NBER Research Proposal

(Un)Informed consumers and (Un)Observable product quality: the effect of on-line information shocks on urban entrepreneurship

Dante Donati
Universitat Pompeu Fabra

1. Introduction

The digital era has transformed the way that consumers learn about product quality. The internet has made information about goods and services more readily available, allowing consumers to make more informed choices. In this respect, online review platforms are one important source of information. For example, TripAdvisor — the world's largest travel review platform — contains 630 million reviews. An extensive and growing body of literature documents the influence of online user reviews on the quantity and price of online transactions (e.g., Resnick et al., 2006; Chevalier and Mayzlin, 2006; Anderson and Magruder, 2012; Berger, Sorensen, and Rasmussen, 2010). More recent studies explicitly analyze the effect of online ratings on travelers' demand and restaurant revenues (e.g. Lewis and Zervas, 2017; Luca, 2016) as well as their interaction with regulatory compliance (Farronato and Zervas, 2018).

However, most of these studies lack exogenous variation in the production or consumption of online reviews. Moreover, they often focus on a narrow range of variables of interest, hence the scope for policy implications remains limited.

The primary goal of this project is to study how increased market transparency, which lowers search costs and allows consumers to learn about product quality through online review platforms, affects service industries. In particular, it will examine how a sharp decrease in the costs of mobile internet caused by the abolition of roaming charges in Europe affected the restaurant industry. The research exploits the approval by the European Parliament of a new policy on the EU roaming regulation ("Roam like at home") which led to the abolition of all roaming charges for temporary roaming within the EEA as of June 15, 2017. In practice, if before that day all EU residents travelling within the EEA were charged high prices for data consumption, after the policy the same home network rate/promotion is applied. This resulted in a sharp increase in data usage for EU residents when visiting other EU countries.

The study will examine the effect of this policy change on the content and usage of online review services, such as TripAdvisor, as well as labor market outcomes of restaurants. Hence, the research aims to contribute to the existing literature in two ways. By exploiting a quasi-experiment based on the exogenous drop in the internet roaming tariffs that EU travelers experienced. By linking the surge in online information availability with the production and consumption of TripAdvisor reviews to study how, eventually, profitability and labor demand of restaurants responded. Therefore, the main contribution comes from the combination of a methodological novelty with a wider range of variables of interest (labor market outcomes) for the literature on digitization economics.

2. Data

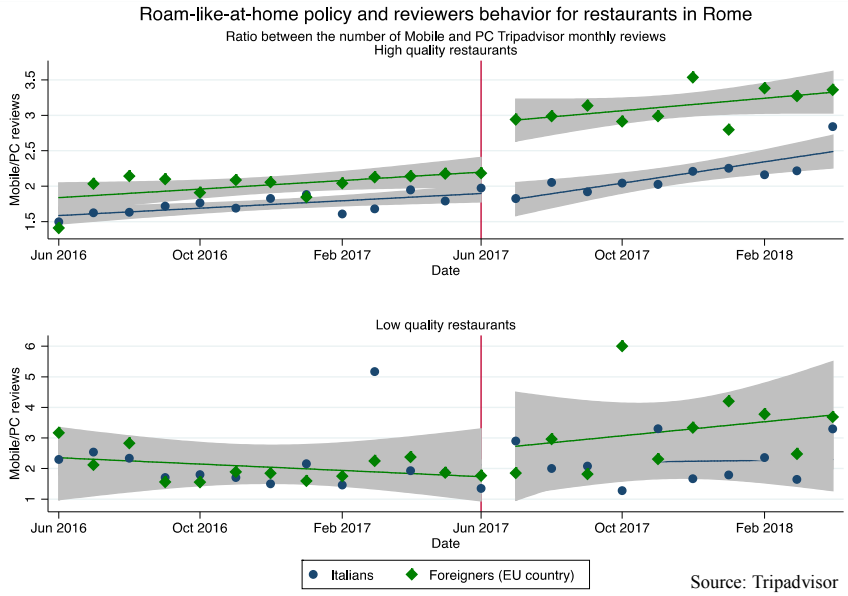
The research focuses on Italy and, more precisely, on medium-large cities with significant touristic flows. I combine two data sources. The first one is TripAdvisor, the most popular review platform in the country and in Europe. I scrape information on restaurants (name, address, category, rating, ranking) as well as, for each restaurant, the entire historical record of reviews (date, device, reviewer information, country/language, text, etc.).

I plan to match each TripAdvisor restaurant with its social security records collected and maintained at the Italian National Social Security Institute (INPS). Data access in only granted to researchers visiting the institute in the context of a specific program, upon approval of the

scientific committee. NBER funding can be fundamental to cover expenses related to data access and processing. These administrative data are extremely rich and come at high frequency: for each establishment, they contain information on location, sector, date of opening and closure, as well as monthly information on number of employees, part time employees and the wage bill. Similar data are increasingly used to study labor market outcomes (e.g., Boeri and Garibaldi, 2018).

3. Research design and outcomes

The research relies on the assumption that the information (internet) shock decreased tourists' uncertainty regarding the quality of different restaurants. As the share of potential customers relying on TripAdvisor increases, online reviews, rating and ranking become more influential. At the same time, the number of online contributions from mobile devices is expected to rise. A preliminary analysis I have conducted on reviewers' behavior for the city of Rome indicates a shift from PC- to Mobile-based contributions following the implementation of the policy. Importantly, this effect is observed only for EU tourists but not for Italians (who were not directly affected by the policy). Interestingly, the jump is much sharper for high-quality restaurants. This suggests that the online platform is used both ex-ante, in the screening process, and ex-post, in the contribution process, and that better quality restaurants may experience higher gains in terms of larger demand.



Therefore, the information shock that tourists faced is useful to isolate the demand-side channel through which online reviews and particularly the salience of quality (rating/ranking) may impact on local entrepreneurship. The channel is also supported by recent literature showing that hotels' advertising spending reactions to reviews operate through the demand-side (e.g. Hollenbeck et al., 2017).

To study the impact of online information availability on labor market outcomes, the research design relies on an Instrumental Variable approach where the instrument is given by the "Post" policy dummy variable. Additional variation may come from the spatial dimension: for instance, one option is the comparison of similar restaurants located in touristic (or more visible) and non-touristic ("hidden") areas of a city. Also, the high quality and frequency of the data (monthly level) together with the establishment-time level of the analysis enrich the identification strategy (city-month fixed-effects/restaurant fixed-effects).

The primary outcome of interest is the (re)allocation of labor before/after the policy across restaurants with different rating (or ranking) as of June 15, 2017. Variables under scrutiny are the overall number of employees, the type of contracts (part- vs. full-time) and wages. In addition, exit of restaurants across the rating distribution will be examined. In particular, I will try to understand if, one year after the policy, better quality providers could drive away low-

quality providers, thus partially solving the classic lemons market problem (Akerlof, 1970).

Moreover, in a context of dynamic quality (such as the one examined by Chevalier et al. (2018)), "managerial responses" to online reviews may also signal a better effort in improving the quality of service. This can also be measured by comparing restaurant ratings before/after the policy, or studying labor force turnover (new staff replacing old staff) by type of restaurant. This analysis would shed light on the hypothetical providers' effort to achieve mobility along the quality distribution as a result of increased market transparency. A final, yet very interesting outcome is provided by the potential redistribution of wages across different positions/tasks within each restaurant.

4. Timeline and budget estimates

This research, along with my other working papers described in my CV, will be part of my PhD dissertation. The visiting program at INSP (namely the "Visitinps scholars of type B") will soon accept new applications. The program requires to follow a strict timeline. Researchers are allowed to spend some months at the institute headquarter in Rome (or Milan, depending on the availability of the archives) to conduct their analysis. Raw data cannot leave the office. My plan is to carry out the majority of the workload within 1 year, so that the release of a first draft is planned by June/July 2019.

As on the budget, the main expenses are planned to be related to the following items.

(1) Travels and accommodations in Rome (or Milan):
3-4 month-stay for the first visit, and 1-2 additional months for a second-third visit. 5-6 months in total. I budget approximately $ 6,000-7,000.

(2) Data storage and processing power (use of servers):
Preliminary analysis seems to suggest interesting outcomes for restaurants on the tails of the quality-distribution. However, these restaurants are not so many when only one city is considered. To ensure enough power for the econometric analysis the sample should be expanded to include all medium-large cities. Given the extremely large amount of online TripAdvisor reviews (only in Rome, for the last 2 years, I scraped approximately 3 million reviews), additional storage and processing power are needed.
I budget approximately $ 2,000.

(3) Research assistants
In addition to the average rating (quality) of a restaurant, another important dimension of the analysis is its ranking on TripAdvisor at the time of the policy. In fact, ranking matters when consumers search for a restaurant, and the fact that each TripAdvisor page contains 30 restaurants may have implications on their visibility before and after the policy. This justify my willingness to study TripAdvisor ranking algorithm so to reconstruct the situation as of June 15, 2017. This way, labor market outcomes will be analyzed across restaurants occupying different rankings (pages) at the time the information shock took place. The study of the algorithm may represent an additional contribution to the literature. In addition, machine learning algorithms will be used to study how reviews' text and managerial responses evolved. For these purposes I need assistantship from a data scientist/IT expert.
I budget approximately $ 3,000, for a 2-month job.

Finally, additional assistantship is expected to be needed for matching TripAdvisor entries with the data from the social security archives. In fact, for a subset of restaurants, the TripAdvisor name is likely to differ from the one on the archives, which usually corresponds to the legal name of the company owning the place. Additional archives will need to be consulted.
I budget approximately $ 3,000, for a 2-month job.

Total approximate budget: $ 14,000-15,000.

DANTE DONATI

Office 20.152
Department of Economics and Business
Universitat Pompeu Fabra, Barcelona

Mobile: +34-657682158
Office: +34-935422696
Email: dante.donati@upf.edu

DOCTORAL STUDIES

2016- Universitat Pompeu Fabra, Barcelona, Spain. PhD candidate in Economics
Affiliations: UPF, BGSE, IPEG and LEAP

Research interests:
Political Economy, Economics of Technology, Media Economics, Development Economics

References:
Maria Petrova, maria.petrova@upf.edu
Ruben Enikolopov, ruben.enikolopov@upf.edu
Marta Reynal-Querol, marta.reynal@upf.edu
Ruben Durante, ruben.durante@upf.edu

PRE-DOCTORAL STUDIES

2013-2016	Bocconi University, Milan, Italy. MSc in Economic and Social Sciences
2014-2015	Chulalongkorn University, Bangkok, Thailand. Visiting program
2013	Northwestern University, Evanston, IL, US. Visiting program
2010-2013	Bocconi University, Milan, Italy. BA in Economic and Social Sciences

References:
Eliana La Ferrara Nicola Gennaioli Francesco Giavazzi

TEACHING AND TUTORING

Development Economics, Master course, Barcelona Graduate School of Economics
Econometrics, Undergraduate course, Universitat Pompeu Fabra
Data Analysis, Undergraduate course, Universitat Pompeu Fabra
Probability and Statistics, Undergraduate course, Universitat Pompeu Fabra

RESEARCH EXPERIENCE AND SKILLS

2017-	Principal Investigator and Consultant, The World Bank, DIME
2016-	Research Assistant, PI: Marta Reynal-Querol, UPF
2015-2016	Research Assistant, PI: Nicola Gennaioli, Bocconi University
2014-2016	Research Assistant, PI: Abhijit Banerjee, Eliana La Ferrara and Victor Orozco

Languages: Italian (native), English (fluent), French (good), Spanish (basic)
Programming: Stata, LaTex, R, GIS, Python, Matlab, SurveyCTO

SCHOLARSHIPS, GRANTS AND AWARDS

2016	LEAP Student Grant, Project: "The effect of digital technologies on behavior"
2016	Fondazione Grazioli, MSc Thesis Award, "The roots of domestic terrorism: evidence from Boko Haram attacks in Nigeria"
2010-2015	Bocconi University Scholarship

PROFESSIONAL ACTIVITIES

2018 NBER Economics of Digitization Tutorial, Stanford; NBER Economics of Digitization Meeting, Stanford; Digital Economics Conference, Toulouse School of Economics; World Bank workshop on Mass Media Entertainment Education Impact Evaluation, New Delhi; GSE-UPF Media Seminar, Barcelona; Workshop on Political Economy of Development and Conflict, Barcelona

2017 Annual Media Economics Workshop, Barcelona; GSE-UPF Media Seminar, Barcelona

Refereed articles for:
Journal of Public Economics

WORKING PAPERS

Mobile Internet access and political outcomes: Evidence from South Africa, 2018, ready for submission.

Does mobile Internet arrival affect political participation and electoral competition in developing countries? I provide an empirical answer to this question looking at the South African municipal elections results between 2000 and 2016. I exploit the temporal and geographical variation in 3G Internet coverage to estimate its impact on (1) the vote shares of the major parties, (2) voter turnout, (3) electoral competition and (4) protests. Using a high-resolution newly constructed dataset along with a Diff-in-Diff and 2SLS estimation, I show that in 2016 Internet availability caused a reduction in the vote share of the ruling party by almost 7 pp. The main opponents have gained from the Internet arrival. Political participation, political competition and number of protests increased. Then, I develop an extensive analysis of the potential mechanisms and show that in localities more exposed to corruption and administrative scandals the impact of 3G arrival was larger. This points out the information role the Internet played in South African politics. Finally, I conduct a spatial analysis to study how opinion diversity in the pre-Internet era interacts with individual information from the web. I show that Internet penetration fostered convergence of preferences over space.

WORK IN PROGRESS

(Un)Informed consumers and (Un)Observable product quality: the effect of on-line information shocks on urban entrepreneurship

The health cost of the digital era: the effect of the Internet on mental disorder, with Ruben Durante and Francesco Sobbrio

Can social media Edu-tainment reduce gender-based violence? An experimental pilot using Facebook Messenger in India, with Victor Orozco (DIME, The World Bank)