to complete this survey voluntarily. You accept that the data you provide will be used for the purpose of academic research. You also accept that we may publish aggregate summaries of your answers in academic documents such as academic papers, thesis, and memoranda.

Q1 How much mobile data (per month) is included in your smartphone contract / prepaid option?

Please refer to your data allowance in your *home country*.

Please give your answer in megabytes. 1 gigabyte = 1000 megabytes. If you have an *unlimited* allowance, please write "99999".

____

Q2 How much do you pay for your smartphone contract / prepaid option (per month)?

Please give your answer in EUR.

____

<u>Scenario 1)</u>

You are traveling abroad for vacation or work. You have your smartphone, but no other mobile devices (laptop, tablet, etc.).

You have access to the internet on your phone, and your data allowance is **140 megabytes** that you can use for free.

Q3 Are you going to use the internet on your phone? Remember - you have 140 megabytes for free.

○ Yes

○ No

Q4 (If selected Yes at Q3) How many megabytes do you use for the following purposes?

In case you do not want to use your entire 140 megabytes, please select "unused megabytes" for the remainder.
Example: You only want to occasionally use Whatsapp. Then you choose 2 MB for "Communication" and 138 MB for "Unused Megabytes".

________ **Communication** (Whatsapp, iMessage...) 1MB: 250 messages or 2 pictures
________ **Getting information** (Google Search, Maps, Blogs...) 1MB: 10 searches
________ **Social media** (posting, reading) such as Instagram, Facebook... 10MB: 1 minute scrolling/posting
________ **Reading the news** (news from home country, international news) 10MB: 5 articles
________ **Music/video** (YouTube, Tiktok, Spotify...) 10MB: 2 songs or a 1 minute video
________ **Transportation** (Uber, Bolt, public transport...) 1MB: 3 uses
________ **Review platforms** (TripAdvisor, Yelp, Zomato...) 1MB: 2 reviews
________ **Unused Megabytes**

Scenario 2)

You are traveling abroad for vacation or work. You have your smartphone, but no other mobile devices (laptop, tablet, etc.).

You have access to the internet on your phone, and your data allowance is **exactly the same as in your home country**.
That is, your data allowance is <Entry value fo Q1> MB.

Q5 Are you going to use the internet on your phone? Remember - you have <Entry value fo Q1>

○ Yes

○ No

Q6 (If selected Yes at Q5)
Your data allowance *while traveling abroad* is <Entry value fo Q1> MB.
How do you use your data allowance for the following purposes?
Please give your answer in percent of your data allowance.

In case you do not want to use your entire data allowance, please select "unused megabytes" for the remainder.

________ **Communication** (Whatsapp, iMessage...) 1MB: 250 messages or 2 pictures
________ **Getting information** (Google Search, Maps, Blogs...) 1MB: 10 searches
________ **Social media** (posting, reading) such as Instagram, Facebook... 10MB: 1 minute scrolling/posting
________ **Reading the news** (news from home country, international news) 10MB: 5 articles read
________ **Music/video** (YouTube, Tiktok, Spotify...) 10MB: 2 songs or a 1 minute video
________ **Transportation** (Uber, Bolt, public transport...) 1MB: 3 uses
________ **Review platforms** (TripAdvisor, Yelp, Zomato...) 1MB: 2 reviews
________ **Unused Megabytes**

Q7 Take a moment to think about the following question:

Assume you did not have a data plan. How much money would you pay to be able to use <Entry value fo Q1> MB of data while you are traveling abroad?

	0 40 80 120 160 200
EUR	

Q8 (If selected 200 in Q7) You selected 200 EUR. Which of the following statements describes you best?

- ◯ I would pay exactly 200 EUR
- ◯ I would pay more than 200 EUR

Q9 How many days did you spend outside of your home country in 2019? Where did you travel to?

The countries of the European Union (EU) are Austria, Belgium, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden.

	Within EU	Outside EU
Number of days spent outside your home country in 2019		

Q10 What do you do when traveling outside of your home country if the internet on your phone is expensive? (Check all that apply)

- ☐ I will not use the mobile internet on my phone.
- ☐ I will use Wifi/WLAN whenever I can.
- ☐ I will buy a local (prepaid) SIM card.
- ☐ I don't think it's too expensive to use the mobile internet on my phone.

Q11 What is your gender?

- ○ Male
- ○ Female
- ○ Other ____
- ○ I don't want to answer

Q12 What is your age?

- ○ Under 18
- ○ 18 - 24
- ○ 25 - 34
- ○ 35 - 44
- ○ 45 - 54
- ○ 55 - 64
- ○ 65 - 74
- ○ 75 - 84
- ○ 85 or older

◯ I don't want to answer

Q13 In which country do you live?

(select Non-EU country if you do not live in one of the member countries of the European Union)

▼ Austria ... Sweden

Q14 What is the highest level of school you have completed or the highest degree you have received?

◯ Less than high school degree

◯ High school graduate (high school diploma or equivalent)

◯ Some college but no degree

◯ Bachelor's degree

◯ Master's degree

◯ Doctoral degree

◯ Professional degree

◯ I don't want to answer

Q15 Please indicate your occupation:

◯ Armed Forces Occupations

◯ Managers

◯ Professionals (e.g. Teaching, Healthcare, Science & Engineering, Legal, Business Administration, ICT)

◯ Clerical Support Worker (e.g. Keyboard Clerks, Secretaries, Customer Service Clerks)

◯ Service and Sales Workers (e.g. Personal Services, Sales Workers, Personal Care, Protective Services)

◯ Agricultural, Forestry and Fishery Workers

◯ Plant and Machine Operators (e.g. Assemblers, Drivers and Mobile Plant Operators, Stationary Plant and Machine Operators)

◯ Craft and Related Trade Workers (e.g. Metal, Machinery, Handicraft, Printing, Electrical and Electronics, Food Processing, Woodworking)

◯ Elementary Occupations (e.g. Cleaners and Helpers, Mining, Construction, Manufacturing and Transport, Food Preparation)

◯ Student

◯ Retired

◯ Unemployed

◯ I don't want to answer

◯ Other ____