1-29-2015

Recreational demand for clean water: evidence from geotagged photographs by visitors to lakes

Bonnie L. Keeler
University of Minnesota

Spencer A. Wood
Stanford Univeristy

Stephen Polasky
University of Minnesota

Catherine L. Kling
Iowa State University, ckling@iastate.edu

Christopher T. Filstrup
Iowa State University

Follow this and additional works at: http://lib.dr.iastate.edu/econ_las_pubs

See next page for additional authors

Part of the Agricultural and Resource Economics Commons, Growth and Development Commons, Health Economics Commons, Natural Resource Economics Commons, and the Public Economics Commons

The complete bibliographic information for this item can be found at http://lib.dr.iastate.edu/econ_las_pubs/557. For information on how to cite this item, please visit http://lib.dr.iastate.edu/howtocite.html.

Authors

Bonnie L. Keeler, Spencer A. Wood, Stephen Polasky, Catherine L. Kling, Christopher T. Filstrup, and John A. Downing

Recreational demand for clean water: Evidence from geotagged photographs by visitors to lakes

6 authors, including:

Bonnie Keeler
University of Minnesota Twin Cities
43 PUBLICATIONS **1,024** CITATIONS

SEE PROFILE

Spencer A Wood
University of Washington Seattle
29 PUBLICATIONS **875** CITATIONS

SEE PROFILE

Christopher Thomas Filstrup
University of Minnesota Duluth
30 PUBLICATIONS **247** CITATIONS

SEE PROFILE

John Downing
University of Minnesota Duluth
150 PUBLICATIONS **13,383** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   Launching the Center for Creative Conservation in Seattle View project

Project   CyHAB - detection of toxigenic freshwater cyanobacteria and their cyanotoxins View project

Recreational demand for clean water: evidence from geotagged photographs by visitors to lakes

Bonnie L Keeler[1*], Spencer A Wood[2,3], Stephen Polasky[4], Catherine Kling[5], Christopher T Filstrup[6], and John A Downing[6]

More than 41 000 water bodies are listed as impaired by the US Environmental Protection Agency under the Clean Water Act. Implementation and enforcement of regulations designed to address these impairments can be costly, raising questions about the value of the public benefits derived from improved surface water quality. Here, we assess the recreational value of changes in water quality using freely available geotagged photographs, taken by members of the public, as a proxy for recreational visits to lakes. We found that improved water clarity is associated with increased numbers of visits to lakes and that lake users were willing to incur greater costs to visit clearer lakes. Lake users were willing to travel 56 minutes farther (equivalent to US$22 in travel costs) for every one-meter increase in water clarity in Minnesota and Iowa lakes, when controlling for other lake attributes. Our approach demonstrates the potential for social-media data to inform social–ecological research, including assessment of the recreational benefits of improvements in water quality.

Lakes, rivers, and streams provide many benefits to the general public, but these services are not captured in markets and have proven difficult to quantify (Brauman *et al.* 2007; Keeler *et al.* 2012). This is problematic because information on the value of water resources is needed in many policy and regulatory contexts. For example, the US Environmental Protection Agency is charged with estimating the benefits and costs associated with major rules and regulations designed to safeguard aquatic habitats (Griffiths *et al.* 2012). Cost–benefit assessments for water-quality changes are also considered in the design of payment and incentive programs, as well as in spatial planning decisions related to investments in conservation or habitat restoration (Olmstead 2010; Griffiths *et al.* 2012). Lack of information about the value of water-quality benefits can complicate justifying major spending on improved water quality.

Despite high demand, estimates of the value of clean water are often not available at the relevant scale for proposed interventions, are time- and resource-intensive to obtain, and are difficult to link to empirical measurements of water quality (Iovanna and Griffiths 2006; Keeler *et al.*

2012). The most common non-market methods for estimating the value of water-quality improvements typically require time-consuming and costly surveys, either to assess respondents' stated willingness-to-pay for improved water quality (Carson and Mitchell 1993) or to gather information on past recreational behavior (sites visited and distances traveled to sites; see Feather *et al.* 1995; Phaneuf and Smith 2005; Egan *et al.* 2009). In the first case, a survey is administered to ask respondents how much they would be willing to pay for a given change in water quality (stated preference approach). In the second case, users are asked information about their past behavior, and values are assessed based on how much (in terms of time and lost wages) they have conceded to obtain a higher-quality resource or experience (revealed preference approach). By design, survey data are often site-specific, limited in temporal and spatial scale, and not easily applicable to other decision-making contexts (Freeman *et al.* 2014). Estimates of benefits are also not typically expressed in terms of changes in water quality that can be linked to pollutant loads or land-use change, making it difficult to compare the costs and benefits associated with additional restoration or protection measures (Wilson and Carpenter 1999; Keeler *et al.* 2012).

We investigated how recreational lake users respond to variations in lake water quality, using data from an online social-media source as an alternative to survey data. Specifically, we used geotagged photographs (images associated with spatial-coordinate metadata) uploaded to the photo-sharing website Flickr (www.flickr.com) to estimate the number of visits to different lakes. We counted the number of uploaded photos taken by individual users on unique days that fell within selected boundaries of each lake

[1]*The Natural Capital Project, Institute on the Environment, University of Minnesota, St Paul, MN* *(keeler@umn.edu); [2]The Natural Capital Project, Woods Institute for the Environment, Stanford University, Stanford, CA; [3]School for Environmental and Forest Sciences, University of Washington, Seattle, WA; [4]Department of Applied Economics and Department of Ecology, Evolution, and Behavior, University of Minnesota, St Paul, MN; [5]Center for Agriculture and Rural Development, Department of Economics, Iowa State University, Ames, IA; [6]Ecology, Evolution, and Organismal Biology Department, Iowa State University, Ames, IA*
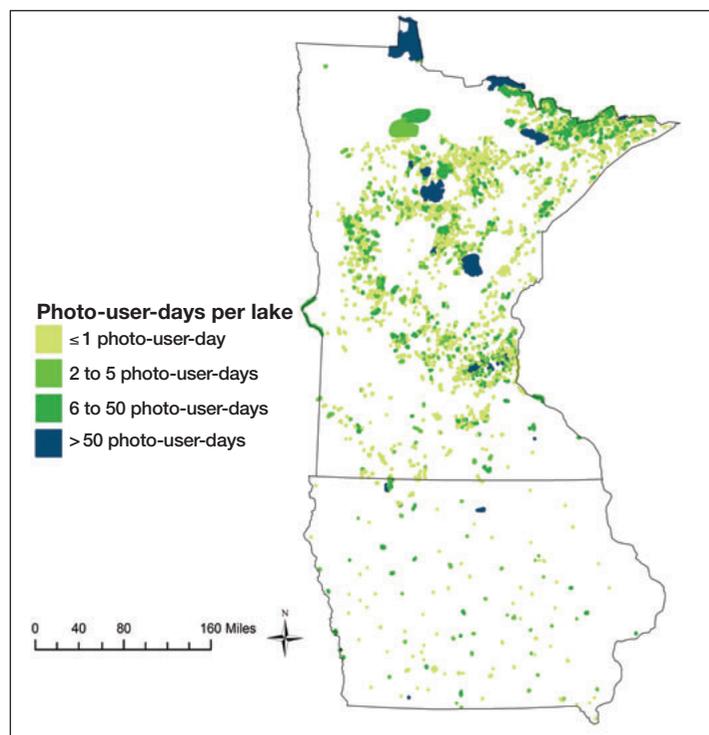
Figure 1. *Distribution of photo-visitations in Minnesota and Iowa lakes as measured by Flickr photographs. Photo-user-days per lake represent the sum of all unique daily lake and user combinations uploaded to Flickr between 2005 and 2012.*

and used this "photo-user-days" measure as a proxy for lake visitation. We coupled these data on lake visits with information provided by Flickr users to estimate the distances that visitors traveled from their self-reported hometown to each photographed lake (assuming transport by passenger vehicle along known road networks). This approach offers some advantages over traditional survey tools in that behavioral data can be collected over longer time periods, across broad spatial scales, and at minimal cost. After extracting and processing the geotagged photographs taken at each lake, we used multiple regression analysis to determine which lake attributes and other factors best explain patterns of lake visitation and travel costs. We then applied the regression model to a scenario involving improved water quality to evaluate observed changes in the numbers of unique visits to lakes and the value that visitors associate with improved water quality.

■ Methods

Lake attributes

Our study assessed the relationship between lake visitation and selected lake attributes for over 1000 lakes in the Midwestern US states of Minnesota and Iowa. We chose these states because of the availability of water-quality data (for lakes in both states) and survey data for Iowa lake users; lakes within this region also represent a gradient of water-quality conditions, from relatively undis-

turbed oligotrophic lakes to lower-quality eutrophic ones. Water clarity in the study region's lakes typically ranges from depths greater than 10 m to less than 0.5 m, encompassing most of the range of lake water clarity observed worldwide (Watson *et al.* 1992) and thereby making this a good study system for our purposes. In addition to water clarity, we also assembled data for various other explanatory variables, including lake-water chemistry, lake depth and size, near-lake human populations, lake amenities such as boat ramps (from which watercraft may enter and be retrieved from the water; also known as slipways or launches) and fishing piers, proximity to state parks and the Boundary Waters Canoe Area Wilderness, and the presence of aquatic invasive species (see WebPanel 1 and WebTable 1 for more details on water quality and other lake attribute data).

Lake visitation

Our study takes advantage of the increase in spatially explicit voluntarily supplied content available online. These shared data are increasing in volume each year and allow researchers to rapidly and inexpensively study user behavior and preferences over space and time (Wood *et al.* 2013). Here we use the photo-sharing website Flickr because it represents one of the largest available datasets of geotagged images and has an application program interface (API) that facilitates data extraction. To assess visitation at Minnesota and Iowa lakes, we queried Flickr for all geotagged images (in our study, these were photographs with latitude and longitude data) taken from January 2005 to December 2012 within the boundaries of over 3000 lakes in Minnesota and over 100 lakes in Iowa. In the corresponding geographic information system (GIS) analysis, we established a 30-m buffer zone around each lake – measured outward from the water's edge – to account for photographs taken along the shoreline. Our search returned a total of 41 852 unique geotagged photographs for Minnesota and Iowa lakes.

For each lake associated with geotagged photographs, we estimated the number of unique photo-user-days per lake (the count of unique combinations of users and lake destinations within a 24-hour period). For instance, if an individual took multiple photos at the same lake on the same day, that would equate to a single photo-user-day. These data were averaged across the 8-year period for which photos were downloaded (2005 to 2012) to derive an average annual number of photo-user-days per lake. In Minnesota, 1079 lakes were visited and photographed by Flickr users; in Iowa, 72 lakes returned geotagged images (Figure 1).

To obtain home location information needed for estimating travel routes, we also downloaded publicly available user-profile information associated with individuals who uploaded the photographs in our sample set. We deleted all personally identifiable information and

assigned each user a numerical identification code that was associated with the number and location (not the content) of their geo-tagged lake photos and their user-specified home location. About 40% of Flickr users who uploaded pictures of lakes provided their home location in their public profile. Flickr users who visited Minnesota lakes came from 47 US states and 36 other countries, with 66% of visitors reporting a home location from Minnesota. There were significantly fewer Flickr users who visited Iowa lakes; these visits originated from 20 US states and there were no international visitors.

Figure 2. *Average visitor numbers per year to Iowa lakes and Minnesota state parks, measured as photo-user-days, as compared with the number of trips per year estimated with traditional surveys. Each observation is a lake in Iowa or a state park in Minnesota. Dotted line is a 1:1 relationship between photo-visitation and surveyed visitation. Minnesota state park data are from Wood et al. (2013). Trendline equations are non-linear fits of the untransformed data plotted on log-log axes. Corresponding R^2 values for each regression are 0.65 (Iowa) and 0.70 (Minnesota).*

Are online photos a proxy for visitation?

Wood *et al.* (2013) compared surveyed data on the number of visits to various sites to visitation estimated by Flickr photo-user-days – using nine datasets consisting of 836 different natural and cultural attractions worldwide – and found this metric to be a good proxy for surveyed visitation rates. To evaluate the applicability of the photo-visitation method to lakes, we obtained data from a statewide survey of Iowa lake users conducted by Iowa State University (Evans *et al.* 2011). Survey information on lake visitation was reported over 5 years (2002–2005 and 2009) for 86 lakes in Iowa. We calculated average annual trips per lake over the 5 years for which data were available and plotted these values against photo-user-days for Iowa lakes estimated from Flickr (Figure 2). We found a significant positive relationship between 2005–2012 photo-user-days and surveyed visitation in Iowa lakes (R^2 = 0.65; Figure 2). This relationship is similar to one presented by Wood *et al.* (2013), between surveyed visits to Minnesota state parks and Flickr photo-user-days (R^2 = 0.70; Figure 2).

Distance analysis

Photograph data can also be used to estimate the distance traveled or time spent traveling from a user's stated home location to a lake destination. For the subset of Flickr users who provided information on the location of their home in their online Flickr profile, we mapped each user's hometown to spatial coordinates in a database of populated places. We considered only users with hometowns in 12 nearby Midwestern US states (CO, IA, IL, IN, KS, MI, MN, MO, ND, NE, SD, WI). Users residing in other states were assumed to have used air travel or other modes of transportation to visit Minnesota and Iowa lakes and were excluded from the distance analysis.

To estimate the distance traveled to visit a lake, we performed a distance analysis in ArcGIS. We used the ESRI ArcGIS Business Analyst Desktop (Redlands, CA) and
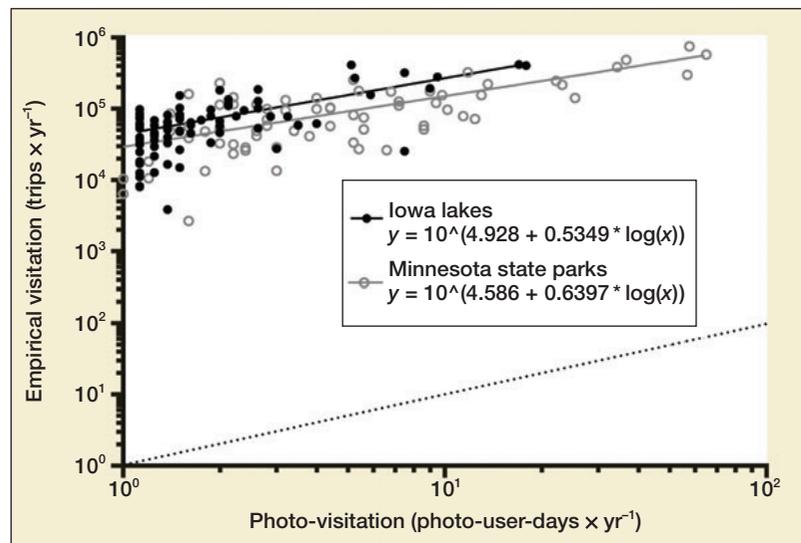
2012 NAVTEQ Street Data (Greenwood Village, CO) to estimate the travel time (accounting for posted speed limits for vehicular traffic along known routes) from each home location to visited lakes (Figure 3). For consecutive-day trips of less than 80 km, we assumed that users returned home between each lake visit. For trips greater than 80 km, we deleted routes where the same lake or different lakes were visited on consecutive days, assuming that the visitor stayed at or near the lake overnight and did not return home between lake visits. After removing consecutive day trips, our database contained 6438 trips to Minnesota and Iowa lakes from 12 neighboring states. For each lake visited by a Flickr user with a known home location, we estimated the average time spent traveling to visit that lake. Of the over 3000 lakes in the dataset, 946 were visited by users with home location information and were assigned average travel-time values.

Regression modeling

We used multiple regression models to identify the relationships between lake attributes, lake visitation, and travel time. For the lake visitation data, we first applied a logistic regression model to identify factors that predicted whether or not a lake was visited (where visitation was defined as lakes being the subject of at least one photograph during the study period; see WebTable 4). We then used linear regression on the subset of lakes that were visited to identify how a hypothetical change in water quality would affect the number of additional visits, assuming that all other lake attributes remained constant. We also applied a multiple linear regres-
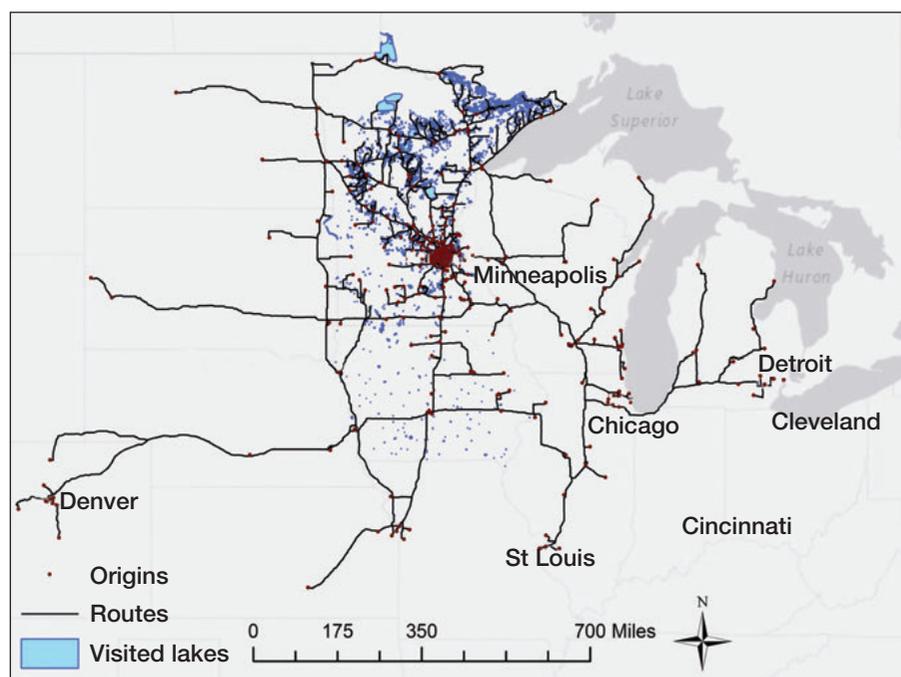
Figure 3. *Map of visited lakes, origins (user hometowns), and routes traveled by recreationists. We derived origins and lake destinations from Flickr user profiles and photographs, respectively, and estimated routes using ESRI ArcGIS Business Analyst.*

sion to the route data to estimate how changes in lake water quality would affect travel times to lakes (a proxy for travel cost). For all models we used backwards stepwise regression with Akaike information criterion (AIC) to select the best-fit model of the relationship between the response variables of lake visits and travel costs and the lake-specific explanatory variables (see WebTables 1–3 for parameter estimates, bivariate regressions, and pairwise correlations).

■ Results

Which factors predict lake visitation?

We found that lake size, water clarity, near-lake population, presence of a boat ramp, and state (Iowa or

Minnesota, represented by a dummy variable) were significant predictors of annual average per-lake visitation. This set of predictors was significant both in the logistic regression model predicting the probability of a lake receiving at least one visit (WebTable 4) and in the linear regression model estimating per-lake visitation for the visited lakes (Table 1). The relationship between visitation and lake clarity was positive, such that lakes with greater water clarity were associated with higher numbers of visits. As expected, larger lakes received more visits than smaller ones, and lakes with a boat ramp attracted more visitors than lakes without one. Lakes in Iowa also had more average visits than those in Minnesota, presumably because there are fewer lakes to visit. Notably, we observed a bimodal distribution of lake visitation whereby lakes in both densely and sparsely populated areas received high numbers of visits (WebFigure 1). To account for this distribution in our regression analysis, we centered the population variable in the multiple linear regression model (subtracted mean population from each lake population estimate) and included a squared term for population (WebPanel 2; Table 1).

Which factors affect the distance traveled to visit lakes?

Preferences for lake attributes can be not only inferred from the number of visits to each lake but also based on how far people are willing to travel to visit each lake (Parsons 2003; Egan *et al.* 2009). By using travel time as a proxy for the value individuals place on various lake attributes, we can infer the amount that individuals are willing to trade off to visit lakes with better water quality, with all other lake attributes being equal. In the non-market valuation literature, travel cost estimates are typically based on site counts of visitors or mailed surveys asking respondents to recall the number of their visits to various destinations. Here we estimated the distance traveled from user-specified home locations to different lakes, based on the locations of their geotagged photographs.

We used multiple linear regression to construct a model of travel time as a function of lake attributes and found a significant positive relationship between lake water clarity and travel time (Table 2),

Table 1. Multiple linear regression for lake attributes and photo-visitation per lake

	Estimate	Standard error (SE)	Effect test (Prob > t)
Intercept	0.143	0.022	<0.0001
Lake size (acres)	4.33×10^{-6}	5.41×10^{-7}	<0.0001
Lake clarity (m)	0.012	0.005	0.0190
Centered population	-3.35×10^{-8}	1.90×10^{-8}	0.0778
Centered population squared	4.39×10^{-14}	1.18×10^{-14}	0.0002
Boat launch (1 = yes)	0.025	0.007	0.0005
Iowa or Minnesota (1 = IA)	0.055	0.015	0.0003

Notes: The response variable is visitation in units log(photo-user-days × yr^{-1}). Each observation refers to a lake located in Minnesota or Iowa that was the subject of at least one recorded photograph during the study period (*n* = 1086 lakes). For explanation and justification of the centered population variable see WebPanel 2 and WebFigure 1, which plots visitation as a function of near-lake population. Effect tests refer to the significance level based on the probability of exceeding the *t* statistic (low values indicate the coefficient is significantly different than zero).

indicating that people are willing to spend more time to travel to clearer lakes. A best-fit model estimates that drivers were willing to spend an additional 56 minutes in round-trip travel time for each additional meter of depth visibility (ie water clarity). This translates to approximately US$22.26 per trip via ground transportation that a given user is willing to trade off for improved water quality, assuming one-third the average hourly wage rate and a transportation cost of US$0.30 per mile (WebTable 5; Parsons 2003). People are also willing to incur greater travel costs to visit larger lakes, lakes in the Boundary Waters Canoe Area Wilderness, and lakes with a boat ramp.

Would improvements to lake water quality increase the number of visits to lakes?

To estimate how a change in water quality would affect the number of visits to lakes, we applied the estimated relationship from the regression equation shown in Table 1 to lakes that received at least one visit over the study period. We estimated the change in photo-visitation between a scenario of baseline water clarity and a scenario where the water clarity of all lakes is increased by one meter. By assuming that the relationship between photo-visitation and surveyed visitation to Iowa lakes holds for all lakes in the sample region, we converted the model estimates of the change in photo-user-days into an estimate of annual trips per lake (based on correlation data presented in Figure 2). Using this approach, we calculated an average increase of 1389 annual trips for an average lake (1305 to 1481, lower and upper 95% mean confidence limits) per one-meter increase in water clarity, if all other variables remained constant. This result reflects the positive relationship between lake water quality and numbers of visitors.

Issues related to how well photo-takers represent all visitors, as well as potential biases in social-media use, contribute to uncertainty in these quantitative estimates. However, the qualitative relationship between visitor numbers and water quality is consistent with survey data and study expectations. That said, regional programs aimed at improving water quality might not generate increased visitation across all affected lakes because users may simply switch from visiting lower-quality lakes to visiting those with improved water quality. A net increase in the numbers of visitors could occur if visiting a clearer lake substituted for leisure time previously spent at a public pool or other attraction.

■ Discussion and conclusions

We used photo-visitation data to understand the effects of improved water quality on the number of visits and the distance traveled to lakes. We found that recreational lake users visit clear lakes more often than less-clear lakes and are willing to incur increased travel costs to visit lakes with better water quality. This conclusion is consistent with stated preference studies using contingent valuation surveys or choice experiments, which have found evidence for a positive relationship between water quality and willingness-to-pay (Carson and Mitchell 1993; Phaneuf 2002; Johnston *et al.* 2003; Banzhaf *et al.* 2006; Viscusi *et al.* 2008; Van Houtven *et al.* 2014). Far fewer studies have evaluated the benefits of improved lake water quality using revealed preference approaches based on the surveyed behavior of recreationists (eg Feather *et al.* 1995; Egan *et al.* 2009).

There are limitations to the use of social-media data to estimate recreational behavior. We recognize that Flickr users are not necessarily representative of all recreationists. We compared demographic characteristics of Flickr users worldwide (reported in Ignite Social Media 2012) with the demographics of Iowa Lakes Survey respondents (reported in Evans *et al.* 2011) and found that Flickr users were more likely to be female and of higher educational status than the Iowa lake users that responded to surveys. Reported income was comparable between the two groups. However, Flickr users were, on average, younger than the Iowa Lakes Survey respondents. At present, little is known about how the behavior and preferences of Flickr users differ from those of other lake users, including information about the lake-related activities that Flickr users participate in and how those activities differ from those of other lake users. Despite these limitations, our comparison of survey data from Iowa (Figure 2) and previous work across multiple sites worldwide (Wood *et al.* 2013) suggest that photo-visitation can be a reliable proxy for actual visitation.

We cannot fully account for the role of bias and representation in our model results due to our reliance on data from social media. Notably, however, similar issues must also be considered in "conventional" approaches to collecting data on preferences and behavior. For example, survey data can be subject to hypothetical bias, non-response and sample selection bias, inattentive or hasty responses to questions, recall errors, order effects, and

Table 2. Multiple linear regression for lake attributes and time spent traveling to each lake

	Estimate	*Standard error (SE)*	*Effect test (Prob > t)*
Intercept	207.50	7.02	<0.0001
Lake size (acres)	0.001	0.00007	<0.0001
Lake clarity (m)	28.07	1.44	<0.0001
Boat launch (1 = yes)	6.51	2.07	0.0017
Boundary waters (1 = yes)	137.83	5.08	<0.0001
Iowa or Minnesota (1 = IA)	43.29	3.77	<0.0001

Notes: The response variable is time spent traveling to each lake (one-way). Observations represent all unique combinations of users and lake destinations on non-consecutive days (*n* = 6438 routes). Effect tests refer to the significance level based on the probability of exceeding the *t* statistic (low values indicate the coefficient is significantly different than zero).

framing effects, all of which can contribute to inaccurate reporting of magnitudes or frequencies (Hanemann 1994; Kling *et al.* 2012). We hope that this paper stimulates further investigations into sources of bias, representation, and interpretation in social-media-related data, just as progress has been made over the past several decades in understanding the strengths and limitations of other revealed and stated preference methodologies (Hanemann 1994; Kling *et al.* 2012).

This paper contributes information on the benefits of improved water quality, which is needed to inform regulatory cost–benefit assessments: particularly where there is uncertainty surrounding the value generated by proposed investments in improving surface water quality (Griffiths *et al.* 2012). Similar methods could be used to evaluate the benefits of other changes in environmental quality, especially where resource and time constraints prevent the use of survey data or where data are required across broad temporal or spatial scales. The next steps for adapting this approach could include scaling up the analysis to link photo-visitation estimates to regional and national databases on lake water quality. These data can be overlaid with data on known impairments to evaluate the return on investments intended to improve surface water quality, from a single lake up to state or regional scales. In the future, we believe that use of social-media-derived geotagged data on recreational demand will help to inform spatial planning and resource investments, as well as to improve our understanding of the behavior and preferences of other users of surface waters.

■ References

Banzhaf HS, Burtraw D, Evans D, and Krupnick A. 2006. Valuation of natural resource improvements in the Adirondacks. *Land Econ* **82**: 445–64.

Brauman KA, Daily GC, Duarte TKE, and Mooney HA. 2007. The nature and value of ecosystem services: an overview highlighting hydrologic services. *Annu Rev Env Resour* **32**: 67–98.

Carson RT and Mitchell RC. 1993. The value of clean water: the public's willingness to pay for boatable, fishable, and swimmable water quality. *Water Resour Res* **29**: 245–54.

Egan KJ, Herriges JA, Kling CL, and Downing JA. 2009. Valuing water quality as a function of water quality measures. *Am J Agr Econ* **91**: 106–23.

Evans KS, Herriges JA, and Kling CL. 2011. Iowa Lakes Survey 2009. www.card.iastate.edu/lakes/iowa-lakes-report-2009.pdf. Viewed 25 Nov 2013.

Feather P, Hellerstein D, and Tomasi T. 1995. A discrete-count model of recreational demand. *J Environ Econ Manag* **29**: 214–27.

Freeman III AM, Herriges JA, and Kling CL. 2014. The measurement of environmental and resource values: theory and methods (3rd edn). Washington, DC: RFF Press.

Griffiths C, Klemick H, Massey M, *et al.* 2012. Valuation of surface water quality improvements. *Rev Environ Econ Pol* **6**: 130–46.

Hanemann WM. 1994. Valuing the environment through contingent valuation. *J Econ Perspect* **8**: 19–43.

Ignite Social Media. 2012. Social network analysis report. www.ignitesocialmedia.com/social-media-stats/2012-social-network-analysis-report/#Flickr. Viewed 24 Apr 2014.

Iovanna R and Griffiths C. 2006. Clean water, ecological benefits, and benefits transfer: a work in progress at the US EPA. *Ecol Econ* **60**: 473–82.

Johnston R, Besedin E, and Wardwell R. 2003. Modeling relationships between use and nonuse values for surface water quality: a meta-analysis. *Water Resour Res* **39**: 1363.

Keeler BL, Polasky S, Brauman KA, *et al.* 2012. Linking water quality and well-being for improved assessment and valuation of ecosystem services. *P Natl Acad Sci USA* **109**: 18619–24.

Kling C, Phaneuf D, and Zhao J. 2012. From Exxon to BP: has some number become better than no number? *J Econ Perspect* **26**: 3–26.

Olmstead SM. 2010. The economics of water quality. *Rev Environ Econ Pol* **4**: 44–62.

Parsons GR. 2003. The travel cost model. In: Champ PA, Boyle KJ, and Brown TC (Eds). A primer on nonmarket valuation. Dordrecht, The Netherlands: Kluwer.

Phaneuf DJ and Smith VK. 2005. Recreation demand models. *Handbook Environ Econ* **2**: 671–761.

Phaneuf DJ. 2002. A random utility model for total maximum daily loads: estimating the benefits of watershed-based ambient water quality improvements. *Water Resour Res* **38**: 1–11.

Van Houtven G, Mansfield C, Phaneuf DJ, *et al.* 2014. Combining expert elicitation and stated preference methods to value ecosystem services from improved lake water quality. *Ecol Econ* **99**: 40–52.

Viscusi WK, Huber J, and Bell J. 2008. The economic value of water quality. *Environ Resour Econ* **41**: 169–87.

Watson S, McCauley E, and Downing JA. 1992. Sigmoid relationships between phosphorus, algal biomass, and algal community structure. *Can J Fish Aquat Sci* **49**: 2605–10.

Wilson MA and Carpenter SR. 1999. Economic valuation of freshwater ecosystem services in the United States: 1971–1997. *Ecol Appl* **9**: 772–83.

Wood SA, Guerry AD, Silver JM, and Lacayo M. 2013. Using social media to quantify nature-based tourism and recreation. *Scientific Reports* **3**: 2976.

WebPanel 1. Lake water-quality and lake attribute data
We assembled lake water-quality data and data on other lake attributes that may be important in explaining variation in lake visitation. WebTable 1 summarizes these data for Minnesota and Iowa lakes. Water clarity, as measured by Secchi depth, was the most commonly reported water-quality attribute across both states. Secchi depth is a standard metric for measuring water transparency and refers to the maximum depth below the lake surface that a researcher can visually see a submerged black and white disk of standardized size and color. For Minnesota lakes missing water-clarity data from Secchi-depth measurements, we substituted remotely sensed water-clarity data from the Minnesota Lake Browser (lakes.gis.umn.edu). These lake clarity estimates are based on the relationship between satellite-derived spectral-radiometric responses (color bands) and empirical measurements of lake clarity, and are available for the years 2005 and 2008 (Kloiber *et al*. 2002; Olmanson *et al*. 2011).

In addition to lake water quality and other lake attributes, we assigned a population proximity estimate to each lake as an indicator of the number of potential lake visitors near each lake. To do this, we created a raster layer representing the population density (people per square mile) from government census data for Iowa, Minnesota, neighboring states, and the provinces of Manitoba and Ontario, Canada, that border northern Minnesota. To estimate the population proximity variable for each lake, we created an 80-km buffer around each lake and summed the number of people living in the buffer zone. A distance of 80 km was selected because it corresponds with the threshold used by the US National Tourism Resources Review Commission to define a tourist (ie tourists are individuals traveling more than 80 km [50 miles] from their hometown).

WebTable 1. Lake attributes for Minnesota and Iowa lakes

	Minnesota	Iowa
	Mean (range)	Mean (range)
Total number of lakes	2120	128
Lake size (acres)	1043 (2–302 822)	350 (7–5300)
Lake clarity (m)	2.7 (0.1–10.2)	1.1 (0.2–5.5)
Lake depth (m)	12.8 (0.3–65.5)	6.5 (1.2–42.1)
Chlorophyll (µg/L)	19.6 (0.5–257.5)	47.9 (2.8–187.7)
Total phosphorus (µg/L)	54.3 (1.0–492.3)	100.8 (11.7–384.8)
Boat launch (1 = yes)	0.57 (0–1)	0.27 (0–1)
Near-lake population (people)	916 862 (28 456–3 460 526)	419 780 (121 512–1 088 103)
Average travel time (min)	180 (2–853)	163 (21–468)
Lake cyanobacteria biomass (mg/L)	No data	138.1 (8.8–892.1)
Lake phytoplankton biomass (mg/L)	No data	148.5 (12.6–8987.7)
Lake temperature (°C)	No data	24.9 (18.1–29.0)
State Park (1 = yes)	0.008 (0–1)	No data
Aquatic invasive species (1 = present)	0.3 (0–1)	No data

Notes: Iowa lake water-quality and clarity data are annual averages from 2005 to 2012, based on three sampling events (early-, mid-, and late summer). Minnesota water-clarity data are annual averages from 2005 to 2012 for multiple sampling events per lake per year from May to September. Iowa water-quality data were provided by Iowa State University as part of the Iowa Lake Survey Project (http://limnology.eeob.iastate.edu/lakereport; J Downing, Iowa State University) and Minnesota water-quality data were provided by the Minnesota Pollution Control Agency (http://cf.pca.state.mn.us/water/watershedweb/wdip/index.cfm; P Anderson and S Heiskary, Minnesota Pollution Control Agency).

WebPanel 2. Model estimation and parameterization
We first evaluated state-specific relationships between photo-visitation and explanatory variables. For each state, we fit each explanatory variable in a bivariate regression with photo-user-days per lake as the response variable (WebTable 2).

For the final regression models considering both Iowa and Minnesota lakes, we excluded variables that were missing from either state (eg data on invasive species were not available for Iowa lakes). We also eliminated variables with significant pairwise correlations (WebTable 3). In the combined dataset, lake water clarity was significantly positively correlated with lake depth (deeper lakes tend to be clearer) and negatively correlated with total phosphorus and chlorophyll (WebTable 3). Of these variables, lake clarity is the one most likely to be perceived by lake visitors and was available for the greatest number of lake destinations. For these reasons, and because previous work found water clarity was a better proxy for perceived water quality than most other physical water-quality attributes (Jeon *et al*. 2005), we selected clarity as the variable to include in the final multiple regression relating lake attributes to lake visitation.

We tested the fully specified regression model for evidence of spatial autocorrelation in the response variable of Flickr-photo-days per lake but found none (ESRI ArcGIS 10.1, Moran's I statistic, $z = 1.12$, $P = 0.26$). Model residuals were also tested for spatial autocorrelation and again we found no evidence of significant spatial clustering.

Next, we evaluated the relationship between near-lake population and lake visitation. As expected, lakes in densely populated areas received the majority of visits (as measured in annual Flickr-photo-user-days). However, lakes in sparsely populated areas also received many visits. This relationship was most evident in Minnesota where lakes within the boundaries of the Twin Cities Metropolitan Area were frequently visited, but lakes in very remote and unpopulated parts of the state, such as lakes in the Boundary Waters Canoe Area Wilderness, were also among the most visited (WebFigure 1). Given the bimodal relationship between lake visitation and near-lake population, we centered the population variable in the multiple linear regression model (subtracted mean population from each lake population estimate) and included a squared term for population. There was not a significant relationship between near-lake population and lake visitation for the subset of Iowa lakes.

WebTable 2. State-specific relationships between photo-visitation and explanatory variables

	Minnesota	Iowa
	$n = 2120$ lakes	$n = 128$ lakes
Lake size (acres)	< 0.0001 (+)	< 0.0001 (+)
Lake clarity (m)	ns	0.0015 (+)
Lake depth (m)	< 0.0001 (+)	< 0.0001 (+)
Chlorophyll (µg/L)	ns	0.0410 (–)
Total phosphorus (µg/L)	ns	0.0622 (–)
Near-lake population (people)	< 0.0001 (+)	ns
Lake cyanobacteria biomass (mg/L)	nd	0.0125 (+)
Lake phytoplankton biomass (mg/L)	nd	0.0167 (+)
Lake temperature (°C)	nd	0.0745 (–)
State Park	ns	nd
Boat launch	< 0.0001 (+)	0.0390 (+)
Number of ramps	< 0.0001 (+)	nd
Number of docks	< 0.0001 (+)	nd
Number of toilets	< 0.0001 (+)	nd
Number of ADA facilities	< 0.0001 (+)	nd
Percent littoral area	ns	nd
Aquatic invasive species (1 = present)	< 0.0001 (+)	nd
Boundary Waters lake	ns	–

Notes: The significance and direction (+/–) of the bivariate relationships between lake visitation and lake attributes for Iowa and Minnesota lakes, where nd = no data and ns = non-significant at $P > 0.10$. ADA facilities are those that are compliant with the Americans with Disabilities Act; Boundary Waters lake is a lake within the Boundary Waters Canoe Area Wilderness. Minnesota classifies "ramps" and "launches" separately, with the latter referring to a boat access point with a dock and a ramp or sloping surface providing water access.

WebTable 3. Pairwise correlations of explanatory variables evaluated in regression models for visitation and travel time

	Lake size (acres)	Lake clarity (m)	Lake depth (m)	Chlorophyll (µg/L)	Total phosphorus (µg/L)	Near lake population (no people in 80-km radius)
Lake size (acres)		–0.01	0.06*	–0.03	–0.01	–0.05*
Lake clarity (m)	–0.01		0.60***	–0.55***	–0.46***	–0.33***
Lake depth (m)	0.06*	0.60***		–0.38***	–0.32***	–0.18***
Chlorophyll (µg/L)	–0.03	–0.55***	–0.38***		0.64***	0.27***
Total phosphorus (µg/L)	–0.01	–0.46***	–0.32***	0.64***		0.21***
Near lake population (no people in 80-km radius)	–0.05*	–0.33***	–0.18***	0.27***	0.21***	

Notes: Data represent all lakes in Minnesota and Iowa with water-quality attributes (2248 lakes). Red and blue values are for correlations greater than 0.5 and greater than 0.25, respectively. As expected, lake clarity is negatively correlated with chlorophyll and total phosphorus and positively correlated with lake depth. Correlations are significant at $P < 0.05^*$, 0.001^{**}, and 0.0001^{***}.
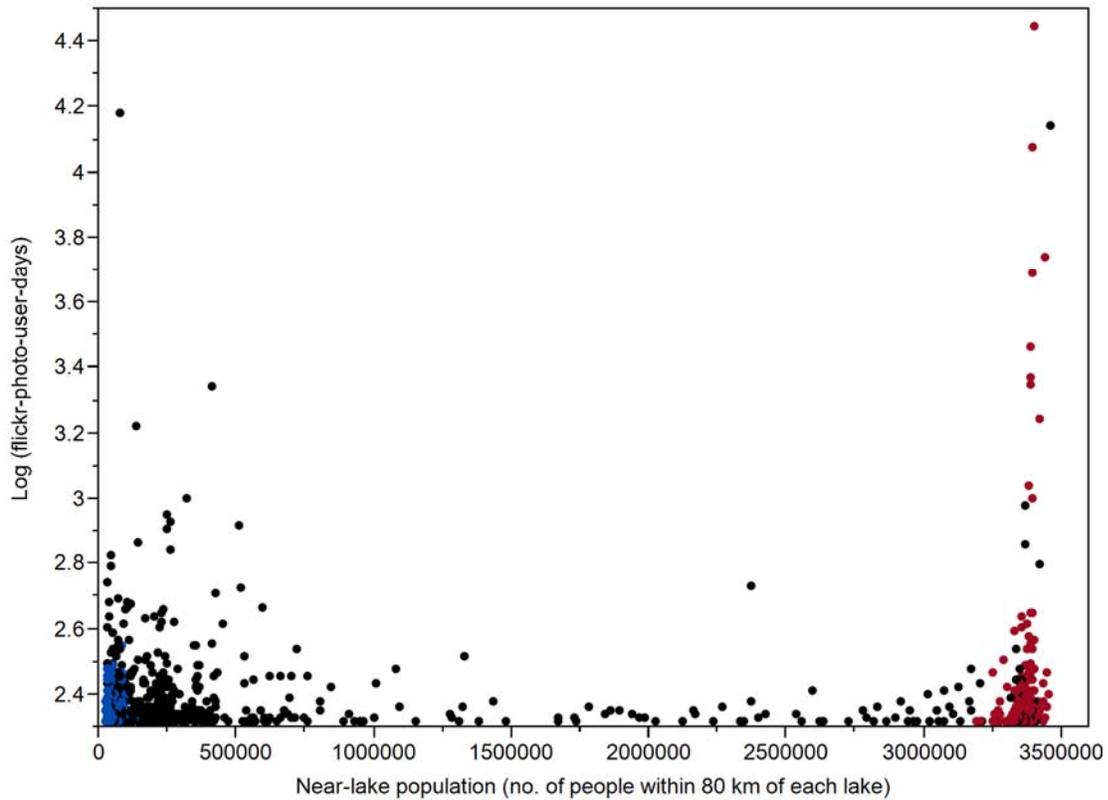
WebFigure 1. Relationship between average photo-user-days and population near each lake. Points represent individual lakes. Population surrounding each lake was measured by summing population density within an 80-km radius around each lake. Red and blue points represent lakes in the Twin Cities metro area and lakes in the Boundary Waters Canoe Area Wilderness, respectively. The figure shows that highly visited lakes occur in both densely populated and unpopulated regions of the study area.

WebTable 4. Parameter estimates and significance tests for the multiple logistic regression

	Estimate	SE	Prob > ChiSq
Intercept	1.12	0.155	< 0.0001
Lake size (acres)	0.001	0.0001	< 0.0001
Lake clarity (m)	0.127	0.033	0.0001
Centered population	-5.57×10^{-7}	1.36×10^{-7}	< 0.0001
Centered population squared	5.19×10^{-13}	8.03×10^{-14}	< 0.0001
Boat launch (1 = yes)	0.207	0.050	< 0.0001
Iowa or Minnesota (1 = IA)	0.651	0.107	< 0.0001

Notes: Parameter estimates and significance tests for the multiple logistic regression for lake attributes and photo visitation per lake where the response is 1 if the lake was visited and 0 if the lake was not visited. Each observation refers to a lake located in Minnesota or Iowa ($n = 2248$ lakes). For explanation and justification of the centered population variable see WebFigure 1, which plots visitation as a function of near-lake population. SE = standard error; Prob = probability; ChiSq = chi-square value.

WebTable 5. Economic valuation of lake clarity

Additional travel time per lake due to a 1-m increase in lake clarity (min per round-trip)	Median hourly wage in MN and IA for 2012	One-third of weighted median hourly wage for MN and IA*	Value of wages (wage × travel time)	Value of mileage ($0.30 per mile × miles traveled)	Total travel costs
56	$17.74 (MN) and $15.33 (IA)	$5.85	$5.46	$16.80	$22.26

Notes: Estimated average per-lake travel costs associated with increased water quality. To translate the travel times estimated in the regression equation specified in Table 2 into cost estimates, we multiplied hourly wages and mileage values by the estimated increase in travel time associated with a one-meter increase in lake clarity. Median hourly wage estimates from MN and IA are averages for all occupations as reported by the US Bureau of Labor Statistics (2012). Value represents a weighted average reflecting the relative proportion of trips to Iowa and Minnesota lakes. All dollar amounts are expressed in US dollars. *The one-third hourly wage adjustment is a commonly used lower-bound estimate used to value time in the recreation literature (Parsons 2003).

WebReferences

Jeon Y, Herriges JA, Kling CL, and Downing J. 2005. The role of water quality perceptions in modeling lake recreation demand. Ames, IA: Iowa State University. www.econ.iastate.edu/research/working-papers/p3840. Viewed 25 Nov 2013.

Kloiber SM, Brezonik PL, and Bauer ME. 2002. Application of Landsat imagery to regional-scale assessments of lake clarity. *Water Res* **36**: 4330–40.

Olmanson LG, Brezonik PL, and Bauer ME. 2011. Evaluation of medium to low resolution satellite imagery for regional lake water quality assessments. *Water Resour Res* **47**: W09515.

Parsons GR. 2003. The travel cost model. In: Champ PA, Boyle KJ, and Brown TC (Eds). A primer on nonmarket valuation. Dordrecht, The Netherlands: Kluwer.

US Bureau of Labor Statistics. 2012. May 2012 National Occupational Employment and Wage Estimates. www.bls.gov/oes/current/oes_nat.htm. Viewed 24 Nov 2013.