Harmful algal blooms and toxic air: The economic value of improved forecasts

Klaus Moeltner[1], Tracy Fanara[2], Hosein Foroutan[3], Regina Hanlon[4], Vince Lovko[5], Shane Ross[6], and David Schmale III[7]

[1]Professor, Department of Agricultural and Applied Economics, Virginia Tech, 208 Hutcheson Hall, Blacksburg, VA 24061, moeltner@vt.edu *
[2]Coastal Modeling Portfolio Manager, U.S. Integrated Ocean Observing System, 1315 East-West Highway 2nd Floor Silver Spring, MD 20910, tracy.fanara@noaa.gov
[3]Assistant Professor, Department of Civil and Environmental Engineering, Virginia Tech, 417 Durham Hall, Blacksburg, VA 24061, foroutan@vt.edu
[4]Research Associate, School of Plant and Environmental Sciences, Virginia Tech, 413 Latham Hall, Blacksburg VA 24061, rhanlon@vt.edu
[5]Senior Scientist and Program Manager, Mote Marine Laboratory & Aquarium, 1600 Ken Thompson Parkway Sarasota, FL 34236, vlovko@mote.org
[6]Professor, Department of Aerospace and Ocean Engineering, Virginia Tech, 495 Old Turner St. Rm 320, Blacksburg VA 24061, sdross@vt.edu
[7]Professor, School of Plant and Environmental Sciences, Virginia Tech, 403 Latham Hall, Blacksburg VA 24061, dschmale@vt.edu

Revised submission: March 21, 2022
Suggested running head: Value of harmful algal bloom forecasts
Keywords: Red tide, Choice experiment, Bayesian analysis
JEL codes: C11, C25, Q51, Q53

Abstract

The adverse economic impacts of harmful algal blooms can be mitigated via tailored forecasting methods. Adequate provision of these services requires knowledge of the losses avoided, or, in other words, the economic benefits they generate. The latter can be difficult to measure for broader population segments, especially if forecasting services or features do not yet exist. We illustrate how Stated Preference tools and Choice Experiments are well-suited for this case. Using as example forecasts of respiratory irritation levels associated with airborne toxins caused by Florida red tide, we show that 24-hour predictions of spatially and temporally refined air quality conditions are valued highly by the underlying population. This reflects the numerous channels and magnitude of red tide impacts on locals' life and activities, which are also highlighted by our study. Our approach is broadly applicable to any type of air quality impediment with risk of human exposure.

*Corresponding author (phone: (540) 231-8249). We acknowledge funding by the Global Change Center and the Institute for Society, Culture, and the Environment at Virginia Tech. The authors declare no conflicts of interest related to this article.

Introduction

Harmful algal blooms (HABs[1]) can cause serious economic damages in the U.S. and around the world by negatively impacting commercial and recreational fisheries, aquaculture, seafood markets, tourism, and public health (Hoagland et al. 2002; C. Adams et al. 2018; Jin et al. 2020; Plaas and Paerl 2021). As discussed in Jin and Hoagland (2008) and Jin et al. (2020), predicting the occurrence and intensity of HABs can alleviate these costs, by allowing decision makers to adjust economic activities to mitigate impacts. Thus, HAB predictions carry economic value, and this value, in turn, needs to be understood by private and public investors for the efficient allocation of funds to develop and refine HAB forecasting systems (Jin et al. 2020).

Jin and Hoagland (2008) present a conceptual framework of deriving the value of HAB predictions in terms of re-optimized harvesting decisions, and apply it to the New England shellfish sector. However, their model requires assumptions on how exactly HABs affect decision makers, and what mitigating actions they take in reaction to the forecast. While this may be feasible for a specific, well-understood sector of the economy, it is a tall order when contemplating forecast values to a broader residential population. This is especially true when the envisioned forecast system does not yet exist, and / or the HAB under consideration has the potential to affect locals via multiple channels, triggering a wide range of mitigating actions.

In such case of non-existing or difficult-to-characterize markets, survey-based Stated Preference (SP) methods constitute an attractive alternative to structural, market-based models to value forecasts. In a nutshell, this approach elicits the monetary value of the commodity or service under consideration (here: HAB forecasts) by building a hypothetical market for the service, and capturing respondents "purchasing" decisions. This can be accomplished without knowing every behavioral nuance that links people with the non-market good under consideration (e.g. Brown 2017; Freeman 2003; Phaneuf and Requate 2017). If properly implemented, following best practice in survey design and estimation, SP methods have been shown to generate realistic estimates of individuals' true value, or "willingness-to-pay" (WTP) for such goods (Johnston, Boyle, et al. 2017).

We illustrate this approach within the context of HABs of *Karenia brevis*, commonly referred to as "Red Tide" (RT), in southwest Florida. These blooms are known for causing respiratory irritation and illness in humans via aerosolized toxins, among other environmental impacts. Specifically, we elicit the economic value to residents of five southwest Florida counties of a hypothetical RT air quality forecast that improves over existing information systems, using an SP approach anchored in a state-of-the-art economic choice experiment (CE). While there exists a considerable body of literature on the value of weather and climate forecasts to agriculture and other sectors (e.g. R. Adams et al. 1995; Johnson and Holt 1997; Solow et al. 1998; Freebairn and Zillman 2002; Lazo and Chestnut 2002; Rollins and Shaykewich 2003), contributions that value air quality predictions, HAB-related or otherwise, are less abundant.[2]

We estimate the annual value of this improved forecasting system at \$17-\$45 per household, depending on stipulated spatial coverage and accuracy. This aggregates to approximately \$14.5 -\$37 million per year for the Five-County Region (5CR) at large. We believe that these figures constitute a useful starting point to assess the potential net benefits to the local population of more refined RT forecasts and information systems. At the same time, the upper end of our estimates for fully accurate forecasts can be considered a (very conservative) lower bound for the economic costs caused by RT air contamination via their interference with human outdoor activities.

To our knowledge this work constitutes the first effort to quantify, in monetary terms, societal benefits of forecasts for aerosolized ecotoxins, or, for that matter, *any type* of air quality problem.[3] In passing, our survey also reveals in more detail than provided by existing studies the diversity and intensity of outdoor activities enjoyed by 5CR residents, and the numerous ways in which past RT episodes have interfered with these endeavors and triggered mitigating actions.

Red tide impacts and existing information sources

In recent years, blooms of RT have increased in frequency, intensity, and geographic spread along the Florida Gulf Coast (FGC) (e.g. Alcock 2007; Brand and Compton 2007; Nierenberg et al. 2009; Fleming et al. 2011; Corcoran et al. 2013; Stumpf et al. 2022). *Karenia brevis* produces powerful neurotoxins ("brevetoxins") that can cause widespread mortality to fish, marine mammals and sea birds, as well as respiratory irritation and illness in humans as RT cells get broken up by wind and wave action and mixed into the ambient aerosol (Backer et al. 2003; Fleming et al. 2011; Kirkpatrick et al. 2006; Kirkpatrick et al. 2011). From October 2017 through the winter of 2018/2019, the FGC experienced the most intense and longest RT bloom in decades, at times stretching over 150 miles of coast line, and producing RT concentrations in the tens of million cells per liter (Wei-Haas 2018; National Ocean Service 2019).[4]

Documented economic consequences of RT events include revenue losses to businesses and tourism, public health costs, beach cleanup expenses, and direct welfare losses to residents and visitors via diminished recreational opportunities and interference with outdoor activities (Larkin and Adams 2007; Morgan, Larkin, and Adams 2008; Hoagland et al. 2009; Morgan, Larkin, and Adams 2009, 2010, 2011). While scientific efforts are ongoing to curb RT blooms through prevention and control methods, the predominant management strategy to date has been mitigation, via early detection and avoidance of human contact (Alcock, 2007; Corcoran et al.,2013). In part, this is achieved through online information free to the public on current and expected RT cell counts and irritation levels. For example, Mote Marine Laboratory & Aquarium (MML) in Sarasota, Florida, manages the Beach Conditions Reporting System (BCRS), which provides twice-daily updates on current RT irritation levels at 37 beach locations along the FGC (Nierenberg et al. 2009; Mote Marine Laboratory & Aquarium 2020). These reports are based on self-experienced or observed frequency of coughing and sneezing by trained beach sentinels, such as lifeguards. Similarly, the Florida Fish & Wildlife Conservation Commission (FWC) maintains a web site that summarizes RT concentrations from water samples, observed over the preceding two weeks. It also gives a link to detailed reports on RT concentrations for a given sampling day and site, and offers a translation of cell counts to potential respiratory irritation levels that may be experienced nearby (Florida Fish and Wildlife Conservation Commission 2020).

While the BCRS and FWC sites focus on past and current RT conditions, the National Oceanic and Atmospheric Administration (NOAA) issues bi-weekly to weekly forecasts for RT blooms via its Gulf of Mexico Harmful Algal Bloom Forecast (HABF) web site. The web site provides a three to four day forecast of potential respiratory irritation for four target regions, including Northwest and Southwest Florida, respectively. Within each region, expected irritation levels are given for smaller sub-areas for each of the following three days. In a similar vein, experimental RT respiratory forecasts are currently underway in selected Gulf coast counties by combining *Karenia brevis* cell counts based on water samples with wind forecasts produced by the National Weather Service to generate predictions of RT irritation levels in three hour increments, up to a 24 hour forecast window (Gulf of Mexico Coastal Ocean Observing System 2020; National Centers for Coastal Ocean Science 2020).

Rather than eliciting values for existing RT information and forecast systems, we propose a hypothetical, more advanced system that fills important shortcomings of existing sources by better capturing the pronounced spatial and temporal variability of RT aerosol concentrations (Nierenberg et al. 2009), and their ability to travel inland for considerable distances (Kirkpatrick et al. 2010). Accordingly, the envisioned system features broader spatial coverage that includes inland areas, and higher spatial and temporal resolution than existing sources. These improvements are essential to allow users to make informed and refined day-to-day plans for outdoor-based activities.

Methods

Conceptual and econometric modeling

Our conceptual modeling framework is anchored in random utility maximization (RUM) theory, which stipulates that individuals gain benefits ("utility") from a given choice option (here a specific forecast scenario), but that the components feeding into the underlying utility function are only partially observed by the researcher. The RUM approach to economic research was pioneered by McFadden (1974), and has since become one of the dominant theoretical foundations for economic valuation.[5]

Following this general approach, we let the indirect utility function (IUF) for person i, choice occasion t, and forecast alternative j be given as:[6]

$$\begin{aligned} &U_{itj} = \mathbf{z}_{itj}'\boldsymbol{\theta} + \lambda\left(m_i - P_{itj}\right) + \epsilon_{itj}, \quad \text{with} \\ &\epsilon_{itj} \sim EV\left(0,1\right), \end{aligned} \tag{1}$$

where vector $\mathbf{z}_{itj}$ comprises forecast attributes, m_i denotes annual income, P_{itj} is the price or "bid" associated with the forecast scenario, and ϵ_{itj} captures all other components that affect utility, but are not visible to the analyst. In essence, this IUF can be interpreted as a reduced-form expression that succinctly captures all utility-enhancing underlying actions that individual would be able to take to mitigate RT impacts, given forecast features $\mathbf{z}_{itj}$.

If the error component of this IUF is identically and independently distributed following a Type-I Extreme Value (EV) distribution with zero mean and unity scale, as noted in the second line of (1), the probability of i selecting option j on occasion t can be conveniently expressed as:

$$\begin{aligned} &\text{prob}\left(y_{itj}=1\right) = \frac{exp\left(\mathbf{x}_{itj}'\boldsymbol{\beta}\right)}{\sum_{j=1}^{J} exp\left(\mathbf{x}_{itj}'\boldsymbol{\beta}\right)}, \quad \text{where} \\ &\mathbf{x}_{itj} = \begin{bmatrix}\mathbf{z}_{itj}' & P_{itj}\end{bmatrix}', \quad \text{and} \quad \boldsymbol{\beta} = \begin{bmatrix}\boldsymbol{\theta}' & -\lambda\end{bmatrix}', \end{aligned} \tag{2}$$

and y_{itj} be a binary indicator that takes the value of 1 if i chooses j on the t^{th} occasion, and a value of zero otherwise.

The choice model characterized by equations (1) and (2) is generally referred to as "Conditional Logit" (CL) in applied economics and related fields (e.g. McFadden 1974; Train 2009; Greene 2012). A further distinction of CL specifications is made based on the nature of the given alternatives. Specifically, our choice options do not represent existing real-world alternatives such as transportation ("bus," "car," "bike,") or food choices ("beef," "pork," "chicken"), but rather hypothetical mixes of forecast attributes that vary in composition across alternatives and individuals. As such, our case constitutes what is generally referred to as an "unlabeled" choice experiment (Henscher, Rose, and Greene 2005; Holmes, Adamowicz, and Carlsson 2017).[7]

Conditional Logit models can also differ via their specification of the "Status Quo" (SQ) alternative, that is a constant "baseline" option given in all choice sets offered to the sample at large. As discussed below in more detail, the inclusion of a SQ alternative is considered "best practice" in choice modeling. It usually captures indirect utility related to some form of non-participation or non-purchase, and is often modeled by setting the price term P_{itj} to zero, and by either setting attribute vector $\mathbf{z}_{itj}$ to fixed levels, or replacing it with a single constant term. The latter strategy is more meaningful in our case, as there are no existing forecasting systems for RT aerosol that fit into the mold of the attribute settings for the new envisioned system. In consequence, the SQ indicator in our model succinctly captures whatever a given respondent perceives as the relevant features of existing RT information sources when deciding between alternatives.[8]

The sample likelihood for $i = 1 \dots N$ independent individuals, each facing T independent choice occasions involving J alternatives,[9] is given by

$$p\left(\mathbf{y}|\boldsymbol{\beta}, \mathbf{X}\right) = \prod_{i=1}^{N} \prod_{t=1}^{T} \prod_{j=1}^{J} \left(\frac{exp\left(\mathbf{x}_{itj}'\boldsymbol{\beta}\right)}{\sum_{j=1}^{J} exp\left(\mathbf{x}_{itj}'\boldsymbol{\beta}\right)} \right)^{y_{itj}}. \tag{3}$$

Traditionally, the model in (3) is estimated via Maximum Likelihood (MLE) using some form of gradient or search algorithm to determine the optimal coefficient vector $\boldsymbol{\beta}^*$ that maximizes the probability of observing the combined choices of the sample at hand. We opt instead for a Bayesian estimation framework for the the reasons of (i) non-dependence on large-sample ("asymptotic") theory, (ii) obtaining full finite-sample distributions for each primary model parameter, and (iii) ease of deriving full distributions for predictive constructs, such as the total value of a forecast with a specific combination of attribute settings.[10]

Bayesian estimation requires the ex-ante choice of a prior distribution for coefficient vector $\boldsymbol{\beta}$. We follow standard practice and specify this density as a multivariate normal with mean $\boldsymbol{\mu}_0$ and variance-covariance matrix $\mathbf{V}_0$ (e.g. Rossi, Allenby, and McCulloch 2005; Train and Sonnier 2005; Daziano 2013). This prior is them combined with the likelihood in (3) to form the joint posterior distribution, the estimation target in Bayesian analysis. Dropping normalization terms that are not needed for parameter estimation, the kernel of this posterior can then be written as:

$$p\left(\boldsymbol{\beta}|\mathbf{y},\mathbf{X}\right) \propto exp\left(-\frac{1}{2}\left(\boldsymbol{\beta}-\boldsymbol{\mu}_0\right)'\mathbf{V}_0^{-1}\left(\boldsymbol{\beta}-\boldsymbol{\mu}_0\right)\right)\prod_{i=1}^{N}\prod_{t=1}^{T}\prod_{j=1}^{J}\left(\frac{exp\left(\mathbf{x}_{itj}'\boldsymbol{\beta}\right)}{\sum_{j=1}^{J} exp\left(\mathbf{x}_{itj}'\boldsymbol{\beta}\right)}\right)^{y_{itj}} \tag{4}$$

Since the expression in (4) does not have a well-understood analytical form, Bayesian inference proceeds by taking draws from this posterior via a Metropolis-Hastings (MH) algorithm based on an auxiliary "proposal function" (Hastings 1970). Several different MH strategies exist for the CL model. Based on results from initial runs with simulated and actual data we opt for the strategy outlined in Rossi, Allenby, and McCulloch (2005) and use a multivariate-t distribution as proposal density with mean equal to the mode, and variance matrix based on the second derivative of (4), respectively. This also allows for the introduction of two tuning parameters to control sampling efficiency, a scalar for the variance matrix (τ) and degree-of-freedom parameter ν. The details of this approach are given in a separate online appendix.

Willingness-to-pay predictions

Formally, the value w_i of a forecasting system with attribute settings $\mathbf{z}_p$ to a given individual is derived implicitly by equating indirect utility for the SQ at full income m_i with indirect utility associated with the forecast and reduced income $m_i - w_i$. In other words, w_i is the maximum amount the individual is willing to trade off for the improved system.[11] Separating the full coefficient vector $\boldsymbol{\theta}$ from (1) into a sub-vector $\boldsymbol{\theta}_z$ that corresponds to actual forecasting attributes, and the SQ coefficient θ_{SQ}, this "willingness-to-pay" (WTP) can then be derived in straightforward fashion as:

$$w_i|\mathbf{z}_p, \boldsymbol{\theta}, \lambda = \frac{1}{\lambda}\left(\mathbf{z}_p'\boldsymbol{\theta}_z - \theta_{SQ}\right). \tag{5}$$

In practice, draws from the full posterior predictive density (PPD) for this WTP measure are conveniently obtained by evaluating (5) for all $r = 1 \ldots R$ draws of $\boldsymbol{\beta}$ flowing from the original posterior sampler. This set of predictive draws, in turn, form the basis for statistical inference. This procedure can then be repeated for other forecast settings, as discussed below.

Survey design and implementation

Focus group sessions

Field research for this project was initiated via three focus group sessions in December 2019 and January 2020, held at the Mote Marine Laboratory & Aquarium (henceforth MML) in Sarasota, Florida. Each session comprised nine participants who were recruited from MML's expansive list of volunteers and citizen scientists. Prospective participants had to be at least 18 years old and year-round residents of the 5CR. Each session lasted 90 minutes, and participants were compensated with $85 for their time. Refreshments and opportunities for breaks were provide throughout each session. All focus group materials (recruitment forms, informed consent, focus group protocols) were approved by Virginia Tech (VT)'s Institutional Review Board (IRB).

The primary aim of these sessions was to learn how recent episodes of RT blooms have affected local residents, with special focus on air quality. Additionally, the sessions provided a first opportunity to gauge the perceived usefulness of an improved air quality forecasting system, prioritize the importance of forecast attributes, and collect initial information on potential willingness to pay for the new system. Focus group moderation followed best practices as outlined inter alia in Johnston et al. (1995) and Nyumba et al. (2017), such as using non-technical language, letting the discussion flow as freely as possible, and probing for participants' experiences rather than opinions.

In sum, the focus group sessions illustrated clearly that RT air contamination causes widespread and severe problems to the local population, including direct health effects and impediments to activities of daily life, not necessarily limited to outdoor recreation. The envisioned forecasting system was generally perceived as a very useful tool that would help to "plan ahead."

Refinement of forecasting system

Given focus group feedback, literature review, team discussions, and the features of existing RT information services, we decided to cast the envisioned system as a 24-hour forecast, with hourly updating, spatial resolution of one square mile, and coverage along the entire 5CR coastline with a band width of up to 12 miles. Following virtually all existing RT monitoring systems, hourly forecasts would be issued in terms of the severity of irritation caused by RT aerosol, from "none" to "high," with intermediate levels of "low," and "moderate," and corresponding color codes of white, yellow, orange, and red. Following MML's Beach Conditions Reporting System (BCRS), we consider a level of "low" as affecting primarily people with chronic respiratory problems, a level of "moderate" as triggering mild symptoms, such as coughing, sneezing, and eye irritation among the majority of the exposed public, and a level of "high" as producing severe respiratory symptoms (heavy coughing or sneezing, severe eye irritation) for the majority of those exposed (Mote Marine Laboratory & Aquarium 2020). In terms of accessibility, we envision the new system as offered online via a designated web site, and as a cell phone application ("app").

Figure 1 depicts the general idea of what users might see first upon accessing the system, using as an example a 6-mile coverage band and the city of Sarasota. Each panel shows a forecast at a different point in time. Each panel also features the one-mile grid of spatial resolution, and the color-coded respiratory irritation levels that govern each square mile at the indicated point in time. Users of the system would then be able to click on any desired square to access the 24-hour by-the-hour forecast for a given location. An example is given in figure 2 for the "St. Armands / Lido Beach" square within the Sarasota grid. As is evident from the figure, forecasted hourly irritation is symbolized as a color-coded horizontal band with hourly cells that span the entire 24-hour window, starting from the most recent hourly update.

In sum, with these envisioned features, the new system would have much higher spatial resolution and much more frequent updating than any of the existing RT information sources. It would also exhibit much broader spatial coverage, even at its smallest band width.[12]

Design of choice experiment

At the heart of the survey instrument lies the actual economic choice experiment (CE). This elicitation tool is based on a sequence of choice sets presented to the respondent. Each set offers several choice options (also referred to as "profiles," or "alternatives"). Each option, in turn is characterized by a set of attribute levels related to the resource or commodity to be valued, plus a "price tag," usually in form of increased taxes or utility bills. Observing each respondent's preferred option across sets then allows the analyst to estimate the dollar-denominated marginal value of each attribute, as well as any desired bundle of attribute settings, to the underlying population of interest.

In addition to the selection of attributes and levels as described above, the process of CE design involves the combination of attribute settings to form a profile, the grouping of profiles into choice sets, and the blocking of choice sets into manageable sequences that can be presented to different sub-sets of survey respondents. Accessible descriptions of these tasks and underlying techniques are given in Henscher, Rose, and Greene (2005), Johnson et al. (2007), and Holmes, Adamowicz, and Carlsson (2017).

The number of attributes and levels is typically chosen to adequately characterize the resource or commodity under consideration while at the same time acknowledging the cognitive limitations of respondents (see e.g. De Shazo and Fermo 2002; Swait and Adamowicz 2001; Boxall, Adamowicz, and Moon 2009). With that in mind, we settled for three attributes that capture the most salient features of the envisioned system: (i) The width of spatial coverage along the coastline, the degree of accuracy of the forecast for the first 12 hours, and the degree of accuracy for the second set of 12 hours, counting from the most recent forecast update available to the user. We express spatial coverage in terms of a band of either six or 12 mile width, stretching along the entire five-county coastline. Figure B1 in the online appendix provides a bird's eye view of the spatial bands, while Figure B2 depicts more localized examples of these bands for selected coastal population hubs. Both figures were also shown to respondents in the final survey questionnaire. As can be seen from Figure B1, both bands cover all existing sampling locations of MML's BCRS, with the 12-mile version extending both further inland and into the Gulf.

Accuracy, in turn, is quantified as the percentage of correctly forecasted degrees of irritation over a 12 hour period, with precision levels set at 50%, 75%, and 100% for both the first and second 12-hour window. An example of the notion of accuracy for the first 12 forecasted hours, also depicted in the final questionnaire, is given in online appendix figure B3. The bottom horizontal color bar gives the (hypothesized) fully accurate forecast, while the bars stacked above depict forecasts with three inaccurate cells (=75% accuracy) and six inaccurate cells (=50% accuracy), respectively. For simplicity, respondents were instructed to assume that false forecasts would deviate from the truth by no more than one irritation level, in either direction.

The set of three forecast attributes is complemented by a price variable, which we specify at four uniformly spaced levels of \$5, \$15, \$25, and \$35. These bids were informed by fees suggested by focus group participants for various stylized sample forecast systems, by WTP estimates produced by a similar CE study on weather forecast improvements (Lazo and Chestnut 2002), and by existing annual subscription fees charged for weather-related online services.[13] These four bid values performed well in pretesting, and were thus retained for final survey implementation. The actual payment vehicle given in the survey was an annual cost in form of "new local taxes and fees," as described below in more detail. Table 1 provides a summary of all CE attributes and corresponding levels.

The full factorial of these attribute settings produces an initial set of 72 choice profiles (also referred to as "options" or "alternatives"), each with a unique combination of attribute levels. Eliminating unrealistic profiles that show higher accuracy for the second set of 12 hour-forecast than for the first set leaves a starting set of 48 unique profiles. These alternatives were then grouped into choice sets of two actual options, plus the SQ option of existing information sources at a cost of \$0. This setup of "two + SQ" is by far the most common template for choice experiments within the realm of environmental economics, balancing a manageable cognitive burden on the respondent with adequate statistical efficiency in estimation (e.g. Ferrini and Scarpa 2007).[14]

The final number of choice sets to be used in the actual survey is driven by two considerations: The number of sets to be shown to a given respondents, and the minimum number of observations the analyst wishes to collect per set. Given expected sample sizes and model parameters to be estimated, we settled for a total of 20 choice sets administered in five different blocks of four sets each. That is, each respondent was randomly assigned to a given block, and presented with four corresponding choice sets. This assures the identification of all model parameters, while keeping cognitive requirements and survey completion time at manageable levels. Blocks of four sets are, again, a common choice in environmental economic research (Ferrini and Scarpa 2007).

From a statistical perspective, the optimal grouping of alternatives into the 20 choice sets follows a "D-optimality" criterion that is based on an optimization algorithm that iteratively and repeatedly groups options into choice sets until the determinant of the covariance matrix of estimated parameters (under specific modeling assumption) is minimized (Johnson et al. 2007; Ferrini and Scarpa 2007; Rose and Bliemer 2009; Holmes, Adamowicz, and Carlsson 2017).[15] During this process, we impose the additional restriction that a choice set cannot contain option pairs such that one option perfectly dominates the other, i.e. by offering wider coverage *and* higher accuracy at an equal or lower price tag. Figure B4 in the online appendix shows an example of a choice set, in the same format as shown to survey respondents.

Survey instrument

The design of the survey instrument was informed by the focus group sessions described above and a pretest (see below). The questionnaire is grouped into four main sections. The first section provides detailed background information on RT blooms along the Southern

Florida Gulf Coast. It then asks respondents how and to what extent they have been affected by poor air quality associated with RT across a broad set of outdoor activities, and what mitigating actions, if any, they have taken in the past.

The second section summarizes existing public information systems for RT cell counts and irritation levels for the region, and elicits familiarity with these (free-of-charge) services. This is followed by an introduction of the envisioned new forecasting system, and preparatory information and primer questions for the actual choice experiment.

The third section presents the four choice sets as described above and collects the four corresponding choice decisions. It also features two important follow-up questions. The first, offered to those who choose the SQ option on all four occasions, identifies problematic responses that are not based on considerations of benefits and costs, and / or indicate disbelief of provided factual information. These "protest votes" can bias estimation of economic values, and are generally excluded from final model estimation (Meyerhoff, Bartczak, and Liebe 2012; Johnston, Boyle, et al. 2017). In our application, protest responses were flagged as those that reflect opinions that the new system ought to be covered by existing fees and taxes, that it will not be scientifically feasible, and / or that it would channel resources away from other efforts to combat RT (despite having been reassured to the contrary earlier in the survey). The second follow-up question asked respondents who chose an actual forecast option on at least one occasion for their primary motivation(s) of doing so.

The final questionnaire section collects standard demographic information, such as location of residence by ZIP code, household size, education level, and income category.[16]

Throughout the development of the survey instrument care was taken to follow "best practices" in stated preference research, as mapped out in Johnston, Boyle, et al. (2017). These interventions are designed to align the choice experiment as closely as possible with a real-world voting context to avoid biased responses that can arise due to the hypothetical nature of the commodity or service in question, or via a number of other undesirable channels that are often related to poor or incomplete survey design (Mitchell and Carson 1989; Holmes, Adamowicz, and Carlsson 2017; Johnston, Boyle, et al. 2017). For example, we presented respondents with a realistic and binding decision rule, that is with a linkage between their vote and the potential implementation of the new forecasting system. Specifically, we explained that (i) their responses would help public officials and organizations understand the type of forecasting system they would support in an actual public vote, (ii) that survey results would be shared with public officials to aid in their decision to implement the forecast, and (iii) implementation decisions, in turn, would hinge on a majority rule for the preferred forecast version (including no new system). In the same vein, we stressed that payment in form of "new local taxes and fees" would be mandatory and binding for all residents of the 5CR for as long as they lived in the region, regardless of their vote, should the new system be implemented.[17]

By the same token, respondents were reminded that the new system, once activated, would be accessible to all residents of the research area, regardless of their original vote. As discussed in Vossler, Doyon, and Rondeau (2012) and Johnston, Boyle, et al. (2017), the combination of assuring consequentiality of voting, ideally via a majority rule, and a credible and binding payment mechanism are essential to avoid "free-riding" and to induce truthful preference revelation.

Furthermore, we mitigate against potential sequencing effects, i.e. undesirable and unobservable influences of earlier choice sets and votes on decisions for subsequent sets for a given individual, by making choice sets "visible" as recommended in Johnston, Boyle, et al. (2017). This entails a verbal preview of the choice tasks to follow, and a reminder to treat each set as fully independent and the only options available on a given voting occasion. In addition, we rotated choice sets across sub-sets of respondents within each block, such that each of the four sets per block had an equal chance to be seen first by a given participant. This provides the option to estimate the full econometric model using only "first-set" observations, which by construction circumvents any potential sequencing effects.

To counter any perceptions of best responses or social norms for selecting a forecast option we assured respondents that there are good reasons for different people to choose different options, and that every vote is legitimate, regardless of the chosen alternative.[18] In the same vein, to induce financial discipline, we reminded survey participants of the budget tradeoffs implied by their vote, potentially leaving less money for other private expenses or public contributions.

Lastly, we follow recommendations in Krupnick and Adamowicz (2007) and Johnston, Boyle, et al. (2017) and collect additional responses at the end of survey section three related to participants' perceptions of (i) the quality and usefulness of background information provided, (ii) the realism and consequentiality of the voting environment, and (iii) freedom to form their own opinion and make their own voting decision. As described below in more detail, we use results from these questions in conjunction with flagged protest votes to identify valid observations that are suitable for econometric estimation.

Survey implementation

The survey questionnaire was pretested online in May 2020, using the original focus group participants as test sample. All 27 individuals provided responses. This led to minor modifications in the questionnaire, but generally provided evidence that the background information was clear, the choice tasks were well understood, survey length was appropriate, and the price / bid amounts in the choice sets were within reasonable range.

The final survey was administered online by a professional polling firm (Qualtrics) between June and September 2020. Respondents had to be at least 18 years of age and permanent residents of the 5CR. Each participant had to agree to an informed consent form prior to accessing the questionnaire, in accordance with VT IRB regulations. Survey participants were recruited from several "opt-in panels" accessible to the firm. To assure representativeness of the resulting sample of participants relative to the underlying target population, the sample was stratified according to predetermined age and income groups, as well as the presence of children under the age of 18, based on official census data for the 5CR. Participants received monetary compensation of undisclosed magnitude. The survey firm provided initial quality control by screening out individuals that skipped 10% or more of the survey questions or completed the survey in a time less than 50% of the sample median. This figure, in turn, was derived from an initial "soft launch" involving 50 individuals to verify correctness of the underlying survey architecture, such as randomization patterns and linked follow-up questions.[19]

Results

Descriptive statistics

The survey generated 502 properly completed questionnaires, slightly exceeding our target sample of 500. Basic sample demographics are given in Table 2. As is evident from a comparison of the first and last column of the table, our sample matches up reasonably well with official population demographics for the 5CR, albeit with a slightly larger contingent of females, individuals age 65 or older, and higher proportions of residents with at least a high school or bachelor degree, respectively.[20]

The table also shows that the average respondent has lived within the 5CR for over 14 years, close to nine of those at the current address. We can therefore expect our sample to be quite familiar with topics related to RT. Approximately 23% of surveyed households have children age 18 or younger, and a sizable share (close to 29%) comprise family members with pre-existing respiratory conditions.

The survey asked respondents about their participation in 11 designated outdoor activities (plus an "other" option) over the past 12 months, ranging from a variety of land and water recreational activities, to school- or work-related time spent outdoors. A follow-up question then elicited the total number of hours spent on each activity by all household members in a typical week without any RT problems. Table 3 captures responses to the participation question. As is clear from the table, a large majority of residents (over 88%) spend some time outside over the course of a year, with the highest percentages observed for beach recreation (close to 73%), followed by trail or road endeavors such as jogging and biking (close to 58%), and water activities other than boating (approximately 57%). Furthermore, over 13% of the sample spends time outdoors as part of their work or volunteer engagements.

Table 4 summarizes actual time spent outside per week, summed over all household members. The first four columns of the table show the share of households that fall into the time bracket given in the row header. For example, over 45% of households spend between one and four hours on the beach in a given week, over 13% clock between five and eight beach hours per week, and over six percent enjoy the beach for nine hours or more. The last column of the table gives the total number of hours per activity, averaged across households, and using category midpoints.[21] We find that the average five-county household spends close to 17 hours per week on the listed activities, across all its members. The largest shares of this total go to trail or road activities (2.73 hours), and beach time (2.70 hours), with the remaining activities distributed across the 0.6 to 1.5 hour/week range.

The survey also inquired about time spent in outdoor areas around the house, such as balcony, deck, or yard. As can be seen from the last row of the table, all but eight percent of households spend at least some time in these locations, with the sample average emerging as close to 17 hours/week. Adding "near home" to "away from home" outdoor time, we find that the typical five-county household spends close to 25 hours outside in a given week. This stresses the importance of outdoor time to the local population, but also its ex-ante susceptibility to RT aerosol exposure during a bloom.

This is confirmed by the next block of questions, which inquired about personally experienced RT air contamination effects and mitigating actions taken during a household's residential tenure in the 5CR. Responses are summarized in Table 5. As captured in the last column of the table, a majority of households (approximately 57%) had to cancel, postpone, or shorten outdoor activities away from home at least occasionally during past blooms, and over 47% were forced to re-locate a planned outdoor event.[22] Over 40% of respondents have been impacted in similar fashion with respect to outside activities around the house. In terms of direct health effects, close to 58% have experienced some respiratory irritation due to RT in the past, with over 11% of residents affected to a severe degree. We also note that a non-negligible share of respondents (three to nine percent) report being impacted virtually every day during a bloom in different ways, with over three percent suffering from severe irritation on a daily basis. A large share of households (over 48%) also indicate that they have been adversely affected by the odor of decaying fish killed by RT.

The last block of rows in the table captures other, perhaps less obvious RT effects that were mentioned during focus group sessions. These include cancellation of visits by family and other out-of-town guests (with over 30% affected to at least some degree in the past), inability to open windows in homes and vehicles (49%), problems letting pets out to play or exercise (over 19%), and even experiencing RT toxins that penetrated through a home or vehicle's climate control system (27%).

Additional, more profound and pervasive actions taken by some households are summarized in Table 6. As can be seen from the table, a considerable share of respondents, over 18%, actually moved away from the coast to escape RT effects during the time they have lived in the 5CR. In addition, five to six percent of survey participants report having sold a boat or other water sport equipment, having recently put their house or condominium on the market in an effort to move out of the area, and / or changed their job or retired early in reaction to RT air contamination.

In sum, this section of our survey demonstrates clearly and quantitatively that a large share of southwest Florida residents have been negatively affected by RT air contamination in the past. These effects manifest themselves in a variety of ways, from interference with outdoor activities, to causing health problems and serious economic consequences, such as forcing residents to move away from the coast, or make career changes. Given locals' pronounced preference for outdoor time, and the many avenues RT interferes with residents' everyday life, we would ex ante expect strong demand for a spatially and temporally refined RT air quality forecasting system, such as described above. This is confirmed by our analysis of the choice experiment data as discussed below in more detail.

The survey also asked about respondents familiarity with existing RT information sources as described above, after describing them in detail and providing web links for further information. As summarized in the first column of Table 7, between one third and one half of sampled households are *unaware* of MML's BCRS (Mote Marine Laboratory & Aquarium 2020), Florida Fish & Wildlife Conservation Commission's RT Current Status web site (Florida Office of Economic & Demographic Research 2020), and NOAA's Harmful Algal Bloom forecast (National Oceanic and Atmospheric Administration 2020), respectively. Another 30-35% of respondents have heard of these sources, but have not used them in the past. This leaves approximately 27% who actually access the BCRS and FWC sites, respectively, with a slightly smaller contingent (close to 22%) visiting NOAA's HAB site at least on few occasions per month (last column of table 7). In terms of usage frequency, 12-16% of the entire sample consults these sites one to four days per month, while eight to ten percent tap into these sources between five and 20 days / month, and a small contingent (one to two percent) rely on these information sources almost daily. In general, there appears to be room to boost public use of online sources related to RT current conditions and forecasts. The envisioned new system with its refined spatial and temporal forecasting features may be well suited to increase information uptake.[23]

All 502 participants completed all four of their assigned voting questions. Overall, 386 survey takers opted for the new system on at least one occasion, providing the following reasons for their vote: Lowered risk of exposure to RT air toxins (54%), ability to spend more time outdoors (32%), ability to better plan for outdoor activities (56%), ability for better timing of outdoor activities (52%), and ability to make better location choices for outdoor endeavors (51%). On the flip side, 134 (close to 27%) of respondents chose the SQ option of "no new system," that is existing information sources at zero cost on all four occasions. Of those, we identify 100 cases (close to 20% of the total sample) as "protest responses," given their follow-up rationale for serial SQ voting as one or more of the following: "The new system should be provided by the government at no cost" (60%), "the new system should be covered from existing taxes and fees" (40%), "the new system will slow down other RT efforts" (8%), and "I do not believe the new system is scientifically feasible" (6%).[24]

Additional invalid observations were identified via the end-of-section auxiliary questions given to all respondents. Results from these seven-tiered Likert-style questions are shown in Table C1 in the online appendix, with original response categories "strongly disagree, disagree, and "somewhat disagree" compiled into a single "disagree to some extent" category. An analogous aggregation was performed on the "agree" side of the response spectrum. As is evident from the table, the vast majority of participants felt that the survey information was useful and sufficient for decision-making, showed confidence in their votes, confirmed their perceived realism of the voting occasion, and did not feel coerced into voting one way or other. We flag the "disagree" responses to the voting realism and absence-of-coercion questions as invalid observations for subsequent analysis. Since three of these 37 cases were also protest responses, we retain a valid sample of 368 individuals for final analysis. With each person providing four voting responses, this translates into 1472 effective observations for econometric processing.

Estimation results

Estimation of the CL model requires the exact specification of the forecast attributes "spatial coverage" (`band`), "accuracy for the first 12 hours " (`acc1`), and "accuracy for the second 12 hours" (`acc2`), as well as the stipulated price (`bid`). We opt to express attributes as binary indicators for each level beyond the lowest setting, while treating the bid variable as a single (pseudo-) continuous regressor. As discussed in Henscher, Rose, and Greene (2005) and Holmes, Adamowicz, and Carlsson (2017), this ex ante allows for potential nonlinear attribute effects, which are confirmed by our estimation results. The implicit baseline policy scenario is thus a forecast system with band width of six miles and 50% accuracy throughout. In accordance with equation (2) we enter bids as positive amounts, which implies that the marginal-utility-of-income parameter λ is estimated with reversed sign.[25]

We choose a multivariate normal prior for coefficient vector $\boldsymbol{\beta}$ with a mean vector of zero and a diagonal covariance matrix with all variance terms equal to 100. This extremely vague prior places the bulk of informational burden on the data, as desired. For the MH step in the draws of $\boldsymbol{\beta}$ we set the scalar for the variance matrix in the multivariate-t proposal density τ to unity, and the degrees of freedom parameter ν to equal 30. These settings yield relatively high acceptance rates (70-75%) while keeping autocorrelation across draws at reasonably low levels.

In all cases, we use Geweke's (1992) split-mean diagnostics to assess convergence. This led to a choice of 150,000 - 200,000 burn-in draws depending on model. In addition, we keep track of autocorrelation via Inefficiency Factors (IEFs), as suggested in Chib (2001). For all estimation runs, IEF scores remain in the one to five range, suggesting acceptably low levels of autocorrelation. In all cases, we retain 100,000 parameter draws for final inference.[26]

Primary estimation results are given in Table 8 for both the full model using all choice sets (columns two to four) and the restricted model using only the choice set encountered first by a given respondent (columns five through seven). For each model the table gives the posterior mean, posterior standard deviation, and the proportion of the posterior distribution to the right of zero ($p > 0$). The latter statistic provides an at-a-glance assessment if a given variable's effect on outcome is predominantly positive ($p > 0$ is close to 1), negative ($p > 0$ is close to 0), or ambiguous ($p > 0$ is close to 0.5). The more posterior mass to one or other side of zero, the clearer is the signal for the direction of the effect of the corresponding coefficient.

The first column of the table lists the estimated coefficients. Recall that the implicit baseline scenario is a forecast with band width of six miles, and 50% accuracy for both the first and second 12-hour time window. Considering first the "all sets" model, we find that respondents overall strongly prefer even the least extensive and accurate new system over what is currently in place, as is evident from the numerically large and unambiguously negative sign of the "SQ" indicator. There is also considerable evidence that a 12-mile band is preferred to the six-mile version (with 87% of the posterior distribution of the corresponding indicator in the positive domain), and compelling evidence that higher accuracy for the first 12 hours is valued more highly than lower accuracy, ceteris paribus. The same holds for the second 12-hour time segment for the 75% accuracy indicator, while no clear signal emerges for the 100% level, with the corresponding posterior distribution centered close to zero ($p > 0 = 0.48$). It is possible that this reflects some degree of doubt that such a flawless forecast this far out is actually feasible. On the other hand, and given the larger coefficient estimates for the `acc1` indicators compared to their `acc2` counterparts, this lack of "significance" (in slight abuse of classical terminology) may simply indicate that accurate near-time forecasts are valued more highly than accuracy for time slots further into the future. The last row of the table summarizes the posterior distribution of the price coefficient. Reassuringly, this distribution is tightly centered around a distinctly negative mean, as dictated by economic theory.

The second set of columns in Table 8 give analogous results for the "first-set-only" model. As expected, with a 75% reduction in sample size (from 1472 to 368 observations) this model produces higher posterior standard deviations, and less discerning $p > 0$ statistics. Nonetheless, the salient estimation patterns observed for the full model remain: A distinctly negative coefficient for the SQ indicator, clear preference for higher accuracy for the first 12 hours, and a clearly negative price coefficient, with almost identical posterior mean as for the "all sets" model. We take this as indication of the absence of serious bias induced by any sequencing effects in the full model.

Role of experience with existing systems

We next examine to what extent experience with one or more of the existing RT information systems, as described above, may shape preferences for an improved forecast. To that effect we generate an aggregate 0/1 variable `use_exist` that indicates if a given household has used one of the existing systems in the past, as captured in Table 7. Overall, approximately 44% of our estimation sample falls into that category, leaving 56% of non-users.

At the same time we need to control for "outdoor avidity," that is preferences for outdoor activities, to guard against omitted variable problems. This is because households with relatively stronger preference for being outside will also be more likely to consult with existing RT information systems.[27] At the same time, more avid households will also have greater need, and thus likely higher WTP for improved forecasts. This would then be misinterpreted as stronger dissatisfaction with existing systems if outdoor avidity is not explicitly captured in the choice model. To provide this control we generate a new variable `hours_outside` that (approximately) measures the total number of hours spent outside by all household members in a typical, RT-free week, based on the binned categories given in Table 4.[28] We then include this variable alongside `use_exist` as interactions with the SQ indicator in an augmented choice model.

Results are given in Table 9. As can be seen from the second and third row of the table, both interaction terms are negative and highly significant, with essentially their entire posterior distribution to the left of zero ($p > 0 = 0$). This suggests that both outdoor avidity and usage of existing RT information systems increase the probability of voting for the new forecast, and thus WTP for the new system. Specifically, an additional hour spent outside in a typical week increases WTP by (0.013/0.047)=\$0.27, while having used the BCRS, FWC, or NOAA system in the past boosts WTP by a large amount relative to non-users, specifically (0.552/0.047)=\$11.75. We take this as direct indication that these existing systems are considered insufficient and less-than-effective in providing local households the information they need to optimize their outdoor activities. Clearly, there is room for improvement along these lines.

We also note that the augmented model leaves all other coefficients essentially unchanged, suggesting that attribute and income effects are not dependent on outdoor avidity or familiarity and / or satisfaction with existing information sources. Since our main interest in terms of welfare predictions rests with the "typical household" of the 5CR with less focus on the role of household characteristics or demographics, we opt for the generic baseline models given in Table 8 as the underlying specifications to generate WTP predictions, as discussed in the next section.

Value predictions

Following the econometric steps outlined above, the posterior distribution of the estimated coefficients was combined with all 12 plausible combinations of settings for bandwidth and accuracy to generate posterior predictive distributions of WTP for each forecast scenario. Results are captured in Table 10 for both the full and the first-set-only model. For each version the table shows the posterior predictive mean, as well as the lower and upper bound of the corresponding highest posterior density interval (HPDI).[29] As is evident from the table, posterior means are generally comparable between the two model versions, with the restricted model exhibiting wider HPDI ranges due to the substantially diminished sample size, as expected.

Focusing on the full model as our overall preferred specification, we observe that the typical five-county household is willing to pay over \$17 per year for even the least refined forecast version with a six-mile band and only 50% accuracy for both 12-hour segments.

This estimate increases to as much as $40-$45 per year for forecasts with higher accuracy, especially for the first 12 hours.[30]

Results including protest responses

As described above, we exclude problematic responses from our main estimation. To recall, these are "serial-SQ" voters that explicitly stated that they did not believe that they should pay for the new system, had doubts that the system was technically feasible (despite being assured otherwise), and/or worried that the new system would slow down other RT mitigation and prevention efforts (despite being assured otherwise). We also dropped individuals that explicitly disagreed that they felt confident about their vote, would vote the same way in an actual public referendum, and/or voted as if costs were real, in addition to those that disagreed with the statement that the survey let them make up their own mind. These exclusion criteria strike us as very reasonable, and actually quite lenient, in an effort to learn about residents' preferences in an unbiased fashion. That said, in the spirit of full transparency, and following recommendations in Johnston, Boyle, et al. (2017) we also report estimation results for the full sample of 502 survey-takers in online appendix Table C2.

As can be seen from the table, including problematic observations biases essentially all coefficients towards zero, as one would expect in this case, since the added "no" votes lack sensitivity to both forecast attributes and price. Most strikingly, the formerly strongly negative and highly significant SQ coefficient is now reduced to a magnitude close to zero, with a less discriminating signal on sign (based on the $p > 0$ statistic). In combination, this leads to substantially lower WTP estimates, given in online appendix Table C3. The most basic forecast is now valued at only $1.62 per year (compared to $17.37), and the most highly valued forecast elicits a WTP of only $26.24 (compared to $44.60).

The nature of the problematic responses captured in our survey (lack of voting realism, disregard of budgets, disagreement with "property rights," etc.) does not give confidence that anything can be learned about true underlying values for this sub-sample. As such, we consider our full-sample estimates (with protest observations included) as biased and uninformative, and the same holds for corresponding WTP predictions. Nonetheless, we will pay attention to the fact that 23% of initially targeted survey-takers were de facto excluded from our main analysis in our aggregation exercise below.

Aggregation

Approximately 835,000 households live in the five-county-area (e.g. U.S. Census Bureau 2021). Aggregating the annual per-household WTP figure of $17.37 for the most basic forecast, as shown in Table 10, produces an aggregate annual value of $14.5 million per year. If we conservatively exclude 23% of these households as potential protest cases whose preferences are simply not captured by our analysis, this reduces to $11.2 million per year.

It is difficult to determine how these estimated annual benefits stack up against expected costs of implementing and running the envisioned forecast, given the lack of comparable systems currently in operation. Petrolia et al. (2019) use an upper estimate for annual costs of $20.1 million for their current beach conditions information system along the entire gulf coast, from Florida to Texas. If we assume that underlying costs per location (e.g a single beach or one of our grid cells) are comparable across systems, and costs scale approximately linearly based on included main beaches, we could derive an upper bound of $20.1*(32/76) = $8.5 million, given that the Petrolia et al. (2019) study includes 76 beaches compared to the 32 beaches currently monitored via the BCRS along the 5CR coastline. In that sense, our envisioned forecast would generate positive societal benefits, even under extremely conservative assumptions.

Perhaps a more meaningful comparison can be made along the lines of research budgets. NOAA recently announced awarding an annual total of $15.2 million for HAB research nationwide (National Oceanic and Atmospheric Administration 2021). Approximately $4 million of this amount went to Florida and the Gulf of Mexico for new and ongoing HAB research.[31] These figures certainly appear low compared to our estimated values for just an RT forecast alone, and a much smaller geographic region.

Conclusion

We implement a survey-based choice experiment to estimate benefits to the residential population of five southwest Florida counties of improved near-term forecasts for air quality levels associated with blooms of RT. We consider this both a practical and feasible approach to obtain these values, given the many ways in which airborne red tide irritants can interfere with human activities and choices, and the corresponding multitude of potential mitigating actions - as also confirmed by our survey. Understanding these benefits, in monetary terms, is essential to determine the optimal allocation of resources for the provision of these services.

We find that annual benefits are of substantial magnitude, even for the least accurate and spatially extended version of our forecast scenarios. This suggests a real potential to reduce economic losses associated with RT blooms by refining existing RT information sources. Naturally, a more conclusive determination of societal net benefits would require better knowledge of the near-term and future costs of implementing this service over the time horizon considered by the policy maker. Nonetheless, our benefit estimates form an important input to this process. Our proposed Stated Preferences framework for the estimation of economic benefits of improved air quality forecasts is directly applicable to other HAB types that generate airborne toxins, and, more generally, any type of air quality problem with human exposure, adding to the paucity of work on this topic.

Naturally, some caveats are in order regarding our empirical findings and their interpretation. First, our forecast scenarios are restricted to deviate in accuracy by no more than one irritation level in either direction. It is possible that forecasts with "really bad errors" would be valued at lower levels by stakeholders. In that sense, our estimates might be considered upper bounds for more error-prone forecast versions. By the same token, the benefits we derive are likely conservative *lower bounds* for the full value of the envisioned forecast system to *all stakeholders* in the 5CR, such as businesses with outdoor seating, outfitters and charter companies, and event planners that rely on outdoor venues. Extending this work to include some of these other sectors of the affected population would be a fruitful extension of this research. We also note that the *upper end* of our estimates for a perfectly accurate forecast can be interpreted as - likely very conservative - lower bound of the actual annual *costs* of RT blooms to the surveyed population, as mitigating impacts with the help of better forecasts will lower, but certainly not erase all costs caused by RT.

On a final note, our proposed forecast was designed to provide an intuitive, user-friendly information environment for human decision-making. From a scientific perspective, the path to its implementation still presents numerous challenges, such as fully understanding the production and movement of *Karenia brevis* in the ocean and air, as well as the aerosolization processes of cells and toxins under varying environmental conditions. In that sense, our results stress the potential benefits to the local population of further investing in these scientific endeavors.

List of acronyms

[table 11 to be inserted in this section]

Notes

[1]A list of acronyms used in this paper is given in table 11.

[2]Petrolia et al. (2019) implement a Contingent valuation (CV) study to elicit the willingness-to-pay by residents of five U.S. gulf states to obtain access to a current-conditions information system for 76 gulf coast beaches. While their envisioned bundle of conditions includes RT respiratory risk, it does not specify any further attributes such as spatial and temporal coverage or accuracy, as is the case in our CE. More importantly, their stipulated service is a "now-cast," giving current conditions, while our system is described as an hourly forecast reaching 24 hours into the future.

[3]The only study we are aware of that focuses on the value of air quality forecasts is Garner and Thompson (2012). However, their stylized loss-cost model requires as input an assumption of lost societal benefits, rather than estimating them.

[4]In comparison Kirkpatrick et al. (2006) suggest a concentration of 100,000 cells per liter (cpl) as the toxin level that affects human health. This level is also referred to as a "typical bloom concentration" in Vargo et al. (2008). Concentrations of under 1,000 cpl are generally considered standard ambient levels of Karenia brevis (Nierenberg et al. 2010).

[5]For accessible introductions to RUM theory and applications see for example Freeman (2003), Train (2009), and Phaneuf and Requate (2017).

[6]The term "indirect utility" refers to an envelope function that reflects maximized utility over all other goods and services, subject to prices and a budget constraint. See for example Mas-Colell, Whinston, and Green (1995).

[7]An important implication of an unlabeled choice setting is that the much cited "Independence of Irrelevant Alternatives" (IIA) concern of CL's with *labeled* alternatives, which refers to the potential correlation of choice probabilities across the J options via unobservable preferences, is less of a concern in our case (Hausman and McFadden 1984; Greene 2012; Train 2009; Holmes, Adamowicz, and Carlsson 2017).

[8]This implies that for our application a significant SQ effect is a valid indication of the relative benefits or disadvantages of existing RT information sources compared to the new system, as perceived by respondents. This contrasts with many existing CL applications with explicit variable settings for the SQ alternative, where the emergence of a significant SQ indicator is often considered a form of undesirable "resistance to change" bias (Boxall, Adamowicz, and Moon 2009; Meyerhoff and Liebe 2009).

[9]Respondents were reminded in the survey instrument to treat each choice occasion as a free-standing, independent event. In addition, we also estimate a model that is based on only the first choice set presented to each individual, and thus free of potential sequencing effects and correlation via unobservables by design. The output from this reduced-sample model essentially mirrors results flowing from the full model, as discussed below.

[10]Existing examples of Bayesian estimation of CL models include Rossi, Allenby, and McCulloch (2005),

Train and Sonnier (2005), and Daziano (2013).

[11]This maximum WTP to obtain an improved level of an environmental commodity or service is also referred to as "Compensating Variation" (CV) in economic jargon (e.g. Freeman 2003; Phaneuf and Requate 2017).

[12]Specifically, the six-mile band comprises approximately 1,300 forecasting squares. This number is doubled for the 12-mile band.

[13]Examples include cell phone applications such as "Weather Radar Live" (Weather or Not Apps 2020), "Weather Live" (Apalon Apps 2020), and "Storm Shield" (E.W. Scripps Company 2020), all of which charge approximately \$20-\$25 per year. Our bid range also overlaps to a large extent with that chosen by Petrolia et al. (2019) for their beach conditions valuation study, though our range extends well beyond their highest bid of \$10. However, given our exclusively local population living near the epicenter of RT impacts as opposed to their more general population of beach-goers residing in five U.S. gulf coast states, this difference in upper cut-offs appears justified.

[14]Inclusion of the SQ option is also an econometric requirement given our random utility setup described above, which rests on differences across utilities, including baseline conditions. It is also consistent with best practices in stated preference research as discussed in Johnston, Boyle, et al. (2017)

[15]We use stata's *dcreate* function to accomplish this task, as well as for the grouping of choice sets into blocks.

[16]The full survey instrument is available from the authors upon request.

[17]As discussed in Johnston, Boyle, et al. (2017), section 4.5, the specification of the payment vehicle is essentially a balancing act of realism, credibility, and familiarity, with the most important feature of being binding for the entire relevant population to assure incentive compatibility. For the latter reason, we decided against the use of (ultimately voluntary) subscription fees, e.g. for phone apps. Furthermore, since Florida does not have a state income tax, sales taxes can be avoided via a change in purchasing behavior, and property taxes may not apply to renters, we opted for a vague definition of "new local taxes and fees," which is well within the scope of recommendations in Johnston, Boyle, et al. (2017) for situations such as ours. Recent examples in the stated preference literature that use similarly broad, but binding, "taxes and/or fees" vehicles include Johnston, Holland, and Yao (2016), Holland and Johnston (2017), Johnston, Schultz, et al. (2017), Parthum and Ando (2020), and Choi and Ready (2021). We further note that our stipulated payment vehicle was not met with resistance or protest in focus groups and pretest.

[18]Mitchell and Carson 1989 refer to such unintended predisposition of respondents towards specific responses as "implied value cues," and list these as one of the main sources of potential bias in stated preference research.

[19]This initial test sample was not used in final analysis.

[20]These deviations from census figures mirror those reported by Morgan, Larkin, and Adams (2010) for

their 2001 survey of 1006 randomly selected Manatee and Sarasota households to study a variety of impacts of RT on the local population.

[21]Specifically, the category of 1-4 hours was recoded to 2.5, the 5-8 hour bracket was recoded to 6.5, and so on. The highest category of >12 hours was conservatively coded as a flat "12."

[22]Morgan, Larkin, and Adams 2010 find RT impacts of comparable magnitude for a 2001 sample of 576 beach-going residents of Sarasota and Manatee counties, 70% of whom report having been forced to cut short, delay, or re-locate outings in the preceding 12 months.

[23]These awareness and usage percentages for Mote's BCRS are pronouncedly higher than those reported in Petrolia et al. (2019), where only 15% were aware of this service, and a mere 7% had actually used it in the past. However, as mentioned before, their underlying survey population comes from a much larger geographic region, and is likely far less familiar with MML and its research and public services related to RT than our sample of local residents, who are much more likely to have visited MML and / or have been exposed to MML outreach efforts regarding RT.

[24]In contrast, permissible "serial NO's" were identified as those related to budget considerations, or perceived relevance / usefulness / value of the new system, captured via one or more of the following stated reasons: "The new systems proposed in the questions were not worth the cost to me and my household" (71 cases), "My household simply can't afford to pay for a new forecasting system at this time" (35 cases), "I / my household rarely spend time outside" (14 cases), "I / my household can be very flexible in making outdoor choices - information on current conditions for local beaches is sufficient" (40 cases), "Florida red tide toxins don't bother me / my household much" (24 cases), "The forecasting system would not cover areas that are important to me" (7 cases), and "I would not use the new system often enough to warrant the cost" (55 cases).

[25]We also considered the inclusion of two-way interactions between attribute levels. However, these emerged as highly insignificant, and are thus omitted in our final specification.

[26]Estimation was performed using Matlab on a single Intel Xeon Gold 6136 3.00GHz processor. Run times ranged from 20 to 30 minutes, depending on model.

[27]This notion is strongly substantiated with a simple binary logit model relating `use_exist` to hours spent outside.

[28]To compute this continuous metric, we follow the same approach as for generating the last column in Table 4, then add up over all outdoor activities.

[29]As described inter alia in Koop (2003) the α-% HPDI is the smallest interval over the range of a given distribution that includes α % of the density mass. It is a common Bayesian statistic used to describe the spread of a given distribution, and/or confidence bounds for a given parameter of predictive construct of interest.

[30]The lower end of our value range mirrors the estimate of $16 per household and year found by Lazo

and Chestnut (2002) for the U.S. population at large for comparable improvements in weather forecasts offered by the National Weather Service. Our welfare estimates are also well-aligned with those produced by Petrolia et al. (2019), who report an annual WTP of approximately $12 per year for their current-conditions beach information system. One would expect our value predictions to be somewhat higher than theirs as our system constitutes a 24-hour forecast, thus providing even more room for adjustments and re-optimization of household outdoor activities compared to their now-cast, at least with respect to RT conditions.

[31]This information comes from an internal NOAA e-mail communication shared with / by co-author Fanara.

References

Adams, C.M., S. Larkin, P. Hoagland, and B. Sancewich. 2018. "Assessing the economic consequences of harmful algal blooms: A summary of existing literature, research methods, data, and information gaps." In *Harmful algal blooms: A compendium desk reference,* edited by S. E. Shumway, J. M. Burkholder, and S.L. Morton, 337–354. Wiley Blackwell.

Adams, R.M., K.J. Bryant, B.A. McCarl, D.M. Legler, J. O'Brian, A.R. Solow, and R.F. Weiher. 1995. "Value of improved long-term weather information." *Contemporary Economic Policy* 13:10–19.

Alcock, F. 2007. *An assessment of Florida red tide: Causes, consequences, and management strategies.* Technical report 1190. Mote Marine Laboratory.

Apalon Apps. 2020. "Weather Live." Accessed January 14, 2021. https://apps.apple.com/us/app/weather-live/id749083919.

Backer, L.C., L.E. Fleming, A. Rowan, Y.-S. Cheng, J. Benson, R.H. Pierce, J. Zaias, et al. 2003. "Recreational exposure to aerosolized brevetoxin during Florida red tide events." *Harmful Algae* 2:19–28.

Boxall, P., W.L. Adamowicz, and A. Moon. 2009. "Complexity in choice experiments: Choice of the status quo alternative and implications for welfare measurement." *Australian Journal of Agricultural and Resource Economics* 53:503–519.

Brand, L.E., and A. Compton. 2007. "Long-term increase in Karenia brevis abundance along the southwest Florida coast." *Harmful Algae* 6:232–252.

Brown, T.C. 2017. "Introduction to stated preference methods." In *A primer on nonmarket valuation,* edited by P.A. Champ, K. Boyle, and T.C. Brown, 133–186. Springer.

Chib, S. 2001. "Markov Chain Monte Carlo metods: Computation and inference." In *Handbook of Econometrics,* edited by J.J. Heckman and E. Leamer, 3569–3649. Elsevier.

Choi, D.S., and R. Ready. 2021. "Measuring benefits from spatially-explicit surface water quality improvements: The roles of distance, scope, scale, and size." *Resource and Energy Economics* 63:101108.

Corcoran, A., M. Dornback, B. Kirkpatrick, and A. Jochens. 2013. *A primer on Gulf of Mexico harmful algal blooms.* Technical report. Gulf of Mexico Alliance & Gulf of Mexico Coastal Ocean Observing System.

Daziano, R.A. 2013. "Conditional-logit Bayes estimators for consumer valuation of electric vehicles driving range." *Resource and Energy Economics* 35:429–450.

De Shazo, J.R., and G. Fermo. 2002. "Designing choice sets for stated preference methods: The effects of complexity on choice consistency." *Journal of Environmental Economics and Management* 44:123–143.

E.W. Scripps Company. 2020. "Storm Shield." Accessed January 14, 2020. https://apps.apple.com/us/app/storm-shield/id526831380.

Ferrini, S., and R. Scarpa. 2007. "Designs with *a priori* information for nonmarket valuation with choice experiments: A Monte Carlo study." *Journal of Environmental Economics and Management* 53:342–363.

Fleming, L.E., B. Kirkpatrick, L.C. Backer, C.J. Walsh, K. Nierenberg, J. Clark, A. Reich, et al. 2011. "Review of Florida red tide and human health effects." *Harmful Algae* 10:224–233.

Florida Fish and Wildlife Conservation Commission. 2020. "Red tide current status." Accessed January 20, 2021. https://myfwc.com/research/redtide/statewide/.

Florida Office of Economic & Demographic Research. 2020. "County Profiles." Accessed January 16, 2021. http://edr.state.fl.us/content/area-profiles/county/index.cfm.

Freebairn, J.W., and J.W. Zillman. 2002. "Economic benefits of meteorological services." *Meteorological Applications* 9:33–44.

Freeman, A.M. III. 2003. *The measurement of environmental and resource values.* 2nd. Resources for the Future Press.

Garner, G.G., and A.M. Thompson. 2012. "The value of air quality forecasting in the Mid-Atlantic region." *Weather, Climate, and Society* Jan.69–79.

Greene, W.H. 2012. *Econometric Analysis.* 7th. Prentice Hall.

Gulf of Mexico Coastal Ocean Observing System. 2020. "Experimental red tide respiratory forecast." Accessed January 20, 2021. https://gcoos.org/red-tide-forecast/.

Hastings, W.K. 1970. "Monte Carlo sampling methods using Markov Chains and their applications." *Biometrika* 57:97–109.

Hausman, J.A., and D. McFadden. 1984. "Specification tests for the multinomial logit model." *Econometrica* 52:1219–1240.

Henscher, D.A., J.M. Rose, and W.H. Greene. 2005. *Applied choice analysis: A primer.* 1st. Cambridge University Press.

Hoagland, P., M. Anderson, Y. Kaoru, and W. White. 2002. "The economic effects of harmful algal bloom in the United States: Estimates, assessment issues, and information needs." *Estuaries* 25:819–837.

Hoagland, P., D. Jin, L.Y. Polansky, B. Kirkpatrick, G. Kirkpatrick, L.E. Fleming, A. Reich, S.W. Watkins, S.G. Ullmann, and L.C. Backer. 2009. "The costs of respiratory illnesses arising from Florida gulf coast Karenia brevis blooms." *Environmental Health Perspectives* 117:1239–1243.

Holland, B.M., and R.J. Johnston. 2017. "Optimized quantity-within-distance models of spatial welfare heterogeneity." *Journal of Environmental Economics and Management* 85:110–129.

Holmes, T.P., W.L. Adamowicz, and F. Carlsson. 2017. "Choice experiments." In *A primer on nonmarket valuation,* edited by P.A. Champ, K. Boyle, and T.C. Brown, 133–186. Springer.

Jin, D., and P. Hoagland. 2008. "The value of harmful algal bloom predictions to the nearshore commercial shellfish fishery in the Gulf of Maine." *Harmful Algae* 7:772–781.

Jin, D., S. Moore, D. Holland, L. Anderson, W.-A. Lim, D. Kim, S. Jardine, S. Martino, F. Gianella, and K. Davidson. 2020. "Evaluating the economic impacts of harmful algal blooms: Issues, methods, and examples." In *GlobalHAB: Evaluating, reducing and mitigating the cost of harmful algal blooms: A compendium of case studies,* edited by V.T. Trainer, 5–41.

Johnson, F.R., B. Kanninen, M. Bingham, and S. Özdemir. 2007. "Experimental design for stated choice studies." In *Valuing environmental amenities using stated choice studies,* edited by B.J. Kanninen, 159–202. Springer.

Johnson, S.R., and M.T. Holt. 1997. "The value of weather information." In *Economic value of weather and climate forecasts,* edited by R.W. Katz and A.H. Murphy, 75–107.

Johnston, R.J., K.J. Boyle, J.B. Adamowicz, J. Bennett, R. Brouwer, T.A. Cameron, W.M. Hanemann, et al. 2017. "Contemporary guide for stated preference studies." *Journal of the Association of Environmental and Resource Economists* 4:319–405.

Johnston, R.J., B.M Holland, and L. Yao. 2016. "Individualized geocoding in stated preference questionnaires: Implications for survey design and welfare estimation." *Land Economics* 92:737–759.

Johnston, R.J., E.T. Schultz, K. Segerson, E.Y. Besedin, and M. Ramachandran. 2017. "Biophysical causality and environmental preference elicitation: Evaluating the validity of welfare analysis over intermediate outcomes." *American Journal of Agricultural Economics* 99:163–185.

Johnston, R.J., T.F. Weaver, L.A. Smith, and S.K. Swallow. 1995. "Contingent valuation focus groups: Insights from ethnographic interview techniques." *Agricultural and Resource Economics Review* 24:56–69.

Kirkpatrick, B., L.E. Fleming, L.C. Backer, J.A. Bean, R. Tamer, G. Kirkpatrick, T. Kane, et al. 2006. "Environmental exposures to Florida red tides: Effects on emergency room respiratory diagnoses admissions." *Harmful Algae* 5:526–533.

Kirkpatrick, B., Fleming L.E., J.A. Bean, K. Nierenberg, L.C. Backer, Y.S. Cheng, R. Pierce, et al. 2011. "Aerosolized red tide toxins (Brevetoxins) and asthma: Continued health effects after 1 hour beach exposure." *Harmful Algae* 10:138–143.

Kirkpatrick, B., R. Pierce, Y.S. Cheng, M.S. Henry, P. Blum, S. Osborn, K. Nierenberg, et al. 2010. "Inland transport of aerosolized Florida red tide toxins." *Harmful Algae* 9:186–189.

Koop, G. 2003. *Bayesian Econometrics.* Wiley.

Krupnick, A., and W.L. Adamowicz. 2007. "Supporting questions in stated-choice studies." In *Valuing environmental amenities using stated choice studies,* edited by B.J. Kanninen, 43–65. Springer.

Larkin, S.L., and C.M. Adams. 2007. "Harmful algal bloom and coastal business: Economic consequences in Florida." *Society and Natural Resources* 20:849–859.

Lazo, J.K., and L.G. Chestnut. 2002. *Economic value of current and improved weather forecasts in the U.S. household sector.* Technical report SC10050. Stratus Consulting Inc.

Mas-Colell, A., M.D. Whinston, and J.R. Green. 1995. *Microeconomic Theory.* Oxford University Press.

McFadden, D. 1974. "Conditional logit analysis of qualitative choice behavior." In *Frontiers in Econometrics,* edited by P. Zarembka, 105–142. Academic Press.

Meyerhoff, J., A. Bartczak, and U. Liebe. 2012. "Protestor or non-protestor: A binary state?" *Australian Journal of Agricultural and Resource Economics* 56:438–454.

Meyerhoff, J., and U. Liebe. 2009. "Status quo effect in choice experiments: Empirical evidence on attitudes and choice task complexity." *Land Economics* 85:515–528.

Mitchell, R.C., and R.T. Carson. 1989. *Using surveys to value public goods: The contingent valuation method.* Resources for the Future, Washington, D.C.

Morgan, K.L., S.L. Larkin, and C.M. Adams. 2008. *Public costs of Florida red tides: A survey of coastal managers.* Technical report 08-1, February, 2008. Florida Agricultural Market Research Center, Food, Resource Economics Department, Institute of Food, and Agricultural Sciences, University of Florida, Gainesville, Florida.

———. 2009. "Firm-level economic effects of HABs: A tool for business loss assessment." *Harmful Algae* 8:212–218.

———. 2010. "Red tide and participation in marine-based activities: Estimating the response of Southwest Florida residents." *Harmful Algae* 9:333–341.

———. 2011. "Empirical analysis of media versus environmental impacts on park attendance." *Harmful Algae* 9:852–859.

Mote Marine Laboratory & Aquarium. 2020. "Beach Conditions Reporting System." Accessed January 14, 2021. https://visitbeaches.org/.

National Centers for Coastal Ocean Science. 2020. "NCCOS Supports Expansion of Red Tide Respiratory Forecast in Florida and Texas." Accessed January 20, 2021. https://coastalscience.noaa.gov/news/nccos-supports-expansion-of-red-tide-respiratory-forecast-in-florida-and-texas/.

National Ocean Service. 2019. "Fall 2018 red tide event that affected Florida and the gulf coast." Accessed January 20, 2021. https://oceanservice.noaa.gov/hazards/hab/florida-2018.html.

National Oceanic and Atmospheric Administration. 2020. "Gulf of Mexico harmful algal bloom forecast." Accessed January 19, 2021. https://tidesandcurrents.noaa.gov/hab/gomx.html.

———. 2021. "NOAA Awards $15.2M for Harmful Algal Bloom Research." Accessed December 13, 2021. https://oceanservice.noaa.gov/news/oct21/2021-hab-awards.html.

Nierenberg, K., M. Byrne, L.E. Fleming, W. Stephan, A. Reich, L.C. Backer, E. Tanga, D.R. Dalpra, and B. Kirkpatrick. 2010. "Florida red tide perceptions: Residents versus tourists." *Harmful Algae* 9:600–606.

Nierenberg, K., A. Reich, R. Currier, B. Kirkpatrick, L. Backer, R. Stumpf, L. Fleming, and G. Kirkpatrick. 2009. "Beaches and HABs: Successful expansion of the Florida red tide reporting system for protection of public health through community education and outreach." *Florida Journal of Environmental Health* 203:18–24.

Nyumba, T.O., K. Wilson, C.J. Derrick, and N. Mukherjee. 2017. "The use of focus group discussion methodology: Insights from two decades of application in conservation." *Methods in Ecology and Evolution* 9:20–32.

Parthum, B., and A.W. Ando. 2020. "Overlooked benefits of nutrient reduction in the Mississippi River Basin." *Land Economics* 96:589–607.

Petrolia, D.R., J. Penn, R. Quainoo, R.H. Caffey, and J.M. Fannin. 2019. "Know thy beach: Values of beach condition information." *Marine Resource Economics* 34:331–359.

Phaneuf, D.J., and T. Requate. 2017. *A course in environmental economics: Theory, policy, and practice.* Cambridge University Press.

Plaas, H.E., and H.W. Paerl. 2021. "Toxic cyanobacteria: A growing threat to water and air quality." *Environmental Science and Technology* 55:44–64.

Rollins, S.K., and J. Shaykewich. 2003. "Using willingness-to-pay to assess the economic value of weather forecasts for multiple commercial sectors." *Meteorological Applications* 10:31–38.

Rose, J.M., and C.J. Bliemer. 2009. "Constructing efficient stated choice experimental designs." *Transport Reviews* 29:587–617.

Rossi, P.E., G.M. Allenby, and R. McCulloch. 2005. *Bayesian Statistics and Marketing.* John Wiley & Sons.

Solow, A.R., R.M. Adams, K.J. Bryant, D.M. Legler, J. O'Brian, B.A. McCarl, and R.F. Weiher. 1998. "The value of ENSO forecasts: The case of U.S. agriculture." *Climatic Change* 39:47–60.

Stumpf, R.P., Yizhen Li, B. Kirkpatrick, R.W. Litaker, K.A. Hubbard, R.D. Currier, K.K. Harrison, and M.C. Tomlinson. 2022. "Quantifying Karenia Brevis bloom severity and respiratory irritation impact along the shoreline of Southwest Florida." *Plos One* 17:e0260755.

Swait, J., and W.L. Adamowicz. 2001. "The influence of task complexity on consumer choice: A latent class model of decision strategy switching." *Journal of Consumer Research* 28:135–148.

Train, K. 2009. *Discrete Choice Methods with Simulation.* 2nd. Cambridge University Press.

Train, K., and G. Sonnier. 2005. "Mixed logit for bounded distributions of correlated partworths." In *Applications of simulation methods in environmental and resource economics,* edited by R. Scarpa and A. Alberini, 117–134.

U.S. Census Bureau. 2021. "Sarasote County households and families." Accessed December 13, 2021. https://data.census.gov/cedsci/table?q=Sarasota%5C%20county%5C%20Households%5C&tid=ACSST1Y2019.S1101.

Vargo, G.A., C.A. Heil, K.A. Fanning, L.K. Dixon, M.B. Neely, K. Lester, D. Ault, et al. 2008. "Nutrient availability in support of Karenia brevis blooms on the central West Florida shelf: What keeps Karenia blooming?" *Continetal Shelf Research* 28:73–98.

Vossler, C.A., M. Doyon, and D. Rondeau. 2012. "Truth in consequentiality: Theory and field evidence on discrete choice experiments." *American Economic Journal: Microeconomics* 4:145–171.

Weather or Not Apps, LLC. 2020. "NOAA Weather Radar Live: Clime." Accessed January 14, 2021. https://apps.apple.com/us/app/noaa-weather-radar-live-clime/id749133753.

Wei-Haas, M. 2018. “Red tide is devastating Florida’s sea life. Are humans to blame?” National Geographic - Science & Innovation News. Accessed January 20, 2021. https://www.nationalgeographic.com/environment/2018/08/news-longest-red-tide-wildlife-deaths-marine-life-toxins/.

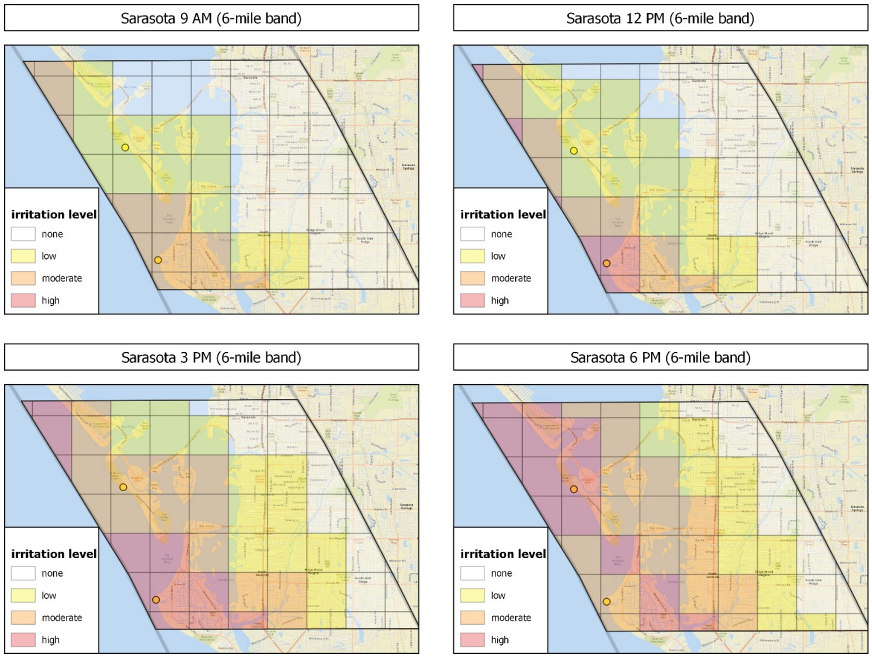

Figure 1: Spatial snapshots of the new forecasting system at different points in time for a given calendar day (1 square mile grid)

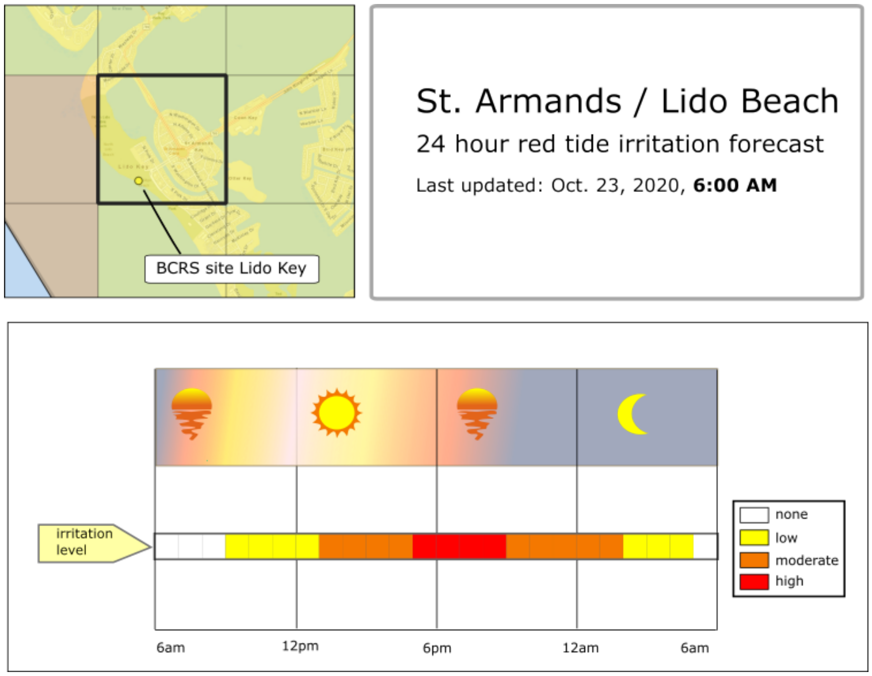

Figure 2: Example of square-specific forecast

Table 1: Attributes and levels

attribute	description	level settings
spatial coverage ("band")	width of coverage band along five-county coast, miles	6, 12
accuracy, first 12 hrs	% of correctly forecasted hourly irritation levels out of 12, first 12 hours counting from most recent update	50, 75, 100
accuracy, second 12 hrs	% of correctly forecasted hourly irritation levels out of 12, second 12 hours counting from most recent update	50, 75, 100
price ("bid")	annual tax per household	\$5, \$15, \$25, \$35

Table 2: Sample demographics

demographic	%	mean	std	min	max	obs.	census*
respondent-specific:							
female	59.76%					502	51.37%
age		56.35	17.85	18	92	502	
18-34	15.94%						19.38%
35-64	40.43%						42.69%
65+	43.63%						37.93%
HS diploma or higher	97.61%					502	89.58%
BA or higher	46.22%					502	30.97%
						502	
years lived at current address		8.74	8.10	0	47	502	
years lived in current county		14.15	11.93	0	50	500	
years lived in 5-county region		14.73	12.38	0	62	501	
HH-specific:							
HH size		2.27	1.25	1	14	501	2.29
family members under age 7	7.77%	0.16	0.76	0	11	502	
family members 7-18	14.74%	0.25	0.74	0	6	502	17.09%**
family members with respiratory conditions	28.86%	0.38	0.70	0	5	499	
HH income <$50,000	36.06%						41.78%
$51,000 - $100,000	39.04%						31.48%
$101,000 - $150,000	14.74%						13.99%
$150,000 - $200,000	4.98%						5.49%
>$200,000	5.18%						7.25%

*source: Florida Office of Economic & Demographic Research.
**% of households with children < 17 years of age.
N = 502.

Table 3: Participation in outdoor activities, past 12 months

activity	% of respondents yes	no	not sure
beach activities (walking, jogging, etc.)	72.71	26.29	1.00
water activities (swimming, snorkeling, etc.)	56.57	42.23	1.20
non-motorized water sports (kayaking, SUP, etc.)	21.91	76.49	1.59
motorized water sports (boat, jet-ski, etc.)	25.50	72.71	1.79
fishing / harvesting	28.69	69.32	1.99
outdoor sports (participant or spectator)	30.28	67.93	1.79
outdoor school activities (participant or spectator)	12.95	84.86	2.19
trail / road activities (jogging, biking, etc.)	57.97	40.04	1.99
park / picnic activities (BBQ, gatherings, etc.)	49.00	49.20	1.79
special outdoor events (participant or spectator)	47.21	51.20	1.59
professional or volunteer outdoor activities	13.35	84.66	1.99
other	4.58	80.68	14.74
any outdoor activity	88.05		

SUP = stand-up paddleboarding.
BBQ = barbecueing.
N = 502.

Table 4: Time spent outside, typical week

activity	hours per week 0	1-4	5-8	9-12	>12 / 13-24	>24	tot. hrs.
	% of respondents						
beach activities (walking, jogging, etc.)	35.06	45.42	13.35	2.99	3.19	-	2.70
water activities (swimming, snorkeling, etc.)	44.42	42.23	9.16	2.19	1.99	-	2.12
non-motorized water sports (kayaking, SUP, etc.)	77.49	16.33	3.98	1.39	0.80	-	0.91
motorized water sports (boat, jet-ski, etc.)	76.29	16.53	4.38	2.19	0.60	-	1.00
fishing / harvesting	71.91	22.51	3.19	1.20	1.20	-	1.04
outdoor sports (participant or spectator)	67.93	20.32	5.58	3.19	2.99	-	1.56
outdoor school activities (participant or spectator)	83.86	10.56	3.59	1.39	0.60	-	0.72
trail / road activities (jogging, biking, etc.)	37.25	40.24	16.73	3.78	1.99	-	2.73
park / picnic activities (BBQ, gatherings, etc.)	55.18	38.65	4.18	1.99	0.00	-	1.45
special outdoor events (participant or spectator)	62.35	33.07	3.19	1.20	0.20	-	1.18
professional or volunteer outdoor activities	84.26	10.96	3.59	0.60	0.60	-	0.64
other outdoor activities	91.63	3.19	1.59	1.39	2.19	-	0.59
all outdoor activities							16.64
balcony, deck, yard	8.17	31.47	22.51	17.33	13.15	7.37	8.21

SUP = stand-up paddleboarding.
BBQ = barbecueing.
tot. hrs. = total hours/week across all activities using category midpoints, averaged across households.
N = 502.

Table 5: RT air quality effects experienced since living in research area

effect	never	sometimes	often	almost daily	can' recall / N/A	affected
outdoor activities:			% of respondents			
cancel / postpone	35.06	40.04	8.57	7.97	8.37	56.58
shorten	35.06	39.84	8.76	8.96	7.37	57.56
re-locate	44.82	30.88	8.76	7.57	7.97	47.21
around house outside activities:						
cancel / postpone	53.98	29.68	6.77	3.98	5.58	40.43
shorten	50.20	29.88	7.97	5.38	6.57	43.23
health effects:						
irritation (but no doctor)	37.05	40.84	9.56	7.57	4.98	57.97
severe irriation, see doctor	82.07	6.37	1.79	3.39	6.38	11.55
bothered / sickened by dead fish smell	45.62	32.87	10.76	4.78	5.97	48.41
other effects:						
guests cancelled visits	60.16	22.71	5.18	2.19	9.76	30.08
unable to open windows (home or car)	45.42	30.08	9.36	9.56	5.57	49.00
unable to let pets out	58.37	10.36	4.98	3.98	22.31	19.32
toxins enered home / car via A/C system	61.55	19.32	4.38	3.39	11.36	27.09

N/A = not applicable.
N = 502.

Table 6: Additional actions taken in response to RT

action	yes	no	can' recall / N/A
	% of respondents		
moved away from coast	18.33	65.54	16.13
sold boat / water sport equipment	6.57	67.33	26.09
put house / condo on market	5.78	75.9	18.32
changed job / retired early	6.37	74.3	19.32

N/A = not applicable.
N = 502.

Table 7: Awareness and usage of existing RT online sources

source	never heard	heard, but never used	useage frequency (days /month) 1-4 days	5-10 days	11-20 days	21-30 days	some use
			% of respondents				
Mote's BCRS	42.43	30.88	16.33	5.98	2.59	1.79	26.69
FWC's RT web site	37.85	34.86	15.94	5.78	3.78	1.79	27.29
NOAA's HAB forecast	45.22	33.27	12.35	4.58	2.99	1.59	21.51

BCRS = Beach Conditions Reporting System.
FWC = Florida Fish & Wildlife Conservation Commission.
RT = RT.
NOAA = National Oceanic and Atmospheric Administration.
HAB = Harmful Algal Bloom.
N = 502.

Table 8: Estimation results

	all sets			first set only		
variable	mean	std.	p(> 0)	mean	std.	p(> 0)
SQ	-0.803	0.114	0.000	-1.053	0.216	0.000
band=12	0.110	0.100	0.866	0.049	0.200	0.595
acc1 = 75%	0.265	0.144	0.968	0.304	0.279	0.860
acc1=100%	0.942	0.166	1.000	1.091	0.330	1.000
acc2=75%	0.223	0.101	0.987	-0.106	0.207	0.308
acc2=100%	-0.007	0.138	0.479	-0.339	0.277	0.111
price (bid)	-0.047	0.005	0.000	-0.049	0.010	0.000

all sets = full valid sample (N=368, n=1472).
first set only: N=n=368.
mean = posterior mean / std. = posterior standard deviation.
$(p > 0)$ = proportion of posterior distribution exceeding zero.

Table 9: Estimation results for augmented choice model

	mean	std.	p(> 0)
SQ	-0.399	0.137	0.002
SQ * use_exist	-0.552	0.146	0.000
SQ * hrs_outside	-0.013	0.006	0.004
band=12	0.114	0.096	0.881
acc1 = 75%	0.265	0.141	0.970
acc1=100%	0.940	0.158	1.000
acc2=75%	0.222	0.099	0.988
acc2=100%	-0.001	0.134	0.503
price (bid)	-0.047	0.005	0.000

full valid sample (N=368, n=1472).
mean = posterior mean / std. = posterior standard deviation.
($p > 0$) = proportion of posterior distribution exceeding zero.

Table 10: WTP estimates ($'s per HH and year)

forecast scenario			all sets			first set only		
band	acc.1	acc.2	mean	low	high	mean	low	high
6	50	50	17.37	11.61	23.08	22.36	10.50	35.79
6	75	50	23.01	17.71	28.45	28.45	16.76	41.20
6	75	75	27.83	22.21	34.21	26.17	14.42	38.43
6	100	50	37.52	32.03	43.44	44.66	32.35	59.71
6	100	75	42.34	36.30	48.92	42.38	30.24	55.63
6	100	100	37.37	31.67	43.24	37.35	26.67	49.03
12	50	50	19.62	14.58	24.78	22.83	12.48	33.14
12	75	50	25.26	20.07	30.33	28.92	18.59	39.92
12	75	75	30.08	25.00	35.09	26.64	16.79	36.86
12	100	50	39.78	34.68	44.97	45.13	33.78	56.86
12	100	75	44.60	39.65	49.85	42.85	32.67	52.76
12	100	100	39.62	34.48	44.98	37.82	27.73	48.37

all sets = full valid sample (N=368, n=1472).
first set only: using only first choice set shown to respondent (N=n=368).
mean = posterior mean.
low [high] = lower [upper] bound of 95% highest posterior density interval.

Table 11: List of acronyms and their definition

acronym	definition
5CR	5 county region (our research area)
BCRS	beach conditions reporting system
CE	choice experiment
CL	conditional logit
CV	compensating variation (a theoretical construct in economic welfare theory)
EV	Exteme Value (a statistical distribution)
FGC	Florida gulf coast
FWC	Florida Fish & Wildlife Conservation Commission
HAB	harmful algal bloom
HPDI	highest posterior density interval (a Bayesian tool to characterize posterior spread)
IEF	inefficiency factor (a Bayesian assessment tool for serial correlation)
IIA	independence of irrlevant alternatives (a comon issue in RUMs)
IRB	institutional review board
IUF	indirect utility function
MH	Metropolis-Hastings (a Bayesian estimation method)
MLE	maximum likelihood estimation
MML	Mote Marine Laboratory and Aquarium
NOAA	National Oceanic and Atmospheric Administration
PPD	posterior predictive density (a Bayesian construct)
RT	red tide
RUM	random utility model
SP	stated preference
SQ	status quo
VT	Virginia Tech
WTP	willingness to pay