The Instagram Effect: Is Social Media Influencing Visitation to Public Land?

Ashley Lowe Mackenzie  Post-Doctoral Fellow, Department of Natural Resources and Environmental Management, University of Hawai‘i, Mānoa; alowemac@hawaii.edu

Steven J. Dundas  Associate Professor, Department of Applied Economics and Coastal Oregon Marine Experiment Station, Oregon State University, Corvallis; steven.dundas@oregonstate.edu

Bo Zhao  Associate Professor, Department of Geography, University of Washington, Seattle; zhaobo@uw.edu

Abstract

Public lands in the United States have recently experienced significant increases in visitation. Journalists and park managers suggest Instagram as a reason for the increase. We explore this issue in the Oregon State Park system by combining visitation data with park-specific georeferenced content and engagement indicators from Instagram. Using several empirical specifications, we show suggestive evidence that Instagram is not likely correlated to increased visitation everywhere, but only in a few locations generating high user participation within the app. We find no contemporary effect and a positive association with cumulative Instagram engagement indicators on visits at this subset of parks.

Keywords: social media, visitation, public land, Instagram, geotags, recreation, online engagement

JEL Classifications: Q26, Q50, Z30
1. Introduction

Starting in the 2000s, there was observable stagnation in visitation to public lands in the United States (U.S.). Research explored how terrorist attacks (McIntosh and Wilmot 2011; Stevens, More, and Markowski-Lindsay 2014; Bergstrom, Stowers, and Shonkwiler 2020), the impacts of recessions (Poudyal, Paudel, and Tarrant 2013), or the effect of entrance fees (Factor 2007; Stevens, More, and Markowski-Lindsay 2014; Bergstrom, Stowers, and Shonkwiler 2020) impacted the overall level of visitation. More recently, visitation trends have increased and shifted back to pre-1990 trends, and U.S national parks, for example, have seen consecutive years of record-high visitation levels (Bergstrom, Stowers, and Shonkwiler 2020). This increase in visitation to public lands produces a tension for state and federal land managers responsible for both providing recreation opportunities and preserving wild places (e.g., Mansfield et al. 2008; Jakus et al. 2010; Dundas, von Haefen, and Mansfield 2018). Journalists often blame the growing crowds on public lands on the introduction of social media photo-sharing apps, most notably Instagram (Figure 1). Park managers attribute social media for an increase in ‘selfie traps,’ or locations of considerable natural beauty that attract large crowds of people using the popular self-portrait technique to post content on these platforms. To explore the potential interrelationship between the rise of social media and increased visitation to public lands, this paper empirically investigates if content and engagement on Instagram has played a role in the recent increase in visits to public land in the U.S. state of Oregon.

[Insert Figure 1 here]

We examine eighteen years (2002 - 2019) of visitation data to 44 Oregon State Parks paired with newly compiled data on park-specific georeferenced content (i.e., spatially identified) and the level of engagement (i.e., likes and comments) with that content on Instagram. Our goal is to
assess the correlation between Instagram features and the observed increase in visitation to Oregon State Parks since the early 2010s. We find that on average across all parks in our sample, Instagram is not likely a contributing factor to increased visitation. Notably, when we subset state parks into groups based on in-app activity, we find that parks with Instagram content that have high user participation are likely to see increases in visits relative to parks with less in-app user activity. To explore this result further, we parameterize Instagram content as the number of geotagged posts at each park per month and engagement as the cumulative set of influential posts per park (i.e., geotagged content in 90th and 95th percentile of engagement per year). Our preferred specification suggests there is not a significant contemporaneous effect associated with content uploaded to Instagram each month but there is a cumulative effect of influential posts on visitation. This latter result suggests influential content may increase visitation up to a 4.2 percent per month. Practically speaking, this percentage increase suggests such content may drive an additional 2,780 vehicles per month (~ 91 per day) to parks with high Instagram activity. Our findings show that specific parks with content that has high levels of Instagram user activity may see visitation increase attributable to within-app behavior. These parks also tend to have more scenery-based amenities than lower activity parks where we did not find an effect of Instagram on visitation. This suggests that any impact Instagram may have on visitation may be limited to areas with picturesque or iconic landscapes which Instagram users find desirable. Lastly, we test that spatially identified geotagged posts are the Instagram feature of interest for estimating potential visitation impacts. A model using the park with the highest user activity in our dataset (Smith Rock State Park) shows a significant positive correlation between geotagged content and visitation while suggesting no effects related to park-specific user-generated posts via organizational hashtags.³
The rise of social media presents another potential factor that may influence recreational visits to public lands (Miller et al. 2019; Ghermandi and Sinclair 2019; Wood et al. 2020) but linking within-app behavior from a social media platform empirically to visitation has not yet been attempted. Prior work focuses on using content from social media to aid in approximating visitation levels and understanding visitor use patterns (Wood et al. 2013; Sessions et al. 2016; Levin, Lechner, and Brown 2017; Tenkanen et al. 2017; Fisher et al. 2018; Walden-Schreiner, Leung, and Tateosian 2018; Walden-Schreiner et al. 2018; Wilkins, Wood, and Smith 2021) or assessing differences in data quality and usage among platforms (van Zanten et al. 2016; Manikonda, Meduri, and Kambhampati 2016; Levin, Lechner, and Brown 2017; Norman and Pickering 2017; Tenkanen et al. 2017; Heikinheimo et al. 2018; Ghermandi and Sinclair 2019; Toivonen et al. 2019). This research is first to our knowledge to empirically link social media content and engagement as potential drivers of increased visitation to public lands. Importantly, our results suggest Instagram is not likely to impact visits to all public lands, just specific locations with iconic landscapes users find favorable and engage with and/or how the Instagram algorithm delivers such content to users. We also provide evidence that within-app engagement, measured as a cumulative total of influential geotagged content, is a potential feature of Instagram that could be associated with some increases in visitation in specific locations. This work contributes to our understanding of how online behavior may translate to changes in visitation to public lands. Our results provide evidence to help land managers understand and adapt to the emerging social media paradigm and improve stewardship of highly used natural resources. Although our findings are suggestive of a correlation between influential content with high user engagement on Instagram and visitation at specific locations, there may be other social, economic, or demographic trends which also contribute to increases in people visiting public
lands in the U.S. Our findings indicate that social media content may be one piece of the puzzle in understanding recent changes in outdoor recreation behavior.

This paper proceeds as follows. Section 2 discusses previous research using social media data in park visitation applications, the advantages and disadvantages of using data from various platforms and a brief history of Instagram in terms of content and user engagement. Section 3 describes our data and the process for collecting and quantifying data from Instagram. Section 4 describes our estimation strategy and outlines our empirical models. Section 5 discusses our results and the final section concludes with policy implications and avenues for future research.

2. Social Media and Park Visitation

To explore the connection between social media and park visitation, prior work has focused on developing visitation estimation methods and tracking visitor-use impacts. The number of geotagged photos posted on social media for a park has been shown to reliably correlate with the number of visitors to a park (Wood et al. 2013; Sessions et al. 2016; Levin, Lechner, and Brown 2017; Tenkanen et al. 2017; Fisher et al. 2018; Wilkens, Wood, and Smith 2021). These findings have important practical applications for reducing the cost of data gathering across managing agencies and extending visitor use estimates to wilderness areas where prior data collection was limited (Wood et al. 2013; Fisher et al. 2018; Hausmann et al 2018; Walden-Schreiner, Leung, and Tateosian 2018; Walden-Schreiner et al. 2018). Data from social media have also helped researchers examine willingness to pay for ecosystem services (Keeler et al. 2015; Ghermandi 2018; Sinclair, Ghermandi, and Sheela 2018) and explore spatial and temporal distribution of visitor-use patterns using location preferences (van Zanten et al. 2016; Heikinheimo et al. 2017, 2018; Levin, Lechner, and Brown 2017; Tenkanen et al. 2017; Walden-Schreiner, Leung, and Tateosian 2018; Walden-Schreiner et al. 2018; Barros Moya-Gómez, and Gutiérrez 2020).
Others have captured imagery within uploaded content and used it to map changes and impacts to natural systems and land use (Antoniou et al. 2016; Silva, Barbieri, and Thomer 2018; Toivonen et al. 2019).

There are, however, limitations when using social media data. It does not perfectly substitute for on-site counts of visitors (Wood et al. 2020) and often requires researchers to aggregate data across many years (Wilkins et al. 2021). The users of social media also are not likely representative of all public land visitors (Wilkins, Wood, and Smith 2021). The accuracy of data from social media used to determine visitation could also be associated with factors such as park popularity, the type of user base of a particular social media platform and how users utilize the platform (van Zanten et al. 2016; Manikonda, Meduri, and Kambhampati 2016; Levin, Lechner, and Brown 2017; Norman and Pickering 2017; Tenkanen et al. 2017). Wilkins, Wood and Smith (2021) provide a systematic review of the relationships between impacts of social media and park visitation. The meta-analysis noted only one paper (Hausmann et al. 2017) that uses in-app engagement behavior (i.e., the number of likes content receives) in their analysis. Twitter and Flickr tend to be the most frequently used data sources in research partly due to the accessibility of information through their respective Application Program Interfaces (APIs) (Heikinheimo et al. 2018; Gherandi and Sinclair 2019; Toivonen et al. 2019; Wilkins, Wood, and Smith 2021). For Instagram, which has been shown to outperform Twitter and Flickr as a proxy for visitation (Tenkanen et al. 2017), data collection is a bit more challenging due to recent changes in access to its APIs (Heikinheimo et al. 2018; Gherandi and Sinclair 2019; Toivonen et al. 2019).

Our research focuses on Instagram due to its large user base (1 billion active monthly users) and to investigate the direct claims made by journalists linking the app to visitation trends.
(Figure 1). The two key features of Instagram often mentioned as potential drivers of visits are geotagging (i.e., providing exact spatial coordinates for beautiful scenery) and the ability of content to “go viral” and be viewed by large sets of users. Instagram was exclusively available for iPhone mobile devices using iOS after its release in October 2010 and grew rapidly from there. It registered one million active users in the first two months, growing to fifty million once it was acquired by Facebook and opened to Android operating systems in April 2012. There are currently over 1 billion monthly active users and it is in the top five most popular websites on the internet as of 2022. The app contains features that organize users’ shared content of photos under geo-referenced locations (geotags), hashtags (January 2011), an Explore page (June 2012), Stories (August 2016) and Reels (August 2020).

When Instagram first launched, content was presented chronologically to its users. Public users could share content with others outside their network by tagging via geotag or hashtag. Each method placed every photo under a searchable URL page containing all other public account users’ content with the same tag. The Explore page, introduced in June 2012, began Instagram’s shift from chronological ordering to an individualized algorithmic curation of content based on the location and behavior of an individual user within the app. Instagram further explored algorithmic ordering in August 2012 with the PhotoMaps feature, which gave users the ability to explore geotagged locations spatially regardless of the time when content was posted to the app. Instagram moved away from PhotoMaps once all of its content became part of its new algorithmic-based searchable system on March 15, 2016. We do not include data for Stories or Reels here. Stories represent a different type of ephemeral content that disappears within 24 hours of posting to the app while Reels were introduced after the timeframe for our analyses.
Data from Instagram provide an opportunity to quantify both content and engagement of place-specific posts and explore how in-app behavior may correlate with outdoor recreation visitation trends. Our hypothesis is content (i.e., photos) uploaded in a month under geotags or hashtags captures individuals that have either recently visited, are sharing memories of the location, or are linking some experience to the location via the tag chosen. Geotags use the GPS coordinates embedded within an uploaded photo and provides a suggested geolocation designated within the app. Geotags can also be created by a user and used by others. Since geotags are creatable, some “points of interest” locations have multiple geotags but are linkable by the name used and the GPS coordinates associated with the location. Geotagged names are searchable under the tab Places within the app. Geotags and the precise locational information they provide to users to find a beautiful view, a hard-to-find trailhead, or other natural features is a primary reason many in the news media are placing blame on Instagram for the influx of visitation rather than other social media platforms.

Each geotagged image in Instagram also contains elements of engagement from the broader user base. The number of likes and comments each post receives reflects app users who enjoyed the content and provides a measurement of the impact it had on users of the platform. Content that has received significantly more engagement than normal posts may indicate viral content. Viral content is more actively shown or suggested to other Instagram users and increases the audience for that content. The ability for content to go viral has likely increased when Instagram moved away from chronological ordering toward its current algorithmic-based organizational structure. The mechanisms at play with viral, or influential, content are that a broad set of users are either exposed to previously unknown locations for outdoor recreation opportunities or are reminded of a location which could induce a potential future visit.
Regardless of the mechanism, land managers are responding to both content and engagement on Instagram. In 2018, the Jackson Hole (Wyoming) Travel & Tourism Board began a campaign directed at Instagram “Influencers”, individuals with thousands to millions of followers that endorse products, places and lifestyles through the content they share, to stop geotagging photographs with exact locations (Holson 2018). Similar campaigns have emerged endorsing tagging responsibly by swapping location-specific geotags with generic or more ambiguous locations (Wastradowski 2019; Merlan 2019). While some groups are attempting to deter visitors because of overcrowding, other tourism boards see opportunities to increase visitation by using Instagram “Influencers.” For example, Travel Oregon launched a “Seven Wonders of Oregon” campaign with both a promotional video and the use of Instagram influencers (McOmie 2014) and Utah had a Mighty 5 ad campaign to increase visitors to its national parks (Drugova, Kim, and Jakus 2021).

3. Data

The Oregon Parks and Recreation Department (OPRD) manages 254 properties and over 100,000 acres of public land in Oregon as of 2022. We first collected monthly visitor counts and details on available amenities for all properties managed by OPRD from January 2002 to August 2019. Each location has a varying number of amenities, including campsites, restrooms, beach access, fishing, viewpoints, surfing, swimming, kayaking, boat ramps, hiking trails, and playgrounds. Within these properties, 50 units hold the designation of state park, the most common classification in the OPRD system. The remaining units are designated as recreation areas and sites, scenic corridors and viewpoints, natural areas, or heritage sites. The focus of our analysis is on state parks as these units, on average, have more amenities, such as developed campgrounds, and may have more reliable visitor counts due to higher staffing levels as
compared to other unit types. It is also advantageous to concentrate our analysis on sites with a similar classification to provide a common pool of substitutable park locations (Weiler and Seidl 2004; Weiler 2006; Fredman, Friberg, and Emmelin 2007; McIntosh and Wilmot 2011; Poudyal, Paudel, and Tarrant 2013; Stevens, More, and Markowski-Lindsay 2014; Bergstrom, Stowers, and Shonkwiler 2020). Visitor counts at state parks are measured by standard traffic monitoring technology counting individual vehicles placed at entrances to a park and recorded by OPRD staff each month. OPRD estimates visitor counts by assuming the same number of passengers per vehicle (4) at all locations. Given this fact and our conversations with OPRD staff, we proceed with car counts as our dependent variable given adjusting to visitors using OPRD’s method is only a rough approximation and a simple scaling of the car count variable would not change our model outcomes or interpretation of those outcomes.

Although there are 50 state parks, three (3) sites are removed from our dataset at the outset. Two parks (Bates and Cottonwood Canyon) were new additions to the OPRD system during our study timeframe (Jan. 2002 to Aug. 2019) and only existed as state parks after the launch of Instagram. Shore Acres State Park is also removed due to large anomalies in the recorded monthly visitation counts relative to all other parks. This narrows our dataset initially to 47 state park locations in Oregon.

A cursory view of the initial visitation data shows that although many of these locations did experience an increase in visitors in the 2010s, visitation at some locations were flat or decreasing (Figure A.2 in online appendix). This suggests that while overall aggregated trends show increased visitation to public lands, there are likely to be park-specific factors, including social media content and engagement, which are important to consider when estimating visitation models. Park visitation models typically assume that visits are a function of travel cost,
population, seasons, and other economic factors. Higher gas prices increase travel costs and are likely to impact an individuals’ budget and recreational decisions. OPRD visitor surveys find a majority of visitors to Oregon State Parks live in Oregon (66 percent) and 65 percent of visitors travel more than 31 miles to reach a park (Bergerson 2019). Since our population of interest is traveling in the U.S. Pacific Northwest, we use regional conventional gasoline prices from the U.S West Coast (PADD5) to control for travel costs (Oh and Hammitt 2011; Poudyal, Paudel, and Tarrant 2013; Bergstrom, Stowers, and Shonkwiler 2020).\textsuperscript{11} Population estimates control for growth in the population of potential recreators (Poudyal, Paudel, and Tarrant 2013, Bergstrom, Stowers, and Shonkwiler 2020). These estimates for Oregon are collected from the Portland State University Population Research Center that provides annual population estimates in the years between the national census conducted by the U.S. Census Bureau. We use a linear interpolation to create monthly time steps for this variable. Economic factors commonly used as controls include median income, personal savings, unemployment rate, business cycle index and consumer confidence index (Oh and Hammitt 2011; Poudyal, Paudel, and Tarrant 2013; Stevens, More, and Markowski-Lindsay 2014; Bergstrom, Stowers, and Shonkwiler 2020). Our focus on a single state would likely benefit from state-specific economic indicators rather than national trends and it is important to have a temporal match between visitor counts (monthly) and the economic indicators collected (Poudyal, Paudel, and Tarrant 2013). Given prior work suggests consumer confidence indicators provide a good fit for predicting park visitation (Poudyal, Paudel, and Tarrant 2013) we use Oregon’s unemployment rate and Zillow’s Housing Value Index for all homes in the state. Both indicators are measured monthly and provide reasonable controls for local economic conditions and consumer confidence.\textsuperscript{12}
Outdoor recreation choices are also affected by weather (Dundas and von Haefen 2020), although weather is not often, if at all, included in visitation models. We obtain park-specific daily observations of temperature and precipitation in 4 km grid cells from Oregon State’s PRISM Climate Group (PRISM 2020). Since our visitation data is reported monthly and our weather observations are daily, we follow Dundas and von Haefen (2020) and create a monthly distribution of daily maximum temperature outcomes using a binning approach. Each bin contains a count of the number of days in each month separated into 10 temperature bins with intervals set at a 10-degree °F scale ranging from less than 30°F to greater than 90°F (e.g., >30°F, 30-39.9°F, 40-49.9°F, etc.). This binning approach allows us to identify a non-linear effect of temperature, or how an additional hot or cold day per month affects visitation.

Precipitation is controlled for by using the monthly average precipitation in inches.¹³

[Insert Figure 2 here]

**Instagram Data**

We collect data on all content posted to Instagram and geotagged to 44 Oregon State Park locations (Figure 2) from the launch of Instagram in October 2010 until August 2019.¹⁴ The geotagged images and metadata are collected from Instagram using a combination of a public domain Python scripts.¹⁵ The application gathers information on the tag of interest (geotag or hashtag) and then crawls the URL where Instagram stores all public posts under the tag of interest visible to public Instagram users.¹⁶ A JavaScript Object Notation (JSON) is produced with information on each post indexed by the date the content was posted, the number of likes and comments each photo received (i.e., engagement) and the photo caption. We do not collect private demographic information on the user from the Instagram bio.¹⁷ When a park had multiple geotag locations associated with it, Instagram posts for each separate geotag were collected and
linked to the specific park location. During the process of obtaining this data, two parks from the initial 47 were dropped (L.L. “Stub” Stewart and Alfred A. Loeb) due to Instagram server-side retrieval errors that prevented collection of geotagged images. A third, Prineville Reservoir State Park, was also dropped from the analysis, as it did not have a park-specific geotag. Of the 44 parks for the analysis, twenty (20) are located on the Oregon coast, eleven (11) are within the Willamette Valley or Columbia River Gorge and thirteen (13) are parks within or east of the Cascade Range.

The collected information allows us to quantify both Instagram content and engagement activity related to each state park. We measure content as the number of geotagged uploads for each park per month. We calculate the engagement variable as the cumulative amount of geotagged content (i.e., photos) within our dataset that received significantly more engagement each year than all other content. This cumulative effect captures how much visitation may change in response to a permanent increase in influential posts for a given location. We measure this variable as follows:

\[
\text{InfPost}_{pT} = \sum_{t=1}^{T} \text{Content}_{pt}^l
\]  

where \(\text{InfPost}_{pT}\) represents the cumulative sum of influential posts at park \(p\) in the current month \(T\) and \(\text{Content}_{pt}^l\) represents the content in each month \(t\) that exceeds a threshold \(l\) for high engagement. We measure high engagement as the observed number of \(\text{likes}\), a tally of the number of users within the app who have interacted with a geotagged photo (i.e., clicked the heart emoji in the post). We identify the threshold for high engagement content as those photos in the 90\(^{th}\) and 95\(^{th}\) percentiles of \(\text{likes}\) for all geotagged photos posted in the given year. This characterization identifies content with high engagement that is likely shown to a broad set of users under Instagram’s evolving content delivery algorithm. Selecting content as high
engagement per year for this variable is important, as it accounts for potential changes over time to both Instagram’s algorithm and general social media trends. A high volume of likes captures a variety of user motivations such as desirable aesthetic or attractive qualities of the content posted, which could be previously known or unknown to the user viewing the content. These variables representing both content and engagement allow us to test for both a contemporaneous and cumulative effect of geotagged Instagram posts on visitation.

Our final dataset is a panel comprising 44 parks across 212 months from January 2002 to August 2019. The panel is slightly unbalanced as there are a few parks that are closed in winter months and a couple instances where monthly visitation data is missing at a park due to random malfunctions of car counters collecting the information. Table 1 panels A and B provide aggregate summary statistics for our sample. Average monthly visitation to these state park units was approximately 31,700 over the nearly eighteen-year timeframe. Post-Instagram (October 2010 - August 2019) visitation was slightly higher (~33,400/month) than pre-Instagram (January 2002 - September 2010) visitation (~29,900/month). During the timeframe of the sample, Oregon’s population ranged from 3.47 million to 4.22 million and the spread of the monthly unemployment rate was 3.3 to 11.3 percent. Both average housing ($154,000 to $299,000) and gas prices ($1.23/gallon to $4.42/gallon) also had significant monthly variation. Summary statistics for visitation, amenities, and Instagram geotagged posts by park unit is provided in Tables A.1 and A.2 in the online Appendix. Table 1 panels C, D, E contain summary statistics for different grouping of parks based on Instagram user activity that are described in the next section.

[Insert Table 1 here]
4. Estimation Strategy

We start with a simple visitation model then proceed by adding a quasi-experimental research strategy and then adding content and engagement variables from Instagram. The first modeling specification examines the impact of Instagram’s launch on visitation using a binary indicator to denote the pre- and post-Instagram periods ($IG_{Launch}$):

$$\ln(Visits_{pt}) = \beta_0 + \beta_1 IG_{Launch} + \beta_2 T_{pt} + \beta_3 Prec_{pt} + \beta_4 \ln(Gas_{t})$$

$$+ \beta_5 X_t + \rho_t + \tau_{y(t)} + \gamma_p + \varepsilon_{pt},$$

(2)

where $Visits_{pt}$ is monthly ($t$) visits to park $p$, $T_{pt}$ is a non-linear function of daily maximum temperatures per month that allows the marginal effect of weather to vary across park locations, $Prec_{pt}$ is the average monthly precipitation (in), and $Gas_{t}$ is the average monthly conventional gasoline prices from the U.S West Coast. The vector $X_t$ contains monthly controls for log transformations of population level, unemployment rate and average housing prices. Our panel data structure allows for a fixed effects specification where we can control for unobservable park-specific characteristics $\gamma_p$, and $\rho_t$ and $\tau_{y(t)}$ represent month and year ($y$) fixed effects, respectively. The estimate of $\beta_1$ suggests an average impact of the launch of Instagram on visitation to Oregon State Parks.

To improve the specification of our model, the first consideration is the timing of the impact of Instagram on park visitation. Our simple specification in eq. (2) assumes that Instagram may have an instantaneous impact on visitation. However, it is not likely that the debut of a smartphone app is an event that would systematically change visitation patterns because Instagram would need time for its user base to grow and develop content. To help determine when Instagram would be likely to start influencing behavior, we plot total geotagged content and influential posts (90th percentile) for all parks over time. In Figure 3 panel A, the first
vertical line is October 2010 (Instagram’s launch) and the second vertical line is April 2012. It is at this second line where we begin to see an increase in app usage associated with Oregon State Parks. This date also coincides with Instagram reaching 50 million active users worldwide, being acquired by Facebook, and releasing the app to Android phone operating systems (instead of just Apple iPhone iOS). It also captures the timing right before Instagram moves from chronological to more algorithmic ordering with the introduction of the Explore page on June 2012. As robustness checks, we also test models on the timing of when Instagram may influence visits including June 2012 (Explore page introduced), August 2012 (PhotoMaps introduced), and March 2016 (switch to all algorithm-based content). Regardless of the date chosen, the impact of Instagram on visitation is likely better represented by these later dates than the debut of the platform in October 2010. We re-specify eq. (2) with $IG_{Use_t}$ as a new indicator variable to account for this timing:

$$
\ln(\text{Visits}_{pt}) = \beta_0 + \beta_1 IG_{Use_t} + \beta_2 T_{pt} + \beta_3 Prec_{pt} + \beta_3 \ln(Gas_t) \\
+ \beta_4 X_t + \rho_t + \tau_{y(t)} + \gamma_p + \epsilon_{pt}.
$$ (3)

We then consider that there may be significant variation among parks in how visitors and others generate and engage with content in Instagram, as prior work has found social media activity is associated with park popularity (Tenkanen et al. 2017). To visualize this potential variation, we plot the user activity (number of likes) for every geotagged image associated with each of the 44 parks in our dataset. This indicator is not necessarily illustrative of visitation but rather represents activity on the Instagram platform by those viewing the content. Figure 3 panel B displays this information, with parks ordered from lowest to highest level of user activity moving from left to right along the x-axis. The 30 parks on the left of the figure do have some associated user activity on Instagram, but the scale needed to capture the activity for the parks on
the right makes it appear to be near zero. Using the information from this graph, we partition the parks into two groups, high and low activity, with the hypothesis that there may be systematic differences in the effect of Instagram on visitation between these park types. Given the natural break after the fourth park starting from the right side of the graph, our preferred designation uses the top 4 parks, displayed as black dots, as the high activity group and the remaining 40 parks, displayed with gray dots, as the low activity group. Importantly, these four parks (Smith Rock, Oswald West, Ecola, and Silver Falls) do not correspond to the top four most visited parks. The most visited park is Valley of the Rogue which ranks near the bottom in Instagram user activity. For examples of geotagged content posted to Instagram from high and low activity parks, please see Figures A.3 and A.4 in the online appendix. We check the robustness of our high and low activity park definitions through iterations that 1) expand the high activity parks to the top 14 (the next natural break in Fig. 3 panel B); 2) drop the “middle 10” parks to compare the top 4 to the bottom 30, and 3) drop the top 4 and define the “middle 10” as high activity.¹⁹

Comparison of summary statistics for this grouping of parks by user activity is shown in Table 1 panels C, D, and E. We see differences between the groups in terms of Instagram geotagged uploads, with high activity parks averaging 388 posts and 50 influential (90th percentile) posts per month compared to 31 and 2 per month in low activity parks. Importantly, the number of available amenities at parks in these groups (panel E) is the same (11.8 on average), with low Instagram activity parks having slightly more activity-based amenities (e.g., hiking, biking, kayaking) and high Instagram activity parks having more scenery-based amenities (e.g., viewpoints). This suggests that high activity parks may contain more “gramable” features.
While the experimental ideal would expose some parks to Instagram and not others to causally identify the impact on visitation, our grouping strategy has the potential to estimate the impact of Instagram on parks with high user activity relative to those with low activity. By combining our park grouping with an appropriate timing when Instagram content and engagement have the potential to impact visitation, we can estimate a difference-in-differences (DiD) specification of our model. A common assumption needed for consistent estimation of DiD parameters is that both groups being compared have parallel pre-trends. This assumption requires that, absent the use of Instagram, the difference in visitation attributable to unobservables between park groups would have remained constant. We assess the validity of this assumption in Figure 4, which plots residual monthly visitation for high and low activity parks before and after April 2012. The residuals arise from a regression of the natural log of monthly visits on all economic and weather controls and seasonal and region-by-year fixed effects that are then aggregated by park group and month. The resulting figure suggest parallel trends are a reasonable assumption and that there may be systematic differences in visitation between high and low activity parks post-April 2012.\textsuperscript{21}

We estimate the DiD specification as follows:

\[
\ln(\text{Visits}_{pt}) = \beta_0 + \beta_1 (\text{High}_p \ast I_{GUse_t}) + \beta_2 I_{GUse_t} + \beta_3 T_{pt} + \beta_4 \text{Prec}_{pt} + \beta_5 \ln(\text{Gas}_t) + \beta_6 X_t + \rho_t + \tau_{y(t)} + \gamma_p + \epsilon_{pt},
\]

where the interaction term of Instagram’s sustained use \((I_{GUse_t} = 1\) post-April 2012) and an indicator for a high activity park \((\text{High}_p)\) estimates the effect of Instagram on visitation trends at these specific locations. All other variables are the same as defined for eq. (2).
Next, we specify additional models that use information about the content and engagement within Instagram, rather than just the existence of the platform itself. First, we use park-specific Instagram data on content, measured as the sum of geotagged photos posted to the platform for each park per month ($Content_{pt}$). We include $IG_{Launch_{t}}$ rather than $IG_{Use_{t}}$ to differentiate the months where a park received zero geotagged posts to Instagram from the time period which had no geotagged posts because Instagram did not yet exist. We specify this model as follows:

$$\ln(Visits_{pt}) = \beta_0 + \beta_1 IG_{Launch_{t}} + \beta_2 Content_{pt} + \beta_3 T_{pt} + \beta_4 Prec_{pt} + \beta_5 \ln(Gas_{t}) + \beta_6 X_t + \rho_t + \tau_{y(t)} + \gamma_p + \epsilon_{pt},$$

(5)

Our next model uses our content variable in a manner analogous to our DiD specification (because a non-zero count of geotagged posts implicitly includes $IG_{Launch_{t}} = 1$) by differentiating content by high and low activity parks:

$$\ln(Visits_{pt}) = \beta_0 + \beta_1 IG_{Launch_{t}} + \beta_2 (Content_{pt} * High_p) + \beta_3 (Content_{pt} * Low_p) + \beta_4 T_{pt} + \beta_5 Prec_{pt} + \beta_6 \ln(Gas_{t}) + \beta_7 X_t + \rho_t + \tau_{y(t)} + \gamma_p + \epsilon_{pt},$$

(6)

where $Low_p$ is an indicator variable defining low activity parks. The next iteration of this model adds our engagement variable, measured as the cumulative count of posts with the number of likes in the 90th and 95th percentiles of all content (eq. 1). This is done to disentangle the contemporaneous effect of content volume from the possible cumulative influence of engagement with influential posts ($InfPost_{pt}$) that can remain visible on Instagram from many months after posting. The next model is specified as follows:

$$\ln(Visits_{pt}) = \beta_0 + \beta_1 IG_{Launch_{t}} + \beta_2 (Content_{pt} * High_p) + \beta_3 (Content_{pt} * Low_p) + \beta_4 (InfPost_{pt} * High_p) + \beta_5 (InfPost_{pt} * Low_p) + \beta_6 T_{pt} + \beta_7 Prec_{pt} + \beta_8 \ln(Gas_{t}) + \beta_9 X_t + \rho_t + \tau_{y(t)} + \gamma_p + \epsilon_{pt}. $$

(7)
Lastly, we use data from the park with the highest Instagram activity (Smith Rock State Park (SRSP)) to explore a more nuanced understanding the relationship between Instagram content and visitation. We collect content from SRSP related to hashtags to compare this type of information to geotagged content. As noted earlier, geotags provide specific geo-referenced locations while hashtags provide the user with the name of the location along with other descriptive information (but not necessarily geo-location information). We focus this final model on determining the impact of each type of tagging of content on visitations:

\[
\ln(V_{\text{sites}}) = \beta_0 + \beta_1 T_{\text{ags}} + \beta_2 T_t + \beta_3 P_{\text{rec}} + \beta_4 \ln(G_{\text{as}}) + \rho + \tau(y(t)) + \epsilon_t,
\]

where \(T_{\text{ags}}\) represents a monthly content count of photos uploaded under either hashtags or geotags and \(\rho\) is a month fixed effect. All models are estimated with robust standard errors.

5. Results

Table 2 columns (1) and (2) show coefficients from estimating eqs. (2) and (3) based on the launch of Instagram (October 2010) and when observable use of Instagram begins (April 2012; Figure 3 panel A), respectively. Both models find no significant effect associated with Instagram. In these models, other variables behave as expected. For example, higher gas prices are likely to reduce visitation and the temperature/recreation response function follows an inverted U shape, suggesting both high and low temperatures may decrease visitation (similar results to Dundas and von Haefen (2020)). See Table A.4 in the online Appendix for full model results.

The null result in col. (1) suggests that the simple debut of a smartphone app is not an event that would systematically change visitation patterns due to the time it takes for the app to be incorporated into a user’s life and influence their behavior. In col. (2), the result suggests that even after Instagram began seeing significant use, there was no significant effect on visitation on
average across all state parks. The hypothesis that any impact Instagram may have is likely to be heterogeneous across locations leads to the difference-in-difference specification in col. (3). The results from estimating eq. (4) find that high activity parks are experiencing significantly more visitation associated with Instagram than low activity parks.\textsuperscript{22} The coefficient estimate on $High_p \times IG_{Use_t}$ suggests high activity parks have experienced a 24.1 percent increase in visitation relative to low activity parks since April 2012.\textsuperscript{23}

This first set of results includes a simple indicator variable for a time when Instagram content specific to Oregon State Parks began increasing. Given the potential correlations between the timing of Instagram’s increasing content and user base with other similar trends occurring across social media platforms (i.e., Twitter, Facebook) and the general overall increase in visitation observed in the 2010s, we next incorporate content and engagement variables to attempt to better understand the potential visitation impacts from Instagram. The effect of the contemporaneous count of a geotagged uploads (i.e., content) for each park per month shows a positive association with visitation (Table 3 col. 1). The coefficient estimate suggests a marginally significant effect per geotagged image of 0.02 percent, which translates to a 1.3 percent increase in visitation each month associated with contemporaneous Instagram content.\textsuperscript{24}

Our results from the initial modeling suggest visitation impacts may be heterogeneous across parks (Table 2 cols. 3-4). To test this, we separate our content variable by high and low activity parks (Table 3 col. 2). This separation suggests the effect of geotagged content is only significant in high activity parks with a similar coefficient estimate. However, in this case a 0.02 percent increase per geotag translates to a total increase in visits to high activity parks of 7.8 percent. This finding supports results from Table 2 that Instagram content is not likely impacting
visitation to all parks but visits to specific parks which have iconic landscapes that may drive user activity.

Next, we add the cumulative effect of influential content with high numbers of likes (i.e., engagement) for our two sets of grouped parks to the model. Table 3 cols. 3 and 4 show results when we measure the cumulative effect of photos which received enough likes to reach the top 90th and 95th percentiles of the entire set of geotagged content in our sample. Results in cols. 5 and 6 estimate the same model but with influential content quantified by all content posted in each given year. Once again, we find that the only significant impacts are associated with content and influential posts from high activity parks. In cols. 3 and 4, the content variable remains marginally significant and suggest a contemporaneous increase in visits of 3.4 percent. This is less than half of the contemporaneous effect captured in the previous model that excludes the engagement variables (Table 3 col. 2). The cumulative effect of influential posts from high activity parks in both the 90th and 95th percentile is, on average, associated with a 3.5 percent increase in visitation per month from all current and past influential photos. Moving to our preferred specifications that measures influential posts relative to all content posted in that year (cols. 5 and 6), we find the contemporaneous effect of content in no longer significant and the cumulative effect of influential posts suggests an increase in visitation by 4 to 4.2 percent per month from all current and past influential photos. See Table A.5 in the online Appendix for full model results. The significance of the cumulative effect suggests the overall impact of continual online exposure of specific parks with iconic landscapes is associated with a small increase in visitation over the last decade. The results overall also highlight how not all public lands are created equal and many may not have qualities or attributes that create visible content and engagement on Instagram. In our study area, both park types have the same available amenities
to recreators on average. The parks where we can estimate a positive association between Instagram activity and visitation differ in having at least one more scenery-based amenity while having less available activity amenities compared to the low activity parks.

[Insert Table 3 here]

Lastly, we tested the differences between content tagged with locational information (geotags) versus organizational information (hashtags) in Smith Rock State Park (SRSP). Data on content tagged with a hashtag for SRSP (#smithrock) allows an opportunity to investigate another potential tagging feature of Instagram that may influence visitation. SRSP is not the most visited park but is the park with the most user activity and influential photos (see Fig. 3, panel B), and a highly photogenic location with a variety of recreational opportunities (i.e., hiking, camping, rock climbing). In our estimation of eq. (8), geotagged content significantly and positively correlates with overall visitation whereas hashtags do not (full results available in online Appendix Table A.6). These results suggest that the provision of locational information may be a potential mechanism driving increased visitation from social media posts. Geotagged photos provide specific information to potential visitor (e.g., location of a trailhead or beautiful vista) whereas hashtags represent many things for Instagram users, including sharing a memory or indicating a desire to go to a location by using a hashtag to categorize their photo. In other words, hashtags encompass abstract organizational patterns when compared to the specific information provision provided by georeferenced location tags.

**Robustness Checks**

To test the robustness of our main findings, we alter the definition of high and low activity parks and the timing of Instagram’s potential influence. First, we examine alternative definitions of our high activity parks. We move the high classification to the next clear visible break in Figure 3
panel B to include ten (10) more parks (Online Appendix Figure A.5 panel A). We then re-estimate all our models with this new definition of high activity. Under this definition, we find results from estimating eq. (4) suggesting a 16 percent increase in visitation to high engagement parks. This result is less in magnitude than our primary specification (24.1 percent) and estimated with less precision (Online Appendix Table A.7). All models that add content and engagement variables yield similar results to our main specifications with high activity defined as the top 4 parks (Online Appendix Table A.8). The question then becomes is the addition of those 10 parks simply attenuating the impact of the top 4 or does it suggest that Instagram may impact visits to more parks? To investigate, we next specify models where the “middle 10” are dropped from the analysis (Online Appendix Figure A.5 panel B) so the models compare the original top four to the bottom thirty (30) low activity parks. These results suggest a slightly higher 27.6 percent increase (table A.9) and a similar result for both content and engagement to our primary models (Online Appendix Table A.10). Lastly, if we drop the top four and treat the “middle 10” as high activity (Online Appendix Figure A.5 panel B), we find no significant effects in the DiD specification (Online Appendix Table A.11) or the models with content and engagement (Online Appendix Table A.12). The combination of these results suggests that it is likely that the potential effect of Instagram on visitation is limited to a few parks with very high activity within the app and supports our primary modeling specifications.

We also explore variations in timing of the potential impact of Instagram’s various features and algorithmic changes and how it may affect our main results presented in Table 2 (Online Appendix Table A.13). Important dates include October 2010 (Instagram launch), June 2012 (Explore page), August 2012 (PhotoMaps), and March 2016 (algorithm-based content). All timing specifications also show significant results for high activity parks only. These finding
support our preferred specification, although also suggesting the timing of potential impacts from Instagram may be difficult to pin down to a single point in time.

6. Discussion

This paper attempts to quantify if and how Instagram has played a part in the observable increase in visitation to state parks in Oregon in the last decade. While our initial models suggest no overall effect, we do find that the introduction of Instagram correlates with increases in visitation to certain Oregon State Parks that have high user activity within the app. We then present additional specifications that find suggestive evidence of a connection between georeferenced content and influential posts posted on Instagram and overall visitation. Georeferenced content suggests a 7.8 percent increase in visitation to high activity parks. Importantly, when we add cumulative influential posts to the model, we find the contemporaneous effect of content is no longer significant and the estimated effect of influential posts suggests an increase in visitation by 4 to 4.2 percent per month from all current and past influential photos. These findings suggest the impact was isolated to certain parks generating high user activity within Instagram and was mostly driven by the influential content at these locations receiving high user engagement within the app. The one observable difference between high and low activity parks suggests scenery-based amenities may play a role. The photogenic qualities of the high activity parks could be attracting Instagram users for their “gramable” iconic viewpoints and landscapes (Figure A.3 in the online Appendix).

There are several potential mechanisms that could explain the estimated effects found here. One avenue may be that a reduction in search and information costs (e.g., Stigler 1961) makes recreators aware of new beautiful locations or entices new recreators to start visiting public lands. Smartphone technologies have embedded GPS functions which enable the relatively
costless discovery of new recreation locations while social media has provided a platform for individuals to share their experiences, discoveries and the location information more freely among a wider audience. However, if reduced information costs were the sole mechanism, we may expect to see impacts at all locations rather than just a small subset. Another potential mechanism is the bandwagon effect (or herd behavior), whereas an individual's demand for a commodity is increased due to the fact that others are also consuming the same good (Leibenstein 1950). This effect has been demonstrated in other contexts in economics (e.g., Biddle 1991), tourism destination preferences (e.g., Liu, Wu, and Li 2019; Pan, Rasouli, and Timmermans 2021; Boto-García and Baños-Pino 2022) and political science (e.g., Barnfield 2020). In this case, social media users would see others at these picturesque landscapes and choose to go there in order to not miss out on the experience and/or to obtain a similar photo of their own. Although this research cannot determine the mechanism, it remains a viable area for future research.

Our findings suggest the anecdotal claims that Instagram is a factor in the recent increase in visitation to public lands may have some validity. However, the act of uploading georeferenced content is not solely responsible for the rapid increase in visitation as some campaigns have claimed. Our models found geotagged content at most parks had little to no significant effect. Instagram’s geotagging feature provides accessible, low-cost, park-specific information to potential visitors but it is not the only social media site providing such information. Websites and other social media platforms (e.g., Youtube, TikTok, AllTrails, Gaia, Twitter, etc.) also provide spatially explicit information for potential recreators. Recent popularity in alternative lifestyles such as #vanlife, a nomadic living situation often involving overnighting in public land, are other potential mechanisms which may have impacted visitation trends (Monroe 2017).
Overall, the increasing accessibility through information and sharing experiences in outdoor recreation can introduce a greater number of individuals to find appreciation for our public lands. Protecting these areas for the benefit of both current and future generations is an important role entrusted to public land managers. Tensions arise when visitors, knowingly or unknowingly, negatively impact these public spaces. Instagram influencers have been blamed for ignoring signs in protected sensitive habitats and not practicing Leave No Trace (LNT) principles in order to get an “gramable” image (Canon 2019). Some influencers have even experienced legal consequences from the National Park Service when they have shared evidence online of their responsibility for resource degradation (Schaffer 2015). Public shaming campaigns have emerged to place pressure on ending the use of geotagging (Merlan 2019). In Deschutes County, Oregon (home of SRSP) the Tag Responsibly, Keep Bend Beautiful campaign focuses on getting outdoor recreators to not reveal location-specific information on social media by using a generic location instead (Wastradowski 2019). Despite these goals, anti-geotagging campaigns have been accused of gatekeeping by suppressing low-cost spatial information to limit public land access to potential new recreators (Mullen 2020; Slepian 2021).

From a policy perspective, our results support the idea of investing in a social media presence for land managers. An online presence can act as an informational pathway to connect with recreational users through education and to potentially mitigate negative outcomes before they occur. Content focusing on best recreational practices, LNT principles, site-specific location updates on resource closures, safe practices for viewing wildlife or trail information may be helpful in combating misuse within the parks they manage. However, the impact is likely dependent on its ability to generate content that would engage users enough to reach a wider audience within social media platforms. Agencies’ awareness and engagement in online
communities could also help them prepare for visitation surges. Identifying abnormally high activity and engagement online (i.e., “going viral”) for content geotagged to their units might be indicative of current and future visitation trends to the area. Public land managers may want to consider commercial filming and still photography permits for media captured within their boundaries, particularly at locations with sensitive resources (e.g., alpine wildflower meadows). Currently this practice is being utilized by the National Park Service, which considers commercial filming to include all film, electronic, magnetic, digital, or other recording of a moving image by a person, business, or other entity for a market audience with the intent of generating income, which includes posting footage on social media sites such as Instagram, YouTube and TikTok (NPS 2022).

Another potential policy takeaway shows a path for social media companies to improve their public relations. Many of the most popular social media companies rely on advertising revenue to keep their platform free for their users. Data collected on the users of these platforms provide an opportunity for hyper-targeted advertising. The platforms in return provide a path for the user to become an influencer and monetize their account. A potential issue arises when influential or viral geotagged content may impact public land use. It generates a conflict between those generating the effect (social media platforms) and those managing the parks (land managers). Overcrowding can cause environmental degradation while exposure to underused parks could help balance use. Unfortunately, the pathway for public agencies and researchers to request online data from many of the leading social media companies on historical and current engagement generated under a public land geotag location is currently unaccommodating. In this paper, content and engagement data were obtained through a computationally intensive Python script that took many weeks to collect and is increasingly challenging given Instagram’s
API permissions. Providing a path to request historical and current de-identified information would be a public service for park managers and researchers examining land use and recreational concerns in public spaces. Other helpful actions for management agencies could include notification services when verified public user accounts post under a given publicly managed geotag location. Verified public user accounts, as well as influencers, tend to have large audiences which can increase the probability uploaded content will be seen by a wide audience and potentially impact future visitation levels.

This paper is an attempt at understanding the complex linkages between social media and visitation to public land. Future research using surveys is likely needed to solidify the link between the timing of learning or being reminded of a specific location and then actually taking a trip to better understand how social media may be inducing new individuals to participate in outdoor recreation activities. Linking individual trip choices, rather than aggregate visitation, with social media indicators would also open the door to a deeper understanding of this new outdoor recreation paradigm.

Acknowledgements
The authors thank two anonymous reviewers and the editor for their helpful comments, as well as Caleb Dickson at Oregon Parks and Recreation Department (OPRD) for data access and valuable feedback on data collection methods used by OPRD.
References


Ghermandi, Andrea. 2018. "Integrating social media analysis and revealed preference methods to value the recreation services of ecologically engineered wetlands." *Ecosystem Services* 31:351-357.


Manikonda, Lydia, Venkata Vamsikrishna Meduri, and Subbarao Kambhampati. 2016. "Tweeting the mind and instgramming the heart: Exploring differentiated content sharing on social media." *Tenth international AAAI conference on web and social media.*


30


Table 1: Descriptive Statistics for 44 Oregon State Park Locations

<table>
<thead>
<tr>
<th>Variables</th>
<th>(1) N</th>
<th>(2) mean</th>
<th>(3) sd</th>
<th>(4) min</th>
<th>(5) max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Visits Per Month</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre Instagram (Jan 2002 - Sep 2010)</td>
<td>4,342</td>
<td>29,862</td>
<td>33,459</td>
<td>30</td>
<td>305,112</td>
</tr>
<tr>
<td>Post Instagram (Oct 2010 - Aug 2019)</td>
<td>4,501</td>
<td>33,391</td>
<td>37,670</td>
<td>42</td>
<td>297,668</td>
</tr>
<tr>
<td><strong>Panel B: Full Data (Jan 2002 - Aug 2019)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visits per Month</td>
<td>8,843</td>
<td>31,658</td>
<td>35,706</td>
<td>30</td>
<td>305,112</td>
</tr>
<tr>
<td>Mean Precipitation (in)</td>
<td>8,843</td>
<td>0.149</td>
<td>0.158</td>
<td>0</td>
<td>1.285</td>
</tr>
<tr>
<td>Max Temperature (°F)</td>
<td>8,843</td>
<td>60.62</td>
<td>12.05</td>
<td>24.81</td>
<td>97.96</td>
</tr>
<tr>
<td>Unemployment Rate (%)</td>
<td>8,843</td>
<td>6.841</td>
<td>2.143</td>
<td>3.300</td>
<td>11.30</td>
</tr>
<tr>
<td>Average Housing Price ($)</td>
<td>8,843</td>
<td>221,769</td>
<td>38,889</td>
<td>154,082</td>
<td>299,160</td>
</tr>
<tr>
<td>Gas Prices ($)</td>
<td>8,843</td>
<td>2.99</td>
<td>0.742</td>
<td>1.23</td>
<td>4.42</td>
</tr>
<tr>
<td>Oregon Population</td>
<td>8,843</td>
<td>3,829,000</td>
<td>210,180</td>
<td>3,472,000</td>
<td>4,215,000</td>
</tr>
<tr>
<td><strong>Panel C: High Activity Parks - Post Instagram Launch (Oct 2010)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visits per Month</td>
<td>426</td>
<td>66,216</td>
<td>48,496</td>
<td>3,874</td>
<td>288,414</td>
</tr>
<tr>
<td>Geotagged Content Uploads</td>
<td>426</td>
<td>387.9</td>
<td>629</td>
<td>0</td>
<td>3,845</td>
</tr>
<tr>
<td>Influential Posts top 90th</td>
<td>426</td>
<td>50</td>
<td>100</td>
<td>0</td>
<td>573</td>
</tr>
<tr>
<td>Cumulative top 90th</td>
<td>426</td>
<td>1,237</td>
<td>2,689</td>
<td>0</td>
<td>13,516</td>
</tr>
<tr>
<td>Influential Posts top 95th</td>
<td>426</td>
<td>27</td>
<td>54</td>
<td>0</td>
<td>344</td>
</tr>
<tr>
<td>Cumulative top 95th</td>
<td>426</td>
<td>680</td>
<td>1,462</td>
<td>0</td>
<td>7,351</td>
</tr>
<tr>
<td><strong>Panel D: Low Activity Parks - Post Instagram Launch (Oct 2010)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visits per Month</td>
<td>4,075</td>
<td>29,960</td>
<td>34,606</td>
<td>42</td>
<td>297,668</td>
</tr>
<tr>
<td>Geotagged Content Uploads</td>
<td>4,075</td>
<td>31</td>
<td>78</td>
<td>0</td>
<td>876</td>
</tr>
<tr>
<td>Influential Posts top 90th</td>
<td>4,075</td>
<td>2</td>
<td>7</td>
<td>0</td>
<td>117</td>
</tr>
<tr>
<td>Cumulative top 90th</td>
<td>4,075</td>
<td>47</td>
<td>141</td>
<td>0</td>
<td>1,426</td>
</tr>
<tr>
<td>Influential Posts top 95th</td>
<td>4,075</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>71</td>
</tr>
<tr>
<td>Cumulative top 95th</td>
<td>4,075</td>
<td>21</td>
<td>71</td>
<td>0</td>
<td>698</td>
</tr>
<tr>
<td><strong>Panel E: Park Amenities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High activity</td>
<td>4</td>
<td>11.8</td>
<td>6.5</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>Low activity</td>
<td>40</td>
<td>11.8</td>
<td>7.5</td>
<td>1.6</td>
<td></td>
</tr>
</tbody>
</table>

Note: High activity parks include Smith Rock, Oswald West, Ecola and Silver Falls. ¹ Amenities include bathrooms, vault toilets, dump station, portable water, as well as scenery- & activity-based amenities. ² Activity amenities include camping, hiking, biking, kayaking, fishing, wind surfing, climbing, surfing, swimming, horses, playground, boat ramp, picnicking, cabin, yurts, yurts with dogs, exhibit information, tepees, amphitheater, disc golf & tours. ³ Scenery amenities are amenities suggesting scenic views and photogenic locations which are listed as viewpoints, beach access, wildlife & waterfalls. Park-specific summary statistics pre- and post-Instagram are shown in Tables A.1 and A.2 in the online Appendix.
Table 2: Results for State Park Visitation Models

<table>
<thead>
<tr>
<th>Log(Monthly Visits)</th>
<th>(1) Launch of Instagram</th>
<th>(2) Use of Instagram</th>
<th>(3) Difference-in-differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Launch of Instagram</td>
<td>-0.042</td>
<td>0.013</td>
<td>-0.007</td>
</tr>
<tr>
<td>(IG_Launch; Oct. 2010)</td>
<td>(0.038)</td>
<td>(0.040)</td>
<td>(0.039)</td>
</tr>
<tr>
<td>Sustained Use of Instagram</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(IG_Use; April 2012)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IG_Use *High Activity Parks</td>
<td></td>
<td></td>
<td>0.216***</td>
</tr>
<tr>
<td>Economic Controls</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Weather Controls</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Year FE</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Month FE</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>N</td>
<td>8,843</td>
<td>8,843</td>
<td>8,843</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.582</td>
<td>0.586</td>
<td>0.588</td>
</tr>
</tbody>
</table>

Note: Panel model include 44 parks in the Oregon State Park system. Col. (1): IG_Launch is equal to 1 for all observations after Instagram’s Launch in October 2010. Col (2): IG_Use is equal to 1 for all observations after April 2012 when Instagram images began being geotagged to Oregon State Parks. Col (3): DiD approach separates parks into high and low Instagram activity. There are 4 high activity parks: Smith Rock, Silver Falls, Oswald West, and Ecola. All specifications include economic controls: gas prices, housing prices, unemployment rate, and population levels. Weather controls include mean monthly precipitation and a non-linear function of daily weather in each month. The panel is unbalanced as there are a few parks that are closed in winter months and a few instances of missing monthly visitation data due to random malfunctions of car counters. Robust standard errors in parentheses. Full model results are displayed in Table A.4 in the online Appendix. * p < 0.1; ** p < 0.05; *** p < 0.01.
Table 3: Results for Visitation Models Using Instagram Content and Engagement

<table>
<thead>
<tr>
<th>Log(Monthly Visits)</th>
<th>(1) Content</th>
<th>(2) Content: High/Low (H/L)</th>
<th>(3) Content: H/L + Engagement (E) (90%)</th>
<th>(4) Content: H/L + E (95%)</th>
<th>(5) Content: H/L + E by Year (90%)</th>
<th>(6) Content: H/L + E by Year (95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IG_Launch</td>
<td>-0.046</td>
<td>-0.048</td>
<td>-0.049</td>
<td>-0.048</td>
<td>-0.048</td>
<td>-0.048</td>
</tr>
<tr>
<td></td>
<td>(0.038)</td>
<td>(0.038)</td>
<td>(0.038)</td>
<td>(0.038)</td>
<td>(0.038)</td>
<td>(0.038)</td>
</tr>
</tbody>
</table>

Content Variables

| # of Geotags        | 0.0002***   |                             |                                        |                            |                                  |                                  |
|                     | (6.26e-05)  |                             |                                        |                            |                                  |                                  |
| # of Geotags High   | 0.0002***   | 0.00009*                   | 0.00009*                              | 0.00008                    | 0.00008                          |                                  |
| Activity (HA) Parks | (5.40e-05)  | (4.81e-05)                  | (5.01e-05)                            | (4.67e-05)                 | (4.90e-05)                       |                                  |
| # of Geotags Low    | 0.0003      | 0.0005                      | 0.0005                                | 0.0004                     | 0.0005                           |                                  |
| Activity (LA) Parks | (0.0003)    | (0.0003)                    | (0.0003)                              | (0.0003)                   | (0.0003)                         |                                  |

Engagement Variables

| Cumulative # of 90th Percentile Posts (HA) | 0.00003*** |                             | 0.00003*** |                             |                                  |                                  |
|                                            | (6.33e-06) |                             | (7.33e-06) |                             |                                  |                                  |
| Cumulative # of 90th Percentile Posts (LA) | -0.0001    |                             | -0.00006   |                             |                                  |                                  |
|                                            | (0.0002)   |                             | (0.0002)   |                             |                                  |                                  |
| Cumulative # of 95th Percentile Posts (HA) | 0.00005*** |                             | 0.00006*** |                             |                                  |                                  |
|                                            | (1.06e-05) |                             | (1.23e-05) |                             |                                  |                                  |
| Cumulative # of 90th Percentile Posts (LA) | -0.0002    |                             | -0.0002    |                             |                                  |                                  |
|                                            | (0.0003)   |                             | (0.0004)   |                             |                                  |                                  |

| Economic Controls | Y           | Y                           | Y           | Y                           | Y                                 | Y                                 |
| Weather Controls  | Y           | Y                           | Y           | Y                           | Y                                 | Y                                 |
| Year FE           | Y           | Y                           | Y           | Y                           | Y                                 | Y                                 |
| Month FE          | Y           | Y                           | Y           | Y                           | Y                                 | Y                                 |

N = 8,843
R² = 0.587

Note: Panel model includes 44 parks in the Oregon State Park system. IG Launch is equal to 1 for all observations after Instagram’s Launch in October 2010. There are 4 high activity parks: Smith Rock, Silver Falls, Oswald West, and Ecola. Economic controls include gas prices, housing prices, unemployment rate, and population levels. Cumulative # of 90th & 95th posts (cols. 3 & 4) are designated by the likes received (135 & 209 likes) on entire set of geotag photos. Cumulative # of 90th & 95th by year (cols. 5 & 6) are designated by likes on set of geotags per year (online Appendix Table A.3). Weather controls include mean monthly precipitation and a non-linear function of daily weather in each month. The panel is unbalanced as there are a few parks that are closed in winter months and a few instances of missing monthly visitation data due to random malfunctions of car counters. Robust standard errors in parentheses. Full model results are displayed in Table A.5 in the online Appendix. * p < 0.1; ** p < 0.05; *** p < 0.01.
Figure 1: Journalists Connect Social Media to Increases in Visitation to Public Land

Note: This figure displays national and regional headlines about social media and growing visitation to public lands.
Figure 2: Map of Oregon State Parks

Note: The U.S. state of Oregon with the location of 44 units in the state park system that are included in our analyses. Cities are labeled and shaded in gray for reference.
Figure 3: Timing of Instagram Content and Engagement and Activity Levels by Park

Note: Panel (A) displays a plot of geotagged content and influential posts in the 90\textsuperscript{th} percentile over time. The first vertical line corresponds to the launch of Instagram in October 2010. The second vertical line corresponds to April 2012, when we begin to see an increase in geotagged uploads for Oregon State Parks. It also is the month where Instagram became available on Android phones, was acquired by Facebook for $1 billion, and reached 50 million monthly active users. Panel (B) plots the Instagram activity (# of likes) for all geotagged content associated with the 44 parks in our analyses. Gray markers are used for 40 parks with low activity relative to the 4 high activity parks (Silver Falls, Ecola, Oswald West, Smith Rock) with black markers. One photo with 1.4 million likes at Smith Rock was removed from this plot to maintain a y-axis scale that allows visual comparison across parks.
Figure 4: Trends in Visitation by Park Group Based on Instagram User Activity

Note: This figure compares the visitation trends of high Instagram activity parks to low activity parks. We first regress log monthly visitation on model covariates and park, quarter, and region by year fixed effects. Next, we separate the residuals by park type and plot them over time with linear trendlines.
Endnotes

1 At the state level, increases in visitation are quickly outpacing budget allocations to maintain recreational facilities and prevent environmental disturbances (Smith, Wilkins, and Leung 2019). Furthermore, many agencies are working with a significant backlog of deferred maintenance, including the National Park Service ($11.6 billion; NPS 2018) and the Oregon Parks and Recreation Department ($59 million; Mukumoto 2019). The concern of negative impacts from overuse, stagnating budgets and staffing levels coupled with aging infrastructure are major concerns facing such agencies in the coming decade.


3 A hashtag (#) is a user-generated classification mechanism where individuals place a pound sign in front of an unspaced phrase. The tag enables cross-referencing and searchable content. It allows the platforms algorithm to push the content to other users outside their network whom might be interested in the content. YouTube is a video sharing platform which also utilizes hashtags to expand viewership, as does the most recent fastest growing social media platform TikTok.

4 For example, Facebook, Twitter and Instagram are in the top five most popular sites on the internet currently but differ in how communities are built and what content is shared among its users. Facebook has the largest number of monthly users and community is built on mutual acceptance of friend requests to access one another’s content. Twitter does not require mutual acceptance if the account is public. The content of Twitter is predominately delivered as tweets, or 280 characters conveying thoughts or ideas, and was the first site that popularized the hashtag (# symbol), a searchable user generated cross-referencing tool. Facebook allows hashtags but searchable content is difficult given community is built on mutual friendship acceptance. Instagram also does not require mutual acceptance to see content if the account is public, has searchable user-generated tagging but differs from Twitter by aiming at primarily sharing media content with a caption. Readers are referred to Toivonen et al. (2019) for an in-depth breakdown of various social media sites and the general habits of the users.

5 “Going viral” refers to when an image, video, advertisement, etc., is circulated rapidly on the internet receiving more engagement from users within the app than normal. Engagement for Instagram are the likes and comments a post receives. There is no predetermined threshold of engagement that qualifies content as “going viral”.

6 Stories are short-lived content typically only visible on a user’s feed for 24 hours unless the user pins it to their profile. Reels are short video clips ranging from 15 seconds to a few minutes.

7 Profiles posting under a hashtag are organized under the public URL (www.instagram.com/explore/tags/<hashtag>). Geotag URLs are organized under (www.instagram.com/explore/location/<LocationID>).

8 Since the acquisition of Instagram by Facebook in 2012, geolocation of posts was enabled trough “Facebook places”, points-of-interests represent the centroid of a known location such as a shop
or a town. Exact locations for geotag photos were shown on PhotoMaps (August 2012-September 2016) which provide searchable geotags on highly personalized map.

Currently, the Leave No Trace Organization is considering an 8th principle regarding responsible geotagging.

Shore Acres hosts several seasonal events not related to outdoor recreation (e.g., holiday lights show) that cause large variation in month-to-month visitation relative to all other parks as shown in the online Appendix Figure A.1 panel A. There are also two state parks directly adjacent to Shore Acres (Sunset Bay and Cape Arago; Figure A.1, panel B) which remain in the dataset and capture recreational visits to this area of the Oregon Coast. Therefore, Shore Acres is removed from the analysis.

Gas prices are obtained from the U.S. Energy Information Administration [https://www.eia.gov/](https://www.eia.gov/). EIA also provides national inflated-adjusted average gasoline prices. Using these national prices instead of regional does not change our results.

Indicators such as median income and personal saving are estimated only quarterly for the state of Oregon. Our outcomes of interest do not change if quarterly income or savings are included in model specifications.

A similar binning approach was attempted for precipitation. However, model results suggested the simpler monthly average was a better predictor for the impact of precipitation on monthly visitation.

The 44 parks included are Beverly Beach, Bob Straub, Brian Booth, Bullards, Cape Arago, Cape Blanco, Cape Lookout, Carl G. Washburne Memorial, Cascadia, Catherine Creek, Collier Memorial, Cove Palisades, Ecola, Elijah Bristow, Fort Stevens, Guy W. Talbot, Harris Beach, Hat Rock, Hilgard Junction, Humbug Mountain, Illinois River Forks, Jessie M. Honeyman Memorial, Lake Owyhee, LaPine, Mayer, Milo McIver, Molalla River, Nehalem Bay, Port Oford Heads, Oswald West, Rooster Rock, Silver Falls, South Beach, Smith Rock, Starvation Creek, Sunset Bay, Tumalo, Umpqua, Valley of the Rogue, Viento, Wallowa Lake, Willamette Mission, William M. Tugman, and White River Falls.

Richard Arcega’s “Instagram Scraper” (no-longer publicly available), Instaloader, Scrape-Instagram-by-Location [https://github.com/timkiely/scrape-instagram-by-location](https://github.com/timkiely/scrape-instagram-by-location)

Recent changes made to Instagram’s API policy requires a user to be logged in to their account to view geotag post information. Until 2020, this information was viewable and accessible without logging in.

An Instagram bio is a 150-character description under your username on your Instagram profile page.

We also characterize influential content as the 90th and 95th percentiles of the entire set of geotagged photos, rather than per year (see Online Appendix Table A.3). For our entire set of
it only takes 135 or 209 \textit{likes} to be considered in the top 90\textsuperscript{th} and 95\textsuperscript{th} percentiles of photos, respectively.

19 The “middle 10” parks are those in gray where the dots begin to appear above the baseline in Figure 3 panel B. These parks include Cape Lookout, Fort Stevens, Rooster Rock, Nehalem Bay, White River Falls, Harris Beach, Cape Arago, Humbug Mountain, Beverley Beach, and Sunset Bay. See Figure A.5 in the online appendix for visualizations of these alternative groupings.

20 Urban Dictionary: Social Media platform-specific adjective relating to Instagram. Something (mostly pictures) is good/fancy/interesting enough to post on Instagram.

21 Parallel trends plots for alternative timing dates (October 2010 - Instagram Launch; June 2012 - \textit{Explore} page; August 2012 - \textit{PhotoMaps}; and March 2016 - algorithm-based content) are provided in the online appendix Figure A.6. These additional figures suggest no effect associated with Instagram’s launch in 2010 or algorithm change in 2016 and shows similar trends in June and August 2012 compared to April 2012. This suggests that an event timing in 2012 is likely a reasonable choice for this model specification.

22 A Hausman test determined a fixed effects model is preferred over a random-effects specification. Nonetheless, when we estimate a random effects model, the coefficient on the interaction of $High_p \times IGUse_t$ is nearly identical to the results in Table 2 col. (3).

23 Percentage effects reported include an adjustment to the coefficient to interpret a dummy variable in a semi-log equation (e.g., Halvorsen and Palmquist 1980).

24 Average geotagged uploads post-Instagram to all parks is 63 post per month.

25 See Table 1 Panel C for average geotagged uploads and influential posts used to calculate the average effect when combined with coefficient estimates from Table 3 cols. 3 and 4.

26 Hypertargeting refers to the ability to deliver advertising content to specific interest-based segments in a network.

27 Prior to our data collection in 2019, the authors applied and requested data from Facebook, now known as Meta, and never received a response. During the review process Meta has added data request methods such as CrowdTangle and Data for Good; however, the list of prioritized research topics is limited.
1. Is Instagram Ruining the Great Outdoors?
   Social media can expose tens of thousands of people to places in an instant. That’s a double-edged sword.

2. Instagram Crowds May Be Ruining Nature

   Like democracy and cute animals before it, The Enchantments mountain range is suffering from the misguided side-effects of social media — in particular, Instagram.

4. Crisis in our national parks: how tourists are loving nature to death
   As thrill seekers and Instagrammers swarm public lands, reporting from eight sites across America shows the scale of the threat.

5. Is Geotagging on Instagram Ruining Natural Wonders? Some Say Yes

6. What’s Being Done to Save Wild Spaces from Instagram
   As outdoor-recreation tourism booms, these places have been forced to find innovative (and sometimes desperate) ways of adapting to and curbing the steady stream of tourists each season.