Universitat Pompeu Fabra
Barcelona

Departament d'Economia i Empresa

DOCTORAT EN ECONOMIA, FINANCES I EMPRESA

Essays in the Economics of Information

Author:
Dante Donati

Thesis Supervisor:
Maria Petrova

TESI DOCTORAL UPF / 2022

To my wife Claudia, and my children Edoardo and Vittorio,
the joy of my life.

Acknowledgements

I thank my advisor and mentor Maria Petrova for her guidance, encouragement, and patience throughout the last 6 years.

I would also like to thank Ruben Durante, Gianmarco León-Ciliotta, Sandro Shelegia, Rosa Ferrer, Giacomo Ponzetto, Albrecth Glitz, Jared Gars, Nandan Rao and Pinar Yildirim for their comments and help, especially throughout the job market. Marta Araque and Laura Agusti provided amazing support throughout the Ph.D.

In the last years, I had the great opportunity to learn from Ruben Enikolopov, Nicola Gennaioli, Matthew Gentzkow, Eliana La Ferrara, David Krisztian Nagy, Marta Reynal-Querol and Alessandro Tarozzi, as well as from all the participants of various UPF seminars and conferences around the world.

I gratefully acknowledge generous financial research support from IPEG, LEAP Bocconi, the World Bank's i2i and Umbrella Facility for Gender Equality trust funds, and great assistance from INPS. I thank the Spanish Ministry of Economy for the generous financial support throughout my studies.

I am indebted to my mother Graziella and my father Danilo, for believing in me and giving me the opportunity to study at prestigious academic institutions. My wife Claudia and my children Edoardo and Vittorio have been a fundamental source of inspiration and support throughout the last years. I dedicate my achievements to them.

Finally, I would like to thank my friends Stefano and Dario, as well as my Ph.D. companions Andrea, Sampreet, Philipp, Laurenz and Andreu for all the amazing memories.

Dante Donati
Pieve Santo Stefano
June 2022

Abstract

This dissertation consists of three essays that empirically test economic theories on the causal effect of information on individuals and firms. The first chapter studies the consequences of lower consumer search costs on restaurants' incentives to upgrade quality. It shows that cheaper access to online information from Tripadvisor reallocates demand to higher-rating establishments. In turn, lower-rating restaurants are more likely to exit, while the surviving ones hire workers with higher wages and better curricula, eventually improving their Tripadvisor ratings. The second chapter provides evidence on the political effects of mobile internet arrival in South Africa. Using granular spatial data on 3G coverage combined with municipal election results, it demonstrates that internet availability increases voter turnout and political competition, eventually damaging the popularity of the incumbent party. The third chapter investigates the potential for social media to deliver educational-entertainment content and change individuals' attitudes and behavior towards gender norms and violence against women (VAW) in India. It shows that short video-clips delivered through Facebook Messenger shift gender norms attitudes towards more progressive stances, reduce social acceptability of VAW, and promote information-seeking and posting behaviors against VAW on the web.

Keywords: Review platforms; Asymmetric information; Search costs; Quality; Media; Mobile internet; Municipal elections; Political outcomes; Edutainment; Gender norms; Social media; Violence against women.

Resumen

Esta tesis consta de tres ensayos que prueban empíricamente teorías económicas sobre el efecto causal de la información en individuos y empresas. El primer capítulo estudia las consecuencias de menores costos de búsqueda del consumidor sobre los incentivos de los restaurantes para mejorar la calidad. Muestra que el acceso más económico a la información en línea de Tripadvisor reasigna la demanda a establecimientos con reseñas más altas. A su vez, los restaurantes con reseñas más bajas tienen más probabilidades de salir, mientras que los que sobreviven contratan trabajadores con salarios más altos y mayor experiencia, lo que eventualmente mejora sus reseñas en Tripadvisor. El segundo capítulo estudia los efectos políticos de la llegada de internet móvil a Sudáfrica. Utilizando datos espaciales granulares sobre la cobertura 3G combinados con los resultados de las elecciones municipales, demuestra que la disponibilidad de internet aumenta la participación electoral y la competencia política, lo que eventualmente daña la popularidad del partido en el poder. El tercer capítulo investiga el potencial de las redes sociales para ofrecer contenido de entretenimiento educativo y así cambiar las actitudes y el comportamiento de las personas hacia las normas de género y la violencia contra las mujeres (VCM) en India. Muestra que vídeos cortos entregados a través de Facebook Messenger cambian las actitudes de las normas de género hacia posturas más progresistas, reducen la aceptabilidad social de la VCM y promueven comportamientos de búsqueda de información y publicación contra la VCM en la web.

Palabras clave: Plataformas de revisión; Información asimétrica; Costos de búsqueda; Calidad; Medios de comunicación; Internet móvil; Elecciones municipales; Resultados políticos; Entretenimiento educativo; Normas de género; Redes sociales; Violencia contra las mujeres.

Preface

This dissertation consists of three essays that empirically test economic theories on the causal effect of information on individuals' and firms' behavior. To do so, I take advantage of natural and randomized controlled experiments, as well as the collection of novel unstructured data from digital sources. The thesis validates the hypothesis that information, even when provided through digital technologies, has profound implications on the economy and society.

In the first chapter, I study the consequences of lower information frictions for consumers on firms' incentives to upgrade product quality in markets under asymmetric information. While economic theories generally predict a positive relationship between consumer information and product quality, the empirical literature faces two main limitations: (1) access to - and the provision of - information is usually endogenous; (2) data on product quality are difficult to get and, if available, they are often incomplete. I overcome these limitations exploiting a natural experiment and a unique newly-assembled dataset. I derive theoretical predictions from a consumer search model and empirically investigate the demand and supply effects of online reviews on the restaurant industry in the province of Rome. For identification, I take advantage of the 2017 policy that abruptly abolished mobile roaming charges in the EU, and assemble a novel dataset combining monthly information on the entire historical records of reviews hand-collected from Tripadvisor with rich administrative establishment-level data. I find that, after the policy, revenues and total employment in mid- and high-rating restaurants grow by 3-10%. In turn, the probability for low-rating restaurants to exit the market doubles compared to the pre-policy period, while surviving low- and mid-rating establishments hire workers with higher wages and better curricula, eventually improving their Tripadvisor ratings. Overall, the share of low-rating restaurants in the most tourist areas decreases by 2.5 percentage points. My findings have implications for the role of review platforms in the performance of offline industries under asymmetric information.

In the second chapter, I study the political consequences of mobile internet arrival in developing countries. Existing literature primarily focused on developed economies finds that broadband internet generally led to lower political participation, most likely due to the increase in entertainment and crowding out of news consumption. I show that these effects can be different in developing countries, where the traditional media sector is overall less competitive and the level of political information in the pre-internet era used to be low. I look at South Africa, and leverage granular spatial data on 3G coverage from the GSMA and combine them with administrative records on political participation, electoral competition, voters' preferences and protests at the municipal level. Using a difference-in-difference strategy, I show that in 2016 mobile internet availability caused a 2 percentage point (p.p) increase in voter turnout and a 3 p.p. reduction in the vote share of the ruling party. The main opponents gained from mobile internet arrival. The number of parties running for election and the number of protests increased. I conclude providing suggestive evidence that both information and coordination mechanisms could explain the observed results.

The third chapter investigates the potential for social media to deliver educational-entertainment (edutainment) content and change individuals' attitudes and behavior towards gender norms and violence against women (VAW). Most development interventions aimed at preventing gender-based violence are delivered through grassroots mobilization campaigns and tend to be resource-intensive. Can social and behavioral change communication campaigns delivered through social media be a cost-effective alternative? We conduct an online RCT to test the impact of two short edutainment campaigns delivered through Facebook Messenger in India. Individuals were randomly assigned to watch video-clips where educational messaging was either implicit (a humorous fake reality TV drama) or explicit (a docu-series about VAW with clear calls to action). We collected self-reported and objective online outcomes to measure impacts after one week and four months. Our findings suggest that edutainment delivered through social media is effective at changing attitudes towards gender norms, reducing social acceptability of VAW, and promoting information-seeking and posting behaviors on the web. Short-term effects of the implicit format on knowledge, norms and clicks on informative links oscillated between 0.16 and 0.21 standard deviations, yet they diminished in the medium term. Medium-term results, on the other hand, show that the explicit format made individuals 91% (7.5 p.p.) more likely to add a frame against VAW in their Facebook profile picture, a public display of their disapproval of this harmful practice. Our findings are particularly promising in settings where resource-intensive campaigns may be costly to scale.

Contents

Chapter 1

The end of tourist traps: a natural experiment on the impact of Tripadvisor on quality upgrading

1.1 Introduction

Asymmetric information can distort market outcomes in different ways. For example, when product quality is imperfectly observed, its equilibrium levels are too low, leading to significant welfare losses for both consumers and producers (Akerlof 1970; Leland 1979). Theoretically, removing information frictions should attenuate market inefficiencies by lowering prices or improving qualities (Chan and Leland 1982; Salop and Stiglitz 1977). In this respect, the arrival of the internet was expected to reduce and homogenize market prices, but, in the end, its effects are thought to be limited (Ellison and Fisher Ellison, 2005).[1] Did the internet then have a more positive influence on firms' decisions over product quality? By helping consumers make more informed choices, online review platforms are expected to create reputation mechanisms and enhance firms' incentives to upgrade product quality (Goldfarb and Tucker 2019; Tadelis 2016). Yet, despite

[1]Empirical studies have documented relatively lower online prices but also the persistence of substantial online price dispersion in the markets for CDs, books, insurance, cars and airlines (Brown and Goolsbee 2002; Brynjolfsson and Smith 2000; Morton et al. 2001; Orlov 2011). More recently, Cavallo (2017) finds that online-offline prices of large multi-channel retailers are identical about 72 percent of the time.

the relevance for regulatory decisions, empirical evidence is scarce[2] due to two main limitations: (1) access to - and the provision of - information is usually endogenous, exacerbating any causal assessment;[3] (2) data on product quality are difficult to obtain and, if available, they often come from inspection records and online ratings, which are not informative about firm-specific investment decisions and industry composition.

In this paper, I exploit a natural experiment and unique data from Tripadvisor matched to confidential administrative employer-employee records to study how lower information costs affect consumers' behavior and firms' incentives to upgrade product quality. I derive theoretical predictions from a consumer-search model and empirically investigate the demand and supply effects of online reviews on the restaurant industry in the province of Rome. I assemble a novel dataset combining monthly information on the entire historical records of reviews collected from Tripadvisor with rich administrative establishment-level data. For identification, I take advantage of the 2017 policy that abruptly abolished mobile roaming charges in the EU, generating an arguably exogenous variation in the costs for travelers to access online reviews. This setting allows me to estimate the aggregate effects of a policy that was not deliberately designed to affect the restaurant industry.[4] Moreover, detailed firm-level data allow me to study firms' response by directly looking at changes in production costs and hiring decisions, which are used as proxies for quality upgrading.[5]

To inform the empirical exercise, I first build a model in which consumers with heterogeneous search costs engage in costly sequential search to buy one unit of a vertically differentiated product while firms with heterogeneous abilities endogenously select into production and compete in quality. The model borrows insights from theoretical papers on search costs and firms' strategic responses (Bar-Isaac et al. 2012; Fishman and Levy 2015; Goldmanis et al. 2010) but, in contrast to them, abstracts from price competition[6]

[2]Some exceptions include Farronato and Zervas (2019) who examine restaurants' incentives to comply with hygiene regulation and Liu et al. (2021) who compare efficiency of Uber and taxi drivers.

[3]This is a well-known challenge to identifying the causal impact of any type of reputation and information disclosure on demand and supply (e.g., Eliashberg and Shugan 1997).

[4]By contrast, Jin and Leslie (2003) and Klein et al. (2016) study very market-specific quality disclosure programs and feedback systems in online platforms.

[5]Others, such as Ananthakrishnan et al. (2019) and Proserpio and Zervas (2017), rely exclusively on online ratings to proxy for quality.

[6]This assumption simplifies the algebra and allows me to solve the model analytically. One potential caveat is that the predictions of the model are just one special case of a more general setting featuring endogenous prices. For instance, Fishman and Levy (2015) study both price and quality competition, yet their framework does not deliver predictions on firm dynamics, which is an important feature of my

and focuses exclusively on firms' incentives to upgrade quality. I show that, when consumers' search costs decrease: (i) the demand faced by firms that were *ex ante* selling high-quality goods increases; (ii) the overall quality level in the industry improves, and this is driven by both the exit of lower-quality providers and the investment in quality upgrading of all surviving firms - and more so of firms selling *ex-ante* lower-quality products.

To empirically study the demand and supply effects of lower consumer search costs, I assemble a novel dataset, which combines information from Tripadvisor, the most widely used review platform in the EU, with administrative employer-employee records maintained at the Italian National Social Security Institute (INPS). For about 5,500 matched restaurants in the province of Rome, the data contain time-invariant information on name, address, price category, type of cuisine, legal status of the firm and additional covariates, as well as time-varying information on number of Tripadvisor reviews (by origin of reviewer and device), average rating, date of opening and closure, number of employees, type of contracts, wages and the full employment history of the workers, observed at monthly intervals between 2007 and 2019. Moreover, for a subset of this sample, the data provide annual information on income statements and balance sheets.

I then use the access to online reviews on Tripadvisor to proxy for reductions in consumer search costs and estimate the impact on the restaurant industry. The identification strategy exploits the approval by the European Parliament of new EU roaming legislation,[7] which abolished all charges for temporary roaming within the European Economic Area (EEA) as of June 15, 2017. In practice, before that day all EU residents traveling within the EEA were charged an additional price for data consumption on top of the home network rate. After the policy, the same home network rate is applied.

I show that, after the policy, Tripadvisor users from EU countries became 1.4 times more likely to post reviews on their mobile devices as opposed to PC. Moreover, the number of mobile reviews written by EU users, as well as their total (mobile+PC) number of

model.

[7]Specifically, the Regulation (EU) 2015/2120 of the European Parliament and of the Council of 25 November 2015 "aims to establish common rules to safeguard equal and non-discriminatory treatment of traffic in the provision of internet access services and related end-users' rights." The subsequent Commission Implementing Regulation (EU) 2016/2286 of 15 December 2016 states that "roaming providers should not levy any surcharge additional to the domestic retail price on roaming customers in any Member State [...]." https://europa.eu/youreurope/citizens/consumers/internet-telecoms/mobile-roaming-costs.

reviews substantially increased after the policy compared to those from the locals. By contrast, reviews from extra-EU and Italian travelers did not exhibit a significant change. Importantly, these results are not driven by an increase in international tourist flows toward Italy. Thus, the policy provides an abrupt and arguably exogenous source of variation in the use of Tripadvisor services by EU travelers, whose reviews constitute about 30% of the total volume in restaurants located in the most tourist areas of Rome.

To identify the parameters of interest, I combine the temporal variation induced by the policy with the spatial variation in tourist demand. In particular, I take advantage of the granularity of my data and construct two measures of restaurants' exposure to tourist clientele that account for the intensity to which each restaurant is potentially affected by the lower information costs induced by the policy. The first measure reflects the probability of finding a restaurant given its location with respect to the closest tourist attraction and the road network around it. In practice, I use information from Google Maps API to define the partial road network that leads to the closest restaurants around each attraction and compute the probability of finding each of these restaurants while walking away from the attraction. I show that higher probability values are positively correlated with the share of reviews from foreigners, while they are negatively correlated with the restaurants' average rating.[8] The second measure exploits the variation in the number of attractions across ZIP codes as a proxy for potential exposure of all restaurants in a ZIP-code to tourist clientele and, therefore, to the change in information costs.

The identification strategy relies on a Difference-in-Differences specification, which compares the changes over time in the outcomes (on a 5-year period before/after the policy) across restaurants that are differentially exposed to tourist clientele. Particularly, in the baseline specification, I use the median value of the previously-described probability measure to create a binary treatment indicator for high *vs.* low tourist exposure. I conduct the analysis on the sample of restaurants with at least one review in the pre-policy period, as well as on three sub-samples of equal size defined using the tertiles of the restaurants' rating at the time of the policy – namely, low (< 3.85), medium ($\in [3.85, 4.25)$) and high-rating (≥ 4.25) categories –, and run the regressions on each group separately. This allows me to assess the presence of differing effects of online information across establishments. The identification assumption requires that, within

[8]As theory predicts (e.g., Chan and Leland 1982) firms that more frequently engage with uninformed consumers tend to under-provide quality.

each group, changes in the outcomes across restaurants with high and low exposure to tourists would have been the same in the absence of the policy.

I then test the theoretical predictions of the model. From the first one, I expect consumers to reallocate their demand toward restaurants with *ex-ante* higher Tripadvisor ratings, which therefore should expand their sales and employment (output). I obtain several empirical results supporting this prediction. After the policy, annual revenues increased by almost 7% in high-rating restaurants, by approximately 3% in mid-rating ones, and remained the same in the low-rating category, with this positive gradient being statistically significant. As a result, revenues increased by almost 5% overall, pointing out an average growth in sales by approximately 32.5 Thousands Euros a year.[9] Average firm size also changed after the policy. Total monthly employment expanded by approximately 4% in more tourist restaurants, compared to less tourist ones. With an increase by 10%, the mid-rating category is mostly responsible for the overall growth. Moreover, the impact appears to be negative for low-rating restaurants, while positive for high-rating ones, yet, in both cases, coefficients are insignificant. The latter indicates the presence of decreasing returns to labor in the restaurant industry, with high-rating establishments already producing at full-capacity in the pre-policy period.

The second set of theoretical predictions concerns the supply side. In particular, I expect to observe higher exit rates for lower-rated establishments, as well as investment into higher quality inputs (such as hiring more qualified workers) for all operating restaurants, and particularly for those with *ex-ante* lower ratings. I start with the analysis of firm exit. I find that, for low-rating establishments, the probability to exit the market doubled after the policy, compared to the baseline period. By contrast, the policy did not significantly impact the probability that mid- and high-rating restaurants left the industry.[10] Moreover, by aggregating observations at the ZIP-code level to study the effect of exit and entry jointly, I find that the share of low-rating firms operating in the most touristy neighborhoods decreased by 2.5 pp after the policy, compared to non-touristy ZIP codes. These results suggest that lower search costs – even when experienced by only a fraction of consumers – can alleviate the adverse selection problem and make the industry more quality-oriented.

Then, I analyze the behavior of operating firms. I consider hiring decisions as a proxy

[9] These results are in line with those from papers studying the impact of review platforms on firms' sales (e.g., Anderson and Magruder 2012; Chevalier and Mayzlin 2006; Luca 2016, among others).

[10] This result is In line with findings by Hui et al. (2018).

for the restaurants' effort to upgrade the service quality through the recruitment of more experienced staff. I find that the probability of hiring a worker with previous experience in the restaurant industry increased overall by 10% with respect to the pre-policy mean. This effect is driven by low- and mid-rating establishments, where such probability went up by 9 and 16%, respectively. By contrast, the coefficient for high-rating restaurants is close to 0 and not statically significant. Moreover, additional evidence suggests that low-rating establishments accumulated human capital at the expenses of high-rating ones. As a result, daily salaries paid by low-rating restaurants increased by more than €1 (2% of their pre-policy mean), while they decreased by a similar amount in high-rating restaurants. Eventually, these opposite recruiting strategies had differing effects on the online reputation of the restaurants, as measured by the average 5-month Tripadvisor rating. In particular, I find that restaurants in the low- and mid-rating groups received better reviews after the policy, as their 5-month Tripadvisor rating improved by 0.09 points (2.5%) and 0.08 points (1.9%), respectively.[11] By contrast, reputation remained unchanged in restaurants that were already at the top of the rating distribution. Estimates tend to be even larger when restaurants that exited the market are excluded from the sample. Overall, these results point out the role of review platforms in alleviating the moral hazard problem in the experience goods market.

I carry out a number of placebo exercises to validate the identifying assumption. For example, I report event-study estimates, which confirm the absence of diverging trends in the outcomes before the roaming regulation. Moreover, a series of policy-permutation tests conducted in the pre-policy period provides further evidence on the exogeneity of the policy date with respect to other potential factors (such as seasonality) that might explain the observed results. Finally, I show that the main estimates are robust to the use of alternative measurements, samples and clustering units.

My results suggest that review platforms have important economy-wide consequences on the whole Italian restaurant industry. Back-of-the-envelope calculations suggest that abating the costs for all consumers to access Tripadvisor leads to an overall increase in restaurant revenues, employment and exit rate by 1.6%, 1.5% and 0.5 pp, respectively. The first two figures correspond to about 12% and 5% of the overall growth in revenue and employment experienced by restaurants between 2016 and 2019, respectively. While the last figure corresponds to almost 3% of the exit rate faced by the industry during the

[11] These results are similar to those by Ananthakrishnan et al. (2019) and Proserpio and Zervas (2017).

first year of the Covid-19 pandemic. All together, these results indicate that facilitating access to review platforms can have real effects on the performance and composition of firms operating in industries generally affected by asymmetric information.

This paper contributes to several strands of literature. First, existing studies estimate the effects of reviews and ratings on sales, and find that online reputation is a significant driver of them (Anderson and Magruder 2012; Cabral and Hortacsu 2010; Chevalier and Mayzlin 2006; Fang 2022; Lewis and Zervas 2019; Luca 2016; Reimers and Waldfogel 2021; Resnick et al. 2006). Others examine the interaction between consumer reviews and firms' advertising decisions (Chen and Xie 2005, 2008; Hollenbeck et al. 2019), as well as the relationship between firms' use of management responses and their online reputation (Chevalier et al. 2018; Proserpio and Zervas 2017; Wang and Chaudhry 2018). A common conclusion of these studies is that review platforms have not only changed how consumers make decisions, but also how firms behave in the marketplace. However, empirical evidence on whether sellers react to reviews by boosting quality is scarce, with the exceptions of Ananthakrishnan et al. (2019), who exclusively rely on online ratings as a proxy for hotel quality, and Farronato and Zervas (2019), who examine restaurants' compliance with hygiene standards looking at inspections records. In contrast, I am able to assess the impact of reviews on firms' subsequent behavior, quality upgrading and industry composition thanks to the richness of my data and the exogenous variation in information.

Second, I contribute to the literature on information and product quality. While there are strong theoretical reasons why information should matter (see Section 1.2), empirical evidence is scarce. Some studies measure the effect of market-specific quality disclosure programs and the introduction of feedback systems in online markets on consumers choices, firms financial performance and incentives to deliver better quality (Bai 2018; Dai and Luca 2020; Elfenbein et al. 2015; Ershov 2020; Hui et al. 2018; Jin and Leslie 2009; Klein et al. 2016). Particularly, Jin and Leslie (2003) find that consumers are sensitive to the information disclosed by restaurants' health inspections, which are shown to be an effective way to incentivize restaurants to be clean. However, their setting does not allow to study general equilibrium effects. In contrast, I can estimate the aggregate effects of increased information. Moreover, compared to all above-mentioned papers, I show that user-generated content on the consumer side provides sufficient incentives for firms to invest in quality upgrading.

Finally, and more broadly, this paper also relates to the literature looking at information frictions as one source of demand constraints impeding firm growth (Aker et al. 2020; Allen 2014; Anderson et al. 2018; Atkin et al. 2017; Bai 2021; Hjort et al. 2020; Jensen 2007; Jensen and Miller 2018; Startz 2018). The general lesson from this work is that improving information in the market enhances growth through a reallocation of market share toward the most productive firms. While this literature has primarily focused on product markets, my paper shows that similar results hold in the service industry.

The remainder of the paper is organized as follows. Section 1.2 presents the theoretical model and Section 1.3 discusses the study setting and data; Section 1.4 describes the empirical strategy and Section 1.5 reports the main results; Sections 3.4.6 and 1.7 show placebo exercises and robustness checks; finally, Section 1.8 discusses the magnitudes and Section 1.9 draws the conclusions.

1.2 Theoretical framework

1.2.1 Experience vs search goods in the digital era

Restaurant meals are a typical example of *experience goods* (Nelson, 1970): their quality can be truly assessed only by consuming them, not before. As a result, because of this type of asymmetric information, equilibrium quality levels tend to be lower than the optimal scenario with perfect information (Riordan, 1986). Repeat purchases, brand reputation, and the use of standards/certifications may offer consumers a way to learn about and exert control over quality. However, in many occasions - such as tourists visiting a city - consumption is transient and regulation only applies to minimum quality standards (e.g. hygiene inspections). Thus, consumers are left with two options to get at least partial information on quality prior to the purchase: the physical inspection of the restaurants (e.g., through publicly displayed hygiene cards and certificates of excellence) and the use of guidebooks or word-of-mouth.

The digital era has changed the way consumers get and share information. The decreasing costs of the internet and the diffusion of review platforms like Yelp and Tripadvisor in the last two decades have made information about quality and other characteristics (e.g., price, location and type of cuisine) of restaurants more readily available, helping consumers to make more informed choices (Ghose et al., 2013). In this respect, the possibility to gather some relevant product's information via online search makes restaurant meals closer to the essence of *search goods*.[12] Because this paper focuses on the effects of a reduction in costs to access information online, throughout the rest of it I consider restaurant meals a search good.

1.2.2 A model of consumer search and quality upgrading

This model extends the work of Goldmanis, Hortaçsu, Syverson and Emre (2010) to account for limited information about quality rather than prices. Such a framework allows to study how reductions in search costs affect the equilibrium quality in the market through both (1) their effect on the behavior of producers (what I refer to as moral hazard) and (2) their effect on the type of producers involved in the market (what I refer to as adverse selection). To do so, I borrow elements from two distinct theoretical literatures.

The first strand of the literature is the set of models on sequential consumer search and endogenous producer choices, such as price, product design and online categorization (Anderson and Renault 1999; Bar-Isaac et al. 2012; Fershtman et al. 2018). In particular, Fishman and Levy (2015) show that lower search costs can have both positive and negative impacts on firms' incentives to invest in quality, because of their differential effects on prices.[13] However, their framework – like most of existing models on consumer search and quality provision (e.g., Moraga-González and Sun 2020; Wolinsky 2005) – does not feature endogenous firm entry and, in turn, is not suitable to assess the effects on industry composition.

To overcome this limitation, I draw from a second strand of the literature, namely, the set of industry equilibrium models with heterogeneous producers and endogenous selection into production (e.g., Hopenhayn 1992; Melitz 2003; Syverson 2004). Endogenizing the set of equilibrium producers allows to study the relationship between producer type and product quality, to eventually assess how a reduction in search costs affects the industry composition (i.e., entry/exit choices by type of producer). This represents an important contribution to the current literature on consumer search and product quality,

[12] According to Klein (1998), it is the relatively higher cost of search with respect to direct purchase that makes a good an experience good. Thus, when consumers can obtain important product information via new interactive media at decreasing costs prior to the purchase, the product can be considered a search good.

[13] While they increase the market shares of high-quality firms, lower search costs also reduce their prices and profits more than those of low-quality firms, hence the effect on incentives to invest in quality is ambiguous.

and allows to investigate the effects of search costs on quality provision through moral hazard and adverse selection, separately.

1.2.2.1 Model setup

There is a continuum of firms, each of which sells one quality $q \in \mathbb{R}_+$ of a vertically differentiated good at a common and exogenous price p to a continuum of consumers whose total mass is fixed and normalized to one.[14] All consumers have perfectly inelastic unit demand but are heterogeneous in their search costs $s \in \mathbb{R}_+$, with $s \sim Z$ (and density z). A consumer that buys one unit of quality q at price p gets utility (net of any search costs) $u = q - 1$, Without loss of generality, I normalize the price p to one, so that utility becomes $u = q - 1$. There is no outside good in the market. Firms are also heterogeneous, differing in their underlying abilities (types), which affect their cost of producing a good of a certain quality. The total mass of firms L is endogenously determined through a zero-profit condition.

The timing is the following. At the beginning of the period, potential firms consider entering the market. If a firm decides to enter, it pays the sunk cost of entry $\kappa \in \mathbb{R}_+$ and learns its own ability parameter $\lambda \in \mathbb{R}_+$, which is drawn i.i.d. from a publicly known probability distribution with cdf Γ and pdf γ. Next, firms decide whether to stay in the market or not. Those firms that choose to stay then decide the quality level of their good and produce. Production requires a fixed cost of operation $C(q, \lambda)$, which depends positively on the chosen quality level and negatively on the exogenous ability parameter of the firm (i.e., $C'_q > 0$ and $C'_\lambda < 0$). This cost can be avoided if the firm chooses to stay out of the market.

1.2.2.2 Consumers' problem

Consumers have full information on the price of the goods being sold. However, they only know the quality distribution, F (with density f) and must engage in costly search to learn the quality provided by any particular firm. This is in line with the idea that

[14] Assuming exogenous prices has two advantages: (1) it keeps the algebra tractable and allows me to solve the model analytically; (2) it excludes the possibility that prices are used by firms to signal quality (as in Wolinsky 1983) and therefore to reduce the asymmetric information problem. At the same time, however, the assumption may sound implausible. To make the model more realistic, one could think of a segmented market (e.g., fast-food vs. starred restaurants) where firms compete in quality and charge the same price within each segment (but not across segments), and consumers search exclusively within a segment. This scenario would not qualitatively change the results of the model.

information on the price of a meal might be gathered before the purchase, for instance, by reading the restaurants' menu on the window. Consumers' search is undirected and sequential: they visit stores one-by-one to learn their quality and after every visit compare the expected benefit and cost of continued search. If the expected quality gain from visiting another store is lower than the marginal cost of search s, the consumer continues to search; otherwise, they buy the product with the highest quality in hand. Following McCall (1970), in this context, the optimal stopping rule is characterized by a reservation quality level $\rho(s)$ (i.e., the minimum acceptable quality of a good) such that a consumer stops searching and buys only if they find a product with quality $q \geq \rho(s)$. Particularly, $\rho(s)$ is implicitly defined by

$$h(\rho, q) \equiv \int_{\rho(s)}^{\infty} [q - \rho(s)]\, f(q)\, dq - s = 0, \tag{1.1}$$

where the integral is the expected quality gain from another search, accounting for the option value of discarding lower quality draws. Using integration by parts, one can rewrite (1.1) as

$$h(\rho, q) \equiv \int_{\rho(s)}^{\infty} [1 - F(q)]\, dq - s = 0. \tag{1.2}$$

Applying the implicit function theorem to (1.2) yields $\rho'(s) = -1/[1 - F[\rho(s)]]$, that is, the reservation quality is strictly decreasing in the search cost (i.e., consumers with lower s are pickier). This also implies that $\rho(s)$ is invertible and its inverse is given by $\rho^{-1}(r) = \int_r^{\infty} 1 - F(q)\, dq$.

1.2.2.3 Producers' problem

Firms do not know the ability parameters and the qualities produced by their rivals in the market, but they do know their distributions (Γ and F). Moreover, firms only know the distribution of search costs Z, and not the search cost of any individual consumer. Each firm takes as given these distributions, and determines its optimal quality based on the demand it faces, characterized by the reservation quality rule $\rho(s)$ implied by (1.1).

I analyze the optimization problem of a firm with ability parameter λ that chooses to stay in the market. To determine the quantity as a function of the quality chosen by the firm, $x(q)$, one should start from the optimal search rule. Only consumers with reservation qualities $\rho(s)$ below q will buy from the firm. Consider a consumer with reservation quality $r < q$. Since the quality distribution in the market is F and the total mass of operating firms is L, the mass of firms producing quality q above r is $L[1 - F(r)]$. The consumer is equally likely to buy from any one of these firms. That is, the probability that they will buy from a particular firm producing quality q is $1/L[1 - F(r)]$. Integrating over all consumers with a reservation quality lower than q yields the following formula for quantity:

$$x(q) = \int_0^q \frac{g(r)}{L[1 - F(r)]}\, dr, \tag{1.3}$$

where $g(r)$ is the pdf of the reservation quality. This formula can be expressed in terms of the search cost and quality distributions, so to obtain the following standard residual demand curve (algebra in Appendix 1.A):

$$x(q) = \frac{1}{L} \int_0^q z[\rho^{-1}(r)]\, dr. \tag{1.4}$$

Equation (1.4) states that the demand faced by a firm is determined by its own quality as well as its competitors' qualities. Note that demand is increasing in quality, since $x'(q) = \frac{1}{L} z[\rho^{-1}(q)] > 0$. However, quality is costly. Higher-quality output requires higher-quality inputs, which come at a cost (e.g., searching for better suppliers and hiring workers with better curricula/experience). I assume that these costs do not depend on the quantity produced, yet they depend on the innate ability of the firm, which is governed by the parameter λ.[15] Hence, the cost function of a firm with ability λ is $C(q, \lambda)$, with $C'_q > 0$, $C''_{qq} > 0$, $C'_\lambda < 0$ and $C''_{q\lambda} < 0$. The last conditions imply that more capable firms (higher λ) are more efficient, so that their fixed cost to produce a given quality is lower or, alternatively, they produce a higher quality product spending the same cost.[16] Hence, the optimization problem of a firm with ability λ choosing to stay in the market is

$$max_q \quad \pi[q(\lambda), \lambda] = x[q(\lambda)] - C[q(\lambda), \lambda]. \tag{1.5}$$

The equilibrium quality function $q(\cdot)$ follows from the first-order condition for an

[15] Assuming quantity-dependent costs does not qualitatively alter the results, yet it makes the algebra more cumbersome.

[16] The positive relationship between managerial skills (human capital) and firm productivity has been well established in the literature (e.g., Gennaioli et al. 2013). In my context, high-skill owners/managers/chefs make a more efficient use of their inputs, and therefore save in costs. For instance, managerial skills help owners self-train their room staff, or cooking abilities allow firms to avoid expensive products and nevertheless make fabulous dishes.

optimum

$$x'_q[q(\lambda)] - C'_q[q(\lambda), \lambda] = 0, \tag{1.6}$$

with the second-order condition for a maximum requiring that

$$x''_{qq}[q(\lambda)] - C''_{qq}[q(\lambda), \lambda] < 0. \tag{1.7}$$

1.2.2.4 Equilibrium

Let $q(\cdot)$ and $x(\cdot)$ be the quality and residual demand function in equilibrium, respectively. Then, the following properties follow (proofs are in Appendix 1.A):

Property 1: The equilibrium quality function is increasing in the ability parameter: $q'_\lambda(\lambda) > 0$, $\forall\lambda$.

Property 2: The demand function is increasing in the ability parameter: $x'_\lambda[q(\lambda)] > 0$, $\forall\lambda$.

Property 3: The profit function is increasing in the ability parameter: $\pi'_\lambda[q(\lambda), \lambda] > 0$, $\forall\lambda$.

From Property 3 it follows that the decision rule for staying in the market or leaving is characterized by a cut-off value $\underline{\lambda}$ such that firms stay in the market if and only if $\lambda \geq \underline{\lambda}$, with $\underline{\lambda}$ satisfying

$$\pi(\underline{\lambda}) = x[q(\underline{\lambda})] - C[q(\underline{\lambda}), \underline{\lambda})] = 0. \tag{1.8}$$

In the initial stage, potential entrants have to decide whether or not to start producing. Assuming unlimited entry into the market, firms keep entering until the expected value of post-entry profits equals the sunk entry cost. That is, the entry condition requires that

$$\kappa = \int_{\underline{\lambda}}^{\infty} \pi(\lambda)\, \gamma(\lambda)\, d\lambda = \int_{\underline{\lambda}}^{\infty} \Big[x[q(\lambda)] - C[q(\lambda), \lambda]\Big] \gamma(\lambda)\, d\lambda. \tag{1.9}$$

Finally, it is possible to express the distribution of qualities F in terms of the distribution of abilities Γ. Property 1 implies that qualities will be distributed with support $[\underline{q}, \bar{q}]$, where $\underline{q} = q(\underline{\lambda})$ and $\bar{q} = q(\infty)$. Thus, for $v \in [\underline{q}, \bar{q}]$, the cdf will be given by

$$F(v) = Pr\{q(\lambda) \leq v \,|\, \pi(\lambda) \geq 0\} = \frac{Pr\{\lambda \leq q^{-1}(v) \;\&\; \lambda \geq \underline{\lambda}\}}{Pr\{\lambda \geq \underline{\lambda}\}} = \frac{\Gamma[q^{-1}(v)] - \Gamma(\underline{\lambda})}{[1 - \Gamma(\underline{\lambda})]}. \tag{1.10}$$

Note that $F(v) = 0$ for $v \leq \underline{q}$ and $F(v) = 1$ for $v \geq \bar{q}$. I can now define the equilibrium in this market.

Definition 1: A search equilibrium is a set $\{\rho : \mathbb{R}_+ \rightarrow \mathbb{R}_+,\ q : \mathbb{R}_+ \rightarrow \mathbb{R}_+,\ x : \mathbb{R}_+ \rightarrow \mathbb{R}_+,\ F : \mathbb{R}_+ \rightarrow [0,1],\ \underline{\lambda} > 0\}$ satisfying (1.2), (1.4), (1.6), (1.7), (1.8), (1.9) and (1.10).

1.2.2.5 Comparative statics

The equilibrium functions and values defined above depend on the search costs that consumers face. My goal is to determine how a decrease in these costs will affect the equilibrium quality schedule $q(\cdot)$ and cost function $C(\cdot)$ of operating firms, their demand as well as the operating cut-off level of ability $\underline{\lambda}$. For this purpose, I impose further assumptions that make the model more aligned with the empirical exercise explained in Section 1.3.

Assumption 1: *The search cost distribution is uniform on $[0, a]$ for $a > 0$.*

This assumption allows to study changes in search costs that are heterogeneous across consumers. In particular, I will focus on a cost reduction that only affect consumers with ex ante the highest costs (a).[17]

Assumption 2: *The firms' cost function takes the form*

$$C(q, \lambda) = \frac{q}{1-q}\frac{1}{\lambda}, \tag{1.11}$$

which satisfies the requirements described in section 1.2.2.3 for $q \in (0,1)$ and $\lambda > 0$.

This assumption has two direct implications. First, the quality level chosen by any firm is bounded between $(0, 1)$, and so is the reservation quality $\rho(s)$. This is consistent with the way individuals value and rate the quality of a meal on review platforms, which exhibit a finite scale (e.g., 1 to 5). In this respect, the upper bound equal to 1 becomes the natural limit of quality via the reputation mechanism, so that firms have no incentive to deliver a level of quality beyond that value. The second implication of equation (1.11) is that costs are increasingly steeper in quality, and tend to infinity as $q \rightarrow 1$, for a given λ. This assumption reflects the idea that superb quality requires the owner to completely rethink and change the business model of the restaurant, which is unfeasible in the short-term.[18]

[17] In the empirical framework these consumers are identified with foreign tourists, which bear the highest costs to browse the internet in the presence of roaming tariffs.

[18] Results are qualitatively unchanged when alternative convex cost functions are used (e.g., $C = 1/q\,\lambda$ or $C = q^2/\lambda$), provided that quality q is bounded between 0 and 1.

From Assumption 1, it follows that for $q \in (\underline{q}, 1)$ the demand function (1.4) becomes

$$x(q) = \frac{1}{L}\int_0^q \frac{1}{a}\mathbb{1}_{\{\rho^{-1}(r)\in[0,a]\}}\,dr = \frac{1}{aL}\int_0^q \mathbb{1}_{\{r\in[\rho(a),\rho(0)]\}}\,dr = \frac{1}{aL}\left[q - \rho(a)\right]. \tag{1.12}$$

Note that $x'(q) = 1/aL > 0$ and $x''(q) = 0$ so that, together with Assumption 2, the second-order condition (1.7) holds. Substituting (1.11) and (1.12) into (1.6), the first-order condition simplifies to

$$q(\lambda; a) = 1 - \sqrt{\frac{aL}{\lambda}}, \tag{1.13}$$

and the additional condition in order for the ability parameter to yield admissible quality levels follows:

$$q(\lambda) \in (0,1) \iff \lambda > aL. \tag{1.14}$$

In other words, firms need to have at least some ability in order to produce positive qualities.

Consistently with Property 1, the equilibrium quality schedule (1.13) is increasing in λ, that is $q'(\lambda) = \sqrt{\frac{aL}{\lambda}}\frac{1}{2\lambda} > 0$. Moreover, the function is concave, that is $q''(\lambda) = -\frac{3}{4\lambda^2}\sqrt{\frac{aL}{\lambda}} < 0$. This is a direct consequence of the functional form of the firm's production cost (1.11). As $q \to 1$, more capable firms will use their ability-advantage mostly to save in costs rather than to deliver a higher quality product. Hence, the quality decision becomes less sensitive to the ability parameter as λ gets larger.

From the equilibrium quality schedule, it follows that the demand, cost and profit functions reduce to

$$x(\lambda; a) = \frac{1}{a\,L}\left[1 - \sqrt{\frac{aL}{\lambda}} - \rho(a)\right], \tag{1.15}$$

$$C(\lambda; a) = \frac{1}{\lambda}\left[\sqrt{\frac{\lambda}{aL}} - 1\right] \text{ and} \tag{1.16}$$

$$\pi(\lambda; a) = \frac{1}{a\,L}\left[1 - \sqrt{\frac{aL}{\lambda}} - \rho(a)\right] - \frac{1}{\lambda}\left[\sqrt{\frac{\lambda}{aL}} - 1\right], \tag{1.17}$$

and the operating threshold value for ability, $\underline{\lambda}$, follows from (1.8). That is,

$$\underline{\lambda} := \pi(\underline{\lambda}) = 0 \iff \underline{\lambda}(a) = \frac{aL}{\left(1 - \sqrt{\rho(a)}\right)^2}, \tag{1.18}$$

such that only for firms with $\lambda \geq \underline{\lambda}$ it is convenient to stay in the market and produce. Note that $\underline{\lambda}$ satisfies condition (1.14). Finally, the lower and upper limits of the support of the equilibrium quality distribution are $\underline{q} = q(\underline{\lambda}) = \sqrt{\rho(a)}$ and $\bar{q} = q(\infty) = 1$.

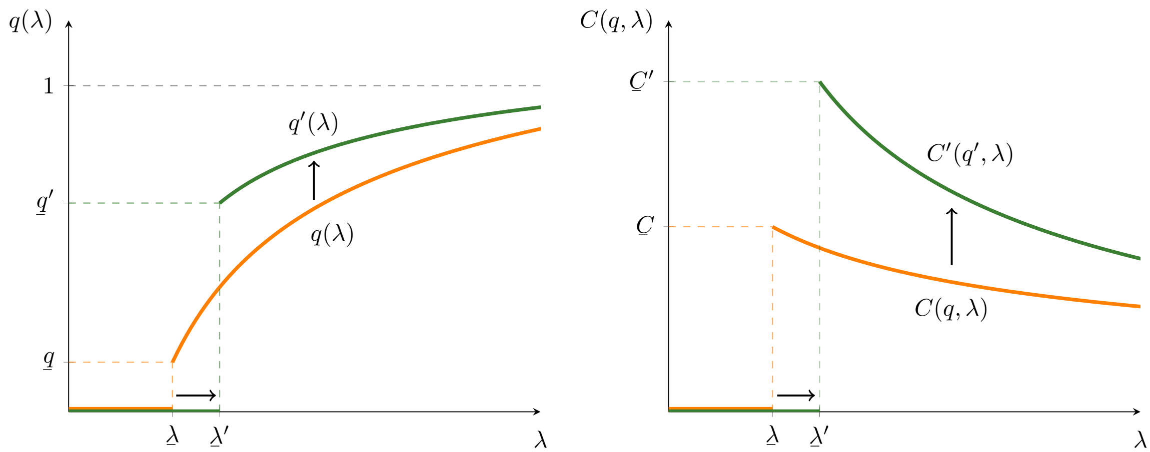
To conclude the comparative statics exercise, it remains to demonstrate how a decrease in search costs for consumers with *ex-ante* the highest costs – i.e., a reduction in a – affects the above quantities, and how these changes depend on the ability parameter of the firm. For this purpose, it is convenient to formalize two preliminary observations that will be used to derive the subsequent results (all proofs are in Appendix 1.A). First, I define the quantity $\delta(a) \equiv a\,L(a)$, where I emphasize the dependence of L on a.
Lemma 1: $\delta(a)$ is increasing in search costs, that is $\delta'_a(a) > 0$.

Note that $\delta(a)$ can be interpreted as the inverse of the per-firm density of consumers with a given level of search costs. Hence, Lemma 1 states that such a density is decreasing in the search costs. The second observation is about the profit function of a firm with ability λ, described in equation (1.17). It is possible to show that, if an increase in search costs reduces the profits of any currently operating firm, it must also reduce those of all firms with higher abilities. Formally,
Lemma 2: If there exists $\lambda_0 \geq \underline{\lambda}(a)$ such that $\pi'_a(\lambda_0; a) \leq 0$, then $\pi'_a(\lambda; a) \leq 0\ \forall \lambda > \lambda_0$.

I can now state the following key results:
Proposition 1: When search costs decrease, the quality $q(\cdot)$ produced by a firm with ability λ increases $\forall \lambda \geq \underline{\lambda}$, and more so for firms with lower ability. That is, $q'_a(\cdot) < 0$ and $q''_{a\lambda}(\cdot) > 0$.
Proposition 2: When search costs decrease, the production costs $C(\cdot)$ of a firm with ability λ increase $\forall \lambda \geq \underline{\lambda}$, and more so for firms with lower ability. That is, $C'_a(\cdot) < 0$ and $C''_{a\lambda}(\cdot) > 0$.
Proposition 3: When search costs decrease, the cut-off ability value $\underline{\lambda}(\cdot)$ increases. That is, $\underline{\lambda}'_a(\cdot) < 0$.
Corollary 1: A decrease in search costs causes the demand $x(\lambda; a)$ faced by all firms with sufficiently high ability to increase: for each a, there exists $\hat{\lambda}(a) \geq \underline{\lambda}(a)$ such that

$x'_a(\lambda; a) < 0 \ \forall \lambda > \hat{\lambda}(a)$.

Figure 1.1: The effect of lower search costs on quality, ability threshold and production costs



Notes: The orange (green) lines report the equilibrium schedules when search costs are higher (lower), that is $a : a \to a' < a$.

Figure 1.1 graphically describes the consequences of a decrease in the parameter $a : a \to a' < a$ on the ability threshold and equilibrium qualities (left panel) as well as on the cost schedule (right panel). The first consequence of a decrease in consumer search costs is that some firms with the lowest abilities exit the industry, as $\underline{\lambda}$ shifts to the right. All other firms that remain in the market (i.e. $\lambda > \underline{\lambda}'$) upgrade their quality when a decreases, and more so those firms that were initially producing lower-quality products (i.e., firms with lower abilities). As a result, production costs change. In fact, all operating firms bear higher costs, indicating that the process of quality upgrading is overall costly. Particularly, firms that were initially producing lower qualities exhibit the largest cost increase.

These results show that two simultaneous mechanisms make the industry more quality oriented when search costs fall. First, the upward shift in the equilibrium quality schedule for all operating firms is a consequence of a reduction in moral hazard: lower search costs make consumers more demanding and their choices more sensitive to the characteristics of the products, hence firms' incentives to upgrade quality increase. Second, the rightward shift in the cut-off ability level is a consequence of a reduction in adverse selection: when search costs fall, demand to lower capable firms decrease, so do their profits, pushing some of them out of the market.

The results of Propositions 1, 2 and 3 together with Corollary 1 constitute the main theoretical predictions of the paper. Declines in search costs in the restaurant industry driven by the advent and diffusion of online review platforms have differing effects across businesses. Low-type sellers are hurt, sometimes to the point of being forced to exit. Higher types, however, gain from the shift as their demand grows. Incentives to upgrade quality arise, resulting in higher quality levels, especially for surviving low-type firms.

1.2.3 Hypotheses

In order to test the model predictions, it is necessary to identify the empirical counterparts of the quantities described in the theoretical framework. Objective measures of quality are difficult to obtain, especially for experience goods and, particularly, for the restaurant industry, where the quality of a meal reflects multiple dimensions (e.g., service and food) whose evaluation is to a large extent subjective. Following Ananthakrishnan et al. (2019), I use Tripadvisor ratings as a proxy for that dimension of quality that is mainly subjective and can be referred to as reputation. Importantly, this is the dimension that is revealed to consumers once they pay the search cost and visit the Tripadvisor profile of the restaurant. Hence, owners and managers care about such a measure and seek to maximize it.

Moreover, since quality decisions are likely to affect firms' production costs through the labor market (e.g., Shin et al. 2021), I rely on hiring choices to capture objective quality upgrading. Particularly, I consider the curriculum of newly-hired employees and their wages to explicitly measure investment into service quality of restaurants. Finally, proxies for output are obtained considering annual revenues as well as the total number of employees, while information on the firm's presence (or not) in the market is retrieved from the official date of opening/closure of the business.

One potential caveat is that the empirical counterpart of the ability parameter λ remains unobserved. Nevertheless, equations (1.10) and (1.13) state that, for operating firms, there exists a one to one mapping between ability and quality. Hence, I can rely on the Tripadvisor rating of the restaurant at baseline (i.e., before the reduction in internet tariffs) to proxy for the underlying ability parameter. Therefore, the above theoretical predictions can be translated in the following testable empirical hypotheses. When search costs for consumers fall (as a consequence of the access to review platforms):

1. The demand faced by firms with *ex-ante* sufficiently high Tripadvisor rating increases: their revenues and number of employees grow.

2. The overall quality level in the industry improves:

   (a) Some of the firms with *ex-ante* the lowest ratings will exit the market: a reduction in adverse selection;

   (b) Surviving firms will invest in quality upgrading (e.g., hiring workers with better curricula) eventually improving their online ratings. These effects will be larger for firms with *ex-ante* lower ratings: a reduction in moral hazard.

1.3 Study setting and data

1.3.1 The EU roaming regulation

Following recent empirical literature on consumer search (e.g., Ershov 2020), I take advantage of online platforms to characterize reductions in search costs. In particular, I rely on an exogenous reduction in the costs of mobile internet caused by the abolition of roaming tariffs in the European Union, which promoted the use of review platforms like Tripadvisor.

International mobile roaming regulations apply when customers use their mobile phones while occasionally travelling outside the country where they live (specifically, outside the geographical coverage area of the home operator's network). This paper exploits the approval by the European Parliament of a new policy on the EU roaming regulation – known as "Roam like at home"[19] –, which led to the abolition of all charges for temporary roaming within the European Economic Area (EEA) as of June 15, 2017. In practice, if before that day all EU residents traveling within the EEA were charged at least € 0.05 per MB of data (on top of the home network rate), after the policy the same home network rate is applied with no additional charges.[20]

The EU roaming regulation consists of a series of policy packages that started in 2007

[19] `https://europa.eu/youreurope/citizens/consumers/internet-telecoms/mobile-roaming-costs`

[20] Note that the policy also affected the prices of SMSs/phone calls, but only toward the country of origin. Calling/texting local restaurants while abroad remained equally expensive, hence this type of communication is unlikely to play any role in the observed results.

and regulate wholesale and retail international roaming tariffs. The policy was initially motivated by the 2006 European Commission impact assessment, which pointed out a large gap between the roaming prices charged to consumers and the actual cost of providing the wholesale service. Therefore, the underlying objectives of the reform were the intensification of the competition among providers and the promotion of market integration (digital single market). The last decisive step took place on June 15, 2017, when wholesale and retail price caps for data were set to 0. Grzybowski and Muñoz (2020) show that the European Commission has succeeded to avoid unintended increases in domestic tariffs and induced operators to absorb the negative effects of the reform. At the same time, Quinn et al. (2021) show that, after the policy, daily mobile data consumption (sum of uploads and downloads) for EU travelers while abroad grew by at least 54MB.

For the purpose of this paper, the reform induced an exogenous shock to the costs of accessing online information for EU travelers while abroad. In particular, information contained in review platforms such as Tripadvisor became available to all EU travelers at virtually zero cost. Hence, the search costs for certain tourists looking for restaurants while visiting a city drastically decreased compared to the pre-policy period.

1.3.2 Data sources

To study the consequences of lower information costs on the restaurant industry, I focus on the whole Province of Rome. Looking at a large geographical area allows me to exploit spatial variation in the intensity of exposure to tourist clientele, an attractive feature for empirical identification. Specifically, I assemble a novel dataset on restaurants in the Province combining three data sources.

The first source is Tripadvisor, the most popular review platform in Italy and Europe. Listing an establishment on the platform is free and can be done either by the clients or by the owners/managers of the firm. I collect information on listed restaurants (e.g., name, address, price category and type of cuisine) as well as, for each restaurant, the entire historical record of reviews (date, rating, device, text, country and language of reviewer, etc.) from 2007 to 2019, for a total of approximately 3 million reviews. Since the format of the data is unstructured, I combine them together to create a panel at the restaurant-month level. Importantly, I used the historical record of reviews to retrieve the average rating of the restaurants in any month between 2007-2019. The Tripadvisor sample contains information on 14,146 establishments with at least one review as of December 2019. Of them, 11,595 had at least one review in May 2017, i.e. the month before the roaming policy was effective.[21] Moreover, from Tripadvisor I also gather information on location and attributes of the top-100 tourist attractions in the Province, according to their total volume of reviews.[22] These will be used in the empirical strategy described in Section 1.4.

The second data source is provided by administrative social security records collected and maintained under restricted-use access at the Italian National Social Security Institute (INPS). For each establishment in the Province of Rome, the records contain information for the last 15 years on location (ZIP code), date of opening and closure, legal status of the firm, monthly number of employees, type of contracts and qualification of the workers, their wage bill and demographics, as well as their full employment history. According to this dataset, 10,391 restaurants operated in the Province and had an active profile at the Social Security Institute in at least one month between 2015 and 2019.[23]

The third data source contains proprietary annual information on income statements and balance sheets originally collected and maintained at the Italian Business Registry (Chamber of Commerce) and accessed through the Cerved database.[24] This dataset provides information on revenues, costs, profits and other financial indicators, and it covers most of those firms with an LLC proprietorship status.[25] In particular, almost 5,000 restaurants in the Province were obliged to report their financial information to the Registry at any point in time between 2015 and 2018, which is the last available year.

[21] It is worth mentioning that having a Tripadvisor profile with a positive number of reviews at a particular point in time does not necessarily imply that a restaurant is an active business at that time. For instance, it might be the case that the restaurant has closed, but the Tripadvisor profile still exists.

[22] I consider as tourist attractions those activities that on Tripadvisor belong to the categories "Sites of interests" and "Monuments". With almost 128,000 as of 2019, the Colosseum is the most-reviewed attraction, while the National Roman Museum ranks 100th on the list, with almost 600 reviews.

[23] I use information on the primary activity of the firm (ATECO code) to identify restaurants. Particularly, I restrict the attention to the following ATECO codes: 56.10.11 (dine-in restaurants), 56.10.12 (agriturismi), 56.10.20 (take-away restaurants), 56.10.30 (bakeries).

[24] Particularly, I access the version of Cerved data that is available at the Social Security Institute in Rome, where the last available year is 2018.

[25] In the restaurant industry in the Province of Rome, LLC companies represent about 57% of the total. These businesses are owned by shareholders, who have limited personal liability for business related debts and are required by the law to report financial statement information at the Chamber of Commerce on an annual basis. By contrast, firms with no financial data are usually unlimited liability partnerships and sole proprietorship businesses, which are generally smaller and more likely to be family-owned restaurants.

1.3.3 Dataset construction

To conduct the empirical analysis, I matched the Tripadvisor sample with social security and financial records. Combining crowd-sourced data with administrative archives is challenging because of the very different nature and confidentiality protocols of the two sources. Specifically, in performing such a matching I faced two main obstacles, namely (1) the anonymity of the administrative records and (2) the lack of official business identifiers in the Tripadvisor data. Regarding the former, access to both INPS and Cerved databases was granted under a specific program (VisitINPS).[26] The program requires researchers to conduct the empirical analysis at the data center in Rome, where they obtain *de-identified* information on employers and employees. Particularly, names, addresses and unique business identifiers of the firms – i.e., the VAT codes – remain unknown to the researchers for confidentiality purposes. On the other hand, Tripadvisor records do contain names and addresses of the restaurants, but not their unique business identifiers. Hence, the information was incomplete on both sides.

To overcome the limitation, I purchased additional data from the Italian Business Registry containing names, addresses and VAT codes for all the restaurants in the Province. I then used the name and address (previously processed by Google API for text harmonization) to assign a VAT code to as many restaurants as possible in the Tripadvisor sample. For restaurants that could not be matched using the name and address, or for which the matching precision was low, I tried to manually collect the VAT codes from their websites or from the pictures of receipts posted by the clients on Tripadvisor/Google. This two-step matching procedure resulted in about 6,000 Tripadvisor restaurants with an associated VAT code. Of them, 80% were matched using the name and address of the firm, while the remaining 20% were matched manually.

Because of their high resolution, the Tripadvisor data could not be imported – as they were – in the servers of the data center. In fact, importing external records is subject to strict rules for confidentiality concerns, to avoid that researchers could identify individual firms from the data. To comply with this requirement, I selected only the most important Tripadvisor variables and simplified the majority of them through categorization (e.g., grouping continuous variables in a limited number of categories).

[26]In compliance with the program requirements, most of the empirical analysis presented in the paper has been carried out at the data center in Rome, and no data has left the center except for the output tables and figures reported in the paper. Official information on the program is available here `https://www.inps.it/dati-ricerche-e-bilanci/attivita-di-ricerca/programma-visitinps-scholars`.

Eventually, this simplified version of the Tripadvisor dataset could be imported in the INSP servers and matched with the social security archives through the previously associated VAT codes.[27] In particular, the final matched Tripadvisor-Social Security sample comprises 5,472 firms that operated in the industry at any point in time between 2015 and 2019.[28] This sample represents almost 53% of the total number of active businesses in that period and it is employed in the market-level analysis to assess how entry/exit dynamics shape the composition of the restaurant industry. Moreover, among the matched restaurants, 4,628 had a Tripadvisor profile with at least one review (and, therefore, a rating) in May 2017. I employ this sample in the firm-level analysis.

1.3.4 Summary statistics

Table 1.1: Summary of firm-level outcomes

		Period: Jan 2015 - Dec 2019						Pre-policy
	Firms	Obs	Mean	SD	Min	Median	Max	Mean
N. of monthly employees	4628	219835	5.69	5.50	0.0	4.0	29.0	5.55
Annual revenues (Thousand, €)	2043	6677	692.18	1065.00	5.0	394.0	8752.0	646.60
Monthly working days	4628	219835	92.13	102.04	0.0	58.2	1922.9	90.96
Working days per worker	4517	197194	15.66	5.84	0.1	15.6	185.9	15.86
1 if firm exits (×100)	4628	219835	0.41	6.36	0.0	0.0	100.0	0.33
1 if firm hires worker w/ previous experience in restaurants	4628	219835	0.08	0.28	0.0	0.0	1.0	0.08
1 if firm hires worker w/o previous experience in restaurants	4628	219835	0.06	0.23	0.0	0.0	1.0	0.06
Months of experience in restaurants of newly-hired employees	3550	30133	14.46	22.96	0.0	3.8	157.0	13.01
Months of experience in restaurants of quitting/fired employees	3584	30911	27.21	29.83	0.0	17.0	152.0	25.12
Average daily salary (€)	4558	200402	66.60	19.12	24.6	61.1	156.8	64.88
Average 5-month Tripadvisor rating	4373	147274	3.96	0.65	1.0	4.0	5.0	3.98
N. of 5-month replies to reviews	4377	146713	2.46	11.47	0.0	0.0	313.0	2.56
N. of monthly Tripadvisor reviews	4572	178425	5.70	12.51	0.0	3.0	1110.0	6.18

Each observation is a restaurant-month-year, with the exception of revenues, which are observed at the restaurant-year level up to 2018. Data on Tripadvisor reviews, rating and replies refer to the period between Jan 2015 and Dec 2018. Daily salary is adjusted for part-time workers so to reflect the full-time equivalent salary.

[27] The main data import process took place in early 2019. For this reason, imported data on reviews, ratings and replies from Tripadvisor cover until December 2018.

[28] To minimize the risk of measurement error due to misreporting in the social security data, before conducting the analysis I trimmed observations with a number of employees above the 98th percentile. The final sample does not include these observations.

Table 1.1 shows descriptive statistics for the 4,628 matched restaurants with available Tripadvisor rating at the time of the roaming policy. For the main outcome variables, the table reports a series of statistics referring to the Jan 2015 - Dec 2019 period, as well as, for the sake of comparison, their mean in the pre-policy period (Jan 2015 - May 2017). The figures indicate that the average restaurant in the sample is a small business, with less than 6 monthly employees and an annual revenue just below 700 Thousand Euros. Its employees work, on average, almost 16 days per month, and their adjusted full-time-equivalent gross salary is about €67 per day. During the period of interest, 8% of the times the average establishment hires a worker with previous experience in the restaurant industry and, when it does, the new employee has worked about 14.5 months in the sector. The 5-month rolling average rating that the typical restaurant obtains on Tripadvisor is almost 4, and the number of 5-month total replies to online reviews is 2.5.

To address concerns on the potential bias in the analysis introduced by the matching procedure, Appendix Table 1.D1 compares the main descriptive statistics of the matched sample with those of the entire Tripadvisor and Social Security datasets, separately. While matched restaurants tend to be slightly larger in size, as well as closer and more exposed to tourist attractions than the average restaurant in the Province, they also appear to have similar Tripadvisor ratings and price categories, recruit equally-experienced workers, and pay comparable salaries. Overall, this evidence seems to discard the possibility that results could be systematically driven by sample selection.

1.3.5 The roaming policy and the use of Tripadvisor

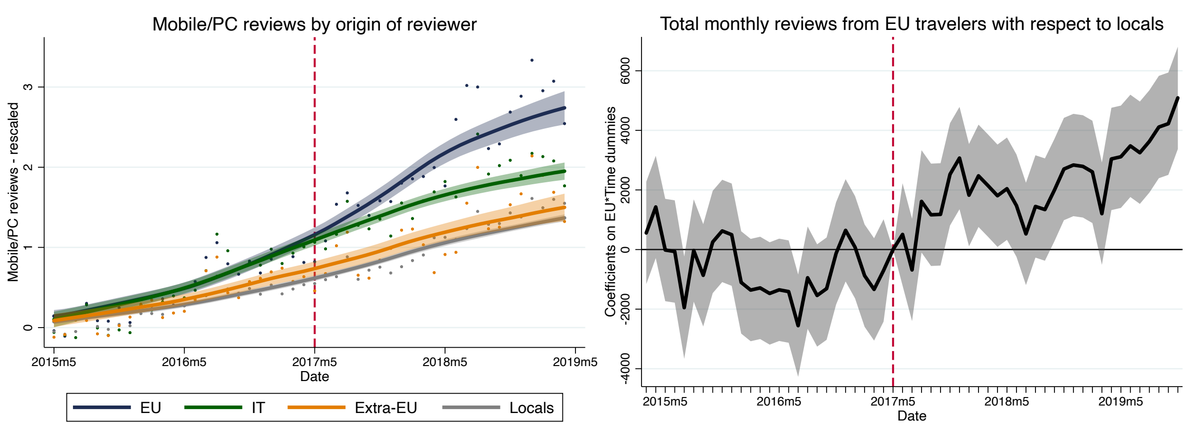
As the costs of mobile internet falls, its use is expected to increase. Consumers with free internet access have the possibility to search and verify products online before purchasing. Tripadvisor data allows to study the reviewers' behavior across types of device and nationality. Although Tripadvisor contributions reflect the supply of reviews and not necessarily their consumption, in the absence of better data on the demand side, I employ them here as a proxy for overall usage of the platform.[29]

Panel (a) of Figure 1.2 shows the change in reviewers' behavior over time across device and origin of the reviewer, which is proxied by the language of the review. The picture

[29]Reviewers are likely to be a subset of the total users (visitors) of Tripadvisor, as posting reviews entails an additional effort that not all users are willing to bear. Hence, finding evidence on changes in reviewers' posting behavior likely implies that similar changes hold, more generally, for the broader set of users.

points out a shift from PC- to mobile-based contributions following the implementation of the roaming policy (the red dashed line). Importantly, this effect is remarkably visible only for EU tourists but not for Italians, extra-EU travelers and locals, who were not deliberately targeted by the new regulation. Regression analysis confirms the visual results. Appendix Table 1.D2 directly compares the posting behavior of EU, extra-EU and Italian travelers with that of the locals. After the policy, users from EU countries became 1.4 times more likely to post reviews on their mobile devices as opposed to PC, while no significant change occurred for extra-EU and Italian users. The table also shows that the absolute number of monthly reviews posted on mobile devices by EU travelers increased by approximately 500 reviews per month after the policy, compared to the locals.

Figure 1.2: The roaming policy and the use of Tripadvisor



(a) (b)

Notes to Panel (a): Data on 14,146 restaurants with at least one review as of December 2019. Dots represent the monthly ratios, lines depict local polynomial fits with 95% confidence intervals. Values are re-scaled so that they are equal to 0 at the beginning of the period.
Notes to Panel (b): The graph reports estimated coefficients on the interactions between EU-dummy and time dummies, from a regression where each observation is a region of origin-month-year. The gray area reports 95% confidence intervals.

These patterns are not necessarily and exclusively driven by PC-to-mobile substitution. Panel (b) of Figure 1.2 shows event-study estimates for the total (mobile+PC) number of reviews posted by EU users compared to the locals. While before the policy total contributions from both categories displayed similar trends, the volume of monthly reviews from EU travelers significantly and steadily increased after May 2017. Importantly, as Appendix Table 1.D3 and Figure 1.C1 show, these results are unlikely to be driven by a discontinuous increase in international tourist flows toward Italy around the time of the policy. Overall, this evidence suggests that the new regulation provided an abrupt and arguably exogenous source of variation in the use of Tripadvisor services by EU travelers.

1.4 Empirical strategy

To estimate the demand and supply effects of increased access to information from online reviews, I combine the temporal variation induced by the policy with the spatial variation in restaurants' exposure to tourist clientele. The basic idea is that restaurants that more frequently cater to tourists are also more likely to be affected by the roaming regulation, as a larger share of their clientele experiences the decrease in mobile internet tariffs. For instance, in the 2.5 years preceding the regulation, Tripadvisor reviews from EU travelers accounted for about 30% of the total volume in restaurants located in the most touristy areas of the Province, while for less than 1.5% in the least touristy ones.

1.4.1 Restaurants' exposure to tourists

I take advantage of the granularity of my data and construct two measures of exposure to tourist clientele that account for the intensity to which each restaurant is potentially affected by the policy. I then use such measures to identify the parameters of interest. In both cases, I rely on the location of a restaurant with respect to tourist attractions to predict the composition of its clientele.

The first measure aims to capture the probability that a tourist finds a restaurant while walking away from an attraction site. To build this measure, I consider the top-100 attractions in the Province of Rome, according to their total volume of Tripadvisor reviews. For each attraction, I identify the closest Tripadvisor restaurants around it, and then use the Google Maps API to construct the shortest walking route from the attraction site to each of these restaurants. The procedure generates the partial road network around every attraction. I then assume that tourists follow a *random walk* while they move away from the attraction site, which allows me to assign equal conditional probabilities to every road at a same junction. Finally, I compute the joint probability to find the restaurant(s) located at any point along the network as a product of the conditional probabilities attached to all the consecutive roads leading to that point. The procedure, which is described in details in Appendix 1.B.1, provides a continuous probability measure $P(i) \in [0, 1]$ that reflects the chances that restaurant i is visited by tourists while they move randomly toward the periphery of the road network, starting from the attraction site. Hence, by construction, this quantity only depends on (1) the location of the restaurant with respect to its closest tourist attraction, and (2) the shape and density of the road network around it.[30]

Appendix 1.B.2 shows that such probability is positively associated with the average share of reviews from tourists, while it is negatively correlated with the average Tripadvisor rating of restaurants, consistently with theories on asymmetric information, repeat purchases and product quality (Cooper and Ross 1984; Riordan 1986). Moreover, I investigate the robustness of my procedure to different data and assumptions. First, instead of focusing only on the shortest path, I also include all alternatives routes provided by the Google Maps API in the computation process. Second, rather than imposing equal conditional probabilities (random walk assumption), I assume tourists form educated guesses on which path to follow, based on importance (frequency) of each road. These procedures are explained in Appendix 1.B.1, and the resulting probability measures will be used in the analysis to conduct robustness checks.

The probability measure described above varies across restaurants, which is an attractive feature to study firm-level response. Such a granular level of variation, however, does not allow to identify the aggregate (market-level) effects of the policy on the industry composition. For this purpose, I focus on the ZIP codes in the province of Rome, and construct an alternative measure of exposure to tourists at that level. Particularly, I focus on the number of top-100 tourist attractions in each ZIP code. In turn, this measure reflects the potential exposure of all restaurants in a ZIP code to tourist clientele and, therefore, to the change in internet tariffs induced by the policy. Among the 127 ZIP codes in which I observe one or more restaurants of my sample, about 25% contain at least one tourist attraction, with the most touristy ZIP codes containing 25 sites.

1.4.2 Identification

The basic idea behind the identification strategy is to compare the evolution of firm-level and ZIP-level outcomes before and after the policy across firms/ZIP codes that are

[30] The advantage of this approach is that it relies exclusively on location parameters. Alternatively, one could use other information contained on Tripadvisor, such as the origin of reviewers, to determine the level of exposure to tourists for each restaurant. However, such information is the result of past consumption patterns and reviewers' behavior that might interplay with the policy, influencing future consumers' decisions on a restaurant regardless of its actual level of exposure to tourists.

differentially exposed to tourists, and therefore to the change in the roaming tariffs.

1.4.2.1 Firm-level outcomes

The firm-level analysis employs the first measure of exposure to tourist clientele, i.e. the probability $P(i)$ previously defied. Appendix 1.B.3 shows that – because of the composition of their clientele – only restaurants with a sufficiently high probability are potentially affected by the policy, as both the pre-policy shares of EU reviews and their change across devices (from PC to mobile) after the policy are significantly higher for restaurants with probability values above the median. Thus, the benchmark empirical specification of the paper relies on these facts to identify two separate groups of restaurants: the treatment group, composed of restaurants with a probability value above or equal to 0.17% – i.e., the median – and the control group with the remaining restaurants. Specifically, I consider observations within a symmetric time-window around the policy (i.e., Jan 2015 - Dec 2019),[31] and estimate the following Difference-in-Differences model:

$$y_{i,t} = \beta\, Tourist_i \times Post_t + \alpha_i + \gamma_t + \phi\, \mathbf{x}_i \times Post_t + \varepsilon_{i,t} \tag{1.19}$$

where i is the restaurant, and t is time. Depending on the outcome, the analysis is conducted at the monthly or yearly level. $Tourist_i$ is a binary variable taking value 1 if the measure of tourist exposure $P(i) \geq 0.17\%$. $Post_t$ takes value 1 after the policy, that is for t after May 2017 when outcomes are observed at monthly frequencies, while for t after 2016 when outcomes are annual.[32]

α_i and γ_t represent restaurant and time fixed effects, respectively. Their inclusion allows controlling for both time-invariant firm-level characteristics and aggregate trends (such as seasonality) that might affect the outcome $y_{i,t}$ while being simultaneously correlated with the main independent variable, $Tourist_i \times Post_t$. Nevertheless, there is still the possibility that some demand- and supply-side factors might influence the outcomes of interest over time, while being simultaneously correlated with the main independent variable. To account for such potential endogeneity issue, vector $\mathbf{x_i}$ includes a series of time-invariant and predetermined restaurant-specific characteristics, which – once

[31] With the exception of financial and rating data, which are only available up to 2018.

[32] In the yearly analysis, year 2017 is assumed to be fully treated even if the policy was effective in June. If anything, this should reduce the size of estimated coefficients, thus providing a lower-bound for the effect of the policy.

interacted with $Post_t$ – are allowed to have different impacts on the outcomes over time. In particular, in all regressions I control for the distance (in km) of the restaurant to its closest attraction to account for factors, other than the presence of tourists, that correlate with proximity and could affect restaurants' and consumers' decisions.[33] It is worth mentioning that, once I control for the proximity to the attraction, the probability measure $P(i)$ mainly captures the "visibility" of a restaurant, i.e. whether the place is easy or difficult to be discovered by a tourist due to the shape and density of the road network. Moreover, other controls include restaurant price categories, a dummy indicating whether its cuisine is Italian or not, indicators for the concentration of restaurants within a 400-meter radius (reflecting the level of competition), indicators for the volume of reviews to the closest attraction (capturing the popularity of the whole area, potential congestion and rental costs), the classification of the main economic activity of the restaurant (Ateco code) and its legal status (e.g., LLC vs sole proprietorship).[34] Appendix Table 1.D4 reports the list of independent and control variables along with their descriptive statistics. Finally, in all regressions, I also include ZIP-code linear time trends, as well indicators for the distance to Rome city-center, to account for potentially diverging patterns in the outcomes across areas of the Province that are subject to different exposure to tourist demand and municipal regulations.[35] I cluster the standard errors at municipality level (86 clusters) to account for serial and spatial dependence in the errors.[36]

Theory described in section 1.2 posits a differential impact of a reduction in search costs on demand and production choices across restaurants selling *ex-ante* different qualities. While demand for higher-quality products is expected to increase more, firm exit and quality-upgrading should be more prevalent among lower-quality restaurants. To empirically analyze the changes in the outcomes along the quality gradient, I estimate model (1.28) on different samples. First, I study the overall impact focusing on all

[33]Examples include rent costs that tend to be higher closer to attraction sites, or congestion (in fact, restaurants in touristy areas can easily be overcrowded, thus leading to longer waiting times, more noise, and worse service).

[34]Although most of the control variables were originally continuous (or categorical with many values), their simplified versions are used in the analysis. In fact, importing external rich data in the INPS servers was not allowed for confidentiality concerns, as the researcher could exploit the high-dimensionality of the dataset to identify specific firms and their records. Therefore, I had to re-categorize the original variables before importing them. Importantly, being these controls, their simplification should not crucially affect the results.

[35]In Italy, sanitary and hygienic regulations of the restaurants as well as their structural standards (such as capacity and equipment) are generally established by the municipal councils.

[36]I provide robustness of the estimates to a different level of clustering, using ZIP codes.

matched restaurants with available Tripadvisor rating at the time of the policy (N=4,628). Second, I use the tertiles of such average rating to split the sample in three sub-samples of equal size and estimate the model on each group, separately. Appendix figure 1.C2 shows the overall rating distribution, and highlights the three subgroups of interest: low-rating restaurants, with rating < 3.85; mid-rating ones, with rating $\in [3.85, 4.25)$; high-rating ones, with rating ≥ 4.25. It is worth mentioning that Tripadvisor does not display the average rating of a restaurant, but rather its rounded value.[37] Therefore, these three groups contain restaurants whose displayed ratings are approximately below, around, and above 4, respectively.

By estimating (1.19) via OLS, the coefficient of interest β reflects the change in the outcomes before/after the policy that restaurant more exposed to tourists experience with respect to those with lower probability values. In order for β to have a causal interpretation, the identification assumption requires that changes in the outcomes across the two groups of restaurants would have been the same in the absence of the policy. This so-called parallel trends assumption entails that the two groups are, on average, comparable over time and that the policy is exogenous with respect to other factors, such as seasonality, tourist flows composition and anticipation effects. I conduct a number of placebo exercises to provide plausible evidence in support of this assumption. These include (i) event study estimates, where the dummy variable $Tourist_i$ is interacted with semester dummies, which allow to both study the timing of the impacts after the policy, and check for the presence of differential trends in the outcomes in the pre-policy period; (ii) a series of permutation tests, where the effect of several placebo policy-dates between 2012-2016 is assessed; (iii) specific placebo policy-dates coinciding with the months of May in years 2013-2016 to explicitly test whether the start of the tourist season could explain the main results.

Finally, there is still the possibility that online information from mapping apps[38] helps tourists navigate the streets around attractions, allowing them to discover less visible restaurants – e.g., those around the corner or in certain hidden alleys of the city center –, which would have not been visited otherwise (e.g., as suggested by Ghose et al. 2013 and Dall'orso et al. 2016). Thus, I take advantage of the granularity of my probability

[37] In both Tripadvisor and Yelp the average rating is rounded at the nearest half integer. So for example, a 3.73 average rating would be rounded to 3.5. Some papers like Farronato and Zervas (2019) and Luca (2016) take explicit advantage of this feature in their identification strategy.

[38] These include Tripadvisor, which has a "find-near-me" option, but also other popular apps such as Google Maps.

measure to study the potential reallocation of consumption over space, from highly visible establishments to more hidden restaurants that are nevertheless easy to reach (i.e., within walking distance) for tourists. To do so, I allow for (1.28) to take a more flexible form, where I use deciles/quantiles of $P(i)$ instead of the dummy variable $Tourist_i$, and interact them with $Post_t$. I display these results in a series a figures.

1.4.2.2 ZIP-level outcomes

To study the aggregate effects of the policy on the distribution of equilibrium qualities in the industry, I group establishments at the ZIP-code level and exploit the variation in the number of tourist attractions to proxy for exposure to tourist clientele (as described in 1.4.1). In this setting, a Diff-in-Diff approach would compare changes over time across ZIP codes with a higher and lower number of attractions. Particularly, I focus on the matched sample of restaurants irrespectively of their presence on Tripadvisor at the time of the policy (N=5,472)[39] and, similarly as before, I consider observations between January 2015 and December 2019 and estimate the following equation:

$$y_{z,t} = \beta \, Attractions_z \times Post_t + \alpha_z + \gamma_t + \phi \, \mathbf{x}_z \times Post_t + \varepsilon_{z,t} \tag{1.20}$$

where z is the ZIP code, and t is time, measured in months. $Post_t$ takes value 1 after May 2017. $Attractions_z$ is a time-invariant variable containing the number of attractions located in z. α_z and γ_t are ZIP-code and month fixed-effects, respectively. To account for potentially diverging trends in the outcomes across different ZIP codes, I also include ZIP-level linear time trends. Moreover, in vector $\mathbf{x}_z$, I include categorical variables reflecting the average distance of restaurants in the ZIP code to Rome city-center, which I interact with $Post_t$ to control for factors, other than the presence of tourists, that correlate with proximity to the main city and might affect consumption and production choices over time. I cluster the standard errors at ZIP-code level (127 clusters) to account for serial correlation in the errors.

By estimating (1.20) via OLS, the coefficient of interest β reflects the change in the outcomes before and after the policy, across ZIP codes that are more and less exposed to tourist clientele. In order for β to have a causal interpretation, the identifying assumption requires that, in the absence of the policy, the outcomes of different ZIP

[39]Note that this allows me to study the effects of the policy not only on the exit but also on the entry-type of new restaurants.

codes would have changed similarly. To check the plausibility of this assumption, I perform a variety of event-study and placebo estimates similar to those described in the firm-level analysis of Section 1.4.2.1.

1.5 Results

This section presents the supply and demand effects of the reduction in consumer information costs from the estimation of the benchmark specifications (1.19) and (1.20). Section 3.4.6 discusses the plausibility of the identifying assumption for all the outcomes and Section 1.7 assesses the sensitivity of the main results to alternative specifications, clustering units and measurements.

1.5.1 Restaurant revenues and size

1.5.1.1 Main results

Theory presented in Section 1.2 predicts that higher-quality firms increase their output as a consequence of lower consumers search costs. To test this hypothesis in the absence of data on quantity, I first rely on restaurant revenues as a proxy for output, using the Tripadvisor rating at the time of the roaming policy to proxy for the baseline quality of the restaurant. Particularly, I consider the sample of restaurants with available annual financial information and estimate equation (1.19). Column (1) of Table 1.2 shows that, after the policy, sales in more touristy restaurants increased by almost 5% compared to less touristy ones. The estimated coefficient is robust to the inclusion of additional controls (column 2), such as the price category of the restaurant and the type of cuisine, which might be correlated with both revenues and the level of exposure to tourists. The most conservative estimates imply an annual average increase in restaurant revenues of approximately 32.5 Thousand Euros, considering that mean revenue in the pre-policy period was around 650 Thousand Euros.

Columns (3-5) of the table analyze differential effects of the policy across restaurants with *ex-ante* different ratings – i.e., low, mid and high rating –, estimating equation (1.19) on three sub-samples containing a similar number of restaurants. Coefficients suggest that the overall increase in revenues is mainly driven by high-rating restaurants, whose sales expanded by almost 7%. Revenues in the mid-rating category also improved but to a smaller extent, by approximately 3%. By contrast, the policy had no impact

Table 1.2: The impact of the roaming policy on restaurant revenues

	Y=log(annual revenues); years 2015-2018				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	0.047***	0.053***	-0.002	0.033**	0.069***
	(0.012)	(0.012)	(0.026)	(0.015)	(0.024)
Restaurant & Year FE	✓	✓	✓	✓	✓
ZIP-code*Year	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	6677	6652	2305	2299	2048
Restaurants	2043	2034	696	697	641
Clusters	57	56	39	40	41
Adj. R-squared	0.846	0.847	0.869	0.849	0.782
Mean Y pre-policy	646.6	648.8	977.4	558.0	360.7
DDD *p-value*				0.962	0.004

Post=1 if date is after 2016. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-year. The sample includes observations between 2015 and 2018. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

on sales in the low-rating category, for which the estimated coefficient is virtually zero and not statistically significant. To test whether the change in revenues in the mid- and high-rating groups is statistically different from the low-rating one, I perform a triple-difference estimation, where dummy variables for mid- and high-rating categories are interacted with $Tourist \times Post$. The last raw of columns (4-5) report the p-values of these coefficients, which confirm a positive revenue gradient along the rating dimension, with the change in high-rating restaurants being significantly different than that in the low-rating ones.

Data on revenues might be subject to measurement error, for instance due to firms misreporting their sales in the attempt to pay lower taxes. Therefore, I complement the analysis of restaurant output using monthly employment records.[40] To some extent, changes over time in the number of employees – i.e., firm size – reflect the variation in the restaurant's ability to attract clientele and fill-up the tables. However, the relationship between output and firm size is not necessarily linear (Basu and Fernald, 1997), especially when firms face capacity constraints, which likely imply decreasing

[40] Labor information is generally more difficult to cover up and misreport to the authorities compared to financial data.

returns to labor. In the case of restaurants, such constraints arise because of the narrow time-windows to serve a meal (launches and dinners) and limited physical space. In practice, an additional worker would not be much productive when all the tables are already filled-up and clients have to wait in line outside of the restaurant for the next available seat. Conscious of this potential limitation of labor data, I estimate the effect of the policy on the (log) number of employees for all restaurants with available Tripadvisor rating at the time of the regulation. Figure 1.3 shows event-study estimates from separate regressions on the three sub-samples corresponding to the different rating categories previously defined. At the same time, Table 1.3 reports the regression output.

Figure 1.3: Event-study estimates for restaurant employment

Notes: The graph reports estimated coefficients on the interactions of Tourist*Semester dummies from three separate regressions where each observation is a restaurant-month-year. All controls and fixed-effects from the main analysis are included. The omitted semester is 2017s1. Tourist restaurants are those with a measure of exposure above the median. The sample includes observations between Jan 2015 and Dec 2019. Shaded areas depict 95% confidence intervals.

In line with findings on restaurant revenues, after the policy, total monthly employment expanded by approximately 4% in more touristy restaurants. The estimated coefficient does not change when additional controls are included in the regression (columns (1-2) of Table 1.3). Figure 1.3 reveals that the mid-rating category is mostly responsible for the overall increase. Here, on average, monthly employment grew by 10% after the policy, implying an average increase in the restaurant size by more than 0.5 workers, when compared to the pre-policy mean. The event-study estimates also suggest that such labor expansion did not take place immediately, but it rather happened around 6-12 months after the policy. This suggests some delay either in the consumer learning

Table 1.3: The impact of the roaming policy on restaurant employment

	Y=log(monthly employees); Jan 2015 - Dec 2019				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	0.043**	0.042**	-0.024	0.103***	0.041
	(0.017)	(0.020)	(0.019)	(0.031)	(0.038)
Restaurant & Time FE	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	219835	217622	72133	76920	68569
Restaurants	4628	4576	1490	1571	1515
Clusters	86	86	59	71	71
Adj. R-squared	0.779	0.778	0.759	0.793	0.769
Mean Y pre-policy	5.5	5.6	6.9	5.7	4.0
DDD *p-value*				0.089	0.065

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

process or in the ability of restaurants to adjust their factors of productions when they experience boost in demand. Moreover, while revenues in the high-rating category expanded, labor did not change significantly. The estimated coefficient is around 4%, yet it is not statistically different from zero at any conventional confidence level. This result can signal the presence of decreasing returns to labor in the restaurant industry. Particularly, high-rating restaurants were likely to be popular among Italian tourists and locals even before the roaming policy, and therefore were already producing at full capacity. The additional demand increase that they get from EU travelers translates into higher revenues, but not into more workers. Finally, and consistently with the revenue analysis, low-rating restaurants do not exhibit any significant change in their size. If anything, the sign of the estimated coefficient is negative (-2.5%) but not statistically different from zero. The last row of columns (4-5) also suggests that the positive trends in employment for mid- and high-rating restaurants are statistically different than those in the low-rating category.

1.5.1.2 Discussion and additional findings

The evidence on revenues and employment presented so far suggests that, because of the cheaper information available to consumers, restaurants with *ex-ante* a better online reputation on Tripadvisor – namely with a rating around 4 or above – attracted new clients and grew in size more than those with worse ratings (below 4). These findings are consistent with those from the previous literature (e.g., Anderson and Magruder 2012; Chevalier and Mayzlin 2006; Lewis and Zervas 2019; Luca 2016),[41] and add to this existing work providing new empirical evidence on the general effects of online word-of-mouth on firm employment.

The fact that gains at the top of the rating distribution are not symmetrically compensated by losses at the bottom could be explained by several reasons. One is market expansion, which is consistent with the average increase in revenues/employment found on the entire sample (e.g., columns (1-2) of Tables 1.2 and 1.3). Aggregate demand expansion could occur, for example, if some consumers start substituting food from supermarkets with meals at the restaurants. However, such dynamic is unlikely to be the exclusive reason underlying the overall positive effects of the policy. Another possible explanation is demand substitution from restaurants with no Tripadvisor account (which are out of the sample and therefore not observed) to those with an active profile (in-sample). Nevertheless, even this type of substitution should not play a major role, as the greatest majority of the restaurants in the Province was already on Tripadvisor around the time of the policy.

Alternative explanations bring into play supply-side dynamics rather than market expansion. For example, firm exit (which is discussed in Section 1.5.2) might lead to a reduction in the total number of players in the market, leaving more clients – and therefore more revenues – to the surviving restaurants even if aggregate demand does not change. Moreover, upward price adjustments in high-rating restaurants could explain the overall larger revenues. This is in line with the evidence on profit margin reported in Appendix Table 1.D5, which shows that profits increased after the policy only in the high-rating category. In fact, these restaurants could benefit from their online reputation to raise their market power – most likely, they had their tables already filled-up before the new regulation –, and eventually increase their markups without loosing much of

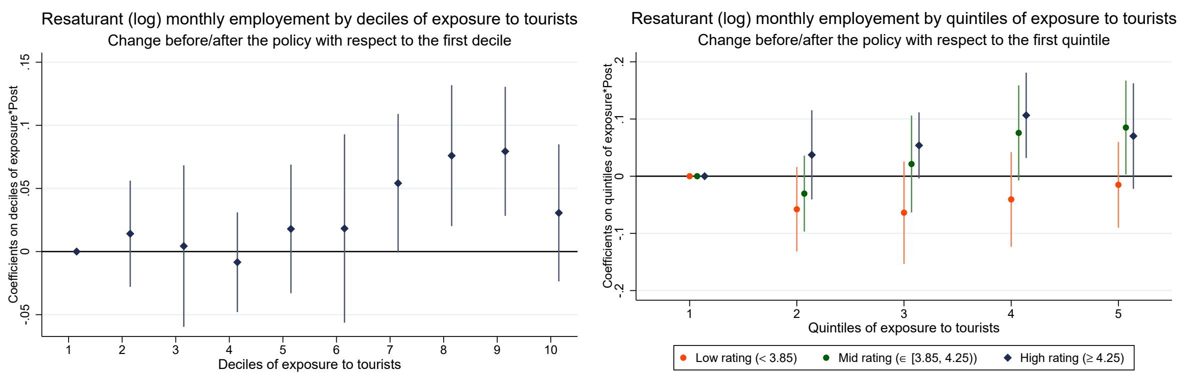
[41] For example, using data on several online platforms such as Yelp, Tripadvisor, Expedia and Hotels.com, Lewis and Zervas (2019) and Luca (2016) have found that a one-star increase in rating leads to a 5-9 percent increase in restaurant/hotel revenues.

their clientele. Nevertheless, the remarkable growth in employment among mid-rating establishments clearly indicates that more than just a price adjustment is going on in this category, and some demand-side dynamics must be driving their expansion. Unfortunately, my data does not allow to disentangle the specific mechanism behind the overall positive effects of the policy, and all these hypotheses remain equally plausible. Regardless of which one prevails, my results speak in favor of the first theoretical prediction of the model: as a consequence of lower costs for consumers to learn about product quality, demand to better-quality producers increases.

So far, I relied on the extensive margins of employment – i.e., the total number of workers – to measure restaurant size and proxy for output. Table 1.D6 in the Appendix presents the results using the (log) number of total working days, which also capture adjustments at the intensive margins. Coefficients are in line with the previous ones, with some additional evidence suggesting negative and significant effects (at the 10% level) in low-rating restaurants, where the number of working days shrank by almost 7% after the policy. Such a reduction might be the result of a negative demand shock that these establishments experienced. Moreover, to isolate changes at the intensive margins only, I consider the number of monthly working days per worker. Appendix Table 1.D7 reveals that, after the policy, each employee in high-rating restaurants worked, on average, 0.4 days more per month, while those in mid-rating ones worked almost half-a-day less. The first finding is consistent with the presence of capacity constraints in high-rating restaurants: rather than hiring additional employees and expand in size, these firms demanded more days of work to their current personnel. The second result is consistent with the view that lower-quality restaurants may attempt at improving their service quality through strategic employment choices, for instance by hiring new dining room staff while guaranteeing them better working conditions (e.g., shorter shifts). Such a mechanism will be covered in detail in Section 1.5.3.

Previous work (e.g., Lewis and Zervas 2019; Luca 2016) has found that online reputation is more important for independent restaurants, where asymmetric information is more severe compared to chains. My data lack information on resultants' affiliation, but they contain details on their price range (i.e., the market segment). In this respect, cheap (e.g., fast food) and fancy starred places are expected to gain less from the information provided by online reviews than those in the middle segment, even when their ratings are high. For instance, low and middle budget tourists (which represent the majority of visitors) are more willing to substitute a low-price restaurant with a medium-price one, once they are reassured about the good quality of the latter. Yet, at the same time, fine-dining restaurants would remain outside of their consideration set. Table 1.D8 in the Appendix replicates the benchmark estimation for three different price categories and provides some evidence in support of this hypothesis. Particularly, it shows that high-rating mid-price restaurants expanded their total employment by approximately 10% after the policy, while the corresponding coefficients for low- and high-price restaurants are negative and not significant. By contrast, employment decreased in low-rating cheap and expensive restaurants after the policy.

Figure 1.4: The impact on restaurant employment across levels of exposure to tourists



(a) (b)

Notes to Panel (a): The graph reports estimates on the interactions of deciles of exposure*Post from a regression where each observation is a restaurant-month-year. The first decile is omitted. All controls and fixed-effects from the main analysis are included. Post takes value 1 after May 2017. The sample includes observations between Jan 2015 and Dec 2019. Heteroskedasticity-robust standard errors are clustered at municipality level. Bars depict 95% confidence intervals.
Notes to Panel (b): The graph reports estimates on the interactions of quintiles of exposure*Post from three separate regressions (low, mid, high-rating restaurants), where each observation is a restaurant-month-year. The first quintile is omitted. All controls and fixed-effects from the main analysis are included. Post takes value 1 after May 2017. The sample includes observations between Jan 2015 and Dec 2019. Heteroskedasticity-robust standard errors are clustered at municipality level. Bars depict 95% confidence intervals.

The benchmark findings are robust to different specifications, measurements and samples. For instance, I examine whether my estimates are driven by the specific choice over the construction of the treatment variable $Tourist$. Particularly, rather than identifying just two groups of restaurants according to the median value of the probability measure, I consider its deciles and estimate a more flexible specification, interacting them with the dummy variable $Post$. Besides providing a robustness for my main results, this approach allows to study demand reallocation over space, e.g., from restaurants located in front of tourist attractions to those "hidden" in the surrounding alleys. To some extent, Panel (a) of Figure 1.4 suggests that such a reallocation is likely to take place. The impact of the roaming policy on employment is not statistically significant for restaurants at the 10th decile of exposure to tourists (the most visible ones), while the coefficients on the 7-8-9th deciles are driving the overall positive results. Moreover, coefficients on lower deciles are remarkably smaller in size, and always insignificant. In a similar fashion, Panel (b) replicates the same exercise across both rating categories and quintiles of exposure to tourists.[42] Point estimates displayed in the figure are qualitatively consistent with those from the main analysis and confirm that more touristy higher-rating restaurants drive the overall results. A similar conclusion hold for revenues, as discussed in Section 1.7. This section also conducts additional sensitivity analysis, such as using different clustering units for the standard errors, focusing only on a balanced sample of restaurants that never exit the market after the policy, and using alternative measures of tourist exposures. My results are generally robust to these alternative specifications, reassuring about potential concerns on selection bias and measurement choices. Finally, placebo exercises carried out in Section 3.4.6 provide additional evidence in favor of the parallel trends assumption for restaurant revenues and employment.

1.5.2 Industry composition

The second set of hypotheses presented in Section 1.2 concerns the supply side, namely: (1) firms' decisions to stay in the market or not and, conditional on staying, (2) their level of investment into quality. This section covers the former, while the latter will be discussed in Section 1.5.3. Firm dynamics represents one potential mechanism through which the cheaper access to information from online reviews could affect the overall quality levels in the industry. Theory predicts that when consumers face lower search costs, those firms producing the lowest-quality products are more likely to be pushed out of the market (i.e., a reduction in the adverse selection problem). To empirically test the effects of the roaming regulation on the industry composition and isolate the role of entry/exit dynamics (as opposed to quality upgrading), I track the presence of restaurants in the market over time by rating category. Particularly, I use the official date of registration and termination of the business as recorded by the INPS database to proxy for firm entry and exit, respectively. The analysis is conducted both at the firm and ZIP-code levels.

[42] In this case, I use quintiles in order to have a sufficient and representative number of observations within each pair *(quintile, rating category)*.

1.5.2.1 Firm exit

The firm-level framework described in equation (1.19) exploits the within-firm variation in the outcomes of interest over time, for the sample of firms with available Tripadvisor rating – and therefore operating in the market – at the time of the policy. As such, this setting only allows to study firm exit, and not entry.[43] Specifically, I construct a dummy variable that takes value 1 when a firm exits the market and 0 otherwise, and estimate (1.19) via OLS. Table 1.4 presents the estimation results of this linear probability model.

Table 1.4: The impact of the roaming policy on restaurant exit

	Y=1 if firm exits the market; Jan 2015 - Dec 2019				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	0.0011*	0.0016**	0.0031***	-0.0000	0.0015
	(0.0006)	(0.0006)	(0.0010)	(0.0009)	(0.0024)
Restaurant & Time FE	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	219835	217622	72133	76920	68569
Restaurants	4628	4576	1490	1571	1515
Clusters	86	86	59	71	71
Adj. R-squared	0.059	0.060	0.058	0.061	0.061
Mean Y pre-policy	0.003	0.003	0.003	0.003	0.004
DDD *p-value*				0.056	0.558

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Columns (1-2) show that, after the policy, monthly exit rate among more touristy restaurants increased by 0.11-0.16 percentage points, with respect to less touristy ones. Comparing these coefficients with the average pre-policy exit rates (0.3%) reveals that the frequency at which firms leave the market went up by approximately 35-55% during the 30 months after the new regulation. Columns (3-5) decompose the effect across different rating categories, providing a direct test for the theoretical predictions. Column

[43] In equation (1.19), the coefficient of interest on $Tourist \times Post$ cannot be identified if firms' outcomes are not observed both before and after the policy. Moreover, if a firm enters the market later on, its Tripadvisor rating at the time of the policy would not be available. For these reasons, firm entry will be studied in the aggregate analysis presented later.

(3) shows that the overall increase in the exit rate is mainly driven by low-rating restaurants: their frequency to exit the market went up more than 0.30 percentage points, which corresponds to doubling the pre-policy average exit rate in this category. At the same time, exit rates in the mid- and high-rating groups were not significantly affected: the estimated coefficients in columns (4-5) are much smaller in size, and they are not statistically different from zero at any conventional confidence level. Overall, this evidence supports the theoretical predictions of the model and points out that cheaper access to information for consumers, made possible through the access to review platforms, has the potential to push some of the lowest-quality providers out of the industry, alleviating the adverse selection problem in the experience goods market.[44]

1.5.2.2 Aggregate effects

Does the above result hold in the aggregate, when firm entry is taken into account? To answer this question, I rely on the ZIP-level framework described in equation (1.20), which provides a more suitable setting to study changes in the industry composition. For each ZIP code/month, I consider the (log) count of active restaurants – of any rating, as well as in the three rating categories previously identified – and regress it on the number of attractions, which is a proxy for exposure to tourist clientele. As in the above analysis, I use the Tripadvisor rating at the time of the policy to proxy for the quality of the restaurant. In addition to that, to measure the quality of the entrants joining the market after the policy and assign them to one of the three rating categories, I use the most recent Tripadvisor rating.[45] Columns (1-4) of Table 1.5 present the results.

I find empirical support for the hypothesis that lower search costs – even when experienced by only a fraction of the total consumers – can make the industry more quality-oriented. Column (1) indicates that, after the policy, one additional tourist attraction in the ZIP code is associated with a reduction in the overall number of active restaurants by 0.4%. Notably, columns (2-4) show that low-rating restaurants are the main drivers of such effect: their number decreased by 0.6% after the policy, for any additional attraction in the ZIP code. By contrast, the count of mid- and high-rating restaurants was not significantly altered. The latter result discards the presence of a "superstar effect" in this context, and challenges the view that cheaper access to information should reduce the number of high-demand firms, while increasing their market shares (Brynjolfsson

[44]This result is in line with what Hui et al. (2018) finds to hold in online marketplaces (eBay).

[45]I focus on the most recent rating so to have a sufficient number of underlying reviews to compute it.

Table 1.5: The aggregate effects of the roaming policy on industry composition

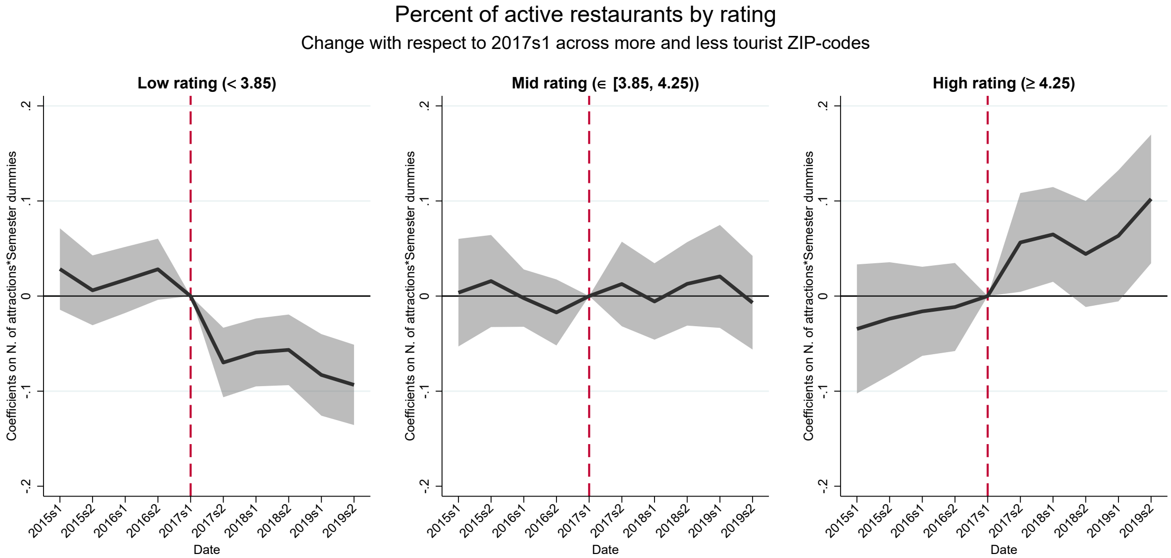
Y=	log(N. of active establishments)				% of active establishments	
	(1) All	(2) Low rating	(3) Mid rating	(4) High rating	(5) Low rating	(6) Mid rating
N. of attractions * Post	-0.004**	-0.006***	-0.003	0.001	-0.102**	0.010
	(0.002)	(0.002)	(0.002)	(0.002)	(0.049)	(0.096)
ZIP-code & Time FE	✓	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓	✓
Observations	7501	7501	7501	7501	7501	7501
ZIP-codes	127	127	127	127	127	127
Adj. R-squared	0.993	0.989	0.985	0.976	0.930	0.882
Mean Y pre-policy	29.90	10.43	10.31	9.16	32.76	34.65

Post=1 if date is after May 2017. Each observation is a ZIP-code-month-year. All regressions include the distance of the ZIP-code to Rome city center interacted with Post. The sample includes observations between Jan 2015 and Dec 2019. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. If the restaurant entered the market after the policy, the most recent rating is considered. Heteroskedasticity-robust standard errors clustered at ZIP-code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

et al., 2010). Most likely, this is due to the peculiarity of the restaurant industry: the presence of (physical) capacity constraints limits the expansion of their output above a certain threshold.

Furthermore, columns (5-6) of Table 1.5 consider the percentage (expressed in 0-100 points) of active restaurants in each rating category, and show that the percent of low-rating restaurants operating in markets more exposed to tourist clientele shrank after the policy. Columns (5) suggests that the presence of one additional tourist attractions in the ZIP code made the proportion of low-rating restaurants decrease by more than 0.10 percentage points. Back-of-the-envelope calculations indicate that the share of low-rated restaurants operating in the most touristy neighborhoods (25 attractions) decreased by 2.5 percentage points after the policy, compared to non-touristy ZIP codes (0 attractions). Figure 1.5 plots the event-study estimates, which confirm the previous findings and show that the quality distribution changed for the better, with low-rating restaurants leaving room for high-rating ones. These patterns also provide visual evidence in favor of the parallel trends assumption, while additional placebo exercises are discussed in Section 3.4.6. Finally, Section 1.7 shows that the above estimates are robust to different specifications and measurements. All together, these findings suggests that the distribution of equilibrium qualities in the restaurant industry improved because some of the existing tourist traps (low-rating restaurants) were forced to leave the market. The next section examines restaurants' production choices as an additional mechanism for achieving better qualities.

Figure 1.5: Event-study estimates for industry composition



Notes: The graph reports estimated coefficients on the interactions of N. of attractions*Semester dummies from three separate regressions where each observation is a ZIP-code-month-year. Controls and fixed-effects from the ZIP-code-level analysis are included. The omitted semester is 2017s1. The sample includes observations between Jan 2015 and Dec 2019. Shaded areas depict 95% confidence intervals. Rating is computed at the time of the policy. If the restaurant entered the market after the policy, the most recent rating is considered.

1.5.3 Restaurant quality upgrading

The last prediction of the theoretical model presented in Section 1.2 states that reductions in consumer search costs affect restaurants' incentives to improve quality, and more so for lower-quality establishments (i.e., a reduction in the moral hazard problem). To test this hypothesis, I consider several proxies for both input and output quality. Producing high-quality outputs typically requires high-quality inputs (see, e.g., Bastos et al. 2018; Halpern et al. 2015; Hansman et al. 2020; Kugler and Verhoogen 2012). This relationship is also visible in the theoretical model, where at better output qualities correspond higher costs of productions (better inputs), holding firm productivity constant. Empirically, I consider restaurants' hiring decisions over workers with different experiences, as well as their wages, as a proxy for changes in the service quality.[46] Moreover, following recent empirical literature (e.g., Ananthakrishnan et al. 2019; Chevalier et al. 2018; Proserpio and Zervas 2017), I use the online reputation of the restaurant as reflected by the average dynamic Tripadvisor rating to proxy for the quality of the output. While hiring decisions capture a more objective dimension of quality – i.e., that of the factors

[46] Along similar lines, Shin et al. (2021) use labor market outcomes to study the impact of the gig economy (Uber and Lyft) on restaurant service quality.

of productions and, specifically, labor –, online ratings reflect the subjective experience of the consumers with the good. Thus, to some extent, the two sources complement each other.

1.5.3.1 Hiring decisions and salaries

I take advantage of the employer-employee matched data and investigate if, in their attempt at improving service quality, restaurants are more likely to hire workers with better curricula, as measured by their previous experience in the restaurant sector.[47] First, I consider the full employment history of every newly-hired employee in my sample of restaurants, and construct dummy variables indicating whether, by the time of their appointment, they had previously worked in the restaurant sector or not.[48] For each outcome, I estimate equation (1.19) on the entire sample of restaurants, as well as on the three sub-groups corresponding to the different rating categories.

In columns (1-4) of Table 1.6 the dependent variable is a dummy taking value 1 in months in which the restaurant hires a new employee who had previously worked in other restaurants, and 0 otherwise.[49] Column (1) indicates that, overall, after the regulation, restaurants more exposed to tourist clientele became almost 1-percentage-point more likely to hire experienced employees, which corresponds to approximately a 10% increase in the pre-policy mean. Consistently with the theoretical prediction, columns (2-4) show that low- and mid-rating restaurants are the drivers of such a change. Their probability to hire workers with better curricula significantly went up by 0.9-1.1 percentage points, an increase of 9-16% with respect to the pre-policy mean. By contrast, the coefficient for high-rating restaurants is considerably smaller in magnitude, close to 0 and not statically significant.

These results could be due to the overall larger employee turnover in lower-rated restaurants, rather than their recruiting strategy being intentionally targeted at more experienced workers. However, columns (5-6) of Table 1.6 indicate that this is unlikely

[47] I exclusively focus on the experience dimension of the worker's curriculum. Unfortunately, other factors (such as education) are not available in the data. Nevertheless, this should not pose a critical obstacle to my analysis, since in the restaurant sector previous experience is likely to be more informative than education to signal the skills of waiters and other room staff.

[48] The restaurant sector is defined by firms with ATECO codes 56.10.11 (dine-in restaurants), 56.10.12 (agriturismi), 56.10.20 (take-away restaurants), 56.10.30 (bakeries).

[49] To facilitate the interpretation of the coefficient estimates, I use a linear probability model (OLS) as the benchmark specification. Coefficients from a Logit model are qualitatively similar and are reported in Table 1.D9 in the appendix.

Table 1.6: The impact of the roaming policy on hiring decisions (extensive margins)

Y=1 if firm hires worker	*with* previous experience in restaurants				*without* experience in restaurants		
	(1) Full sample	(2) Low rating	(3) Mid rating	(4) High rating	(5) Low rating	(6) Mid rating	(7) High rating
Tourist*Post	0.009***	0.009***	0.011***	0.002	-0.006**	0.011**	0.007**
	(0.002)	(0.003)	(0.003)	(0.004)	(0.002)	(0.004)	(0.003)
Restaurant & Time FE	✓	✓	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓
Observations	217622	72133	76920	68569	72133	76920	68569
Restaurants	4576	1490	1571	1515	1490	1571	1515
Clusters	86	59	71	71	59	71	71
Adj. R-squared	0.124	0.143	0.104	0.116	0.049	0.043	0.037
Mean Y pre-policy	0.08	0.10	0.07	0.08	0.06	0.06	0.06
DDD *p-value*			0.259	0.047		0.000	0.002

The sample includes observations between Jan 2015 and Dec 2019. Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (3-4) and (6-7) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

to be the case. In fact, the probability of hiring workers with no experience in the industry significantly decreased by 10% after the policy among more touristy low-rating restaurants, while it increased for the mid- and high-rating ones by almost 18 and 12%, respectively, compared to their pre-policy values. All together, these estimates also suggest that while low- and high-rating restaurants seem to focus their recruiting efforts on distinct and opposite types of workers (experienced vs. inexperienced, respectively), mid-rating ones hire from a more heterogeneous pool of candidates.

While the above results assess the impact of the policy on the extensive margins of targeted recruiting strategies (i.e., whether or not restaurants hire experienced workers), consistent findings are obtained when considering the intensive margins of worker experience. In this case, I restrict the attention only to those months in which the restaurant hires/fires an employee, the contract terminates, or the employee voluntarily quits the job. I then measure the cumulative experience of such workers by counting the total number of months they have been employed in the restaurant sector in the past. This way, I can quantify the impact of the policy on the gain/loss in human capital that restaurants face.

Table 1.7 shows the results. Column (1) indicates that, after the policy, restaurants

Table 1.7: The impact of the roaming policy on hiring decisions (intensive margins)

Y=Months of experience in restaurants of	newly-hired employees				quitting/fired employees		
	(1) Full sample	(2) Low rating	(3) Mid rating	(4) High rating	(5) Low rating	(6) Mid rating	(7) High rating
Tourist*Post	1.469*	2.977***	0.789	0.465	0.177	-0.440	2.375**
	(0.834)	(1.083)	(1.160)	(0.544)	(1.034)	(1.192)	(1.181)
Restaurant & Time FE	✓	✓	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓
Observations	30059	11318	10205	8536	12281	10395	8131
Restaurants	3531	1163	1220	1148	1197	1226	1136
Clusters	76	53	59	61	51	57	58
Adj. R-squared	0.117	0.109	0.117	0.127	0.190	0.170	0.183
Mean Y pre-policy	13.0	13.5	13.5	11.8	25.8	27.0	21.5
DDD *p-value*			0.034	0.000		0.598	0.460

The sample includes observations between Jan 2015 and Dec 2019. Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (3-4) and (6-7) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

in more touristy areas hire workers with additional 1.5 months of previous experience in the industry, compared to less touristy ones. Importantly, columns (2-4) show that low-rating restaurants are the drivers of such a change: after the policy, they hire workers with 3 additional months of experience in the industry, corresponding to a 22% increase with respect to the pre-policy mean. The coefficients for the mid- and high-rating categories are positive but not statistically significant, suggesting that the accumulation of human capital mainly takes place in lower-rated restaurants. By contrast, high-rating establishments appear to loose human capital. Columns (5-7) consider the employment history of those workers who left the firm, either because they decided to quit, their contract expired or they got fired.[50] While no effect is detectable for low- and mid-rating establishments, employees that are let go by more touristy high-rating restaurants are, on average, 2.4 months more experienced compared to less touristy ones in the pre-policy period.

[50]Note that this definition is intentionally broad, for instance it also includes workers who reached their retirement age. This is done to capture the overall loss in human capital that restaurants experience, by looking at any worker who left the firm, irrespectively of the reason.

Changes in the composition of the labor input quality should be reflected in the firms' production costs. In a competitive labor market, firms must pay higher wages in order to attract employees with better skills. Figure 1.6 provides graphical evidence on the evolution of average gross daily salaries paid by the restaurants in the three rating categories.[51] Consistently with the evidence on workers experience presented so far, these event-study estimates point out an increase in the salaries paid to employees of more touristy low-rating restaurants by more than €1 a day. By contrast, salaries in high-rating restaurants decreased by a similar amount, on average, while they did not change in the mid-rating category. At the same time, the figures provide reassuring evidence on the absence of diverging trends in the outcome across more and less touristy restaurants in the pre-policy period. Regression estimates reported in Table 1.8 confirm the graphical analysis. Salaries in low-rating (high-rating) establishments grew (shrank) by almost 2% (1.8%) with respect to their pre-policy mean. By contrast, no significant change in average salaries is detected in the overall sample and in the mid-rating category.[52]

All together, these findings provide persuasive evidence that lower information costs for consumers can affect firms' incentives to upgrade product quality through strategic employee turnover, especially for those producers with the highest margins of improvement. In particular, restaurants with initially lower qualities targeted their hiring efforts at better skilled and experienced workers (both at the intensive and extensive margins) and ended-up paying higher salaries. By contrast, restaurants that were already selling higher qualities took advantage of their established online reputation to divest in human capital and save in production costs. Evidence reported in Appendix Table 1.D5 indicates that these decisions eventually impact on restaurants profitability, since – at least among restaurants with available financial records – the profit margin of high-rating establishments grew by 2.5 points after the policy. By contrast, profits in low- and mid-rating restaurants decreased. Finally, it is worth mentioning that these opposite recruiting strategies might also generate human capital flows from high- to low-rating establishments. As Table 1.D11 in the Appendix shows, low- and mid-rating restaurants became more likely to hire workers previously employed in establishments with better Tripadvisor ratings. This evidence offers an optimistic assessment of the possibility of

[51] To make the salaries of full- and part-time employees comparable, I compute the full-time equivalent salary for part-time employees, using the percentage of the part-time as reported in their contract.

[52] These results are robust to the use of a logarithmic scale of salaries, as shown in Appendix Table 1.D10.

Figure 1.6: Event-study estimates for restaurant daily salaries (€)

Notes: The graph reports estimated coefficients on the interactions of Tourist*Semester dummies from three separate regressions where each observation is a restaurant-month-year. Full-time equivalent salary is computed for part-time employees, according to the percentage of the part-time as reported in their contract. All controls and fixed-effects from the main analysis are included. The omitted semester is 2017s1. Tourist restaurants are those with a measure of exposure above the median. The sample includes observations between Jan 2015 and Dec 2019. Shaded areas depict 95% confidence intervals.

Table 1.8: The impact of the roaming policy on restaurant daily salaries

	Y=Average daily salary (€); Jan 2015 - Dec 2019				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	-0.010	0.038	1.312***	-0.120	-1.125**
	(0.245)	(0.257)	(0.374)	(0.440)	(0.448)
Restaurant & Time FE	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	200402	199026	67507	70593	60926
Restaurants	4558	4512	1492	1538	1482
Clusters	86	86	59	71	70
Adj. R-squared	0.467	0.469	0.485	0.465	0.451
Mean Y pre-policy	64.9	64.9	66.0	65.0	63.5
DDD *p-value*				0.206	0.015

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

quality upgrading in *ex-ante* worse restaurants through human capital externalities from better producers. The next section provides additional evidence in support of this view.

1.5.3.2 Tripadvisor rating

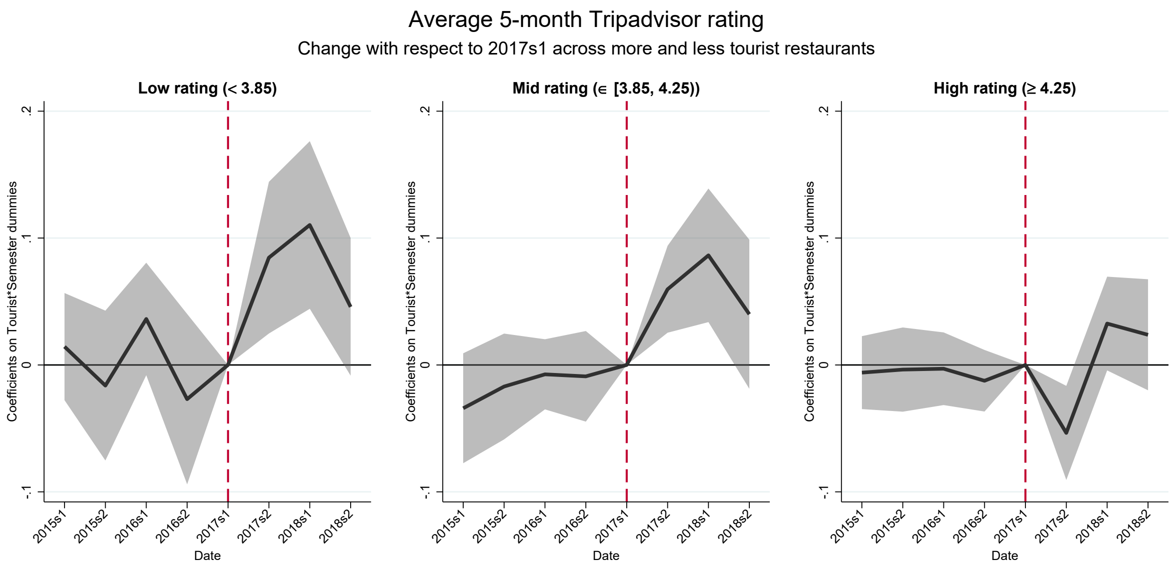
In the absence of objective measures of restaurant output quality,[53] I rely on Tripadvisor rating as a proxy for its reputation dimension. In particular, I compute the moving average of the monthly Tripadvisor rating over dynamic 5-month windows for all the restaurants in my sample. Figure 1.7 displays event-study estimates across the three rating categories, and shows that restaurants in the low- and mid-rating groups received better ratings after the policy, with peaks reaching almost 0.11 and 0.09 points, respectively. Coefficient estimates of model (1.19) reported in Table 1.9 confirm these patters. After the regulation, the 5-month Tripadvisor rating of more touristy restaurants improved by almost 0.05 points overall (i.e., a 1.3% increase with respect to the pre-policy mean), and by 0.09 points (2.5%) and 0.08 points (1.9%) in low- and mid-rating establishments, respectively. By contrast, the coefficient for the high-rating category is virtually zero and not statistically significant, indicating no change in the online reputation of those restaurants already at the top of the rating distribution.

These findings confirm the theoretical predictions of the model and are generally consistent with the evidence on the hiring decisions presented above. Particularly, Appendix Table 1.D12 shows that recruiting workers with previous experience in the restaurant sector is associated with more positive Tripadvisor reviews in the subsequent months. The correlation is even larger when the new employee comes from a higher-rating establishment. By contrast, hiring non-experienced employees has no impact on subsequent rating. These results corroborate the view that better online reputation can be achieved through labor market choices (as in Shin et al. 2021), which are, in turn, a consequence of the lower information costs for consumers.

Nevertheless, two facts might appear less obvious. First, both low- and mid-rating restaurants were successful at improving their online reputation by a similar amount despite their diverse recruiting strategies and salaries, which highlighted that low-rating restaurants more intensively targeted experienced workers. A potential explanation behind the rating improvement in the mid category might be the strategic use of

[53]Previous work (e.g., Farronato and Zervas 2019; Jin and Leslie 2003) relied on health inspection scores to measure the hygiene dimension of quality. Unfortunately, these data are not available for Italy.

Figure 1.7: Event-study estimates for restaurant Tripadvisor rating



Notes: The graph reports estimated coefficients on the interactions of Tourist*Semester dummies from three separate regressions where each observation is a restaurant-month-year. All controls and fixed-effects from the main analysis are included. The omitted semester is 2017s1. Tourist restaurants are those with a measure of exposure above the median. The sample includes observations between Jan 2015 and Dec 2018. Shaded areas depict 95% confidence intervals.

Table 1.9: The impact of the roaming policy on restaurant Tripadvisor rating

	Y=Average 5-month Tripadvisor rating; Jan 2015 - Dec 2018				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	0.040***	0.049***	0.087***	0.077***	-0.003
	(0.012)	(0.012)	(0.017)	(0.019)	(0.012)
Restaurant & Time FE	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	147274	146620	48577	53659	44384
Restaurants	4373	4330	1413	1499	1418
Clusters	86	86	59	70	70
Adj. R-squared	0.503	0.504	0.324	0.251	0.297
Mean Y pre-policy	3.98	3.98	3.51	4.05	4.43
DDD *p-value*				0.000	0.000

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

management responses to consumer reviews as a way to obtain more positive feedback.[54] To assess this hypothesis, I construct the dynamic count of online replies over 5-month windows and analyze whether restaurants started interacting more frequently with their reviewers in the attempt to upgrade their subsequent online reputation. Appendix Table 1.D13 shows a significant increase in these replies among restaurants in the mid-rating category, pointing out a potential reason for their reputation upgrading. Another possible explanation for their success is the growth in personnel (more employees) and improvements in working conditions (less working days per worker) discussed in Section 1.5.1, both of which could have positive effects on the service quality and, eventually, on the customer experience. Finally, a further possibility is that these restaurants started using better raw materials in their kitchens. Although the data do not allow me test this explicitly, I find that the (log) annual net purchases increased in mid-rating restaurants after the policy (Appendix Table 1.D14), potentially indicating the use of higher-quality ingredients in preparing the recipes. On the other hand, purchases remained the same in the other two rating categories.[55]

The second empirical fact that might be puzzling is the absence of a decline in the online reputation for high-rating restaurants, despite the documented loss in human capital. For them, replies to reviews, expansions in their personnel and the use of better ingredients are not plausible explanations.[56] An alternative reason, which is in line with the theoretical model, is that these establishments are very capable (efficient) at managing their factors of productions. This efficiency advantage allows them to employ less skilled workers, save in costs and increase profits without compromising output quality.

Finally, for all outcomes discussed above, Section 3.4.6 provides evidence in favor of the parallel trends assumption. Moreover, it is worth mentioning that all the above results are not driven by firm selection into exit: even larger coefficients are obtained when restricting the attention to restaurants that did not leave the market after the policy. These sensitivity checks are discussed in Section 1.7, which also shows that the

[54]For instance, Proserpio and Zervas 2017 find that after responding to reviews on Tripadvisor, hotels' ratings increase by 0.12 points, an effect that is comparable to my estimates.

[55]Net purchases reflect any expenditure in inputs other than labor. As such, the variable might also include the purchase of specialized services from online advertising and customer management companies.

[56]Appendix Table 1.D13 shows that, if anything, high-rating restaurants engaged less with reviewers, after the policy. Moreover, as discussed in section 1.5.1, these restaurants did not significantly employ more workers. Finally, as shown in Appendix Table 1.D14, annual purchases remained the same.

previous estimates are robust to alternative specifications, clustering units and different measurements.

1.6 Placebos

This section carries out a series of placebo exercises to assess the plausibility of the identifying assumptions. The first set of exercises aims to address potential concerns about the correlation between seasonality and the timing of the policy. In fact, the new regulation was effective in June 2017, which coincides with the beginning of the tourist season in Italy. It might be that the corresponding change in the volumes and composition of tourist flows drives the above estimates, invalidating their interpretation. In order to investigate this possibility, I focus on the pre-policy period (between May 2012 and May 2017) and perform a variety of regressions using placebo policy-dates coinciding with the month of June – i.e., the same of the regulation – but in the four years preceding the roaming policy. In practice, I regress all the above outcomes on the *Tourist* dummy variable used throughout the paper, interacted here with these four placebo policy-dates, separately. For each policy-date, the sample includes observations within a 24-month window around the placebo policy.[57] Moreover, to replicate the results across different rating-categories, I consider the Tripadvisor rating of the restaurant at the time of the respective placebo policy-date. Table 1.10 shows an example of the output of this procedure for the (log) number of employees. The estimated coefficients in column (1) are generally small in magnitude and not statistically significant. Moreover, their signs across different rating categories reported in columns (2-4) do not exhibit any systematic and significant pattern. This evidence discards the possibility that results on employment are driven by seasonal peaks.

Appendix Tables 1.D15 to 1.D21 replicate the same exercise for the other outcomes, namely, restaurant revenues, exit, industry composition, hiring decisions, salaries and rating. For them, the estimated coefficients on the entire sample of restaurants are generally small and insignificant. The same is true for coefficients in the three rating categories, although there are a few exceptions with some estimates resulting statistically different from zero. However, the signs of significant coefficients have consistently

[57]For revenues, which are observed annually, I consider 4-year windows around the placebo policy-years 2013, 2014 and 2015. In addition, for industry composition, I conduct the analysis at the ZIP-code level and use the number of tourist attractions in the ZIP code to proxy for exposure to tourists.

Table 1.10: Placebo policies and restaurant employment

	Y=log(monthly employees)			
	(1)	(2)	(3)	(4)
	Full sample	Low rating	Mid rating	High rating
Tourist*Post May2013	-0.007	-0.029	0.040	-0.006
	(0.014)	(0.018)	(0.050)	(0.025)
Observations	38862	16647	12921	9294
Tourist*Post May2014	0.015	0.030	0.011	0.006
	(0.020)	(0.029)	(0.017)	(0.029)
Observations	48795	19572	16664	12559
Tourist*Post May2015	-0.013	0.007	-0.024	-0.023
	(0.011)	(0.015)	(0.015)	(0.030)
Observations	58028	22394	19658	15976
Tourist*Post May2016	0.005	-0.018	0.007	0.044
	(0.015)	(0.017)	(0.019)	(0.028)
Observations	59571	21195	20219	18157

Every row/column is the output of a separate regression where each observation is a restaurant-month-year. Post May *year*=1 if date is after May of the respective *year*. Tourist restaurants are those with a measure of exposure above the median. Heteroskedasticity-robust standard errors clustered at municipality level. Each regression includes all controls and fixed effects from the main analysis. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the placebo policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

opposite directions with respect to those estimated in the main analysis. Overall, this evidence rules out the possibility that tourism seasonality could drive the main results and corroborates their interpretation as the consequence of lower information costs induced by the policy.

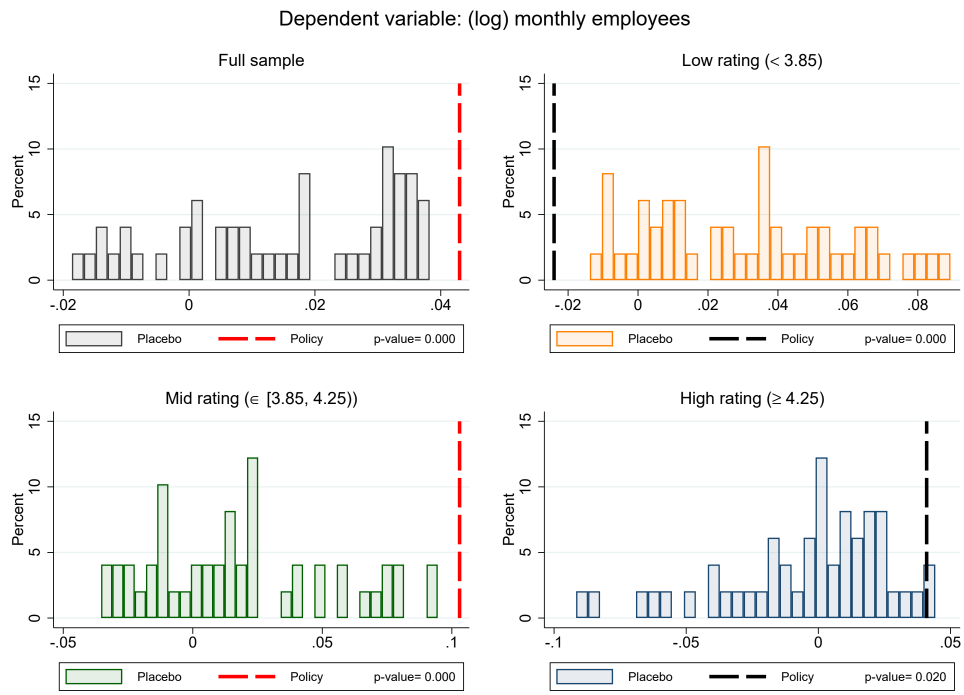
The second set of placebo exercises consists of a series of policy-permutation tests conducted in the period before the roaming regulation (Jan 2012 - Dec 2016) to assess its exogeneity with respect to other potential factors or existing pre-trends in the outcomes that might explain the observed results. In practice, I replicate the previous exercises for all placebo policy-dates between May 2012 and May 2016 – for a total of 49 regressions for each outcome –, following the approach used to carry out randomization inference in experiments (e.g., Gerber and Green 2012).[58] Then, I plot the histograms of all estimated placebo coefficients for the whole sample of restaurants, as well as for the three different rating categories considered in the main analysis.[59] For instance, Figure 1.8 reports the results for restaurant employment, where the vertical dashed lines depict the respective policy coefficient estimated in Section 1.5.1. Red (black) lines indicate

[58]In this case, rather than varying the composition of the control group, I modify the time dummy *Post*.

[59]I consider the Tripadvisor rating of the restaurant at the time of the respective placebo policy-date.

that the coefficient was significant (insignificant).

Figure 1.8: Permutation test for restaurant employment



Notes: Each panel plots the distribution of coefficients on Tourist*Post-Month, where Month is between May 2012 and May 2016, estimated on a sample of observations between Jan 2012 and Dec 2016. Tripadvisor rating and the respective category are calculated in each month. The vertical dashed lines report the actual policy coefficients, as estimated in the main analysis. The line is red when the respective coefficient is significant at least at the 10% confidence level, and black otherwise.

The comparison of the distribution of placebo coefficients with the policy point estimate would speak in favor of the identifying assumption if the policy estimates lie at the extremes of the distributions of placebo coefficients. More formally, I compute the p-value of the permutation test by counting the number of times the placebo coefficients are more extreme than the policy estimate, and dividing it by the total number of coefficients. In this context, low p-values imply that, most of the times, the policy estimates are larger in magnitude than the placebos, alleviating concerns on the endogeneity of the roaming regulation. For example, in each subplot of Figure 1.8 the p-value is always lower than the conventional 5% level. This means that the estimated 4.2% increase in overall restaurant employment and the 10% effect found in the mid-rating category can be plausibly attributed to the roaming regulation.

I summarize the results for the other outcomes in a series of figures reported in the Appendix (Figures 1.C3 to 1.C8). When the policy estimates are significant, they are always at the extremes of the distribution of placebo coefficients. The associated p-values are consistently small and always below the conventional 5% level, implying that only in very few cases the placebos are larger in magnitude than the estimated policy coefficients.

Finally, additional placebo exercises are reported in the event-study estimates. These graphs can be used to investigate the presence of potentially diverging trends in the outcomes of interest before the regulation was effective. As a result, they provide some evidence on the plausibility of the parallel-trend assumption underlying the empirical strategy of the paper. Figures 1.3, 1.5, 1.6, and 1.7 reported in the text suggest that this assumption is likely to hold for restaurant employment, industry composition, salaries and rating, respectively. Additional figures reported in the Appendix (1.C9 to 1.C12) confirm that similar conclusions generally hold for restaurant revenues, exit and hiring decisions, although in some cases these graphs are less clear because of the nature of certain outcomes (e.g., binary variables for exit and hiring decisions). Nevertheless, all together, the above placebo exercises corroborate the identifying assumptions and validate the use of the Difference-in-Differences strategy in this context.

1.7 Robustness

This section carries out a series of additional estimations to investigate the sensitivity of the main results to different measurements, samples and clustering units. First, a potential concern with the firm-level estimates is that they might be driven by both (1) the definition of the binary variable *Tourist* and (2) the construction of the underlying measure of exposure to tourist clientele $P(i)$, defined in Section 1.4. To address the first point, I replicate the baseline estimation using a more flexible specification, in which I consider dummies for quintiles of exposure to tourists – rather than the median value – interacted with the variable *Post*. This procedure generalizes the results to study the effects of the policy along a more continuous gradient of exposure and, at the same time, guarantees that estimates remain easy to interpret. Particularly, I show the results using a series of figures, reporting the effects at each quintile (with respect to the first one, which is the omitted category) as well as across the three rating categories. Panel (b) of Figure 1.4 already discussed in Section 1.5.1 presents the results for restaurant employment. Consistently with the main analysis, better-rated restaurants at the higher quintiles of the tourist-exposure distribution are the drivers of employment growth. Appendix Figures 1.C13 to 1.C18 show that similar conclusions hold for restaurant revenues, exit, hiring decisions, salaries and rating, with point estimates being consistently larger for the highest quintiles. This evidence corroborates the benchmark specification and alleviates the concern that the main results are driven by the specific definition of the binary variable $Tourist$.

Moreover, the procedure I adopted to construct the probability measure $P(i)$ could also influence the firm-level outcomes. To address this concern, I study the sensitivity of the main analysis to the use of alternative measures of tourist exposures, which rely on different data and assumptions as explained in Appendix 1.B.1. First, instead of focusing only on the shortest path, I also include all alternatives routes provided by the Google Maps API in the computation process. Second, rather than imposing equal conditional probabilities (random walk assumption), I assume tourists form educated guesses on which path to follow, based on importance (frequency) of each road. I then replicate the main estimations using newly-defined $Tourist$ dummy variables based on these two alternative measures (as in the main analysis, I use the median value to create the binary treatment indicator). Tables 1.D22 to 1.D33 in the Appendix show that in both cases estimated coefficients are always qualitatively, and often quantitatively, similar to those from the benchmark specification.

In addition, another potential concern is that sorting of restaurants into exit might bias the baseline estimates. In fact, the latter are based on a sample that includes the approximately 560 firms that, at some point after the policy, ceased their operations and left the market. Hence, I replicate the estimations on the sample of restaurants that survived throughout the whole 30-month period after the roaming regulation. Appendix Tables 1.D34 to 1.D39 show that coefficients are very similar to those from the main analysis. Certain effects – such as the hiring of experienced workers and improvements in ratings for low-rating restaurants – are even more remarked, suggesting that, if anything, the presence of exiting firms might attenuate the results.

Finally, my estimates are generally robust to different clustering units. Appendix Tables 1.D40 to 1.D46 use the 127 ZIP codes (that are smaller than municipalities) to cluster the standard errors, and show that – with the exception of restaurant revenues, which are observed at annual frequencies – the significance of the coefficients is not remarkably altered.

1.8 Economy-wide effects of Tripadvisor

The goal of this section is to recover the economy-wide effects of increasing access to information from Tripadvisor in the experience goods market. In fact, the Diff-in-Diff estimation provides only reduced-form evidence on the impact of the roaming policy on restaurants in the Province of Rome. On the one hand, the regression coefficients represent intent-to-treat effects of the provision of information, because in the treatment group (identified by the *Tourist* binary variable) only a fraction of the clientele – namely, the EU users – benefited of the cheaper internet costs. Therefore, recovering the treatment-on-the-treated effect is the first step to assess the importance of expanding the access to Tripavisor for the whole customer base. On the other hand, the benchmark estimation was conducted only on a sample of restaurants, which poses a limit to the generalization of the effects. Thus, re-weighting the estimates is needed to contextualize their magnitudes in the entire Italian restaurant industry.

The procedure requires three additional assumptions:

1. Among *Tourist* restaurants, take-up of the policy was 23%, which corresponds to the percent of Tripadvisor reviews from EU travelers in the post-policy period.

2. Non-tourist restaurants are not affected by the policy.

3. The share of tourist restaurants in Italy is 8%, which corresponds to the fraction of establishments located in ZIP-codes with at least one top-tourist attraction.

Assumption (1) relies on Tripadvisor contributions from Europeans to proxy for their usage of the platform in the post-policy period.[60] Although a gap between demand and supply of reviews plausibly exists, it is unlikely that it depends on the origin of the reviewers. Hence, the ratio EU/Total contributions should provide a reasonable approximation for the relative usage among Europeans. Moreover, assumption (2) requires that outcomes in the control group (i.e., non-tourist restaurants) do not change after the policy. This assumption is likely to hold for at least two reasons. First, estimates of the policy by deciles/quintiles of exposure to tourists show that effects are generally driven by restaurants at higher levels of exposure. Second, as shown in Appendix Figure 1.C19, de-trended average employment in non-tourist restaurants remains stable after the roaming regulation. Regarding assumption (3), I collect information from

[60]Disaggregated Tripadvisor usage statistics are not available.

Tripadvisor on the top-100 tourist attractions in Italy (based on their total volume of reviews) and then compute the fraction of Italian restaurants that are located in the ZIP-codes with at least one of such attractions.

Finally, I consider the most conservative estimates of the policy on revenues, employment and exit as reported in columns (1-2) of Tables 1.2, 1.3 and 1.4. I divide these coefficients by 0.23 and multiply them by 0.08, and compare them with the aggregate trends in the industry. Results are reported in Table 1.11, and suggest that promoting access to review platforms has relevant economy-wide consequences on the whole Italian restaurant industry. Back-of-the-envelop calculations point out that reducing the costs for consumers to access Tripadvisor leads to an overall increase in restaurant revenues, employment and exit rate by 1.6%, 1.5% and 0.5 pp, respectively. The first two figures correspond to about 12% and 5% of the overall growth in revenue and employment experienced by restaurants between 2016 and 2019, respectively. While the last figure corresponds to almost 3% of the exit rate faced by the industry during the first year of the Covid-19 pandemic. All together, these results indicate that lower consumer information costs due to review platforms can have real effects on the performance and composition of firms operating in industries generally affected by asymmetric information.

Table 1.11: Economy-wide effects of Tripadvisor

	Adjusted effect of Tripadvisor	2016-2019 growth rate	Percent of growth explained by Tripadvisor
Annual revenues	1.6%	13.2%	**12.1%**
Monthly employment	1.5%	29.7%	**5.1%**
	Adjusted effect of Tripadvisor	Exit rate in Covid year 2019-2020	Percent of exit rate explained by Tripadvisor
Annual exit rate	0.46 pp	15.7%	**2.9%**

Notes: Revenue growth rate refers to the 2016-2018 period.

1.9 Conclusions

The digital era has changed the way consumers get and share information. Yet, it is not (fully) clear what are consequences of this phenomenon for markets with information asymmetries, such as the service sector. While there is general optimism around the possibility for online review platforms to create reputation and feedback mechanisms that attenuate adverse selection and moral hazard on the producer side, empirical evidence is scarce. This paper shows that lower information frictions for consumers - caused by an exogenous abolition of internet tariffs - have the potential to change how firms in these markets operate and make the service industry more quality oriented.

First, I built a model in which consumers with heterogeneous search costs engage in sequential search to buy one unit of a vertically differentiated product, while firms with heterogeneous abilities endogenously select into production and compete in quality. The model predicts that lower search costs positively affect the equilibrium quality levels but have differing effects across businesses. In fact, some of the lowest-quality firms exit the market while the surviving ones increase their effort to upgrade product quality.

To test these hypotheses, I focused on the restaurant industry in the province of Rome and assembled a unique dataset which combines restaurants' information from Tripadvisor with rich administrative establishment-level data. I took advantage of a plausibly exogenous reduction in the costs of mobile internet - caused by the abolition of roaming charges for tourists in the European Union - to identify the effects of lower search costs on consumers' behavior, restaurants' incentives to upgrade their quality, as well as changes in the industry composition.

Using a Difference-in-Differences strategy, I compared the variation (before/after policy) in the outcomes across restaurants that are differentially exposed to tourist clientele. I estimated the model on the whole sample of restaurants with available Tripadvisor rating at the time of the policy, as well as on three sub-samples containing restaurants with different ratings: namely, low, mid and high rating.

I showed that, after the policy, revenues increased in mid- and high-rating restaurants, while employment grew only in the mid-rating category, suggesting that high-rating establishments were already producing at full capacity. I then analyzed the supply side. First, I showed that for low-rating restaurants, the probability to exit the market double after the policy compared to the pre-policy period. Moreover, by aggregating observations at the ZIP-code level, I found that the share of low-rating firms operating in the most touristy neighborhoods decreased by 2.5 pp after the policy, compared to non-touristy ZIP codes. Then, I analyzed the behavior of surviving firms. In particular, I showed that low-rating restaurants focused their recruiting efforts on workers with previous experience in the restaurant industry and ended-up paying higher salaries. Eventually, low- and mid-rating improved their online reputation, as their dynamic Tripadvisor rating increased after the policy.

All together, my findings indicate that lower information costs for consumers create the conditions that push some low-quality providers out of the market and encourage others to produce higher-quality goods. These results offer an optimistic assessment of the possibilities of quality-upgrading in the restaurant industry through policies that reduce information costs for consumers and facilitate the use of review platforms. More generally, these results imply that ICT policies can improve welfare by reducing asymmetric information problems.

1.10 Appendices

1.A Model

Derivation of equation (1.4)*:* Let $g(r)$ be the pdf of the reservation quality. Then, using equation (1.2), the corresponding cdf can be expressed as:

$$G(r) = 1 - Z[\rho^{-1}(r)] = 1 - Z\left[\int_r^\infty [1 - F(q)]\, dq\right]$$

Taking the derivative of $G(r)$ with respect to r yields:

$$G'(r) = g(r) = -z\left[\int_r^\infty [1 - F(q)]\, dq\right][F(r) - 1] = z[\rho^{-1}(r)]\,[1 - F(r)]$$

Finally, replacing $g(r)$ into equation (1.3) yields equation (1.4).

Proof of Property 1: Applying the Implicit Function Theorem to the first-order condition (1.6) yields:

$$\begin{aligned}\frac{\partial q(\lambda)}{\partial \lambda} &= -\frac{x''_{qq}[q(\lambda)]\, q'_\lambda(\lambda) - C''_{q\lambda}[q(\lambda), \lambda] - C''_{qq}[q(\lambda), \lambda]\, q'_\lambda(\lambda)}{x''_{qq}[q(\lambda)] - C''_{qq}[q(\lambda), \lambda]} \\ &= -q'_\lambda(\lambda) + \frac{C''_{q\lambda}[q(\lambda), \lambda]}{x''_{qq}[q(\lambda)] - C''_{qq}[q(\lambda), \lambda]}\end{aligned}$$

$$\iff q'_\lambda(\lambda) = \frac{1}{2}\, \frac{C''_{q\lambda}[q(\lambda), \lambda]}{x''_{qq}[q(\lambda)] - C''_{qq}[q(\lambda), \lambda]} > 0$$

The latter inequality holds because the numerator is negative by assumption, while the denominator is negative by the second-order condition (1.7). ■

Proof of Property 2:

$$\frac{\partial x[q(\lambda)]}{\partial \lambda} = x'_q[q(\lambda)]\; q'_\lambda(\lambda) > 0$$

The latter inequality holds because $x'_q > 0$ as the demand function is upward sloping in quality, and $q'_\lambda > 0$ by Property 1. ■

Proof of Property 3: Applying the Envelope Theorem to the profit function (1.5) yields:

$$\frac{\partial \pi[q(\lambda), \lambda]}{\partial \lambda} = -C'_\lambda(q, \lambda) > 0$$

The latter inequality holds because of the assumption on the cost function. ■

Proof of Lemma 1: Replacing the equilibrium profit schedule (1.17) into the entry condition (1.9) yields the following identity

$$\Theta(\rho, L; a) \equiv \int_{\underline{\lambda}(a)}^{\infty} \left[\frac{1}{\lambda} - 2\sqrt{\frac{1}{aL\lambda}} + \frac{1}{aL}(1-\rho) \right] \gamma(\lambda)\, d\lambda - \kappa = 0, \tag{1.21}$$

where both ρ and L are functions of a, and $\underline{\lambda}(a) = aL(1-\rho)^{-2}$. Implicitly differentiating the identity with respect to a yields

$$\Theta'_a + \Theta'_\rho \frac{\partial \rho(a)}{\partial a} + \Theta'_L \frac{\partial L(a)}{\partial a} = 0. \tag{1.22}$$

These partial derivatives are

$$\begin{aligned}
\Theta'_a &= \int_{\underline{\lambda}(a)}^{\infty} \frac{1}{a^2} \left[\sqrt{\frac{a}{L\lambda}} - \frac{(1-\rho)}{L} \right] \gamma(\lambda)\, d\lambda < 0; \\
\Theta'_\rho &= \int_{\underline{\lambda}(a)}^{\infty} -\frac{1}{aL} \gamma(\lambda)\, d\lambda < 0; \\
\Theta'_L &= \int_{\underline{\lambda}(a)}^{\infty} \frac{1}{L^2} \left[\sqrt{\frac{L}{a\lambda}} - \frac{(1-\rho)}{a} \right] \gamma(\lambda)\, d\lambda = \frac{a}{L} \Theta'_a < 0.
\end{aligned}$$

The latter equality together with equation (1.22) yields

$$\frac{\partial \rho}{\partial a} = -\frac{1}{\Theta'_\rho} \left[\Theta'_a + \frac{\partial L}{\partial a} \Theta'_L \right] = -\frac{\Theta'_a}{\Theta'_\rho} \left[1 + \frac{a}{L} \frac{\partial L}{\partial a} \right]. \tag{1.23}$$

Since $\rho'_a < 0$ from equation (1.2), $\Theta'_a < 0$ and $\Theta'_\rho < 0$, then (1.23) implies that

$$1 + \frac{a}{L}\frac{\partial L}{\partial a} > 0 \iff a\frac{\partial L}{\partial a} + L > 0 \implies \delta'_a > 0. \ \blacksquare$$

Proof of Lemma 2: Taking the derivative of the equilibrium profit function (1.17) with respect to a yields

$$\pi'_a(\lambda; a) = \frac{1}{\delta(a)}\left[\frac{\delta'_a(a)}{\sqrt{\delta(a)}}\frac{1}{\sqrt{\lambda}} - \frac{(1-\rho)}{\delta(a)} - \frac{\partial \rho(a)}{\partial a}\right].$$

Thus, the sign of $\pi'_a(\lambda; a)$ depends on the sign of the term in brackets. Since $\delta'_a > 0$ by Lemma 1, this term is decreasing in λ. This implies that, if the term is negative for $\lambda_0 \geq \underline{\lambda}$, then it will also be negative $\forall\, \lambda > \lambda_0$. $\blacksquare$

Proof of Proposition 1: Lemma 1 implies that

$$q'_a(\lambda; a) = -\left[\frac{\delta'_a(a)}{2\sqrt{\delta(a)\lambda}}\right] < 0. \tag{1.24}$$

Moreover, taking the derivative of (1.24) with respect to λ, yields

$$q''_{a\lambda}(\lambda; a) = \frac{1}{4}\frac{\delta'_a(a)}{\sqrt{\delta(a)}}\lambda^{-3/2} > 0.$$

That is, the negative change in quality predicted by (1.24) becomes smaller (i.e., closer to zero) for larger values of λ. $\blacksquare$

Proof of Proposition 2: Lemma 1 implies that

$$C'_a(\lambda; a) = -\frac{1}{2\sqrt{\lambda}}\delta(a)^{-3/2}\delta'_a(a) < 0. \tag{1.25}$$

Moreover, taking the derivative of (1.25) with respect to λ, yields

$$C''_{a\lambda}(\lambda;a) = \frac{1}{4}[\lambda\delta(a)]^{-3/2}\delta'_a(a) > 0.$$

That is, the negative change in costs predicted by (1.25) becomes smaller (i.e., closer to zero) for larger values of λ. ■

Proof of Proposition 3: Consider the entry condition

$$\int_{\underline{\lambda}(a)}^{\infty} \pi(\lambda;a)\,\gamma(\lambda)\,d\lambda = \kappa.$$

Differentiating this with respect to a, yields

$$\int_{\underline{\lambda}(a)}^{\infty} \pi'_a(\lambda;a)\,\gamma(\lambda)\,d\lambda = 0. \tag{1.26}$$

Together with Lemma 2, this implies that $\pi'_a[\underline{\lambda}(a);a] > 0$, otherwise the integrand in (1.26) would be negative $\forall\,\lambda > \lambda$ (by Lemma 2), which would contradict (1.26). To see how the ability threshold $\underline{\lambda}(a)$ changes as search costs decreases, consider a shift in a: a_1 to $a_2 < a_1$. Then

$$\pi[\underline{\lambda}(a_2), a_2] = 0 = \pi[\underline{\lambda}(a_1), a_1] > \pi[\underline{\lambda}(a_1), a_2],$$

where the two equalities follow from the definition of $\underline{\lambda}$, and the inequality follows from $\pi'_a[\underline{\lambda}(a);a] > 0$. Since (by Property 3) $\pi'_\lambda > 0\ \forall\,\lambda$, it follows that $\underline{\lambda}(a_2) > \underline{\lambda}(a_1)$. ■

Proof of Corollary 1: Taking the derivative of $x(\lambda;a)$ with respect to a, yields

$$x'_a(\lambda;a) = \frac{1}{\delta(a)}\left[\frac{\delta'_a\,(\rho-1)}{\delta(a)} + \frac{\delta'_a}{2\sqrt{\lambda\delta(a)}} - \rho'_a\right]. \tag{1.27}$$

The sign of (1.27) equals the sign of the expression in brackets. In particular, the expression is negative for sufficiently high values of λ, that is

$$x'_a(\lambda;a) < 0 \iff \lambda > \frac{\delta(a)}{4\left[1-\rho+\frac{\delta(a)\rho'_a}{\delta'_a}\right]^2}.$$

Hence, there exists a $\hat{\lambda} > \underline{\lambda}$ such that $x'_a(\hat{\lambda};a) < 0$. Since from (1.27) it is clear that x'_a is decreasing in λ (as $\delta'_a > 0$ by Lemma 1), this implies that $x'_a(\lambda;a) < 0 \; \forall \lambda > \hat{\lambda}$. ■

1.B Measurements

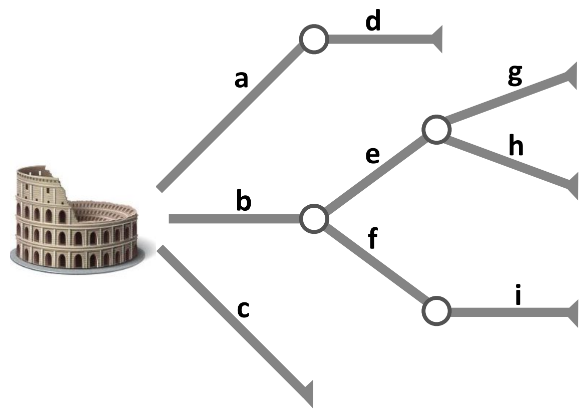
1.B.1 Defining exposure to tourists

I assume that search is sequential over space and bounded by the structure of the road network around a tourist site. In particular, the probability of coming across a restaurant is equal to the probability of ending up on the street where the restaurant is, taking into account the previous path. Hence, two or more restaurants located on the same street have the same probability of being found, but this probability depends on the path to their location. Tourists start inspecting high-visible places around them - such as those in front of a tourist attraction - and then move to other less visible places until the marginal expected cost (time and fatigue of walking) becomes larger than the marginal expected benefit of finding a good deal.

In practice, I use information from Google Maps and construct the partial road network that leads to the Tripadvisor restaurants around each attraction, and eventually compute the probabilities to find them while walking away from the attraction. The procedure works as follows. First, for each restaurant i, I consider its closest (shortest distance) tourist attraction t. Then, for each identified pair (t, i), I use the Google Maps API to find the directions (street names) of all the paths that lead from t to i on foot. In case more than one path is suggested by Google, I consider the shortest distance path to build the benchmark measure, while I also provide robustness to the use of all alternative paths. Then, I construct the partial road network around attraction t using the street names from Google Maps and, for each i, I compute the conditional probabilities of being in any of the roads that form the path from t to i (taking into account competing roads). Finally, I multiply them and compute the probability of finding i. In the process, I assume that all roads are equally weighted. Figure 1.B1 shows a simplified example of this calculation.

Figure 1.B1: Example of partial road network

Segments represent roads, circles represent junctions



Assuming equal conditional probabilities:

$p(a) = p(b) = p(c) = 1/3$

$p(e|b) = p(f|b) = p(g|e) = p(h|e) = 1/2 \Rightarrow p(e) = p(f) = 1/6$ and $p(g) = p(h) = 1/12$

$p(d|a) = p(i|f) = 1 \Rightarrow p(d) = 1/3$ and $p(i) = 1/6$

More formally, the probability that a tourist moving away from attraction t comes across restaurant i is equal to the joint probability of traveling the path defined by a vector of streets $(s_1, ..., s_{Ni})$ connecting t to i. Hence,

$$P(i) = P(s_1 \,\&\, s_2 \,\&\, ... \,\&\, s_{Ni}) = P(s_1|t) \prod_{j=2}^{N_i} P(s_j|s_{j-1}) \tag{1.28}$$

This probability measure reflects the chances that a restaurant is visited by a tourist and thereby the extent to which it is exposed to the policy. In particular, this measure not only reflects the "visibility" of restaurant i from attraction t, but also the effect of proximity of i to t. In fact, an increasing number of streets compete in the road network as the radius enlarges by moving farther away from the attraction, and this naturally drives down the estimated probability. Hence, any observed differential impact of the policy along the probability measure could be in part explained by factors - other than the presence of tourists - that correlate with proximity and affect restaurants' decisions.[61] For this reason, in the empirical analysis, I always control for the distance to the attraction and the distance to Rome city center.

I use equation 1.28 and compute this probability for all restaurants in the Tripadvisor sample.[62] The empirical distribution of the probability measure is right-skewed. About 50% of the restaurants have roughly a 0 probability of being found by the tourists (specifically, $P(i) < 0.17\%$, which is the median). Most likely, distance to the attraction explains the fact. Because these establishments are usually located too far from tourist paths, their chance of being affected by the policy is very limited: even if tourists were aware of their existence, they would not visit them as the cost to reach them is too high. For this reason, I consider them as a control group in the baseline specification. Among the remaining 50%, the top 10% most visible restaurants have a probability larger than 12.5%, meaning that about 40% of the restaurants are left with a probability between 0.17% and 12.5%.

Finally, I also use an alternative approach to measure tourist exposure that lifts the assumption of equal conditional probabilities and instead allows for weighting roads based on their "importance". Particularly, weights reflect the number of times that a road appears on every (sub)path over the total number of (sub)paths. For instance, considering the example reported on figure 1.B1, I get:

$p(a) = 2/9, p(b) = 6/9, p(c) = 1/9$

$p(e|b) = 3/5, p(f|b) = 2/5$

$p(g|e) = p(h|e) = 1/2$

$p(d|a) = p(i|f) = 1$

1.B.2 Exposure to tourists, clientele and rating

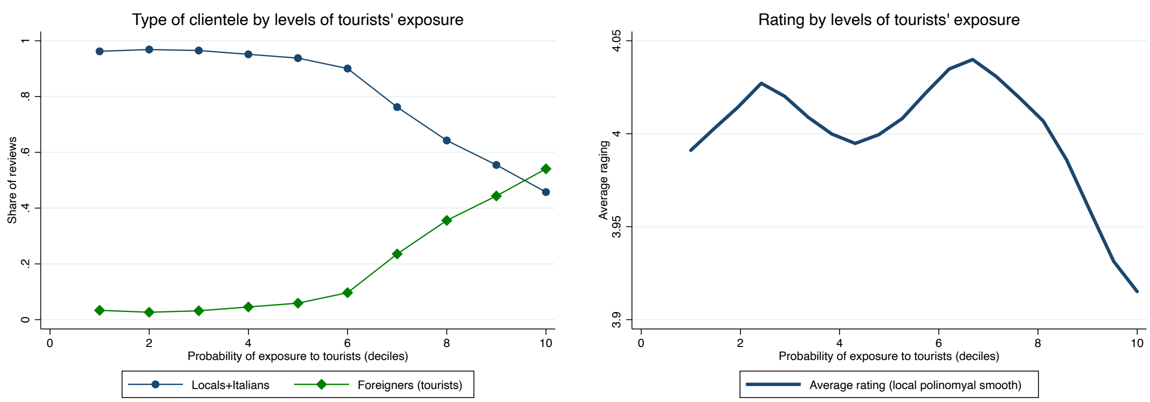
Using the probability measure defined above, I examine how restaurants' type of clientele and ratings vary with the restaurant potential exposure to tourist demand. This exercise helps me to both validate the constructed measure as well as provide a description

[61] Examples include rent costs that tend to be higher closer to attraction sites, or congestion - restaurants near touristic locations can easily be overcrowded, thus leading to longer waiting times, more noise, and worse service.

[62] Note that this measure could not be computed for the entire universe of restaurants in the Social Security database, as their exact location remains unknown to the researcher for confidentiality reasons.

of the restaurant's industry. The left panel of figure 1.B2 exploits the origin of the reviewers to distinguish them between locals and tourists. I identify as foreign tourists all those reviewers writing in a language different than Italian. This means that the green line in the graph is probably under-reporting the share of total tourists, as Italian travelers are not accounted.[63] However, since the roaming policy did not affect Italians directly, excluding them from this group provides a more conservative picture of the potential effect of the policy across different levels of exposure to tourists. Particularly, for restaurants whose probability is below or equal to the median, the share of foreign clientele remains quite stable and below 10%. By contrast, this increases rapidly afterwards, and reaches almost 60% for restaurants at the top probability-decile.

Figure 1.B2: Type of clientele and rating by exposure to tourists



The right panel of figure 1.B2 sows how Tripadvisor rating varies across levels of exposure to tourists. For each decile of probability, it reports the mean of all restaurants' average rating at the time of the policy. Restaurants more exposed to tourist demand have, on average, poorer Tripadvisor ratings, in the order of 0.10-0.15 on a scale from 1 to 5. In line with the theory, restaurants that rely more on repeated and informed clientele sell higher quality meals.[64] The explanation is two-fold. First, locals are more likely to be informed. Restaurants located in areas where the share of informed consumers is higher

[63]The main problem here is that, among Italians, I can identify the "locals" only for a subset of reviewers who explicitly indicate their town in their profile. The rest of Italians can either be tourists or locals, yet it is impossible to distinguish them.

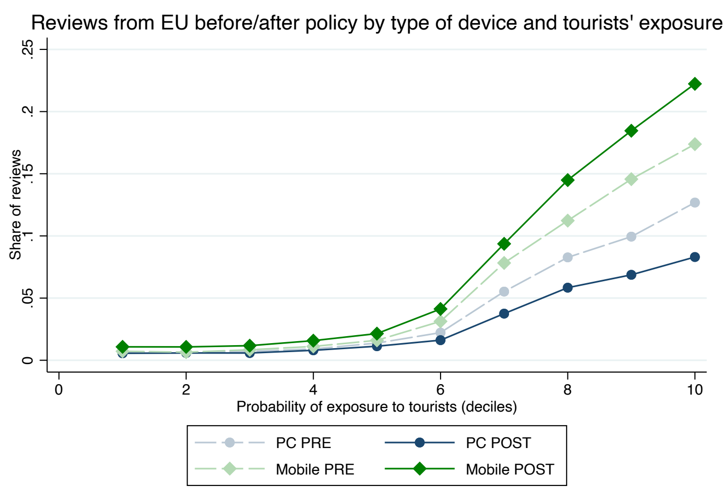
[64]These results are in line with those of Dall'orso et al. (2016), who provide evidence on the existence of quality differential across more and less visible restaurants using data from Yelp on 10 large cities in Europe and North America. In particular, they show that restaurants with higher visibility from tourists - i.e. those located at street intersections - consistently exhibit a lower Yelp rating.

(e.g. those with a lower probability measure), have incentive to provide a better quality product/service to stay in the market (Cooper and Ross, 1984). Second, locals exert control over quality through repeated purchases. Then, quality provision becomes a way to establish reputation in the market (Riordan, 1986).

1.B.3 The roaming policy and exposure to tourists

Does the roaming policy affect the composition of demand across restaurants based on their level of exposure to tourists? To answer this question, in figure 1.B3, I plot the share of reviews from Europeans by different type of device across deciles of probability, before and after the policy. The graph shows that, after the policy, the share of mobile (PC) reviews from Europeans increased (decreased). However, all the largest change took place in restaurants with higher exposure to tourists, namely, those with a probability above the median, while virtually nothing happened for other restaurants.

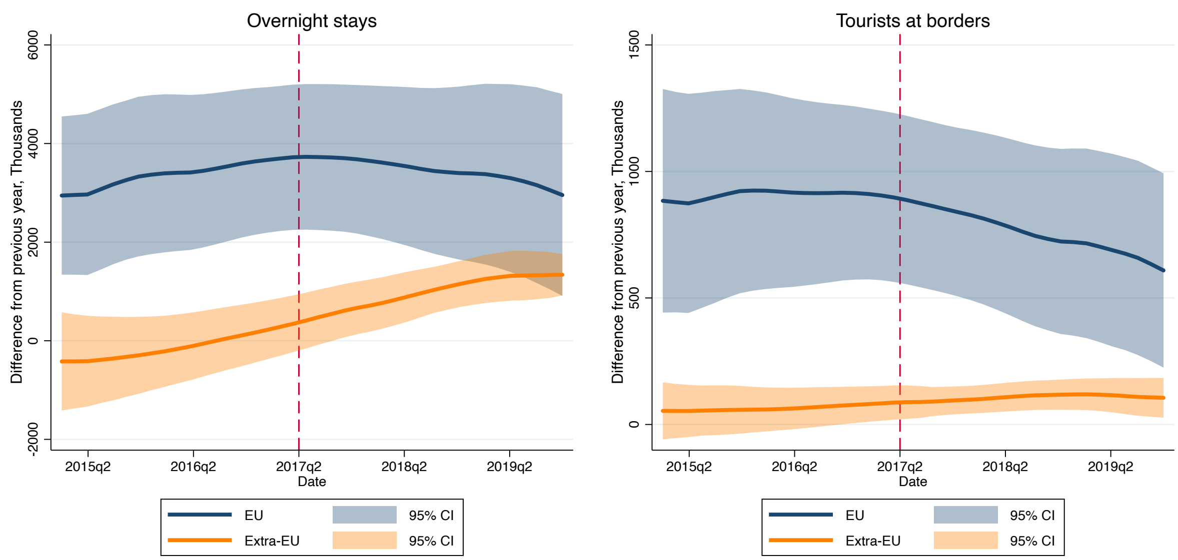
Figure 1.B3: Roaming and reviews from Europeans



PRE-policy period goes from Jan 2015 to May 2017. POST-policy period goes from June 2017 to December 2018.

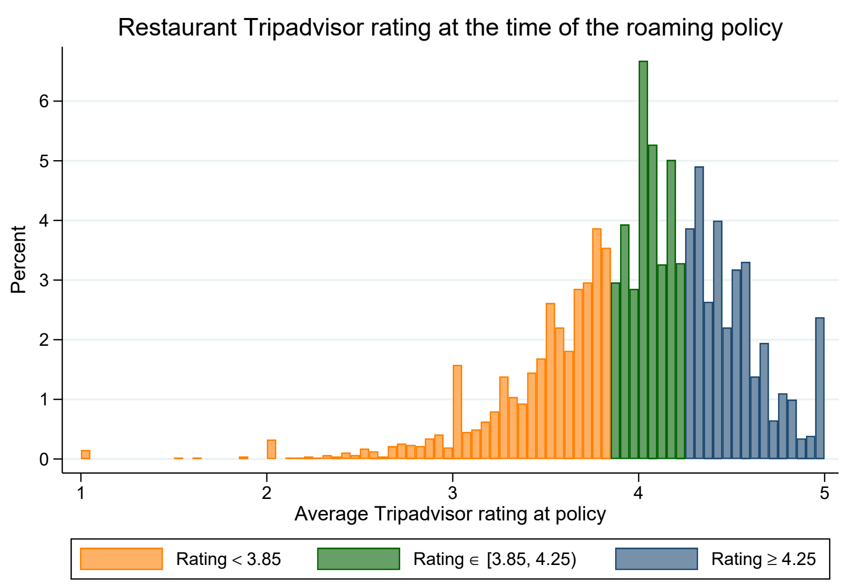
1.C Figures

Figure 1.C1: International travelers to Italy (Δ from previous year, Thousands)



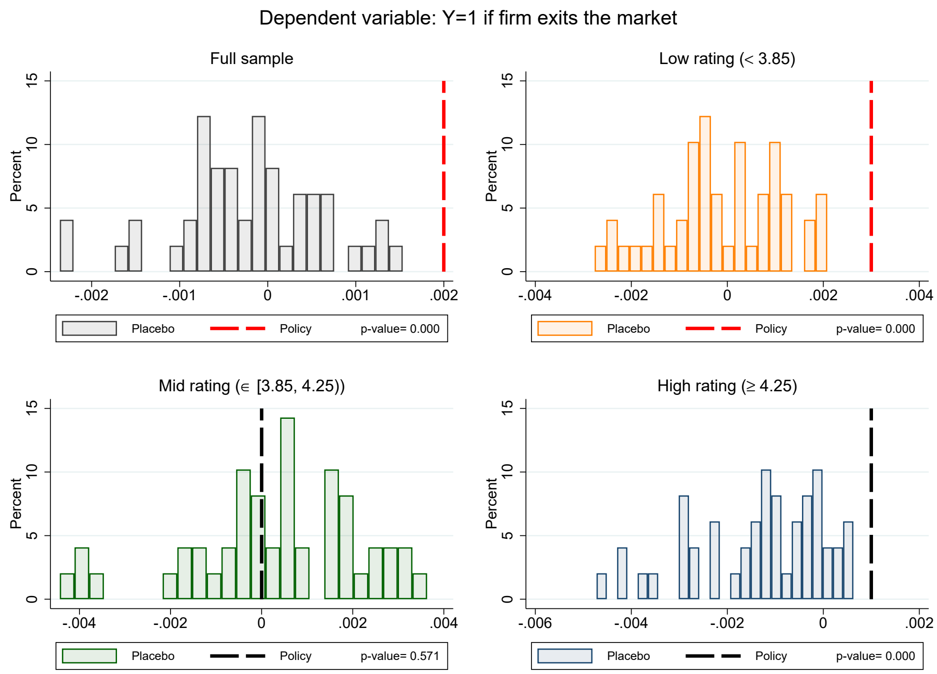
The data report international tourists traveling to Italy from different regions/countries of the world, in Thousdands. The lines depict local polynomial fits of quarterly observations reporting the difference from the quarter of the previous year. EU refer to tourists from a EU country, and Extra-EU refer to all other tourists. Source: Bank of Italy.

Figure 1.C2: Distribution of Tripadvisor average rating



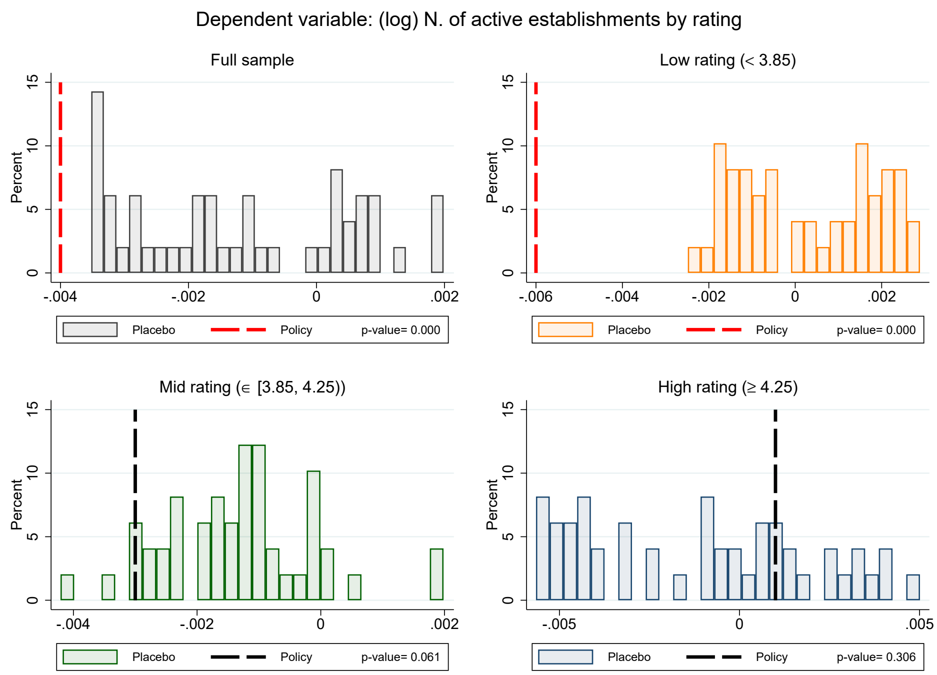
The figure shows the histogram of Tripadvisor average ratings of restaurants at the time of the policy, for the 4,628 matched restaurants with available information. Different colors split the overall sample in subgroups based on tertiles of ratings.

Figure 1.C3: Permutation test for restaurant exit



Notes: Each panel plots the distribution of coefficients on Tourist*Post-Month, where Month is between May 2012 and May 2016, estimated on a sample of observations between Jan 2012 and Dec 2016. Tripadvisor rating and the respective category are calculated in each month. The vertical dashed lines report the actual policy coefficients, as estimated in the main analysis. The line is red when the respective coefficient is significant at least at the 10% confidence level, and black otherwise.

Figure 1.C4: Permutation test for industry composition

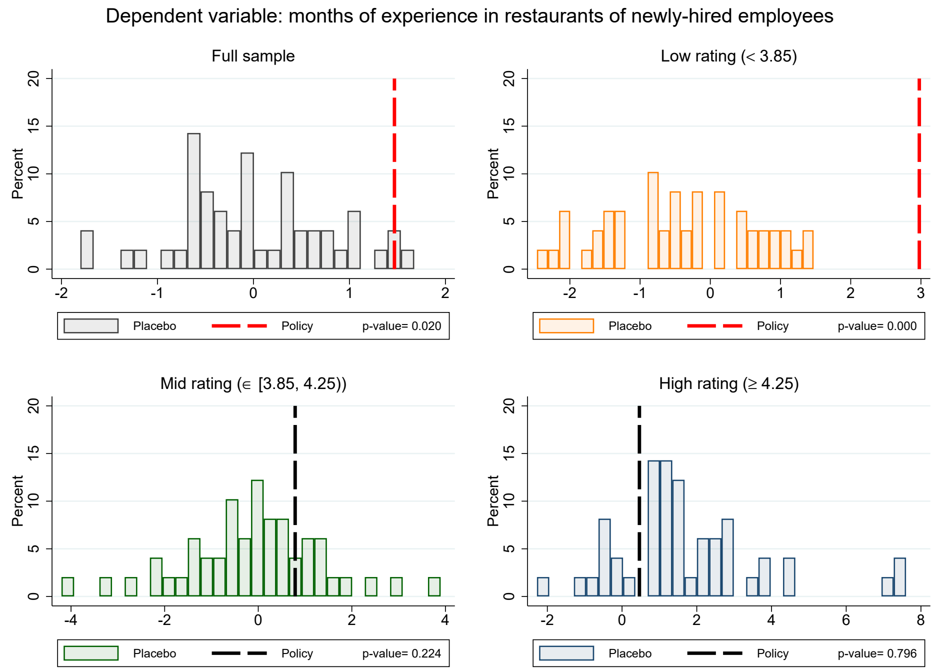


Notes: Each panel plots the distribution of coefficients on Tourist*Post-Month, where Month is between May 2012 and May 2016, estimated on a sample of observations between Jan 2012 and Dec 2016. Tripadvisor rating and the respective category are calculated in each month. The vertical dashed lines report the actual policy coefficients, as estimated in the main analysis. The line is red when the respective coefficient is significant at least at the 10% confidence level, and black otherwise.

Figure 1.C5: Permutation test for restaurant hiring decisions (extensive margins)

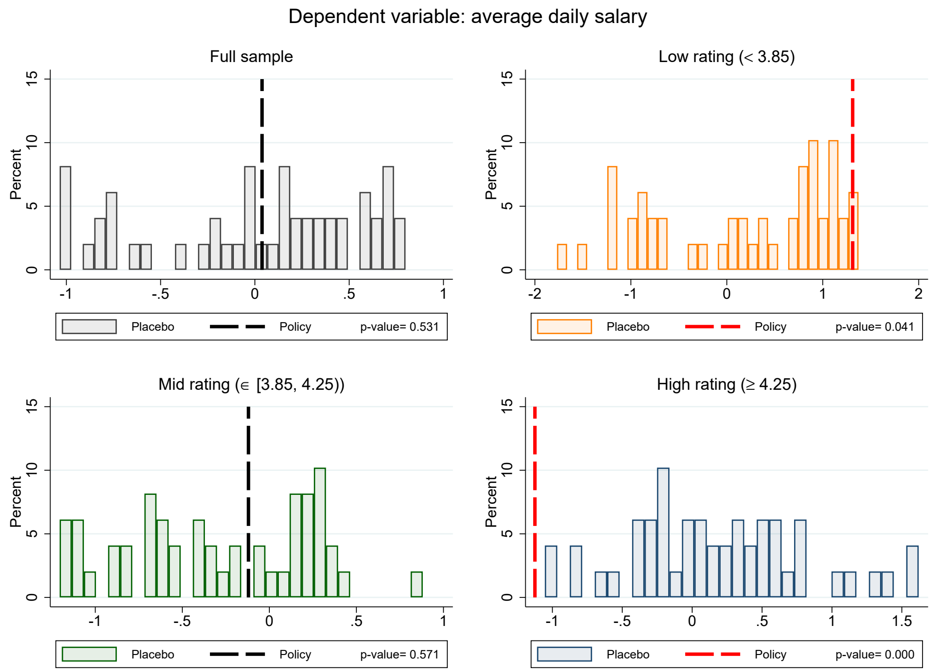
Notes: Each panel plots the distribution of coefficients on Tourist*Post-Month, where Month is between May 2012 and May 2016, estimated on a sample of observations between Jan 2012 and Dec 2016. Tripadvisor rating and the respective category are calculated in each month. The vertical dashed lines report the actual policy coefficients, as estimated in the main analysis. The line is red when the respective coefficient is significant at least at the 10% confidence level, and black otherwise.

Figure 1.C6: Permutation test for restaurant hiring decisions (intensive margins)



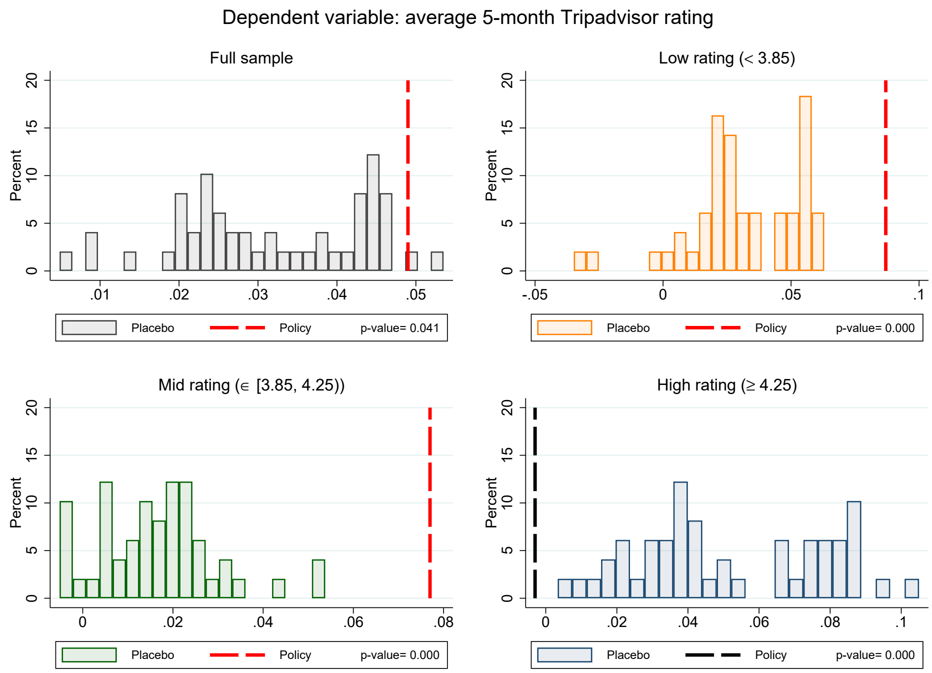
Notes: Each panel plots the distribution of coefficients on Tourist*Post-Month, where Month is between May 2012 and May 2016, estimated on a sample of observations between Jan 2012 and Dec 2016. Tripadvisor rating and the respective category are calculated in each month. The vertical dashed lines report the actual policy coefficients, as estimated in the main analysis. The line is red when the respective coefficient is significant at least at the 10% confidence level, and black otherwise.

Figure 1.C7: Permutation test for restaurant daily salaries



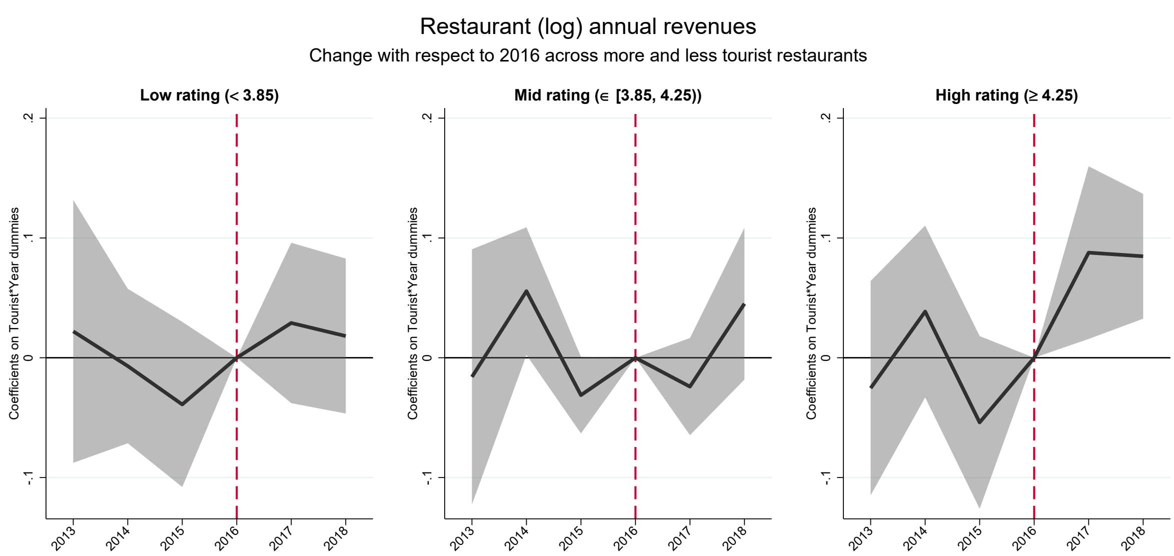
Notes: Each panel plots the distribution of coefficients on Tourist*Post-Month, where Month is between May 2012 and May 2016, estimated on a sample of observations between Jan 2012 and Dec 2016. Tripadvisor rating and the respective category are calculated in each month. The vertical dashed lines report the actual policy coefficients, as estimated in the main analysis. The line is red when the respective coefficient is significant at least at the 10% confidence level, and black otherwise.

Figure 1.C8: Permutation test for restaurant Tripadvisor rating



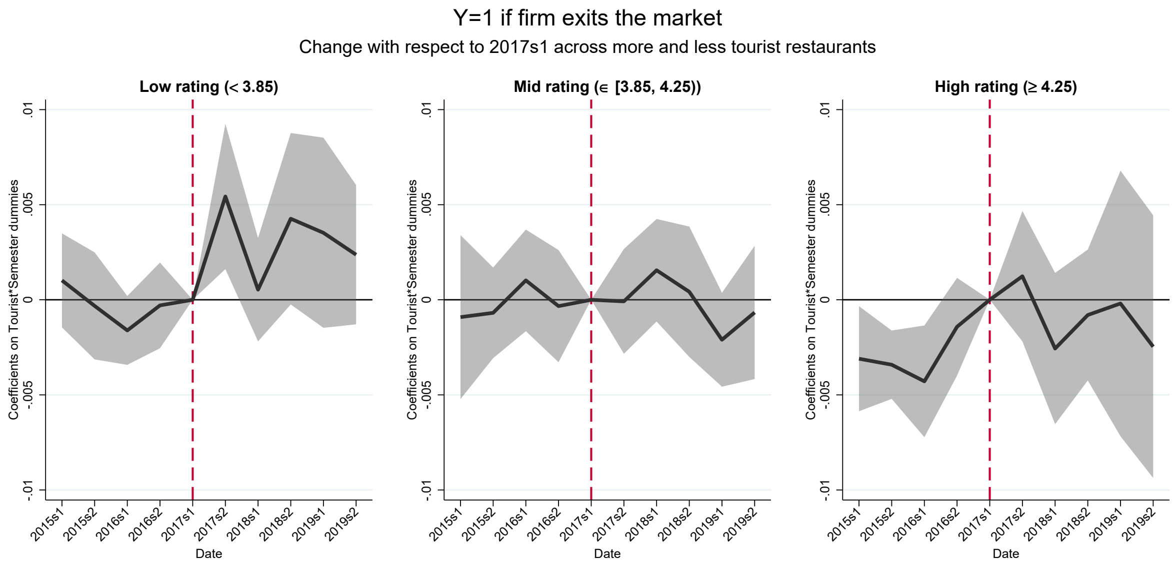
Notes: Each panel plots the distribution of coefficients on Tourist*Post-Month, where Month is between May 2012 and May 2016, estimated on a sample of observations between Jan 2012 and Dec 2016. Tripadvisor rating and the respective category are calculated in each month. The vertical dashed lines report the actual policy coefficients, as estimated in the main analysis. The line is red when the respective coefficient is significant at least at the 10% confidence level, and black otherwise.

Figure 1.C9: Event-study estimates for restaurant revenues



Notes: The graph reports estimated coefficients on the interactions of Tourist*Year dummies from three separate regressions where each observation is a restaurant-year. All controls and fixed-effects from the main analysis are included. The omitted year is 2016. Tourist restaurants are those with a measure of exposure above the median. The sample includes observations between 2013 and 2018. Shaded areas depict 95% confidence intervals.

Figure 1.C10: Event-study estimates for restaurant exit



Notes: The graph reports estimated coefficients on the interactions of Tourist*Semester dummies from three separate regressions where each observation is a restaurant-month-year. All controls and fixed-effects from the main analysis are included. The omitted semester is 2017s1. Tourist restaurants are those with a measure of exposure above the median. The sample includes observations between Jan 2015 and Dec 2018. Shaded areas depict 95% confidence intervals.

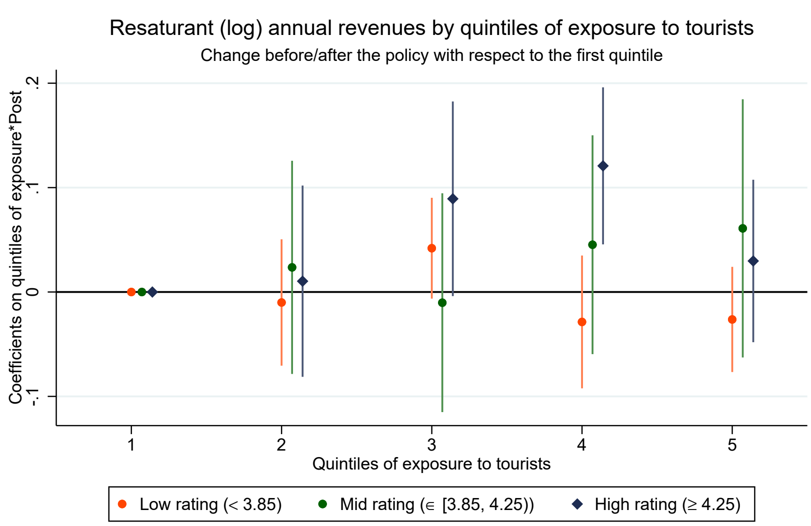
Figure 1.C11: Event-study estimates for hiring decisions (extensive margins)

Notes: The graph reports estimated coefficients on the interactions of Tourist*Semester dummies from three separate regressions where each observation is a restaurant-month-year. All controls and fixed-effects from the main analysis are included. The omitted semester is 2017s1. Tourist restaurants are those with a measure of exposure above the median. The sample includes observations between Jan 2015 and Dec 2018. Shaded areas depict 95% confidence intervals.

Figure 1.C12: Event-study estimates for hiring decisions (intensive margins)

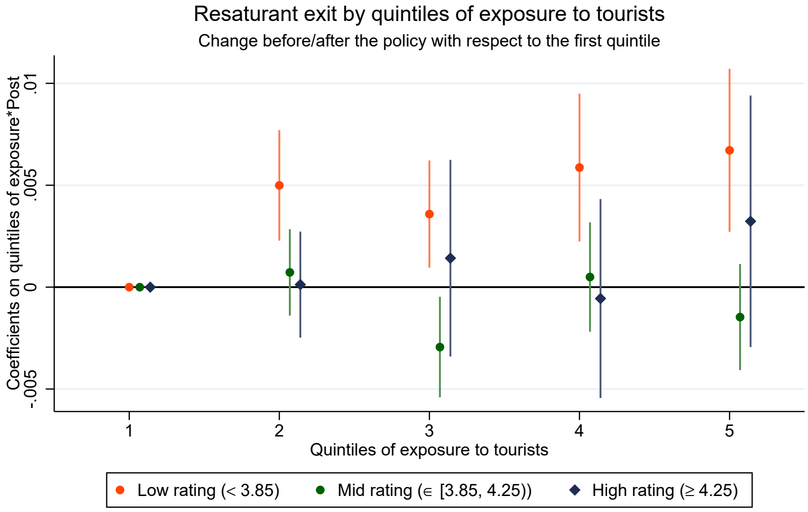
Notes: The graph reports estimated coefficients on the interactions of Tourist*Semester dummies from three separate regressions where each observation is a restaurant-month-year. All controls and fixed-effects from the main analysis are included. The omitted semester is 2017s1. Tourist restaurants are those with a measure of exposure above the median. The sample includes observations between Jan 2015 and Dec 2018. Shaded areas depict 95% confidence intervals.

Figure 1.C13: The impact on restaurant revenues across quintiles of exposure



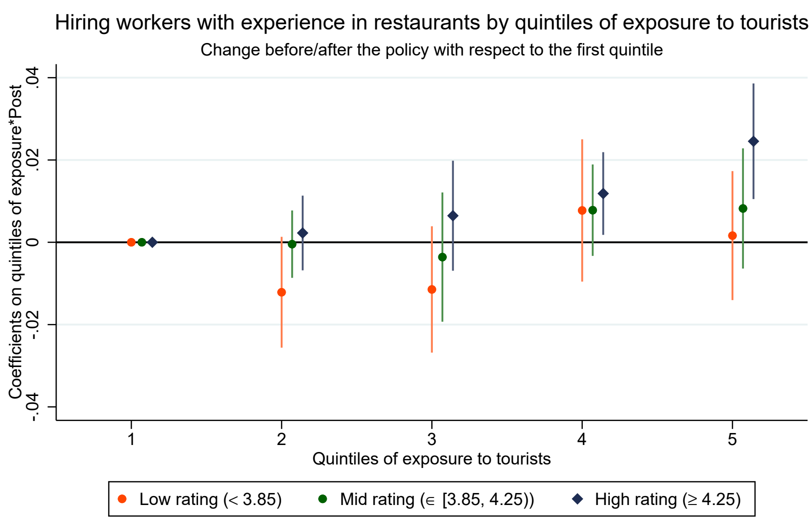
Notes: The graph reports estimates on the interactions of quintiles of exposure*Post from three separate regressions (low, mid, high-rating restaurants), where each observation is a restaurant-year. The first quintile is omitted. All controls and fixed-effects from the main analysis are included. Post takes value 1 after 2016. The sample includes observations between 2015 and 2018. Heteroskedasticity-robust standard errors are clustered at municipality level. Bars depict 95% confidence intervals.

Figure 1.C14: The impact on restaurant exit across quintiles of exposure



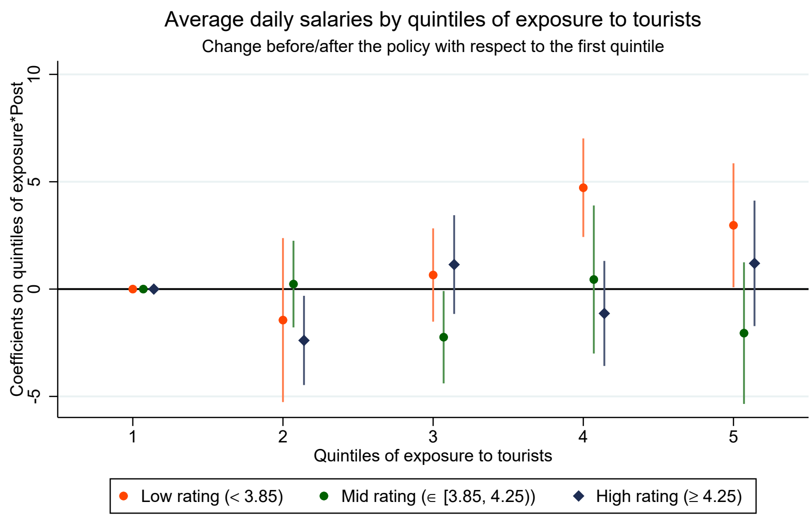
Notes: The graph reports estimates on the interactions of quintiles of exposure*Post from three separate regressions (low, mid, high-rating restaurants), where each observation is a restaurant-month-year. The first quintile is omitted. All controls and fixed-effects from the main analysis are included. Post takes value 1 after May 2017. The sample includes observations between Jan 2015 and Dec 2019. Heteroskedasticity-robust standard errors are clustered at municipality level. Bars depict 95% confidence intervals.

Figure 1.C15: The impact on hiring decisions (extensive margins) across quintiles of exposure



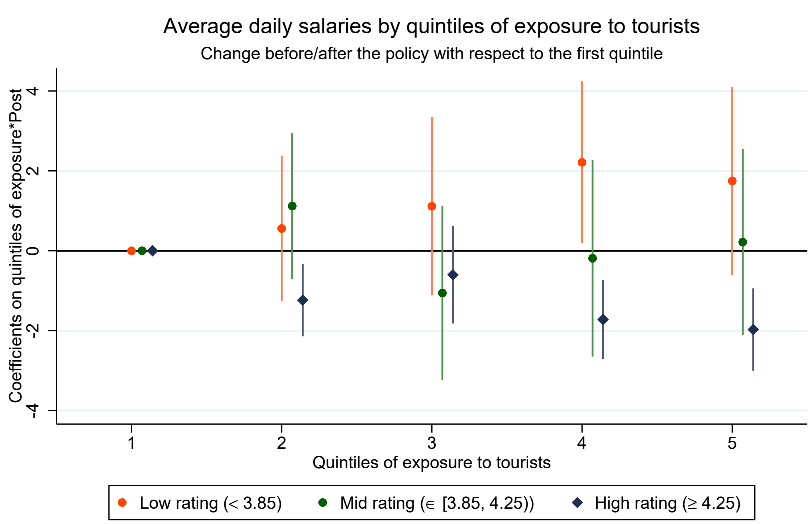
Notes: The graph reports estimates on the interactions of quintiles of exposure*Post from three separate regressions (low, mid, high-rating restaurants), where each observation is a restaurant-month-year. The first quintile is omitted. All controls and fixed-effects from the main analysis are included. Post takes value 1 after May 2017. The sample includes observations between Jan 2015 and Dec 2019. Heteroskedasticity-robust standard errors are clustered at municipality level. Bars depict 95% confidence intervals.

Figure 1.C16: The impact on hiring decisions (intensive margins) across quintiles of exposure



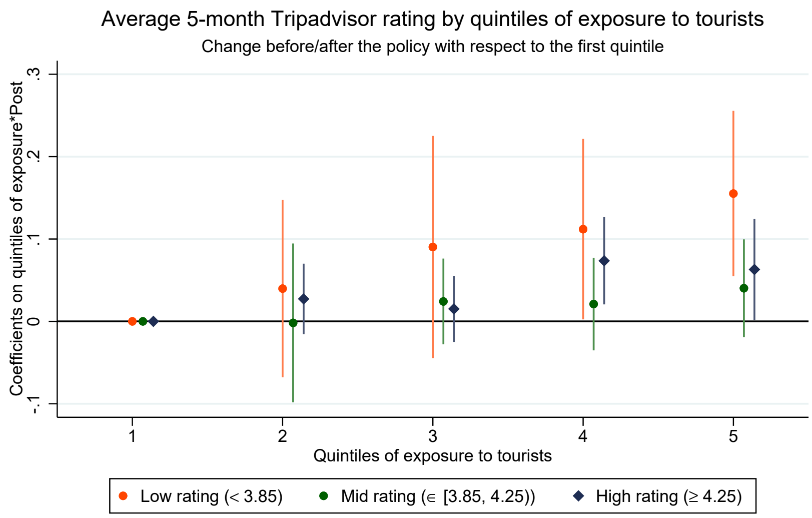
Notes: The graph reports estimates on the interactions of quintiles of exposure*Post from three separate regressions (low, mid, high-rating restaurants), where each observation is a restaurant-month-year. The first quintile is omitted. All controls and fixed-effects from the main analysis are included. Post takes value 1 after May 2017. The sample includes observations between Jan 2015 and Dec 2019. Heteroskedasticity-robust standard errors are clustered at municipality level. Bars depict 95% confidence intervals.

Figure 1.C17: The impact on salaries across quintiles of exposure



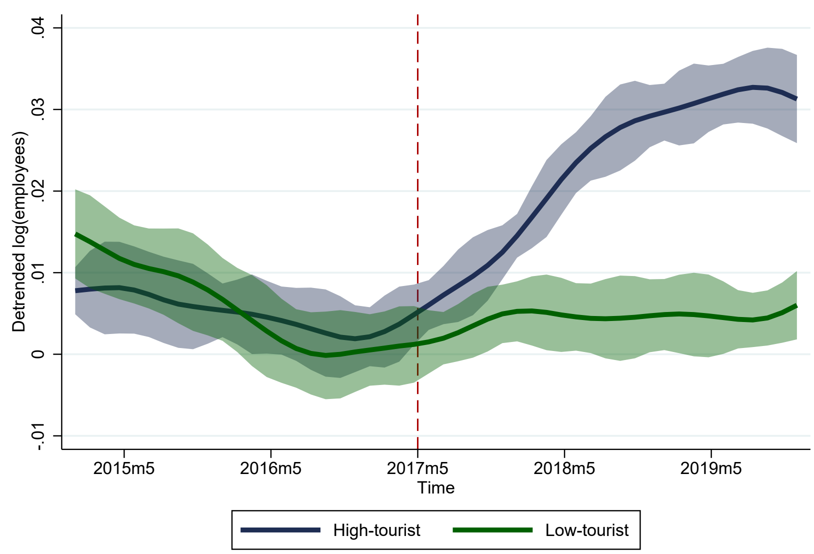
Notes: The graph reports estimates on the interactions of quintiles of exposure*Post from three separate regressions (low, mid, high-rating restaurants), where each observation is a restaurant-month-year. The first quintile is omitted. All controls and fixed-effects from the main analysis are included. Post takes value 1 after May 2017. The sample includes observations between Jan 2015 and Dec 2019. Heteroskedasticity-robust standard errors are clustered at municipality level. Bars depict 95% confidence intervals.

Figure 1.C18: The impact on Tripadvisor rating across quintiles of exposure



Notes: The graph reports estimates on the interactions of quintiles of exposure*Post from three separate regressions (low, mid, high-rating restaurants), where each observation is a restaurant-month-year. The first quintile is omitted. All controls and fixed-effects from the main analysis are included. Post takes value 1 after May 2017. The sample includes observations between Jan 2015 and Dec 2019. Heteroskedasticity-robust standard errors are clustered at municipality level. Bars depict 95% confidence intervals.

Figure 1.C19: The impact of the roaming policy on de-trended monthly employment



Notes: High-tourist restaurants are those for which the binary variable *Tourist*=1, while *Tourist*=0 for low-tourist restaurants.

1.D Tables

Table 1.D1: Comparison of main statistics across samples

	Tripadvisor sample	Matched sample	INPS sample
	Mean/(SD)	Mean/(SD)	Mean/(SD)
N. of tourist attractions in the ZIP code	3.01	3.10	2.59
	(6.65)	(6.59)	(6.13)
Probability of exposure to tourists	0.04	0.04	
	(0.12)	(0.10)	
Distance (km) from closest attraction	9.02	8.19	
	(11.14)	(10.71)	
Average 5-month Tripadvisor rating	3.97	4.00	
	(0.65)	(0.56)	
Average N. of 5-month replies to reviews	1.79	2.38	
	(8.98)	(9.58)	
Total Tripadvisor reviews	134.62	174.46	
	(320.06)	(297.80)	
Price €	0.25	0.25	
	(0.43)	(0.43)	
Price €€– €€€	0.71	0.71	
	(0.45)	(0.45)	
Price €€€€	0.04	0.04	
	(0.20)	(0.19)	
Average N. of monthly employees		5.24	4.31
		(5.02)	(4.60)
1 if firm exits market in Jan2015-Dec2019		0.21	0.29
		(0.41)	(0.45)
1 if firm exits after policy (Jun2017-Dec2019)		0.13	0.15
		(0.34)	(0.36)
1 if firm enters market in Jan2015-Dec2019		0.41	0.44
		(0.49)	(0.50)
1 if firm enters after policy (Jun2017-Dec2019)		0.15	0.17
		(0.36)	(0.38)
1 if firm hires workers w/ experience in restaurants at least once in Jan2015-Dec2019		0.76	0.69
		(0.43)	(0.46)
Average months of experience of newly-hired employees		13.64	12.72
		(14.55)	(14.45)
Average daily salaries (€)		65.98	65.88
		(10.48)	(11.41)
Observations	14146	5472	10391

Each observation is a restaurant. Data refer to the period between Jan 2015 - Dec 2019, unless otherwise specified. Data on Tripadvisor reviews, rating and replies refer to the period between Jan 2015 and Dec 2018. The matches sample is used in the market-level analysis.

Table 1.D2: The roaming policy and the use of Tripadvisor by nationality of the reviewer

	Ratio Mobile/PC monthly reviews			Total monthly reviews from Mobile devices		
	(1)	(2)	(3)	(4)	(5)	(6)
EU	0.33***			-2301.21***		
	(0.08)			(111.86)		
EU*Post	0.47***			496.37***		
	(0.12)			(155.62)		
Extra-EU		0.18***			-2916.72***	
		(0.06)			(141.85)	
Extra-EU*Post		-0.12			129.76	
		(0.09)			(197.34)	
IT			0.99***			5617.90***
			(0.11)			(300.21)
IT*Post			0.02			-714.22*
			(0.15)			(417.65)
Month*Year FE	✓	✓	✓	✓	✓	✓
Observations	120	120	120	120	120	120
Adj. R-squared	0.840	0.824	0.764	0.923	0.908	0.880
Mean EU pre-policy	1.14			3162.97		
Mean Extra-EU pre-policy		1.00			2547.45	
Mean IT pre-policy			1.80			11082.07

Post takes value 1 after May 2017. Each observation is a region of origin-month-year. The regions of origin are EU, Extra-EU, IT and locals, which is the comparison (omitted) category in every column. The panel includes observations between 2015 and 2019. Standalone Post is absorbed by the Month*Year FE. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D3: The roaming policy and international travelers to Italy

	Δ from previous year, Thousdands					
	Overnight stays			Tourists at borders		
	(1)	(2)	(3)	(4)	(5)	(6)
EU	506.212***	530.153***		126.723***	133.083***	
	(83.752)	(118.615)		(17.343)	(24.561)	
Post	154.244*	172.200*	310.167	0.800	5.570	55.051
	(81.093)	(102.724)	(204.811)	(16.793)	(21.270)	(41.994)
EU*Post		-47.883	-47.883		-12.720	-12.720
		(167.747)	(161.642)		(34.734)	(33.143)
Origin FE			✓			✓
Year and quarter FE			✓			✓
Observations	320	320	320	320	320	320
Adj. R-squared	0.107	0.104	0.168	0.139	0.136	0.214

The data contain the number of international tourists traveling to Italy from different regions/countries of the world, in Thousdands. Each observation is a region of origin-year-quarter. The panel includes observations between 2015 and 2019. EU takes value 1 when the region of origin of the tourists is a EU country, and 0 otherwise. Post takes value 1 after 2017q2. Source: Bank of Italy. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D4: Summary statistics of independent and control variables

	Firms	Mean	SD	Min	Median	Max
Probability of exposure to tourists (×100)	4628	4.21	10.27	0.0	0.17	100.0
1 if tourist restaurant	4628	0.49	0.50	0.0	0.00	1.0
Tripadvisor rating at policy	4628	4.01	0.51	1.0	4.05	5.0
1 if low-rating ($\in [1, 3.85)$)	4628	0.33	0.47	0.0	0.00	1.0
1 if mid-rating ($\in [3.85, 4.25)$)	4628	0.34	0.47	0.0	0.00	1.0
1 if high-rating ($\in [4.25, 5]$)	4628	0.33	0.47	0.0	0.00	1.0
1 if restaurant is LLC	4628	0.56	0.50	0.0	1.00	1.0
1 if sole proprietorship	4628	0.23	0.42	0.0	0.00	1.0
1 if dine-in restaurant/bar	4628	0.86	0.34	0.0	1.00	1.0
1 if food truck	4628	0.04	0.19	0.0	0.00	1.0
1 if take-away only	4628	0.07	0.25	0.0	0.00	1.0
Distance (km) from closest attraction	4628	8.10	10.71	0.0	3.00	55.0
1 if distance to Rome city center <6 km	4628	0.49	0.50	0.0	0.00	1.0
1 if distance to Rome city center 6-15 km	4628	0.20	0.40	0.0	0.00	1.0
1 if distance to Rome city center >15 km	4628	0.32	0.47	0.0	0.00	1.0
1 if price is €	4576	0.26	0.44	0.0	0.00	1.0
1 if price is €€– €€€	4576	0.70	0.46	0.0	1.00	1.0
1 if price is €€€€	4576	0.04	0.19	0.0	0.00	1.0
1 if cuisine is Italian	4628	0.76	0.43	0.0	1.00	1.0
1 if no other restaurant in 400 m radius	4628	0.05	0.23	0.0	0.00	1.0
1 if 1-10 restaurants in 400 m radius	4628	0.25	0.43	0.0	0.00	1.0
1 if 11-30 restaurants in 400 m radius	4628	0.19	0.39	0.0	0.00	1.0
1 if more than 30 restaurants in 400 m radius	4628	0.51	0.50	0.0	1.00	1.0
1 if closest attraction has <1,000 reviews	4628	0.37	0.48	0.0	0.00	1.0
1 if closest attraction has 1,000-5,000 reviews	4628	0.36	0.48	0.0	0.00	1.0
1 if closest attraction has >5,000 reviews	4628	0.26	0.44	0.0	0.00	1.0

Each observation is a restaurant.

Table 1.D5: The impact of the roaming policy on restaurant profit margin

	Y=Annual profit margin; years 2015-2018				
	(1) Full sample	(2) Full sample	(3) Low rating	(4) Mid rating	(5) High rating
Tourist*Post	-0.013***	-0.013***	-0.017***	-0.031***	0.025**
	(0.004)	(0.004)	(0.006)	(0.009)	(0.010)
Restaurant & Year FE	✓	✓	✓	✓	✓
ZIP-code*Year	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	6614	6591	2291	2283	2017
Restaurants	2026	2018	693	693	632
Clusters	57	56	39	40	41
Adj. R-squared	0.345	0.349	0.326	0.352	0.363
DDD *p-value*				0.061	0.096

Post=1 if date is after 2016. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-year. The sample includes observations between 2015 and 2018. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Additional results

Table 1.D6: The impact of the roaming policy on total working days

	Y=log(monthly working days); Jan 2015 - Dec 2019				
	(1) Full sample	(2) Full sample	(3) Low rating	(4) Mid rating	(5) High rating
Tourist*Post	0.087***	0.081**	-0.066*	0.188***	0.100
	(0.031)	(0.035)	(0.039)	(0.055)	(0.098)
Restaurant & Time FE	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	219835	217622	72133	76920	68569
Restaurants	4628	4576	1490	1571	1515
Clusters	86	86	59	71	71
Adj. R-squared	0.714	0.713	0.701	0.728	0.700
Mean Y pre-policy	91.0	91.7	113.8	95.4	63.0
DDD *p-value*				0.034	0.104

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D7: The impact of the roaming policy on working days per worker

	Y=N. of working days per worker; Jan 2015 - Dec 2019				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	0.085	0.008	0.038	-0.466***	0.371**
	(0.094)	(0.092)	(0.133)	(0.142)	(0.159)
Restaurant & Time FE	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	196749	195348	64454	70021	60873
Restaurants	4517	4471	1454	1537	1480
Clusters	86	86	59	71	70
Adj. R-squared	0.727	0.727	0.741	0.742	0.695
Mean Y pre-policy	15.8	15.8	16.0	16.2	15.2
DDD *p-value*				0.263	0.294

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D8: The impact of the roaming policy on restaurant employment by price category

	Y=log(monthly employees); Jan 2015 - Dec 2019								
	Low price (€)			Medium price (€€– €€€)			High price (€€€€)		
	(1) Low rating	(2) Mid rating	(3) High rating	(4) Low rating	(5) Mid rating	(6) High rating	(7) Low rating	(8) Mid rating	(9) High rating
Tourist*Post	-0.057**	0.088**	-0.025	-0.008	0.117***	0.103**	-0.361***	0.113	-0.015
	(0.027)	(0.034)	(0.032)	(0.026)	(0.036)	(0.046)	(0.074)	(0.071)	(0.110)
Restaurant & Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	16919	17819	22560	54187	56170	42228	1027	2931	3781
Restaurants	337	365	493	1130	1149	935	23	57	87
Clusters	29	33	45	55	64	66	6	12	17
Adj. R-squared	0.762	0.781	0.769	0.754	0.792	0.756	0.801	0.806	0.791
Mean Y pre-policy	5.4	4.7	3.2	7.4	5.8	4.2	8.5	9.6	7.3

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D9: Logit estimates: the impact of the roaming policy on hiring decisions

Y=1 if firm hires worker	*with* previous experience in restaurants				*without* experience in restaurants		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Full sample	Low rating	Mid rating	High rating	Low rating	Mid rating	High rating
Tourist*Post	0.175***	0.150**	0.235***	0.032	-0.028	0.227**	0.160
	(0.048)	(0.074)	(0.087)	(0.095)	(0.084)	(0.093)	(0.100)
Restaurant FE	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓
Observations	176898	61956	62097	52845	61361	63352	55925
Restaurants	3568	1229	1221	1118	1178	1221	1157
Pseudo R-squared	0.014	0.017	0.015	0.018	0.013	0.014	0.012
Mean Y pre-policy	0.12	0.14	0.10	0.11	0.14	0.08	0.07
DDD *p-value*			0.376	0.013		0.409	0.378

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (3-4) and (6-7) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D10: The impact of the roaming policy on restaurant daily salaries

	Y=log(average daily salary (€)); Jan 2015 - Dec 2019				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	-0.000	0.000	0.016**	-0.001	-0.014**
	(0.003)	(0.003)	(0.006)	(0.006)	(0.005)
Restaurant & Time FE	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	200402	199026	67507	70593	60926
Restaurants	4558	4512	1492	1538	1482
Clusters	86	86	59	71	70
Adj. R-squared	0.495	0.496	0.511	0.491	0.482
Mean Y pre-policy	64.9	64.9	66.0	65.0	63.5
DDD *p-value*				0.477	0.102

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D11: The impact of the roaming policy on hiring from Tripadvisor restaurants

Y=1 if firm hires worker from	Tripadvisor restaurants with any rating				Tripadvisor restaurants with mid/high rating			
	(1) Full sample	(2) Low rating	(3) Mid rating	(4) High rating	(5) Full sample	(6) Low rating	(7) Mid rating	(8) High rating
Tourist*Post	0.004***	0.004	0.006**	-0.001	0.006***	0.008***	0.005**	0.001
	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.001)
Restaurant & Time FE	✓	✓	✓	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Observations	217622	72133	76920	68569	217622	72133	76920	68569
Restaurants	4576	1490	1571	1515	4576	1490	1571	1515
Clusters	86	59	71	71	86	59	71	71
Adj. R-squared	0.086	0.102	0.071	0.082	0.067	0.076	0.054	0.070
Mean Y pre-policy	0.04	0.05	0.04	0.04	0.03	0.03	0.03	0.03
DDD *p-value*			0.142	0.184			0.649	0.004

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (3-4) and (7-8) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D12: Correlation between Tripadvisor rating and restaurant hiring decisions

	Y=Average 5-month Tripadvisor rating; Jan 2012 - Dec 2018					
	(1)	(2)	(3)	(4)	(5)	(6)
Hire worker w/ experience in restaurants	0.0116***	0.0116***	0.0112***	0.0054*	0.0097***	
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	
Hire worker w/o experience in restaurants		-0.0002	0.0001	0.0001	-0.0013	
		(0.003)	(0.003)	(0.003)	(0.003)	
Hire worker from higher-rating restaurant				0.0162***		
				(0.003)		
log(monthly employees)					0.0061**	
					(0.003)	
Years of experience in restaurants						0.0043***
						(0.001)
Restaurant & Time FE	✓	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓	✓
Controls			✓	✓	✓	✓
Observations	223494	223494	222405	222405	222405	30015
Restaurants	5147	5147	5089	5089	5089	3737
Clusters	89	89	89	89	89	76
Adj. R-squared	0.480	0.480	0.482	0.482	0.482	0.557
Mean Y	3.95	3.95	3.95	3.95	3.95	3.91

Each observation is a restaurant-month-year. The sample includes observations between Jan 2012 and Dec 2018. Heteroskedasticity-robust standard errors clustered at municipality level. Controls include distance (km) to closest attraction and indicators for distance to Rome city center, restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with time trends. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D13: The impact of the roaming policy on restaurant replies to Tripadvisor reviews

	Y=N. of 5-month replies to reviews; Jan 2015 - Dec 2018				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	0.156*	0.308***	0.008	1.013**	-0.444*
	(0.083)	(0.094)	(0.247)	(0.442)	(0.258)
Restaurant & Time FE	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	146713	145085	48937	52172	43976
Restaurants	4377	4328	1412	1499	1417
Clusters	86	86	59	70	70
Adj. R-squared	0.704	0.704	0.663	0.647	0.752
Mean Y pre-policy	2.56	2.59	1.58	2.51	3.90
DDD *p-value*				0.006	0.000

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D14: The impact of the roaming policy on restaurant net purchases

	Y=log(annual net purchases); years 2015-2018				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	0.056***	0.060***	-0.034	0.115***	0.019
	(0.012)	(0.010)	(0.029)	(0.041)	(0.026)
Restaurant & Year FE	✓	✓	✓	✓	✓
ZIP-code*Year	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	6677	6652	2305	2299	2048
Restaurants	2043	2034	696	697	641
Clusters	57	56	39	40	41
Adj. R-squared	0.858	0.859	0.882	0.865	0.801
Mean Y pre-policy	255.9	256.8	369.1	228.9	154.7
DDD *p-value*				0.839	0.027

Post=1 if date is after 2016. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-year. The sample includes observations between 2015 and 2018. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Placebo policy-dates

Table 1.D15: Placebo policies and restaurant revenues

	Y=log(annual revenues)			
	(1)	(2)	(3)	(4)
	Full sample	Low rating	Mid rating	high rating
Tourist*Post 2013	0.010	-0.037	0.047	-0.028
	(0.023)	(0.030)	(0.038)	(0.023)
Observations	4173	1763	1393	1017
Tourist*Post 2014	-0.003	0.010	-0.025	-0.024
	(0.016)	(0.020)	(0.027)	(0.043)
Observations	5029	2040	1662	1327
Tourist*Post 2015	0.026	0.077	-0.022	0.003
	(0.019)	(0.046)	(0.015)	(0.048)
Observations	4351	1706	1397	1248

Every row/column is the output of a separate regression where each observation is a restaurant-year. Post *year*=1 if date is after *year*. Tourist restaurants are those with a measure of exposure above the median. Heteroskedasticity-robust standard errors clustered at municipality level. Each regression includes all controls and fixed effects from the main analysis. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the placebo policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D16: Placebo policies and restaurant exit

	Y=1 if firm exits the market			
	(1)	(2)	(3)	(4)
	Full sample	Low rating	Mid rating	High rating
Tourist*Post May2013	-0.001	0.002	-0.003***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.002)
Observations	39012	16693	12963	9356
Tourist*Post May2014	0.001	0.002	0.001	-0.001
	(0.001)	(0.002)	(0.002)	(0.003)
Observations	48795	19572	16664	12559
Tourist*Post May2015	-0.000	-0.000	-0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.002)
Observations	58028	22394	19658	15976
Tourist*Post May2016	0.000	0.000	0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.002)
Observations	58893	21014	20066	17813

Every row/column is the output of a separate regression where each observation is a restaurant-month-year. Post May *year*=1 if date is after May of the respective *year*. Tourist restaurants are those with a measure of exposure above the median. Heteroskedasticity-robust standard errors clustered at municipality level. Each regression includes all controls and fixed effects from the main analysis. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the placebo policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D17: Placebo policies and industry composition

	Y=log(N. of active establishments)			
	(1)	(2)	(3)	(4)
	All	Low rating	Mid rating	High rating
N. of attractions*Post May2013	0.001	0.001	-0.001	0.003
	(0.001)	(0.001)	(0.001)	(0.002)
Observations	2712	2712	2712	2712
N. of attractions*Post May2014	-0.001	0.000	-0.001	-0.001
	(0.001)	(0.001)	(0.002)	(0.002)
Observations	2841	2841	2841	2841
N. of attractions*Post May2015	-0.002**	-0.001	-0.001	-0.002
	(0.001)	(0.002)	(0.001)	(0.001)
Observations	2881	2881	2881	2881
N. of attractions*Post May2016	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.002)
Observations	2240	2240	2240	2240

Every row/column is the output of a separate regression where each observation is a ZIP code-month-year. Post May *year*=1 if date is after May of the respective *year*. Heteroskedasticity-robust standard errors clustered at ZIP-code level. Each regression includes all controls and fixed effects from the main analysis. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the placebo policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. If the restaurant entered the market after the placebo policy-date, the most recent rating is considered. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D18: Placebo policies and restaurant hiring decisions (extensive margins)

	Y=1 if firm hires worker with previous experience in restaurants			
	(1)	(2)	(3)	(4)
	Full sample	Low rating	Mid rating	High rating
Tourist*Post May2013	0.006	0.014	-0.009	0.015
	(0.004)	(0.008)	(0.008)	(0.016)
Observations	38862	16647	12921	9294
Tourist*Post May2014	-0.001	-0.005	0.002	-0.003
	(0.004)	(0.004)	(0.006)	(0.010)
Observations	48795	19572	16664	12559
Tourist*Post May2015	0.003	-0.001	-0.004	0.019***
	(0.003)	(0.005)	(0.004)	(0.004)
Observations	58028	22394	19658	15976
Tourist*Post May2016	0.005	-0.009	-0.001	0.026***
	(0.003)	(0.006)	(0.004)	(0.005)
Observations	60147	21359	20354	18434

Every row/column is the output of a separate regression where each observation is a restaurant-month-year. Post May *year*=1 if date is after May of the respective *year*. Tourist restaurants are those with a measure of exposure above the median. Heteroskedasticity-robust standard errors clustered at municipality level. Each regression includes all controls and fixed effects from the main analysis. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the placebo policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D19: Placebo policies and restaurant hiring decisions (intensive margins)

	Y=Months of experience in restaurants of newly-hired employees			
	(1)	(2)	(3)	(4)
	Full sample	Low rating	Mid rating	High rating
Tourist*Post May2013	1.391	0.054	1.884	3.791*
	(1.080)	(1.314)	(1.924)	(2.147)
Observations	6270	3028	2001	1241
Tourist*Post May2014	-0.842	-2.531***	-0.085	-0.562
	(0.574)	(0.919)	(0.984)	(1.409)
Observations	7431	3499	2338	1594
Tourist*Post May2015	-0.547	0.425	-2.142	-1.574
	(0.578)	(0.921)	(1.643)	(2.583)
Observations	10770	4922	3351	2497
Tourist*Post May2016	0.652	-0.631	0.799	3.394***
	(0.458)	(0.545)	(1.390)	(0.967)
Observations	11592	4957	3579	3056

Every row/column is the output of a separate regression where each observation is a restaurant-month-year. Post May *year*=1 if date is after May of the respective *year*. Tourist restaurants are those with a measure of exposure above the median. Heteroskedasticity-robust standard errors clustered at municipality level. Each regression includes all controls and fixed effects from the main analysis. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the placebo policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D20: Placebo policies and restaurant daily salaries

	Y=Average daily salary (€)			
	(1)	(2)	(3)	(4)
	Full sample	Low rating	Mid rating	High rating
Tourist*Post May2013	-0.747	-0.730	-0.465	-0.466
	(0.456)	(0.457)	(0.944)	(0.870)
Observations	35759	15321	11943	8495
Tourist*Post May2014	0.336	0.761	0.145	0.031
	(0.440)	(0.759)	(0.429)	(0.453)
Observations	44362	18103	15201	11058
Tourist*Post May2015	0.322	0.486	0.480	-0.515
	(0.306)	(0.396)	(0.421)	(0.492)
Observations	52527	20429	17934	14164
Tourist*Post May2016	0.167	0.690	-0.233	0.072
	(0.338)	(0.564)	(0.364)	(0.630)
Observations	53148	18996	18298	15854

Every row/column is the output of a separate regression where each observation is a restaurant-month-year. Post May *year*=1 if date is after May of the respective *year*. Tourist restaurants are those with a measure of exposure above the median. Heteroskedasticity-robust standard errors clustered at municipality level. Each regression includes all controls and fixed effects from the main analysis. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the placebo policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D21: Placebo policies and restaurant Tripadvisor rating

	Y=Average 5-month Tripadvisor rating			
	(1)	(2)	(3)	(4)
	Full sample	Low rating	Mid rating	High rating
Tourist*Post May2013	0.016	0.005	0.040	0.008
	(0.024)	(0.051)	(0.025)	(0.040)
Observations	36043	15470	12016	8557
Tourist*Post May2014	0.005	0.001	-0.006	0.007
	(0.022)	(0.026)	(0.018)	(0.037)
Observations	47105	18979	16161	11965
Tourist*Post May2015	0.010	0.004	0.009	0.008
	(0.011)	(0.026)	(0.026)	(0.016)
Observations	56176	21665	19262	15249
Tourist*Post May2016	0.007	-0.028	0.026	0.034**
	(0.010)	(0.018)	(0.022)	(0.015)
Observations	56540	20096	19572	16872

Every row/column is the output of a separate regression where each observation is a restaurant-month-year. Post May *year*=1 if date is after May of the respective *year*. Tourist restaurants are those with a measure of exposure above the median. Heteroskedasticity-robust standard errors clustered at municipality level. Each regression includes all controls and fixed effects from the main analysis. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the placebo policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Robustness: exposure to tourist including alternative routes

Table 1.D22: The impact of the roaming policy on restaurant revenues

	Y=log(annual revenues); years 2015-2018				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	0.048***	0.057***	0.010	0.001	0.063**
	(0.011)	(0.011)	(0.022)	(0.016)	(0.026)
Restaurant & Year FE	✓	✓	✓	✓	✓
ZIP-code*Year	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	6677	6652	2305	2299	2048
Restaurants	2043	2034	696	697	641
Clusters	57	56	39	40	41
Adj. R-squared	0.846	0.847	0.869	0.849	0.782
Mean Y pre-policy	646.6	648.8	977.4	558.0	360.7
DDD *p-value*				0.669	0.020

Post=1 if date is after 2016. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-year. The sample includes observations between 2015 and 2018. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D23: The impact of the roaming policy on restaurant employment

	Y=log(monthly employees); Jan 2015 - Dec 2019				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	0.041***	0.042***	-0.008	0.069***	0.052*
	(0.012)	(0.015)	(0.019)	(0.024)	(0.030)
Restaurant & Time FE	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	219835	217622	72133	76920	68569
Restaurants	4628	4576	1490	1571	1515
Clusters	86	86	59	71	71
Adj. R-squared	0.779	0.778	0.759	0.793	0.769
Mean Y pre-policy	5.5	5.6	6.9	5.7	4.0
DDD *p-value*				0.100	0.026

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D24: The impact of the roaming policy on hiring decisions (extensive margins)

Y=1 if firm hires worker	*with* previous experience in restaurants				*without* experience in restaurants		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Full sample	Low rating	Mid rating	High rating	Low rating	Mid rating	High rating
Tourist*Post	0.007***	0.009***	0.004	0.005	-0.006**	0.006	0.006**
	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.004)	(0.003)
Restaurant & Time FE	✓	✓	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓
Observations	217622	72133	76920	68569	72133	76920	68569
Restaurants	4576	1490	1571	1515	1490	1571	1515
Clusters	86	59	71	71	59	71	71
Adj. R-squared	0.124	0.143	0.104	0.116	0.049	0.043	0.037
Mean Y pre-policy	0.08	0.10	0.07	0.08	0.06	0.06	0.06
DDD *p-value*			0.985	0.113		0.006	0.013

The sample includes observations between Jan 2015 and Dec 2019. Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (3-4) and (6-7) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D25: The impact of the roaming policy on hiring decisions (intensive margins)

Y=Months of experience in restaurants of	newly-hired employees				quitting/fired employees		
	(1) Full sample	(2) Low rating	(3) Mid rating	(4) High rating	(5) Low rating	(6) Mid rating	(7) High rating
Tourist*Post	0.678	2.478***	0.060	-0.780	-0.424	-0.051	1.412
	(0.851)	(0.802)	(1.264)	(0.656)	(1.389)	(1.113)	(1.855)
Restaurant & Time FE	✓	✓	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓
Observations	30059	11318	10205	8536	12281	10395	8131
Restaurants	3531	1163	1220	1148	1197	1226	1136
Clusters	76	53	59	61	51	57	58
Adj. R-squared	0.117	0.109	0.117	0.127	0.190	0.170	0.183
Mean Y pre-policy	13.0	13.5	13.5	11.8	25.8	27.0	21.5
DDD *p-value*			0.038	0.000		0.976	0.428

The sample includes observations between Jan 2015 and Dec 2019. Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (3-4) and (6-7) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D26: The impact of the roaming policy on restaurant daily salaries

	Y=Average daily salary (€); Jan 2015 - Dec 2019				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	-0.121	-0.116	0.930***	-0.657	-0.680
	(0.232)	(0.259)	(0.343)	(0.680)	(0.431)
Restaurant & Time FE	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	200402	199026	67507	70593	60926
Restaurants	4558	4512	1492	1538	1482
Clusters	86	86	59	71	70
Adj. R-squared	0.467	0.469	0.485	0.465	0.451
Mean Y pre-policy	64.9	64.9	66.0	65.0	63.5
DDD *p-value*				0.181	0.095

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D27: The impact of the roaming policy on restaurant Tripadvisor rating

	Y=Average 5-month Tripadvisor rating; Jan 2015 - Dec 2018				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	0.040***	0.040***	0.036*	0.076***	0.009
	(0.009)	(0.010)	(0.019)	(0.017)	(0.011)
Restaurant & Time FE	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	147274	146620	48577	53659	44384
Restaurants	4373	4330	1413	1499	1418
Clusters	86	86	59	70	70
Adj. R-squared	0.503	0.504	0.324	0.251	0.297
Mean Y pre-policy	3.98	3.98	3.51	4.05	4.43
DDD *p-value*				0.007	0.000

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Robustness: exposure to tourist weighted by road importance

Table 1.D28: The impact of the roaming policy on restaurant revenues

	Y=log(annual revenues); years 2015-2018				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	0.052***	0.054***	-0.003	0.025	0.061**
	(0.008)	(0.008)	(0.017)	(0.021)	(0.026)
Restaurant & Year FE	✓	✓	✓	✓	✓
ZIP-code*Year	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	6677	6652	2305	2299	2048
Restaurants	2043	2034	696	697	641
Clusters	57	56	39	40	41
Adj. R-squared	0.846	0.847	0.869	0.849	0.782
Mean Y pre-policy	646.6	648.8	977.4	558.0	360.7
DDD *p-value*				0.862	0.005

Post=1 if date is after 2016. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-year. The sample includes observations between 2015 and 2018. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D29: The impact of the roaming policy on restaurant employment

	Y=log(monthly employees); Jan 2015 - Dec 2019				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	0.030***	0.035***	-0.002	0.058**	0.044
	(0.010)	(0.011)	(0.018)	(0.023)	(0.027)
Restaurant & Time FE	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	219835	217622	72133	76920	68569
Restaurants	4628	4576	1490	1571	1515
Clusters	86	86	59	71	71
Adj. R-squared	0.779	0.778	0.759	0.793	0.769
Mean Y pre-policy	5.5	5.6	6.9	5.7	4.0
DDD *p-value*				0.153	0.056

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D30: The impact of the roaming policy on hiring decisions (extensive margins)

Y=1 if firm hires worker	*with* previous experience in restaurants				*without* experience in restaurants		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Full sample	Low rating	Mid rating	High rating	Low rating	Mid rating	High rating
Tourist*Post	0.006**	0.007**	0.002	0.006	-0.002	0.011***	0.005**
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.002)
Restaurant & Time FE	✓	✓	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓
Observations	217622	72133	76920	68569	72133	76920	68569
Restaurants	4576	1490	1571	1515	1490	1571	1515
Clusters	86	59	71	71	59	71	71
Adj. R-squared	0.124	0.143	0.104	0.116	0.049	0.043	0.037
Mean Y pre-policy	0.08	0.10	0.07	0.08	0.06	0.06	0.06
DDD *p-value*			0.954	0.330		0.005	0.085

The sample includes observations between Jan 2015 and Dec 2019. Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (3-4) and (6-7) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D31: The impact of the roaming policy on hiring decisions (intensive margins)

Y=Months of experience in restaurants of	newly-hired employees				quitting/fired employees		
	(1) Full sample	(2) Low rating	(3) Mid rating	(4) High rating	(5) Low rating	(6) Mid rating	(7) High rating
Tourist*Post	0.339	2.313***	-0.880	-0.295	0.280	2.240	3.064**
	(0.661)	(0.571)	(1.296)	(0.516)	(1.272)	(1.749)	(1.333)
Restaurant & Time FE	✓	✓	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓
Observations	30059	11318	10205	8536	12281	10395	8131
Restaurants	3531	1163	1220	1148	1197	1226	1136
Clusters	76	53	59	61	51	57	58
Adj. R-squared	0.117	0.109	0.117	0.127	0.190	0.170	0.183
Mean Y pre-policy	13.0	13.5	13.5	11.8	25.8	27.0	21.5
DDD *p-value*			0.011	0.000		0.345	0.564

The sample includes observations between Jan 2015 and Dec 2019. Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (3-4) and (6-7) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D32: The impact of the roaming policy on restaurant daily salaries

	Y=Average daily salary (€); Jan 2015 - Dec 2019				
	(1) Full sample	(2) Full sample	(3) Low rating	(4) Mid rating	(5) High rating
Tourist*Post	-0.153	-0.128	1.036**	-0.623	-1.023*
	(0.269)	(0.269)	(0.439)	(0.621)	(0.561)
Restaurant & Time FE	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	200402	199026	67507	70593	60926
Restaurants	4558	4512	1492	1538	1482
Clusters	86	86	59	71	70
Adj. R-squared	0.467	0.469	0.485	0.465	0.451
Mean Y pre-policy	64.9	64.9	66.0	65.0	63.5
DDD *p-value*				0.224	0.114

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D33: The impact of the roaming policy on restaurant Tripadvisor rating

	Y=Average 5-month Tripadvisor rating; Jan 2015 - Dec 2018				
	(1) Full sample	(2) Full sample	(3) Low rating	(4) Mid rating	(5) High rating
Tourist*Post	0.055***	0.064***	0.087***	0.065**	0.044***
	(0.019)	(0.018)	(0.020)	(0.030)	(0.014)
Restaurant & Time FE	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	147274	146620	48577	53659	44384
Restaurants	4373	4330	1413	1499	1418
Clusters	86	86	59	70	70
Adj. R-squared	0.503	0.504	0.324	0.251	0.297
Mean Y pre-policy	3.98	3.98	3.51	4.05	4.43
DDD *p-value*				0.000	0.000

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Robustness: excluding firms that exited the market after the policy

Table 1.D34: The impact of the roaming policy on restaurant revenues

	Y=log(annual revenues); years 2015-2018				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	0.068***	0.067***	0.033	0.043**	0.074**
	(0.016)	(0.017)	(0.030)	(0.016)	(0.030)
Restaurant & Year FE	✓	✓	✓	✓	✓
ZIP-code*Year	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	6176	6158	2148	2130	1880
Restaurants	1872	1865	642	637	586
Clusters	56	55	35	38	39
Adj. R-squared	0.852	0.853	0.871	0.858	0.790
Mean Y pre-policy	664.8	666.7	1011.7	563.7	366.1
DDD *p-value*				0.830	0.024

Post=1 if date is after 2016. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-year. The sample includes observations between 2015 and 2018. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D35: The impact of the roaming policy on restaurant employment

	Y=log(monthly employees); Jan 2015 - Dec 2019				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	0.050**	0.047*	-0.017	0.101***	0.047
	(0.022)	(0.026)	(0.020)	(0.034)	(0.051)
Restaurant & Time FE	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	199012	197346	65432	70772	61142
Restaurants	4071	4033	1315	1413	1305
Clusters	85	85	58	71	70
Adj. R-squared	0.784	0.783	0.764	0.797	0.777
Mean Y pre-policy	5.7	5.7	7.1	5.8	4.1
DDD *p-value*				0.099	0.063

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D36: The impact of the roaming policy on hiring decisions (extensive margins)

Y=1 if firm hires worker	*with* previous experience in restaurants				*without* experience in restaurants		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Full sample	Low rating	Mid rating	High rating	Low rating	Mid rating	High rating
Tourist*Post	0.009***	0.010***	0.008**	0.002	-0.005**	0.010*	0.009***
	(0.002)	(0.003)	(0.004)	(0.003)	(0.002)	(0.005)	(0.003)
Restaurant & Time FE	✓	✓	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓
Observations	197346	65432	70772	61142	65432	70772	61142
Restaurants	4033	1315	1413	1305	1315	1413	1305
Clusters	85	58	71	70	58	71	70
Adj. R-squared	0.125	0.144	0.107	0.118	0.050	0.043	0.038
Mean Y pre-policy	0.08	0.10	0.07	0.07	0.06	0.06	0.06
DDD *p-value*			0.598	0.073		0.021	0.008

The sample includes observations between Jan 2015 and Dec 2019. Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (3-4) and (6-7) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D37: The impact of the roaming policy on hiring decisions (intensive margins)

Y=Months of experience in restaurants of	newly-hired employees				quitting/fired employees		
	(1) Full sample	(2) Low rating	(3) Mid rating	(4) High rating	(5) Low rating	(6) Mid rating	(7) High rating
Tourist*Post	1.677*	3.294***	1.252	0.174	0.409	0.142	2.561**
	(0.911)	(1.156)	(1.270)	(0.612)	(1.034)	(1.195)	(1.103)
Restaurant & Time FE	✓	✓	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓
Observations	27448	10396	9414	7638	11228	9436	7171
Restaurants	3129	1035	1098	996	1056	1093	975
Clusters	74	53	57	59	51	56	55
Adj. R-squared	0.116	0.105	0.122	0.122	0.181	0.169	0.182
Mean Y pre-policy	13.1	13.7	13.4	11.8	26.0	27.4	21.8
DDD *p-value*			0.115	0.000		0.866	0.362

The sample includes observations between Jan 2015 and Dec 2019. Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (3-4) and (6-7) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D38: The impact of the roaming policy on restaurant daily salaries

	Y=Average daily salary (€); Jan 2015 - Dec 2019				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	-0.007	-0.014	1.022**	-0.033	-1.025*
	(0.256)	(0.274)	(0.387)	(0.404)	(0.516)
Restaurant & Time FE	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	181903	180953	61649	64930	54374
Restaurants	4011	3979	1319	1384	1276
Clusters	85	85	58	71	69
Adj. R-squared	0.474	0.475	0.489	0.475	0.458
Mean Y pre-policy	64.8	64.9	65.9	65.0	63.5
DDD *p-value*				0.375	0.061

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D39: The impact of the roaming policy on restaurant Tripadvisor rating

	Y=Average 5-month Tripadvisor rating; Jan 2015 - Dec 2018				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	0.053***	0.063***	0.112***	0.091***	0.011
	(0.012)	(0.012)	(0.018)	(0.022)	(0.013)
Restaurant & Time FE	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	131302	130941	43522	48710	38709
Restaurants	3849	3818	1248	1351	1219
Clusters	85	85	58	70	69
Adj. R-squared	0.503	0.504	0.324	0.252	0.300
Mean Y pre-policy	3.97	3.97	3.51	4.05	4.43
DDD *p-value*				0.000	0.000

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at municipality level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Robustness: clustering standard errors at the ZIP-code level

Table 1.D40: The impact of the roaming policy on restaurant revenues

	Y=log(annual revenues); years 2015-2018				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	0.047	0.053	-0.002	0.033	0.069
	(0.030)	(0.033)	(0.075)	(0.052)	(0.071)
Restaurant & Year FE	✓	✓	✓	✓	✓
ZIP-code*Year	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	6677	6652	2305	2299	2048
Restaurants	2043	2034	696	697	641
Clusters	113	113	101	98	99
Adj. R-squared	0.846	0.847	0.869	0.849	0.782
Mean Y pre-policy	646.6	648.8	977.4	558.0	360.7
DDD *p-value*				0.965	0.419

Post=1 if date is after 2016. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at ZIP-code level. Each observation is a restaurant-year. The sample includes observations between 2015 and 2018. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D41: The impact of the roaming policy on restaurant employment

	Y=log(monthly employees); Jan 2015 - Dec 2019				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	0.043**	0.042*	-0.024	0.103***	0.041
	(0.021)	(0.022)	(0.043)	(0.036)	(0.033)
Restaurant & Time FE	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	219835	217622	72133	76920	68569
Restaurants	4628	4576	1490	1571	1515
Clusters	127	127	115	119	119
Adj. R-squared	0.779	0.778	0.759	0.793	0.769
Mean Y pre-policy	5.5	5.6	6.9	5.7	4.0
DDD *p-value*				0.078	0.051

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at ZIP-code level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D42: The impact of the roaming policy on restaurant exit

	Y=1 if firm exits the market; Jan 2015 - Dec 2019				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	0.0011	0.0016*	0.0031**	-0.0000	0.0015
	(0.0008)	(0.0008)	(0.0015)	(0.0012)	(0.0018)
Restaurant & Time FE	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	219835	217622	72133	76920	68569
Restaurants	4628	4576	1490	1571	1515
Clusters	127	127	115	119	119
Adj. R-squared	0.059	0.060	0.058	0.061	0.061
Mean Y pre-policy	0.003	0.003	0.003	0.003	0.004
DDD *p-value*				0.365	0.621

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at ZIP-code level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D43: The impact of the roaming policy on hiring decisions (extensive margins)

Y=1 if firm hires worker	*with* previous experience in restaurants				*without* experience in restaurants		
	(1) Full sample	(2) Low rating	(3) Mid rating	(4) High rating	(5) Low rating	(6) Mid rating	(7) High rating
Tourist*Post	0.009**	0.009	0.011**	0.002	-0.006	0.011*	0.007
	(0.004)	(0.007)	(0.005)	(0.007)	(0.006)	(0.006)	(0.006)
Restaurant & Time FE	✓	✓	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓
Observations	217622	72133	76920	68569	72133	76920	68569
Restaurants	4576	1490	1571	1515	1490	1571	1515
Clusters	127	115	119	119	115	119	119
Adj. R-squared	0.124	0.143	0.104	0.116	0.049	0.043	0.037
Mean Y pre-policy	0.08	0.10	0.07	0.08	0.06	0.06	0.06
DDD *p-value*			0.538	0.363		0.047	0.115

The sample includes observations between Jan 2015 and Dec 2019. Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at ZIP-code level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (3-4) and (6-7) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D44: The impact of the roaming policy on hiring decisions (intensive margins)

Y=Months of experience in restaurants of	newly-hired employees				quitting/fired employees		
	(1) Full sample	(2) Low rating	(3) Mid rating	(4) High rating	(5) Low rating	(6) Mid rating	(7) High rating
Tourist*Post	1.469	2.977**	0.789	0.465	0.177	-0.440	2.375
	(0.924)	(1.149)	(1.584)	(1.680)	(1.548)	(1.928)	(2.135)
Restaurant & Time FE	✓	✓	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓
Observations	30059	11318	10205	8536	12281	10395	8131
Restaurants	3531	1163	1220	1148	1197	1226	1136
Clusters	121	108	111	114	107	110	113
Adj. R-squared	0.117	0.109	0.117	0.127	0.190	0.170	0.183
Mean Y pre-policy	13.0	13.5	13.5	11.8	25.8	27.0	21.5
DDD *p-value*			0.031	0.000		0.434	0.466

The sample includes observations between Jan 2015 and Dec 2019. Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at ZIP-code level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (3-4) and (6-7) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D45: The impact of the roaming policy on restaurant daily salaries

	Y=Average daily salary (€); Jan 2015 - Dec 2019				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	-0.010	0.038	1.312**	-0.120	-1.125*
	(0.384)	(0.376)	(0.555)	(0.528)	(0.638)
Restaurant & Time FE	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	200402	199026	67507	70593	60926
Restaurants	4558	4512	1492	1538	1482
Clusters	125	125	114	118	118
Adj. R-squared	0.467	0.469	0.485	0.465	0.451
Mean Y pre-policy	64.9	64.9	66.0	65.0	63.5
DDD *p-value*				0.346	0.079

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at ZIP-code level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.D46: The impact of the roaming policy on restaurant Tripadvisor rating

	Y=Average 5-month Tripadvisor rating; Jan 2015 - Dec 2018				
	(1)	(2)	(3)	(4)	(5)
	Full sample	Full sample	Low rating	Mid rating	High rating
Tourist*Post	0.040*	0.049**	0.087**	0.077**	-0.003
	(0.021)	(0.020)	(0.040)	(0.030)	(0.035)
Restaurant & Time FE	✓	✓	✓	✓	✓
ZIP-code*Time	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Observations	147274	146620	48577	53659	44384
Restaurants	4373	4330	1413	1499	1418
Clusters	127	127	115	119	119
Adj. R-squared	0.503	0.504	0.324	0.251	0.297
Mean Y pre-policy	3.98	3.98	3.51	4.05	4.43
DDD *p-value*				0.128	0.015

Post=1 if date is after May 2017. Tourist restaurants are those with a measure of exposure above the median. Restaurants in low, mid and high categories had their Tripadvisor rating at the time of the policy $\in [1, 3.85), [3.85, 4.25), [4.25, 5]$, respectively. Heteroskedasticity-robust standard errors clustered at ZIP-code level. Each observation is a restaurant-month-year. Distance (km) to closest attraction and indicators for distance to Rome city center interacted with Post are included in all regressions. Controls include indicators for restaurant price category, Italian cuisine, concentration of restaurants in 400m radius, n. of reviews to closest attraction, type of economic activity and legal status of the firm, all interacted with Post. Columns (4) and (5) report the p-values from a triple-difference estimation testing whether the DDD coefficients for mid and high ratings are equal to low rating. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Chapter 2

Mobile internet access and political outcomes: evidence from South Africa

2.1 Introduction

Many developing countries have poor governance, often driven by low political participation and low electoral competition. International observers and reporters believed that the arrival of mobile internet could provide new opportunities to solve some of the problems of the developing world. However, little is known on the impact of the internet on political participation and electoral competition, especially in developing countries where the topic is particularly relevant. Existing literature primarily focused on developed economies finds that broadband internet generally led to lower political participation, most likely due to the increase in entertainment and crowding out of news consumption (Prior, 2005). These effects could be drastically different in developing countries, where the traditional media sector is overall less competitive and the level of political information in the pre-internet era used to be low. In this environment, internet arrival can provide citizens with new sources of information about politics, help individuals to overcome collective action and coordination problems by reducing communication costs, and foster political change.

This paper studies whether the arrival of mobile internet in South Africa promotes political participation and electoral competition, and affects voting behavior. Specifically, I focus on a period during which the country was experiencing corruption scandals and socio-economic turmoil, and look at the impact of 3G coverage on voter turnout and number of parties, as well as the allocation of the vote shares among the major parties. In contrast to findings for developed countries (Falck et al., 2014; Gavazza et al., 2019), I find that internet arrival in fact increased people's political participation in South Africa.[1] I also find a positive effect on political competition, reflected in a larger number of parties running for elections. In turn, after more than two decades, the undisputed dominance of the incumbent party was damaged. All together, these results suggest that mobile internet can have the potential to improve the quality of governance in developing countries.

Identifying the causal impact of mobile internet access on political outcomes is challenging. Compared to the literature on radio and TV coverage,[2] in this case it is harder to exploit technological features of the transmitters to find an exogenous source of variation. Internet coverage is far from random, as it reflects private operators' decisions on where to install the technology (Buys et al., 2009). I address endogeneity issues related to internet coverage adopting two complementary strategies: a Difference-in-Difference estimation and an Instrumental Variable approach. In the first case, I exploit variation over time and space in the expansion of 3G coverage along with the change in political outcomes across 35,000 partitions of voting districts between the 2006 and 2016 municipal elections. The second strategy exploits the variation in terrain ruggedness and its differential impact (pre- vs post-2005, i.e. the year in which 3G became available) on 3G coverage. In both cases, I rely on a granular measure of mobile internet coverage and administrative data sources. In particular, I assemble a new dataset containing geo-referenced information at the voting (sub)district-year-level on the share of area covered by 3G, political outcomes, protests, population density, luminosity as a proxy for GDP, and a variety of socio-demographic characteristics, infrastructure and geography.

In the most conservative specification, my findings show that a unitary increase in 3G coverage (i.e. from 0 to 100% of area covered) led to a 2 percentage point increment in voter turnout (approximately 4% of its value in 2006) in 2011 and 2016. At the same time, mobile internet was detrimental for the popularity of the incumbent party (the African National Congress) whose vote share dropped by more than 3 percentage points

[1] Campante et al. (2018) show that in Italy broadband internet positively affected turnout only after 10 years from its arrival.

[2] See Olken (2009), Enikolopov et al. (2011) , La Ferrara et al. (2012) and Durante et al. (2019), among others.

in 2016. By contrast, other parties gained. Moreover, as a result of 3G, competition at ballots intensified: the number of parties running for elections in more covered localities went up by 10% compared to 2006. Lastly, the number of protests against political institutions increased, suggesting that mobile internet arrival exacerbated political discontent or facilitated mass mobilization. The results are robust to different specifications, estimators and units of observations. Moreover, various placebo exercises support the validity of the identifying assumptions.

I complement these findings with suggestive evidence on the possible mechanisms. Following Manacorda and Tesei (2020), I investigate the information and coordination potential of mobile internet for local politics.[3] For voters, access to mobile internet can significantly decrease the cost of obtaining information and, eventually, affect their voting behavior. In line with this hypothesis, I find a more detrimental impact of 3G coverage on the performance of the incumbent party in municipalities with worse administration of government finances and in localities more exposed to socio-economic turmoil. At the same time, for (potential) candidates, digital technologies could drastically reduce communication and coordination costs with voters, thus making barriers to entry into politics easier to overcome, especially for smaller groups that lack significant financial resources. In this respect, I find that mobile internet coverage facilitated political turnover and fostered political competition by favoring the emergence of new parties.

My paper primarily contributes to the literature on political participation and competition in developing countries. Models of electoral participation suggest that more informed individuals are more likely to vote (Feddersen and Pesendorfer, 1996; Matsusaka, 1995). Other authors theoretically analyze information and political competition (Baron, 1994; Grossman and Helpman, 1996; Lindbeck and Weibull, 1987; Lohmann, 1998). Empirical exercises mainly focus on the role of awareness campaigns in local communities (Banerjee, 2006; Goetz and Jenkins, 2001; Jenkins and Goetz, 1999; Paul, 2002) and NGO activities to enhance individuals' public engagement (Boulding, 2010). However, empirical literature on the determinants of political competition and participation in the developing world is scarce, and this paper attempts to fill the existing gap.

My work is also close to the literature in the political economy of media that studies the

[3]Unlike me, their focus is mainly on traditional 2G cellphone-related technologies (e.g., SMS) as their analysis covers a period in which 3G variation in the African context was extremely low. A similar study is also provided by Christensen and Garfias (2018).

impact of broadband internet and digital ICTs on various forms of political participation and mobilization in developed countries (Campante et al., 2018; Falck et al., 2014; Gavazza and Lizzeri, 2009) and developing economies (Manacorda and Tesei, 2020; Miner, 2015). This paper contributes to the latter, bringing novel insights on voter turnout and electoral competition, as well as on the potential underlying mechanisms. In this respect, the paper also adds to the literature on the role of information in selecting better politicians, promoting political accountability and shaping public policies (Besley, 2005; Besley and Burgess, 2002; Besley et al., 2005; Besley and Prat, 2006; Eisensee and Strömberg, 2007; Ferraz and Finan, 2008; Gentzkow et al., 2007; Olken, 2007; Reinikka and Svensson, 2005; Snyder Jr and Strömberg, 2010; Strömberg, 2004).

Finally, this paper also takes inspiration from the literature on digital media and their influence on politics (Allcott and Gentzkow, 2017; Bond et al., 2012; Chen and Yang, 2019; Enikolopov et al., 2020, 2018; Guriev et al., 2019; Petrova et al., 2021; Zhuravskaya et al., 2020) as well as the literature on traditional media and its impact on voting behavior (Casey, 2015; Chiang and Knight, 2011; DellaVigna et al., 2014; DellaVigna and Kaplan, 2007; Enikolopov et al., 2011; Gentzkow et al., 2015, 2011).

The remainder of this paper is structured as follows. Section 2.2 describes the political scenario and the media landscape. Section 2.3 outlines the empirical strategy, while section 2.4 describes the data. Section 2.5 shows the main results and section 2.6 investigates the mechanisms. Finally, section 2.7 draws the conclusions.

2.2 Background

2.2.1 Political scenario

South Africa is the second largest economy and the most industrialized country of the African continent. However, poverty, unemployment and income inequality remain widespread. Government corruption, inefficient bureaucracy and political instability are among the top five challenges of doing business in the country.[4] Local and international observers argue that the search for a policy compromise within the African National Congress (ANC) - the country's dominant party - has often resulted in policy paralysis

[4]Global Competitiveness Report 2015-2016. World Economic Forum.

and the adoption of ineffective policies, hampering economic development.[5] Since the end of apartheid in 1994, the ANC has dominated South Africa's politics. It is the ruling party in eight of the nine provinces of the country, and in most of the municipal councils. Currently, its main challengers are the Democratic Alliance (DA) and the recently formed Economic Freedom Fighters (EFF).

Local government in South Africa consists of municipalities of various types. The largest metropolitan areas are governed by 8 metropolitan municipalities, while the rest of the country is divided into 226 local municipalities. Their councils are elected by a system of mixed-member proportional representation every 5 years. Table 2.1 depicts the overall results of the three major parties in the last municipal ballots, as well as voter turnout and the number of political parties running for election. The striking feature that emerges from the table is the gradual decline in the ANC vote share after 2006, which has been accompanied by an increase in popularity of the DA and the newly formed EFF. At the same time, both voters' participation and electoral competition intensified.

Table 2.1: South African municipal election results

	2000	2006	2011	2016
African National Congress (%)	59.4	64.8	61.9	53.9
Democratic Alliance (%)	22.1	16.2	23.9	26.9
Economic Freedom Fighters (%)	*Formed in 2013*			8.2
Voter turnout (%)	48.1	48.4	57.5	57.7
N. of parties	81	97	122	206

Source: Independent Electoral Commission of South Africa

South Africa's political scene has undergone a substantial transformation over the last 20 years. Possible explanations can be found in the increasing socio-economic and political discontent resulting from a series of administrative controversies and corruption scandals that directly involved the ANC party and the former President Zuma.[6] For instance, in 2012, at the Marikana platinum mine, rock drillers began a series of wildcat strikes seeking for a pay raise. In one of these occasions, the police opened fire on a group of strikers, killing 34 miners and wounding other 78. The incident was considered to be the single most lethal use of force by the police against civilians since the apartheid era.

[5] "ANC corruption is a major cause of South Africa's failure - and the polls will show it". The Guardian. May 8, 2019.

[6] For an overview, see "The trials of Jacob Zuma". BBC. December 15, 2017.

Its impact on the economy was severe, and media repercussions were large, both at the national and international level. Opposition parties and leaders criticized the police and called for Zuma to resign because of the controversy over the shooting, meanwhile black people felt betrayed by "their" party.[7] Other scandals have emerged more recently. A major campaign issue during the 2016 election was corruption within the ANC, in particular President Zuma's relationship with the Gupta family[8] and funding for the construction of his home. Moreover, after the elections, the ANC was accused to spend almost $3.8 million on a covert campaign targeting opposition parties. In the first months of 2018, growing pressure on Zuma led him to resign as President of South Africa.

2.2.2 Media landscape

South Africa has a fragile media independence. For instance, "coverage of certain subjects involving the ruling ANC, government finances, or state-funded improvements to President Zuma's personal home are either off limits or provoke a hostile reaction from the authorities".[9] Similar restrictions are less likely to hold for new media, which are promoted by the spread of the internet. In this respect, the country has one of the most technologically advanced mobile-internet infrastructure on the African continent. The market is mainly dominated by three private internet service providers (ISPs) with more or less homogeneous market shares. Internet users as percentage of the population went up from 5% in 2006 to almost 55% in 2016.[10] This growth was achieved through a significant expansion of the 3G internet infrastructure between 2006 and 2016, which is depicted in Figure 2.C1. By contrast, 2G (GSM) coverage - used for calls and SMSs - has remained mostly stable over the same period.[11] In 2016, about 60% of the adult population owned a smartphone.[12] The cheapest price for 1GB of prepaid data was about $7, which was almost 1.5% of the average monthly per capita income, and no sharp change in data prices has occurred.[13] Fixed broadband internet subscribers remain

[7] "Marikana mine shootings revive bitter days of Soweto and Sharpeville". The Guardian. September 7, 2012.

[8] An Indian-South African business family which owns a business empire spanning computer equipment, media and mining.

[9] The country is also ranked 31^{st} out of 180 on the World Press Freedom Index. http://www.rsf.org

[10] The World Bank.

[11] http://www.collinsbartholomew.com

[12] https://www.wearesocial.com

[13] This makes South Africa one of the country with the lowest tariffs in the Southern African Development Community.

below 4%, and they mostly live in large cities where landline is available.

Previous literature (Falck et al., 2014; Gavazza et al., 2019; George, 2008) has documented that broadband internet arrival in developed economies partially crowded-out the use of traditional media in order to access information. Survey data on South Africa from Afrobarometer (Figure 2.C2) show that although access to online information more than doubled between 2011 and 2015, traditional media still remain the most popular source of information. However, among regular internet users only, the decline in newspapers readership is more visible, indicating that, to some extent, habits towards news consumption gradually shifted from printed papers to their online versions. Circulation numbers from the Audit Bureau of Circulations lend support to this hypothesis (Figure 2.C3). Between 2010 and 2018, circulation dropped relatively more for national outlets (47%) than local newspapers (20%). This makes South Africa at odds with the evidence that the internet has been more detrimental for those media with a greater amount of local news content (Gavazza et al., 2019).[14] Moreover, while hard copies sales decreased substantially, these outlets managed to boost readership in their online versions. In only three years (2016-2018), daily unique browsing for national and local newspapers increased by 65 and 34%, respectively.

Politicians themselves are also aware of the importance of mobile technologies to communicate with voters, and they have increased their presence in social media platforms over the last decade.[15] During the 2014 national election campaign, platforms such as Facebook, YouTube and Twitter were actively used (Malherbe, 2015). While in that year the ANC, DA and EFF all together had almost 300 thousand followers on their Twitter accounts, in 2019 they reached nearly 2 million followers.[16]

2.3 Empirical strategy

2.3.1 Determinants of 3G coverage

The rollout of the 3G infrastructure in South Africa started around 2005. At that time, existing antennas were already supporting the 2G (GSM) wireless communication

[14] See also "Can local newspapers survive in the internet age?". BBC News. 11 May 2012.

[15] "Political parties and social media: Does it make a difference?" City Press. 29 April 2019. https://city-press.news24.com

[16] "Political parties can't ignore social media power in election campaigning: Expert". SABC News. 18 February 2019. http://www.sabcnews.com

technology, which enabled mobile phone calls and SMSs (Minges et al., 2008). 3G technology is an extension of GSM and is designed to offer faster data access speeds for mobile internet. As described by Harris (2011) and exemplified in Figure 2.C4, the primary job of a cell tower is to elevate antennas that transmit and receive radio-frequency signals from mobile devices. Wires run from the antenna to base station equipment (typically located at ground level), which is then connected through fiber optic cables to the backbone network. Three main differences exist between a 2G and a 3G mobile network infrastructure. First, more 3G antennas are needed to cover the same number of connections that one 2G repeater could support alone. Second, the equipment to be installed in the base station to support 3G services (UMTS) is different than the one used for 2G (EDGE). Third, 3G base stations need to be connected to the backbone national network infrastructure via fiber optic cables, which are usually installed underground. Therefore, supplanting 2G with 3G technology requires: (1) the installation of additional towers; (2) updating the base station equipment in existing sites; (3) laying down fiber optic cables.

In this paper I seek to identify the causal impact of 3G internet on political outcomes. However, 3G coverage is far from random, and it reflects market-based calculations and profit optimization choices performed by the private internet Service Providers. In particular, the decisions over the particular site in which the transmitter is installed is likely endogenous.[17] Buys et al. (2009) study the determinants of disparities in cell phone coverage in Sub-Saharan Africa and highlight that both demand and supply side factors play a role. Among other things, they show that places with larger market size, lower elevation and smoother terrain characteristics positively predict coverage. Similarly, in Table 2.B1, I analyze the determinants of mobile internet coverage in South Africa and show that the share of area in a locality that is covered by 3G signal is positively associated with luminosity (income), urbanization, education, presence of a road, and past cellphone penetration, among other things. At the same time, terrain ruggedness, average population age, distance from main cities and to the fiber backbone network negatively predict 3G.

[17] To overcome endogeneity concerns, Manacorda and Tesei (2020) exploit the slower adoption of mobile technology in areas subject to a higher incidence of lightning strikes in the African continent. However, in my case using lightning strikes as an instrumental variable leads to the weak instrument problem.

2.3.2 Difference-in-Difference

The first empirical strategy of this paper is a Difference-in-Difference. To addresses potential endogeneity concerns, this approach relies on the high-resolution features of the data and compares the within-district change over time in the outcomes, across districts experiencing a differential 3G expansion. The main specification I estimate is the following:

$$y_{it} = \beta_0\, 3\mathrm{G}_{it} + \beta_1\, 3\mathrm{G}_{it} \times \theta_{2011,t} + \beta_2\, 3\mathrm{G}_{it} \times \theta_{2016,t} + \mathbf{x}'_{it}\beta_3 + (\mathbf{w}_i \times \theta_t)'\beta_4 + \mu_i + \theta_t + \varepsilon_{it} \quad (2.1)$$

where i is the voting sub-district, and t is the year, such that $t = 2006, 2011, 2016$. This approach uses the period 2006 to 2016 to carry out causal inference, and exploits the first two electoral rounds (2000-2006) to conduct placebo checks. 3G is the share of area covered by 3G signal in each voting sub-district/year. θ_{2011} and θ_{2016} take value 1 when t is 2011 or 2016, respectively. Interacting 3G with year dummies while keeping standalone 3G into the equation allows me to account for pre-existing differences between more and less covered localities in 2006, without attributing a casual meaning to the standalone coefficient. Vector $\mathbf{x}$ contains time-variant controls, while $\mathbf{w}$ contains their values in 2000 as well as time-invariant variables, which I interact with year dummies. I group controls into two categories, namely socio-economic and geographic variables. The former includes luminosity, population density, urbanization rate, average education, share of youths (aged between 14 and 30), average age and indicators for ethnicity. The latter contains indicators for the presence of mines, roads and rivers, slope index, elevation, ruggedness, area, distance to closest city and to the fiber backbone. I assess the sensitivity of the regression coefficients to the inclusion of different sets of controls, as well as to a 5^{th}-order polynomial function of a subset of them when explicitly indicated. Finally, μ and θ represent voting sub-district and year fixed effects, respectively.

Specification reported in equation (2.1) allows me to account for various sources of potential endogenity. In particular, voting sub-district fixed effects account for time invariant unobserved factors that may affect the outcomes and may also be correlated with 3G coverage. The year dummies capture instead the time trend in the outcomes that is common to all localities. I also include province linear time trends to control for unobserved and time-varying province-level characteristics. Finally, I allow demand and supply side factors to have different impacts on the outcome over time by interacting pre-determined controls with time dummies. I cluster standard errors at the level of the smallest (geographically) stable union of voting sub-districts to account for both cross-sectional and temporal correlation in the errors.[18]

Identifying assumption: In equation (2.1), coefficient β_0 captures the average difference in y between more and less covered localities at the baseline year, i.e. 2006. In order for β_1 and β_2 to consistently identify the average treatment effect of 3G coverage on the outcomes of interest, the parallel trend assumption must hold. In other words, in the absence of the 3G technology, outcomes in more and less covered localities should exhibit common trends. I will carry out a series of exercises to provide evidence in favor of this assumption.

2.3.3 Instrumental variable

The second approach exploits the variation in terrain characteristics and their differential impact on 3G coverage before vs. after 2005, i.e. the year in which the mobile internet technology became available. In practice, I instrument 3G coverage with plausibly exogenous variation in (the log of) terrain ruggedness interacted with a dummy equals to 1 if $year > 2005$. To justify the approach, I rely on several technical papers highlighting how the site where an antenna is built must be adjacent to a road for physical access, with availability of electrical power and telecommunications network connectivity (Aker and Mbiti, 2010; GSMA, 2015; Harris, 2011). When the rollout of the 3G started in South Africa, i.e. around 2005, phone companies had to verify that locations met these physical requirements before building antennas. At the same time, they had to expand the fiber optic network to reach additional areas. *Ceteris paribus*, phone companies initially avoided localities with geographical characteristics associated with higher costs - namely, irregular terrain and distance from a main road and major urban centers. This is confirmed by the exploratory analysis reported in Table 2.B1, which also shows that, controlling for other factors, the negative impact of $log(ruggedness)$ on the expansion of 3G coverage enlarges over time. Therefore, I rely on the following 2SLS procedure:

$$
\begin{aligned}
1^{st}\ Stage: &\quad \mathrm{3G}_{it} = \pi_0[log(ruggedness) \times \mathbb{1}(Year > 2005)]_{it} + \mathbf{z}_{it}^{'}\pi_1 + \mu_i + \theta_t + u_{it} \\
2^{nd}\ Stage: &\quad y_{it} = \gamma_0 \widehat{\mathrm{3G}_{it}} + \mathbf{z}_{it}^{'}\gamma_1 + \mu_i + \theta_t + \varepsilon_{it}
\end{aligned}
\tag{2.2}
$$

[18] Details on the construction of the dataset are provided in section **??**.

where $t = 2000,\ 2006,\ 2011,\ 2016$ and 3G is the share of area covered by 3G signal in each voting sub-district/year. *Ruggedness* is measured by the standard deviation of elevation. Vector $\mathbf{z}$ contains the same controls as those in the Diff-in-Diff analysis, with the exception of the presence of a mine, slope index and (log) elevation, which significantly reduce the predictive power of the instrument because of their high correlation with ruggedness.

Identifying assumption: In order for the instrument to be valid, two conditions have to be satisfied. First, it has to be relevant. One can easily investigate relevance by looking at the *Wald F-statistic* from the 1^{st} stage. The second condition is exogeneity. In this context, the identifying assumption requires that the correlation between ruggedness and any relevant omitted voting district's characteristics did not change around 2005, other than through the availability of 3G internet.[19] I will provide evidence on the plausibility of this assumption.

2.4 Data

2.4.1 Dataset construction

I leverage administrative and spatial data sources to assemble a new dataset containing time-varying and geo-referenced information on political outcomes (vote shares, voter turnout, number of parties and protests), share of area covered by 3G, economic development (as measured by luminosity at night), population density, urbanization and a variety of socio-demographic indicators (education, age, youth population and ethnicity) measured at the voting sub-district level. In addition, each observation contains time constant information on average elevation (m), terrain ruggedness (m^2), average terrain slope, area (km^2), a dummy variable indicating the presence of a major road, a river and a mine, the distance (km) from the closest provincial capital or city with more than 1 million inhabitants, and the distance (km) from the closest fiber optic national backbone infrastructure. The construction of the dataset involves two steps.

Step 1. I create new geographical units of analysis whose boundaries are stable over time.

[19] In this respect, terrain ruggedness has been shown to have positive indirect consequences on economic development in Africa, for instance because it afforded protection to those being raided during the slave trades (Nunn and Puga, 2012). However, results also show no differential role of ruggedness for economic development among South African countries. Moreover, to mitigate concerns about potential endogeneity of the standalone ruggedness variable, I use its interaction with a time dummy.

Data on political outcomes come at a very disaggregated level, i.e. the voting district. Districts are small geographical entities with no political significance, which contain voting stations. Each voting district has only one voting station. In 2000, the average municipality contained 57 voting districts. However, the boundaries and the number of these districts changed over time. The Municipal Demarcation Board of South Africa, which provided the demarcation data, is in charge of the determination of districts' boundaries for the elections.[20] The main driver of demarcation redeterminations is the increasing or decreasing number of voters. In 2000 there were 14,988 voting districts, 18,872 in 2006, 20,857 in 2011 and 22,612 in 2016. Hence, using the intersecting procedure explained in details in appendix 2.A, I create about 43,000 geographically stable sub-districts per year. These new entities constitute my observational units.

Step 2. For each of these 43,000 units I calculate zonal statistics (mean/standard deviation) of the above mentioned variables using the GIS toolbox. Appendix 2.A reports a detailed description of the variables and their source. The analysis is conducted after cleaning the dataset and considering only observations that are farther than 15 km from the closest provincial capital or city with more that 1 million inhabitants,[21] and such that their population density in 2000 was lower than the 95^{th} percentile. Excluding the major urban agglomerations mitigates potential confounding effects that may bias the results due to (1) the possible expansion of the landline broadband internet, and (2) the presence of Wi-Fi connections.[22] Finally, to increase the precision of my measures, I neglect localities with no population and exclude the largest 1% in terms of area (km^2). Hence, out of a total of 43,014 observations, 8,026 are dropped.

2.4.2 Descriptive statistics

Table 2.2 provides the main descriptive statistics (mean and standard deviations) for the independent and dependent variables, while Table 2.B2 in the appendix expands

[20] The Board is an independent authority. Its status is protected by section 3 of "The Local Government: Municipal Demarcation Act", 1998, and various judgments by the Constitutional Court. http://www.demarcation.org.za

[21] These are Bhisho, Bloemfontein, Cape Town, Durban, East London, Johannesburg, Kimberley, Nelspruit, Pietermaritzburg (Ulundi), Pietersburg (Polokwane), Port Elizabeth, Pretoria and Richards Bay.

[22] For instance, improvements in access to landline internet connection in cities after 2006 increased employment in higher-skill occupations (Hjort and Poulsen, 2019). If these changes affect people's voting preferences, and landline connectivity is somehow correlated with 3G, then any estimates of the impact of 3G on political outcomes in cities would be biased.

the analysis to the entire set of controls. Between 2006 and 2016, the ANC vote share dropped by 5 percentage points (pp), while voter turnout increased by 4 pp and the number of parties running for elections grew by 3.5 units, on average. During the same period, the average share of area covered by 3G rose by 50 pp. Figure 2.C5 shows the spatial and temporal expansion of 3G signal over the entire country.[23]

Table 2.2: Descriptive statistics

Variable \Year	2000		2006		2011		2016	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
3G (share of area covered)	0	*0*	0.10	*0.29*	0.22	*0.38*	0.60	*0.41*
ANC share	0.60	*0.29*	0.68	*0.28*	0.68	*0.26*	0.63	*0.24*
DA share	0.15	*0.23*	0.10	*0.19*	0.13	*0.22*	0.14	*0.23*
Other parties share	0.26	*0.28*	0.20	*0.24*	0.17	*0.21*	0.21	*0.17*
Turnout	0.51	*0.15*	0.53	*0.13*	0.57	*0.12*	0.57	*0.11*
N. of parties	4.56	*2.54*	6.57	*3.78*	7.96	*4.23*	9.92	*5.48*
N. of protests	0	*0.02*	0.001	*0.04*	0.003	*0.07*	0.019	*0.36*

2.5 Main results

2.5.1 Voter turnout and incumbent party vote share

Table 2.3 reports the results for voter turnout and the ANC share from the estimation of the Diff-in-Diff model described in equation (2.1). The coefficients of interest are those on the interaction terms ($3G \times 2016\,(2011)$). The table also reports the standalone coefficient on 3G, which captures average differences in the dependent variables between more and less covered localities at baseline (2006). For each outcome, a sensitivity analysis of the estimated coefficients across different specifications is reported in the table.

Column (1) of Table 2.3 shows that a unitary increase in coverage - i.e. moving from 0 to 100% of the area covered by the 3G - is associated with a 9.5 percentage point (pp) growth in voter turnout in 2016. The magnitude of the point estimate drops to 2.5 pp when socio-economic controls are included, yet the coefficient remains statistically significant at the 5% confidence level. This suggests that omitting relevant socio-economic characteristics, which are likely correlated with both 3G coverage and political

[23]The geographical entities in the picture correspond to the level at which standard errors are clustered in the regressions

outcomes, could yield upward-biased estimates. However, adding geographic controls as well as a polynomial function of a subset of them does not significantly alter the results. To some extent, this is reassuring of the fact that observable socio-economic variables account for most of the potential endogeneity of 3G in explaining voter turnout. Moreover, the effect of 3G on turnout remains positive and significant also in 2011. In the most conservative specification (columns 3-4) a unitary increase in 3G coverage leads to a 2.1 pp (1.6 pp) growth in turnout in 2016 (2011), which is approximately 4% (3%) of its mean in 2006.

To corroborate these findings, Figure 2.1 depicts the average trends in the outcomes of interest for two subgroups of localities: those with most of their area covered by 3G in 2016 - i.e. 3G$\geq$ 50% - and those with less than 50% of their area covered. To inspect their trends in the pre-internet era, I exclude localities that were covered by 3G in 2006 and re-scale the lines so that in 2006 the averages for the two subgroups exactly coincide. Moreover, to be consistent with the empirical analysis, I drop metropolitan areas and provincial capitals. The left panel of the figure supports the regression estimates: localities that are more exposed to 3G coverage exhibit higher voter turnout in 2011 and 2016, compared to localities with lower 3G coverage. Importantly, turnout does not show significant diverging trends across the two subgroups before the arrival of 3G. This provides visual evidence in support of the parallel trend assumption.

Columns (5-8) of Table 2.3 look at the vote share of the incumbent party. Similar to voter turnout, the coefficients of interest are sensitive to the inclusion of socio-economic variables, yet they remain stable after the addition of geographic controls and the polynomial function. The most conservative specification (column 7) suggests that a unitary increase in 3G coverage causes a 3.1 pp (4.5%) drop in the ANC vote share in 2016, while the impact in 2011 is still negative but not statistically significant. The right panel of Figure 2.1 provides supporting evidence for these estimates. Differently from turnout, the effect of 3G on the incumbent party is observable only in 2016. A potential reason, which is discussed more in details later, is that many administrative controversies and corruption scandals involving the ANC party emerged after 2011, and the availability of 3G exacerbated the socio-economic and political discontent resulting from them. Finally, similarly to turnout and in line with the parallel-trend assumption, the figure shows no significant diverging trends between the two subgroups of localities before 2006.

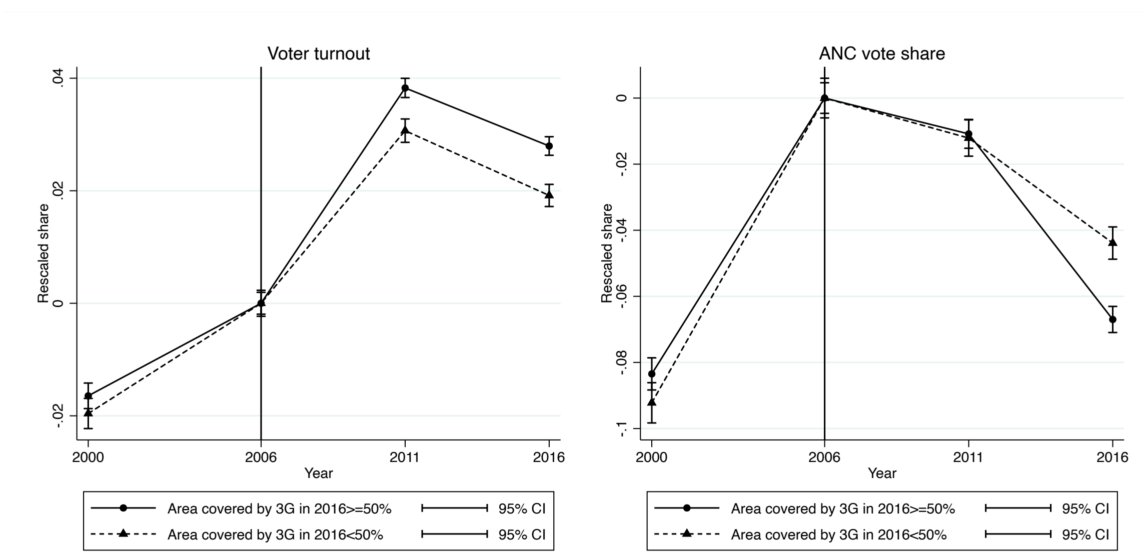
Table 2.3: Diff-in-Diff estimates of the impact of 3G on turnout and ANC vote share

OLS; years 2006 to 2016; whole sample

	Turnout				ANC share			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
3G*Year 2016	0.094***	0.025**	0.021**	0.021**	-0.099***	-0.035***	-0.031**	-0.035***
	(0.007)	(0.011)	(0.010)	(0.010)	(0.014)	(0.012)	(0.013)	(0.013)
3G*Year 2011	0.095***	0.021**	0.016**	0.016**	-0.024***	-0.006	-0.011	-0.013
	(0.006)	(0.008)	(0.008)	(0.008)	(0.007)	(0.010)	(0.010)	(0.010)
3G	-0.084***	-0.028***	-0.021**	-0.021**	0.037***	-0.003	0.002	0.004
	(0.006)	(0.009)	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)
Voting Dist FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Province trends	✓	✓	✓	✓	✓	✓	✓	✓
Socio-economic		✓	✓	✓		✓	✓	✓
Geographic			✓	✓			✓	✓
Polynomial				✓				✓
Mean in 2006	0.529	0.529	0.529	0.529	0.684	0.684	0.684	0.684
Observations	103925	103925	103925	103925	103901	103901	103901	103901
Adj. R-squared	0.455	0.507	0.514	0.514	0.734	0.753	0.756	0.757

Notes: 3G is the share of area covered by 3G signal. Provincial capitals and larger cities are excluded. Standard errors in parentheses clustered at the smallest stable aggregation of voting districts. Socio-economic variables include contemporaneous information on (log) luminosity, (log) population density, urbanization rate, years of schooling, ethnicity, age and share of youth, as well as their 2000 values interacted with year dummies. Geographic variables include presence of mine, road, river, slope index, (log) elevation, (log) ruggedness, (log) area, (log) distance to closest city and (log) distance to fiber backbone interacted with year dummies. Polynomial is a 5^{th}-order polynomial function of (log) ruggedness, (log) area and (log) population density interacted with year dummies. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

Figure 2.1: Trends in voter turnout and ANC vote share by 3G coverage



Notes: Provincial capitals and larger cities are excluded, as well as localities that were covered by 3G in 2006. Lines are re-scaled so that in 2006 the averages for the two subgroups exactly coincide and are equal to 0.

I complement these findings estimating model (2.2) via 2SLS. Table 2.4 shows the results. For each outcome, the first pair of columns reports OLS coefficients while the last pair provides the 2SLS estimates, with and without control variables. This analysis takes all four electoral waves into account (from 2000 to 2016) so to exploit the differential impact - around 2005 - that terrain ruggedness has on political outcomes through the arrival of the 3G technology. Coefficients on the instrumental variable from the 1^{st} stage regressions and the Wald F-statistics are also reported.

Table 2.4: IV estimates of the impact of 3G on turnout and ANC vote share

OLS and 2SLS; years 2000 to 2016; whole sample

	Turnout				ANC share			
	(1) OLS	(2) OLS	(3) 2SLS	(4) 2SLS	(5) OLS	(6) OLS	(7) 2SLS	(8) 2SLS
3G	0.014***	0.006**	0.074***	0.058**	-0.051***	-0.028***	-0.169***	-0.233***
	(0.003)	(0.003)	(0.019)	(0.028)	(0.008)	(0.006)	(0.042)	(0.058)
Voting Dist FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Controls		✓		✓		✓		✓
Mean in 2000	0.51	0.51	0.51	0.51	0.60	0.60	0.60	0.60
log(ruggedness)*Post			-0.09***	-0.08***			-0.09***	-0.08***
1st stage F-stat			136.2	177.2			135.8	176.9
Observations	138335	138335	138335	138335	138269	138269	138269	138269
Adj. R-squared	0.332	0.384	0.047	0.123	0.692	0.720	0.035	0.012

Notes: 3G is the share of area covered by 3G signal. Provincial capitals and larger cities are excluded. Standard errors in parentheses clustered at the smallest stable aggregation of voting districts. Controls include contemporaneous information on (log) luminosity and (log) population density, urbanization rate, years of schooling, age and share of youths, as well as presence of road and river, (log) area, (log) distance to closest city, (log) distance to fiber backbone interacted with year dummies. Coefficient on the instrumental variable from the first stage regression and the respective Kleibergen-Paap rk Wald F-statistic are reported. Post is a dummy which equals 1 if Year>2005. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

While the signs of the OLS and 2SLS coefficients are always internally consistent, their magnitudes differ substantially across estimators, with 2SLS coefficients being generally larger than the OLS ones, in the order of approximately 3-4 times. Columns (1-2) show an overall positive and significant relationship between 3G and turnout, ranging between 0.6 and 1.4 pp, depending on the specification. By contrast, the IV approach (columns 3-4) suggests that a unitary increase in 3G coverage leads to a 5.8-to-7.4 pp rise in voter turnout, a much larger estimated impact than the one predicted by the OLS. Similar patterns hold for the vote share of the incumbent party. Columns (5-6) point out a negative impact of 3G on the ANC vote share, in the order of -2.8 pp in the most conservative case. The 2SLS estimation yields larger coefficients, ranging between -17 and -23 pp.

The fact that the 2SLS estimates are larger than the OLS ones is common in the empirical literature, and there might be several explanations for this: (i) the OLS is biased, (ii) the instrument is not exogenous and so the 2SLS results are biased, (iii) there is measurement error in the independent variable and (iv) Average Treatment Effect (ATE) and Local Average Treatment Effect (LATE) do not coincide. Given the granularity of the data and their reliable source, measurement error in 3G coverage is improbable. Moreover, the exogeneity of the instrument is arguable, and I will show supporting evidence on this in the robustness section. Omitted variable bias in the OLS regressions is plausible. Nevertheless, the OLS estimates are also very aligned with the Diff-in-Diff results, providing a mutually reinforcing validation. Hence, the divergence in the estimates is most likely explained by the fact that the IV approach measures the LATE of 3G coverage, while the OLS regressions (as well as the Diff-in-Diff approach) capture the ATE, and the two do not coincide.

To put the Diff-in-Diff and IV estimates into perspective and reconcile their magnitudes, consider that average mobile internet coverage increased by 50 pp in the 2006-2016 period, while voter turnout increased by 4 pp and the ANC share dropped by 5 pp (Table 2.2), on average. The Diff-in-Diff estimates then imply that the expansion in 3G internet coverage accounted for about one fourth of the increase in voter turnout and one third of the vote swing against the ANC. By contrast, the 2SLS coefficients suggest that the rise in 3G coverage accounted for three fourths of the increase in turnout and for the entire loss in the ANC share during the same period. Hence, because the Diff-in-Diff results measure the ATE of 3G coverage, and this tends to be overall more conservative than the LATE - as estimated by the 2SLS procedure -, I consider the Diff-in-Diff approach to be the baseline specification of the empirical analysis.

2.5.2 Other political outcomes

I now describe the results for the vote share of the Democratic Alliance (DA) - the ANC's main political opponent - as well as the aggregate vote share of all remaining parties, the number of parties running for election and the number of protests against political institutions. I show the baseline results both with and without controls, and interpret the most conservative estimates.

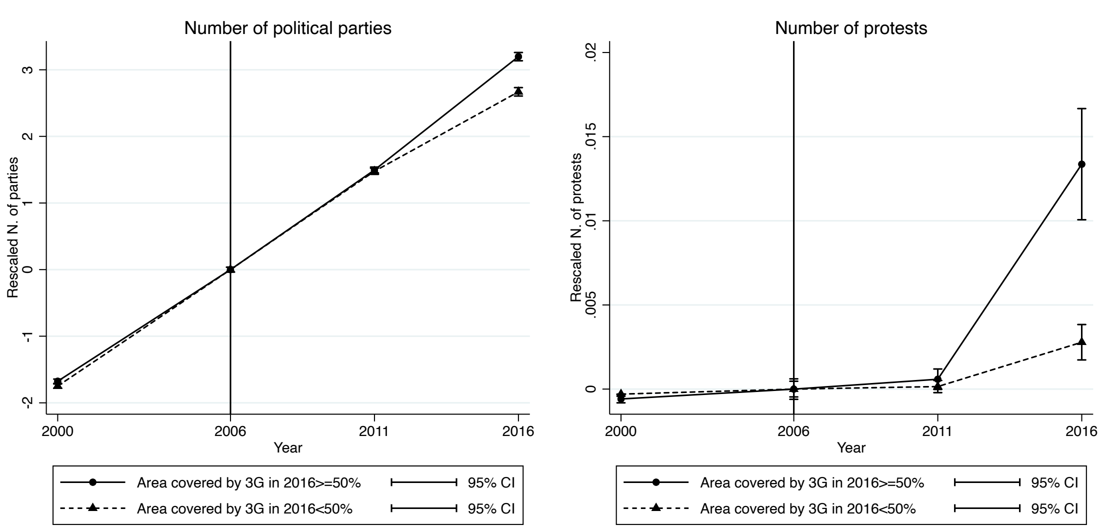
Table 2.5: Diff-in-Diff estimates of the impact of 3G on other political outcomes

OLS; years 2006 to 2016; whole sample

	DA share		Other parties share		N. of parties		N. of protests	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
3G*Year 2016	0.040***	0.027***	0.059***	0.004	1.681***	0.689**	0.077***	0.033***
	(0.007)	(0.007)	(0.014)	(0.013)	(0.304)	(0.293)	(0.014)	(0.010)
3G*Year 2011	0.050***	0.022***	-0.021***	-0.010	0.314	0.553**	0.027***	0.015**
	(0.006)	(0.006)	(0.007)	(0.009)	(0.218)	(0.248)	(0.007)	(0.006)
3G	-0.039***	-0.020***	0.004	0.020**	-1.022***	-0.318	-0.062***	-0.031***
	(0.006)	(0.006)	(0.008)	(0.009)	(0.253)	(0.251)	(0.014)	(0.011)
Voting Dist FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Province trends	✓	✓	✓	✓	✓	✓	✓	✓
Socio-economic		✓		✓		✓		✓
Geographic		✓		✓		✓		✓
Mean in 2006	0.099	0.099	0.203	0.203	6.559	6.559	0.001	0.001
Observations	99033	99033	103321	103321	103925	103925	104292	104292
Adj. R-squared	0.837	0.847	0.686	0.717	0.836	0.853	0.129	0.132

Notes: 3G is the share of area covered by 3G signal. Provincial capitals and larger cities are excluded. Standard errors in parentheses clustered at the smallest stable aggregation of voting districts. Socio-economic variables include contemporaneous information on (log) luminosity, (log) population density, urbanization rate, years of schooling, ethnicity, age and share of youths, as well as their 2000 values interacted with year dummies. Geographic variables include presence of mine, road, river, slope index, (log) elevation, (log) ruggedness, (log) area, (log) distance to closest city and (log) distance to fiber backbone interacted with year dummies. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

Figure 2.2: Trends in number of parties and protests by 3G coverage



Notes: Provincial capitals and larger cities are excluded, as well as localities that were covered by 3G in 2006. Lines are re-scaled so that in 2006 the averages for the two subgroups exactly coincide and are equal to 0.

Column (2) of Table 2.5 shows that the arrival of 3G benefits the Democratic Alliance, whose vote share rises by almost 2.2 and 2.7 percentage points in 2011 and 2016, respectively, which correspond to about 22 and 27% of the mean in 2006. By contrast, column (4) shows that the effect of 3G on the aggregate share of other parties is much smaller and not statistically significant. Columns (5-6) investigate the effects of the 3G technology on the number of parties running for election as a proxy for political competition. The most conservative estimates suggest that mobile internet coverage intensifies competition, leading to an increase of about 8 and 10% in the number of parties in 2011 and 2016, with respect to their mean in 2006. The result suggests that the new technology may lower entry costs into politics, eventually promoting the proliferation of new parties or helping existing parties to run in new districts. Finally, columns (7-8) analyze the effect of 3G availability on protests against a typically political entity as a proxy for social and political discontent. The coefficients reported in column (8) indicate that in 2011 and 2016, 3G coverage causes the number of protests to rise by 15 and 30 times their baseline value in 2006. Figure 2.2 depicts the patterns for the last two outcomes and provides visual evidence on their parallel trends across more and less covered localities before 3G became available. Finally, IV estimates are reported in the appendix Table 2.B3. The 2SLS coefficients are generally larger and, in contrast to the Diff-in-Diff results, they show that the DA did not benefit from 3G coverage, while other smaller parties did.

To contextualize the Diff-in-Diff results presented in Table 2.5, consider that the DA vote share as well as the average number of parties and protests against political institutions increased by 4 pp, 3.3 and 0.02 in the 2006-2016 period, respectively. The estimates then imply that the rise in 3G coverage accounted for about one third of the DA vote gain, for 10% of the growth in the number of parties and for 75% of the upsurge in protests.

Overall, my findings on the decline in the ANC share and the simultaneous gain by the opponents are in line with those obtained by Miner (2015) on the decrease in popularity of the ruling party in Malaysia. However, in contrast to Miner (2015), I find empirical support for the hypothesis that internet coverage may in fact promote political participation in developing countries. In this respect, my findings partially reconcile with the theory by Campante et al. (2018), suggesting that in the long-run the internet can become a political tool to reach out and recruit new individuals by newly formed parties. More generally, my results on voter turnout are at odds with what established literature (Falck et al., 2014; Gavazza et al., 2019) has found in developed countries.[24] Finally, my estimates on political competition and protests reconcile with the findings of Manacorda and Tesei (2020), who show that traditional cellphone technologies are instrumental to mass mobilization during economic downturns in Africa.

2.5.3 Robustness checks

I test the robustness of my estimates to different estimators, clustering, observational units and samples. First, I use a Propensity Score Matching estimation in which I rely on socio-economic and geographic variables measured in 2000 to predict the probability of being sufficiently covered by 3G in the following years (i.e. having more than 50% of the area covered). I then compute the average treatment effects (ATE) of 3G by taking the average difference in the outcomes between covered and non covered localities with similar estimated probabilities (propensity score). Table 2.B4 shows the results. Magnitudes of the ATE are generally aligned but more attenuated than those from the Diff-in-Diff analysis, yet they remain statistically significant.

Second, I check the sensitivity of the standard errors to a different level of clustering. Particularly, I use municipal boundaries to cluster the errors and report the results in Table 2.B5. With the exception of the results on the number of parties, coefficients for all other outcomes remain statistically significant.

Third, I attempt to alleviate concerns about endogenous changes in boundaries and potential measurement error from the construction of the units of analysis. In fact, one potential issue is that 3G coverage may be correlated with strategic re-demarcation interventions, which, in turn, affect political outcomes. To investigate this, first, I replicate the baseline estimation only for those localities whose demarcations remained fairly stable over time. Particularly, I restrict the attention to those voting districts that experienced at maximum 3 changes in their borders between 2000 and 2016.[25] These represent around 8% of the sample. Results from Panel A of Table 2.B6 show that all estimated coefficients are larger than those from the baseline analysis. This points out that places where boundary modifications occurred more frequently - i.e. those where concerns on endogeneity are higher - are not biasing the results upwards.

[24] An exception is Rotesi (2019), who finds a positive impact of Twitter penetration on turnout in the US.

[25] I choose this threshold in order to have sufficient power (observations) for the tests. However, restricting the attention to districts that did not experience any change yields even larger coefficients.

Second, I replicate the estimation using larger observational units. In particular, I aggregate voting districts until their boundaries remain unchanged over time and use these aggregations as units of observations.[26] Panel B of Table 2.B6 shows that when these wider units are employed, estimated coefficients are consistently larger than those from the baseline analysis.

Fourth, I assess the sensitivity to the use of different samples. In particular, Table 2.B7 shows the results including provincial capitals and larger cities in the sample. In this case, coefficients of interest are very similar to baseline estimates. Moreover, Table 2.B8 reports the estimates including year 2000 in the sample, which therefore becomes the baseline year. In this case, coefficients on the interactions of 3G with the 2011 and 2016 dummies generally confirm the baseline results. By contrast, coefficients on the interaction of 3G with year 2006 provide insights on the early effects of mobile internet on politics, yet these estimates tend to be at odds with the longer-term effects and should be interpreted with caution. The potential caveat is that very selected localities were covered by 3G in 2006. Most likely, these were cities and their surrounding areas, whose political outcomes could have changed regardless of the arrival of mobile internet. Unfortunately, given the absence of electoral data before 2000, it is impossible to conduct a placebo exercise or show pre-3G trends in the outcomes for these early-covered localities. For this reason, the benchmark Diff-in-Diff analysis considers 2006 as the baseline year, without attributing a causal meaning to the standalone coefficient on 3G.

Finally, with regards to the IV approach, I analyze its robustness to the use of a different explanatory variable. Specifically, I exploit the same instrumental variable to predict the number of years (since 2005) in which each district had more than 50% of its land covered by 3G signal. Table 2.B9 shows the results, which are closely aligned to those presented above.

2.5.4 Validation of identifying assumptions

I perform a number of exercises to check the plausibility of the assumptions underlying the identification strategies. Figures 2.1 and 2.2 provide some reassurance on the validity of the parallel trend assumption of the Diff-in-Diff approach. More formally, I conduct placebo exercises where I regress all electoral outcomes in the years before the expansion

[26] These are the same units that have been used to cluster the standard errors in the baseline specification.

of the mobile technology (2000-2006) on the leads of 3G coverage (2006-2016). If the parallel trend assumption holds, then one should find no effects of future 3G on past political variables. Table 2.B10 provides evidence in favor of this assumption: for all the outcomes, coefficients on the leads of 3G interacted with the time dummy are never statistically different from zero. This evidence mitigates concerns on the presence of divergent trends in political outcomes in the pre-internet era across localities that will end up being more and less covered by 3G afterwards.

Then, I perform a test following the procedure developed by Altonji et al. (2005). This strategy is useful in cases in which doubt remains about the exogeneity of the treatment variable. The approach involves two steps. First, I regress 3G coverage on a bunch of potentially relevant predictors (socio-economic and geographic variables) and then I calculate its fitted values. Second, I regress political outcomes between 2006 and 2016 on the predicted 3G coverage and its interaction with time dummies to assess the extent to which its plausibly endogenous component may affect these outcomes. Table 2.B11 shows the results from the second step.[27] The coefficients on predicted 3G interacted with time dummies have opposite signs with respect to those from the baseline analysis. Hence, the procedure suggests that the component of 3G coverage explained by the observables is not driving the baseline estimates, alleviating the concern that omitted variables could bias the results upwards in the baseline specification.

Finally, I conduct a placebo exercise to assess the plausibility of the exogeneity of the instrument in the IV approach. In particular, a reasonable claim is that ruggedness interacted with time may affect political outcomes through different channels, other than 3G coverage. To investigate this, I consider localities with no 3G in 2006 - so to rule out the effect of mobile internet - and regress the outcomes of interest between 2000 and 2006 on $log(Ruggedness) * 2006$. For the exogeneity assumption to be valid, the coefficient should not be statistically different from zero. Table 2.B12 provides evidence supporting the exogeneity of the instrument, as ruggedness does not predict any change in political outcomes. The result partially invalidates the possibility that terrain ruggedness affects politics through different channels, other than 3G.

[27] The output of the first step is shown in Table 2.B1.

2.6 Mechanisms

To provide suggestive evidence on the potential mechanisms, I follow Manacorda and Tesei (2020) and investigate the role of information and coordination as possible drivers of the observed outcomes. Empirically, I exploit additional administrative and survey data, and rely on a triple-difference estimation where, depending on the channel under investigation, specific variables will be interacted with 3G coverage. In particular, I estimate the following regression:

$$\begin{aligned} y_{it} = \; & \beta_0\, 3\mathrm{G}_{it} + \beta_1\, 3\mathrm{G}_{it} \times \theta_t + \mathbf{x}_{it}^{'}\beta_2 + (\mathbf{w}_i \times \theta_t)^{'}\beta_3 + \\ & \beta_4\, 3\mathrm{G}_{it} \times \theta_t \times E_i + \beta_5\, 3\mathrm{G}_{it} \times E_i + \beta_6 \theta_t \times E_i + \mu_i + \theta_t + \varepsilon_{it} \end{aligned} \tag{2.3}$$

where the notation used in section 2.3.2 applies, 3G is the share of area of voting sub-district i covered by 3G and E is a time invariant variable that represents the level of exposure of i to certain socio-economic conditions, depending on the specific channel analyzed. Note that standalone E is omitted from equation (2.3) because of the inclusion of voting sub-district fixed effects. The main coefficient of interest is β_4, the one associated to the triple-interaction term.

2.6.1 Information

In order to investigate the information role of mobile internet, I exploit the spatial variation in the probability of being exposed to the various administrative controversies described in section 2.2.1. Specifically, I take advantage of two sources of variation. First, I use differences in irregular expenditure across municipalities as a proxy for the quality of the administration of the municipal finances. In places where this expenditure is higher, voters have more reasons for political grievance as they attribute it to mismanagement of public funds (Dlomo, 2017). Second, I use variation in the economic relevance of the mining industry as a proxy for the exposure to mining strikes and the subsequent socio-economic turmoil. Localities where mining plays a significant economic role were more likely to be affected by the long-lasting negative consequences of the strikes. In both cases, most of the times the ANC party was considered accountable for the inappropriate administration. Although mismanagement of public funds and mining strikes are only two potential sources of political discontent, if mobile internet is used as a source of information, one should expect the 3G technology to play a bigger role in eroding the ANC popularity in places more exposed to these events.

Table 2.6: The information channel

Panel A: Diff-in-Diff-in-Diff, years 2011 to 2016; whole sample

	ANC share	DA share	Other parties	Turnout	N. of parties	N. of protests
	(1)	(2)	(3)	(4)	(5)	(6)
Irregular Exp*3G*2016	-0.009***	-0.001	0.009**	0.001	0.231**	0.003
	(0.003)	(0.002)	(0.004)	(0.001)	(0.114)	(0.003)
Total Exp*3G*2016	0.011	0.009*	-0.014	-0.008*	-0.100	-0.019**
	(0.011)	(0.005)	(0.011)	(0.005)	(0.324)	(0.009)
Voting Dist FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Province trends	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Observations	69488	67658	68730	69528	69528	69528
Adj. R-squared	0.793	0.903	0.705	0.602	0.870	0.211

Notes: 3G is the share of area covered by 3G signal. Provincial capitals and larger cities are excluded. Standard errors in parentheses clustered at municipality level. Irregular Exp is the log of total irregular and unauthorized municipal expenditure per capita from 2012 to 2015. Total Exp is the log of income and capital municipal expenditure per capita from 2012 to 2015. All other interactions are included but not reported for conciseness. Controls of column (4) of Table 2.3 are included here. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

Panel B: Diff-in-Diff-in-Diff, years 2011 to 2016, whole sample

	ANC share	DA share	Other parties	Turnout	N. of parties	N. of protests
	(1)	(2)	(3)	(4)	(5)	(6)
Mining GVA*3G*2016	-0.030***	0.022*	0.007	-0.021*	-0.178	0.012
	(0.011)	(0.013)	(0.010)	(0.011)	(0.328)	(0.014)
GVA*3G*2016	0.001	0.001	-0.002**	-0.002**	0.024	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.017)	(0.002)
Voting Dist FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Province trends	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Observations	69488	67658	68730	69528	69528	69528
Adj. R-squared	0.792	0.903	0.704	0.603	0.866	0.211

Notes: 3G is the share of area covered by 3G signal. Provincial capitals and larger cities are excluded. Standard errors in parentheses clustered at the smallest stable aggregation of voting districts. Mining is the per capita Gross Value Added from mining and quarrying of the locality. GVA is the per capita Gross Value Added from all sectors. All other interactions are included but not reported for conciseness. Controls of column (4) of Table 2.3 are included here. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

Empirically, I consider equation (2.3) and replace E with the exposure to either mismanagement of public funds or mining strikes, alternatively.[28] In the first case, E is

[28]In this case I focus only on the 2011 and 2016 electoral waves. In fact, data on municipal expenditure are available for years after 2011, and information on the gross value added (GVA) from the mining sector is measured in 2009. At the same time, it is true that mining accidents and scandals occurred much more frequently after 2011. Therefore, in equation (2.3), θ_t is a dummy which equals 1 if the

the sum of irregular and unauthorized audited municipal expenditure per capita over the 2012-2015 period. In the second case, E is the per capita gross value added (GVA) from mining and quarrying, for each voting sub-district. I also include the per capita aggregate municipal expenditure and GVA as controls, respectively. A negative and significant β_4 when y is the ANC share would provide suggestive evidence in favor of the information channel.

Table 2.6 shows the results. Column (1) of panel A suggests that, among localities with a higher level of irregular expenditure, a unitary increase in coverage causes an additional reduction in the votes for the ANC party of almost 1 percentage point.[29] At the same time, columns (2-3) show that the main gainers from the poor financial administration when 3G is present are smaller and newly formed parties as opposed to the DA, whose vote share is not significantly affected. Finally, while columns (4) and (6) indicate that political participation and protests do not change, column (5) shows that political competition intensifies, as the number of parties running for election increases further.[30]

Panel B of Table 2.6 looks at the interaction of 3G with variation in per-capita GVA from mining.[31] Column (1) shows that in localities where the mining industry is more prominent, a unitary increase in 3G coverage causes an additional 3 pp reduction in the ANC vote share. Differently from the case of municipal financial irregularities, socio-economic turmoil following the mining strikes favor the second biggest party, i.e. the DA, whose vote share increases by an additional 2.2 pp in mining sites more covered by 3G. This may suggest that individuals are more likely to trust consolidated parties when they face economic downturn, while they prefer new or smaller political entities when mismanagement of public funds is seen as a major problem. Finally, the number of parties and protests remains unchanged, yet overall disaffection towards politics grows, as voter turnout further decreases by 2 pp in localities exposed to mining activities when 3G is present.

In line with the findings by Ferraz and Finan (2008), the above analysis suggests that a

year is 2016, meaning that the baseline is set in 2011.

[29] If one exclusively focuses on municipalities where the ANC was ruling (which are anyway the majority) the magnitude of the coefficient does not significantly vary.

[30] Appendix Table 2.B13 shows that localities with low traditional media (radio and TV) penetration drive the negative effect of 3G on the ANC share and the gain of other parties.

[31] Figure 2.C6 shows the geographical dispersion of mining employment and GVA over the country in 2009.

segment of the population is informed through mobile internet about the unsatisfactory administration of public funds and socio-economic issues. Exposure to these facts makes individuals more likely to change their political opinions and subsequent behavior at ballots. The incumbent party is consistently damaged, while different opponents gain depending on the source of the event. Similarly, political participation and electoral competition react differently.

I complement this analysis with descriptive evidence from survey data and information on online activity from Google Trends. First, I use individuals' answers from the 2015 Afrobarometer survey to show how opinions towards the President, the ruling party and the political process differ between internet users non-users. One potential caveat is that self-reported internet usage is likely to be correlated with unobserved individual characteristics that directly affect political opinions. To partially address this concern, I combine individuals' exact location from the survey[32] with information on whether the respondent is covered by 3G or not, and instrument self-reported internet usage with 3G signal reception.[33] I run regressions both in reduced-form and implementing the 2SLS estimation. In both cases, I control for a variety of socio-economic individual-level covariates and province fixed effects.

Table 2.7 shows that general affection towards politics is not influenced by internet usage, while ANC affection is negatively associated with it. Moreover, internet users are less likely to vote for the ANC and more likely to distrust the ruling party. They are also more likely to think that most of the people in the President's office are involved in corruption, that opposition parties are silenced by the incumbent and that elections are unfair. Figure 2.C7 combines different survey waves to show that divergence in opinions (on trust and corruption) across individuals with and without 3G access enlarges over time. This is consistent with the fact that many corruption scandals in politics emerged in more recent years and a share of individuals could be exposed to them through mobile internet.[34]

Finally, I follow Gonzales et al. (2019) and exploit trends in search terms such as "news" and "corruption" relative to a set of similarly common words in the Google search

[32]In particular, I use GPS coordinates that I obtained from Afrobarometer upon request.

[33]Although 3G coverage is not exogenous, the variation that this variable provides at the individual level is preferable than self-reported measures of usage or access, as done by Guriev et al. (2019).

[34]Additional evidence from Afrobarometer reported in Table 2.B14 points out that frequent internet usage is indeed positively correlated with a more intense access to information online and through social media platforms.

Table 2.7: Internet use and political opinions (Afrobarometer, 2015)

	Any party affection			ANC affection			Vote for ANC		
	OLS	OLS	2SLS	OLS	OLS	2SLS	OLS	OLS	2SLS
Internet user	-0.04		-0.31	-0.09**		-0.55*	-0.10***		-0.63**
	(0.02)		(0.27)	(0.04)		(0.31)	(0.04)		(0.32)
Covered by 3G		-0.04			-0.06*			-0.08**	
		(0.03)			(0.03)			(0.03)	
Wald F-stat			19.4			18.3			19.0
Observations	2349	2349	2349	2349	2349	2349	2360	2360	2360

	Distrust president			Distrust ANC			Disapprove president		
	OLS	OLS	2SLS	OLS	OLS	2SLS	OLS	OLS	2SLS
Internet user	0.07**		0.30	0.07		0.57*	0.11***		0.52
	(0.03)		(0.33)	(0.04)		(0.30)	(0.03)		(0.36)
Covered by 3G		0.04			0.07*			0.05	
		(0.04)			(0.04)			(0.04)	
Wald F-stat			19.5			19.3			17.7
Observations	2345	2345	2345	2339	2339	2339	2313	2313	2313

	President is corrupt			Opposition is silenced			Elections are unfair		
	OLS	OLS	2SLS	OLS	OLS	2SLS	OLS	OLS	2SLS
Internet user	0.05*		0.77**	0.04		0.91***	0.01		0.51**
	(0.03)		(0.30)	(0.04)		(0.34)	(0.03)		(0.25)
Covered by 3G		0.08**			0.09**			0.06**	
		(0.04)			(0.04)			(0.03)	
Wald F-stat			19.1			17.6			19.2
Observations	2313	2313	2313	2166	2166	2166	2196	2196	2196

Notes: Standard errors in parentheses clustered at the village level. Internet user is a dummy which takes value 1 if the respondent reported to use the internet for any purpose at least a few times a month. Covered by 3G is a dummy which takes value 1 if the respondent's GPS location was reached by 3G signal at the beginning of 2015. In the 2SLS estimation this variable is used to instrument internet usage. The Kleibergen-Paap rk Wald F statistic is reported. All regressions include province fixed effects as well as individual-level controls for TV, radio and newspaper use to access information, occupational status, age, education, religion, distance from closest provincial capital/major city. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

engine to provide evidence on how online news consumption changed over the last 20 years. Figure 2.C8 shows that online searches for the general term "news" increased over time, especially after the municipal elections that took place in 2011. At the same time, Google queries for the specific word "corruption" became more frequent in certain periods around the 2014 national elections and in some of the subsequent months, reflecting the upsurge in political scandals involving the ANC party. All together, these findings provide additional evidence in support of the information channel.

2.6.2 Coordination

Another mechanism through which mobile internet could affect politics is the reduction in communication costs among users, which may ultimately facilitate new-party formation. More specifically, the 3G technology can be used by potential candidates as a coordination device to mobilize and communicate with voters at low cost. This can make barriers to entry into politics easier to overcome, especially for parties that lack significant financial resources.[35] In this respect, one should expect to find a higher number of newly formed parties running in places where 3G is present.[36] To empirically investigate this, I construct new dependent variables that allow me to distinguish whether the increase in political competition discussed in section 2.5.2 is driven by new parties or old coalitions running in new places. To do so, I consider a party to be new in years 2006 to 2016 if it did not run in any voting districts during the 2000 municipal election and replicate the Diff-in-Diff and IV estimations.

Columns (1-3) of Table 2.8 show a positive impact of mobile internet on the number of new parties running for elections, in the order of about 1 new entrant in the most conservative specification. By contrast, columns (4-6) indicate that old parties are less likely to stand for elections in places more covered by 3G. Figure 2.C9 corroborates the Diff-in-Diff estimates and provides visual evidence in support of the parallel trend assumption. These results suggest that, by reducing coordination and communication costs, the mobile technology can facilitate political turnover in developing countries, a finding that is in line with the results of Campante et al. (2018) for Italy.

Finally, I investigate the heterogeneity of these estimates across different local situations that may facilitate the formation of new political parties. For instance, when deciding whether to step in or not, entrants may take into account the historical level of political engagement of a locality or the adoption of traditional technologies such as cellphones. Moreover, potential entrants expect higher returns from running in places where the reputation of established parties is questioned by the local community. If mobile internet is used as a communication tool, then online mobilization campaigns should be more effective in these places, eventually enhancing new party formation even further. To test

[35] For instance, Petrova et al. (2021) show that joining Twitter helps politicians in the US raise political donations.

[36] For instance, as discussed by Malherbe (2015), during the 2014 national election campaign, the newly formed party EFF had a much more integrated and deliberate online communication strategy than the ANC. The EFF was also active when it came to building online communities, which are an essential novel tool for political mobilization.

Table 2.8: The coordination channel

Diff-in-Diff (2006 to 2016), OLS and 2SLS (2000 to 2016), whole sample

	N. of new parties			N. of old parties		
	(1) Diff-in-Diff	(2) OLS	(3) 2SLS	(4) Diff-in-Diff	(5) OLS	(6) 2SLS
3G*Year 2016	1.076***			-0.387**		
	(0.293)			(0.166)		
3G*Year 2011	0.898***			-0.345*		
	(0.188)			(0.197)		
3G	-0.672***	0.972***	3.823***	0.354**	-0.289***	-0.681
	(0.232)	(0.115)	(0.612)	(0.161)	(0.088)	(0.494)
Voting Dist FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Province trends	✓			✓		
Controls	✓	✓	✓	✓	✓	✓
1st stage Wald F-stat			178.5			178.5
Observations	103925	139227	139227	103925	139227	139227
Adj. R-squared	0.793	0.715	0.610	0.777	0.793	0.188

Notes: 3G is the share of area covered by 3G signal. Provincial capitals and larger cities are excluded. Standard errors in parentheses clustered at the smallest stable aggregation of voting districts. Controls of Table 2.5 are included in columns (1) and (4), while controls of Table 2.B3 are included in columns (2-3) and (5-6). For every voting district-year, new and old parties are defined with respect to those parties existing in year 2000. A party is considered to be new if it did not run in any districts during the 2000 municipal elections. In columns (3) and (6) the instrumental variable is (log) ruggedness*Post, where the latter is a dummy which equals 1 if Year>2005. The associated Kleibergen-Paap rk Wald F-statistic is reported. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

this, I use a triple difference estimation as described in equation (2.3), where E now reflects pre-determined (i.e. measured in 2000) specific local circumstances that may favor new party formation. These are voter turnout, cellphone penetration, the share of votes to ANC opponents, as well as a measure of individuals' ethnic affinity with the ANC leader Jacob Zuma as a proxy for political grievance.[37]

Table 2.9 provides the results. Columns (1-2) point out that in places where political participation and cellphone penetration rates are historically larger (above the median in 2000), 3G availability plays an additional role in encouraging new parties to stand for elections.[38] By contrast, columns (5-6) indicate that the number of established parties

[37]Zuma belongs to the Zulu ethnicity. During his political career, the leader extensively exploited his Zulu traditions to mobilize voters and gain local support. However, more recently he took advantage of his ethnic origins to cover-up poor personal choices, indiscretions and wrong behavior ("Jacob Zuma President of South Africa". Encyclopedia Britannica. Last update: 19 March 2018).

[38]Interacting 3G coverage with cellphone penetration allows to disentangle the role of new media (e.g., social networks) vs. traditional ICTs (e.g., SMSs) in facilitating coordination and communication.

Table 2.9: Heterogeneous results on party formation

Diff-in-Diff-in-Diff, years 2006 to 2016, whole sample

	N. of new parties				N. of old parties			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High turnout*3G*2016	0.68**				-0.22			
	(0.28)				(0.14)			
High turnout*2016	-0.32**				0.08			
	(0.14)				(0.06)			
High cellphone*3G*2016		0.42**				-0.10		
		(0.19)				(0.13)		
High cellphone*2016		-0.18				0.03		
		(0.12)				(0.06)		
%Anti-ANC*3G*2016			4.69***				-1.19***	
			(0.56)				(0.36)	
%Anti-ANC*2016			-3.01***				0.26	
			(0.40)				(0.24)	
Zulu*3G*2016				3.26***				-1.76***
				(0.97)				(0.37)
Zulu*2016				-2.04***				0.22
				(0.41)				(0.20)
Voting Dist FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Province trends	✓	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Observations	103086	103925	102960	103925	103086	103925	102960	103925
Adj. R-squared	0.80	0.79	0.81	0.80	0.78	0.78	0.78	0.78

Notes: 3G is the share of area covered by 3G signal. Provincial capitals and larger cities are excluded. Standard errors in parentheses clustered at the smallest stable aggregation of voting districts. High turnout is a dummy variable which equals 1 if the voting district's turnout in 2000 was above the median. High cellphone is a dummy variable which equals 1 if the cellphone penetration in the district in 2000 was above the median. %Anti-ANC is the aggregate share of ANC opponents in 2000. Zulu is a dummy which equals 1 if Zulu is the main ethnic group of the voting district. All other interactions are included but not reported for conciseness. Controls of Table 2.5 are included here. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

is not further affected by these circumstances. Finally, columns (3-4) show that new parties are more likely to run for elections if mobile internet is available in places where the support for the ANC is lower. In fact, coefficients on the interaction of 3G with the share of anti-ANC votes and the indicator for the Zulu ethnicity - President Zuma's ethnic group - are positive and statistically significant. By contrast, columns (7-8) show that opposite results hold for established coalitions, which are more likely to abandon the political arena when mobile coverage is present and they are more likely to face unfavorable circumstances. Overall, these results provide additional evidence on the role of the 3G technology in reducing communication and coordination costs especially

In fact, cellphone penetration is measured in 2000, when no smartphone was available. In this respect, results in column (2) of Table 2.9 suggest that, in the absence of 3G, standard cellphone availability is not sufficient to trigger new party formation.

in places where the political context is more favorable to new party formation.

2.6.3 Additional channels

Other mechanisms could be in place. For instance, 3G access can impact on political outcomes through the availability of new monitoring technologies at the voting stations and their effect on vote buying. Although I cannot completely rule out this hypothesis, self-reported voting intentions described in Table 2.7 showed that ANC support was much smaller among internet users. This fact should partially invalidate the proposed channel because voting intentions are independent of possible monitoring at the voting stations.

Another possibility is that the rise of mobile internet might result in better investigative journalism. This, in turn, may spill over to other traditional media outlets and affect their quality also in localities with low internet coverage. Insufficient data availability does not allow me to assess whether the overall quality of TV news programs or newspapers improved over time. However, descriptive evidence from surveys might help. If traditional media benefited from improved investigative journalism, then opinions towards the President across traditional media users and non-users should diverge over time also as a result of higher quality. Figure 2.C10 shows patterns of political opinions (on trust and corruption) for TV and newspapers users and non-users in localities not covered by 3G. Individuals' opinions do not exhibit significant divergent trends over time, providing no evidence to support the hypothesized channel.

Finally, mobile internet could affect economic development, for instance through better labor market matching or agglomeration economies and, in turn, this could explain changes in voting behavior. My data do not allow to directly test for this channel.[39] However, because of labor mobility, better economic outcomes in certain localities would attract more workers, thus resulting in larger population. By contrast, results form Table 2.B15 provide no evidence on the impact of 3G on population density. This might indicate that any change in economic development due to mobile internet arrival is not substantial, thus partially invalidating the proposed channel.[40]

[39]Indeed, luminosity is missing for recent years and I linearly extrapolated the information for 2016 to use it as a control variable.

[40]Alternatively, individuals might still live in low-coverage regions but commute for working in 3G-covered and more developed areas. However, this would most likely imply that my estimates represent a lower bound of the true effect.

2.7 Conclusions

In this paper, I analyzed the impact of mobile internet coverage on political outcomes looking at the South African municipal election results between 2000 and 2016. I mitigated concerns on potential endogeneity exploiting a newly constructed high-resolution dataset along with two alternative empirical approaches (Diff-in-Diff and IV).

My findings demonstrated that the arrival of fast mobile internet led to a decline in the popularity of the incumbent party (-3 pp), and the simultaneous gain of its rivals. These results were accompanied by increased political participation and electoral competition. In fact, 3G coverage positively affected voter turnout (2 pp) and the number of parties running for election (10%). Furthermore, 3G-covered localities experienced an upsurge in the number of riots and protests against political institutions. These findings are robust to various model specifications, as well as to different estimators and samples.

The overall effect of mobile internet on political outcomes could be the consequence of increased discontent towards politics amplified by the exposure to online information, as well as the result of better coordination facilitated by lower communication costs. Thus, I complemented the main results with suggestive evidence on these proposed mechanisms. This analysis revealed that (1) the mobile technology provided voters with additional information and this led to a change in their political behavior; (2) the presence of 3G in a locality facilitated political turnover by lowering entry costs into politics for new political parties, as it most likely decreased communication costs with their potential voters.

My paper indicates that mobile internet might be an effective tool to monitor politicians' performance and promote political accountability and participation in developing countries. More generally, the increasing availability of 3G and 4G technologies suggest that the potential for digital ICTs to foster electoral competition and improve the democratic process in the developing world might be sustained in the long term. Taken together, the observed results offer an optimistic assessment of the possibility of political change through digital media in countries with historically fragile traditional media independence. However, understanding how politicians react to the new technology and whether they become more attentive to the voters' demands remains unclear and is left for future work.

2.8 Appendices

2.A Data

Construction of observational units

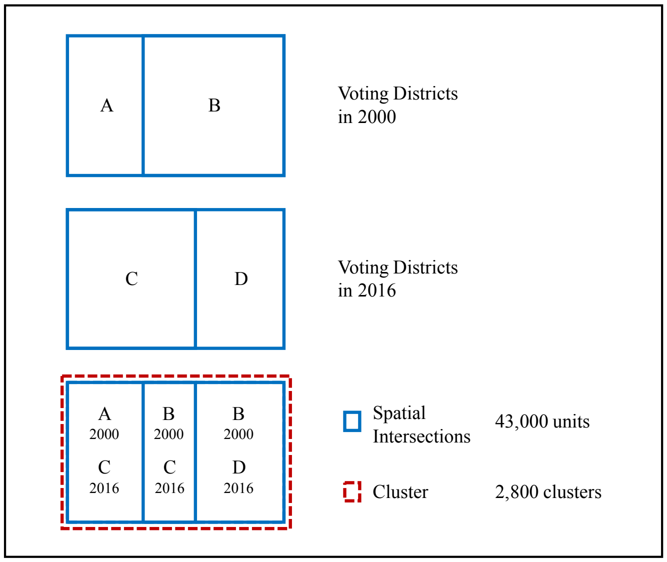
The goal of this procedure is to create stable (geographically and time invariant) units of observation. This has two purposes: first, it allows comparability of electoral outcomes over time; second, it solves the problem of endogenous change in the district boundaries. There are to possible ways of creating new units of observations.

The first method focuses on the union of all neighboring districts whose boundaries change over time: I use an algorithm that combine neighboring and mutable voting districts until their spacial union reaches a stable aggregation. The output provides new entities that represent the smallest aggregations of districts whose borders are constant between 2000 and 2016. These are approximately 3,800 clusters, much less than the initial number of voting districts.

The second method, instead, is based on the intersection of initial districts. The procedure is exemplified in Figure 2.A1. The spatial intersection of two initial voting districts whose common border changes over time gives birth to three new entities. For each entity/year I assign the political outcome of the voting district from which it originated, for the respective year. This procedure creates about 43,000 observations per year.

The two methodologies produce two distinct datasets, whose number and size of the observational units largely differ. The first method has the advantage of measuring the political outcome more accurately, as it is simply the sum of the outcome in the initial districts. However, the dimension of the new entities is critical. Most of them are large in size and, in turn, internet coverage ends up being imprecisely measured. In particular, for each large observation, it becomes impossible to distinguish who is covered and who is not. To some extent, in this case the granularity of my measure of 3G coverage becomes useless and misleading. Therefore, to better take advantage of the high-resolution of the data, I decide to focus on the second procedure. The 43,000 observations/year have much smaller dimensions, hence it becomes possible to precisely attribute average internet coverage to each of them. Then, to account for correlation in the errors induced by the intersecting procedure, I cluster the standard errors at the

Figure 2.A1: Construction of observational units



level of the smallest stable aggregation, as exemplified by the dashed line in Figure 2.A1. The procedure provides larger, hence more conservative, standard errors. At the same time, in order to alleviate the suspect of endogenous change in the borders, I replicate the baseline estimation only for those districts whose boundaries remained almost constant over time. For the sake of comparison, I also replicate the analysis using as units of observations the aggregations of voting districts (i.e. those produced by the first method described above). In both cases, estimated coefficients are larger or similar in magnitude to those from the baseline analysis.

Variables description and sources

Variables: Turnout, Vote shares, Vote margins, Number of parties, Population

Source: Independent Electoral Commission[41]

Description: Voting data contains information on the total number of registered voters, the number of those who actually voted, and the number of votes each party got in each voting district for all municipal elections (2000, 2006, 2011, 2016). I used the number of registered voters as a proxy for population for each voting district.

[41] http://www.elections.org.za

Variable: Mobile internet coverage

Source: Collins Mobile Coverage Explorer[42]

Description: The dataset comes in GIS vector format and for each country provides 2G, 3G and 4G coverage, separately: each pixel has value 1 if covered, 0 otherwise. In South Africa the geographical precision varies from year to year, with a maximum pixel size of 1km by 1km (up to almost 200m by 200m in the most recent version). I exploit only 3G coverage data since there is practically no variation in 2G or 4G technologies between 2007 and 2015: almost all places had 2G already before 2007, almost no place had 4G in 2015 (nor before). To proxy coverage in 2016 I use internet coverage in 2015, as this depicts the situation up to December of that year. Similarly, I use coverage in 2007 to proxy for coverage in 2006. Measurement error associated to this approximations should be small. On the one hand, municipal elections in 2016 were held on the 3rd of August. The assumption is that 3G coverage did not change abruptly in the months right before elections. If it did, my estimate would represent a lower-bound for the actual effect, as coverage likely expands over time. On the other hand, in the initial years (2007-2009), changes in coverage were extremely marginal. Therefore, is it reasonable to assume that infrastructure status between 2006 and 2007 does not differ significantly.

Variable: Luminosity

Source: National Centers for Environmental Information[43]

Description: Each pixel (1 square kilometer) in the luminosity data is assigned a digital number (DN) representing its luminosity. The DNs are integers that range from 0 to 63. The higher this number, the greater is the economic activity in the pixel. The reader may look at Pinkovskiy and Sala-i Martin (2016) for a recent application of this dataset as a proxy for GDP in Africa. Unfortunately, the last available year for the data is 2013. Hence, I predicted night-light for 2015/16. I exploited linear extrapolation by assuming for each observation a constant growth rate between 2011 and 2015, using the observed growth experienced between 2011 and 2013.

[42] http://www.collinsbartholomew.com/mobile-coverage-maps/mobile-coverage-explorer/
[43] http://www.ngdc.noaa.gov/eog/dmsp/downloadV4composites.html

Variable: Number of protests

Source: PRIO/Uppsala Armed Conflict Location and Event (ACLED)[44]

Description: The dataset provides information on a variety of characteristics for any kind of conflict event. I restrict the attention to unilateral actions perpetrated by rioters/protesters. This category of events includes demonstrations against a typically political entity, such as a government institution. The event is coded as involving protesters when it is non-violent, and as involving rioters if the demonstrators employ violence. However, I disregard this distinction: I merge the two categories and refer to them as protests. Moreover, I consider only events that took place in the election years and in the preceding one. So, for instance, for year 2000 I use events that happened in 1999 and 2000. Finally, for each locality/year, I compute the total sum of protests.

Variables: Age, Urbanization rate, Years of schooling, Phone/TV/Radio ownership

Source: 2000 and 2011 Census[45]

Description: Information comes at a very disaggregated unit of analysis, called Small Area level: in 2001 and 2011 there were approximately 56,000 and 85,000 of such Small Areas, respectively. I use zonal statistics to compute the average quantity of the variables of interest for each observational unit/year. Since information on 2006 and 2016 does not exist, I interpolate and extrapolate the quantities, respectively, so to create a balanced panel for each locality. Note that I use the 2001 census wave as a proxy for socio-economic characteristics in year 2000.

Variables: Total expenditure, Irregular expenditure

Source: Department of National Treasury[46]

Description: Total expenditure is the sum of municipal income and capital expenditure. For income, I use the Statement of Financial Performance: how a municipality has

[44]http://www.acleddata.com
[45]http://www.statssa.gov.za
[46]https://www.municipaldata.treasury.gov.za

spent money and received income; for capital, I use expenditure on purchase, repair and renewal of capital assets. I use the sum of irregular and unauthorized expenditure at the municipality level to proxy for the quality of local governance. These are specific amounts from audited financial results, recorded in the notes to the annual financial statements of each municipality. Data are available from 2012 to 2015. For each variable/municipality I construct aggregate measures by summing up the quantities over the years.

Variables: Various geographic information
Data on elevation and ruggedness (as measured by the standard deviation of the elevation) at high resolution come from the Global Multi-resolution Terrain Elevation Data 2010. The dataset is hosted by the Earth Resources Observation and Science (EROS).[47] Moreover, shapefiles containing information on major cities, main roads and waterways come from Open Street Map.[48] Data on mining presence come from the Mineral Resources Data System (MRDS).[49] Information on Gross Value Added from mining and quarrying in 2009 comes from the economic sector maps provided by Quantec and accessible online on the Geospatial Analysis Platform.[50] I follow Hjort and Poulsen (2019) to retrieve information on the terrestrial national fiber backbone infrastructure.[51] Finally, the terrain slope index is borrowed from Nunn and Puga (2012) and geographic data on major ethnic groups come from the GREG database.[52]

[47] https://www.topotools.cr.usgs.gov/gmted_viewer/
[48] https://www.openstreetmap.org
[49] https://www.mrdata.usgs.gov/mrds/
[50] https://www.gap.csir.co.za/download-maps-and-data
[51] https://www.africabandwidthmaps.com and https://afterfiber.nsrc.org
[52] https://www.icr.ethz.ch/data/greg/

2.B Tables

Table 2.B1: Determinants of 3G coverage

OLS; years 2000 to 2016; whole sample

	3G (share of area covered)	
log(Pop. Density)	0.006**	(0.002)
log(Nightlight)	0.010*	(0.006)
Urbanization	0.062***	(0.015)
Schooling	0.013***	(0.004)
Age	-0.004***	(0.001)
Youths share	-0.014	(0.061)
Mining*2006	0.008	(0.029)
Mining*2011	0.060**	(0.024)
Mining*2016	0.001	(0.020)
Road*2006	0.027***	(0.010)
Road*2011	0.023**	(0.011)
Road*2016	0.034***	(0.012)
Waterway*2006	0.020*	(0.011)
Waterway*2011	0.029**	(0.011)
Waterway*2016	-0.002	(0.017)
log(Elevation)*2006	0.007	(0.007)
log(Elevation)*2011	-0.006	(0.008)
log(Elevation)*2016	0.010	(0.008)
log(Ruggedness)*2006	-0.008**	(0.003)
log(Ruggedness)*2011	-0.014***	(0.005)
log(Ruggedness)*2016	-0.027***	(0.005)
Slope*2006	0.002*	(0.001)
Slope*2011	-0.002	(0.002)
Slope*2016	-0.009***	(0.002)
log(Distance to city)*2006	-0.048***	(0.012)
log(Distance to city)*2011	-0.036**	(0.014)
log(Distance to city)*2016	0.019	(0.015)
log(Area)*2006	0.014**	(0.005)
log(Area)*2011	0.025***	(0.006)
log(Area)*2016	0.005	(0.009)
log(Distance to fiber)*2006	0.007	(0.007)
log(Distance to fiber)*2011	-0.030***	(0.009)
log(Distance to fiber)*2016	-0.006	(0.010)
Phone (in 2000)*2006	0.272***	(0.064)
Phone (in 2000)*2011	0.237***	(0.051)
Phone (in 2000)*2016	0.023	(0.061)
Cellphone (in 2000)*2006	0.118***	(0.037)
Cellphone (in 2000)*2011	0.051	(0.047)
Cellphone (in 2000)*2016	0.160**	(0.068)
ANC share (in 2000)*2006	-0.070***	(0.016)
ANC share (in 2000)*2011	-0.031*	(0.019)
ANC share (in 2000)*2016	0.146***	(0.031)
Voting Dist FE	✓	
Year FE	✓	
Socio-economic trend	✓	
Observations	137472	
Adj. R-squared	0.667	

Notes: Provincial capitals and larger cities are excluded. Standard errors in parentheses clustered at the smallest stable aggregation of voting districts. Socio-economic trend includes (log) luminosity, (log) population density, years of schooling, ethnicity, age and share of youths measured in 2000 and interacted with time dummies. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

Table 2.B2: Descriptive statistics: final sample

Variable	Year	Obs	Mean	Median	Std. Dev.	Min	Max
3G (share of area covered)	2000	34988	0.00	0	0	0	0
3G (share of area covered)	2006	34988	0.10	0	0.29	0	1
3G (share of area covered)	2011	34988	0.22	0	0.38	0	1
3G (share of area covered)	2016	34988	0.60	0.71	0.41	0	1
ANC share	2000	34590	0.60	0.70	0.29	0	1
ANC share	2006	34615	0.68	0.80	0.28	0	1
ANC share	2011	34986	0.68	0.78	0.26	0	1
ANC share	2016	34970	0.63	0.69	0.24	0	1
DA share	2000	27789	0.15	0.04	0.23	0	1
DA share	2006	31187	0.10	0.02	0.19	0	1
DA share	2011	34071	0.13	0.03	0.22	0	1
DA share	2016	34970	0.14	0.03	0.23	0	1
Other parties share	2000	33071	0.26	0.13	0.28	0	1
Other parties share	2006	34437	0.20	0.10	0.24	0	1
Other parties share	2011	34604	0.17	0.09	0.21	0	.99
Other parties share	2016	34970	0.21	0.16	0.17	0	1
Turnout	2000	34632	0.51	0.51	0.15	0	1
Turnout	2006	34619	0.53	0.53	0.13	0	1
Turnout	2011	34988	0.57	0.58	0.12	0	1
Turnout	2016	34988	0.57	0.57	0.11	0	1
N. of parties	2000	34632	4.56	4	2.54	2	14
N. of parties	2006	34619	6.57	6	3.78	2	23
N. of parties	2011	34988	7.96	7	4.23	2	33
N. of parties	2016	34988	9.92	9	5.48	3	36
N. of protests	2000	34988	0.000	0	0.02	0	2
N. of protests	2006	34988	0.001	0	0.04	0	4
N. of protests	2011	34988	0.003	0	0.07	0	6
N. of protests	2016	34988	0.019	0	0.36	0	27
Luminosity	2000	34988	9.09	2.65	15.09	0	63
Pop. density	2000	34988	725.35	275.20	1001.46	.002	4797
Urbanization rate	2001	34988	0.12	0	0.30	0	1
Years of schooling	2001	34988	4.95	4.75	1.79	0	16
Age	2001	34988	26.28	25.80	5.55	0	82.5
Youth share (14<age<30)	2001	34988	0.27	0.27	0.08	0	1
Zulu main ethnic group		34764	0.21	0	0.41	0	1
Phone (share of households)	2001	34988	0.11	0.04	0.16	0	1
Cellphone (share of households)	2001	34988	0.20	0.16	0.16	0	1
TV (share of households)	2011	34988	0.45	0.45	0.28	0	1
Radio (share of households)	2011	34988	0.52	0.57	0.24	0	1
GVA from mining (Millions, Rpc)	2009	34988	0.04	0.00	0.40	0	28
Total GVA (Millions, Rpc)	2009	34988	0.47	0.08	1.86	0	110
Irregular expenditure (Rpc)	2012-2015	234	3961	1510	13701	0	204008
Total expenditure (Rpc)	2012-2015	234	18416	15866	15079	2257	100333
Elevation (m)		34988	921.62	976.08	469.29	0	2546
Ruggedness (m^2)		34988	38.79	23	45.39	0	551
Area (km^2)		34988	20.14	2.58	61.52	.25	654
Distance to city (km)		34988	123.05	112.13	74.61	15	646
Distance to fiber (km)		34988	11.24	7.68	12.37	0	218
Road (yes/no)		34988	0.11	0	0.31	0	1
River (yes/no)		34988	0.04	0	0.20	0	1
Mine (yes/no)		34988	0.01	0	0.10	0	1

Notes - Outcome variables are displayed for the years 2000, 2006, 2011 and 2016, while statistics for socio-economic variables refer to 2001. Moreover, the mean total Gross Value Added (GVA) and GVA from mining and quarrying are measured in 2009, while data on total and irregular expenditure come at the municipality level (234 municipalities) and refer to the 2012-2015 period. Population is proxied by the number of registered voters. Rpc stands for Rounds per capita.

Table 2.B3: IV estimates of the impact of 3G on other political outcomes

OLS and 2SLS; years 2000 to 2016; whole sample

	DA share			Other parties share		
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	2SLS	2SLS	OLS	2SLS	2SLS
3G	-0.000	-0.028	-0.026	0.030***	0.208***	0.270***
	(0.003)	(0.018)	(0.027)	(0.006)	(0.042)	(0.061)
Voting Dist FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Controls	✓		✓	✓		✓
Mean in 2000	0.15	0.15	0.15	0.26	0.26	0.26
log(ruggedness)*Post		-0.11***	-0.09***		-0.09***	-0.08***
1st stage F-stat		141.0	194.9		130.9	167.1
Observations	126609	126609	126609	136187	136187	136187
Adj. R-squared	0.805	0.037	0.053	0.716	-0.008	-0.065
	N. of parties			N. of protests		
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	2SLS	2SLS	OLS	2SLS	2SLS
3G	0.537***	3.687***	3.234***	-0.001	0.034***	0.030***
	(0.119)	(0.595)	(0.660)	(0.002)	(0.009)	(0.011)
Voting Dist FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Controls	✓		✓	✓		✓
Mean in 2000	4.55	4.55	4.55	0.00	0.00	0.00
log(ruggedness)*Post		-0.09***	-0.08***		-0.09***	-0.08***
1st stage F-stat		136.2	177.2		138.1	179.0
Observations	138335	138335	138335	139056	139056	139056
Adj. R-squared	0.842	0.521	0.571	0.091	0.000	0.004

Notes: 3G is the share of area covered by 3G signal. Provincial capitals and larger cities are excluded. Standard errors in parentheses clustered at the smallest stable aggregation of voting districts. Controls include contemporaneous information on (log) luminosity and (log) population density, urbanization rate, years of schooling, age and share of youths, as well as presence of road and river, (log) area, (log) distance to closest city and (log) distance to fiber backbone interacted with year dummies. Coefficient on the instrumental variable from the first stage regression and the respective Kleibergen-Paap rk Wald F-statistic are reported. Post is a dummy which equals 1 if Year>2005. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

Table 2.B4: Average Treatment Effects of 3G on political outcomes

Propensity Score Matching; years 2006 to 2016; whole sample						
	ANC share (1)	DA share (2)	Other parties (3)	Turnout (4)	N. of parties (5)	N. of protests (6)
ATE in 2016						
Covered vs. Not	-0.014***	0.006**	0.006	0.007*	0.246**	0.015***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.107)	(0.002)
ATE in 2011						
Covered vs. Not	-0.033***	0.015***	0.019***	0.006	-0.058	0.002**
	(0.010)	(0.004)	(0.007)	(0.006)	(0.064)	(0.001)
Observations	34219	30424	33698	34224	34224	34368

Notes: Provincial capitals and larger cities are excluded. Robust Abadie-Imbens standard errors in parentheses. A locality is considered Covered in a given year if the share of area covered by 3G$\geq$0.5 in that year. In all columns the propensity of being covered is estimated using (log) luminosity, (log) population density, urbanization rate, years of schooling, ethnicity, age, share of youths and ANC vote share all measured in 2000; as well as an indicator variable for the presence of a road, (log) elevation, (log) ruggedness, (log) area and (log) distance to closest city.* Significant at 10%, ** significant at 5%, and *** significant at 1%.

Table 2.B5: Diff-in-Diff estimates with standard errors clustered at municipality level

OLS; years 2006 to 2016; whole sample						
	ANC share (1)	DA share (2)	Other parties (3)	Turnout (4)	N. of parties (5)	N. of protests (6)
3G*Year 2016	-0.031*	0.027**	0.004	0.021*	0.689	0.033***
	(0.018)	(0.011)	(0.019)	(0.011)	(0.661)	(0.011)
3G*Year 2011	-0.011	0.022**	-0.010	0.016	0.553	0.015**
	(0.015)	(0.010)	(0.015)	(0.011)	(0.585)	(0.006)
3G	0.002	-0.020*	0.020	-0.021**	-0.318	-0.031***
	(0.014)	(0.011)	(0.015)	(0.009)	(0.612)	(0.012)
Voting Dist FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Province trends	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Observations	103901	99562	103344	103925	103925	104292
Adj. R-squared	0.756	0.847	0.717	0.514	0.853	0.132

Notes: 3G is the share of area covered by 3G signal. Provincial capitals and larger cities are excluded. Standard errors in parentheses clustered at the municipality level. Controls include: contemporaneous information on (log) luminosity and (log) population density, urbanization rate, years of schooling, ethnicity, age and share of youths, as well as their 2000 values interacted with year dummies, indicators variables for the presence of mine, road and river, slope index, (log) elevation, (log) ruggedness, (log) area, (log) distance to closest city and (log) distance to fiber backbone interacted with year dummies. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

Table 2.B6: Sensitivity of the Diff-in-Diff estimates to different observational units

Panel A: Results using voting districts with mostly stable demarcations; OLS; years 2006 to 2016

	ANC share (1)	DA share (2)	Other parties (3)	Turnout (4)	N. of parties (5)	N. of protests (6)
3G*Year 2016	-0.043**	0.027*	0.010	0.022**	2.001***	0.116*
	(0.019)	(0.014)	(0.016)	(0.010)	(0.428)	(0.066)
3G*Year 2011	0.026	0.031**	-0.055***	0.010	0.716**	0.063**
	(0.017)	(0.012)	(0.014)	(0.011)	(0.334)	(0.026)
3G	-0.023	-0.025*	0.055***	-0.013	-1.341***	-0.108*
	(0.016)	(0.013)	(0.013)	(0.009)	(0.329)	(0.059)
Voting Dist FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Province trends	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Observations	8236	8064	8218	8236	8236	8256
Adj. R-squared	0.819	0.917	0.757	0.621	0.902	0.167

Notes: 3G is the share of area covered by 3G signal. The sample includes only localities that experienced at maximum 3 changes in their demarcations between 2000-2016. These are almost 8% of the initial sample. Standard errors in parentheses clustered at the smallest stable aggregation of voting districts. Controls include: contemporaneous information on (log) luminosity and (log) population density, urbanization rate, years of schooling, ethnicity, age and share of youths, as well as their 2000 values interacted with year dummies, indicators variables for the presence of mine, road and river, slope index, (log) elevation, (log) ruggedness, (log) area, (log) distance to closest city and (log) distance to fiber backbone interacted with year dummies. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

Panel B: Results using aggregations of voting districts; OLS; years 2006 to 2016

	ANC share (1)	DA share (2)	Other parties (3)	Turnout (4)	N. of parties (5)	N. of riots (6)
3G*Year 2016	-0.139***	0.095***	0.039	0.050***	3.883***	0.473**
	(0.036)	(0.030)	(0.034)	(0.016)	(1.431)	(0.229)
3G*Year 2011	-0.055*	0.089***	-0.036	0.036**	2.008	0.236**
	(0.029)	(0.033)	(0.033)	(0.015)	(1.381)	(0.109)
3G	0.048	-0.074**	0.033	-0.043***	-3.226**	-0.359*
	(0.029)	(0.032)	(0.029)	(0.015)	(1.359)	(0.184)
Voting Dist FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Province trends	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Observations	8831	8831	8831	8831	8831	8868
Adj. R-squared	0.837	0.923	0.774	0.720	0.904	0.288

Notes: 3G is the share of area covered by 3G signal. The units of observations are the smallest aggregations of voting districts whose boundaries do not change over time. These units have been used to cluster the standard errors in the baseline specification. Provincial capitals and larger cities are excluded. Standard errors in parentheses clustered at municipality level. Controls include: contemporaneous information on (log) luminosity and (log) population density, urbanization rate, years of schooling, ethnicity, age and share of youths, as well as their 2000 values interacted with year dummies, indicators variables for the presence of mine, road and river, slope index, (log) elevation, (log) ruggedness, (log) area, (log) distance to closest city and (log) distance to fiber backbone interacted with year dummies. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

Table 2.B7: Diff-in-Diff estimates including provincial capitals and larger cities

OLS; years 2006 to 2016; whole sample including provincial capitals and larger cities

	ANC share (1)	DA share (2)	Other parties (3)	Turnout (4)	N. of parties (5)	N. of protests (6)
3G*Year 2016	-0.031**	0.034***	-0.003	0.018*	0.863***	0.057**
	(0.012)	(0.006)	(0.012)	(0.009)	(0.279)	(0.023)
3G*Year 2011	-0.012	0.027***	-0.014*	0.013*	0.627**	0.037***
	(0.009)	(0.005)	(0.008)	(0.007)	(0.263)	(0.012)
3G	0.005	-0.027***	0.024***	-0.018**	-0.559**	-0.051***
	(0.009)	(0.005)	(0.008)	(0.008)	(0.246)	(0.020)
Voting Dist FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Province trends	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Mean in 2006	0.682	0.113	0.191	0.525	7.191	0.002
Observations	127584	122026	126915	127617	127617	128100
Adj. R-squared	0.778	0.870	0.702	0.524	0.889	0.294

Notes: 3G is the share of area covered by 3G signal. Standard errors in parentheses clustered at the smallest stable aggregation of voting districts. Controls include: contemporaneous information on (log) luminosity and (log) population density, urbanization rate, years of schooling, ethnicity, age and share of youths, as well as their 2000 values interacted with year dummies, indicators variables for the presence of mine, road and river, slope index, (log) elevation, (log) ruggedness, (log) area, (log) distance to closest city and (log) distance to fiber backbone interacted with year dummies. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

Table 2.B8: Diff-in-Diff estimates including 2000 as baseline year

OLS; years 2000 to 2016; whole sample

	ANC share (1)	DA share (2)	Other parties (3)	Turnout (4)	N. of parties (5)	N. of protests (6)
3G*Year 2016	-0.044***	0.013***	0.033***	0.008**	0.707***	0.012***
	(0.010)	(0.004)	(0.010)	(0.004)	(0.193)	(0.002)
3G*Year 2011	-0.008	0.009**	0.005	0.025***	0.557***	-0.012***
	(0.007)	(0.004)	(0.007)	(0.005)	(0.187)	(0.004)
3G*Year 2006	-0.002	-0.016***	0.024***	-0.046***	0.366**	-0.027***
	(0.007)	(0.005)	(0.007)	(0.006)	(0.157)	(0.007)
Voting District FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Province time trends	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Mean in 2000	0.604	0.149	0.257	0.514	4.561	0.000
Observations	138269	126609	136187	138335	138335	139056
Adj. R-squared	0.723	0.807	0.724	0.388	0.833	0.093

Notes: 3G is the share of area covered by 3G signal. Standard errors in parentheses clustered at the smallest stable aggregation of voting districts. Controls include: contemporaneous information on (log) luminosity and (log) population density, urbanization rate, years of schooling, ethnicity, age and share of youths, as well as their 2000 values interacted with year dummies, indicators variables for the presence of mine, road and river, slope index, (log) elevation, (log) ruggedness, (log) area and (log) distance to fiber backbone interacted with year dummies. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

Table 2.B9: IV estimates of the impact of number of years with 3G on political outcomes

OLS and 2SLS; years 2000 to 2016; whole sample

	ANC share			DA share			Other parties share		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	2SLS	2SLS	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Years with 3G	-.005***	-.029***	-.046***	.002***	-.005	-.005	.003***	.035***	.053***
	(0.001)	(0.007)	(0.011)	(0.001)	(0.003)	(0.005)	(0.001)	(0.007)	(0.012)
Voting Dist FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls	✓		✓	✓		✓	✓		✓
Mean in 2000	0.60	0.60	0.60	0.15	0.15	0.15	0.26	0.26	0.26
log(ruggedness)*Post		-0.54***	-0.39***		-0.61***	-0.44***		-0.56***	-0.40***
1st stage F-stat		128.6	185.0		129.1	221.6		126.7	176.9
Observations	138269	138269	138269	126609	126609	126609	136187	136187	136187
Adj. R-squared	0.720	0.051	0.005	0.805	0.029	0.044	0.715	-0.012	-0.123

	Turnout			N. of parties			N. of protests		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	2SLS	2SLS	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Years with 3G	.005***	.013***	.011**	.093***	.631***	.638***	.005***	.006***	.006***
	(0.001)	(0.003)	(0.006)	(0.026)	(0.101)	(0.132)	(0.001)	(0.002)	(0.002)
Voting Dist FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls	✓		✓	✓		✓	✓		✓
Mean in 2000	0.51	0.51	0.51	4.55	4.55	4.55	0.00	0.00	0.00
log(ruggedness)*Post		-0.54***	-0.39***		-0.54***	-0.39***		-0.54***	-0.38***
1st stage F-stat		129.0	185.3		129.0	185.3		130.8	186.8
Observations	138335	138335	138335	138335	138335	138335	139056	139056	139056
Adj. R-squared	0.387	0.075	0.133	0.842	0.544	0.567	0.092	0.006	0.007

Notes: Years with 3G is a variable measuring the number of years (since 2005) in which the voting sub-district had more than 50% of its area covered by 3G. Provincial capitals and larger cities are excluded. Standard errors in parentheses clustered at the smallest stable aggregation of voting districts. Controls include: contemporaneous information on (log) luminosity, (log) population density, urbanization rate, years of schooling, age and share of youths, as well as presence of road and river, (log) area, (log) distance to closest city and (log) distance to fiber backbone interacted with year dummies. Coefficient on the instrumental variable from the first stage regression and the respective Kleibergen-Paap rk Wald F-statistic are reported. Post is a dummy which equals 1 if Year>2005. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

Table 2.B10: Diff-in-Diff placebo regressions

OLS; years 2000 to 2006; whole sample

	ANC share (1)	DA share (2)	Other parties (3)	Turnout (4)	N. of parties (5)	N. of protests (6)
Lead 3G*2006	-0.004	0.003	0.005	-0.014	0.302	-0.003
	(0.015)	(0.007)	(0.014)	(0.010)	(0.225)	(0.002)
Lead 3G	-0.019	0.011	0.006	0.003	-0.338*	0.002
	(0.012)	(0.007)	(0.011)	(0.008)	(0.197)	(0.002)
Voting Dist FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Province trends	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Observations	68440	54256	65156	68532	68532	69528
Adj. R-squared	0.793	0.804	0.827	0.378	0.905	0.001

Notes: Provincial capitals and larger cities are excluded. To compute Lead 3G, I assign the value of 3G coverage in 2006 to year 2000, and the value of 3G coverage in 2016 to year 2006. Standard errors in parentheses clustered at the smallest stable aggregation of voting districts. Controls include: contemporaneous information on (log) luminosity, (log) population density, urbanization rate, years of schooling, ethnicity, age, share of youths, and their 2000 values interacted with year dummies, as well as indicators for the presence of mine, road, river, slope index, (log) elevation, (log) ruggedness, (log) area, (log) distance to closest city and (log) distance to fiber backbone interacted with year dummies. They also include a 5^{th}-order polynomial function of (log) ruggedness, (log) area and (log) population density interacted with year dummies. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

Table 2.B11: Diff-in-Diff using the Altonji, Elder and Taber (2005) procedure

OLS; years 2006 to 2016; whole sample

	ANC share (1)	DA share (2)	Other parties (3)	Turnout (4)	N. of parties (5)	N. of protests (6)
Predicted 3G*2016	0.231***	-0.353***	0.094*	-0.573***	-1.294	-0.295***
	(0.046)	(0.029)	(0.049)	(0.031)	(1.189)	(0.048)
Predicted 3G*2011	0.458***	-0.284***	-0.189***	-0.499***	-5.007***	-0.260***
	(0.052)	(0.033)	(0.060)	(0.031)	(1.235)	(0.042)
Predicted 3G	-0.349***	0.268***	0.122**	0.518***	0.466	0.187***
	(0.051)	(0.027)	(0.049)	(0.030)	(1.019)	(0.030)
Voting Dist FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Province trends	✓	✓	✓	✓	✓	✓
Observations	103901	99033	103321	103925	103925	104292
Adj. R-squared	0.734	0.840	0.688	0.468	0.834	0.129

Notes: Provincial capitals and larger cities are excluded. Standard errors in parentheses clustered at the smallest stable aggregation of voting districts. Predicted 3G is the linear prediction of the share of area covered by 3G in each district. The regression output of the linear prediction is reported in Table 2.B1. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

Table 2.B12: IV placebo regressions

2SLS; years 2000 to 2006; only localities with no 3G coverage in 2006

	ANC share (1)	DA share (2)	Other parties (3)	Turnout (4)	N. of parties (5)	N. of protests (6)
log(Ruggedness)*2006	-0.003	-0.002	0.005	-0.002	-0.043	-0.000
	(0.003)	(0.002)	(0.004)	(0.002)	(0.033)	(0.000)
Voting Dist FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Observations	59354	45356	56150	59442	59442	60428
Adj. R-squared	0.783	0.722	0.830	0.345	0.859	0.001

Notes: Provincial capitals and larger cities are excluded. Standard errors in parentheses clustered at the smallest stable aggregation of voting districts. Localities covered by 3G in 2006 are excluded from the analysis. Controls include: contemporaneous information on (log) luminosity and (log) population density, urbanization rate, years of schooling, age and share of youths, as well as presence of road and river, (log) area, (log) distance to closest city and (log) distance to fiber backbone interacted with year dummies. They also include a 5^{th}-order polynomial function of (log) area, (log) population density and (log) luminosity interacted with year dummies. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

Table 2.B13: The information channel: heterogeneity by traditional media penetration

Diff-in-Diff-in-Diff, years 2011 to 2016; whole sample

	Low radio/TV penetration				High radio/TV penetration			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ANC share	Other parties	Turnout	N. of parties	ANC share	Other parties	Turnout	N. of parties
Irregular Exp*3G*2016	-0.015***	0.017***	0.002	0.116	-0.004	0.004	0.000	0.102
	(0.005)	(0.005)	(0.003)	(0.164)	(0.004)	(0.005)	(0.002)	(0.146)
Total Exp*3G*2016	0.015	-0.013	-0.008	-0.050	0.031***	-0.035***	-0.009*	-0.134
	(0.015)	(0.014)	(0.006)	(0.441)	(0.010)	(0.012)	(0.005)	(0.379)
Voting Dist FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Province trends	✓	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Observations	26930	26878	26956	26956	26606	26030	26614	26614
Adj. R-squared	0.685	0.667	0.543	0.786	0.853	0.720	0.649	0.900

Notes: 3G is the share of area covered by 3G signal. Provincial capitals and larger cities are excluded. Standard errors in parentheses clustered at municipality level. Localities with low (high) traditional media adoption are places where both radio and TV penetration in 2011 was below (above) the respective median value. Irregular Exp is the log of total irregular and unauthorized municipal expenditure per capita from 2012 to 2015. Total Exp is the log of income and capital municipal expenditure per capita from 2012 to 2015. All other interactions are included but not reported for conciseness. Controls of column (4) of Table 2.3 are included here. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

Table 2.B14: internet use and information acquisition (Afrobarometer, 2015)

Panel A: Extensive margin - whole sample

	Y=1 if respondent regularly accesses to information through:				
	Internet	Social media	Newspapers	Television	Radio
Internet user	0.605***	0.548***	0.212***	0.056***	0.039*
	(0.022)	(0.029)	(0.030)	(0.017)	(0.024)
Observations	2357	2358	2362	2364	2364

Notes: Standard errors in parentheses clustered at the village level. Province fixed effects included in all regressions. Controls include: occupational status, age, education, religion, distance from closest provincial capital/major city. The dependent variables take value 1 if the respondent reported to use the respective media to access information at least a few times a month. Internet user is a dummy which takes value 1 if the respondent reported to use the internet for any purpose at least a few times a month.* Significant at 10%, ** significant at 5%, and *** significant at 1%.

Panel B: Intensive margin - only respondents who use the internet at least a few times a week

	Y = N. of days per month respondent accesses to information through:				
	Internet	Social media	Newspapers	Television	Radio
Use Internet everyday = 1 (vs. a few times a week = 0)	7.879***	7.061***	1.025	-0.777*	-0.490
	(0.681)	(0.844)	(0.937)	(0.462)	(0.758)
Observations	1108	1109	1109	1109	1109

Notes: Standard errors in parentheses clustered at the village level. Province fixed effects included in all regressions. Controls include: occupational status, age, education, religion, distance from closest provincial capital/major city. The dependent variables count the number of days in a month the respondent reported to use the respective media to access information. These are computed from the categorical variables as follows: Everyday = 28 days, A few times a week = 12 days, A few times a month = 4 days, Less than once a month = 1 day, Never = 0 days. The main independent variable is a dummy which takes value 1 if the respondent reported to use the internet for any purpose everyday, and 0 if she uses the internet only a few times a week.* Significant at 10%, ** significant at 5%, and *** significant at 1%.

Table 2.B15: Internet coverage and population density

Dependent variable is: log(population density)			
	Diff-in-Diff: 2006-2016 (1)	OLS: 2000-2016 (2)	2SLS: 2000-2016 (3)
3G*Year 2016	0.032		
	(0.030)		
3G*Year 2011	-0.000		
	(0.028)		
3G	-0.023	0.012	0.086
	(0.027)	(0.012)	(0.102)
Voting Dist FE	✓	✓	✓
Year FE	✓	✓	✓
Province trends	✓		
Controls	✓	✓	✓
1st stage Wald F-stat			180.7
Observations	104292	139952	139952
Adj. R-squared	0.951	0.930	0.018

Notes: 3G is the share of area covered by 3G signal. Provincial capitals and larger cities are excluded. Standard errors in parentheses clustered at the smallest stable aggregation of voting districts. Controls of Table 2.B5 are included in column (1), while controls of Table 2.B3 are included in columns (2-3), with the exception of population density, which is the dependent variable. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

2.C Figures

Figure 2.C1: Mobile internet coverage 2007-2015 (Raw data from Collins Bartholomew)

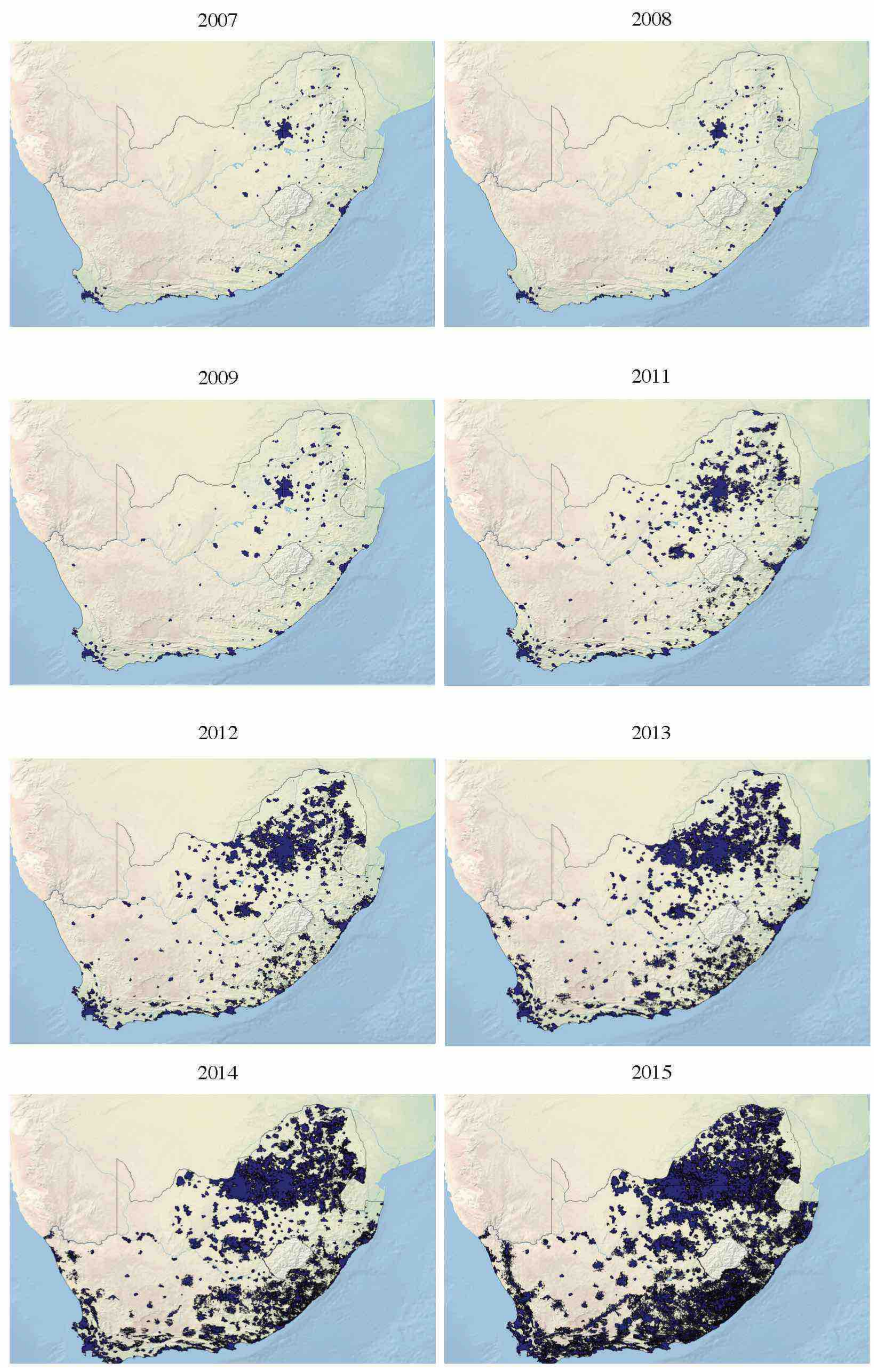


Figure 2.C2: Access to information by type of media

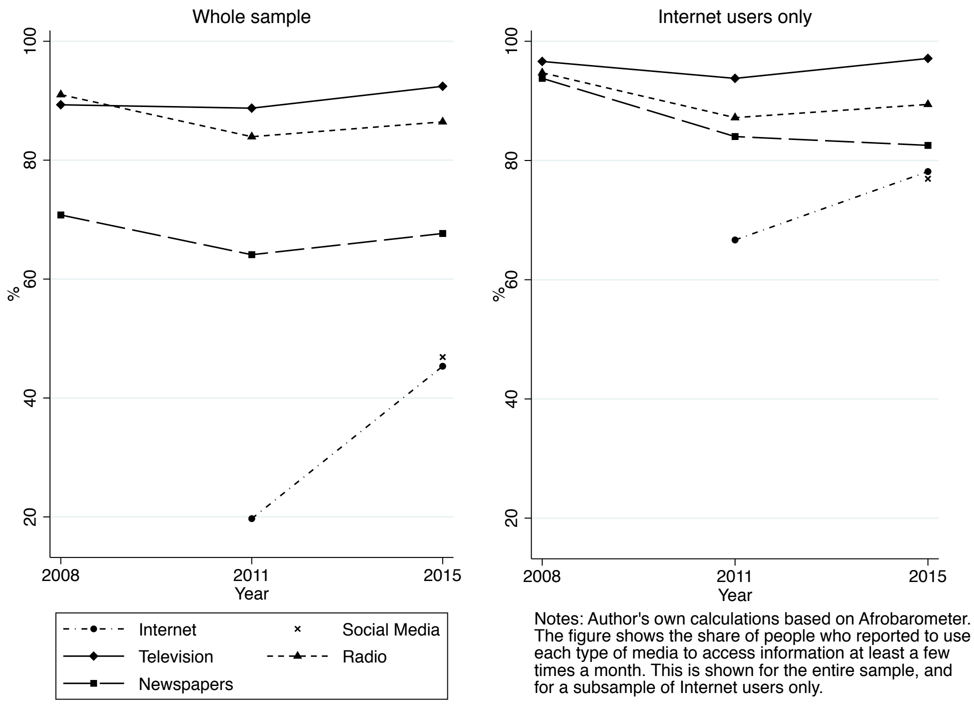
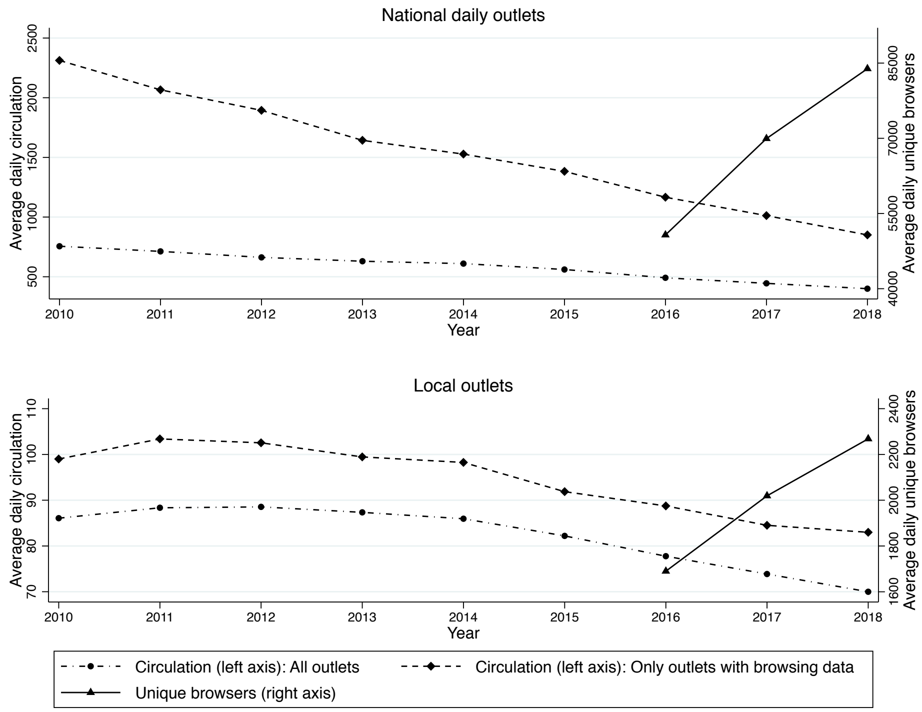


Figure 2.C3: National and local newspapers circulation



Notes: Author's own calculations based on data from the Audit Bureau of Circulations of South Africa (http://abc.org.za). Average daily circulation in a year is computed from quarterly information as follows. For each quarter, I divide the total number of printed copies sold by the total number of days in the quarter, irrespectively of the frequency of publication. I then take the average of the resulting number over the year.

Figure 2.C4: Mobile phone technology

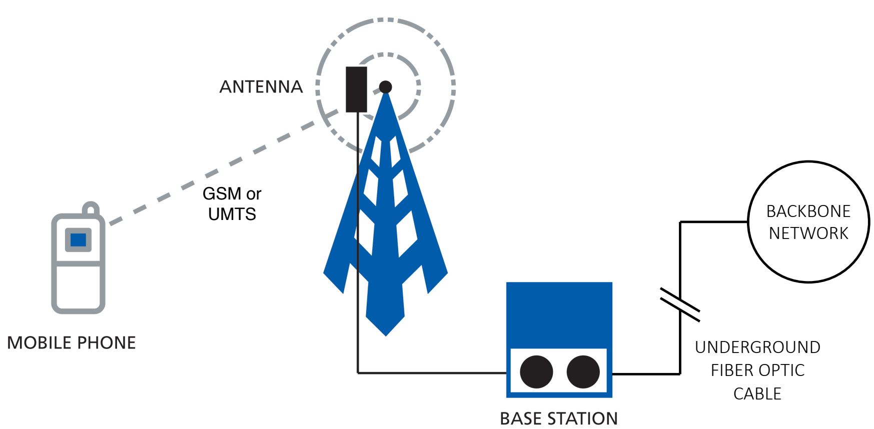


Figure 2.C5: Share of area covered by 3G

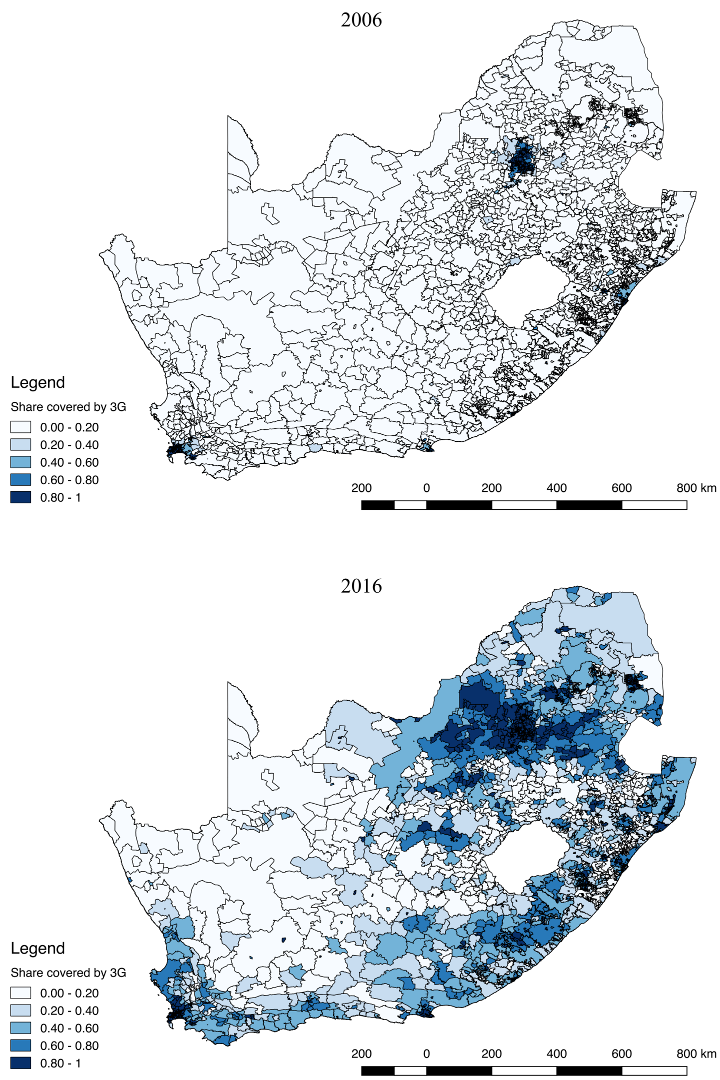
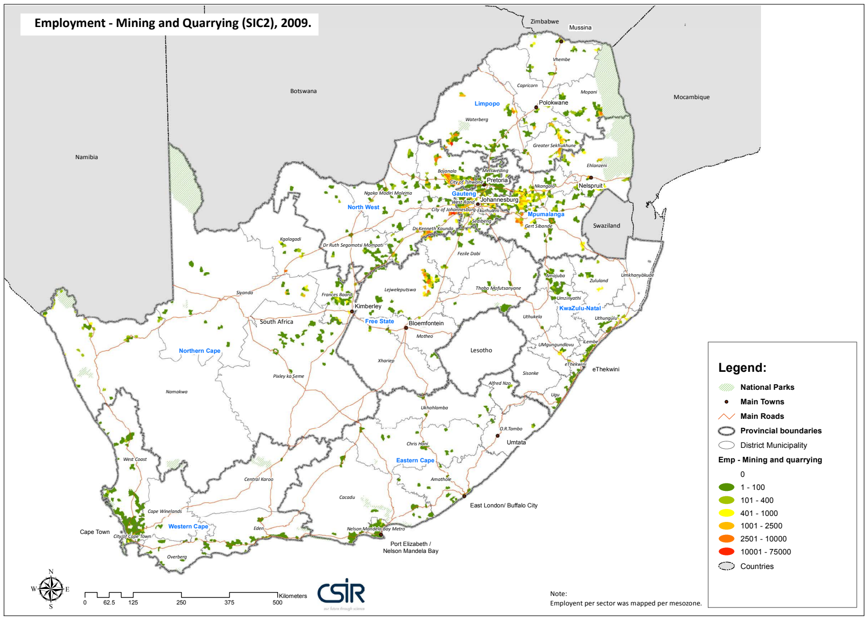


Figure 2.C6: Economic relevance of the mining industry (2009)



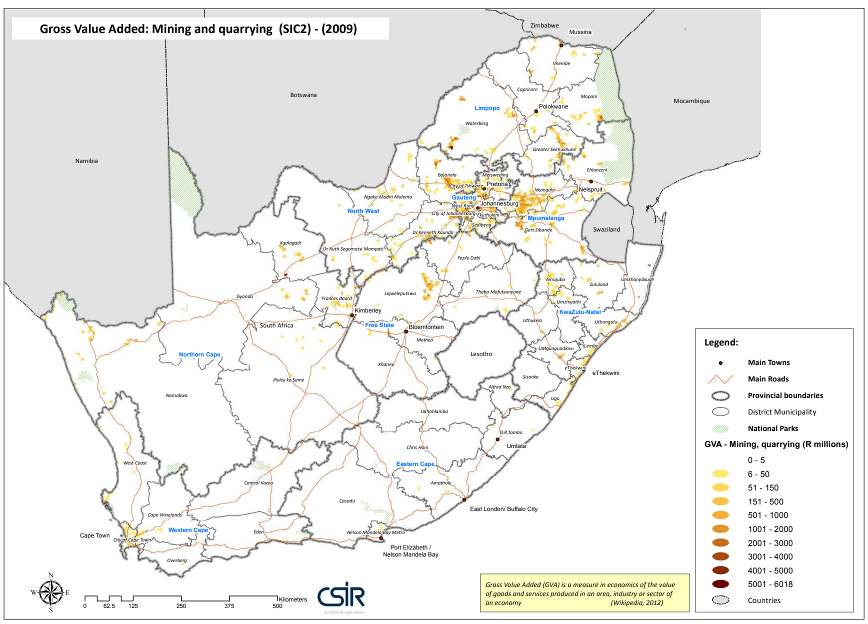


Figure 2.C7: 3G coverage and political opinions over time (Afrobarometer)

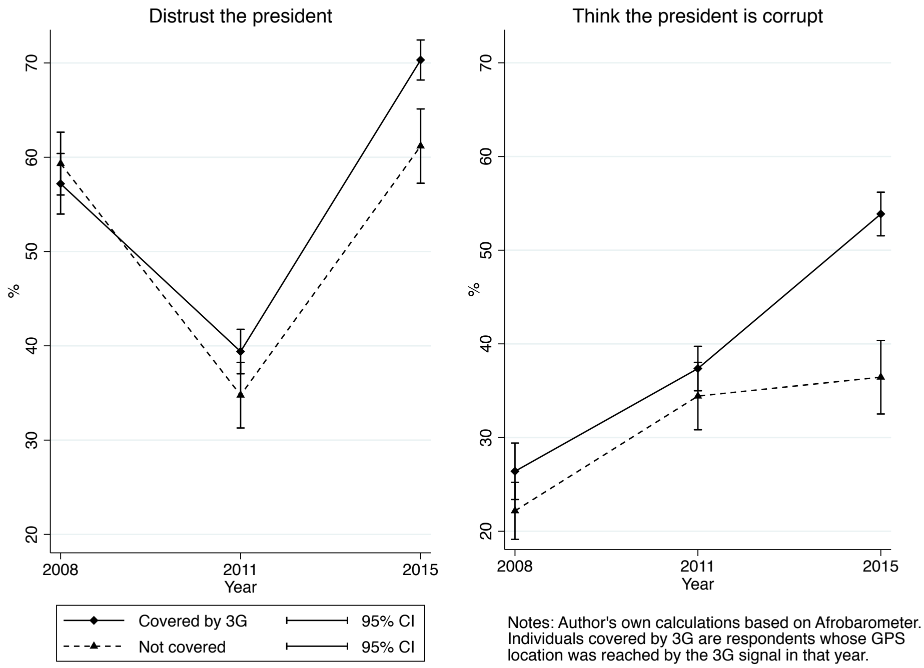
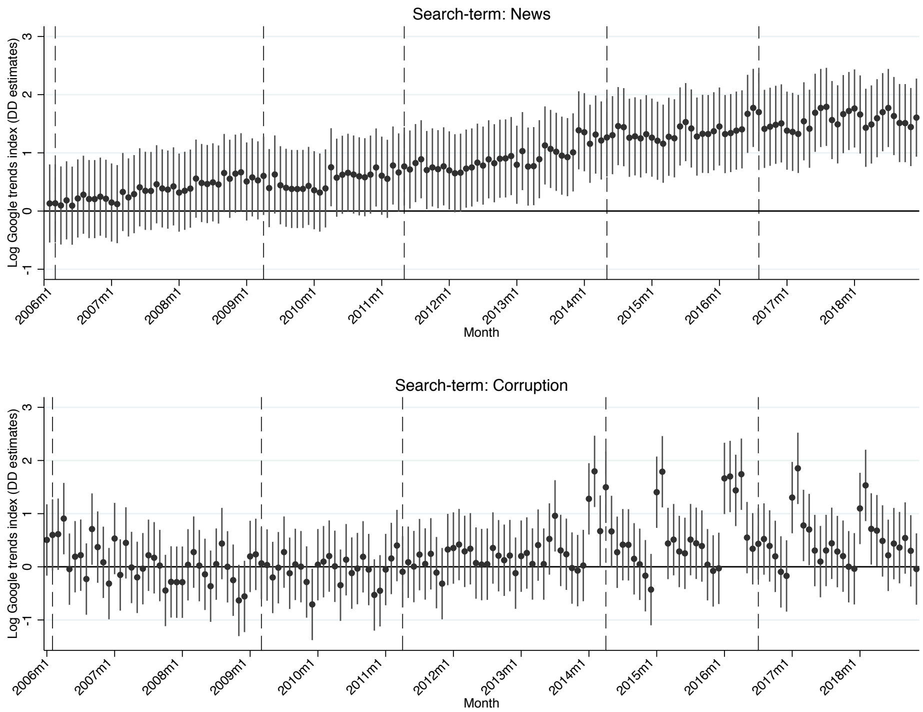
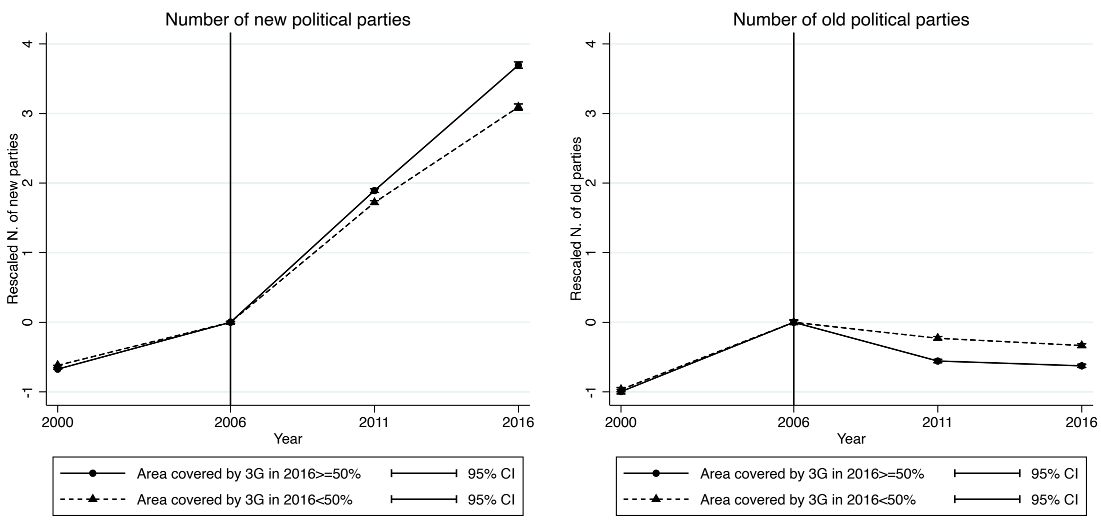


Figure 2.C8: Online information acquisition (monthly level)



Notes: The graph shows point estimates and 95% confidence intervals of a regression of the natural log of a search-term popularity index from Google trends on a full set of month dummies interacted with an indicator for search terms "news" (top panel) or "corruption" (bottom panel). Regression includes search-term and month fixed effects. The omitted month is January 2006. Regressions includes 3,432 observations from 22 search terms. The most conservative standard errors are used to construct the confidence intervals. The vertical dashed lines indicate the months in which national and municipal elections took place. The list of search terms included in the reference group reflects the general categories identified by Google. These are: Arts, Beauty, Books, Business, Cars, Computers, Finance, Fitness, Food, Games, Health, Hobbies, Home, Internet, Jobs, Law, Pets, Science, Shopping, Sports and Travel.

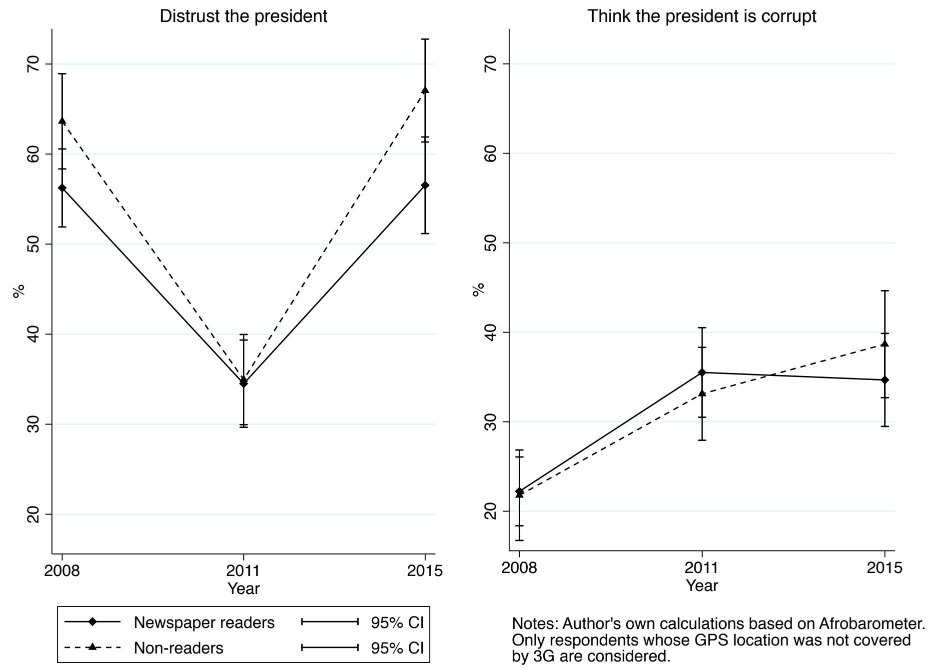
Figure 2.C9: Trends in number of new and old parties by 3G coverage



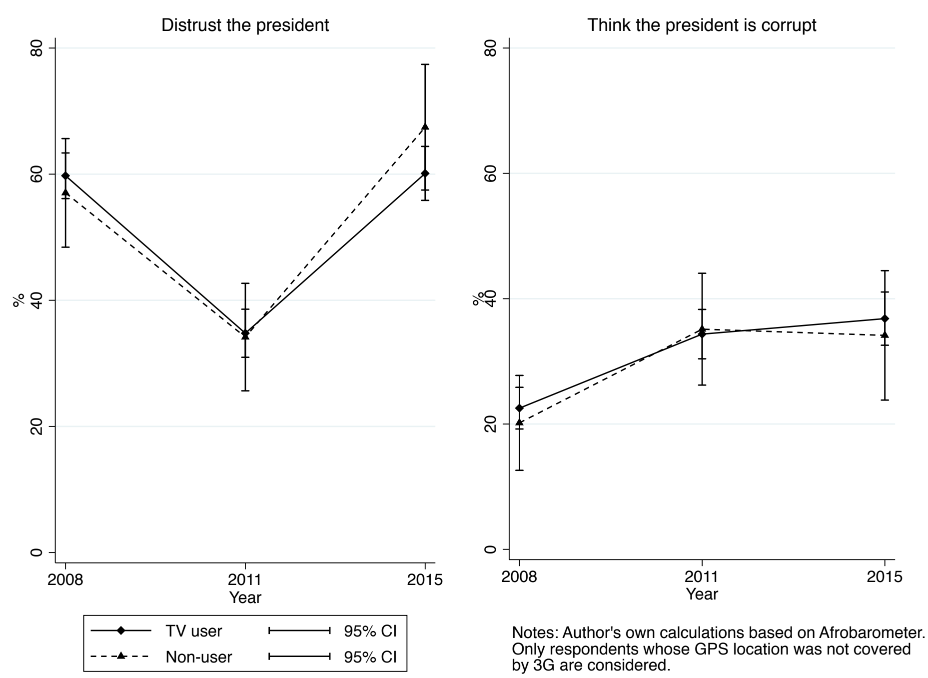
Notes: Provincial capitals and larger cities are excluded, as well as localities that were covered by 3G in 2006. Lines are re-scaled so that in 2006 the averages for the two subgroups exactly coincide and are equal to 0. A party is considered to be new if it did not run in any districts during the 2000 municipal elections.

Figure 2.C10: Traditional media use and political opinions over time (Afrobarometer)

(a) Opinions by newspaper readership to access information



(b) Opinions by TV use to access information



Chapter 3

Using social media to change gender norms: an experimental evaluation within Facebook Messenger in India

Joint with Victor Orozco-Olvera and Nandan Rao

3.1 Introduction

Violence against women (VAW) is a global epidemic, with thirty-five percent of women worldwide having experienced physical or sexual violence in their lives (WHO, 2013). Its adverse effects range from physical and mental health issues for women and their children to broader social and economic losses (Raghavendra et al., 2019). For instance, Duvvury et al. (2013) find that VAW's annual economic costs can reach up to 1-2 percent of GDP due to healthcare costs, productivity loss, and losses in future human capital formation. Permissive attitudes towards VAW are widely accepted, especially in developing countries. Sardinha and Catalán (2018) estimate that four in ten women (and three in ten men) justify VAW in about 50 low and middle-income countries. Such a widespread acceptance of domestic violence is a risk factor for its incidence (Abramsky et al. 2011; Flood and Pease 2009).

Most VAW prevention programs are delivered through grassroots mobilization campaigns and tend to be resource-intensive (Green et al., 2020). Evidence on these interventions is mixed (Abramsky et al. 2014, 2016; Bourey et al. 2015; Dhar et al. 2022; Jewkes et al. 2020; Kerr-Wilson et al. 2020; Wagman et al. 2015) and their scaled implementation may be prohibitively costly in low-resource settings. Social and behavior change communication campaigns (SBCC) can provide a cheaper alternative. Their messaging can directly reshape individuals' attitudes and behaviors (individual channel), and by reaching many community members, mass media can also potentially update people's perceptions of prevalent social norms (Akerlof and Kranton 2000; Mackie 1996). This in turn can reshape individual beliefs and attitudes (social channel).

Entertainment-education – also known as edutainment – is the use of entertainment media to increase audience members' knowledge about an educational issue, create favorable attitudes, shift social norms, and change overt behavior (Brown and Singhal 1999; Singhal and Rogers 2012). Recent field experiments demonstrate that even low doses of edutainment programming (under three hours) can effectively reshape gender norms and reduce the social acceptability and incidence of VAW (Arias 2019; Banerjee et al. 2019*b*; Green et al. 2020). This research, which tested radio, television and film dramas in open community settings (e.g., community screenings), experimentally shows that the social channel greatly drove these impacts, and that private viewings of edutainment media had limited effects on VAW outcomes (Arias 2019; Green et al. 2020). A natural question, then, is whether these findings can be generalized to social media campaigns, where the influence of individual and social channels are less clear.

In this paper, we experimentally test the effectiveness of edutainment at reshaping gender norms and reducing social acceptability of VAW when delivered individually through Facebook Messenger. The tested campaigns consisted of short video clips (that in total amounted to 25 minutes), commonly used in social media distribution. We recruited 18-to-24-years-old individuals living in New Delhi and 6 other northern Indian cities using a Facebook ad campaign. We collected self-reported and objective online outcomes using a newly-developed chatbot and measured impacts one week and four months after program exposure. To understand whether implicit messaging was more effective than explicit messaging, the treatment group was randomly exposed to either a humorous fake reality TV drama (implicit) or a docu-series about VAW with clear calls to action (explicit).

Our findings show that the content of behavior change campaigns is very important, with edutainment formats helping increase both take up rates and program effectiveness. Our objective viewership data shows that take-up was twice as high for the humorist drama, *Sex Ki Adalat*,[1] compared to the more information-focused docu-series. On effectiveness, however, neither format dominated, yet different outcomes were affected consistently with program content. While the drama was more effective at raising knowledge and reshaping attitudes related to gender norms and VAW, the docu-series was more impactful in increasing willingness to share video-clips with friends, promoting online information-seeking behaviors, and making social media users more likely to add the frame “End Violence Against Women” to their Facebook picture profile. The latter effects are important outcomes for edutainment campaigns that seek to influence social norms and perceptions of baseline beliefs within communities.

In particular, our study shows that edutainment had economically and statistically significant effects in the short-term (one week after program exposure) on awareness and knowledge, attitudes related to gender norms and roles, attitudes condoning VAW, and information-seeking behaviors. Campaign impacts on indexes oscillated between 0.16 and 0.21 standard deviations, which are substantial considering the low dosage of the intervention (25 minutes or less). Several of these effects diminished or fully dissipated in the medium term (four months after program exposure), suggesting a time-decay in effects. Only the gender norm/role attitudes index of the entertainment drama was statistically significant in the medium term, with its magnitude decreasing by a fourth compared with its short-term effect (0.15 vs 0.20 SDs). These findings are strengthened by evidence of baseline balance, no evidence of attrition bias or placebo effects, and are robust to the addition of controls across practically all outcomes of interest.

In our post-hoc heterogenous analysis,[2] consistent with theories suggesting that people’s beliefs and behaviors are influenced by their perception of prevalent social norms (e.g., Bicchieri 2005, 2016; Miller and McFarland 1987), we observe short-term smaller effects for individuals that perceived their Facebook friends to have more conservative gender views at baseline, yet these effects vanished in the medium term. While short-term impacts could potentially be explained by the docu-series message of how certain practices are widespread in Indian society (and by consequence, in users’ social circles), as time passed and respondents re-engaged with their social circles, pre-treatment beliefs may have re-emerged.

[1] Hindi name for "The Court of Sex".

[2] The study analysis was not pre-registered, including the variables used in our heterogenous analysis. Thus, we refer to the latter as post-hoc.

With the exception of females, who generally experienced smaller effects than males,[3] we find no evidence that the intervention had differential effects for a series of social and demographic indicators, including age, caste, membership to a social organization, and educational achievement of parents and of individuals. The general lack of heterogeneous effects for different social status indicators suggests that social media may be an effective medium for delivering social norms campaigns to vulnerable populations, who may find it harder to participate and have a voice in community events.

Most of the short-term effects disappeared in the medium term. However, the medium-term analysis revealed that the docu-series was successful in encouraging "public champions", by making individuals 91% (7.5 p.p.) more willing to take a public stance against VAW by adding a frame to their Facebook profile picture. For this objective measure, we did not find significant effects for the implicit format, a result that suggests the explicit format was more effective in promoting individuals' public commitment to social action. Moreover, we documented persisting effects in the long-term. We found that after about 1 year and a half from the moment the frame was presented to the respondents, it was used by more than 34 thousand people all over the world, an amplification of about 55 times the initial audience size.

Our findings provide new evidence that social norms marketing campaigns that use edutainment formats can trigger immediate shifts in individuals' core values related to gender norms, roles and VAW attitudes, even when delivered individually through Facebook Messenger. Impacts on this individual channel are also observed by the study of the television drama MTV Shuga in Nigeria (Banerjee et al., 2019*a*), but not in the studies of a radio program in Mexico and movie ads in Uganda (Arias, 2019; Green et al., 2020). Our study design cannot precisely disentangle the role that the social channel played in mediating program impacts. However, because study participants did not know if their friends were part of the study or not, the data suggests the social channel is not a necessary condition for changing individuals' attitudes and behaviors. Moreover, considering the power of social media in amplifying content and opinions through online sharing, our results lend support to the possibility of achieving a self-sustained virtuous circle of social change.

Our study contributes to various streams of economic literature. The importance of

[3]Attitudinal impacts across genders are mixed in previous VAW edutainment studies, with the Nigeria (Banerjee et al., 2019*a*) and Uganda trials (Green et al., 2020) generally finding stronger impacts for men and women, respectively.

social norms in perpetuating gender gaps has received recent attention by economists (Alesina et al. 2013; Bertrand 2020; Bertrand et al. 2015. Some studies emphasize the impact of gender stereotypes on educational outcomes (Alan et al. 2018; Carlana 2019; Lavy and Sand 2015; Terrier 2015), on belief distortions (Bordalo et al., 2019), and the importance of correcting gender misperceptions for female labor participation (Bursztyn et al., 2020). We add to this literature by showing that edutainment interventions which correct perceptions of social norms can be effective at changing existing permissive attitudes towards VAW.

This study also informs the literature on VAW prevention programs in low-income and middle-income countries by providing further evidence that social media-delivered edutainment can also be effective in combating this global epidemic (WHO, 2013). Empirical work within the social media literature have thus far focused on health, crime, political and wellbeing outcomes (Alatas et al. 2019; Allcott et al. 2020; Allcott and Gentzkow 2017; Banerjee et al. 2020; Bond et al. 2012; Donati 2019; Enikolopov et al. 2020; Petrova et al. 2020). To the best of our knowledge, this is the first study that experimentally tests an edutainment campaign aimed at reshaping gender norms and VAW attitudes via a social media platform. Finally, we demonstrate the scale potential of online research to study the effects of social media campaigns beyond short-term impacts (e.g., clicks and online posting behavior). In particular, we use a newly-developed survey chatbot called Virtual Lab (Rao et al., 2020) that can be integrated into social media platforms to deliver and measure the impact of online campaigns.

The remainder of this paper is structured as follows: Section 3.2 discusses the evidence base of edutainment and our study's research questions, Section 3.3 describes the study design and intervention, Section 3.4 shows the results, and Section 3.5 draws policy conclusions.

3.2 Theoretical framework

3.2.1 Edutainment and Social Norms

Systematic reviews of information-only campaigns tend to show limited effectiveness on behavior change (e.g., Ferri et al. 2013; McKenzie-Mohr 2000). Communication researchers argue that information-only campaigns often fail because their explicit messaging may trigger counter-arguing (Nyhan et al., 2014) especially for sensitive issues that require individuals to revisit their core values. Their lack of engaging narratives and identifiable role models may also prevent individuals from enhancing their self-efficacy beliefs (Singhal et al., 2003).

Retrograde gender norms perpetuate VAW and gender discrimination, since the acts are justified by established beliefs and attitudes (Abramsky et al. 2011; Flood and Pease 2009. Such effects may be particularly strong for subgroups with lower social status (Goffman 1963; Hoff and Stiglitz 2010; Hoff and Walsh 2018; World-Bank 2014). Gender norms, driven by men and women's motivation to adjust their self-view to what seems socially appropriate (Akerlof and Kranton, 2000), can be important barriers in reducing gender gaps by becoming internalized into individual preferences. Social norms marketing campaigns[4] to improve women's economic, political and social status are increasingly using edutainment. Through vicarious learning, people may acquire new information about social norms as well as ways of responding to social situations based on behaviors modeled by program characters (Bandura, 2004).

Field experiments of edutainment demonstrate that dramatized narratives are effective in promoting attitudinal and behavioral change across development sectors, including improving financial decision-making (Berg and Zia, 2017) increasing willingness to report corruption (Blair et al., 2019), reducing deference to authority (Paluck and Green, 2009), improving educational outcomes (Kearney and Levine, 2019), and promoting safer sexual behaviors (Banerjee et al. 2019*b*; Orozco-Olvera et al. 2019; W. Vaughan 2000; Wang and Singhal 2016).

The evidence base of edutainment in development economics has recently expanded to gender norms and VAW. In Nigeria, Banerjee et al. (2019*a*) showed that a short storyline on domestic violence embedded in the television drama MTV Shuga was effective in changing attitudes and behaviors related to domestic violence. In Uganda, Green et al. (2020) found that short advertisement clips during a film festival were effective in increasing audiences' willingness to report violence to police or community leaders and in decreasing reported incidence. In Mexico, Arias (2019) finds that a radio drama increased rejection of VAW and increased support for gender equality.[5] The

[4] Social norms marketing refers to traditional marketing techniques, including mass media and face to face campaigns, that are designed to alter individuals' perceptions about which attitudes and behaviors are typical or desirable in their community (Burchell et al., 2013; McKenzie-Mohr, 2000).

[5] These findings confirm quasi-experimental evidence that suggested that by exposing viewers to outside views and lifestyles in a dramatized format, communities' access to cable or to the "soap opera"

Uganda and Mexico studies provide experimental evidence that effects were driven by the social channel, facilitated by the communal delivery of public information. This is in line with Bursztyn et al. (2020), who show that pluralistic ignorance – i.e., when most group members privately reject a group norm but publicly follow it as they believe that most members accept it (Miller and McFarland, 1987) – might be affecting men's willingness to allow women's participation in the labor force in Saudi Arabia.

The potential of social media platforms in delivering SBCC is greatly untapped in development. For instance, in India, where our study takes place, over 70 percent of individuals between 18 and 34 years used Facebook in 2018. Most of them spent between two and four hours on social media every day (Statista 2020). Despite their potential to reach many at low costs, social media are still overlooked by development programs aimed at reshaping attitudes towards gender norms and combating VAW. A potential reason is the general lack of empirical evidence on their effectiveness at achieving such development goals.

3.2.2 Research questions

Our study aims to address two questions. First, can the above findings from edutainment be generalized to social media?[6] Theoretical arguments can be made for social media to be less or more effective than television, radio or film. On the one hand, the need for shorter clips for social media consumption may prevent users from effectively immersing in a program and identifying with characters.[7] The lack of a shared community viewing experience may prevent activating the social channel. Moreover, social media friends may not be a relevant group if users perceive them as too detached and unlikely to take any credible actions against VAW in their communities.

On the other hand, social media campaigns may theoretically be more effective than "offline" media. Online campaigns can be more effective in encouraging people to seek further information or take action (e.g., visit a website or donate to a social cause) as internet sites are a few clicks away. Social media usually exposes users to a larger

channel were effective in improving gender outcomes in Brazil (La Ferrara et al., 2012) and decreasing acceptability of VAW in India (Jensen and Oster, 2009).

[6]With social media platforms engaging one in two people worldwide (Digital 2020), delivering edutainment campaigns through this medium could potentially be a scalable and cost-effective approach for reshaping gender-equality and VAW outcomes.

[7]Berg and Zia (2017) and Banerjee et al. (2019*a*,*b*) provide suggestive evidence that program impacts were mediated by program immersion and identification with characters.

number of friends and acquaintances, which can facilitate the spread of gender-equality norms, especially among youth. The ease with which users can publicly display their views (e.g., by posting on their Facebook walls) and share information online can help "normalize" new views within online communities. Online campaigns that are effective in encouraging users to publicly show their support for a social cause can potentially trigger a cascade of broader social support, as shown by online movements such as #MeToo (Levy and Mattsson, 2021).

The second question our study addresses is whether formats that deliver messaging in more implicit ways are more effective in influencing gender norms and VAW attitudes. Implicit formats such as fictional and humorous narratives could potentially reduce counterarguing, the thoughts that may dispute persuasive arguments (Benoit, 1987), and may create a safer space for audiences to consider new views.[8] Lab studies, mostly conducted in US colleges, show that people are more likely to remember and internalize messages when presented in a narrative format (Frank et al. 2015; Ochoa et al. 2020; Oliver et al. 2012). On the other hand, dramas could also trivialize social issues (Moyer-Gusé, 2008). Because documentaries are usually based on real-world people and situations, their "call for action" messaging could theoretically be more effective in influencing the social channel.

3.3 This study

3.3.1 Study Design

The study is a randomized control trial of short clips of edutainment campaigns designed to reshape gender norms, roles and VAW attitudes. The study, a partnership with the Population Foundation of India and the World Bank, sought consent from all study participants and received ethical clearance from Solutions IRB (IORG0007116).

We recruited 18-24 years olds residing in New Delhi and six other large cities in northern India. Study participants were recruited on Facebook and Instagram through a 1-week geo-targeted advertising campaign.[9] As participation incentives, individuals that

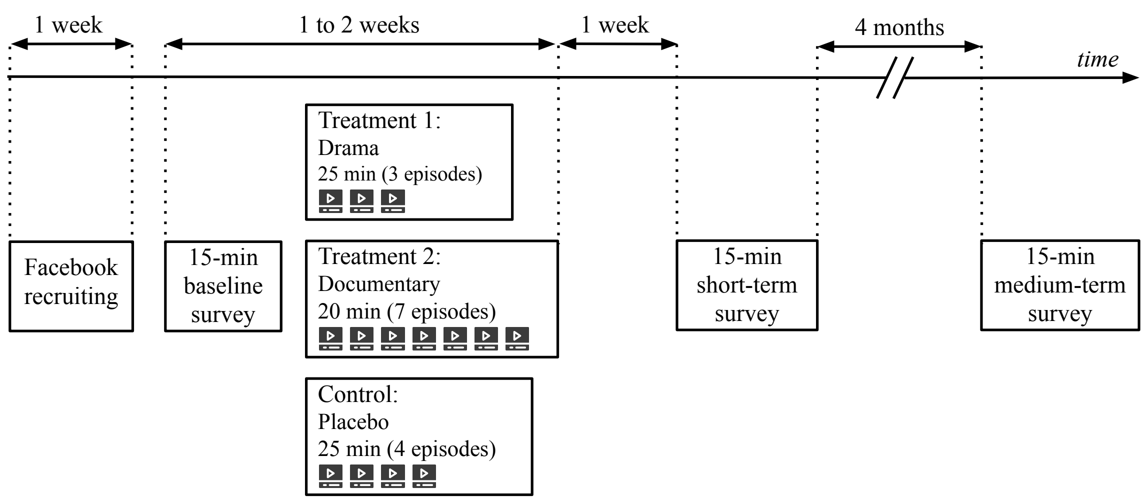
[8] Lab studies indicate that audiences enjoying an entertainment program are less likely to question and rebut program messages through increased program transportation or immersion (Hall and Bracken 2011; Moyer-Gusé 2008).

[9] Figure 3.B1 in the Appendix shows the geographic targeting and the ad content used in the recruiting campaign.

completed the baseline and at least one follow up survey were eligible for a lottery to win Samsung Galaxy smartphones or a "selfie" picture with a Bollywood celebrity.[10] Individuals who clicked on the ad banner were redirected to Facebook Messenger, where both the intervention and data collection surveys were delivered through Virtual Lab, a newly-developed open-source automated chatbot described by Rao et al. (2020).[11] This platform also allowed the research team to directly measure treatment adherence (i.e., viewership rates, an interesting outcome for edutainment campaigns) and therefore allowed for more reliable estimation of the treatment effect on the treated.

Baseline respondents (n=5,229) were randomized into treatment and control conditions. In both conditions, participants were shown a series of short edutainment video clips totalling less than 25 minutes. We measured program impacts one-week and four-months after the end of the intervention, what we refer to in the rest of the paper as short-term (n=606) and medium-term impacts (n=619), respectively. The timeline and structure of the study are described in Figure 3.1.

Figure 3.1: Timeline of the study



3.3.2 Interventions

The studied edutainment content was produced by WEvolve, a multi-donor initiative supporting innovative campaigns against VAW, and Population Foundation of India,

[10]The celebrity was Farhan Akhtar, a popular Indian actor, director, screenwriter and producer.
[11]Here is the website of the platform, while here is the GitHub repository of its code. Moreover, Appendix Figure 3.B2 shows an example of its functioning.

an NGO that advocates for the formulation and implementation of gender-sensitive development policies and programs.[12] Focus group research informed campaign content and celebrities, including actor and screenwriter Farhan Akhtar, make appearances. To understand if explicit or implicit formats were more effective in influencing gender norms and VAW attitudes, we selected the following edutainment programs.

Treatment 1 (implicit) was a fake reality TV web series called Sex Ki Adalat (Court of Sex). The program takes place in a fictitious court where myths and misconceptions around gender norms are discussed often in a humorous way. We showed 3 episodes of this series, for a total length of 25 minutes. These episodes focused on the determinants of the child's gender (sex selection at birht), female and male virginity at marriage, and the "menstruation ritual", which bans women from entering the kitchen or household shrine during their period.

Treatment 2 (explicit) was a series of WEvolve clips that aimed to raise awareness on VAW prevalence in India, including real-life stories and experiences from people similar to our target population (i.e., young, middle-class, city-dwelling Indians), with clear calls to action. The clips heavily used music and editing to be both contemporary and emotionally impactful and were often character-driven. We showed 7 episodes for a total length of 20 minutes. The control group was exposed to a "placebo". In particular, respondents in this condition were invited to watch Carbon, a short and engaging edutainment movie on climate change. The movie was delivered in 4 episodes, for a total length of 25 minutes.

Treatment 1 and control video-clips were in Hindi with English subtitles, while Treatment 2 was mainly in English.[13] For each arm, episodes were released in a staggered way. In particular, each new episode was delivered two hours after the individual self-confirmed to having watched the previous episode. Participants received reminders to encourage viewership and were free to choose when to watch the episodes. While this approach increased the duration of the total intervention – which took an average of 7-to-10 days for completion, depending on the arm – , it mimicked the way competing content is commonly consumed and shared in social media platforms (e.g., Instagram reels),

[12] Specifically, the material is the following: Treatment 1: Sex Ki Adalat (E1) Male child, (E2) Virginity, (E3) Menstruation; Treatment 2: WEvolve clips; Control: Carbon.

[13] We discard the possibility that content comprehension drives the results, since our targeted population (i.e., youths living in cities) can speak and understand both Hindi and English. Our respondents' self-evaluation of their English comprehension was 5.5 on a 1-10 scale. We also control for this covariate in all the regressions.

making the study more generalizable.

3.3.3 Outcomes and measures

We measure program impacts on three categories of outcomes: (1) creating awareness on gender norms and VAW in India, (2) changing attitudes regarding gender norms and condoning VAW, and (3) online information-seeking and posting behaviors related to gender issues. The first two sets of outcomes are self-reported and measured through the survey instrument. Specifically, awareness and knowledge questions cover issues explicitly discussed by the series. Attitudinal items aimed to measure core values were derived from the India's National Family Health Survey.

In addition to self-reported data, we independently measured two online outcomes. First, we measured clicks and visit durations to gender- and environment-related website links provided in both short- and medium-term follow-up surveys. We hypothesized that the treatment group would be more likely to click and spend a longer time browsing gender websites as opposed to other topics.[14] Second, study participants were provided with the option to publicly display their disapproval of VAW through their Facebook profiles. This measure was operationalized by giving participants in the medium-term survey the opportunity to add a frame against VAW in their Facebook picture profile (frame in Appendix Figure 3.B3).

Practically most attitudinal and behavioral measures were originally coded using a 5-point agreement scale. To facilitate the interpretation of the results, we transformed them into binary indicators, thus program impacts are reported as percentage point changes. To address the issue of multiple hypothesis testing, we group individual level outcomes into four topic indexes: (i) knowledge and awareness of existing gender norms

[14] Time spent visiting a website is extrapolated from data on link clicking collected by the Virtual Lab platform. Appendix Figure 3.B2 shows an example of the provided buttons. Visit duration is calculated from the moment the respondents click on one link until the moment they click on the subsequent link. For this reason, we could not measure visit duration for the last link the respondent clicked. Since missing duration tends to be more frequent on the last websites, we focused the attention on those links which were provided at the beginning, namely the PFI and Delhi Green websites, and exclude from the analysis the UN Women India and the UN Environment Program India websites. An additional problem with measuring visit duration arises when the respondents click on one link, then close the browser (or phone) without any further action, and then return to visit another link at a later moment in time. In this case, our measure of visit duration would count the entire difference between the two visits as a visit to the first website. To correct for the related measurement error we assumed that all visits that lasted more than 2 hours were capturing time spent away from our websites. Hence, we discarded them by replacing the values with missing data.

and VAW, (ii) attitudes towards gender norms and roles, (iii) attitudes condoning VAW, and (iv) beliefs on others. We additionally constructed a general index that aggregates all these topic indexes, except for the last one. In fact, beliefs on others capture users' perception of the attitudes of their closest Facebook friends. As such, the index is not a final outcome, but rather a potential mediator – namely, the social channel – of the effect of the intervention on users' self-attitudes.

Individual items were aggregated into indexes following Kling et al. (2007), i.e., we constructed equally weighted averages of the z-scores of the variables that enter each index.[15] For robustness, we also used a second method based on principal component analysis. Appendix Tables 3.A32 and 3.A33 describe the individual items used per index, their factor loadings and the Cronbach's alpha. Variables were oriented so that the impact of treatments on each component of the index should be positive. Therefore, consistent with the gender objectives of the videoclips, higher values of each index reflect more progressive views. To facilitate interpretation of impacts on the outcome indexes, we also report their standard deviations in the control group at the respective follow-up.[16]

3.3.4 Empirical specification

We conducted two separate analyses for the short-term (n=606) and medium-term (n=619) follow-up samples. It is worth mentioning that individuals in these two samples were not necessarily the same. In particular, 42% of those who completed the medium-term survey also filled the short-term survey. This should mitigate the concern that the results could be driven by over-exposure to the questionnaire and its interaction with the treatment (e.g., recall bias). To recover the average treatment effects of the two series in the short and medium terms, we estimated the following linear model via OLS:

$$Y_{i,t=\{1,2\}} = \alpha + \beta\,\text{Treatment1}_{i,t=0} + \gamma\,\text{Treatment2}_{i,t=0} + \delta\,\mathbf{X'}_{i,t=0} + \varepsilon_{i,t=\{1,2\}} \tag{3.1}$$

[15]For missing values, we also followed Kling et al. (2007). Particularly, if a respondent has a non-missing value for at least one of the variables in an index, we impute any missing values for the other variables using the random assignment group mean. This implies that differences between treatment and control means of an index coincide with the average of treatment and control means of the variables in that index (when divided by their standard deviations).

[16]Impact effects are divided by standard deviations derived from the control group at follow-up. An increase of two standard deviations, for example, would move someone from having average knowledge to being in the top 5 percent of the group.

where i is the individual, and t stands for the survey wave, namely, baseline ($t = 0$), short term ($t = 1$) and medium term ($t = 2$). Treatment 1 is a dummy indicator equal to 1 if the individual was assigned to the humorous drama with implicit messaging and 0 otherwise, while Treatment 2 equals 1 if the individual was assigned to the docu-series with explicit messaging and 0 otherwise. Y indicates the outcomes of interest, specifically, the general and topic indexes described previously as well as variables measuring information-seeking and posting behaviors.

Most of our outcomes were only collected at follow-ups. Given the experimental design, the lack of baseline values should not affect the causal interpretation of the results as both the treatment and control groups are identical in expectations (Bruhn and McKenzie, 2009). As shown below, this is further confirmed by observed balance for the pre-treatment values of observable characteristics and attitudinal self-reported outcomes. As such, we show plain estimates from a parsimonious specification with no controls. Nevertheless, to improve the efficiency of the estimator, we also report the results controlling for a series of socio-economic indicators measured at baseline, represented by vector **X'** in (3.1).[17] Among them, we always include an index measuring the baseline stance of the individual towards gender norms and VAW. Specifically, we rely on those few knowledge and attitudinal variables that were measured at baseline and aggregate them into an index – i.e., the Baseline Stance Index –, which captures the ex-ante progressiveness of the respondents' view.[18] Finally, when the baseline values of standalone outcomes of interest were available, we also included them on the RHS of equation (3.1) to explicitly account for potential pre-existing imbalances in those outcomes among arms.

For individuals assigned to the control group, both the drama and documentary dummies are simultaneously equal to 0. Hence, the β and γ coefficients capture the ATE of being assigned to the drama or documentary condition on the outcome of interest, respectively (or, alternatively, the ITT effects of the interventions). In the tables, we also report the p-value of a Wald test on the hypothesis that $\beta = \gamma$. As randomization was done at the

[17]Selected controls were theoretically associated with outcomes or showed pretreatment imbalances and included age, gender, education, education of household head, religion, caste, occupation, relationship status, self-assessment of the English language, frequency of watching videos online, indicators for being a student, having sisters, having male friends beating partner or female friends beaten, and city-of-residence fixed effects.

[18]Table 3.A31 describes the individual items used to construct this index, their factor loadings and the Cronbach's alpha.

individual level, standard errors were not clustered, yet we adjusted them to account for heteroscedasticity. Individuals who completed the survey too fast, or whose responses on gender and age were not consistent across surveys were excluded from the analysis.[19]

The primary analysis was done through Intention-to-Treat (ITT) estimates. ITT analysis replicates better what happens in the "real world", incorporating individuals' non-compliance or poor adherence to the program. As a result, ITT estimates provide a lower-bound of program impacts. Treatment-on-the-Treated effects (ToT) were also estimated using the objective measure of viewership. For this purpose we defined as compliers those individuals in the treatment arms who clicked play on at least half of assigned videoclips, as recorded by the Virtual Lab platform. We instrumented this measure of viewership using the random assignment to the treatment conditions and estimated ToT effects. Given that no individual in the control group received the treatment,[20] our ToT estimates equal the Local Average Treatment Effect (LATE) of the interventions. The analysis reports effect sizes and two-sided p-values. The text discusses only results that are statistically significant at the conventional level of $p < 0.10$.

We also conducted heterogeneous effects analysis. For this purpose, the moderator variables and their interactions with the treatment variable were added to the specification in (3.1) (Gerber and Green, 2012). Given that the analysis was not pre-registered, we restricted the number of variables to those with a relatively clear theoretical relationship. We hypothesized that study participants would benefit less from the program if (i) they perceived their friends to be more conservative at baseline, through the incentive for public compliance with social references (Miller and McFarland, 1987); and (ii) if they belonged to groups with lower social status, who may experience the constraining influence of social norms particularly strongly (Goffman 1963; Hoff and Stiglitz 2010; Hoff and Walsh 2018; World-Bank 2014). The latter included the participants' gender, age, caste, their and their parents' educational achievement and membership to a social group.

[19] Specifically, 21 and 32 observations were dropped in the short- and medium-term analysis, respectively.

[20] This was ensured by preventing participants from being able to share the video-clips. Moreover, for every video-clip the bot collected all Messenger IDs of the viewers. This gave the research team full control of compliance for all individuals and all video-clips.

3.4 Results

3.4.1 Response rates, randomization check and sample characteristics

Panel A of Table 3.1 shows that although 5,229 individuals filled the baseline survey, only about 12% of them completed either or both of the follow up surveys. Attrition in the medium-term sample affected each arm equally. A similar reasoning applies to the short-term sample, with the only difference that in this case the explicit (Treatment 2) arm exhibits a lower response rate due to a technical constraint triggered by a policy change in Facebook Messenger while we were conducting the study.[21]

The low follow-up response rates naturally affect the study's statistical power, especially for the medium-term survey, where the magnitude of the effects are generally smaller.[22] Low response rates could also jeopardize the external validity of the study if individuals with certain characteristics have a higher propensity to complete the study. Annex Table 3.A1 compares pre-intervention characteristics of those respondents who filled the follow up surveys with those who only completed the baseline and never continued. We find evidence of self-selection into study completion with respect to gender-related outcomes, though the different samples are generally similar across most socio-demographic characteristics.[23] Since respondents in the final samples had ex-ante more progressive

[21]The new policy, which has since been modified again, prevented chatbots from sending automated follow up messages. It went into effect one week after we began our study, which forced us to compress the timing and send the follow up survey exactly one week after the study began, before the new policy prevented the follow up from being sent. This disproportionately affected the documentary watchers, as they had more episodes to watch. While many were contacted manually later, they were still less likely to respond.

[22]The study's initial power calculations estimated a two-sided test with power of 0.8, alpha of 0.05 and no intra-cluster correlation for the question "Do you think a husband is justified in hitting or beating his wife if he suspects her of being unfaithful" among 18-24 individuals living in urban India (NFHS-4). To detect a six-percentage point increase in this outcome, each treatment arm required around 500 observations. However, each follow up survey had approximately 200 observations per arm.

[23]The two panel subsamples tend to exhibit overall more progressive attitudes than the only-baseline subsample, and the differences are generally statistically significant. For instance, panel subsamples are less likely to justify domestic violence and more willing to report it. This is a consequence of this study trying to replicate the real world, where media consumption is an individual choice and tends to be biased. At the same time, it also suggests that program content matters and raises the potential concern that even the implicit format (a humorous fake reality TV show) could fail to attenuate self-selection into viewership by reaching and engaging individuals with different interests. On the other hand, the different subsamples are generally similar across socio-demographic characteristics, with the exception of female respondents, respondents with sisters and respondents with high media consumption who are all more likely to complete the two follow ups.

views, we expect our estimates on knowledge and attitudes to represent a lower-bound of the campaign impact, because of the lower margins for improvements in the higher baseline values.

Table 3.1: Survey response, treatments' take-up and objective compliance

	All arms	*Treatment 1* Drama	*Treatment 2* Documentary	*Control* Placebo
	(1)	(2)	(3)	(4)
Panel A: Survey response				
Baseline respondents	5,229	1,791	1,783	1,655
Share (% out of baseline)	*100*	*100*	*100*	*100*
Short-term survey (1 week) respondents	606	258	128	220
Share (% out of baseline)	*11.59*	*14.41*	*7.18*	*13.29*
Medium-term survey (4 months) respondents	619	212	200	207
Share (% out of baseline)	*11.84*	*11.84*	*11.22*	*12.51*
Panel B: Overall take-up and compliance				
Self-reported to complete intervention	2,328	939	584	805
Share (% out of baseline)	*44.52*	*52.43*	*32.75*	*48.64*
Self-reported to watch half or more	3,153	1,169	831	1,153
Share (% out of baseline)	*60.30*	*65.27*	*46.61*	*69.67*
Actually played half or more	1,849	762	345	742
Share (% out of baseline)	*35.36*	*42.55*	*19.35*	*44.83*
Panel C: Short-term sample compliance				
Share of respondents who played half or more (%)	78.38	81.01	67.97	81.36
Mean duration of intervention (days)	8.9	8.4	10.0	8.8
Panel D: Medium-term sample compliance				
Share of respondents who played half or more (%)	66.07	75.00	49.00	73.43
Mean duration of intervention (days)	8.2	7.0	10.6	7.2

Notes: Panel B reports numbers for all respondents who completed the baseline survey. Panels C and D restrict the attention to respondents who completed the short-term and medium-term surveys, respectively. Treatment compliance is objectively measured by the bot. In particular, for each respondent and video, the bot recorded specific browsing events such as play, pause and end.

To assess whether attrition invalidated the randomization strategy, which may pose a threat to internal validity, Annex Tables 3.A2 and 3.A3 show differences in sample means between experimental groups of outcomes and covariates measured at baseline, only for those respondents who completed the short-term and medium-term follow up surveys, respectively. Overall, we find no evidence that the high attrition rates led to a differential self-selection into study completion across treatment assignments along observable characteristics. In fact, we generally observe baseline balance in both outcome and control variables between treatment and control groups for both follow-up samples, although this does not rule out potential imbalances across experimental groups in latent characteristics.[24] Pre-treatment differences between experimental groups are generally small and not statistically significant, including outcomes such as justification of violence and acceptance of pre-marital sex. For the few statistically significant differences, these tend to be small in magnitude and are accounted for in the analysis by the inclusion of the Baseline Stance Index.[25] Taken together, we find no evidence that the internal validity of the study is jeopardized by attrition bias (Dumville, Torgerson and Hewitt 2006).

Annex Table 3.A1 also provides a description of the short-term and medium-term samples. Overall, these samples consist of highly educated and typically unmarried individuals. More than four in ten report belonging to socially and economically disadvantaged castes. Individuals in the sample hold mixed attitudes towards gender norms and VAW. For example, while a large majority believes that women should be able to wear clothing of their choice, almost a quarter thinks that women should be banned from the kitchen/shrine during menstruation or justifies VAW in cases of unfaithfulness.

3.4.2 Take up rates and objective compliance

Panel B of Table 3.1 shows viewership statistics of the media campaigns for the full baseline sample. The data suggests that the drama (treatment 1) and placebo movie experienced higher viewership rates compared to the documentary (treatment 2), potentially due to their higher entertainment content. While 65% of treatment 1 viewers self-reported watching half or more clips, only 47% of treatment 2 individuals reported watching half or more clips. The objective metrics, as measured by click data, confirmed higher take-up rates for treatment 1, compared to treatment 2, and provided overall insights into intervention compliance. These show that around 25% of people in any arm over-reported watching more than half of the videoclips (35 vs. 60%), though over-reporting was more-or-less similar across treatment arms.[26]

[24] Annex Table 3.A4 shows balance among experimental conditions in the full baseline sample too.

[25] For standalone outcome variables that were also collected at baseline, we include their baseline values (instead of the Baseline Stance Index) on the RHS of equation (3.1) to control for potential pre-existing differences across treatment arms.

[26] We cannot fully attribute the take up difference to the entertainment format as the number of video-clips for the documentary was 7 as opposed to 3 for the drama, and video-clips were released every 2 hours (conditional on self-reporting that the previous one had been watched). Thus, the take up differences may be partly explained by a larger number of episodes. On the other hand, the total

Panels C and D of Table 3.1 provide information on compliance with the intervention for the two follow-up samples, using click data from the video "play" events. These data will also be used to estimate the Treatment-on-the-Treated effects, which thus will account for any differences in viewership across experimental arms. About 78% and 66% of respondents in the short-term and medium-term samples played half or more of the assigned series, respectively. Viewership patterns across treatment arms follow the general trends, with the docu-series being the least watched. In particular, information on mean duration of the intervention reported in panels C and D indicate that this explicit format lasted about 2 to 3 days more than the others.

3.4.3 Short-term impacts (one-week after exposure)

Panels A and B of Table 3.2 presents the Intent-to-Treat (ITT) and Treatment-on-the-Treated (ToT) estimates of the two series on the five indexes previously described. Higher values of the indexes indicate more progressive stances. Odd columns report plain estimates, while even columns add control variables described in Section 3.3.4 to the specification. For almost all outcome indexes, point-estimates for both specifications are very similar, which gives reassurance that selection issues are unlikely to affect our estimates. For readability, the text discusses results of the specification with control variables unless noted otherwise.

3.4.3.1 Self-reported outcomes

The implicit humorous drama (treatment 1) had economically and statistically significant effects on all outcome indexes in the short-term. The positive effect on the global index indicates a progressive overall shift of about 0.25 standard deviations (SDs). With respect to the control group, we observe improvements on knowledge and awareness of 0.21 SDs, improvements on gender norm/role attitudes of 0.20 SDs, and decreases in condoning violence against women of 0.16 SDs. Moreover, ToT estimates reported in Panel B of Table 3.2 exhibit even larger effects on program viewers, as coefficients for having played half or more of the assigned video-clips are about 25% higher than ITT estimates. This is especially important because if the effects were entirely driven by social desirability bias (i.e. respondents pleasing the researcher as they realized the link between the survey and exposure to an anti-VAW campaign), one would not necessarily

time was longer for the drama compared to the documentary (25 and 20 minutes respectively).

Table 3.2: Short-term impacts on outcome indexes

Panel A: ITT estimates

Dep. Var. (Y):	*Global index*		*Knowledge*		*Gender norms/roles*		*VAW attitudes*		*Beliefs on others*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Drama	0.188***	0.143***	0.199***	0.168***	0.182***	0.130***	0.192**	0.154**	-0.088	-0.117
	(0.051)	(0.035)	(0.068)	(0.064)	(0.058)	(0.042)	(0.089)	(0.074)	(0.086)	(0.086)
Documentary	0.091	0.068*	0.071	0.064	0.083	0.055	0.146	0.119	-0.242**	-0.307***
	(0.060)	(0.038)	(0.077)	(0.074)	(0.070)	(0.047)	(0.103)	(0.086)	(0.107)	(0.100)
Controls		✓		✓		✓		✓		✓
R-squared	0.020	0.585	0.012	0.122	0.014	0.516	0.005	0.338	0.005	0.086
P-value equal coef.	0.088	0.043	0.067	0.125	0.127	0.099	0.639	0.672	0.151	0.060
Observations	606	606	606	606	606	606	606	606	606	606
Mean Y (Control)	-0.059	-0.059	-0.034	-0.034	-0.062	-0.062	-0.084	-0.084	0.106	0.106
SD Y (Control)	0.582	0.582	0.785	0.785	0.660	0.660	0.991	0.991	0.904	0.904

Panel B: ToT estimates

Dep. Var. (Y):	*Global index*		*Knowledge*		*Gender norms/roles*		*VAW attitudes*		*Beliefs on others*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Play Drama	0.232***	0.179***	0.245***	0.210***	0.225***	0.163***	0.237**	0.193**	-0.109	-0.150
	(0.063)	(0.042)	(0.083)	(0.078)	(0.071)	(0.052)	(0.109)	(0.090)	(0.106)	(0.105)
Play Documentary	0.134	0.104*	0.105	0.097	0.122	0.083	0.215	0.180	-0.356**	-0.460***
	(0.087)	(0.056)	(0.112)	(0.108)	(0.101)	(0.068)	(0.150)	(0.126)	(0.163)	(0.154)
Controls		✓		✓		✓		✓		✓
R-squared	0.037	0.587	0.017	0.122	0.025	0.515	0.018	0.341	−0.025	0.049
P-value equal coef.	0.209	0.130	0.147	0.226	0.252	0.198	0.872	0.906	0.114	0.034
Observations	606	606	606	606	606	606	606	606	606	606
Wald F-statistic	265.5	231.3	265.5	231.3	265.5	231.3	265.5	231.3	265.5	231.3
Mean Y (Control)	-0.059	-0.059	-0.034	-0.034	-0.062	-0.062	-0.084	-0.084	0.106	0.106
SD Y (Control)	0.582	0.582	0.785	0.785	0.660	0.660	0.991	0.991	0.904	0.904

Notes: Heteroscedasticity-robust standard errors in parentheses. Higher values of the indexes in (1) to (8) stand for more progressive stances. The index "Beliefs on others" reported in (9) and (10) takes higher values when the respondents believe that a larger share (out of 10) of their closest Facebook friends have progressive attitudes. Controls include: Baseline Stance Index, age, gender, education, education of household head, religion, caste, occupation, relationship status, self-assessment of the English language, frequency of watching videos online, indicators for being a student, having sisters, having male friends beating partner or female friends beaten, and city-of-residence fixed effects. All controls are measured at baseline. In Panel B, independent variables Play Drama and Play Documentary take value 1 if the respondent has played half or more of the assigned video-clips, as objectively recorded by the bot. These variables are instrumented using the random assignment indicators to the treatment groups. The first-stage Wald F-statistic is reported below. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

expect them to be different based on actual viewing time.[27]

In the case of the explicit docu-series (treatment 2), the coefficient of interest is only statistically significant for the global index, again indicating a shift towards more progressive attitudes. However, its magnitude (0.12 SDs) is half compared to the coefficient of treatment 1 and the difference is statistically significant. For the other

[27]However, we cannot discard the possibility that observed impacts may be partly driven by our chatbot technology: having a chatbot inviting study participants to watch the different video-clips and then follow up with questions about the topic could be itself an effective mode of reflecting on the watched content and changing viewers' attitudes.

three subindexes, the coefficients are in the intended direction but are not statistically significant. This may be partly driven by the program's adverse effects on the social norm index (columns 9-10 of Table 3.2): for individuals in treatment 2, the program increased the perception that more of their Facebook friends had conservative stances by approximately 0.34 SDs, while the coefficient is statistically insignificant for treatment 1.[28] This unintended effect may potentially be explained by the awareness goal of the docu-series, which may have "normalized" the prevalence of conservative and masculine behaviors in India. As expected, ToT estimates indicate larger (adverse) effects on the social norm channel. Finally, Annex Tables 3.A15 shows that quantitatively similar coefficients are obtained when using indexes constructed with principal component analysis.

Annex Tables 3.A5 to 3.A8 present ITT results for the individual outcomes that compose the above indexes.[29] Specifically, Annex Table 3.A5 shows that the main messages delivered per treatment arm were absorbed by their respective audiences in a consistent manner with the content. While treatment 1 increased knowledge that fathers determine the sex of children (an increase of 12 percentage points (p.p.) or 38% with respect to the control's mean) and awareness of the prevalence of specific gender rituals like the menstruation and virginity rituals (an increase of 7.3 p.p. or 25%), treatment 2 made study participants aware that VAW is a major issue in India by 8.3 p.p., an increase of 11%.

While neither intervention had impacts on the broader and potentially harder to change attitudes towards gender roles (Annex Table 3.A6), treatment 1 impacted most outcomes related to challenging existing gender norms (Annex Table 3.A7). In particular, treatment 1 made respondents more critical of blindly following adverse social norms by 7.2 p.p. (or 14% with respect to the control), less likely by 7 p.p. (9% decrease) in believing that women should be virgin until marriage, less likely by 8.3 p.p. (24% decrease) to believe women should be banned from the kitchen or household shrine during menstruation, and more likely by 6 p.p. (7% increase) to believe that women should be able to wear whatever they want without fear of sexual harassment. The latter outcome was also

[28] Annex Table 3.A9 reports impacts on individual items that make up the social norms index. It shows that an increased number of individuals in treatment 2 believed that their Facebook friends were against pre-marital sex for men, and that their friends would condone domestic violence when wives were either unfaithful or went out without permission of their husbands.

[29] For the sake of conciseness, we omit results for individual items when they are statistically insignificant.

affected by treatment 2 (an increase of 6.8 p.p., or 8%), which had an explicit episode on women's freedom of dressing. Yet, treatment 2 had no other short-term impacts on outcomes related to challenging existing gender norms.

On outcomes measuring attitudes that condone VAW (Annex Table 3.A8), treatment 1 decreased the likelihood of individuals justifying domestic violence if a wife went out without her husband's permission by 7.5 p.p (27% decrease), and increased the likelihood that individuals would not be passive bystanders. The treatment group was 9.2 p.p more willing to report if a friend experienced physical violence, an increase of 11% with respect to the control. Again, treatment 2 had no effects on attitudes that condone VAW.

3.4.3.2 Content sharing intentions

In line with its call-to-action messages, treatment 2 was effective in increasing participants' willingness to share the campaign videoclips with their Facebook friends.[30] Specifically, columns (1-2) of Table 3.3 show that individuals assigned to treatment 2 were 5 p.p. more likely to report a greater willingness to share the docu-series with their Facebook friends right after the intervention, an increase of 9.4% with respect to the control. The significant differences between treatment arms indicate that only the docu-series was effective, while treatment 1 was not. ToT estimates reported in columns (3)-(4) are in line with the previous findings and show that the effect of the docu-series on viewers precisely doubled.

3.4.3.3 Information-seeking behaviors

Panel A of Table 3.4 shows that both treatment arms were effective in promoting information-seeking behaviors, with effects generally being larger for treatment 2 (though impact differences between treatment arms are not statistically significant). Individuals in treatment 1 and treatment 2 were respectively 7.3 and 10 p.p. more likely to click on both gender-equality-related websites provided to them at the end of the survey, which represented increases of 85 and 116%, respectively.

The impact on time spent visiting these websites was almost twice as large for those treated with treatment 2 compared to treatment 1. While individuals in treatment 1 spent on average an additional 2 minutes on the websites (compared to the control's

[30]Note that this is just a measure of willingness to share the clips, and not actual sharing/posting.

Table 3.3: Impact on willingness to share the videos right after the intervention

Dep. Var. (Y):	*Willing to share videos with Facebook friends*			
	(1)	(2)	(3)	(4)
	ITT	ITT	ToT	ToT
Drama	-0.023	-0.025		
	(0.024)	(0.024)		
Documentary	0.060**	0.050*		
	(0.027)	(0.027)		
Play Drama			-0.035	-0.037
			(0.037)	(0.036)
Play Documentary			0.125**	0.105*
			(0.057)	(0.057)
Controls		✓		✓
R-squared	0.003	0.033	−0.001	0.029
P-value equal coef.	0.002	0.005	0.002	0.005
Observations	2269	2269	2269	2269
Wald F-statistic			444.5	432.9
Mean Y (Control)	0.529	0.529	0.529	0.529

Notes: The sample considered here is made of all participants who self-reported to complete the intervention. Heteroscedasticity-robust standard errors in parentheses. Controls are described in the notes to 3.2. In columns (3)-(4), independent variables Play Drama and Play Documentary take value 1 if the respondent has played half or more of the assigned video-clips, as objectively recorded by the bot. These variables are instrumented using the random assignment indicators to the treatment groups. The first-stage Wald F-statistic is reported below. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

average of half a minute), individuals in treatment 2 spent an additional 3.4 minutes, which is 5.6 times longer than the average visit duration in the control group. As expected, we observe no significant effects on the likelihood of clicking climate-change-related websites nor on the time study participants spent on these websites. These findings are confirmed in the ToT analysis reported in Panel B, where the magnitudes of the estimated coefficients are about 25 to 50% higher than ITT results.

3.4.4 Medium-term impacts (four months after exposure)

3.4.4.1 Self-reported outcomes

Table 3.5 presents medium-term results of ITT and ToT estimates. Most coefficients are in the direction of more progressive attitudes towards gender norms. However, the data suggests a time-decay in effects, with program impacts decreasing over time in

Table 3.4: Short-term impacts on clicks on informative links

Panel A: ITT estimates

Dep. Var. (Y):	*Click gender-links*		*Click climate-links*		*Duration gender-link*		*Duration climate-link*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Drama	0.080***	0.073**	0.028	0.024	108.802*	118.461*	2.892	-1.403
	(0.030)	(0.031)	(0.031)	(0.032)	(60.060)	(62.303)	(68.719)	(56.626)
Documentary	0.101**	0.100***	0.041	0.044	181.815**	204.119**	29.871	17.582
	(0.039)	(0.039)	(0.040)	(0.039)	(91.531)	(103.121)	(89.654)	(84.350)
Controls		✓		✓		✓		✓
R-squared	0.012	0.022	−0.001	0.024	0.005	0.002	−0.003	0.035
P-value equal coef.	0.617	0.512	0.745	0.595	0.482	0.451	0.746	0.787
Observations	606	606	606	606	543	543	554	554
Mean Y (Control)	0.086	0.086	0.123	0.123	36.371	36.371	116.112	116.112

Panel B: ToT estimates

Dep. Var. (Y):	*Click gender-links*		*Click climate-links*		*Duration gender-link*		*Duration climate-link*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Play Drama	0.099***	0.092**	0.035	0.030	136.594*	153.456**	3.642	-1.554
	(0.037)	(0.037)	(0.039)	(0.038)	(75.036)	(76.877)	(86.312)	(70.611)
Play Documentary	0.149***	0.150***	0.061	0.067	261.637**	300.831**	43.312	25.846
	(0.057)	(0.056)	(0.058)	(0.057)	(130.735)	(145.898)	(129.589)	(121.172)
Controls		✓		✓		✓		✓
R-squared	0.021	0.032	0.007	0.032	0.012	0.009	−0.003	0.035
P-value equal coef.	0.391	0.297	0.641	0.492	0.383	0.339	0.734	0.782
Observations	606	606	606	606	543	543	554	554
Wald F-statistic	265.5	231.3	265.5	231.3	309.5	253.3	292.3	252.5
Mean Y (Control)	0.086	0.086	0.123	0.123	36.371	36.371	116.112	116.112

Notes: Heteroscedasticity-robust standard errors in parentheses. In columns (1)-(2), the dependent variable takes value 1 if the respondent clicked on both gender links (PFI and UN women India). In columns (3)-(4), the dependent variable takes value 1 if the respondent clicked on both climate links (Delhi Green and UN environment program India). Visit duration refers to the PFI website in columns (5)-(6) and to the Delhi Green website in columns (7)-(8). Controls are described in the notes to 3.2. In Panel B, independent variables Play Drama and Play Documentary take value 1 if the respondent has played half or more of the assigned video-clips, as objectively recorded by the bot. These variables are instrumented using the random assignment indicators to the treatment groups. The first-stage Wald F-statistic is reported below. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

both magnitude and statistical significance.[31]

For treatment 1, we only observe statistically significant effects for the index for attitudes toward gender norms/roles. The medium-term ITT estimate indicates an increase of 0.15 standard deviations with respect to the control, approximately three-fourths of the short-term magnitude. The ToT estimate indicates that the effect of the treatment on those who played half or more of the series is 0.19 SDs. For treatment 2, similar to our short-term results, we see no statistical effects on topic indexes and the global index.

[31] Because the short- and medium-term panels have very similar sample sizes (n=606 and 619 respectively), a larger sample size would be required to statistically detect the smaller effect sizes observed in the medium term survey.

Table 3.5: Medium-term impacts on outcome indexes

Dep. Var. (Y):	*Global index*		*Knowledge*		*Gender norms/roles*		*VAW attitudes*		*Beliefs on others*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	ITT	ToT	ITT	ToT	ITT	ToT	ITT	ToT	ITT	ToT
Drama	0.046		-0.031		0.086**		0.028		0.042	
	(0.035)		(0.061)		(0.042)		(0.068)		(0.089)	
Documentary	0.028		0.026		0.035		0.015		0.097	
	(0.034)		(0.060)		(0.040)		(0.072)		(0.085)	
Play Drama		0.061		-0.041		0.114**		0.037		0.056
		(0.045)		(0.079)		(0.054)		(0.088)		(0.115)
Play Documentary		0.056		0.053		0.070		0.030		0.197
		(0.068)		(0.118)		(0.080)		(0.144)		(0.169)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	0.537	0.542	0.129	0.129	0.480	0.485	0.290	0.292	0.074	0.079
P-value equal coef.	0.583	0.929	0.367	0.386	0.191	0.520	0.856	0.958	0.529	0.342
Observations	619	619	619	619	619	619	619	619	619	619
Wald F-stat		116.9		116.9		116.9		116.9		116.9
Mean Y (Control)	-0.019	-0.019	0.018	0.018	-0.028	-0.028	-0.033	-0.033	-0.049	-0.049
SD Y (Control)	0.530	0.530	0.628	0.628	0.594	0.594	0.839	0.839	0.903	0.903

Notes: Heteroscedasticity-robust standard errors in parentheses. Higher values of the indexes in (1) to (8) stand for more progressive stances. The index "Beliefs on others" reported in (9) and (10) takes higher values when the respondents believe that a larger share (out of 10) of their closest Facebook friends have progressive attitudes. Controls are described in the notes to 3.2. Independent variables Play Drama and Play Documentary take value 1 if the respondent has played half or more of the assigned video-clips, as objectively recorded by the bot. These variables are instrumented using the random assignment indicators to the treatment groups. The first-stage Wald F-statistic is reported below. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Interestingly, the social norms index is no longer negative (as in the short-term survey), though the effect is not statistically significant. Finally, Annex Tables 3.A16 shows that using indexes constructed with principal component analysis yields quantitatively similar results.

Annex Tables 3.A10 to 3.A13 present ITT results for the individual outcomes that compose the above topic indexes. Results are consistent with the time-decay explanation, with the size of coefficients and their statistical significance decreasing in the medium term. Nevertheless, some impacts persist. In particular, the edutainment program is found to make viewers less likely by 13 p.p. (which is equivalent to 33% of the mean in the control) to believe women should be banned from the kitchen or household shrine during menstruation and 7.3 p.p. less likely (24% decrease) to think that it is more important that a boy goes to school than a girl. Treatment 2 also impacted the latter outcome by 8.8 p.p. (30% decrease) and increased the awareness of VAW being an issue in India by 7.8 p.p. (10% increase). For the latter, the coefficient is similar in magnitude to the short-term estimate.

3.4.4.2 Information-seeking and posting behaviors

The short-term effects on information-seeking behaviors disappeared in the medium term (Annex Table 3.A14).[32] On the other hand, treated participants were more willing than control participants to publicly display their disapproval of VAW in the medium term.[33] Columns (1)-(3) of Table 3.6 show that treatment 2 made participants more likely to intend to add the frame against VAW in their Facebook profile picture. While ITT estimates indicate an impact of 7.9 p.p. (32% increase), ToT estimates are twice as large (16 p.p.). Results are much smaller in magnitude and not statistically significant for treatment 1.

Most importantly, treatment 2 was effective in making individuals actually update their profile picture. ITT estimates indicate that participants in treatment 2 were 7.5 p.p. more likely than the control to add the VAW frame "End violence against women" to their picture, an increase of 91% ($p<0.05$). ToT estimates were twice as large: individuals who watched at least half of the documentaries were 15.3 p.p. more likely to add the banner, an increase of almost 190% ($p<0.05$). For treatment 1, the observed increases are not statistically significant.

Although these impacts were recorded only four months after program exposure, we were able to observe their potential cascade effects within the social network in the long-run. We found that after about 1 year and a half from the medium-term survey – i.e., the first time the frame was presented to the respondents –, the frame was used by more than 34 thousand people all over the world.[34] This is 55 times larger than the sample receiving the frame at first, and almost 500 times larger than the number of respondents who used it initially. Back-of-the-envelope calculations suggest that treatment 2 alone was responsible for about 6,300 of such uses. These results demonstrate the power of social media in amplifying content and opinions through online sharing, and lend support to the possibility of achieving a self-sustained virtuous circle of social change.

[32] We discard the possibility that recall bias from the short-term survey drives this result. In fact, we find no medium-term information-seeking effects even when focusing on the subsample of individuals who were not interviewed at the short-term followup (58%).
[33] This outcome was only measured in the four-month data collection.
[34] The figure was objectively measured by the Facebook Frame Manager.

Table 3.6: Medium-term impacts on updating profile picture

Dep. Var. (Y):	*Intent to update*			*Actual picture update*		
	(1)	(2)	(3)	(4)	(5)	(6)
	ITT	ITT	ToT	ITT	ITT	ToT
Drama	0.056	0.054		0.026	0.029	
	(0.044)	(0.044)		(0.029)	(0.029)	
Documentary	0.079*	0.079*		0.068**	0.075**	
	(0.045)	(0.047)		(0.032)	(0.034)	
Play Drama			0.071			0.039
			(0.057)			(0.037)
Play Documentary			0.161*			0.153**
			(0.093)			(0.067)
Controls		✓	✓		✓	✓
R-squared	0.002	0.011	0.012	0.004	0.001	0.002
P-value equal coef.	0.614	0.588	0.281	0.211	0.192	0.069
Observations	619	619	619	619	619	619
Wald F-stat			117.0			117.0
Mean Y (Control)	0.246	0.246	0.246	0.082	0.082	0.082

Notes: Heteroscedasticity-robust standard errors in parentheses. Controls are described in the notes to 3.2. Independent variables Play Drama and Play Documentary take value 1 if the respondent has played half or more of the assigned video-clips, as objectively recorded by the bot. These variables are instrumented using the random assignment indicators to the treatment groups. The first-stage Wald F-statistic is reported below. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.4.5 Treatment effect heterogeneity

Annex Tables 3.A22 and 3.A29 present our post-hoc heterogenous analysis for participants' perceptions of the conservatism of their Facebook friends prior to the intervention. We study these effects to test our hypothesis that perceptions of public acceptance of retrograde norms affects treatment impacts (Miller and McFarland, 1987). For the short- and medium-term samples, we generally find that both treatments had smaller impacts on gender norms attitudes and VAW indexes for individuals who reported having friends with more conservative views on gender-related topics. This result is consistent with theories suggesting that people's beliefs and behaviors are influenced by their perception of prevalent social norms (e.g., Bicchieri 2005, 2016).

Annex Tables 3.A17 to 3.A30 present heterogeneous effects for a series of social status indicators, where we hypothesized that effects would be smaller for groups with lower social status (World-Bank 2014). We find no evidence of heterogeneous effects by individuals' age, caste, educational achievement or membership to a social group in the short-term. However, both treatment arms generally were less effective for women and for individuals whose household-heads were on average more educated. Heterogeneous effects across gender persisted in the medium term, while those across household-head education vanished. Contrary to our hypothesis, in the medium-term we observe larger effects for less educated individuals in terms of the global index as well as the gender norms/roles and VAW indexes. This could be explained by the lower baseline values of the outcomes for these groups of individuals, which therefore had higher margins of improvements.

3.4.6 Placebo estimates

To discard social desirability affecting the responses of study participants, both follow up surveys included a series of placebo questions that should not be affected by our treatments (unless respondents were trying to please the researchers with an expected answer). Table 3.7 presents impacts on placebo outcomes, such as thinking that climate change is a threat to humankind, willingness to vote for better fuel-efficient cars, thinking to be working in a paid job in two years' time, or thinking that corruption is an issue in India. We find no evidence of social desirability bias for the short-term (Panel A) and the medium term (Panel B) samples.

3.5 Discussion

Social media platforms engaged around 4.1 billion users in 2020, more than half of the world's population (Digital 2020). By reaching large segments of the population, social media could potentially be used to correct beliefs' distortions, challenge gender stereotypes, and discuss misconceptions about socially harmful practices at scale and at low cost. Complementing evidence that shows the effectiveness of edutainment at reshaping VAW attitudes and behaviors in community settings, our study shows that social media, with its more private consumption of information, can also be an effective medium. In addition to testing a new delivery mechanism – namely, Facebook Messenger versus community screenings or radio –, we are able to objectively measure new behavioral outcomes (i.e., likelihood to click and time spent on pro-gender-equality websites and public displays indicating disapproval of VAW by adding a frame to Facebook profile pictures).

Our study provides empirical evidence that edutainment delivered through social media

Table 3.7: Placebo outcomes

Panel A: Short-term sample

Dep. Var. (Y):	*Climate change is threat*		*Vote for fuel efficiency*		*Think will be working*	
	(1)	(2)	(3)	(4)	(5)	(6)
Drama	-0.011	-0.015	0.003	-0.001	0.008	0.013
	(0.035)	(0.035)	(0.046)	(0.046)	(0.034)	(0.034)
Documentary	-0.034	-0.038	0.019	0.018	-0.001	0.018
	(0.045)	(0.044)	(0.056)	(0.055)	(0.043)	(0.042)
$Y_{baseline}$	0.501***	0.467***			0.577***	0.576***
	(0.038)	(0.041)			(0.040)	(0.041)
Controls		✓		✓		✓
R-squared	0.271	0.280	−0.003	0.021	0.310	0.313
P-value equal coef.	0.606	0.591	0.772	0.720	0.832	0.908
Observations	606	606	606	606	606	606
Mean Y (Control)	0.741	0.741	0.505	0.505	0.709	0.709

Panel B: Medium-term sample

Dep. Var. (Y):	*Climate change is threat*		*Vote for fuel efficiency*		*Think will be working*		*Corruption is an issue*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Drama	-0.000	-0.011	-0.007	-0.010	0.056	0.045	0.012	0.008
	(0.040)	(0.039)	(0.049)	(0.050)	(0.046)	(0.047)	(0.031)	(0.033)
Documentary	0.021	0.029	0.065	0.072	0.005	0.005	0.001	0.007
	(0.041)	(0.040)	(0.050)	(0.050)	(0.047)	(0.048)	(0.032)	(0.032)
$Y_{baseline}$	0.424***	0.380***			0.001	-0.012		
	(0.040)	(0.044)			(0.041)	(0.043)		
Controls		✓		✓		✓		✓
R-squared	0.180	0.212	0.001	0.009	−0.002	0.022	−0.003	0.032
P-value equal coef.	0.597	0.328	0.147	0.108	0.267	0.393	0.727	0.987
Observations	617	617	617	617	617	617	617	617
Mean Y (Control)	0.714	0.714	0.490	0.490	0.650	0.650	0.879	0.879

Notes: Heteroscedasticity-robust standard errors in parentheses. Controls are described in the notes to 3.2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

can be an effective tool for reshaping gender norms and VAW attitudes. Our results show that in the short-term (1 week after the intervention), the drama was effective at increasing knowledge and awareness of gender practices and shifting gender norms towards more progressive stances. Moreover, both implicit and explicit formats increased short-term information seeking behaviors on the web, yet most of our outcomes experience time-decay effects in the medium-term (four months after the intervention). On the other hand, the docu-series was successful in encouraging social media users to take a public stance within their online communities against VAW in the medium-term. This treated group was more likely (i) to report a willingness to share the video-clips with their Facebook friends; and (ii) to add the frame "End violence against women" to their

Facebook profile picture. In our post-hoc analysis of treatment effect heterogeneity, we observe smaller effects for females and individuals that at baseline perceived their Facebook friends to be more conservative. However, we find no evidence that the intervention had differential effects for a series of social status indicators, including age, caste, membership to a social organization, and individual and parental educational achievement.

Some knowledge gaps that follow from this study's findings would benefit from further research. This study incentivized individuals to watch edutainment videos, whereas a real-world campaign would rely on users discovering the content through ad campaigns. Future experimental research should address this gap by randomizing and evaluating social media campaigns at a level more relevant for marketing campaigns (e.g., neighborhoods or media markets). Future VAW research should also scale up the use of innovative online measurements, such as crowdsourcing safety data from mobile applications (e.g., Borker et al. 2021; Kondylis et al. 2020). In light of our findings that large effects diminished in the medium-term, greater research is needed to understand how best to design reinforcer campaigns for long-term impacts.

To conclude, this study provides experimental evidence that social media campaigns that used "low-touch" edutainment were effective at influencing gender norms and acceptability of VAW This evidence is particularly promising in low-resource settings, where resource-intensive campaigns may be costly to scale. The heterogeneous analysis provides indicative evidence that vulnerable populations are not systematically less likely to benefit from such campaigns. Considering the power of social media in amplifying content and opinions through online sharing, our results lend support to the possibility of achieving a self-sustained virtuous circle of social change.

3.6 Appendices

3.A Tables

Table 3.A1: Self-selection into completion of follow-up surveys

	Mean Complete only baseline N=4232 (1)	*Mean* Complete short term N=606 (2)	*Mean* Complete medium term N=619 (3)	*Norm.Diff.* Complete ST vs. baseline (4)	*Norm.Diff.* Complete MT vs. baseline (5)	*Diff=0 (p-value)* Complete ST vs. baseline (6)	*Diff=0 (p-value)* Complete MT vs. baseline (7)
Panel A: Outcomes							
Baseline index	-0.029	0.091	0.124	0.164	0.209	0.000	0.000
Father determines sex	0.268	0.281	0.320	0.020	0.081	0.511	0.009
Stricter control daugthers	0.556	0.491	0.464	-0.092	-0.131	0.003	0.000
Women should be virgin	0.742	0.720	0.706	-0.036	-0.056	0.281	0.088
Justify beating if unfaith	0.370	0.294	0.266	-0.115	-0.160	0.000	0.000
Women wear whatever	0.952	0.965	0.956	0.048	0.013	0.103	0.665
Ban kitchen during period	0.382	0.298	0.305	-0.126	-0.116	0.000	0.000
Tell anyone if friend beat	0.854	0.895	0.916	0.088	0.138	0.004	0.000
Climate change is a threat	0.598	0.670	0.696	0.105	0.146	0.000	0.000
Work in the future	0.720	0.744	0.721	0.038	0.000	0.212	0.998
Panel B: Controls							
Age (years)	20.939	20.969	20.889	0.011	-0.018	0.725	0.547
English self-assess (0-10)	5.449	5.576	5.667	0.029	0.050	0.338	0.092
Survey in english	0.372	0.368	0.449	-0.006	0.111	0.833	0.000
Female	0.240	0.272	0.275	0.052	0.056	0.097	0.073
Primary	0.073	0.069	0.050	-0.010	-0.067	0.754	0.019
Secondary	0.437	0.413	0.425	-0.036	-0.018	0.246	0.557
University	0.462	0.492	0.498	0.042	0.050	0.177	0.102
HH-head primary	0.186	0.208	0.183	0.039	-0.006	0.206	0.849
HH-head secondary	0.338	0.325	0.321	-0.020	-0.025	0.522	0.408
HH-head university	0.357	0.353	0.384	-0.005	0.041	0.869	0.182
Hindu	0.804	0.809	0.827	0.008	0.042	0.783	0.156
Muslim	0.134	0.149	0.116	0.029	-0.038	0.352	0.198
Christian	0.012	0.003	0.008	-0.068	-0.025	0.004	0.376
Sikh	0.031	0.023	0.024	-0.034	-0.029	0.238	0.318
General caste	0.534	0.526	0.564	-0.010	0.043	0.742	0.157
OBC caste	0.285	0.284	0.252	-0.001	-0.052	0.963	0.082
SC caste	0.133	0.152	0.152	0.039	0.039	0.214	0.209
Student	0.622	0.652	0.667	0.043	0.066	0.156	0.028
Employed	0.132	0.130	0.124	-0.004	-0.016	0.906	0.589
Self-employed	0.093	0.081	0.079	-0.031	-0.035	0.305	0.235
Live in Delhi	0.562	0.587	0.559	0.036	-0.005	0.237	0.882
Member of organization	0.180	0.135	0.155	-0.087	-0.047	0.003	0.112
Currently dating	0.334	0.333	0.309	-0.000	-0.038	0.988	0.208
Married	0.065	0.061	0.050	-0.011	-0.045	0.724	0.125
With sisters	0.675	0.733	0.709	0.090	0.053	0.003	0.078
Daily freq. social media	15.558	15.917	15.845	0.023	0.018	0.486	0.578
Daily freq. watch videos	2.288	2.389	2.388	0.067	0.066	0.023	0.026
Male friend beating	0.146	0.140	0.157	-0.012	0.021	0.691	0.503
Female friend beated	0.146	0.155	0.158	0.019	0.025	0.542	0.415

Notes: Table shows sample means at baseline for different categories of respondents: those who only completed baseline, those who completed the short-term survey and those who completed the medium-term survey.

Table 3.A2: Balance of baseline outcomes and covariates for those completing the short-term survey

	Mean Drama N=258 (1)	*Mean* Document. N=128 (2)	*Mean* Control N=220 (3)	*Norm.Diff.* Drama vs.Control (4)	*Norm.Diff.* Document. vs.Control (5)	*Diff=0 (p-value)* Drama vs.Control (6)	*Diff=0 (p-value)* Document. vs.Control (7)
Panel A: Outcomes							
Baseline index	0.130	0.085	0.049	0.110	0.049	0.091	0.534
Father determines sex	0.267	0.312	0.277	-0.016	0.054	0.810	0.490
Stricter control daugthers	0.439	0.520	0.533	-0.134	-0.018	0.044	0.822
Women should be virgin	0.689	0.728	0.751	-0.099	-0.037	0.153	0.657
Justify beating if unfaith	0.286	0.306	0.296	-0.015	0.017	0.821	0.836
Women wear whatever	0.976	0.959	0.956	0.076	0.011	0.260	0.894
Ban kitchen during period	0.284	0.291	0.318	-0.053	-0.043	0.435	0.603
Tell anyone if friend beat	0.866	0.913	0.917	-0.115	-0.010	0.087	0.902
Climate change is a threat	0.663	0.641	0.695	-0.049	-0.082	0.446	0.299
Work in the future	0.752	0.750	0.732	0.032	0.029	0.618	0.709
Panel B: Controls							
Age (years)	20.938	21.062	20.950	-0.004	0.040	0.946	0.616
English self-assess (0-10)	5.651	5.656	5.441	0.049	0.049	0.452	0.532
Survey in english	0.399	0.352	0.341	0.085	0.016	0.188	0.841
Female	0.287	0.258	0.264	0.037	-0.009	0.572	0.905
Primary	0.062	0.102	0.059	0.009	0.110	0.894	0.174
Secondary	0.368	0.422	0.459	-0.131	-0.053	0.045	0.501
University	0.535	0.453	0.464	0.101	-0.015	0.121	0.850
HH-head primary	0.209	0.250	0.182	0.049	0.117	0.450	0.143
HH-head secondary	0.318	0.289	0.355	-0.055	-0.099	0.399	0.205
HH-head university	0.372	0.367	0.323	0.073	0.066	0.259	0.403
Hindu	0.810	0.797	0.814	-0.006	-0.030	0.921	0.706
Muslim	0.147	0.148	0.150	-0.005	-0.003	0.934	0.969
Christian	0.000	0.000	0.009	-0.096	-0.096	0.157	0.158
Sikh	0.016	0.047	0.018	-0.015	0.114	0.822	0.169
General caste	0.531	0.570	0.495	0.050	0.106	0.439	0.177
OBC caste	0.271	0.273	0.305	-0.052	-0.048	0.426	0.537
SC caste	0.171	0.125	0.145	0.049	-0.042	0.453	0.589
Student	0.612	0.727	0.655	-0.062	0.110	0.341	0.158
Employed	0.159	0.102	0.114	0.093	-0.027	0.149	0.725
Self-employed	0.078	0.062	0.095	-0.045	-0.086	0.490	0.261
Live in Delhi	0.593	0.602	0.573	0.029	0.041	0.655	0.599
Member of organization	0.136	0.117	0.145	-0.020	-0.059	0.760	0.447
Currently dating	0.357	0.344	0.300	0.085	0.066	0.189	0.403
Married	0.070	0.039	0.064	0.017	-0.079	0.789	0.303
With sisters	0.756	0.734	0.705	0.082	0.047	0.210	0.550
Daily freq. social media	16.426	14.747	15.975	0.029	-0.079	0.675	0.354
Daily freq. watch videos	2.356	2.275	2.493	-0.097	-0.151	0.137	0.058
Male friend beating	0.163	0.133	0.118	0.091	0.031	0.160	0.694
Female friend beated	0.147	0.133	0.177	-0.057	-0.087	0.378	0.263

Notes: Table shows sample means at baseline for respondents who completed the short-term survey. The normalized difference in columns 4 and 5 is the difference in the sample means of treatment and control groups divided by the square root of the sum of the sample variances.

Table 3.A3: Balance of baseline outcomes and covariates for those completing the medium-term survey

	Mean Drama N=212 (1)	*Mean* Document. N=200 (2)	*Mean* Control N=207 (3)	*Norm.Diff.* Drama vs.Control (4)	*Norm.Diff.* Document. vs.Control (5)	*Diff=0 (p-value)* Drama vs.Control (6)	*Diff=0 (p-value)* Document. vs.Control (7)
Panel A: Outcomes							
Baseline index	0.161	0.076	0.131	0.040	-0.075	0.562	0.287
Father determines sex	0.316	0.335	0.309	0.010	0.039	0.880	0.578
Stricter control daugthers	0.423	0.513	0.457	-0.048	0.079	0.496	0.269
Women should be virgin	0.715	0.724	0.681	0.052	0.066	0.477	0.374
Justify beating if unfaith	0.250	0.286	0.262	-0.019	0.039	0.792	0.593
Women wear whatever	0.980	0.941	0.944	0.137	-0.007	0.055	0.920
Ban kitchen during period	0.273	0.305	0.337	-0.098	-0.049	0.178	0.512
Tell anyone if friend beat	0.918	0.879	0.948	-0.085	-0.174	0.249	0.021
Climate change is a threat	0.670	0.700	0.720	-0.077	-0.031	0.267	0.661
Work in the future	0.708	0.750	0.705	0.003	0.071	0.960	0.312
Panel B: Controls							
Age (years)	20.816	20.960	20.894	-0.029	0.024	0.680	0.731
English self-assess (0-10)	5.552	5.775	5.681	-0.030	0.022	0.663	0.754
Survey in english	0.443	0.505	0.401	0.061	0.148	0.380	0.035
Female	0.264	0.265	0.295	-0.048	-0.047	0.487	0.506
Primary	0.042	0.065	0.043	-0.004	0.067	0.959	0.340
Secondary	0.382	0.415	0.478	-0.138	-0.090	0.047	0.200
University	0.542	0.500	0.449	0.132	0.072	0.057	0.307
HH-head primary	0.208	0.160	0.179	0.052	-0.035	0.456	0.615
HH-head secondary	0.354	0.315	0.295	0.089	0.031	0.197	0.657
HH-head university	0.344	0.420	0.391	-0.069	0.041	0.320	0.557
Hindu	0.844	0.810	0.826	0.035	-0.029	0.616	0.675
Muslim	0.099	0.120	0.130	-0.070	-0.022	0.315	0.751
Christian	0.014	0.005	0.005	0.068	0.002	0.325	0.981
Sikh	0.014	0.045	0.014	-0.002	0.127	0.977	0.072
General caste	0.557	0.595	0.541	0.022	0.077	0.750	0.273
OBC caste	0.269	0.235	0.251	0.028	-0.027	0.681	0.704
SC caste	0.151	0.135	0.169	-0.035	-0.067	0.614	0.339
Student	0.651	0.720	0.633	0.027	0.132	0.700	0.060
Employed	0.137	0.120	0.116	0.044	0.009	0.522	0.899
Self-employed	0.113	0.045	0.077	0.086	-0.095	0.211	0.174
Live in Delhi	0.571	0.530	0.575	-0.006	-0.064	0.932	0.364
Member of organization	0.160	0.170	0.135	0.050	0.068	0.470	0.332
Currently dating	0.307	0.275	0.343	-0.055	-0.104	0.428	0.138
Married	0.042	0.050	0.058	-0.050	-0.025	0.469	0.723
With sisters	0.703	0.740	0.686	0.026	0.084	0.709	0.229
Daily freq. social media	15.924	14.896	16.721	-0.050	-0.113	0.492	0.126
Daily freq. watch videos	2.368	2.296	2.496	-0.089	-0.138	0.200	0.050
Male friend beating	0.160	0.205	0.106	0.113	0.194	0.104	0.006
Female friend beated	0.132	0.210	0.135	-0.007	0.140	0.924	0.047

Notes: Table shows sample means at baseline for respondents who completed the medium-term survey. The normalized difference in columns 4 and 5 is the difference in the sample means of treatment and control groups divided by the square root of the sum of the sample variances.

Table 3.A4: Balance of baseline outcomes and covariates for those who completed baseline survey

	Mean Drama N=1791	*Mean* Document. N=1783	*Mean* Control N=1655	*Norm.Diff.* Drama vs.Control	*Norm.Diff.* Document. vs.Control	*Diff=0* *(p-value)* Drama vs.Control	*Diff=0* *(p-value)* Document. vs.Control
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Outcomes							
Baseline index	0.002	-0.007	-0.008	0.014	0.002	0.552	0.942
Father determines sex	0.266	0.278	0.280	-0.022	-0.004	0.356	0.858
Stricter control daugthers	0.543	0.544	0.540	0.004	0.005	0.873	0.832
Women should be virgin	0.730	0.727	0.758	-0.045	-0.050	0.081	0.053
Justify beating if unfaith	0.356	0.362	0.346	0.015	0.025	0.548	0.325
Women wear whatever	0.950	0.956	0.955	-0.017	0.007	0.507	0.792
Ban kitchen during period	0.359	0.371	0.370	-0.016	0.001	0.528	0.965
Tell anyone if friend beat	0.859	0.856	0.870	-0.023	-0.028	0.380	0.272
Climate change is a threat	0.602	0.609	0.627	-0.035	-0.025	0.146	0.291
Work in the future	0.706	0.744	0.716	-0.015	0.044	0.530	0.068
Panel B: Controls							
Age (years)	20.915	20.980	20.915	-0.000	0.024	0.997	0.327
English self-assess (0-10)	5.496	5.478	5.468	0.006	0.002	0.792	0.925
Survey in english	0.366	0.397	0.379	-0.018	0.026	0.446	0.273
Female	0.242	0.246	0.248	-0.009	-0.002	0.712	0.918
Primary	0.067	0.079	0.067	-0.000	0.031	0.994	0.196
Secondary	0.430	0.421	0.451	-0.031	-0.044	0.206	0.070
University	0.477	0.472	0.454	0.033	0.026	0.165	0.278
HH-head primary	0.188	0.189	0.178	0.019	0.021	0.425	0.389
HH-head secondary	0.332	0.337	0.341	-0.013	-0.006	0.595	0.791
HH-head university	0.374	0.348	0.358	0.024	-0.015	0.318	0.541
Hindu	0.813	0.794	0.811	0.003	-0.031	0.912	0.202
Muslim	0.132	0.144	0.126	0.014	0.037	0.561	0.124
Christian	0.013	0.009	0.011	0.013	-0.014	0.594	0.575
Sikh	0.025	0.033	0.030	-0.024	0.012	0.312	0.630
General caste	0.529	0.546	0.535	-0.009	0.015	0.723	0.543
OBC caste	0.290	0.278	0.275	0.023	0.005	0.333	0.831
SC caste	0.135	0.123	0.146	-0.021	-0.046	0.376	0.057
Student	0.611	0.646	0.627	-0.023	0.028	0.342	0.249
Employed	0.130	0.125	0.140	-0.021	-0.032	0.387	0.192
Self-employed	0.106	0.085	0.083	0.056	0.005	0.019	0.840
Live in Delhi	0.555	0.581	0.554	0.001	0.038	0.957	0.111
Member of organization	0.175	0.178	0.168	0.014	0.019	0.568	0.422
Currently dating	0.332	0.333	0.331	0.001	0.003	0.973	0.900
Married	0.068	0.063	0.062	0.019	0.003	0.440	0.886
With sisters	0.685	0.698	0.663	0.034	0.054	0.164	0.026
Daily freq. social media	15.576	15.382	15.943	-0.022	-0.034	0.400	0.198
Daily freq. watch videos	2.314	2.283	2.318	-0.003	-0.022	0.914	0.355
Male friend beating	0.142	0.158	0.143	-0.000	0.030	0.985	0.218
Female friend beated	0.136	0.149	0.161	-0.051	-0.025	0.035	0.304

Notes: Table shows sample means at baseline for respondents who completed (at least) the baseline survey. The normalized difference in columns 4 and 5 is the difference in the sample means of treatment and control groups divided by the square root of the sum of the sample variances.

Table 3.A5: Short-term impact on knowledge and awareness

Dep. Var. (Y):	*Father determines sex*		*Know GBV incidence*		*Mention gender rituals*		*VAW is an issue in India*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Drama	0.123***	0.117***	0.030	0.034	0.085**	0.073*	0.070*	0.050
	(0.036)	(0.036)	(0.045)	(0.046)	(0.043)	(0.044)	(0.039)	(0.037)
Documentary	-0.023	-0.026	0.037	0.034	-0.010	-0.012	0.076*	0.083*
	(0.038)	(0.040)	(0.055)	(0.055)	(0.050)	(0.052)	(0.046)	(0.046)
$Y_{baseline}$	0.620***	0.591***						
	(0.036)	(0.042)						
Controls		✓		✓		✓		✓
R-squared	0.346	0.356	−0.002	0.006	0.006	0.016	0.004	0.072
P-value equal coef.	0.000	0.001	0.894	0.994	0.059	0.101	0.882	0.446
Observations	606	606	606	606	606	606	606	606
Mean Y (Control)	0.314	0.314	0.377	0.377	0.291	0.291	0.736	0.736

Notes: Heteroscedasticity-robust standard errors in parentheses. Controls are described in the notes to 3.2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A6: Short-term impact on attitudes towards gender roles

Dep. Var. (Y):	*Stricter control daughters*		*Women takes rights away*		*Important boys to school*		*Men should participate*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Drama	-0.034	-0.028	0.013	0.006	-0.025	-0.021	0.023	0.022
	(0.041)	(0.041)	(0.048)	(0.045)	(0.042)	(0.038)	(0.032)	(0.033)
Documentary	-0.024	-0.023	0.001	0.006	-0.051	-0.072	0.062*	0.057
	(0.049)	(0.048)	(0.058)	(0.056)	(0.050)	(0.046)	(0.035)	(0.035)
$Y_{baseline}$	0.429***	0.256***						
	(0.036)	(0.049)						
Controls		✓		✓		✓		✓
R-squared	0.203	0.245	−0.004	0.144	−0.002	0.192	0.001	0.045
P-value equal coef.	0.826	0.911	0.839	0.993	0.586	0.248	0.235	0.296
Observations	567	567	542	542	585	585	572	572
Mean Y (Control)	0.377	0.377	0.586	0.586	0.297	0.297	0.858	0.858

Notes: Heteroscedasticity-robust standard errors in parentheses. Controls are described in the notes to 3.2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A7: Short-term impact on attitudes towards gender norms

Dep. Var. (Y):	*Wrong to follow norms*		*Women virgin at marriage*		*Ban kitchen during period*		*Wear whatever they want*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Drama	0.081*	0.072*	-0.086**	-0.070*	-0.091**	-0.083**	0.076***	0.060**
	(0.045)	(0.043)	(0.040)	(0.040)	(0.040)	(0.039)	(0.029)	(0.028)
Documentary	-0.024	-0.033	-0.059	-0.055	-0.030	-0.022	0.084***	0.068**
	(0.056)	(0.054)	(0.048)	(0.048)	(0.050)	(0.050)	(0.031)	(0.031)
$Y_{baseline}$			0.536***	0.434***	0.452***	0.337***	0.155	0.182*
			(0.042)	(0.052)	(0.044)	(0.053)	(0.101)	(0.100)
Controls		✓		✓		✓		✓
R-squared	0.005	0.109	0.280	0.288	0.206	0.270	0.024	0.060
P-value equal coef.	0.052	0.047	0.569	0.753	0.202	0.199	0.732	0.777
Observations	606	606	509	509	531	531	546	546
Mean Y (Control)	0.532	0.532	0.744	0.744	0.347	0.347	0.867	0.867

Notes: Heteroscedasticity-robust standard errors in parentheses. Controls are described in the notes to 3.2. * p < 0.10, ** p < 0.05, *** p < 0.01

Table 3.A8: Short-term impact on attitudes towards VAW

Dep. Var. (Y):	*Justify beating if goes out*		*Justify beating if unfaith*		*Tell anyone if friend beaten*	
	(1)	(2)	(3)	(4)	(5)	(6)
Drama	-0.075*	-0.075**	-0.032	-0.033	0.090***	0.092***
	(0.041)	(0.038)	(0.041)	(0.040)	(0.033)	(0.033)
Documentary	-0.050	-0.041	-0.003	-0.004	0.052	0.045
	(0.050)	(0.048)	(0.050)	(0.051)	(0.040)	(0.040)
$Y_{baseline}$			0.407***	0.273***	0.445***	0.391***
			(0.044)	(0.054)	(0.068)	(0.069)
Controls		✓		✓		✓
R-squared	0.003	0.132	0.159	0.216	0.152	0.207
P-value equal coef.	0.586	0.462	0.537	0.553	0.293	0.202
Observations	568	568	542	542	522	522
Mean Y (Control)	0.277	0.277	0.318	0.318	0.815	0.815

Notes: Heteroscedasticity-robust standard errors in parentheses. Controls are described in the notes to 3.2. * p < 0.10, ** p < 0.05, *** p < 0.01

Table 3.A9: Short-term impact on beliefs about Facebook friends

Dep. Var. (Y):	*Others: women virginity*		*Others: men virginity*		*Others: beat unfaith*		*Others: beat goes out*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Drama	0.164	0.224	0.464	0.490	0.353	0.434	-0.035	0.074
	(0.271)	(0.276)	(0.302)	(0.302)	(0.288)	(0.289)	(0.321)	(0.319)
Documentary	0.385	0.510	0.959***	1.079***	1.056***	1.270***	0.586	0.831**
	(0.329)	(0.324)	(0.367)	(0.357)	(0.356)	(0.349)	(0.398)	(0.379)
$Y_{baseline}$	0.413***	0.401***			0.393***	0.373***		
	(0.038)	(0.039)			(0.039)	(0.042)		
Controls		✓		✓		✓		✓
R-squared	0.178	0.211	0.008	0.060	0.167	0.196	0.002	0.073
P-value equal coef.	0.500	0.373	0.184	0.103	0.049	0.019	0.113	0.044
Observations	605	605	605	605	598	598	605	605
Mean Y (Control)	4.731	4.731	3.877	3.877	3.512	3.512	3.461	3.461

Notes: Heteroscedasticity-robust standard errors in parentheses. Controls are described in the notes to 3.2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A10: Medium-term impacts on knowledge and awareness

Dep. Var. (Y):	*Father determines sex*		*Know GBV incidence*		*Mention gender rituals*		*VAW is an issue in India*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Drama	0.074*	0.066	-0.053	-0.068	-0.027	-0.036	-0.000	-0.008
	(0.042)	(0.043)	(0.048)	(0.050)	(0.046)	(0.047)	(0.038)	(0.037)
Documentary	0.038	0.017	-0.046	-0.063	-0.025	-0.030	0.065*	0.078**
	(0.042)	(0.043)	(0.049)	(0.050)	(0.046)	(0.047)	(0.035)	(0.035)
$Y_{baseline}$	0.493***	0.447***						
	(0.038)	(0.044)						
Controls		✓		✓		✓		✓
R-squared	0.223	0.242	−0.001	0.024	−0.003	0.023	0.004	0.062
P-value equal coef.	0.419	0.267	0.887	0.920	0.966	0.896	0.062	0.016
Observations	617	617	617	617	617	617	617	617
Mean Y (Control)	0.350	0.350	0.451	0.451	0.335	0.335	0.820	0.820

Notes: Heteroscedasticity-robust standard errors in parentheses. Controls are described in the notes to 3.2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A11: Medium-term impacts on attitudes towards gender norms and gender roles

Dep. Var. (Y):	*Women virgin at marriage*		*Ban kitchen during period*		*Wear whatever they want*		*Important boys to school*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Drama	-0.101*	-0.041	-0.112**	-0.126***	0.006	-0.011	-0.071	-0.073*
	(0.052)	(0.046)	(0.044)	(0.045)	(0.019)	(0.019)	(0.044)	(0.042)
Documentary	-0.050	-0.025	-0.049	-0.063	0.020	0.016	-0.075*	-0.088**
	(0.052)	(0.047)	(0.047)	(0.046)	(0.019)	(0.019)	(0.044)	(0.043)
$Y_{baseline}$	0.423***	0.139***	0.443***	0.348***	0.164**	0.127*		
	(0.041)	(0.047)	(0.044)	(0.054)	(0.077)	(0.072)		
Controls		✓		✓		✓		✓
R-squared	0.136	0.355	0.193	0.236	0.030	0.062	0.003	0.122
P-value equal coef.	0.341	0.753	0.160	0.169	0.420	0.146	0.927	0.736
Observations	461	461	532	532	578	578	597	597
Mean Y (Control)	0.471	0.471	0.401	0.401	0.953	0.953	0.296	0.296

Notes: Heteroscedasticity-robust standard errors in parentheses. Controls are described in the notes to 3.2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A12: Medium-term impacts on attitudes towards VAW

Dep. Var. (Y):	*Justify beating if goes out* (1)	(2)	*Justify beating if unfaith* (3)	(4)	*Justify beating if neglect* (5)	(6)	*Justify beating if disrespect* (7)	(8)
Drama	0.026	0.025	-0.032	-0.041	-0.029	-0.044	0.014	0.010
	(0.041)	(0.039)	(0.046)	(0.047)	(0.049)	(0.046)	(0.050)	(0.047)
Documentary	0.004	-0.004	0.001	0.002	0.039	0.032	0.013	-0.002
	(0.041)	(0.040)	(0.047)	(0.047)	(0.050)	(0.048)	(0.051)	(0.050)
$Y_{baseline}$			0.292***	0.154***				
			(0.047)	(0.057)				
Controls		✓		✓		✓		✓
R-squared	−0.003	0.092	0.075	0.118	−0.000	0.148	−0.003	0.125
P-value equal coef.	0.598	0.477	0.486	0.361	0.168	0.110	0.984	0.806
Observations	599	599	543	543	589	589	578	578
Mean Y (Control)	0.200	0.200	0.317	0.317	0.400	0.400	0.432	0.432

Notes: Heteroscedasticity-robust standard errors in parentheses. Controls are described in the notes to 3.2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A13: Medium-term impacts on beliefs about Facebook friends

Dep. Var. (Y):	*Others: women virginity* (1)	(2)	*Others: beat unfaith* (3)	(4)	*Others: beat goes out* (5)	(6)
Drama	-0.531*	-0.478	-0.078	0.027	-0.217	-0.095
	(0.299)	(0.302)	(0.309)	(0.313)	(0.321)	(0.322)
Documentary	-0.219	-0.105	0.003	-0.046	-0.473	-0.503
	(0.298)	(0.295)	(0.309)	(0.313)	(0.319)	(0.311)
$Y_{baseline}$	0.349***	0.327***	0.325***	0.293***		
	(0.040)	(0.041)	(0.040)	(0.042)		
Controls		✓		✓		✓
R-squared	0.124	0.151	0.117	0.138	0.000	0.064
P-value equal coef.	0.290	0.212	0.792	0.816	0.421	0.201
Observations	616	616	615	615	619	619
Mean Y (Control)	5.000	5.000	3.546	3.546	3.208	3.208

Notes: Heteroscedasticity-robust standard errors in parentheses. Controls are described in the notes to 3.2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A14: Medium-term impacts on clicks on informative links

Dep. Var. (Y):	*Click gender-links*		*Click climate-links*		*Duration gender-link*		*Duration climate-link*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ITT	ToT	ITT	ToT	ITT	ToT	ITT	ToT
Drama	-0.002		0.010		-1.059		-1.846*	
	(0.017)		(0.014)		(1.206)		(1.028)	
Documentary	-0.018		0.000		-0.967		-0.732	
	(0.015)		(0.013)		(1.174)		(1.373)	
Play Drama		-0.003		0.013		-1.407		-2.411*
		(0.022)		(0.018)		(1.563)		(1.316)
Play Documentary		-0.036		0.000		-2.005		-1.420
		(0.031)		(0.026)		(2.382)		(2.643)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	−0.001	−0.007	−0.001	−0.003	−0.022	−0.031	0.015	0.003
P-value equal coef.	0.298	0.197	0.544	0.637	0.889	0.660	0.199	0.588
Observations	619	619	619	619	589	589	584	584
Wald F-statistic		116.9		116.9		107.1		112.8
Mean Y (Control)	0.029	0.029	0.014	0.014	1.612	1.612	2.269	2.269

Notes: Heteroscedasticity-robust standard errors in parentheses. In columns (1)-(2), the dependent variable takes value 1 if the respondent clicked on both gender links (PFI and UN women India). In columns (3)-(4), the dependent variable takes value 1 if the respondent clicked on both climate links (Delhi Green and UN environment program India). Visit duration refers to the PFI website in columns (5)-(6) and to the Delhi Green website in columns (7)-(8). Controls are described in the notes to 3.2. Independent variables Play Drama and Play Documentary take value 1 if the respondent has played half or more of the assigned video-clips, as objectively recorded by the bot. These variables are instrumented using the random assignment indicators to the treatment groups. The first-stage Wald F-statistic is reported below. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A15: Short-term impacts on outcome indexes constructed using principal component

Panel A: ITT estimates

Dep. Var. (Y):	*Global index*		*Knowledge*		*Gender norms/roles*		*VAW attitudes*		*Beliefs on others*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Drama	0.735***	0.555***	0.386***	0.325**	0.577***	0.423***	0.314**	0.246*	-0.177	-0.236
	(0.223)	(0.152)	(0.135)	(0.128)	(0.187)	(0.139)	(0.154)	(0.127)	(0.172)	(0.172)
Documentary	0.337	0.240	0.164	0.152	0.200	0.113	0.229	0.187	-0.486**	-0.617***
	(0.265)	(0.172)	(0.154)	(0.148)	(0.231)	(0.161)	(0.176)	(0.148)	(0.214)	(0.201)
Controls		✓		✓		✓		✓		✓
R-squared	0.015	0.573	0.011	0.124	0.013	0.502	0.004	0.342	0.005	0.086
P-value equal coef.	0.112	0.060	0.112	0.205	0.081	0.046	0.615	0.682	0.150	0.060
Observations	606	606	606	606	606	606	606	606	606	606
Mean Y (Control)	-0.238	-0.238	-0.068	-0.068	-0.188	-0.188	-0.123	-0.123	0.212	0.212
SD Y (Control)	2.516	2.516	1.574	1.574	2.141	2.141	1.682	1.682	1.810	1.810

Panel B: ToT estimates

Dep. Var. (Y):	*Global index*		*Knowledge*		*Gender norms/roles*		*VAW attitudes*		*Beliefs on others*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Play Drama	0.908***	0.693***	0.476***	0.406***	0.712***	0.528***	0.388**	0.308**	-0.219	-0.302
	(0.272)	(0.185)	(0.166)	(0.156)	(0.229)	(0.169)	(0.189)	(0.155)	(0.213)	(0.211)
Play Documentary	0.496	0.365	0.241	0.231	0.294	0.174	0.336	0.282	-0.714**	-0.925***
	(0.386)	(0.249)	(0.226)	(0.216)	(0.336)	(0.234)	(0.257)	(0.216)	(0.326)	(0.309)
Controls		✓		✓		✓		✓		✓
R-squared	0.036	0.579	0.013	0.122	0.031	0.506	0.017	0.345	−0.025	0.049
P-value equal coef.	0.233	0.146	0.227	0.351	0.162	0.093	0.827	0.894	0.113	0.033
Observations	606	606	606	606	606	606	606	606	606	606
Wald F-statistic	265.5	231.3	265.5	231.3	265.5	231.3	265.5	231.3	265.5	231.3
Mean Y (Control)	-0.238	-0.238	-0.068	-0.068	-0.188	-0.188	-0.123	-0.123	0.212	0.212
SD Y (Control)	2.516	2.516	1.574	1.574	2.141	2.141	1.682	1.682	1.810	1.810

Notes: Heteroscedasticity-robust standard errors in parentheses. Higher values of the indexes in (1) to (8) stand for more progressive stances. The index "Beliefs on others" reported in (9) and (10) takes higher values when the respondents believe that a larger share (out of 10) of their closest Facebook friends have progressive attitudes. Controls are described in the notes to 3.2. In Panel B, independent variables Play Drama and Play Documentary take value 1 if the respondent has played half or more of the assigned video-clips, as objectively recorded by the bot. These variables are instrumented using the random assignment indicators to the treatment groups. The first-stage Wald F-statistic is reported below. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A16: Medium-term impacts on outcome indexes constructed using principal component

Dep. Var. (Y):	*Global index*		*Knowledge*		*Gender norms/roles*		*VAW attitudes*		*Beliefs on others*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	ITT	ToT	ITT	ToT	ITT	ToT	ITT	ToT	ITT	ToT
Drama	0.249		-0.039		0.287**		0.075		0.064	
	(0.165)		(0.121)		(0.131)		(0.154)		(0.155)	
Documentary	0.160		0.095		0.199		-0.014		0.176	
	(0.167)		(0.117)		(0.127)		(0.169)		(0.149)	
Play Drama		0.330		-0.051		0.381**		0.099		0.085
		(0.212)		(0.156)		(0.168)		(0.199)		(0.201)
Play Documentary		0.325		0.194		0.405		-0.030		0.358
		(0.332)		(0.232)		(0.251)		(0.337)		(0.294)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	0.504	0.510	0.138	0.139	0.495	0.504	0.278	0.278	0.073	0.078
P-value equal coef.	0.599	0.986	0.281	0.248	0.496	0.913	0.592	0.667	0.463	0.294
Observations	619	619	619	619	619	619	619	619	619	619
Wald F-stat		116.9		116.9		116.9		116.9		116.9
Mean Y (Control)	-0.114	-0.114	0.014	0.014	-0.094	-0.094	-0.061	-0.061	-0.080	-0.080
SD Y (Control)	2.447	2.447	1.248	1.248	1.901	1.901	1.925	1.925	1.570	1.570

Notes: Heteroscedasticity-robust standard errors in parentheses. Higher values of the indexes in (1) to (8) stand for more progressive stances. The index "Beliefs on others" reported in (9) and (10) takes higher values when the respondents believe that a larger share (out of 10) of their closest Facebook friends have progressive attitudes. Controls are described in the notes to 3.2. Independent variables Play Drama and Play Documentary take value 1 if the respondent has played half or more of the assigned video-clips, as objectively recorded by the bot. These variables are instrumented using the random assignment indicators to the treatment groups. The first-stage Wald F-statistic is reported below. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A17: Short-term heterogeneous impacts by gender

Dep. Var. (Y):	*Global index*	*Knowledge*	*Gender norms/roles*	*VAW attitudes*	*Beliefs on others*	*Click gender-links*	*Willing to share video*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Drama	0.140***	0.123	0.129**	0.198**	-0.088	0.116***	-0.003
	(0.043)	(0.076)	(0.053)	(0.089)	(0.100)	(0.034)	(0.027)
Documentary	0.093**	0.047	0.088	0.169	-0.312***	0.141***	0.061**
	(0.044)	(0.086)	(0.055)	(0.105)	(0.119)	(0.044)	(0.031)
Drama * Female	0.011	0.166	0.002	-0.166	-0.106	-0.161**	-0.098*
	(0.067)	(0.142)	(0.084)	(0.147)	(0.195)	(0.074)	(0.056)
Documentary * Female	-0.096	0.065	-0.131	-0.195	0.020	-0.161*	-0.051
	(0.083)	(0.164)	(0.104)	(0.177)	(0.233)	(0.091)	(0.065)
Female	0.139***	-0.067	0.152**	0.368***	-0.104	0.119**	-0.045
	(0.053)	(0.114)	(0.067)	(0.118)	(0.144)	(0.057)	(0.042)
R-squared	0.585	0.121	0.515	0.338	0.083	0.028	0.034
Observations	606	606	606	606	606	606	2269

Notes: Heteroscedasticity-robust standard errors in parentheses. Higher values of the indexes in (1) to (4) stand for more progressive stances. The index "Beliefs on others" reported in (5) takes higher values when the respondents believe that a larger share (out of 10) of their closest Facebook friends have progressive attitudes. Baseline controls described in the notes to 3.2 are included in all regressions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A18: Short-term heterogeneous impacts by age

Dep. Var. (Y):	*Global index* (1)	*Knowledge* (2)	*Gender norms/roles* (3)	*VAW attitudes* (4)	*Beliefs on others* (5)	*Click genderlinks* (6)	*Willing to share video* (7)
Drama	0.111***	0.152**	0.090*	0.128	-0.036	0.071**	0.002
	(0.040)	(0.077)	(0.047)	(0.085)	(0.100)	(0.035)	(0.028)
Documentary	0.053	0.044	0.039	0.115	-0.274**	0.061	0.047
	(0.045)	(0.085)	(0.055)	(0.100)	(0.115)	(0.044)	(0.032)
Drama * Below 20	0.115	0.056	0.146	0.093	-0.298	0.008	-0.094*
	(0.078)	(0.144)	(0.098)	(0.174)	(0.196)	(0.069)	(0.053)
Documentary * Below 20	0.059	0.078	0.060	0.033	-0.111	0.136	0.013
	(0.088)	(0.166)	(0.109)	(0.193)	(0.231)	(0.094)	(0.060)
Below 20	-0.082	-0.095	-0.057	-0.148	0.015	-0.030	0.012
	(0.070)	(0.130)	(0.087)	(0.160)	(0.168)	(0.056)	(0.048)
R-squared	0.585	0.118	0.515	0.336	0.087	0.022	0.034
Observations	606	606	606	606	606	606	2269

Notes: Heteroscedasticity-robust standard errors in parentheses. Higher values of the indexes in (1) to (4) stand for more progressive stances. The index "Beliefs on others" reported in (5) takes higher values when the respondents believe that a larger share (out of 10) of their closest Facebook friends have progressive attitudes. Baseline controls described in the notes to 3.2 are included in all regressions. * p < 0.10, ** p < 0.05, *** p < 0.01

Table 3.A19: Short-term heterogeneous impacts by caste

Dep. Var. (Y):	*Global index* (1)	*Knowledge* (2)	*Gender norms/roles* (3)	*VAW attitudes* (4)	*Beliefs on others* (5)	*Click genderlinks* (6)	*Willing to share video* (7)
Drama	0.091**	0.091	0.094*	0.081	-0.133	0.106**	-0.009
	(0.046)	(0.088)	(0.054)	(0.099)	(0.110)	(0.043)	(0.032)
Documentary	0.039	0.050	0.002	0.150	-0.455***	0.095*	0.029
	(0.051)	(0.098)	(0.064)	(0.115)	(0.126)	(0.050)	(0.036)
Drama * Low Caste	0.119*	0.174	0.084	0.163	0.036	-0.073	-0.033
	(0.069)	(0.127)	(0.084)	(0.152)	(0.174)	(0.060)	(0.048)
Documentary * Low Caste	0.065	0.019	0.130	-0.088	0.358*	0.016	0.052
	(0.076)	(0.148)	(0.093)	(0.172)	(0.206)	(0.081)	(0.056)
Low Caste	-0.072	-0.075	-0.080	-0.038	-0.087	0.018	-0.000
	(0.052)	(0.100)	(0.065)	(0.113)	(0.126)	(0.041)	(0.036)
R-squared	0.586	0.124	0.515	0.339	0.087	0.023	0.034
Observations	606	606	606	606	606	606	2269

Notes: Heteroscedasticity-robust standard errors in parentheses. Higher values of the indexes in (1) to (4) stand for more progressive stances. The index "Beliefs on others" reported in (5) takes higher values when the respondents believe that a larger share (out of 10) of their closest Facebook friends have progressive attitudes. Baseline controls described in the notes to 3.2 are included in all regressions. * p < 0.10, ** p < 0.05, *** p < 0.01

Table 3.A20: Short-term heterogeneous impacts by education

Dep. Var. (Y):	*Global index*	*Knowledge*	*Gender norm-s/roles*	*VAW attitudes*	*Beliefs on others*	*Click gender-links*	*Willing to share video*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Drama	0.129***	0.106	0.142**	0.118	-0.133	0.041	-0.017
	(0.049)	(0.097)	(0.061)	(0.109)	(0.128)	(0.044)	(0.033)
Documentary	0.079	0.028	0.073	0.167	-0.441***	0.156***	0.067*
	(0.054)	(0.102)	(0.067)	(0.127)	(0.147)	(0.059)	(0.037)
Drama * University	0.024	0.125	-0.026	0.060	0.044	0.066	-0.018
	(0.068)	(0.131)	(0.081)	(0.154)	(0.172)	(0.063)	(0.048)
Documentary * University	-0.031	0.071	-0.044	-0.126	0.305	-0.120	-0.038
	(0.078)	(0.146)	(0.096)	(0.172)	(0.207)	(0.078)	(0.055)
University	-0.019	-0.083	0.029	-0.093	-0.130	-0.047	0.002
	(0.050)	(0.102)	(0.062)	(0.114)	(0.123)	(0.038)	(0.037)
R-squared	0.584	0.122	0.515	0.336	0.087	0.029	0.033
Observations	606	606	606	606	606	606	2269

Notes: Heteroscedasticity-robust standard errors in parentheses. Higher values of the indexes in (1) to (4) stand for more progressive stances. The index "Beliefs on others" reported in (5) takes higher values when the respondents believe that a larger share (out of 10) of their closest Facebook friends have progressive attitudes. Baseline controls described in the notes to 3.2 are included in all regressions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A21: Short-term heterogeneous impacts by education of household-head (HH)

Dep. Var. (Y):	*Global index*	*Knowledge*	*Gender norm-s/roles*	*VAW attitudes*	*Beliefs on others*	*Click gender-links*	*Willing to share video*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Drama	0.184***	0.246***	0.154***	0.203**	-0.132	0.068*	-0.030
	(0.042)	(0.078)	(0.051)	(0.095)	(0.108)	(0.036)	(0.030)
Documentary	0.081*	0.082	0.037	0.229**	-0.488***	0.096**	0.074**
	(0.048)	(0.090)	(0.059)	(0.112)	(0.129)	(0.047)	(0.034)
Drama * HH university	-0.121*	-0.247*	-0.063	-0.149	0.056	0.010	0.014
	(0.071)	(0.137)	(0.089)	(0.145)	(0.172)	(0.068)	(0.050)
Documentary * HH university	-0.052	-0.095	0.045	-0.318*	0.498**	0.006	-0.067
	(0.081)	(0.154)	(0.102)	(0.171)	(0.199)	(0.087)	(0.057)
HH-university	0.084	0.195*	0.019	0.152	0.152	0.052	0.046
	(0.056)	(0.112)	(0.072)	(0.110)	(0.121)	(0.047)	(0.039)
R-squared	0.581	0.123	0.509	0.335	0.091	0.020	0.034
Observations	606	606	606	606	606	606	2269

Notes: Heteroscedasticity-robust standard errors in parentheses. Higher values of the indexes in (1) to (4) stand for more progressive stances. The index "Beliefs on others" reported in (5) takes higher values when the respondents believe that a larger share (out of 10) of their closest Facebook friends have progressive attitudes. Baseline controls described in the notes to 3.2 are included in all regressions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A22: Short-term heterogeneous impacts by perceptions of social norms (SN)

Dep. Var. (Y):	*Global index*	*Knowledge*	*Gender norms/roles*	*VAW attitudes*	*Beliefs on others*	*Click gender-links*	*Willing to share video*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Drama	0.192***	0.199**	0.150**	0.323***	-0.116	0.050	-0.007
	(0.049)	(0.092)	(0.060)	(0.101)	(0.107)	(0.045)	(0.034)
Documentary	0.130**	0.021	0.146**	0.225*	-0.353***	0.103*	0.019
	(0.053)	(0.098)	(0.065)	(0.115)	(0.120)	(0.056)	(0.039)
Drama * Strict SN perceptions	-0.094	-0.057	-0.038	-0.328**	0.025	0.044	-0.034
	(0.068)	(0.130)	(0.083)	(0.145)	(0.154)	(0.059)	(0.048)
Documentary * Strict SN perceptions	-0.134*	0.093	-0.201**	-0.212	0.017	-0.009	0.061
	(0.076)	(0.145)	(0.095)	(0.169)	(0.187)	(0.082)	(0.054)
Strict SN perceptions	0.038	-0.036	0.025	0.180	-0.771***	-0.015	0.035
	(0.050)	(0.102)	(0.061)	(0.111)	(0.112)	(0.041)	(0.035)
R-squared	0.586	0.120	0.517	0.341	0.233	0.018	0.035
Observations	606	606	606	606	606	606	2269

Notes: Heteroscedasticity-robust standard errors in parentheses. Higher values of the indexes in (1) to (4) stand for more progressive stances. The index "Beliefs on others" reported in (5) takes higher values when the respondents believe that a larger share (out of 10) of their closest Facebook friends have progressive attitudes. Baseline controls described in the notes to 3.2 are included in all regressions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A23: Short-term heterogeneous impacts by membership of social organization

Dep. Var. (Y):	*Global index*	*Knowledge*	*Gender norms/roles*	*VAW attitudes*	*Beliefs on others*	*Click gender-links*	*Willing to share video*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Drama	0.111***	0.141**	0.088*	0.149*	-0.110	0.073**	-0.030
	(0.038)	(0.069)	(0.046)	(0.080)	(0.092)	(0.033)	(0.026)
Documentary	0.076*	0.065	0.048	0.182**	-0.286***	0.095**	0.039
	(0.041)	(0.080)	(0.051)	(0.091)	(0.107)	(0.041)	(0.030)
Drama * ORG member	0.230***	0.192	0.305***	0.035	-0.053	0.005	0.022
	(0.088)	(0.185)	(0.112)	(0.214)	(0.243)	(0.089)	(0.066)
Documentary * ORG member	-0.096	-0.040	0.016	-0.548**	-0.168	0.048	0.065
	(0.100)	(0.198)	(0.126)	(0.257)	(0.303)	(0.119)	(0.073)
ORG member	-0.082	-0.085	-0.131	0.087	0.071	-0.005	0.047
	(0.066)	(0.149)	(0.087)	(0.143)	(0.174)	(0.052)	(0.050)
R-squared	0.590	0.120	0.520	0.341	0.082	0.017	0.035
Observations	606	606	606	606	606	606	2269

Notes: Heteroscedasticity-robust standard errors in parentheses. Higher values of the indexes in (1) to (4) stand for more progressive stances. The index "Beliefs on others" reported in (5) takes higher values when the respondents believe that a larger share (out of 10) of their closest Facebook friends have progressive attitudes. Baseline controls described in the notes to 3.2 are included in all regressions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A24: Medium-term heterogeneous impacts by gender

Dep. Var. (Y):	*Global index*	*Knowledge*	*Gender norms/roles*	*VAW attitudes*	*Beliefs on others*	*Intent to update picture*	*Actual picture update*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Drama	0.099**	0.028	0.145***	0.064	0.097	0.100*	0.006
	(0.042)	(0.072)	(0.050)	(0.082)	(0.104)	(0.053)	(0.036)
Documentary	0.051	0.073	0.067	0.001	0.192*	0.113**	0.048
	(0.041)	(0.071)	(0.049)	(0.084)	(0.101)	(0.055)	(0.040)
Drama * Female	-0.188**	-0.209	-0.210**	-0.129	-0.193	-0.149	0.089
	(0.078)	(0.128)	(0.088)	(0.154)	(0.193)	(0.093)	(0.056)
Documentary * Female	-0.079	-0.164	-0.112	0.056	-0.343*	-0.140	0.090
	(0.076)	(0.129)	(0.086)	(0.162)	(0.179)	(0.098)	(0.067)
Female	0.181***	0.091	0.213***	0.189*	0.115	0.027	-0.095***
	(0.055)	(0.097)	(0.065)	(0.106)	(0.131)	(0.067)	(0.032)
R-squared	0.540	0.130	0.483	0.290	0.076	0.023	0.003
Observations	619	619	619	619	619	619	619

Notes: Heteroscedasticity-robust standard errors in parentheses. Higher values of the indexes in (1) to (4) stand for more progressive stances. The index "Beliefs on others" reported in (5) takes higher values when the respondents believe that a larger share (out of 10) of their closest Facebook friends have progressive attitudes. Baseline controls described in the notes to 3.2 are included in all regressions. * p < 0.10, ** p < 0.05, *** p < 0.01

Table 3.A25: Medium-term heterogeneous impacts by age

Dep. Var. (Y):	*Global index*	*Knowledge*	*Gender norms/roles*	*VAW attitudes*	*Beliefs on others*	*Intent to update picture*	*Actual picture update*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Drama	0.040	-0.055	0.064	0.069	0.034	0.045	0.053
	(0.043)	(0.073)	(0.049)	(0.079)	(0.106)	(0.053)	(0.035)
Documentary	0.011	0.049	0.023	-0.045	0.052	0.082	0.081**
	(0.042)	(0.069)	(0.047)	(0.086)	(0.098)	(0.055)	(0.038)
Drama * Below 20	0.024	0.077	0.077	-0.125	0.044	0.048	-0.072
	(0.080)	(0.137)	(0.096)	(0.162)	(0.194)	(0.096)	(0.063)
Documentary * Below 20	0.064	-0.085	0.047	0.217	0.174	-0.027	-0.029
	(0.073)	(0.135)	(0.090)	(0.156)	(0.193)	(0.101)	(0.076)
Below 20	-0.058	0.003	-0.083	-0.056	-0.238	-0.064	-0.001
	(0.063)	(0.114)	(0.078)	(0.128)	(0.168)	(0.083)	(0.057)
R-squared	0.535	0.127	0.479	0.293	0.074	0.019	0.001
Observations	619	619	619	619	619	619	619

Notes: Heteroscedasticity-robust standard errors in parentheses. Higher values of the indexes in (1) to (4) stand for more progressive stances. The index "Beliefs on others" reported in (5) takes higher values when the respondents believe that a larger share (out of 10) of their closest Facebook friends have progressive attitudes. Baseline controls described in the notes to 3.2 are included in all regressions. * p < 0.10, ** p < 0.05, *** p < 0.01

Table 3.A26: Medium-term heterogeneous impacts by caste

Dep. Var. (Y):	*Global index*	*Knowledge*	*Gender norms/roles*	*VAW attitudes*	*Beliefs on others*	*Intent to update picture*	*Actual picture update*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Drama	0.077*	0.003	0.124**	0.041	0.045	-0.012	0.006
	(0.045)	(0.079)	(0.051)	(0.084)	(0.116)	(0.056)	(0.034)
Documentary	0.043	0.028	0.069	0.003	0.076	0.063	0.064
	(0.043)	(0.080)	(0.049)	(0.090)	(0.107)	(0.058)	(0.040)
Drama * Low Caste	-0.071	-0.081	-0.090	-0.023	0.005	0.175*	0.065
	(0.073)	(0.120)	(0.085)	(0.145)	(0.174)	(0.089)	(0.062)
Documentary * Low Caste	-0.033	-0.001	-0.083	0.040	0.066	0.023	0.025
	(0.070)	(0.119)	(0.084)	(0.153)	(0.171)	(0.093)	(0.070)
Low Caste	-0.054	-0.058	-0.028	-0.101	-0.055	-0.005	0.021
	(0.051)	(0.085)	(0.063)	(0.103)	(0.124)	(0.062)	(0.041)
R-squared	0.534	0.127	0.479	0.287	0.067	0.023	0.000
Observations	619	619	619	619	619	619	619

Notes: Heteroscedasticity-robust standard errors in parentheses. Higher values of the indexes in (1) to (4) stand for more progressive stances. The index "Beliefs on others" reported in (5) takes higher values when the respondents believe that a larger share (out of 10) of their closest Facebook friends have progressive attitudes. Baseline controls described in the notes to 3.2 are included in all regressions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A27: Medium-term heterogeneous impacts by education

Dep. Var. (Y):	*Global index*	*Knowledge*	*Gender norms/roles*	*VAW attitudes*	*Beliefs on others*	*Intent to update picture*	*Actual picture update*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Drama	0.072	0.024	0.111*	0.033	0.063	0.007	0.028
	(0.048)	(0.084)	(0.061)	(0.097)	(0.128)	(0.062)	(0.041)
Documentary	0.084*	0.097	0.058	0.125	0.135	0.053	0.087*
	(0.045)	(0.082)	(0.055)	(0.098)	(0.118)	(0.067)	(0.048)
Drama * University	-0.068	-0.111	-0.064	-0.040	-0.050	0.095	-0.001
	(0.072)	(0.119)	(0.084)	(0.139)	(0.175)	(0.087)	(0.057)
Documentary * University	-0.128*	-0.141	-0.064	-0.245*	-0.084	0.032	-0.035
	(0.068)	(0.118)	(0.081)	(0.141)	(0.168)	(0.091)	(0.064)
University	0.058	0.127	0.048	0.023	0.052	-0.043	-0.007
	(0.053)	(0.084)	(0.061)	(0.107)	(0.122)	(0.064)	(0.041)
R-squared	0.533	0.128	0.473	0.287	0.073	0.013	−0.004
Observations	619	619	619	619	619	619	619

Notes: Heteroscedasticity-robust standard errors in parentheses. Higher values of the indexes in (1) to (4) stand for more progressive stances. The index "Beliefs on others" reported in (5) takes higher values when the respondents believe that a larger share (out of 10) of their closest Facebook friends have progressive attitudes. Baseline controls described in the notes to 3.2 are included in all regressions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A28: Medium-term heterogeneous impacts by education of household-head (HH)

Dep. Var. (Y):	*Global index*	*Knowledge*	*Gender norms/roles*	*VAW attitudes*	*Beliefs on others*	*Intent to update picture*	*Actual picture update*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Drama	0.048	-0.010	0.087	0.016	0.035	0.090	0.001
	(0.045)	(0.076)	(0.055)	(0.088)	(0.114)	(0.055)	(0.035)
Documentary	0.030	0.046	0.037	0.003	0.111	0.138**	0.088*
	(0.044)	(0.079)	(0.053)	(0.095)	(0.114)	(0.061)	(0.045)
Drama * HH university	0.003	-0.035	-0.002	0.045	0.030	-0.086	0.083
	(0.073)	(0.126)	(0.084)	(0.144)	(0.178)	(0.091)	(0.061)
Documentary * HH university	0.003	-0.033	-0.002	0.039	-0.013	-0.162*	-0.039
	(0.069)	(0.120)	(0.081)	(0.144)	(0.166)	(0.091)	(0.065)
HH-university	0.062	0.025	0.076	0.065	0.238*	0.039	0.005
	(0.054)	(0.084)	(0.063)	(0.105)	(0.123)	(0.066)	(0.042)
R-squared	0.534	0.118	0.479	0.287	0.067	0.024	0.005
Observations	619	619	619	619	619	619	619

Notes: Heteroscedasticity-robust standard errors in parentheses. Higher values of the indexes in (1) to (4) stand for more progressive stances. The index "Beliefs on others" reported in (5) takes higher values when the respondents believe that a larger share (out of 10) of their closest Facebook friends have progressive attitudes. Baseline controls described in the notes to 3.2 are included in all regressions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A29: Medium-term heterogeneous impacts by perceptions of social norms (SN)

Dep. Var. (Y):	*Global index*	*Knowledge*	*Gender norms/roles*	*VAW attitudes*	*Beliefs on others*	*Intent to update picture*	*Actual picture update*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Drama	0.072	-0.090	0.132**	0.080	0.133	-0.005	0.012
	(0.048)	(0.084)	(0.056)	(0.094)	(0.116)	(0.061)	(0.040)
Documentary	0.077*	-0.052	0.106**	0.121	0.145	-0.007	0.077*
	(0.044)	(0.082)	(0.050)	(0.097)	(0.105)	(0.061)	(0.046)
Drama * Strict SN perceptions	-0.054	0.122	-0.096	-0.109	-0.182	0.131	0.040
	(0.073)	(0.119)	(0.084)	(0.145)	(0.164)	(0.090)	(0.059)
Documentary * Strict SN perceptions	-0.108	0.168	-0.153*	-0.238*	-0.155	0.178*	-0.012
	(0.068)	(0.117)	(0.084)	(0.143)	(0.157)	(0.091)	(0.066)
Strict SN perceptions	0.049	-0.108	0.105*	0.062	-0.439***	-0.063	-0.016
	(0.051)	(0.082)	(0.061)	(0.102)	(0.118)	(0.063)	(0.041)
R-squared	0.536	0.128	0.481	0.291	0.166	0.025	−0.001
Observations	619	619	619	619	619	619	619

Notes: Heteroscedasticity-robust standard errors in parentheses. Higher values of the indexes in (1) to (4) stand for more progressive stances. The index "Beliefs on others" reported in (5) takes higher values when the respondents believe that a larger share (out of 10) of their closest Facebook friends have progressive attitudes. Baseline controls described in the notes to 3.2 are included in all regressions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A30: Medium-term heterogeneous impacts by membership of social organization

Dep. Var. (Y):	*Global index*	*Knowledge*	*Gender norms/roles*	*VAW attitudes*	*Beliefs on others*	*Intent to update picture*	*Actual picture update*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Drama	0.039	-0.031	0.086*	0.002	-0.000	0.033	0.024
	(0.038)	(0.065)	(0.046)	(0.075)	(0.096)	(0.048)	(0.032)
Documentary	0.036	0.018	0.054	0.014	0.106	0.058	0.076**
	(0.036)	(0.062)	(0.043)	(0.079)	(0.091)	(0.050)	(0.037)
Drama * ORG member	0.049	-0.015	0.006	0.189	0.288	0.173	0.054
	(0.107)	(0.182)	(0.107)	(0.203)	(0.228)	(0.123)	(0.074)
Documentary * ORG member	-0.052	0.045	-0.129	0.025	-0.038	0.111	-0.016
	(0.104)	(0.186)	(0.110)	(0.197)	(0.235)	(0.127)	(0.076)
ORG member	-0.015	0.140	-0.019	-0.130	-0.016	-0.058	-0.063
	(0.077)	(0.142)	(0.075)	(0.150)	(0.166)	(0.087)	(0.046)
R-squared	0.535	0.132	0.481	0.289	0.074	0.021	0.002
Observations	619	619	619	619	619	619	619

Notes: Heteroscedasticity-robust standard errors in parentheses. Higher values of the indexes in (1) to (4) stand for more progressive stances. The index "Beliefs on others" reported in (5) takes higher values when the respondents believe that a larger share (out of 10) of their closest Facebook friends have progressive attitudes. Baseline controls described in the notes to 3.2 are included in all regressions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A31: Index at baseline

	Loading factor
Baseline Stance Index *(Cronbach's alpha = 0.56)*	
Knows that father determines sex of child	0.176
Agreement (1-5) on "women should be virgin until marriage"	0.456
Agreement (1-5) on "a woman should be banned from entering the kitchen or household shrine during her period"	0.442
Agreement (1-5) "parents should maintain stricter control over their daughters than their sons"	0.497
Agreement (1-5) "girls should be allowed to wear whatever they want without being harassed"	0.251
Agreement (1-5) on "a husband is justified in hitting or beating his wife if he suspects her of being unfaithful"	0.458
Would tell anyone if finds out that a friend beats or physically hurts his partner	0.214

Notes: All variables were re-oriented so that the impact of treatments on each component of the index should be positive. The column on the right reports the loading factors used for the construction of the indexes with principal component.

Table 3.A32: Indexes at short-term follow-up

	Loading factor
Knowledge and awareness *(Cronbach's alpha = 0.31)*	
Knows that father determines sex of child	0.455
Knows that 1/3 of women in the world experience violence	0.524
Thinks that GBV is an issue in India	0.579
Mentions either virginity ritual or menstruation rituals as prevalent gender norms	0.428
Attitudes toward gender norms/roles *(Cronbach's alpha = 0.61)*	
Thinks that is it wrong to follow the above-mentioned gender norms	0.352
Agreement (1-5) on "women should be virgin until marriage"	0.384
Agreement (1-5) on "a woman should be banned from entering the kitchen or household shrine during her period"	0.385
Agreement (1-5) "women should be able to marry whomever they want, regardless of their parents' views"	0.072
Agreement (1-5) "girls should be allowed to wear whatever they want without being harassed"	0.190
Agreement (1-5) "parents should maintain stricter control over their daughters than their sons"	0.454
Agreement (1-5) "when women get rights they are taking rights away from men"	0.345
Agreement (1-5) "it is more important that a boy goes to school than a girl"	0.389
Agreement (1-5) "nowadays men should participate in child rearing and household chores rather than leaving it all to the women"	0.031
Thinks that husband and wife should have the equal say in deciding how many children to have	0.249
Attitudes toward VAW *(Cronbach's alpha = 0.54)*	
Agreement (1-5) on "a husband is justified in hitting or beating his wife if she goes out without telling him"	0.660
Agreement (1-5) on "a husband is justified in hitting or beating his wife if he suspects her of being unfaithful"	0.656
Would tell anyone if finds out that a friend beats or physically hurts his partner	0.367
Beliefs on others *(Cronbach's alpha = 0.79)*	
Imagine to pick 10 of your closest Facebook friends. According to you, how many of them think that women should be virgins till marriage?	0.484
Imagine to pick 10 of your closest Facebook friends. According to you, how many of them think that men should be virgins till marriage?	0.504
Imagine to pick 10 of your closest Facebook friends. According to you, how many of them think that a husband is justified in hitting or beating his wife if he suspects her of being unfaithful?	0.511
Imagine to pick 10 of your closest Facebook friends. According to you, how many of them think that a husband is justified in hitting or beating his wife if she goes out without telling him?	0.501

Notes: The *Global Index* (Cronbach's alpha = 0.71) is created using all variables reported above with the exception of those contained in the *Beliefs on others* index. All variables were re-oriented so that the impact of treatments on each component of the index should be positive. The column on the right reports the loading factors used for the construction of the indexes with principal component.

Table 3.A33: Indexes at medium-term follow-up

	Loading factor
Knowledge and awareness *(Cronbach's alpha = 0.24)*	
Knows that father determines sex of child	0.540
Knows that 1/3 of women in the world experience violence	0.433
Thinks that GBV is an issue in India	0.596
Mentions either virginity ritual or menstruation rituals as prevalent gender norms	0.406
Attitudes toward gender norms/roles *(Cronbach's alpha = 0.60)*	
Thinks that is it wrong to follow the above-mentioned gender norms	0.367
Agreement (1-5) on "women should be virgin until marriage"	0.359
Agreement (1-5) on "a woman should be banned from entering the kitchen or household shrine during her period"	0.338
Agreement (1-5) "women should be able to marry whomever they want, regardless of their parents' views"	0.087
Agreement (1-5) "girls should be allowed to wear whatever they want without being harassed"	0.232
Agreement (1-5) "parents should maintain stricter control over their daughters than their sons"	0.430
Agreement (1-5) "when women get rights they are taking rights away from men"	0.370
Agreement (1-5) "it is more important that a boy goes to school than a girl"	0.444
Agreement (1-5) "nowadays men should participate in child rearing and household chores rather than leaving it all to the women"	0.091
Thinks that husband and wife should have the equal say in deciding how many children to have	0.181
Attitudes toward VAW *(Cronbach's alpha = 0.71)*	
Agreement (1-5) on "a husband is justified in hitting or beating his wife if she goes out without telling him"	0.479
Agreement (1-5) on "a husband is justified in hitting or beating his wife if he suspects her of being unfaithful"	0.467
Agreement (1-5) on "a husband is justified in hitting or beating his wife if she neglects the house or the children"	0.522
Agreement (1-5) on "a husband is justified in hitting or beating his wife if she shows disrespect for in-laws"	0.504
Would tell anyone if finds out that a friend beats or physically hurts his partner	0.159
Beliefs on others *(Cronbach's alpha = 0.69)*	
Imagine to pick 10 of your closest Facebook friends. According to you, how many of them think that women should be virgins till marriage?	0.496
Imagine to pick 10 of your closest Facebook friends. According to you, how many of them think that a husband is justified in hitting or beating his wife if he suspects her of being unfaithful?	0.630
Imagine to pick 10 of your closest Facebook friends. According to you, how many of them think that a husband is justified in hitting or beating his wife if she goes out without telling him?	0.598

Notes: The *Global Index* (Cronbach's alpha = 0.74) is created using all variables reported above with the exception of those contained in the *Beliefs on others* index. All variables were re-oriented so that the impact of treatments on each component of the index should be positive. The column on the right reports the loading factors used for the construction of the indexes with principal component.

3.B Figures

Figure 3.B1: Geographic targeting and ad banner used to recruit study participants

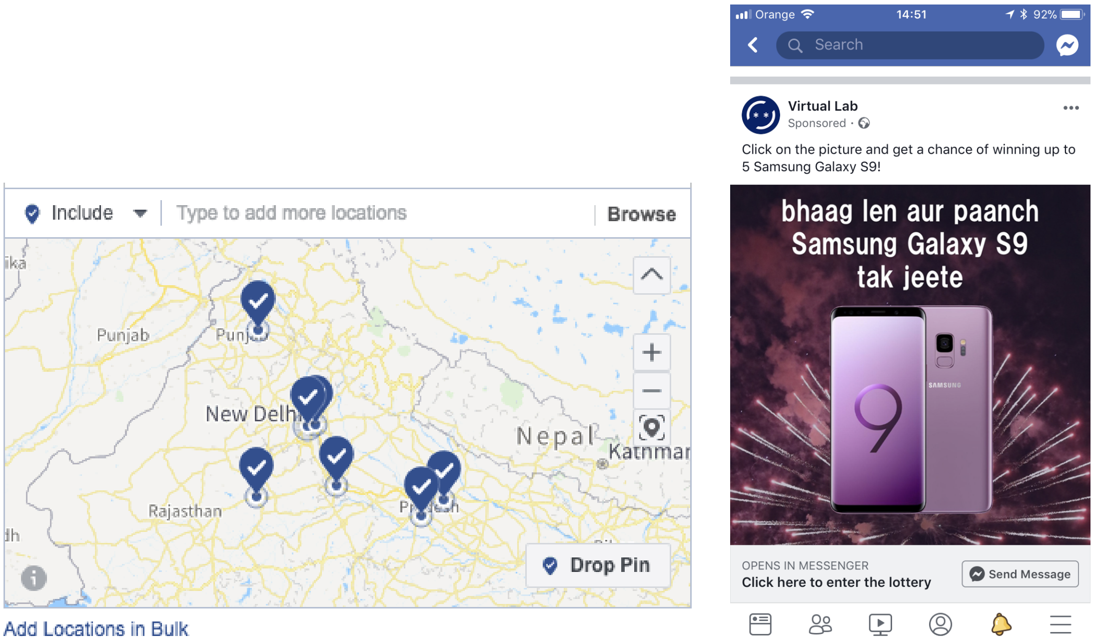


Figure 3.B2: Example of mobile video screening and surveying within FB Messenger

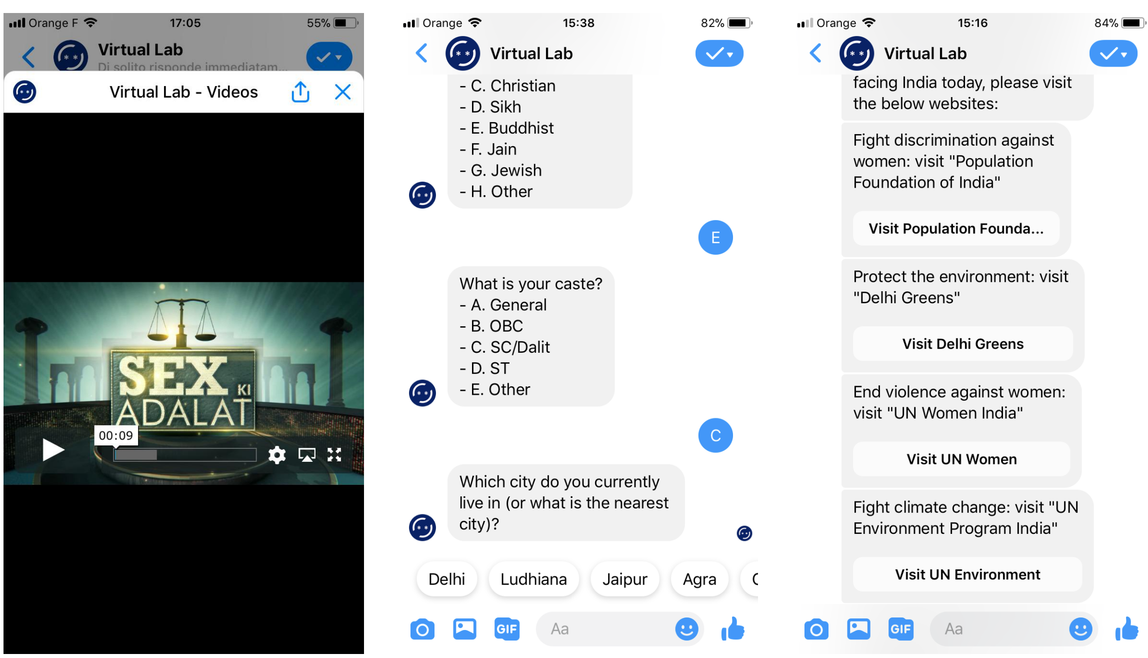
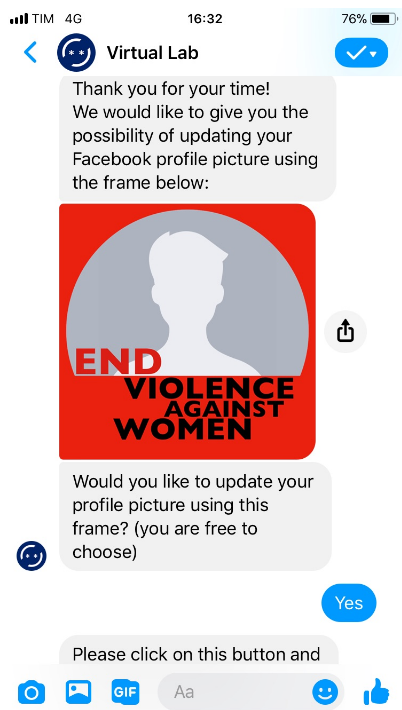


Figure 3.B3: VAW frame to be added on the FB profile picture



Bibliography

Abramsky, T., Devries, K., Kiss, L., Nakuti, J., Kyegombe, N., Starmann, E., Cundill, B., Francisco, L., Kaye, D., Musuya, T. et al. (2014), 'Findings from the sasa! study: a cluster randomized controlled trial to assess the impact of a community mobilization intervention to prevent violence against women and reduce hiv risk in kampala, uganda', *BMC medicine* **12**(1), 1–17.

Abramsky, T., Devries, K. M., Michau, L., Nakuti, J., Musuya, T., Kyegombe, N. and Watts, C. (2016), 'The impact of sasa!, a community mobilisation intervention, on women's experiences of intimate partner violence: secondary findings from a cluster randomised trial in kampala, uganda', *J Epidemiol Community Health* **70**(8), 818–825.

Abramsky, T., Watts, C. H., Garcia-Moreno, C., Devries, K., Kiss, L., Ellsberg, M., Jansen, H. A. and Heise, L. (2011), 'What factors are associated with recent intimate partner violence? findings from the who multi-country study on women's health and domestic violence', *BMC public health* **11**(1), 1–17.

Aker, J. C., Blumenstock, J. E. and Dillon, B. (2020), 'How important is the yellow pages? experimental evidence from tanzania'.

Aker, J. C. and Mbiti, I. M. (2010), 'Mobile phones and economic development in africa', *Journal of economic Perspectives* **24**(3), 207–32.

Akerlof, G. A. (1970), 'The market for "lemons": Quality uncertainty and the market mechanism', *The Quarterly Journal of Economics* **84**(3), 488–500.
URL: *http://www.jstor.org/stable/1879431*

Akerlof, G. A. and Kranton, R. E. (2000), 'Economics and identity', *The quarterly journal of economics* **115**(3), 715–753.

Alan, S., Ertac, S. and Mumcu, I. (2018), 'Gender stereotypes in the classroom and effects on achievement', *Review of Economics and Statistics* **100**(5), 876–890.

Alatas, V., Chandrasekhar, A. G., Mobius, M., Olken, B. A. and Paladines, C. (2019), When celebrities speak: A nationwide twitter experiment promoting vaccination in indonesia, Technical report, National Bureau of Economic Research.

Alesina, A., Giuliano, P. and Nunn, N. (2013), 'On the origins of gender roles: Women and the plough', *The quarterly journal of economics* **128**(2), 469–530.

Allcott, H., Braghieri, L., Eichmeyer, S. and Gentzkow, M. (2020), 'The welfare effects of social media', *American Economic Review* **110**(3), 629–76.

Allcott, H. and Gentzkow, M. (2017), 'Social media and fake news in the 2016 election', *Journal of economic perspectives* **31**(2), 211–36.

Allen, T. (2014), 'Information frictions in trade', *Econometrica* **82**(6), 2041–2083.

Altonji, J. G., Elder, T. E. and Taber, C. R. (2005), 'Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools', *Journal of political economy* **113**(1), 151–184.

Ananthakrishnan, U. M., Proserpio, D. and Sharma, S. (2019), 'I hear you: Does quality improve with customer voice?', *Available at SSRN 3467236* .

Anderson, M. and Magruder, J. (2012), 'Learning from the crowd: Regression discontinuity estimates of the effects of an online review database', *The Economic Journal* **122**(563), 957–989.

Anderson, S. J., Chandy, R. and Zia, B. (2018), 'Pathways to profits: The impact of marketing vs. finance skills on business performance', *Management Science* **64**(12), 5559–5583.

Anderson, S. P. and Renault, R. (1999), 'Pricing, product diversity, and search costs: A bertrand-chamberlin-diamond model', *The RAND Journal of Economics* pp. 719–735.

Arias, E. (2019), 'How does media influence social norms? experimental evidence on the role of common knowledge', *Political Science Research and Methods* **7**(3), 561–578.

Atkin, D., Khandelwal, A. K. and Osman, A. (2017), 'Exporting and firm performance: Evidence from a randomized experiment', *The Quarterly Journal of Economics* **132**(2), 551–615.

Bai, J. (2018), Melons as lemons: Asymmetric information, consumer learning and quality provision, Technical report, Working paper.

Bai, J. (2021), 'Melons as lemons: Asymmetric information, consumer learning and seller reputation', *CID Working Paper Series* .

Bandura, A. (2004), 'Health promotion by social cognitive means', *Health education & behavior* **31**(2), 143–164.

Banerjee, A. (2006), *Can Information Campaigns Spark Local Participation and Improve Outcomes?: A Study of Primary Education in Uttar Pradesh, India*, Vol. 3967, World Bank Publications.

Banerjee, A., Alsan, M., Breza, E., Chandrasekhar, A. G., Chowdhury, A., Duflo, E., Goldsmith-Pinkham, P. and Olken, B. A. (2020), Messages on covid-19 prevention in india increased symptoms reporting and adherence to preventive behaviors among 25 million recipients with similar effects on non-recipient members of their communities, Technical report, National Bureau of Economic Research.

Banerjee, A., La Ferrara, E. and Orozco-Olvera, V. (2019*a*), Entertainment, education, and attitudes toward domestic violence, *in* ‘AEA Papers and Proceedings’, Vol. 109, pp. 133–37.

Banerjee, A., La Ferrara, E. and Orozco-Olvera, V. H. (2019*b*), The entertaining way to behavioral change: Fighting hiv with mtv, Technical report, National Bureau of Economic Research.

Bar-Isaac, H., Caruana, G. and Cuñat, V. (2012), ‘Search, design, and market structure’, *American Economic Review* **102**(2), 1140–60.

Baron, D. P. (1994), ‘Electoral competition with informed and uninformed voters’, *American Political Science Review* **88**(1), 33–47.

Bastos, P., Silva, J. and Verhoogen, E. (2018), ‘Export destinations and input prices’, *American Economic Review* **108**(2), 353–92.

Basu, S. and Fernald, J. G. (1997), ‘Returns to scale in us production: Estimates and implications’, *Journal of political economy* **105**(2), 249–283.

Benoit, W. L. (1987), ‘Argumentation and credibility appeals in persuasion’, *Southern Journal of Communication* **52**(2), 181–197.

Berg, G. and Zia, B. (2017), ‘Harnessing emotional connections to improve financial decisions: Evaluating the impact of financial education in mainstream media’, *Journal of the European Economic Association* **15**(5), 1025–1055.

Bertrand, M. (2020), Gender in the twenty-first century, *in* ‘AEA Papers and proceedings’, Vol. 110, pp. 1–24.

Bertrand, M., Kamenica, E. and Pan, J. (2015), ‘Gender identity and relative income within households’, *The Quarterly Journal of Economics* **130**(2), 571–614.

Besley, T. (2005), ‘Political selection’, *Journal of Economic perspectives* **19**(3), 43–60.

Besley, T. and Burgess, R. (2002), ‘The political economy of government responsiveness: Theory and evidence from india’, *The quarterly journal of economics* **117**(4), 1415–1451.

Besley, T., Pande, R. and Rao, V. (2005), ‘Participatory democracy in action: Survey evidence from south india’, *Journal of the European Economic Association* **3**(2-3), 648–657.

Besley, T. and Prat, A. (2006), 'Handcuffs for the grabbing hand? media capture and government accountability', *American economic review* **96**(3), 720–736.

Bicchieri, C. (2005), *The grammar of society: The nature and dynamics of social norms*, Cambridge University Press.

Bicchieri, C. (2016), *Norms in the wild: How to diagnose, measure, and change social norms*, Oxford University Press.

Blair, G., Littman, R. and Paluck, E. L. (2019), 'Motivating the adoption of new community-minded behaviors: An empirical test in nigeria', *Science advances* **5**(3), eaau5175.

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D., Marlow, C., Settle, J. E. and Fowler, J. H. (2012), 'A 61-million-person experiment in social influence and political mobilization', *Nature* **489**(7415), 295–298.

Bordalo, P., Coffman, K., Gennaioli, N. and Shleifer, A. (2019), 'Beliefs about gender', *American Economic Review* **109**(3), 739–73.

Borker, G. et al. (2021), *Safety first: Perceived risk of street harassment and educational choices of women*, World Bank.

Boulding, C. E. (2010), 'Ngos and political participation in weak democracies: Subnational evidence on protest and voter turnout from bolivia', *The Journal of Politics* **72**(2), 456–468.

Bourey, C., Williams, W., Bernstein, E. E. and Stephenson, R. (2015), 'Systematic review of structural interventions for intimate partner violence in low-and middle-income countries: organizing evidence for prevention', *BMC public health* **15**(1), 1–18.

Brown, J. R. and Goolsbee, A. (2002), 'Does the internet make markets more competitive? evidence from the life insurance industry', *Journal of political economy* **110**(3), 481–507.

Brown, W. J. and Singhal, A. (1999), 'Entertainment-education media strategies for social change: Promises and problems', *Mass media social control and social* pp. 263–280.

Bruhn, M. and McKenzie, D. (2009), 'In pursuit of balance: Randomization in practice in development field experiments', *American economic journal: applied economics* **1**(4), 200–232.

Brynjolfsson, E., Hu, Y. and Smith, M. D. (2010), 'Research commentary—long tails vs. superstars: The effect of information technology on product variety and sales concentration patterns', *Information Systems Research* **21**(4), 736–747.

Brynjolfsson, E. and Smith, M. D. (2000), 'Frictionless commerce? a comparison of internet and conventional retailers', *Management science* **46**(4), 563–585.

Burchell, K., Rettie, R. and Patel, K. (2013), 'Marketing social norms: social marketing and the 'social norm approach'', *Journal of Consumer behaviour* **12**(1), 1–9.

Bursztyn, L., González, A. L. and Yanagizawa-Drott, D. (2020), 'Misperceived social norms: Women working outside the home in saudi arabia', *American economic review* **110**(10), 2997–3029.

Buys, P., Dasgupta, S., Thomas, T. S. and Wheeler, D. (2009), 'Determinants of a digital divide in sub-saharan africa: A spatial econometric analysis of cell phone coverage', *World Development* **37**(9), 1494–1505.

Cabral, L. and Hortacsu, A. (2010), 'The dynamics of seller reputation: Evidence from ebay', *The Journal of Industrial Economics* **58**(1), 54–78.

Campante, F., Durante, R. and Sobbrio, F. (2018), 'Politics 2.0: The multifaceted effect of broadband internet on political participation', *Journal of the European Economic Association* **16**(4), 1094–1136.

Carlana, M. (2019), 'Implicit stereotypes: Evidence from teachers' gender bias', *The Quarterly Journal of Economics* **134**(3), 1163–1224.

Casey, K. (2015), 'Crossing party lines: The effects of information on redistributive politics', *American Economic Review* **105**(8), 2410–48.

Cavallo, A. (2017), 'Are online and offline prices similar? evidence from large multi-channel retailers', *American Economic Review* **107**(1), 283–303.

Chan, Y.-S. and Leland, H. (1982), 'Prices and qualities in markets with costly information', *The Review of Economic Studies* **49**(4), 499–516.

Chen, Y. and Xie, J. (2005), 'Third-party product review and firm marketing strategy', *Marketing science* **24**(2), 218–240.

Chen, Y. and Xie, J. (2008), 'Online consumer review: Word-of-mouth as a new element of marketing communication mix', *Management science* **54**(3), 477–491.

Chen, Y. and Yang, D. Y. (2019), 'The impact of media censorship: 1984 or brave new world?', *American Economic Review* **109**(6), 2294–2332.

Chevalier, J. A., Dover, Y. and Mayzlin, D. (2018), 'Channels of impact: User reviews when quality is dynamic and managers respond', *Marketing Science* **37**(5), 688–709.

Chevalier, J. A. and Mayzlin, D. (2006), 'The effect of word of mouth on sales: Online book reviews', *Journal of marketing research* **43**(3), 345–354.

Chiang, C.-F. and Knight, B. (2011), 'Media bias and influence: Evidence from newspaper endorsements', *The Review of economic studies* **78**(3), 795–820.

Christensen, D. and Garfias, F. (2018), 'Can you hear me now? how communication technology affects protest and repression', *Quarterly journal of political science* **13**(1), 89.

Cooper, R. and Ross, T. W. (1984), 'Prices, product qualities and asymmetric information: The competitive case', *The Review of Economic Studies* **51**(2), 197–207.

Dai, W. and Luca, M. (2020), 'Digitizing disclosure: The case of restaurant hygiene scores', *American Economic Journal: Microeconomics* **12**(2), 41–59.

Dall'orso, J., Gauriot, R. and Page, L. (2016), Disappointment looms around the corner: Visibility and local businesses' market power, Technical report, QUT Business School.

DellaVigna, S., Enikolopov, R., Mironova, V., Petrova, M. and Zhuravskaya, E. (2014), 'Cross-border media and nationalism: Evidence from serbian radio in croatia', *American Economic Journal: Applied Economics* **6**(3), 103–32.

DellaVigna, S. and Kaplan, E. (2007), 'The fox news effect: Media bias and voting', *The Quarterly Journal of Economics* **122**(3), 1187–1234.

Dhar, D., Jain, T. and Jayachandran, S. (2022), 'Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in india', *American Economic Review* **112**(3), 899–927.

Dlomo, P. A. (2017), The impact of irregular expenditure in the South African public finance with specific reference to the National Department of Public Works, PhD thesis, Cape Peninsula University of Technology.

Donati, D. (2019), 'Mobile internet access and political outcomes: Evidence from south africa', *Universitat Pompeu Fabra, Mimeo* .

Durante, R., Pinotti, P. and Tesei, A. (2019), 'The political legacy of entertainment tv', *American Economic Review* **109**(7), 2497–2530.

Duvvury, N., Callan, A., Carney, P. and Raghavendra, S. (2013), 'Intimate partner violence: Economic costs and implications for growth and development'.

Eisensee, T. and Strömberg, D. (2007), 'News droughts, news floods, and us disaster relief', *The Quarterly Journal of Economics* **122**(2), 693–728.

Elfenbein, D. W., Fisman, R. and McManus, B. (2015), 'Market structure, reputation, and the value of quality certification', *American Economic Journal: Microeconomics* **7**(4), 83–108.

Eliashberg, J. and Shugan, S. M. (1997), 'Film critics: Influencers or predictors?', *Journal of marketing* **61**(2), 68–78.

Ellison, G. and Fisher Ellison, S. (2005), 'Lessons about markets from the internet', *Journal of Economic perspectives* **19**(2), 139–158.

Enikolopov, R., Makarin, A. and Petrova, M. (2020), 'Social media and protest participation: Evidence from russia', *Econometrica* **88**(4), 1479–1514.

Enikolopov, R., Petrova, M. and Sonin, K. (2018), 'Social media and corruption', *American Economic Journal: Applied Economics* **10**(1), 150–74.

Enikolopov, R., Petrova, M. and Zhuravskaya, E. (2011), 'Media and political persuasion: Evidence from russia', *American Economic Review* **101**(7), 3253–85.

Ershov, D. (2020), 'Consumer product discovery costs, entry, quality and congestion in online markets', *Unpublished manuscript* .

Falck, O., Gold, R. and Heblich, S. (2014), 'E-lections: Voting behavior and the internet', *American Economic Review* **104**(7), 2238–65.

Fang, L. (2022), 'The effects of online review platforms on restaurant revenue, consumer learning, and welfare', *Management Science* .

Farronato, C. and Zervas, G. (2019), 'Consumer reviews and regulation: evidence from nyc restaurants', *Working Paper* .

Feddersen, T. J. and Pesendorfer, W. (1996), 'The swing voter's curse', *The American economic review* pp. 408–424.

Ferraz, C. and Finan, F. (2008), 'Exposing corrupt politicians: the effects of brazil's publicly released audits on electoral outcomes', *The Quarterly journal of economics* **123**(2), 703–745.

Ferri, M., Allara, E., Bo, A., Gasparrini, A. and Faggiano, F. (2013), 'Media campaigns for the prevention of illicit drug use in young people', *Cochrane Database of Systematic Reviews* (6).

Fershtman, C., Fishman, A. and Zhou, J. (2018), 'Search and categorization', *International Journal of Industrial Organization* **57**, 225–254.

Fishman, A. and Levy, N. (2015), 'Search costs and investment in quality', *The Journal of Industrial Economics* **63**(4), 625–641.

Flood, M. and Pease, B. (2009), 'Factors influencing attitudes to violence against women', *Trauma, violence, & abuse* **10**(2), 125–142.

Frank, L. B., Murphy, S. T., Chatterjee, J. S., Moran, M. B. and Baezconde-Garbanati, L. (2015), ‘Telling stories, saving lives: creating narrative health messages’, *Health communication* **30**(2), 154–163.

Gavazza, A. and Lizzeri, A. (2009), ‘Transparency and economic policy’, *The Review of Economic Studies* **76**(3), 1023–1048.

Gavazza, A., Nardotto, M. and Valletti, T. (2019), ‘Internet and politics: Evidence from uk local elections and local government policies’, *The Review of Economic Studies* **86**(5), 2092–2135.

Gennaioli, N., La Porta, R., Lopez-de Silanes, F. and Shleifer, A. (2013), ‘Human capital and regional development’, *The Quarterly journal of economics* **128**(1), 105–164.

Gentzkow, M., Glaeser, E. L. and Goldin, C. (2007), *6. The Rise of the Fourth Estate: How Newspapers Became Informative and Why It Mattered*, University of Chicago Press.

Gentzkow, M., Petek, N., Shapiro, J. M. and Sinkinson, M. (2015), ‘Do newspapers serve the state? incumbent party influence on the us press, 1869–1928’, *Journal of the European Economic Association* **13**(1), 29–61.

Gentzkow, M., Shapiro, J. M. and Sinkinson, M. (2011), ‘The effect of newspaper entry and exit on electoral politics’, *American Economic Review* **101**(7), 2980–3018.

George, L. M. (2008), ‘The internet and the market for daily newspapers’, *The BE Journal of Economic Analysis & Policy* **8**(1).

Gerber, A. S. and Green, D. P. (2012), *Field experiments: Design, analysis, and interpretation*, WW Norton.

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.

Goetz, A. M. and Jenkins, R. (2001), ‘Hybrid forms of accountability: citizen engagement in institutions of public-sector oversight in india’, *Public Management Review* **3**(3), 363–383.

Goffman, E. (1963), ‘Stigma englewood cliffs’, *NJ: Spectrum* pp. 127–128.

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

Goldmanis, M., Hortaçsu, A., Syverson, C. and Emre, Ö. (2010), ‘E-commerce and the market structure of retail industries’, *The Economic Journal* **120**(545), 651–682.

Gonzales, M., León Ciliotta, G. and Martínez, L. R. (2019), 'How effective are monetary incentives to vote? evidence from a nationwide policy', *Evidence from a Nationwide Policy (July 2019). University of Chicago, Becker Friedman Institute for Economics Working Paper* (2019-101).

Green, D. P., Wilke, A. M. and Cooper, J. (2020), 'Countering violence against women by encouraging disclosure: A mass media experiment in rural uganda', *Comparative Political Studies* **53**(14), 2283–2320.

Grossman, G. M. and Helpman, E. (1996), 'Electoral competition and special interest politics', *The Review of Economic Studies* **63**(2), 265–286.

Grzybowski, L. and Muñoz, A. (2020), 'Impact of roaming regulation on revenues and prices of mobile operators in the eu', *Unpublished manuscript* .

GSMA, I. (2015), 'The mobile economy (2013)', *White Paper* .

Guriev, S., Melnikov, N. and Zhuravskaya, E. (2019), '3g internet and confidence in government', *Available at SSRN 3456747* .

Hall, A. E. and Bracken, C. C. (2011), '"i really liked that movie": Testing the relationship between trait empathy, transportation, perceived realism, and movie enjoyment.', *Journal of Media Psychology: Theories, Methods, and Applications* **23**(2), 90.

Halpern, L., Koren, M. and Szeidl, A. (2015), 'Imported inputs and productivity', *American Economic Review* **105**(12), 3660–3703.

Hansman, C., Hjort, J., León-Ciliotta, G. and Teachout, M. (2020), 'Vertical integration, supplier behavior, and quality upgrading among exporters', *Journal of Political Economy* **128**(9), 3570–3625.

Harris, M. (2011), How cell towers work, *in* 'UNISON'.

Hjort, J., Iyer, V. and De Rochambeau, G. (2020), Informational barriers to market access: Experimental evidence from liberian firms, Technical report, National Bureau of Economic Research.

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

Hoff, K. and Stiglitz, J. E. (2010), 'Equilibrium fictions: A cognitive approach to societal rigidity', *American Economic Review* **100**(2), 141–46.

Hoff, K. and Walsh, J. (2018), 'The whys of social exclusion: Insights from behavioral economics', *The World Bank Research Observer* **33**(1), 1–33.

Hollenbeck, B., Moorthy, S. and Proserpio, D. (2019), 'Advertising strategy in the presence of reviews: An empirical analysis', *Marketing Science* **38**(5), 793–811.

Hopenhayn, H. A. (1992), 'Entry, exit, and firm dynamics in long run equilibrium', *Econometrica: Journal of the Econometric Society* pp. 1127–1150.

Hui, X., Saeedi, M. and Sundaresan, N. (2018), 'Adverse selection or moral hazard, an empirical study', *The Journal of Industrial Economics* **66**(3), 610–649.

Jenkins, R. and Goetz, A. M. (1999), 'Accounts and accountability: theoretical implications of the right-to-information movement in india', *Third world quarterly* **20**(3), 603–622.

Jensen, R. (2007), 'The digital provide: Information (technology), market performance, and welfare in the south indian fisheries sector', *The quarterly journal of economics* **122**(3), 879–924.

Jensen, R. and Miller, N. H. (2018), 'Market integration, demand, and the growth of firms: Evidence from a natural experiment in india', *American Economic Review* **108**(12), 3583–3625.

Jensen, R. and Oster, E. (2009), 'The power of tv: Cable television and women's status in india', *The Quarterly Journal of Economics* **124**(3), 1057–1094.

Jewkes, R., Willan, S., Heise, L., Washington, L., Shai, N., Kerr-Wilson, A. and Christofides, N. (2020), 'Effective design and implementation elements in interventions to prevent violence against women and girls', *What works to prevent VAWG* .

Jin, G. Z. and Leslie, P. (2003), 'The effect of information on product quality: Evidence from restaurant hygiene grade cards', *The Quarterly Journal of Economics* **118**(2), 409–451.

Jin, G. Z. and Leslie, P. (2009), 'Reputational incentives for restaurant hygiene', *American Economic Journal: Microeconomics* **1**(1), 237–67.

Kearney, M. S. and Levine, P. B. (2019), 'Early childhood education by television: Lessons from sesame street', *American Economic Journal: Applied Economics* **11**(1), 318–50.

Kerr-Wilson, A., Gibbs, A., McAslan Fraser, E., Ramsoomar, L., Parke, A., Khuwaja, H. M. and Jewkes, R. (2020), 'A rigorous global evidence review of interventions to prevent violence against women and girls', *What Works to prevent violence among women and girls global Programme, Pretoria, South Africa* .

Klein, L. R. (1998), 'Evaluating the potential of interactive media through a new lens: Search versus experience goods', *Journal of business research* **41**(3), 195–203.

Klein, T. J., Lambertz, C. and Stahl, K. O. (2016), 'Market transparency, adverse selection, and moral hazard', *Journal of Political Economy* **124**(6), 1677–1713.

Kling, J. R., Liebman, J. B. and Katz, L. F. (2007), 'Experimental analysis of neighborhood effects', *Econometrica* **75**(1), 83–119.

Kondylis, F., Legovini, A., Vyborny, K., Zwager, A. M. T. and Cardoso De Andrade, L. (2020), 'Demand for safe spaces: Avoiding harassment and stigma', *World Bank Policy Research Working Paper* (9269).

Kugler, M. and Verhoogen, E. (2012), 'Prices, plant size, and product quality', *The Review of Economic Studies* **79**(1), 307–339.

La Ferrara, E., Chong, A. and Duryea, S. (2012), 'Soap operas and fertility: Evidence from brazil', *American Economic Journal: Applied Economics* **4**(4), 1–31.

Lavy, V. and Sand, E. (2015), On the origins of gender human capital gaps: Short and long term consequences of teachers' stereotypical biases, Technical report, National Bureau of Economic Research.

Leland, H. E. (1979), 'Quacks, lemons, and licensing: A theory of minimum quality standards', *Journal of political economy* **87**(6), 1328–1346.

Levy, R. and Mattsson, M. (2021), 'The effects of social movements: Evidence from# metoo', *Available at SSRN 3496903* .

Lewis, G. and Zervas, G. (2019), 'The supply and demand effects of review platforms', *Available at SSRN 3468278* .

Lindbeck, A. and Weibull, J. W. (1987), 'Balanced-budget redistribution as the outcome of political competition', *Public choice* **52**(3), 273–297.

Liu, M., Brynjolfsson, E. and Dowlatabadi, J. (2021), 'Do digital platforms reduce moral hazard? the case of uber and taxis', *Management Science* .

Lohmann, S. (1998), 'An information rationale for the power of special interests', *American Political Science Review* **92**(4), 809–827.

Luca, M. (2016), 'Reviews, reputation, and revenue: The case of yelp. com', *Com (March 15, 2016). Harvard Business School NOM Unit Working Paper* (12-016).

Mackie, G. (1996), 'Ending footbinding and infibulation: A convention account', *American sociological review* pp. 999–1017.

Malherbe, D. (2015), The political use of new media in the 2014 South African national election, PhD thesis, Stellenbosch: Stellenbosch University.

Manacorda, M. and Tesei, A. (2020), 'Liberation technology: Mobile phones and political mobilization in africa', *Econometrica* **88**(2), 533–567.

Matsusaka, J. G. (1995), 'Fiscal effects of the voter initiative: Evidence from the last 30 years', *Journal of political Economy* **103**(3), 587–623.

McCall, J. J. (1970), 'Economics of information and job search', *The Quarterly Journal of Economics* pp. 113–126.

McKenzie-Mohr, D. (2000), 'Fostering sustainable behavior through community-based social marketing.', *American psychologist* **55**(5), 531.

Melitz, M. J. (2003), 'The impact of trade on intra-industry reallocations and aggregate industry productivity', *econometrica* **71**(6), 1695–1725.

Miller, D. T. and McFarland, C. (1987), 'Pluralistic ignorance: When similarity is interpreted as dissimilarity.', *Journal of Personality and social Psychology* **53**(2), 298.

Miner, L. (2015), 'The unintended consequences of internet diffusion: Evidence from malaysia', *Journal of Public Economics* **132**, 66–78.

Minges, M., Briceño-Garmendia, C., Williams, M., Ampah, M., Camos, D. and Shkratan, M. (2008), 'Information and communications technology in sub-saharan africa: A sector review', *Background paper* **10**.

Moraga-González, J. L. and Sun, Y. (2020), 'Product quality and consumer search'.

Morton, F. S., Zettelmeyer, F. and Silva-Risso, J. (2001), 'Internet car retailing', *The Journal of Industrial Economics* **49**(4), 501–519.

Moyer-Gusé, E. (2008), 'Toward a theory of entertainment persuasion: Explaining the persuasive effects of entertainment-education messages', *Communication theory* **18**(3), 407–425.

Nelson, P. (1970), 'Information and consumer behavior', *Journal of political economy* **78**(2), 311–329.

Nunn, N. and Puga, D. (2012), 'Ruggedness: The blessing of bad geography in africa', *Review of Economics and Statistics* **94**(1), 20–36.

Nyhan, B., Reifler, J., Richey, S. and Freed, G. L. (2014), 'Effective messages in vaccine promotion: a randomized trial', *Pediatrics* **133**(4), e835–e842.

Ochoa, C. Y., Murphy, S. T., Frank, L. B. and Baezconde-Garbanati, L. A. (2020), 'Using a culturally tailored narrative to increase cervical cancer detection among spanish-speaking mexican-american women', *Journal of Cancer Education* **35**(4), 736–742.

Oliver, M. B., Dillard, J. P., Bae, K. and Tamul, D. J. (2012), 'The effect of narrative news format on empathy for stigmatized groups', *Journalism & Mass Communication Quarterly* **89**(2), 205–224.

Olken, B. A. (2007), 'Monitoring corruption: evidence from a field experiment in indonesia', *Journal of political Economy* **115**(2), 200–249.

Olken, B. A. (2009), 'Do television and radio destroy social capital? evidence from indonesian villages', *American Economic Journal: Applied Economics* **1**(4), 1–33.

Orlov, E. (2011), 'How does the internet influence price dispersion? evidence from the airline industry', *The Journal of Industrial Economics* **59**(1), 21–37.

Orozco-Olvera, V., Shen, F. and Cluver, L. (2019), 'The effectiveness of using entertainment education narratives to promote safer sexual behaviors of youth: A meta-analysis, 1985-2017', *PloS one* **14**(2), e0209969.

Paluck, E. L. and Green, D. P. (2009), 'Deference, dissent, and dispute resolution: An experimental intervention using mass media to change norms and behavior in rwanda', *American political Science review* **103**(4), 622–644.

Paul, S. (2002), *Holding the state to account: Citizen monitoring in action*, Public Affairs Centre.

Petrova, M., Bursztyn, L., Egorov, G. and Enikolopov, R. (2020), 'Social media and xenophobia: Evidence from russia'.

Petrova, M., Sen, A. and Yildirim, P. (2021), 'Social media and political contributions: the impact of new technology on political competition', *Management Science* **67**(5), 2997–3021.

Pinkovskiy, M. and Sala-i Martin, X. (2016), 'Lights, camera, income! illuminating the national accounts-household surveys debate', *The Quarterly Journal of Economics* **131**(2), 579–631.

Prior, M. (2005), 'News vs. entertainment: How increasing media choice widens gaps in political knowledge and turnout', *American Journal of Political Science* **49**(3), 577–592.

Proserpio, D. and Zervas, G. (2017), 'Online reputation management: Estimating the impact of management responses on consumer reviews', *Marketing Science* **36**(5), 645–665.

Quinn, M., Godinho de Matos, M. and Peukert, C. (2021), 'The economic effects of mobile internet access – evidence from roam-like-at-home', *Unpublished manuscript* .

Raghavendra, S., Kim, K., Ashe, S., Chadha, M., Asante, F. A., Piiroinen, P. T. and Duvvury, N. (2019), The macroeconomic loss due to violence against women and girls: the case of ghana, Technical report, Working Paper.

Rao, N., Donati, D. and Orozco-Olvera, V. (2020), 'Conducting surveys and interventions entirely online: a virtual lab practitioner's manual', *Working Paper* .

Reimers, I. and Waldfogel, J. (2021), 'Digitization and pre-purchase information: the causal and welfare impacts of reviews and crowd ratings', *American Economic Review* **111**(6), 1944–71.

Reinikka, R. and Svensson, J. (2005), 'Fighting corruption to improve schooling: Evidence from a newspaper campaign in uganda', *Journal of the European economic association* **3**(2-3), 259–267.

Resnick, P., Zeckhauser, R., Swanson, J. and Lockwood, K. (2006), 'The value of reputation on ebay: A controlled experiment', *Experimental economics* **9**(2), 79–101.

Riordan, M. H. (1986), 'Monopolistic competition with experience goods', *The Quarterly Journal of Economics* **101**(2), 265–279.

Rotesi, T. (2019), Do social media matter? the impact of twitter on political participation, Technical report, Mimeo.

Salop, S. and Stiglitz, J. (1977), 'Bargains and ripoffs: A model of monopolistically competitive price dispersion', *The Review of Economic Studies* **44**(3), 493–510.

Sardinha, L. and Catalán, H. E. N. (2018), 'Attitudes towards domestic violence in 49 low-and middle-income countries: A gendered analysis of prevalence and country-level correlates', *PloS one* **13**(10), e0206101.

Shin, M., Shin, J., Ghili, S. and Kim, J. (2021), 'The impact of gig economy on product quality through the labor market: Evidence from ride-sharing and restaurant quality'.

Singhal, A., Cody, M. J., Rogers, E. M. and Sabido, M. (2003), *Entertainment-education and social change: History, research, and practice*, Routledge.

Singhal, A. and Rogers, E. (2012), *Entertainment-education: A communication strategy for social change*, Routledge.

Snyder Jr, J. M. and Strömberg, D. (2010), 'Press coverage and political accountability', *Journal of political Economy* **118**(2), 355–408.

Startz, M. (2018), 'The value of face-to-face: Search and contracting problems in nigerian trade', *Available at SSRN 3096685* .

Strömberg, D. (2004), 'Radio's impact on public spending', *The Quarterly Journal of Economics* **119**(1), 189–221.

Syverson, C. (2004), 'Market structure and productivity: A concrete example', *Journal of Political Economy* **112**(6), 1181–1222.

Tadelis, S. (2016), 'Reputation and feedback systems in online platform markets', *Annual Review of Economics* **8**, 321–340.

Terrier, C. (2015), 'Giving a little help to girls? evidence on grade discrimination and its effect on students' achievement'.

W. Vaughan, Everett M. Rogers, A. S. R. M. S. P. (2000), 'Entertainment-education and hiv/aids prevention: A field experiment in tanzania', *Journal of health communication* **5**(sup1), 81–100.

Wagman, J. A., Gray, R. H., Campbell, J. C., Thoma, M., Ndyanabo, A., Ssekasanvu, J., Nalugoda, F., Kagaayi, J., Nakigozi, G., Serwadda, D. et al. (2015), 'Effectiveness of an integrated intimate partner violence and hiv prevention intervention in rakai, uganda: analysis of an intervention in an existing cluster randomised cohort', *The Lancet Global Health* **3**(1), e23–e33.

Wang, H. and Singhal, A. (2016), 'East los high: Transmedia edutainment to promote the sexual and reproductive health of young latina/o americans', *American journal of public health* **106**(6), 1002–1010.

Wang, Y. and Chaudhry, A. (2018), 'When and how managers' responses to online reviews affect subsequent reviews', *Journal of Marketing Research* **55**(2), 163–177.

WHO (2013), *Global and regional estimates of violence against women: prevalence and health effects of intimate partner violence and non-partner sexual violence*, World Health Organization.

Wolinsky, A. (1983), 'Prices as signals of product quality', *The review of economic studies* **50**(4), 647–658.

Wolinsky, A. (2005), 'Procurement via sequential search', *Journal of Political Economy* **113**(4), 785–810.

World-Bank (2014), *World development report 2015: Mind, society, and behavior*, The World Bank.

Zhuravskaya, E., Petrova, M. and Enikolopov, R. (2020), 'Political effects of the internet and social media', *Annual Review of Economics* **12**, 415–438.