ICIS 2022 Proceedings

IS in Healthcare

Dec 12th, 12:00 AM

# Healthy reviews!- Online physician ratings reduce healthcare interruptions

Michael Kummer
*University of East Anglia*, M.Kummer@uea.ac.uk

Ulrich Laitenberger
*Telecom Paris*, laitenberger@enst.fr

Danny Hughes
*Georgia Institute of Technology*, danny.hughes@econ.gatech.edu

Cyrus Rich
*Georgia Institute of Technology*, crich18@gmail.com

Turgay Ayer
*Georgia Institute of Technology*, ayer@isye.gatech.edu

# Healthy reviews!

# Online physician ratings reduce healthcare interruptions

*Completed Research Paper*

**Michael Kummer**
UEA
Department of Economics
M.kummer@uea.ac.uk

**Ulrich Laitenberger**
Telecom Paris,
Institut Polytechnique de Paris
ulrich.laitenberger@telecom-paris.fr

**Cyrus Rich**
Georgia Institute of Technology
H. Milton Stewart School of Industrial and Systems Engineering,
crich@gatech.edu

**Danny Hughes**
Georgia Institute of Technology
School of Economics
danny.hughes@econ.gatech.edu

**Turgay Ayer**
Georgia Institute of Technology
Scheller College of Business
ayer@isye.gatech.edu

## Abstract

*We show that review platforms reduce healthcare interruptions for patients looking for a new physician. We employ a difference-in-differences strategy using physician retirements as a "disruptive shock" that forces patients to find a new physician. We combine insurance claims data with web-scraped physician reviews and highlight a substantial care-gap resulting from a physician's retirement. We then show that online physician reviews reduce this gap and help patients find a new physician faster. Our results are robust to including a variety of controls and various instruments for the availability of physician reviews, but are not found for patients of nonretiring physicians. By reducing interruptions in care, reviews can improve clinical outcomes and lower costs.*

**Keywords:** Healthcare, Online physician ratings, Online physician reviews, Care-gap.

## Introduction

Finding good health-care providers is not easy, because health-care services are credence goods (Darby and Karni, 1973). Patients traditionally lack objective information about the quality of physicians. Many patients rely therefore on word-of-mouth recommendations from friends and relatives (Tu and Lauer, 2008) when selecting new primary care physicians, and even physicians rely on their inside information for referrals (Hackl et al., 2015). The lack of information is especially problematic for patients who need to choose a new provider because they changed insurance plans, their residence, or experience other life changes. Such patients might delay a visit to seek care and treatment, which can result in suboptimal health outcomes (Barach et al., 2020). The retirement of a primary care physician decreases their patients' future primary care utilization, and increases specialty care, emergent care, and charges (Sabety et al., 2021).

In this paper, we show that health-related online reviews help patients to overcome the friction associated with choosing a new physician. We analyze patient behavior and document a substantial care-gap - a delay in finding care - before patients visit a new physician. We further show that online physician reviews shorten this care-gap and help patients to find new physicians faster. We leverage the retirement of physicians as a "disruptive shock" that induces a patient to search for a new physician. We analyze the time lapse until patients visit a new physician after their physician has retired. We identify the effect of reviews using a difference-in-differences (DiD) strategy. In this strategy, we compare cities with a large accumulation of online reviews to cities with very few online reviews.

We conduct our analysis by augmenting a unique data set on individual-level patient claims with information from a major review platform for physicians (yelp.com). The combined database is ideally suited for studying our key research question. The medical claims data covers the entire population of patients associated with the included insurance carriers (over 55 million unique individuals). Through the pattern of insurance claims we can track whether patients experience an interruption to their care after their physician's retirement. The data about online reviews indicates the availability of information about physicians in a patient's city, and how that availability changes over time.

The empirical challenge when measuring the effect of reviews on patient behavior is that it is difficult for a researcher to observe when patients want to find a new physician. Most patients typically have stable relationships with their physicians, making a patient searching for a new physician a low-frequency event. Focusing on patients of retiring physicians allows overcoming this challenge: First, the retirement of a physician requires patients to search for a new physician. Second, researchers can observe retirements. Our strategy focuses on a specific type of interruption, but we expect other major life changes of patients to induce similar interruptions.

We document two main findings. First, physician retirements cause disruptions in care. Specifically, physician retirements increase the gap between two visits by more than 140 days on average. Second, this gap is shorter when online reviews are available. In particular, the gap in time between two visits is 30.7 days shorter on average in the presence of online reviews, and patients are 7 percentage points more likely to visit a new physician within 15 months of their current physician retiring.

We demonstrate the robustness of our findings by showing that the effect becomes insignificant for patients of nonretiring physicians, highlighting that no shifts in underlying trends occurred in the pretrend analysis and by controlling for various alternative and potentially omitted variables (including income, education, broadband availability, and the number of physicians in the county). Moreover, we apply an instrumental variable (IV) strategy that exploits reviews in the same city for other professions to rule out the possibility that the effect is driven by physician-specific technology adoption.

Our results have several implications for research and policy. Review platforms appear to have a decisive impact on patients' health-related decision making when patients lose their physician. Alas, health related online reviews are scattered across various platforms in the United States and are often completely absent in many other countries. Our results suggest that insurers and public health organizations can motivate patients to receive more continuous care when their relationships get disrupted. Scaling online review platforms could be a way to reduce disruptions in care, which can have severe negative effects on a variety of clinical outcomes (Barach et al., 2020). Moreover, high quality online reviews promise to be a cost-effective measure to improve access to care since review content is user generated and can be made widely available at low cost. Lastly, we demonstrate that it is feasible to identify a precise mechanism through which online reviews affect patient outcomes, and future research can build on this step to quantify the link between online reviews and clinical outcomes.

## Literature review and study contributions

By highlighting the care-gap after physician retirements and showing that reviews quicken the process in which patients choose a new physician we make several contributions.

Online and medical reviews. We contribute to the literature that has shown how online reviews affect the behavior of individuals and firms (Chevalier and Mayzlin, 2006; Duan et al., 2008; Forman et al., 2008; Cabral and Hortacsu, 2010; Luca, 2011; Helmers et al., 2019; Reimers and Waldfogel, forthcoming). They achieve this effect by conveying accurate information on the quality of goods or services and informing consumer choices (Hu et al., 2006; Lu and Rui, 2018; Yin et al., 2016; Howard et al., 2017; Sahoo et al., 2018; Choi et al., 2019). Our paper contributes by highlighting the ability of online reviews to increase overall economic efficiency in health care.

We also contribute to the literature on how online health-related information can affect patient choices and health outcomes (Billari et al., 2019; Amaral-Garcia et al., forthcoming). A series of papers have highlighted the correlation between physician reviews and physician quality and have shown that physician demand is affected by online reviews, which can induce congestion, concentration and – ultimately - inefficient allocation of patients (Greaves et al., 2012; Gao et al., 2012; Emmert and Meier, 2013; Lu and Rui, 2018; Saifee et al., 2020; Chen and Lee, 2021; Luca and Vats, 2014; Lu and Wu, 2019; Kaye, 2020; Bensnes and Huitfeldt, 2021; Chen and Lee, 2021; Xu et al., 2021). Shukla et al. (2021) have shown that online health reviews reduce patients' search cost and influence their choices. Our findings expand beyond prior knowledge, because we highlight that reviews do not only influence which physicians are chosen, but they also help patients to choose any physician and to make their choice faster.

## Background: physician reviews in the US

**On Physician review sites.** Patients increasingly rely on physician review websites to find healthcare practitioners. Customer surveys suggest that about 72% of consumers use physician rating sites as the first step in finding a new physician,1 80% trusted online reviews as much as personal recommendations from acquaintances, and 47% preferred out-of-network physicians with better reviews over in-network physicians with comparable qualifications but poorer reviews. Holliday et al. (2017) find that 53% of physicians have visited a physician review website, potentially in an effort to improve patient satisfaction.

Accordingly, over 70 different websites host reviews of physicians in the US. This increased availability of reviews is driven by both new dedicated sites (e.g. Healthgrades.com, RateMDs.com, and ZocDoc.com), and established platforms that added a platform for physician reviews (e.g. Google, Facebook, and Yelp). In 2010 the Centers for Medicare & Medicaid Services launched "Physician Compare," which saw lower usage than expected. While it would be ideal to observe the other big platforms as well, we focus on yelp as the most widely used review site.[1] Including only yelp potentially leads to under-reported estimates, if the review availability of other platforms across cities correlates negatively with review availability on yelp.

**Crowd-based ratings on Yelp.** Yelp is a website where consumers can leave reviews for a variety of businesses. It is widely available across the US and has high coverage of physicians. On Yelp, users can freely read and write reviews of physicians. Users can give ratings from 1 to 5 stars and add a narrative review. By mid 2020, Yelp users have contributed over 21 million health-related reviews.

## Empirical Strategy

**Hypothesis development:** Our main hypothesis is that online reviews affect patient choices by rendering the task of searching for a new physician of high quality easier (Shukla et al., 2021). We expect that online reviews decrease search cost and reduce patients' uncertainty about the quality of physicians.

To formalize this idea, consider time-constrained patients forced to find a new physician. Denote the net benefit of searching for and contacting a physician as vp. The net benefit vp depends negatively on the patient's search cost cs and positively on the new patient's expected match value E(q) with the physician.

[1] In future work we plan to widen our pool of online reviews.

Assume (without loss of generality) that patients do one task per day and have an alternative action of value v, which is randomly drawn from a continuous distribution. Each day, patients choose the task with the highest net value out of v, $v_p$. The probability of a patient doing something other than finding a physician is given by P($v_p \leq$ v). The expected number of days until a patient searches for a physician can be calculated as a probability-weighted average of the different infinite possible outcomes:

$$E(DAY\,S) = (1 - P(v_p \leq v)) \cdot 1 + P(v_p \leq v) \cdot (E(DAY\,S) + 1) \tag{1}$$

The expected number of days is given by E(DAY S) = 1/(1 − P($v_p$ ≤ v). The probability that the physician search is less attractive than the alternative action enters negatively in the denominator, such that a higher net-value of finding a physician $v_p$ implies a lower expected number of days until the patient engages in search and, consequently, a shorter inter-visit duration. Reviews increase $v_p$ by reducing the search cost (cs), and by raising patients' expected match value E(q). This mechanism leads to fewer days until the next visit. We derive two hypotheses from this mechanism:

- **H1:** If online reviews are available, then patients who need to find a new physician are more likely to do so in any given time window.
- **H2**: If online reviews are available, then patients who need to find a new physician will find their physician faster, conditional on searching.

Note that our hypotheses on the idea that patients play an active role in their choice of specialists, which we think is true for routine visits. However, a similar dynamic would ensue if GPs considered the information form health reviews when deciding their referrals.

**Estimation approach:** Our identification strategy relies on physician retirements as a disruptive shock that forces patients to search a new physician. We use this approach, because patients will generally stick to their existing relationships. Physician retirements have been shown to a negative effect on patients (Lam et al., 2020), and we will document that physician retirements induce a gap in treatment. While a physician's retirement is a specific shock, similar disruptive events can be observed when patients move to a new location, change their insurance plan, or develop a new condition. We would expect to find similar effects of online reviews in patients that experience such life changes.

To analyze whether online reviews shorten the care-gap, we use a DiD approach that quantifies the effect of having a high number of online reviews for physicians in a given city. We estimate linear models with observations at the level of patient p in period t of the form

$$y_{p,t} = \alpha\, post_t + \beta\, did_{p,t} + \gamma\, X_{p,t} + \xi_c + \varepsilon_{p,t}. \tag{2}$$

We estimate this model for two dependent variables yp,t: first, a binary variable that takes a value of 1 if a patient saw a new physician within 15 months (fup15), and, second, the time until the next physician visit measured in days (DAY S), using city-level fixed effects ($\xi_c$). Although fup15 is a binary variable, we use a linear probability model (LPM) because this allows us to obtain consistent estimates while including a large number of fixed effects and interaction terms (Ai, 2003). Note that the fixed effects at the city level replace the usual indicator for the treated group, as this dummy would get absorbed by the fixed effects.

Treated patients live in cities with a large increase in the availability of online reviews between the pre and post period. The precise definition provided in Section 5.1. The term *post* is an indicator variable that takes a value of 1 if the retirement is in the treatment period. The DiD-term $did_{p,t}$ takes a value of 1 if an observation was made in a city with reviews and in the period after treatment (post period). This indicator is built using only reviews before the patients choice to avoid reverse causality. The control variables $X_{p,t}$ capture patient (age, gender) and physician (specialty) characteristics.

Note that we cannot match reviews to physicians to preserve anonymity, but we observe the presence of reviews at the city level. Nevertheless, even our city-level measure of reviews can capture various channels of the usefulness of reviews in reducing the care gap after a physician retires. We further explored measurement concerns in our (omitted) online Appendix C.[2] In this appendix we show that our results are

[2] The Online Appendix had to be omitted for reasons of space but is available upon request.

robust to controlling for additional demographics such as education, highspeed internet, income or number of doctors in the area.

Greater availability of Yelp reviews could be associated with a higher degree of education or engagement of citizens in managing their health care. These factors are likely to be relatively constant over time. To account for unobserved time-constant heterogeneity among cities, we include fixed effects for the city where the patient lives $\xi_c$. We discuss potential unobserved heterogeneity in cities over time in the following subsection.

**Identification** The key challenge in our analysis is identifying whether a causal link exists between the presence of online physician reviews and the time it takes for patients to select their new physicians. The main hindrance to identification is endogeneity caused by i) omitting important explanatory variables and ii) selection into treatment, which we discuss below. In an (omitted) online Appendix C we discuss these and other concerns in greater detail.

**Omitted variable bias.** We use four strategies to address a potential OVB: (1) cityfixed effects in the DiD regression control for any time-invariant factors (geography, city layout, or physical infrastructure,...). (2) we need to control for time-varying factors that may affect the time it takes patients to find a new physician after a retirement. Such factors could be changes in the composition of the population (e.g., age, education, income), technological progress (e.g., improved infrastructure, broadband adoption, adoption of ICT), asymmetric government or insurance policies to shorten the care gap (e.g. reduced entry barriers for physicians), and changes in health-related attitudes. We therefore add control variables for broadband availability, income, education, and the number of physicians. (3) we run a placebo test on patients of non retiring physicians. This placebo analysis documents that the time between visits has not generally decreased in treated cities. The control variable analysis and the placebo analysis are further supported by our pre-trend analysis (Figure 2), which highlights no shifts in the underlying trends until 2012. (4) we take additional precaution and run an IV two-stage least squares (2SLS) estimation based on the presence of reviews in other Yelp categories. This IV addresses any omitted variable bias or reverse causation due to physician-driven factors, that might have begun after 2012 and changed more quickly in treated cities than untreated cities.

Because of these four measures, any remaining omitted variable bias would have to be due to a variable that (1) is time-varying, (2) is correlated with online reviews and the time between visits, (3) is not driven by physicians directly, (4) would not be seen in the pre-trends until 2012 and, (5) influences the time between visits for patients of retiring physicians, but not for patients of non retiring physicians.

**Selection into treatment.** Two more concerns arise because treatment was not randomly assigned. First, unobserved factors might partially govern which cities received treatment (e.g. physicians' technology adoption such as appointment booking systems). This concern is akin to OVB, just with a focus on drivers of review provision. It is addressed by our robustness checks in Table 3. Second, treatment could be more effective in treated places. And less effective in untreated places – even if they were created exogenously. Hence, we interpret our results as Average Treatment Effect on the Treated (ATT), but we argue that this issue will disappear as adoption progresses.

**Measurement Error.** While it would be ideal to observe the other big platforms as well, we focus on yelp as the most widely used review site. Including only yelp potentially induces a downward bias in our effect estimates, if other platforms have better review availability in places in which yelp is underrepresented and vice-verca. In the Online Appendix[3] we provide a more extensive discussion how our identification strategy addresses reverse causality, OVB, simultaneity, measurement error and selection into treatment.

[3] The Online Appendix had to be omitted for reasons of space, but is available upon request.

# Data

## *Sources and preparation*

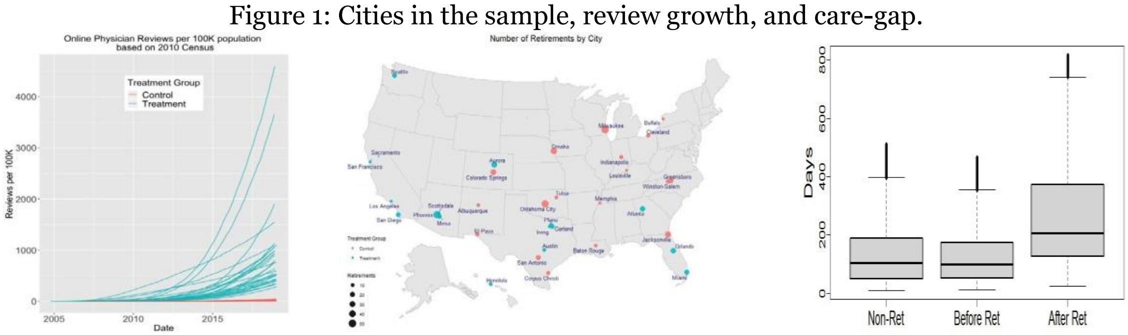
Our analysis combines patient-physician-level data with city-level data on online reviews. We use data from the website Yelp.com to measure the availability of physician reviews across the United States and Optum's deidentified Clinformatics R Data Mart Database (CDM) to observe patients' physician choices.

**Physician review data and the definition of the control and treatment group.** Starting from the top 100 US cities by population, we selected the 30 cities with the lowest and highest review accumulation per capita. The full list of cities can be found in Tables A2 and A3 in an (omitted) online Appendix A.2 which is available upon request. In this Appendix we also describe the details of our procedure. Until the late 2000s – our pre-treatment period – hardly any reviews existed in any given city.

The left panel of Figure 1 shows that our procedure was effective, as reviews indeed became widely available in the treated cities but not in the cities in our control group. The central panel of Figure 1 shows the cities in which we identified physician retirements, and the boxplot on the right highlights the care-gap that results from a physician's retirement. The average number of days until the next visit after a retirement exceeds 275 days, whereas the average time between visits without a physician retirement is 137.75 days. Similarly, the average time between visits before a retirement is 125.02 (see Figure 1)

**Patient data.** Optum's deidentified Clinformatics Data Mart Database (CDM), is a commercial and Medicare Advantage claims database with beneficiaries in all 50 US states from 2007 to 2018. Claims are created whenever patients visit their physician and the physician charges the patients' health insurance for the payment (in full or in part). From this information we can observe patients' physician choices and the time until the next visit. The database includes over 55 million unique patients over the 10+ years captured, and roughly 15-18 million patients in a given year.

Figure 1: Cities in the sample, review growth, and care-gap.




**Notes**: The left-hand figure compares the growth in the number of reviews per 100,000 inhabitants in treated and untreated (control) cities. The central figure shows the location of the treated and untreated cities in our sample; the circle's size reflects the number of retirements detected. The boxplot on the right compares the distribution of the time between patient visits for nonretiring physicians (left) to retiring physicians before (middle) and after retirement (right).

For our analysis we focused on specializations where patients can easily delay a visit to the new doctor. In other words we picked specializations that involve an element of trust, or have a strong element of routine visits and preventive care, so that they are often considered 'non-urgent,' thus creating a more imminent risk of a care gap when the relationship is disrupted. Our most preferred discipline would have been dentist visits, but – because they are not available in our data – we focused on cardiology, dermatology, infectious disease, gastroenterology, and psychiatry.

We aggregate patient claims to daily visits, using data from October 2007 to March 2010, and from October 2015 until March 2018. We focus our analysis to specialties that are considered to benefit from regular visits and preventative care and that are frequently considered 'non-urgent,' because we suspect a more imminent risk of a care gap when the relationship is disrupted in such a discipline. We chose cardiology, dermatology, infectious disease, gastroenterology, and psychiatry, which gave us a sizeable but manageable amount of data to work with.[4] We also use the claims data to infer when a physician retires and identify retirements that occur in the twelve central months of both data windows (2008Q2 to 2009Q1 vs. 2016Q2 to 2017Q1).

We retain only cities with observations before and after treatment, which leaves us with 16 treated and 18 untreated cities. The location of the cities in this data set can be seen on the right-hand side of Figure 1, where blue bullets indicate cities in the treated group, red bullets indicate cities in the control group. The size of the bullet indicates the number of physician retirements we observe in each city.

## *Descriptive Statistics*

The unit of observation in our main dataset is a patient of a retiring physician, either at the time of the first visit to a new physician, or the end of our observation period. Table 1 shows the main variables in our study. Panel A gives information about the size of our treatment group and sample period. About 44% of the patients that are affected by a retirement are observed in 2014-2017, which we define as our treatment period (post), and 45% of the patients in our sample live in the cities in the treatment group (treat). Our data cover more patients from untreated cities than treated cities, and 15% of patients are in the treat group in the post period (DiD).

**Patient characteristics.** In our data, 55% of the patients are female, and patients are on average 65 years old when their physician retires. Only 36% of the patients in our data are under 65, 27% are between 65 and 75, and 37% are 75 and over. Patients between 20 and 49 are the reference category in our regression analysis. Consistent with the age structure, most visits concern cardiology (53%) and dermatology (29%). The remaining visits concern infectious diseases and gastroenterology (15%) and psychiatry (3%).

**Time between visits and "care-gap" after retirements.** Our main outcome variable is the time that elapses between two visits. Only 40% of patients in our data follow up with a new physician within 15 months of their last visit to a retiring physician (fup15). We observe a large care-gap after a physician retires. Note that we chose 15 months to allow patients a little over a year to find a new doctor. Our main results are robust to using 3, 6, 9, or 12.[5]

The average number of days until the next visit after a retirement exceeds 275 days, whereas the average time between visits without a physician retirement is 137.75 days. Similarly, the average time between visits before a retirement is 125 days (see also Figure 1). Further details are available upon request.

**The impact of reviews:** descriptive evidence. In Table 2, we compare the relative frequency of a followup visit and the days between visits for patients in the four treatment groups (pre- vs. post and treatment vs. control group). Although follow-up visits within 15 months seem to be less likely over time, the time between visits increases in the control group but decreases in the treatment group. In Figure A1 in the (omitted) online Appendix A we analyzed the distribution of the time between visits for each group of patients, confirming that the mass of long time between visits decreases. [6]

[4] Our ideal candidate discipline would be dentist visits, but these are not in our data. Future work aims to expand the scope of our analysis to GPs and other suitable disciplines.
[5] The effect is strongest for 12 months. 0.075 vs. the reported 0.07. These results are shown in an online appendix (E ) which is not included for reasons of space, but is available upon request.
[6] Table A5 of our online Appendix A.4 (not included) shows a comprehensive analysis of intervisit times.

Table 1: Descriptive statistics.

| | Mean | Min. | Max. |
|---|---|---|---|
| Panel A: Treatment definition | | | |
| post | 0.44 | 0 | 1 |
| treat | 0.45 | 0 | 1 |
| DiD | 0.15 | 0 | 1 |
| Panel B: Patient, city and physician characteristics | | | |
| Female | 0.55 | 0 | 1 |
| Age (years) | 64.7 | 0 | 89 |
| Age below 20 years (*age_u20*) | 0.03 | 0 | 1 |
| Age 20 to 39 years (*age_2039*) | 0.09 | 0 | 1 |
| Age 40 to 49 years (*age_4049*) | 0.07 | 0 | 1 |
| Age 50 to 64 years (*age_5064*) | 0.17 | 0 | 1 |
| Age 65 to 74 years (*age_6574*) | 0.27 | 0 | 1 |
| Age 75+ years (*age_75*) | 0.37 | 0 | 1 |
| Cardiology (*card*) | 0.53 | 0 | 1 |
| Dermatology (*derm*) | 0.29 | 0 | 1 |
| Infectious Diseases & Gastroenterology (*infc_gast*) | 0.15 | 0 | 1 |
| Psychiatry (*psych*) | 0.03 | 0 | 1 |
| Panel C: Main outcome variables | | | |
| Follow-up visit in 15 month (*fupm15*) | 0.398 | 0 | 1 |
| Time between visits (*days*) | 275.32 | 3 | 1,401 |

**Notes**: The table shows the summary statistics of the main estimation sample used in Table B6. The number of observations is 27,113.

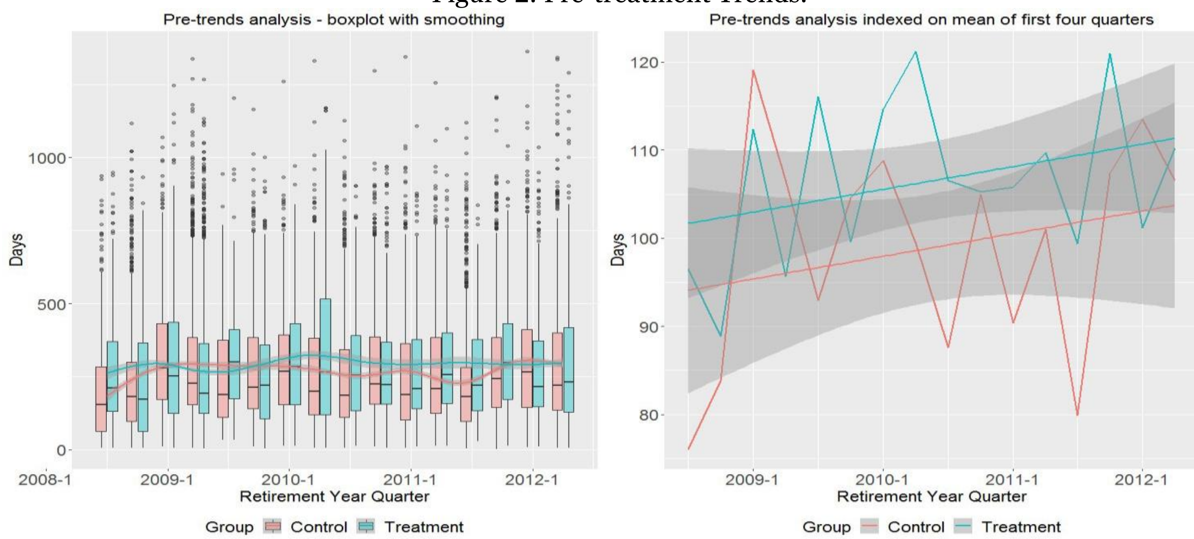
Table 2: Relative frequency of follow-up visit and time between visits.

| | All | Control | | Treatment | |
|---|---|---|---|---|---|
| | | Pre | Post | Pre | Post |
| Follow-up visit (15m) | 0.398 | 0.445 | 0.435 | 0.353 | 0.336 |
| Intervisit time (days) | 275.32 | 271.71 | 282.18 | 274.18 | 268.48 |

**Notes**: The table shows the mean frequency of a follow-up visit within 15 months as well as the mean time between visits in days for both treatment groups in each period.

**Analysis of the pre-treatment trends.** Our DiD estimation requires verification of the parallel trends assumption. The left side of Figure 2 compares the quarterly averages of the time between visits after a retirement in treated and untreated cities. We provide this comparison for the year used in our pretreatment period (2008Q2 to 2009Q1) and the subsequent three years. The right panel applies linear smoothing and shows confidence intervals. The average time between visits is slightly longer in treated cities, and grows slightly faster in cities that receive treatment by reviews, but the difference is small and insignificant. The visual impression is confirmed by a regression analyzing interactions of the treatment dummy with a time trend or retirement quarters (analyzed in Table A6 in the omitted online Appendix A). We found no significant difference in the development of the number of days between visits over time.

Figure 2: Pre-treatment Trends.




**Notes**: This figure analyses whether a patient's time between visits after their physician retires evolves similarly over time for patients in cities with and without reviews. The left figure shows quarterly box plots of the distribution of these patients' time between visits by treatment and control groups. The right figure shows the linear trend of the average time between visits per patient, indexed by the average time between visits in the pre-period.

# Results

## *Effect of reviews on physician visits*

In columns 1-2 of Table 3, we document our main regression result: Online reviews reduce the care gap that patients experience after their physician retires. We document this for two outcome variables: the probability of a follow-up visit within 15 months (col. 1) and the number of days until the next visit (conditional on observing a visit; col. 2). All specifications include specialty and city fixed effects, and control for a patient's age and gender. Before turning to our main finding, we note that patients in the postperiod are less likely to pursue continuous care within 15 months and they have their follow-up visits later across all specifications.

When reviews are available, patients in cities in our treatment group are on average 7 percentage points (p.p.) more likely to have a follow-up visit in the next 15 months (Col. 1) Column 2 of Table 3 summarizes the analysis of our second main outcome variable: days between visits. The result is consistent with the findings in column 1. The main effect in column 2 shows that the time between visits decreases by 30.7 days with reviews. We undertook additional analyses and robustness checks in the (omitted) online Appendix B. Our results Table 3: Effect of reviews on follow-up visit (LPM) and time between visits remain consistent if we (1) vary the pre-treatment time-window (Section B.1), (2) use different time fixed effects (Section B.2), or (3) use shorter cutoffs for the follow-up regression (Section B.3). We also analyze how the effects vary by age (Section B.4), specialty (Sections B.4 and B.5) and by population growth

(Section B.6). The effect is driven by cities with faster population growth, and by visits concerning cardiology and gastroenterology.

Finally, the effects are strongest for young patients (20-49) and for very old patients (75+) who might receive help from younger friends and relatives. reviews on follow-up visit (LPM) and time between visits.

Table 3: Main Results

| | *Base model* | | *OVB Controls* | | *IV (plumber rev.)* | | *Placebo* | |
|---|---|---|---|---|---|---|---|---|
| | Follow-up | DAYS | Follow-up | DAYS | Follow-up | DAYS | Follow-up | DAYS |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| DiD | 0.070*** | −30.685*** | 0.047 *** | −28.656 *** | 0.116*** | −55.459*** | −0.008** | −3.574 |
| | (0.014) | (9.131) | (0.019) | (13.063) | (0.020) | (12.685) | (0.004) | (2.244) |
| post | −0.057 *** | 24.943 *** | −0.056** | 32.816** | −0.060 *** | 25.739 *** | 0.029 *** | 2.722 * |
| | (0.009) | (5.376) | (0.023) | (16.376) | (0.010) | (5.844) | (0.003) | (1.637) |
| Education (%) | | | 0.001 | 2.081 | | | | |
| | | | (0.002) | (1.492) | | | | |
| Income p.c. | | | −2.285** | 817.864 | | | | |
| | | | (1.039) | (864.906) | | | | |
| Poverty (%) | | | −0.011*** | 6.411*** | | | | |
| | | | (0.003) | (2.280) | | | | |
| # Physicians | | | 0.00001* | −0.003 | | | | |
| | | | (0.00000) | (0.002) | | | | |
| Broadband (%) | | | 0.040 | −22.086 | | | | |
| | | | (0.062) | (48.084) | | | | |
| Ref. Category | 20-49 | 20-49 | 20-49 | 20-49 | 20-49 | 20-49 | 20-49 | 20-49 |
| Dem. Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Specialty FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| City FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 27,113 | 13,007 | 23,038 | 10,719 | 27,113 | 13,007 | 288,436 | 160,935 |
| IV 1st Stage | *(OLS coefficients)* | | | | | | | |
| #Plumber Revs. | | | | | 0.051*** | 0.059*** | | |
| | | | | | (0.0002) | (0.0004) | | |
| marg. adj. $R^2$ | | | | | 0.155 | 0.164 | | |

**Notes**: In this table, we analyze the effect of reviews on a patient's likelihood to follow up with a new physician (odd columns) and the number of days until this follow-up visit occurs (even columns) after their old physician retires. Columns 1 and 2 show the baseline result (reference category: patients age 20-49 years). In columns 3-4 we include controls that could be omitted confounding factors. Columns 5-6 show an instrumental variable regression using a city's number of online reviews for plumbers as the IV. Columns 7-8 show the results of a placebo test for patients of nonretiring physicians. White's (1980) heteroscedasticity-robust standard errors are shown in parentheses. *p<0.1, **p<0.05, ***p<0.01.

## *Identification and robustness checks*

Our main argument highlights that reviews are particularly useful when existing physician-patient relationships are disrupted. In the subsequent sections, we examine the possibility that our results are driven by factors that are correlated with the presence of reviews and also affect the time between visits.

First, to mitigate the concern that our results are due to the potential selection into treatment, we report instrumental variable estimations based on (1) Yelp reviews for other professions and (2) the share of young physicians in a city. Second, to rule out the possibility that our results are driven by an omitted variable, we (1) add further control variables and show that the results do not qualitatively change, and (2) conduct a placebo test in which we estimate the main regressions for patients of nonretiring physicians. This test allows us to rule out any other omitted factors that generally decrease the time between visits in a location, regardless of retirements.

### Instrumental variable specification

Reviews for physicians could emerge due to unobserved technology adoption by physicians that coincides with improved follow-up rates and shorter intervals between visits. We address this concern using reviews that are not health related as an instrument. Reviews in other categories are not driven by health-specific developments but are predictive of reviews for physicians. We collected the number of Yelp reviews for plumbers and hairdressers at the city level. Tables D7 and D8 in an (omitted) online Appendix D we verified that reviews for plumbers and hairdressers are highly predictive of reviews for physicians. The exclusion restriction is easily defended, because reviews for plumbers plausibly do not affect the time between physician visits directly.

Columns 5 and 6 in Table 3 report the results when instrumenting the availability of physician reviews with the number of plumber reviews in the same city. Our main results are confirmed: reviews have a strong positive effect on the probability of follow-up, and shorten the time between visits. The respective coefficients are larger in magnitude than with ordinary least squares (OLS). With reviews, the follow-up probability increases by 11.6 p.p., and the time between visits reduces by roughly 55 days.

In an online Appendix,[7] we also analyzed descriptive statistics for the instruments and compare the first- and second-stage results when reviews for hairdressers, rather than plumbers, are used as instruments. We also use reviews per capita, instead of the absolute number. The results consistently point to a higher likelihood of a follow-up visit and a shorter time between visits. In an alternative approach, we used the share of physicians under 45 as an instrument.[8] The effects remain consistent in sign, but the specification appears to suffer from a weak-IV problem.

### Controlling for potentially omitted factors

A second concern for our identification strategy is the potential omission of variables that are correlated with the availability of reviews and drive a reduction in the care gap. To mitigate this concern, we added additional relevant control variables in columns 3-4 of Table 3. Specifically, we control for broadband internet usage, demographics (income, poverty, education), and the availability of physicians. After controlling for these variables, we observe that reviews still increase the likelihood of a follow-up visit (+4.7 p.p.) and reduce the time between visits (by 28.7 days).

We explored more detailed descriptive statistics of this data, summary statistics, and regression results in Tables E1 and E2 of the (omitted) online Appendix E. In Tables E3 and E4, we re-estimated our main specification from Table 3, but sequentially add potentially omitted control variables. As expected, we observe that the likelihood of a follow-up visit is negatively and significantly associated with income and poverty level, and internet adoption has a positive and significant effect, though its significance diminishes when all variables are controlled for jointly.

[7] The online appendices were omitted to save space but are available upon request.

[8] The online appendices were omitted to save space but are available upon request. These results can be found in Tables D1, D3 and D4 of Online Appendix D. The following results are in Tables D5-D8.

### Placebo test: Patients of nonretiring physicians

Finally, we use a placebo test to rule out confounding factors that generally affect the time between visits. Such confounding factors should also affect patients of nonretiring physicians. Columns 7-8 of Table 3 replicate the regressions in columns 1-2 for patients of nonretiring physicians.

Column 7 analyzes the probability of a follow-up visit within 15 months. In the placebo condition, this probability is estimated to decrease by 0.8% in treated cities. Column 8 analyzes the number of days until a patient's next visit to a physician with the same specialty. The coefficient is 3.57 days and is not statistically significant, despite the larger sample size and increased statistical power. We conclude that the background tendency of seeing physicians more frequently in treated cities is negligible.

In the (omitted) online Appendix F we verified the consistency and robustness of our results. First, (omitted) Appendix Tables F1 and F2 we analyzed the decomposition of the placebo effect by specialty, age, and city population growth. Second, we made sure that the placebo results are not driven by how we impose placebo retirement dates by using a patient's average time between visits over the full period of observation as dependent variable (in Appendix Table F3).

### Remaining limitations

These three robustness analyses together reduce various identification concerns. First, the IV approach rules out that selection into treatment (review generation) is driven by any physician-specific developments. Second, the additional controls account for other internet-driven factors (broadband) and important demographic developments. Third, the placebo test highlights that the pattern of a reduction in the care gap is not generally observed in treated cities, but only for patients whose relationship ended. We show the absence of a pre-trend and, by design, always control for time-invariant factors. Moreover, we always control for age-specific factors and show in a robustness check that the pattern is driven by dynamic places with higher population growth.

Any remaining confounding factors that drive the effects we observe would have to be (1) different from the ones we controlled for and (2) limited only to patients of retiring physicians, rather than all patients in a treated city. However, we cannot address one remaining source of bias. Adoption of online reviews is stronger in dynamic places with a younger population, and our effects are strongest in places where reviews are widely used and weakest where adoption is low. In other words, we estimate a (highly relevant) average treatment effect on the treated (ATT). However, this issue would disappear once review platforms are widely adopted everywhere. Until then, the concern merely confirms the need to encourage greater availability and adoption of online reviews of physicians by designing an engaging and comprehensive review platform.

## Conclusion

**Summary.** We show that online reviews help patients to find a new physician when they need one. We use physicians' retirements as a disruptive shock that requires patients to find a new physician. We employ a DiD strategy in which we compare cities with and without many online reviews in the years of 2008 and 2017. When online reviews are available, patients are 7 p.p. more likely to follow up with their care within 15 months. Furthermore, if a patient follows up with a new physician, the time until this visit is on average about one month (30.7 days) shorter than otherwise.

**Interpretation.** Our findings highlight the potential of online reviews to improve efficiency and welfare when patients need to find a new physician. Patients consider online reviews when searching for a physician and these reviews affect patients' behavior. Online reviews seem to "make it easier" to visit a physician, and to reduce the risk of a gap in care whenever physician-patient relationships are interrupted. This finding is of particular importance in the medical sector and for patients where continued preventative care is important, and it also has a distributional dimension which is more prevalent than in other domains of digital WOM: Societies or large health-care systems where information about good physicians is "implicit knowledge" might be able to harness online reviews to achieve more equity in health care by making access to information easier and more widespread.

**Societal and Managerial Implications.** Our findings have important implications for managers of healthcare systems and platform owners because they show that enhancing the flow of information about doctors can lower the hurdle for patients to seek advice from a doctor. This is particularly important in countries where reviews are less widely available than in the United States. Given their value to patients, health insurers, whether public or private, should pay close attention to the availability of online health-related reviews. Moreover, stakeholders should try to make the information on existing platforms more accessible to their clients, and consider providing carefully designed and incentive-compatible platforms that motivate patients to provide informative and accurate online reviews about physicians.

**Further research.** By showing that online reviews help to reduce the risk of interruptions to continuous care when physicians retire, we provide a first step to understanding the potential of online reviews to directly benefit patients. However, we believe that the same mechanism applies to many other life changes, especially for patients – for example, when they move to a new city (or even to a new neighborhood), need to see a new specialist, or change their insurance carrier. Further research should explore these other potential causes of interruptions in care. A second fruitful avenue for further research is studying patient-physician matches. We hypothesize that online reviews help patients to find better matches, and this could be tested by analyzing the number of physicians visited by patients before they commit to a particular one. Third, a long-term study could leverage the ongoing COVID pandemic to evaluate the risk of care gaps to patients' health. Such a study could determine the long-term health risks associated with a gap in care when treating cardiovascular disease or other serious chronic conditions.

## Acknowledgements

This work was supported by the grant "Review Platforms & Improved Health Care" (ICOL - 512660E) by the American College of Radiology, Reston VA. We thank participants of the seminar of the CBESS group and in particular Chris Forman and Michelangelo Rossi for helpful comments. The authors alone are responsible for the content.

## References

Amaral-Garcia, Sofia, Mattia Nardotto, Carol Propper, and Tommaso Valletti, "Mums Go Online: Is the Internet Changing the Demand for Healthcare?," Review of Economics and Statistics, forthcoming.

Ai C, Norton EC, "Interaction terms in logit and probit models." Economics Letters, 2003, 80(1):123-129.

Barach, Paul, Stacy D. Fisher, M. Jacob Adams, Gale R. Burstein, Patrick D. Brophy, Dennis Z. Kuo, and Steven E. Lipshultz, "Disruption of healthcare: Will the COVID pandemic worsen non-COVID outcomes and disease outbreaks?," dec 2020.

Bensnes, Simon and Ingrid Huitfeldt, "Rumor has it: How do patients respond to patient-generated physician ratings?," Journal of Health Economics, 2021, 76, 102415.

Billari, Francesco C, Osea Giuntella, and Luca Stella, "Does broadband Internet affect fertility?," Population studies, 2019, 73 (3), 297–316.

Cabral, Luis and Ali Hortacsu, "The dynamics of seller reputation: Evidence from eBay," The Journal of Industrial Economics, 2010, 58 (1), 54–78.

Chen, Yiwei and Stephanie Lee, "User-Generated Physician Ratings and Their Effects on Patients' Physician Choices: Evidence from Yelp," SSRN Discussion Paper, 2021.

Chevalier, Judith A and Dina Mayzlin, "The effect of word of mouth on sales: Online book reviews," Journal of Marketing Research, 2006, 43 (3), 345–354.

Choi, Angela Aerry, Daegon Cho, Dobin Yim, Jae Yun Moon, and Wonseok Oh, "When seeing helps believing: The interactive effects of previews and reviews on e-book purchases," Information Systems Research, dec 2019, 30 (4), 1164–1183.

Darby, Michael R and Edi Karni, "Free competition and the optimal amount of fraud," The Journal of Law and Economics, 1973, 16 (1), 67–88.

Duan, Wenjing, Bin Gu, and Andrew B Whinston, "Do online reviews matter?—An empirical investigation of panel data," Decision support systems, 2008, 45 (4), 1007–1016.

Duflo, Esther, "Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment," American economic review, 2001, 91 (4), 795–813.

Emmert, Martin and Florian Meier, "An analysis of online evaluations on a physician rating website:

evidence from a German public reporting instrument," Journal of medical Internet research, 2013, 15 (8), e157.

Forman, Chris, Anindya Ghose, and Batia Wiesenfeld, "Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets," Information systems research, 2008, 19 (3), 291–313.

Gao, Guodong Gordon, Jeffrey S McCullough, Ritu Agarwal, and Ashish K

Jha, "A changing landscape of physician quality reporting: analysis of patients' online ratings of their physicians over a 5-year period," Journal of medical Internet research, 2012, 14 (1), e38.

Greaves, Felix, Utz J Pape, Dominic King, Ara Darzi, Azeem Majeed, Robert M Wachter, and Christopher Millett, "Associations between webbased patient ratings and objective measures of hospital quality," Archives of internal medicine, 2012, 172 (5), 435–436.

Hackl, Franz, Michael Hummer, and Gerald J. Pruckner, "Old boys' network in general practitioners' referral behavior?," Journal of Health Economics, sep 2015, 43, 56–73.

Helmers, Christian, Pramila Krishnan, and Manasa Patnam, "Attention and saliency on the internet: Evidence from an online recommendation system," Journal of Economic Behavior & Organization, 2019, 161, 216–242.

Holliday, Alison M., Allen Kachalia, Gregg S. Meyer, and Thomas D. Sequist, "Physician and Patient Views on Public Physician Rating Websites: A Cross-Sectional Study," Journal of General Internal Medicine, jun 2017, 32 (6), 626–631.

Howard, Paul, Senior Fellow, and Yevgeniy Feyman, "Yelp for health using the wisdom of crowds to find high-quality hospitals," 2017.

Hu, Nan, Paul A Pavlou, and Jennifer Zhang, "Can online reviews reveal a product's true quality? Empirical findings and analytical modeling of online word-of-mouth communication," in "Proceedings of the 7th ACM conference on Electronic commerce" 2006, pp. 324–330.

Kaye, Aaron, "Digitizing Doctor Demand: The Impact of Online Reviews in the Market for Primary Care Services," in "9th Annual Conference of the American Society of Health Economists" ASHECON 2020.

Lam, Kenneth, Cameron G Arnold, Rachel D Savage, Nathan M Stall, Lynn Zhu, Wei Wu, Katrina Piggott, Susan E Bronskill, and Paula A Rochon,

"Does physician retirement affect patients? A systematic review," Journal of the American Geriatrics Society, 2020, 68 (3), 641–649.

Lu, Susan F and Huaxia Rui, "Can we trust online physician ratings? Evidence from cardiac surgeons in Florida," Management Science, 2018, 64 (6), 2557–2573.

Lu, Wei and Hong Wu, "How online reviews and services affect physician outpatient visits: content analysis of evidence from two online health care communities," JMIR medical informatics, 2019, 7 (4), e16185.

Luca, Michael, "Reviews, reputation, and revenue: The case of Yelp. com (No. 12-016)," Harvard Business School, 2011.

and Sonal Vats, "Digitizing doctor demand: The impact of online reviews on doctor choice," 2014.

Reimers, Imke C and Joel Waldfogel, "Digitization and pre-purchase information: the causal and welfare impacts of reviews and crowd ratings," American Economic Review, forthcoming.

Sabety, Adrienne H., Anupam B. Jena, and Michael L. Barnett, "Changes in Health Care Use and Outcomes after Turnover in Primary Care," JAMA Internal Medicine, feb 2021, 181 (2), 186–194.

Sahoo, Nachiketa, Chrysanthos Dellarocas, and Shuba Srinivasan, "The impact of online product reviews on product returns," Information Systems Research, sep 2018, 29 (3), 723–738.

Saifee, Danish H, Zhiqiang Zheng, Indranil R Bardhan, and Atanu Lahiri, "Are Online Reviews of Physicians Reliable Indicators of Clinical Outcomes? A Focus on Chronic Disease Management," Information Systems Research, 2020, 31 (4), 1282– 1300.

Samuelson, William and Richard Zeckhauser, "Status quo bias in decision making," Journal of risk and uncertainty, 1988, 1 (1), 7–59.

Shukla, A. D., Gao, G., & Agarwal, R. (2021). How digital word-of-mouth affects consumer decision making: Evidence from doctor appointment booking. Management Science, 67(3), 1546-1568.

Tu, Ha T and Johanna R Lauer, Word of mouth and physician referrals still drive health care provider choice, Center for Studying Health System Change Washington, DC, 2008.

White, Halbert. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." Econometrica, 1980, vol. 48, no. 4, pp. 817–38.

Xu, Yuqian, Mor Armony, and Anindya Ghose, "The Interplay Between Online Reviews and Physician Demand: An Empirical Investigation," Management Science, 2021.

Yin, Dezhi, Sabyasachi Mitra, and Han Zhang, "When do consumers value positive vs. negative reviews? An empirical investigation of confirmation bias in online word of mouth," Information Systems Research, mar 2016, 27 (1), 131–144.