Title
Exploring Data in Environmental Information Programs

Author
Kohli, Aanchal

Publication Date
2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Exploring Data in Environmental Information Programs

A dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Environmental Science and Engineering

by

Aanchal Kohli

2017

ABSTRACT OF THE DISSERTATION

Exploring Data in Environmental Information Programs

by

Aanchal Kohli

Doctor of Environmental Science and Engineering

University of California, Los Angeles, 2017

Professor Magali A. Delmas, Chair

Environmental information programs have the potential to increase public awareness of environmental pollutants and associated health risks, and ultimately lead to the adoption of healthier behaviors and environmentally-friendly practices. In this dissertation, I study the role of information from such programs to mitigate environmental pollution and improve public health protection. My research explores information from two environmental information programs developed by the Environmental Protection Agency (EPA) – air quality (AQ) information from the AirNow program and toxic chemical releases information from the Toxic Release Inventory (TRI). One focus of this research is to better understand how the public engages with AQ information from the AirNow program through a mobile application (commonly called app). We developed an AQ app, AirForU, using data from the AirNow program, recruited about 3000 app users and studied their engagement with the app. Groups that are disproportionately affected by air pollution were more engaged with the app than the general public, however, engagement

dropped over time for most app users. Highly engaged app users adopted health protective measures against air pollution directly as a result of the information provided in the app. We tested the effect of various air pollution messages to boost engagement with AQ information using two methods -through an online survey and through a field experiment conducted via the app. Survey results suggest that messages with a strong fear appeal might not be effective at engaging the public in the case of air pollution. For the app experiment, the messages were effective to re-engage less engaged app users, but the content of the message mattered less. Results from this study suggest that targeted messages and timely reminders positively influence engagement with AQ information. Another focus of this research was to develop a robust environmental and economic performance index to evaluate facilities reporting to the TRI. A novel method, data envelopment analysis, capable of addressing irregularities in TRI data was successfully used to develop this index and allow easy comparison of facility performance within industrial sectors. This method of analysis could be used to easily communicate and evaluate environmental performance for individual facilities by both the public and the facilities themselves. Both research efforts were aimed at improving the effectiveness of environmental information programs.

The dissertation of Aanchal Kohli is approved.

Keith D. Stolzenbach

Adrienne G. Lavine

Yifang Zhu

Magali A. Delmas, Committee Chair

University of California, Los Angeles

2017

This dissertation is dedicated to

My parents, Radha and Krishan Kohli, for giving me the opportunity to pursue my dreams

My sisters, Sonia, Saveena, Neha and Ankita, for their incredible love and support

and

My mentor, Alexandra Smithmixter, for being my guiding light

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ALA	American Lung Association
AP	Air Pollution
App	Mobile Application
AQ	Air Quality
AQI	Air Quality Index
CDC	Centers for Disease Control and Prevention
CEI	Composite Environmental Index
DEA	Data Envelopment Analysis
EPCRA	Emergency Planning and Community Right-to-Know Act
GR	Gross Revenue
MTurk	Amazon Mechanical Turk
NAAQS	National Ambient Air Quality Standards
PWM	Percentage of Waste Managed through Recycling, Energy Recovery and Treatment
QTTR	Quantity of Total Toxic Releases
TRI	Toxics Release Inventory
TTTR	Toxicity of Total On-site Toxic Releases
US EPA	United States Environmental Protection Agency
UN	United Nations
WHO	World Health Organization

Acknowledgments

I would like to thank my advisor and my committee members for their guidance and support. I stumbled upon the Environmental Science and Engineering (ESE) program unexpectedly and I will be forever grateful to Professor Stolzenbach for giving me the opportunity to be part of this program. Professor Lavine supported me when I had given up on graduate school and helped me get into the ESE program. She has continued to support me in the ESE program and I am extremely grateful for that. She is an inspiring role model. I would also like to thank for Professor Zhu for her excellent feedback and comments.

I am incredibly grateful to my advisor, Professor Magali Delmas. Magali always made time to meet with me; even while I was San Francisco for my internship. She encouraged me to achieve much more than I thought I was capable of. While working with her I have learned many valuable skills beyond research such as collaborating with different scholars, presenting my research to different audiences, building partnerships, managing teams and writing grants. Chapter 2, based on the AirForU air quality app research study, is a manuscript in preparation for publication to be co-authored with my advisor. My advisor led the AirForU research study. We developed the conceptual design of the AirForU app and the email experiment. I managed a team of two app developers and oversaw the development of the iPhone and Android versions, managed the data, conducted the email experiment and analyzed the data.

I would like to thank Dr. Peng Zhou, professor at Nanjing University of Aeronautics and Astronautics, Nanjing, China. While he was a visiting scholar at UCLA, we had the opportunity

to collaborate on a research project, which is presented in this dissertation. Chapter 3 is a version of our recently published work (Zhou, P., Delmas, M. A., & Kohli, A. (2017). Constructing meaningful environmental indices: A nonparametric frontier approach, *Journal of Environmental Economics and Management, 85*, 21–34. http://doi.org/10.1016/j.jeem.2017.04.003). Approval has been obtained from Dr. Zhou and Dr. Delmas to include this paper in my dissertation. My advisor, Dr. Delmas, and I led a team of undergraduate steniors as part of the EPA Toxic Release Inventory (TRI) University Challenge where we gathered and analyzed toxic chemicals data and economic performance data to develop an environmental performance index. Dr. Zhou built upon and refined the index using a new method. I conducted a literature review of over 50 studies based on TRI data.

Besides my professors, I also received academic guidance from a number of postdocs working at the Institute of the Environmental and Sustainability. I would like to thank Jinghui Lim, Stephen Locke and Victor Chen for answering many of my questions. In particular, Jinghui's support with Stata and data analysis was very helpful. I would also like to thank Qingwei Lan and Walter Qian for being amazing app developers.

I would like to thank UCLA, the ESE department and my advisor for the financial support that I have received over the course of my education. I would like to thank Professor Donald Browne for giving me the opportunity to be a Teaching Assistant for Engineering Ethics. I would also like to thank the UCLA Center for Occupational and Environmental Health for the grants I received from them in 2016 and 2017.

Lastly, I would like to thank my parents, Radha and Krishan Kohli who have always believed in me and supported me unconditionally. They have given me the greatest gift in life, an excellent education and the means to be an independent woman. They sacrificed a lot for my sisters and I to have opportunities in life that they never had. I would like to thank my sisters, Savy, Anki, Golu and Sonia for their incredible support throughout my education. They inspired, supported, encouraged and uplifted me through the ups and downs of my doctorate. I would also like to thank my mentor, Alex, for guiding me through this process. I would also like to thank my friends at UCLA for their incredible support and for making this journey enjoyable. A special thanks to Jon and Gaby who were always there for me, and to my ESE cohort members, Yareli and Kazue, who supported me during my struggles.

VITA

EDUCATION

M.S. Mechanical Engineering, UCLA 2013

B.S. Mechanical Engineering, UCLA 2009

PUBLICATIONS

Constructing Meaningful Environmental Indices: A Nonparametric Frontier Approach, Journal of Environmental Economics and Management, 85 (2017) 21-34, Authors: Peng Zhou, Magali Delmas and Aanchal Kohli

Environmental Protection Agency (EPA) Toxic Release Inventory University Challenge Report 2014-2015, Authors: Magali Delmas, Aanchal Kohli et al.

Environmental Protection Agency (EPA) Toxic Release Inventory University Challenge Report 2013-2014, Authors: Magali Delmas, Aanchal Kohli et al.

PRESENTATIONS

AirForU: Using an App to Increase Awareness of Air Pollution and Drive Health Protection; Southern California Regional Conference and General Assembly: Palm Springs, May 4-5, 2017

AirForU: Using an App to Increase Awareness of Air Pollution and Drive Health Protection; Going Global: Collaborative Visions for a Sustainable Future, UCLA, April 21, 2017

AirForU: Using mobile technology to Increase Awareness of Air Pollution and Drive Health Protection; Earth Now: Earth 2050, UCLA, October 19, 2016

An Air Quality App for Increasing Awareness of Air Pollution and Driving Health Protection; Spring Green Expo, Metropolitan Water District of Southern California, April 28, 2016

AirForU, An Air Quality Mobile Application; Youth for the Environment and Sustainability, Berkeley, 2016

HONORS AND AWARDS

UCLA Center for Environmental and Occupational Health Grant, *2016 & 2017*

Malcolm-Stacey Fellowship Award, *2015*

Barbara Maida Award for Outstanding Female Graduate Student, *2014*

UCLA Graduate Student Fellowship, *2013*

PROFESSIONAL EXPERIENCE

Project Coordinator, The Energy Coalition, 2015

Air Quality Engineer, Bay Area Air Quality Management District, 2015-2016

I. Introduction

1 Environmental Legislation

Decades of industrialization, population growth and increases in the demand for basic resources such as water, food and energy have led to the deterioration of the environment in developed and developing nations. In the US, many cities are still designated a non-attainment status with respect to national ambient air quality standards (NAAQS) (US EPA, 2017a), water bodies are designated an impaired status because of their polluted status (US EPA, 2017c), and several ecosystems are at risk due to anthropogenic activities (US EPA, 2017b). Environmental pollution not only adversely affects human health but it also threatens peace and security (UN Environment Programme, 2016). In this dissertation, I study the role of information to mitigate environmental pollution and improve public health protection.

The 1960 s and 1970s saw a wave of legislation to protect human health and the environment. Some of the most important pieces of environmental legislation established in this last century are – Clean Air Act (1970), Clean Water Act (1972), Endangered Species Act (1973) and the Toxic Substances Control Act (1976). The United States Environmental Protection Agency (US EPA) is charged with the task of implementing such types of environmental legislation.

Most legislative measures fall under the category of command and control regulation i.e. regulations that permit/forbid certain activities. For example, setting pollutant emission concentration, pollutant ambient levels or technology standards for different industries. While traditional command and control regulations have been very successful at curbing environmental degradation, in recent years they have been supplemented with new approaches to environmental legislation and protection such as information disclosure programs.

2 Environmental Information Programs

Environmental information programs, or information disclosure programs, provide information to the public about pollutants in the environment and in some instances information about the polluters. They usually include information about the health risk associated with the pollutants as well. The EPA developed the slogan “A Right to Know, A Basis to Act” to communicate the purpose of these information programs (US EPA, 2017d). These programs were established to empower the citizenry to learn about pollution sources within their community, take action to better protect their health, to communicate with local lawmakers and facilities producing pollution (Delmas, Shimshack, & Montes-Sancho, 2010). The hope was that negatively publicity generated by the information revealed by these programs would encourage companies to become more environmentally-friendly. Information disclosure programs also increase transparency and trust within the community.

One such program is the Toxic Release Inventory (TRI) Program. After the chemical plant disaster in Bhopal, India, which is considered to be the worst chemical disaster in history, Congress created the TRI program under the Emergency Planning and Community Right-to-Know (EPCRA) Act in 1986 (US EPA, 2017d). All large industrial facilities are required to report their releases of hazardous and toxic chemicals through the TRI program on an annual basis. There are currently 700+ chemicals for which reporting is required. These chemicals have been linked to serious health effects. Besides reporting chemical releases to the environment (land, air, water), companies are also required to report the amount of chemicals are treated or recycled and hence prevented from being released into the environment.

Another type of informational program is the AirNow program. The AirNow program was established by the EPA to report the state of the air to the public and allow them to protect their health against air pollution. According to the 40 CFR (Code of Federal Regulations) 58.50 and 40 CFR Appendix G to Part 58, regulatory air quality (AQ) agencies are required to report the daily AQI for large cities. The AirNow program doesn't identify specific sources of pollution unlike the TRI and primarily enables the public to protect their health against air pollution (AP) by providing information on the levels of pollution and the associated health effects at those levels.

Similar governmental and non-governmental programs exist to increase the public's awareness of pollution. The Beach Report Card, established by Heal the Bay, a nonprofit organization reports weekly real-time water quality for California beaches.

The success of such informational programs hinges on the effective communication of relevant information and the ability to reach a large proportion of the target audience. Based on the gaps identified in relevant literature, my research explores different communication aspects of the AirNow and TRI program in chapters 2 and 3 respectively.

In the second chapter, I discuss how people engage with data from the AirNow program. Despite its importance as a public health issue, we don't have a very good understanding of how people engage with AQ information, which populations use this information regularly and whether it results in behavior changes that reduce health risk against air pollution (Mansfield 2006). While some studies have assessed response to next-day smog alerts published in newspapers (M. Neidell, 2006; M. Neidell, 2004), there have no studies that investigate response to real-time AQ information. My research explores how people engage with real-time AQ information via a mobile application (commonly called app). I investigate which groups respond to this information. I also discuss the use of strategies to increase that engagement based on a survey and field experiment.

In the third chapter, I discuss a new method for analyzing TRI data at the facility level. The complexity of TRI data has prevented it from being used as an assessment tool by facilities and the public more widely. Compared to prior research in this field, my research enhances TRI data analysis on two accounts – simplifying complex data to develop an index using a new method and providing information at the facility level rather than the corporate level. In this chapter, I discuss the development of an environmental performance index from TRI data to identify the "best" and "worst" facilities in the Los Angeles Basin from an environmental and financial perspective. The index was developed using data envelopment analysis (DEA), a novel method that works very well with the unique data irregularities present in TRI data. Even though TRI data is available at the facility level, most prior analyses have been conducted at the corporate level because it is difficult to find financial data at the facility level. My research provides an additional level of granularity since it utilizes financial data at the facility level.

In the fourth and final chapter, I present concluding remarks and future work.

References

Delmas, M. A., Shimshack, J. P., & Montes-Sancho, M. (2010). Information Disclosure Policies : Evidence from the Electricity Industry UC Santa Barbara. *Economic Inquiry*, *48*(2), 483–498. http://doi.org/10.1111/j.1465-7295.2009.00227.x

Mansfield, C., Reed Johnson, F., & Van Houtven, G. (2006). The missing piece: Valuing averting behavior for children's ozone exposures. *Resource and Energy Economics*, *28*(3), 215–228. http://doi.org/10.1016/j.reseneeco.2006.02.002

Neidell, M. (2006). Public Information and Avoidance Behavior: Do People Respond to Smog Alerts? *Center of Statistical and Environmental Science Technical Report*, (24).

Neidell, M. J. (2004). Air pollution, health, and socio-economic status: the effect of outdoor air quality on childhood asthma. *Journal of Health Economics*, *23*(6), 1209–36. http://doi.org/10.1016/j.jhealeco.2004.05.002

UN Environment Programme. (2016). *Annual Report 2016: Engaging People To Protect The Planet*.

US EPA. (2017a). Current Nonattainment Counties for All Criteria Pollutants. Retrieved September 25, 2017, from https://www3.epa.gov/airquality/greenbook/ancl.html

US EPA. (2017b). EnviroAtlas Benefit Category: Biodiversity Conservation. Retrieved September 25, 2017, from https://www.epa.gov/enviroatlas/enviroatlas-benefit-category-biodiversity-conservation

US EPA. (2017c). The Impaired Waters and TMDLs Program in your EPA Region, State or Tribal Land. Retrieved September 25, 2017, from https://www.epa.gov/tmdl/impaired-waters-and-tmdls-program-your-epa-region-state-or-tribal-land

US EPA. (2017d). Toxics Release Inventory (TRI) Program. Retrieved July 9, 2017, from https://www.epa.gov/toxics-release-inventory-tri-program

II. Engagement with Air Pollution Information through an App

1 Introduction

Air Pollution (AP) affects most people; air is vital for survival and yet poor air quality is ubiquitous. The World Health Organization declared that AP is the single largest environmental health risk globally (UN WHO, 2014); particularly in urban areas (Bickerstaff & Walker, 2001). The varied and numerous adverse health effects of AP are well established (Brunekreef and Holgate 2002; Pope and Dockery 2006; Curtis et al. 2006). Besides the human suffering, the associated health care costs run into billions of dollars annually just in the United States (US) (CDC, 2014; Colls, 2002).

Currently, more than half of all Americans – 166 million people – live in areas that don't meet national air quality standards (ALA, 2016). And in the coming years, air quality (AQ) is expected to worsen with climate change (Bergquist et al., 2012; Jacob & Winner, 2008; Mickley, Jacob, Field, & Rind, 2004). Given the health burden of AP, it is not surprising that governments have developed extensive AQ monitoring and reporting programs. The underlying basis of these programs is to increase the public's awareness of the state of the air, especially with regards to health effects so that individuals can adjust their behavior to protect their health (Ruggieri & Plaia, 2012). Individuals can choose to limit time spend outdoors, reschedule outdoor activities, use air conditioning or air filters (US EPA, 2014). As forecasting and modeling technology has progressed, real-time (hourly) AQ updates for most cities are now available. At the same time, AQ information previously published in newspapers and TV, is now rapidly disseminated through websites and mobile applications (commonly called apps) owing to advances in information technology. Apps are third-party software downloaded onto smart phones.

The success of AQ informational programs hinges on their effective communication; yet we have a limited understanding of how people respond to the information in these programs or which groups of people even use this information (Mansfield 2006). Neidell (2004; 2006) found that people protect their health against smog alerts on the day following the alert issued in the newspaper, by reducing outdoor recreational activities but this effect wanes for alerts issues on consecutive days (Zivin and Neidell 2009). Beyond that, there is little information on how people engage with real-time information or their behavior with regards to non-recreational activities.

One may argue that people use their perception or sensory organs to evaluate pollution levels; hence there isn't a strong need for AQ information to protect health. Indeed there is evidence that even if people haven't received information about AP they engage in averting behaviors in response to acute symptoms (Bresnahan, Dickie, & Gerking, 1997). This argument only holds if AP is perceptible and people's perception of it is accurate. On both accounts this argument is not a strong one; AP is often imperceptible (Bickerstaff & Walker, 2001) and when it is, people's perception is often inaccurate and not correlated with actual AP levels (Semenza et al., 2008). AP is an invisible and silent killer. There is a need to encourage people to protect their health based on actual AQ information rather than their perception alone. Thus, also making it important to study the public's response to AQ information and improve it over time to better protect public health.

In addition, AP disproportionately affects a large proportion of the general population; young children, the elderly, pregnant women, asthmatics heart and lung disease patients and those with a compromised immune system (Brook et al., 2004; Pope & Dockery, 2006). AP related health conditions are on the rise (D'amato, Liccardi, D'amato, & Cazzola, 2001) particularly the incidence of asthma in children and exacerbation of asthma symptoms due to traffic-related AP even in developed nations (Gehring et al. 2010; McConnell et al. 2010). Research also indicates that the health effects of AP cannot be reduced by public policy alone, individual action is also required (Laumbach & Kipen, 2012). Those with chronic health conditions such as asthma don't engage in averting behavior based on their assessment of pollution levels (Bresnahan et al., 1997) but no study has assessed their response based on AQ information or for other vulnerable groups.

This research focuses on gaining a better understanding of which groups engage with AQ information and analyze the user pattern to help improve AQ reporting programs. This research study is focused on that. We developed an air quality app, AirForU, with a built-in research study to answer these questions. Through an intake survey within the app, we collected demographic and medical condition information about app users, conducted messaging interventions and tracked how users engaged with AQ information through the app. The app relied on data from the Environmental Protection Agency's (EPA) AirNow program; users could access real-time i.e hourly AQ information and next-day forecasts.

Messaging interventions are widely used in public health and environmental programs. Text message interventions for improving health behaviors (e.g. smoking cessation, diabetes management) have been successful (Fjeldsoe, Marshall, & Miller, 2009). Both voice and text message reminders and educational messages have improved health outcomes for thousands of patients with different health conditions (Krishna, Austin Boren, & Balas, 2009). The effectiveness of internet-based health interventions has been demonstrated through numerous studies (Bennett & Glasgow, 2009). The use of smartphones and smartphones apps in healthcare has burgeoned because of their potential in improving access to healthcare information (Ozdalga, Ozdalga, & Ahuja, 2012; Terry, 2010). Apps can encourage participatory healthcare; allowing many patients to actively participate in their own healthcare and access healthcare information (Kailas, Chong, & Watanabe, 2010; Terry, 2010). Participatory healthcare is especially important for chronic conditions (Boulos, Wheeler, Tavares, & Jones, 2011) such as asthma.

After having built a significant app user base (~3000 users), we then tested the effect of 12 AP messages on engagement with AP information within the app. We wanted to better understand how different groups respond to different health messages because AP information needs are personal and varied depending on an individual's health status and location (Bush, Moffatt, & Dunn, 2001). According to a survey conducted in the UK (which has a similar history of AP programs to the US), the public indicates a greater desire for personalized information relating to their health and their family's health (Bickerstaff & Walker, 2001). The AP messages developed in this study were first tested through an online survey (via Amazon's Mechanical Turk or MTurk) and then delivered via email to the app users. In the survey, hypothetical engagement with AQ information was tested while in the app actual engagement was measured. Engagement for app users was measured before and after the messages were sent using Google Mobile

Analytics. Hypothetical engagement measured in the MTurk survey differed from actual engagement measured through the app. Understanding which groups respond (or don't respond) to AQ information via an app, how they respond and how their engagement can be influenced will contribute towards improving AQ reporting programs.

2 Literature Review and Scope of Study

Two facets of the AP messages were tested – the type of health impact and the framing i.e. wording of the message. We referenced the AP health literature in developing the content of these messages. The wording was based on the message framing literature commonly used in public health campaigns and behavioral interventions. Because the AP health literature is extremely vast, only part of the literature pertaining to specific health issues was referenced in depth.

2.1 Air Pollution Health Literature

2.1.1 Outdoor Exercise

Outdoor exercises such was walking, biking and running are the most common and accessible forms of exercise and thus allowed us to target a wide range of the population. While exercise has many benefits, exercising outdoors during high AP can have a detrimental impact (Giles & Koehle, 2014).

2.1.2 Child Asthma

Asthma is the leading chronic condition affecting children (Neidell 2004) and is the main reason for school absenteeism and hospital admissions among children (US Dept. of Health and Human Services, 2012). Child asthma is also on the rise (Centers For Disease Control and Prevention, 2015). We developed a message geared towards asthmatics and/or caretakers for asthmatic children.

2.1.3 Child Cognition

High levels of AP have been linked with cognitive decline in children (Calderón-Garcidueñas et al., 2008; Freire et al., 2010; Suglia, Gryparis, Wright, Schwartz, & Wright, 2008). We developed two messages focused on children because they tend to illicit a stronger response (Davis, 1995) and the public is highly concerned about the health impact of AP on their family (Bickerstaff & Walker, 2001).

2.1.4 Alzheimer's

More recently, AP has been linked to brain damage and neurodegenerative disorders like Alzheimer's disease (M. L. Block & Calderón-Garcidueñas, 2009; Calderón-Garcidueñas, 2002; Kampa & Castanas, 2008; Levesque, Surace, McDonald, & Block, 2011; Moulton & Yang, 2012; Weuve et al., 2012). We developed a message geared at the elderly (~ 60 years) who are at higher risk (Brookmeyer, Gray, & Kawas, 1998). Early AP health studies were focused on respiratory ailments but in the recent decades AP has linked to end-points where the connection to AP is more surprising.

2.1.5 Perception of AQ

The last category is a general message not geared at a specific subgroup. This message aims to make the invisible AP visible by highlighting its health effects. This is also linked to the idea of encouraging people to check AQ information rather than relying on their perception of AP which tends to be inaccurate (Semenza et al., 2008).

2.2 Message Framing literature

There is an extensive and conflicting literature investigating the effect of judgments and decision-making of different but equivalent descriptions of the same statement or commonly known as “framing effects”. Framing effects are commonly used in the fields of public and social health to influence behavioral change interventions. Much of the literature derives from Kahneman and Tversky’s (1979) prospect theory, which suggests that people prefer taking risk over certainty when considering losses and people prefer certainty over risk when considering gains. Alternatively, loss (vs. gain) framing increases motivation for risk-seeking behavior while gain (vs. loss) framing increases motivation for risk-averse behavior. They also theorized that positively or negatively framed information affects the perception of risk (Tversky & Kahneman, 1985).

Adding to the framing effects introduced by the prospect theory, there is a vast literature on negativity bias and fear appeal, while sometimes inconsistent, the literature indicates that threatening messages with high-efficacy messages induce behavior change while threatening messages with low-efficacy messages induce a defensive reaction (Witte & Allen, 2000).

Negatively framed messages are more effective than positively framed messages for those who have a high degree of involvement with that issue (Greenwald and Leavitt 1984; Chaiken 1980).

In the public health domain, studies have assessed the impact of gain/loss or positive/negative language for a number of behaviors. Positive framing was found to be more effective at encouraging sunscreen use (Detweiler, Bedell, Salovey, Pronin, & Rothman, 1999) and the purchase of lean meat (Levin, 1987). Negative framing was more effective at promoting breast-self-examinations (Meyerowitz & Chaiken, 1987), mammography examinations (Banks et al., 1995) and recycling (Davis, 1995). Another factor further complicates the results – issue involvement. Negative messages were effective when detailed processing of the message was required e.g. heart disease (Maheswaran & Meyers-levy, 1990) and skin cancer/sexually transmitted diseases (L. G. Block & Keller, 1995) and neither framing was important for messages with low processing.

In yet another category of studies that assessed the effect of combined positive and negative framing compared to positive or negative framing alone, combined messages were effective at increasing the use of car seats for infants (Treiber, 1986) and in smoking cessation (Wilson, Wallston, & King, 1990).

A number of factor seem to be at play – efficacy of behavior, issue involvement, processing depth. No studies have assessed the effect of positively/negatively framed AP information and given the conflicting results obtained in other studies it is unclear whether positively or negatively framed information will be more effective. It is also possible that different groups will respond differently. This research study will be first one to assess the effect of valence framing for AP-based messages.

2.3 Hypotheses

Based on the message framing literature, which often shows mixed results, a reasonable guess can be made about the type of message framing that is likely to be more effective. Positive framing for low-risk behaviors (avoiding exposure to outdoor air pollution is a relatively low-risk behavior) tends to be more effective. On the other hand, negative framing is more effective among those with a high degree of issue involvement (app users can be considered to have a high involvement since they voluntary downloaded the app and are thus interested in AP). Since no study has considered AP messaging thus far, we develop the following hypothesis:

A. Positively framed messages are more effective at engaging people with AQ information

Through the last hypothesis, we test whether targeted messages are more effective. Personalized messaging health-based interventions tend to be even more effective than general ones (Lustria et al., 2013; Noar, Benac, & Harris, 2007). Similarly, tailored internet-delivered pro-environmental interventions have also been more successful (Abrahamse, Steg, Vlek, & Rothengatter, 2007; Asensio & Delmas, 2015). For this study, this includes the effect of children-based messages among those with young children living in their household, the effect of the Alzheimer's message among the elderly, the effect of messages among asthmatics and the effect of the exercise message among those that exercise more frequently.

B. Messages aimed at specific groups are more effective. Children-based messages are more effective among parents/guardians, the Alzheimer's message is more effective among the elderly and the exercise message is more effective among those who exercise regularly.

3 Methods

To test our hypotheses, we conducted two different types of field experiments. First, we tested the AP messages through an MTurk survey and measured hypothetical engagement with AQ information. Second, we tested the same messages among the app users and measured actual engagement within the app. We had the unique advantage of comparing survey data to a field experiment designed to be as similar as possible to the survey.

We choose to use and develop an app to conduct our research study for a number of reasons. New technologies such as smartphones and smartphone apps offer an innovative platform to conduct research due to their functionality and ubiquity. They offer unprecedented opportunities to engage a large number of people and collect extensive data; 77% of adults in the US use smartphones (Pew Research Center, 2017). We partnered with UCLA Health to recruit participants because we wanted to focus on sensitive populations that are disproportionately affected by AP.

The AP messages were finalized after conducting surveys and focus groups. Initially, we developed about 10-15 messages that focused on the health content, the first facet of the messages. The messages targeted different groups of the population – parents of children, pregnant women, the elderly and those who exercise outdoors frequently. We tested these messages through surveys and focus groups multiple times among groups of 15-20 university students. Some of the messages were based on common health conditions of AP such as asthma while others were based on more threatening conditions such as cancer, Alzheimer's and cognitive development. Not all conditions, for example Alzheimer's, were relevant to university students and thus we asked them to also consider these messages in relation to their family members; in surveys on air quality information people indicated an interest in health messages that pertained not only to themselves but also their family's health (Bickerstaff & Walker, 2001). Other were factual messages based on the end result of AP; these included numbers of school absences due to AP, shortened of life span due to AP and deaths linked to AP in the US and worldwide. Some of the messages were general (e.g. children are more vulnerable) and some were more specific (e.g. cancer). Based on the feedback and results we received from the initial surveys, we selected the following five categories of health effects – outdoor exercise, child asthma, child cognition, Alzheimer's, invisibility of AP – and a baseline message for comparison. Valence framing, the second facet of the message, was then added.

3.1 MTurk survey

3.1.1 MTurk Survey Details and Administration

The goal of the MTurk survey was to identify AP statements that influence (hypothetical) engagement with an air quality app. Engagement in this context refers to checking AQ levels on the app before engaging in outdoor activities. Several statements about AP and health impacts were presented and respondents were asked rate each statement for its comprehensibility, realism, relevance and whether they would check AQ on an app before engaging in outdoor activities. Each statement was rated on a seven-point scale ranging from strongly disagree to strongly agree on these four aspects. In addition, demographic information and respondents' knowledge of common AQ terms was also part of the survey (Refer to

Appendix 3 – Complete MTurk Survey).

Two facets of the AP messages were tested – the type of health impact and the framing of the message using positive/negative wording. Five categories of health impact – outdoor exercise, child asthma, child cognition, Alzheimer's, invisibility of AP - and for each category the wording was modified without altering the content as much as possible to have a positively and negatively framed version of the message. Each respondent received only one question from each category randomly assigned; either positively or negatively framed. One of the categories, outdoor exercise, had a combined positive and negative message in addition to the individually framed messages. Besides these 5 categories there was a control statement without any specific framing that was presented to all respondents. All messages are listed in Appendix 1.

Below, we provide an example of positively and negatively framed air pollution statements from the child asthma category respectively.

1. Do you know that high air pollution can cause or worsen childhood asthma? **Avoiding air pollution can reduce this risk.** Check your local air quality on AirForU today before engaging in outdoor activities!
2. Do you know that high air pollution can cause or worsen childhood asthma? **Exposure to air pollution can increase this risk**. Check your local air quality on AirForU today before engaging in outdoor activities!

The survey was conducted via MTurk, an online survey service frequently used by researchers (Buhrmester, Kwang, & Gosling, 2011). A total of 835 responses were collected. The survey was conducted in 2 parts. The AP questions for both parts remained the same but for the second part additional demographic information - frequency of outdoor exercise, education level, annual household income and race/ethnicity - was collected. Respondents received $1 for their participation. The response rate was 100%; the high response rate can be owed to the fact that the respondents only received payment after completion of the survey.

3.1.2 MTurk Survey Data Analysis

Linear OLS regressions were used to analyze the MTurk data. There were 4 dependent variables (y) of interest – comprehensibility of the statement, realism of the statement, relevance of the statement and whether the statement would lead to checking the AQ on the app before engaging in outdoor activities heretofore referred as comprehension, realism, relevance and check AQ. Of these, the two most important ones are relevance and check AQ. Relevance indicates which messages are more relevant for which groups and thus helps develop targeted messages. Checking AQ, perhaps the most important dependent variable as it indicates a behavior change: checking AQ on the app before going outdoors. This is one of the effective steps that people can take to protect their health against AP.

The treatment variables (x) were the question category (e.g. exercise, baseline) and the question framing (e.g. positive, negative). The controls were: age, gender, asthma, children living in household, children living in household with asthma, income, household income, frequency of outdoor exercise and education. Another variable of interest was the knowledge of AQI and $PM_{2.5}$. The last four controls were used only for 430 responses since these data weren't collected from all respondents.

$$y_i = \beta_{0i} + \beta_{treatment,i}x_{treatment,i} + \beta_{control1,i}x_{control1,i} \qquad (1)$$

y_i = comprehensibility, realism, relevance and check AQ

$x_{treatment}$ = question category and question type

$x_{control,n}$

= age, gender, asthma, children in HH, child asthma, income, exercise freq., education, AQ knowledge

The following interactions were added to the regressions – question category X children, question category X asthma, question category X age and question category X frequency of outdoor exercise.

3.2 AirForU App

3.2.1 Air Quality Data, App Development and Recruitment

The fundamental feature of an air quality (AQ) app is that it provides information about the state of the air i.e. how clean or polluted the air is and, in most cases, associated health risk information. Global Positioning System (GPS) capability in smartphones allows users to access AQ based on their current location with little effort. While many other AQ apps exist on the market, AirForU is specifically and uniquely designed as a research tool to characterize engagement with AQ information. AirForU also contains supplementary AQ-related features not found other AQ apps.

We developed two versions of the AirForU app – one for iPhones and one for Android devices;

both versions are exactly the same besides cosmetic differences due to the native development platforms. Development for the AirForU app began towards the end of 2014. Testing began a few months later and the final version was launched in October 2015 under the UCLA Health brand in the Google Play (for Android devices) and App Store (for iPhones) (together these heavily dominate the market (Statista, 2017)). The app is available for free to the general public. The study was approved by the IRB (Protocol ID #15-000215).

While ownership varies by age, income and education level, adoption rates are still high across the spectrum. Adoption rate among Whites, Blacks and Hispanics are also fairly equal; 77%, 72% and 75% respectively (Pew Research Center, 2017). Being able to reach minorities and those with a lower education is important because while those with higher education take more steps to protect their health (Bresnahan et al., 1997), less educated people (which tend to belong to lower socio-economic status communities) are the ones that are usually disproportionately impacted by AP (Krewski et al., 2000, 2004; Laumbach & Kipen, 2012) particularly African Americans (Mannino et al., 2002; Sly, 1988; Weiss & Wagener, 1990) and Hispanics (AAFA & NFC, 2005).

AQ and health data is obtained from EPA's AirNow program, which publishes real-time hourly updates of AQ as well as next-day forecasts on their website. The AQ is reported to the public based on EPA's guidelines in the form of an Air Quality Index (AQI), which accounts for the ambient concentrations of several pollutants. The AQI communicates how clean or polluted the outdoor air is along with the associated health effects that may be of concern at those levels (US EPA, 2006).

To understand how app users interact with AQ information, two primary methods are used collect data – surveys and Google Mobile Analytics. Users also have the option to allow notifications from the app; if they sign up for notifications they receive weekly alerts encouraging them to check the AQ. When the app is first downloaded onto the phone, users need to complete an intake survey before they can begin accessing the app's features. In the intake survey demographic, medical history and other related information is collected. Wherever the order of answers is not important, the answer options are randomized. Users had the choice to not respond to the questions but they prompted to do so. Through Google Mobile analytics each user's movements with the app are tracked.

A number of avenues (social media, newsletters, websites, flyers) were used to market the app. The UCLA Health Media and Marketing team provided invaluable support in marketing the app and recruiting users through their extensive health network. Partnering with UCLA Health facilitated contact with a larger number of sensitive groups. They marketed the app through their website, social media and health newsletters which have over 650,000 subscribers consisting of healthcare professionals.

3.2.2 Air Pollution Messages Email Experiment

The same statements tested in the survey were tested with the app users. Users were randomly assigned 1 of 13 groups (~200-250 users in each group), 1 control and 12 treatment groups each receiving one statement delivered via email for 5 consecutive weeks. Emails were collected in the app's intake survey of the app. The emails were sent via MailChimp, an email service.

3.2.3 Measuring Engagement for App Users

One challenge is to develop a relevant metric for user engagement for AirForU. Engagement can be generally defined as a user's level of involvement with a product; for technological items it usually refers to behavioral proxies such as the frequency, intensity, or depth of interaction over some time period (Rodden, Hutchinson, & Fu, 2010). Mobile/app behavior is a relatively new phenomenon and, while different from online behavior, knowledge of knowledge of web analytics can be extended to mobile analytics. Engagement with technology is multi-faceted and highly dependent on the technology (Attfield, Kazai, & Lalmas, 2011; Lehmann, Lalmas, Yom-Tov, & Dupret, 2012), hence it is important to define engagement based on application's objectives (Fagan, 2014; Lalmas, O'Brien, & Yom-Tov, 2014). For AirForU, the only "critical" objective is to check AQ (either current or forecast) and hence engagement is defined as simply opening the app. The first screen the user is led to it the current AQ for that purpose. We didn't use the duration of the app visit since a visit may last only a few seconds yet the user might have accessed "critical" content and be "satisfied" with the information.

A pooled ordinary least squares (OLS) with basic difference-in-differences model was selected for measuring engagement before and after the email experiment and testing the following hypothesis. Differences-in-differences is a common method in social sciences research for measuring the effects of an experiment or quasi-experiment. Equation (2) below describes our model:

$$y = \beta_0 + \beta_1 dB + \delta_0 d2 + \delta_1 d2 \cdot dB + \beta_{control1,i} x_{control1,i} + \varepsilon \qquad (2)$$

d2 is a dummy variable for the second-time period and captures any changes that would have occurred over time even in the absence of a treatment, dB captures differences between the control and treatment groups prior to the treatment and the coefficient of interest δ_1 is a measure of the change in y (dependent variable or outcome of interest) due to the treatment. The control variables included week dummies to control for seasonality or other time-related factors potentially affecting engagement, demographic controls from the intake survey and also a control to account for the user's activity with the app. The activity control accounted for whether a person had been active (or inactive) with the app for a number of weeks (5 weeks, 10 weeks, 15 weeks and 20 weeks) since they downloaded it. New users were more engaged in general and this factor accounted for that as well.

The unit of observation was person-week. For each person, the total number of observations equaled the number of weeks since they downloaded the app until the end of the data collection for this experiment i.e. since the app has been available for about 88 weeks, the maximum number of observations for one person is 88. This was selected the daily numbers were infrequent while the monthly data would not have adequately captured the changes from the email experiment.

4 Results and Discussion

4.1 MTurk Survey

4.1.1 Summary Statistics

The purpose of the MTurk survey was to identify the most effective messages (i.e. messages that encourage people to check air quality through an AQ app, in this case AirForU app specifically). One of the first concerns is the applicability of the data to the general population. MTurk is a frequently used online survey service; data collected through MTurk is at least as reliable as data collected through traditional methods (Buhrmester et al., 2011). Based on the demographics, the survey was overrepresented by men (Table 1). Mean household income for the US population ($79,263) was higher compared to the survey respondents ($54,563) (US Census Bureau, 2015). This was not surprising considering that respondents were paid for their participation. Ethnicity/Race proportions are close to the general population. The complete results can be found in the Appendix (Table 19 and Table 20).

Asthma was higher among respondents and their children compared to CDC data. One reason for this may be that asthma reported by respondents could be self-diagnosed rather than diagnosed by a medical professional, which is the basis for CDC numbers. The same trend was also observed among the app users (Table 11). Since sensitive groups are more susceptible to the health effects of AP, understanding their response can be more advantageous at improving AQ communication systems.

Most AP messages received a high comprehensibility score implying that most people understood the messages clearly (Table 20). Realism of statements also received a high score for the most part; the child cognition and Alzheimer's messages received the lowest scores (Table 20). This is possibly because these effects are as common or as intuitively linked to AP asthma or exercising outdoors. Another reason could be because Alzheimer's and cognition are relatively serious conditions and messages that induce a strong fear appeal and are low-efficacy tend to produce defensive actions (Witte & Allen, 2000). Messages geared at children (child asthma, child cognition) and the elderly (Alzheimer's) received the lowest score on the relevance scale. This is expected because only about 30% of the respondents have children and only about 7% of the respondents are 55 years or older. The exercise and invisibility questions received the highest score in encouraging people to check AQ through an app. These results are further analyzed through regression analysis.

Table 1: MTurk Survey Summary Statistics for Demographics (N=835 or N=430)

	Survey Respondents					Population	
	N	M	SD	Min	Max	M	SD/MoE
Female	835	0.430	0.495	0	1	0.508	± 0.1
Age (years)	835	35.3	11.1	18-24	≥ 65	37.8	± 0.1
Income[a]	430	54563[b]	37518[b]	≤ 24999	≥150000	79263	403
Frequency of outdoor exercise[a]	430	3.68[c]	1.46[c]	1	6		
Education[a]	430	4.05[d]	1.31[d]	1	6		
Race (White or Caucasian)	430	0.726	-	-	-	0.769	
Have Asthma	835	0.115	0.319	0	1	0.074	
Children (<18 yrs.) living in HH	835	0.295	0.456	0	1	0.250	
Children (<18 yrs.) with asthma	835	0.236	0.425	0	1	0.086	
Have Knowledge of AQI	835	0.181	0.385	0	1		
Have Knowledge of $PM_{2.5}$	835	0.403	0.491	0	1		

[a] N for these demographics is 430 because these questions were only asked in the second set of survey responses
[b] Responses ranged less than $24,999 to $150,000 or more coded as values 1 to 6
[c] Responses ranged from once a year or less to 5 or more times a week coded as values 1 to 6
[d] Responses ranged from less than high school to graduate degree coded as values 1 to 6

The mean and standard deviation for age, income, education and frequency of outdoor exercise (Table 1) were calculated by using the midpoint of each category and thus they are estimates.

4.1.2 Regressions

The purpose of the regressions was to identify the most effective AP message categories and the most effective message framing within those categories. The regressions provide evidence for which groups respond most to which categories suggesting the important of targeted AP messages. People desire personalized AP info (Beaumont, Hamilton, Machin, Perks, & Williams, 1999; Bickerstaff & Walker, 1999).

In the first set of regressions, the four dependent variables considered were comprehensibility, realism, relevance and checking AQ with the question category as the treatment variable (Table 3). All coefficients are interpreted compared to the baseline message which was used as the control. The total number of observations was 5010 i.e. 835 respondents each of whom received 6 AP messages. Standard errors were clustered and heteroscedasticity and robustness checks were added to all regressions.

Correlation coefficients for the four dependent variables indicate that they are all moderately-strongly correlated (Table 2). For a statement to influence one to check AQ, it must be relevant to the individual and it should also be realistic and comprehensible (lowest correlation). This pattern is reflected in the regression results as well. Messages with a high (or low) coefficient for realism and relevance tend to follow the same trend for checking AQ. A complete coefficient matrix can be found in the appendix (Table 21).

Table 2: MTurk Survey: Correlation Coefficients for the four dependent variables (N=5010)

Variables	Check AQ	Realism	Relevance	Comprehend
Check AQ	1			
Realism	0.466	1		
Relevance	0.696	0.462	1	
Comprehend	0.243	0.532	0.259	1

The exercise and AP invisibility category were the most effective categories (Column 4 Table 3) relative to the baseline message. Older people (> 55 years) are more likely to find the messages relevant and are more likely to check AQ compared to younger people (< 55 years). Females are also likely to find the messages relevant and also more likely to check AQ compared to males. Asthmatics are also more likely to check AQ as well as people with children. Interestingly, those with children with asthma are not more likely to check AQ compared to people with children without asthma. Those who are more knowledgeable about AQ are more likely to check AQ. This is in line with previous research about education (Bresnahan et al., 1997; Krewski et al., 2004).

As far as the relevance of the messages are concerned (Column 4 Table 3), the childhood asthma and cognition messages and the Alzheimer's messages were not relevant to all users. This makes sense because these messages are targeted towards specific subgroups and hence they were not among the most effective messages.

The Alzheimer's message received the lowest score for comprehensibility (Table 20). (Column 1 Table 3). Compared to the baseline message it was not as easily comprehensible (Column 1 Table 3). One reason for this could be that the connection between AP and Alzheimer's is not obvious or common knowledge. It is obvious that respiratory conditions such as asthma are linked to AP while the link to Alzheimer's disease is obvious. This message was not as realistic relative to the baseline (Column 2 Table 3) and it follows from its low comprehensibility score.

While the R values were small (0.14 and 0.05) for relevance and check AQ, the independent variables explain a non-trivial component of the dependent variables based on the F-values (P

<0.000).

Table 3: Regressions for message categories relative to the baseline category for MTurk survey respondents (N=835)

Study Characteristic	**(1) Comprehend**	**(2) Realism**	**(3) Relevance**	**(4) Check AQ**
Category Treatment				
Exercise	0.04	0.36***	0.11**	0.38***
	(0.03)	(0.05)	(0.05)	(0.05)
Child asthma	0.12***	0.55***	-1.12***	-0.03
	(0.03)	(0.04)	(0.08)	(0.06)
Child Cognition	-0.04	-0.05	-1.07***	0.01
	(0.03)	(0.05)	(0.07)	(0.06)
Alzheimer’s	-0.11***	-0.53***	-0.66***	0.00
	(0.03)	(0.06)	(0.06)	(0.06)
AP Invisibility	0.03	0.38***	0.18***	0.42***
	(0.03)	(0.04)	(0.04)	(0.05)
Controls				
Age > 55 years	0.08	0.22**	0.38**	0.28
	(0.10)	(0.11)	(0.17)	(0.19)
Female	0.18***	0.25***	0.19**	0.25***
	(0.05)	(0.06)	(0.08)	(0.09)
Asthma	0.08	0.24***	0.53***	0.66***
	(0.07)	(0.09)	(0.10)	(0.11)
Children	0.14**	0.18**	0.60***	0.34***
	(0.07)	(0.07)	(0.09)	(0.11)
Children with asthma	-0.14	-0.15	0.08	0.07
	(0.13)	(0.15)	(0.16)	(0.18)
Knowledge of AQ	0.18***	0.22***	0.33***	0.39***
	(0.07)	(0.08)	(0.09)	(0.11)
Constant	6.06***	5.29***	4.94***	4.41***
	(0.05)	(0.06)	(0.07)	(0.08)
Observations	5,010	5,010	5,010	5,010
Adjusted R-squared	0.02	0.10	0.14	0.05
F	9.080	51.82	52.40	22.44

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The second set of regressions considered relevance and checking AQ as dependent variables with the message framing as the treatment variable (Table *4*) for all respondents.

Besides the exercise category, the type of framing did not have a significant impact on effectiveness. For the exercise category, the negative framing and combined positive and negative framing were statistically significant and of the two, the combined exercise message was more effective (Column 2, Table 4). For the invisibility category, there was no statistically significant difference between the two types of framing. These two categories – AP invisibility and exercise – may be more effective compared to the other categories is because they are less threatening i.e lower fear appeal compared to Alzheimer's, child brain cognition and child asthma. They are also more relevant (Column 1 Table 4) since most people go outdoors (whether they exercise or not) and everyone breathes in air pollution that is invisible. Thus, hypothesis A did not hold for the AP messages in the MTurk Survey.

Table 4: Regressions for message framing relative to the baseline question for MTurk survey respondents (N=835)

Study Characteristic	**(1)** **Relevance**	**(2)** **Check AQ**
Question Treatment		
Exercise positive	-0.04	0.32***
	(0.09)	(0.09)
Exercise negative	0.16*	0.14
	(0.09)	(0.09)
Exercise positive and negative	0.23***	0.66***
	(0.08)	(0.08)
Child asthma positive	-1.05***	-0.05
	(0.10)	(0.08)
Child asthma negative	-1.19***	0.00
	(0.10)	(0.09)
Child cognition positive	-1.00***	0.02
	(0.10)	(0.09)
Child cognition negative	-1.13***	0.00
	(0.10)	(0.09)
Alzheimer's positive	-0.77***	-0.06
	(0.09)	(0.09)
Alzheimer's negative	-0.55***	0.06
	(0.09)	(0.08)
AP Invisibility positive	0.20***	0.44***
	(0.06)	(0.07)
AP Invisibility negative	0.17***	0.39***
	(0.06)	(0.07)
Controls		

Study Characteristic	(1) Relevance	(2) Check AQ
Age > 55 years	0.38**	0.29
	(0.17)	(0.19)
Female	0.19**	0.25***
	(0.08)	(0.09)
Asthma	0.53***	0.65***
	(0.10)	(0.11)
Children	0.59***	0.34***
	(0.09)	(0.11)
Children with asthma	0.08	0.06
	(0.16)	(0.18)
Knowledge of AQ	0.33***	0.38***
	(0.09)	(0.11)
Constant	4.94***	4.41***
	(0.07)	(0.08)
Observations	5,010	5,010
Adjusted R-squared	0.14	0.06
F	34.89	16.26

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The previous regressions (Table 3 and Table 4) did not have added controls for frequent outdoor exercise, college education, income and race but the results with the added controls are in Table 5. Since this information was not collected for some of the survey respondents the number of observations if lower (N=430). There is no difference in the categories as expected but non-white races/ethnicities are more likely to check AQ compared to whites. Minorities are often exposed to higher levels of pollution so this information may suggest that they are aware of the discrepancy and take measures to protect against it. Those who exercise outdoors are more likely to check AQ compared to those who don't but this result is significant only at the 10 % level. The trends in other controls remained the same, while age and knowledge of AQI lost significance. Regressions for the message framing with the new set of controls are in the appendix (Table 23). Not much new information is gleaned from those regressions.

Table 5: Regression results for message categories relative to the baseline for MTurk survey respondents (N=430)

Study Characteristic	**(1) Relevance**	**(2) Check AQ**
Category Treatment		
Exercise	0.08	0.40***
	(0.06)	(0.06)
Child asthma	-1.14***	-0.03
	(0.10)	(0.08)
Child Cognition	-1.08***	-0.04
	(0.10)	(0.09)
Alzheimer’s	-0.70***	0.03
	(0.09)	(0.09)
AP Invisibility	0.13**	0.34***
	(0.06)	(0.06)
Controls		
Age > 55 years	0.42*	0.41
	(0.22)	(0.26)
Frequent Outdoor Exercise	0.27**	0.13
	(0.11)	(0.13)
College education	-0.00	0.05
	(0.11)	(0.13)
Above median income	0.16	0.03
	(0.11)	(0.13)
Non-white	0.27**	0.41***
	(0.12)	(0.14)
Female	0.23**	0.30**
	(0.11)	(0.13)
Asthma	0.55***	0.77***
	(0.15)	(0.16)
Children	0.55***	0.28*
	(0.13)	(0.15)
Children with asthma	-0.01	-0.04
	(0.22)	(0.28)
Knowledge of AQ	0.12	0.20
	(0.15)	(0.17)
Constant	4.77***	4.23***
	(0.13)	(0.15)
Observations	2,580	2,580
Adjusted R-squared	0.15	0.06
F	20.97	9.430

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The next set of regressions included interactions to expand the understanding of which messages were more effective among which groups. The following interaction terms were included – message category with children (Table 6), message category with asthma (Table 7), message category with age (Table 8) and message category with exercise frequency. Including an interaction term for children and treatment category (Table 6), significantly changes the results. People with children are significantly more likely to respond to messages geared at children irrespective of whether their children have asthma. There is also no significant difference between the asthma and the cognition messages so they are likely to respond to each equally. Similarly, those with asthma are more likely to check AQ in response to the message about child asthma even if they don't have children (Table 7).

In contrast, people above 55 years don't show a significant response to the Alzheimer's message (Table 8). They are also unlikely to respond to messages based on child conditions which is not surprising. They don't find the Alzheimer's relevant and they might not believe that this message is true (based on negative coefficients for realism). There is evidence indicating that people know about the general health impacts of AP but don't know as much as the specific impacts (Bickerstaff & Walker, 1999). The Alzheimer's message is not appealing to its target audience. Similarly, including a term for exercise frequency interacted with message category does not indicate a strengthened response for the exercise message among those who exercise more frequently. Additional regressions included in the appendix (Table 22 and Table 23). Thus, hypothesis B, was true for most of the targeted groups except for the elderly group with regards to the Alzheimer's message.

Table 6: Regression with interactions for children and message category for MTurk survey respondents (N=835)

Study Characteristic	(1) Relevance	(2) Check AQ
Category Treatment		
Exercise	0.11**	0.42***
	(0.05)	(0.06)
Child asthma	-1.55***	-0.31***
	(0.09)	(0.07)
Child Cognition	-1.63***	-0.28***
	(0.08)	(0.08)
Alzheimer's	-0.65***	0.01
	(0.07)	(0.07)
AP Invisibility	0.17***	0.41***
	(0.05)	(0.06)
Interactions with Message Category		
Exercise*Have children	0.02	-0.14
	(0.11)	(0.10)
Child asthma*Have children	1.48***	0.98***
	(0.15)	(0.13)
Child Cognition*Have children	1.91***	1.00***
	(0.14)	(0.13)
Alzheimer's*Have children	-0.04	-0.04
	(0.14)	(0.13)
AP Invisibility*Have children	0.05	0.02
	(0.10)	(0.10)
Controls		
Age > 55 years	0.38**	0.28
	(0.17)	(0.19)
Gender	0.19**	0.25***
	(0.08)	(0.09)
Asthma	0.53***	0.66***
	(0.10)	(0.11)
Children	0.03	0.03
	(0.12)	(0.13)
Child Asthma	0.08	0.07
	(0.16)	(0.18)
Knowledge of AQ	0.33***	0.38***
	(0.06)	(0.06)
Constant	5.11***	4.50***
	(0.07)	(0.08)
Observations	5,010	5,010
Adjusted R-squared	0.19	0.07
F	46.76	22.67

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Regressions with interactions for asthma for MTurk survey respondents (N=835)

Study Characteristic	**(1) Relevance**	**(2) Check AQ**
Category Treatment		
Exercise	0.11**	0.36***
	(0.05)	(0.05)
Child asthma	-1.29***	-0.13**
	(0.08)	(0.07)
Child Cognition	-1.11***	-0.03
	(0.08)	(0.07)
Alzheimer's	-0.70***	-0.05
	(0.07)	(0.07)
AP Invisibility	0.18***	0.39***
	(0.04)	(0.05)
Interactions with Message Category		
Exercise*Have asthma	0.05	0.16
	(0.15)	(0.17)
Child Asthma*Have asthma	1.47***	0.93***
	(0.22)	(0.19)
Child Cognition*Have asthma	0.36	0.33
	(0.24)	(0.20)
Alzheimer's*Have asthma	0.34*	0.41**
	(0.19)	(0.17)
AP Invisibility*Have asthma	0.02	0.19
	(0.16)	(0.16)
Controls		
Age > 55 years	0.38**	0.28
	(0.17)	(0.19)
Female	0.19**	0.25***
	(0.08)	(0.09)
Asthma	0.16	0.32*
	(0.14)	(0.16)
Children	0.60***	0.34***
	(0.09)	(0.11)
Child Asthma	0.08	0.07
	(0.16)	(0.18)
Knowledge of AQ	0.33***	0.39***
	(0.09)	(0.11)
Constant	4.99***	4.45***
	(0.07)	(0.08)
Observations	5,010	5,010
Adjusted R-squared	0.15	0.06
F	40.07	19.47

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 8: Regressions with interactions for age > 55 years for MTurk survey respondents (N=835)

Study Characteristic	**(1) Relevance**	**(2) Check AQ**
Category Treatment		
Exercise	0.13**	0.38***
	(0.05)	(0.05)
Child Asthma	-1.10***	0.01
	(0.08)	(0.06)
Child Cognition	-1.05***	0.05
	(0.08)	(0.07)
Alzheimer's	-0.66***	0.00
	(0.07)	(0.06)
AP Invisibility	0.19***	0.41***
	(0.04)	(0.05)
Interactions with Message Category		
Exercise* Age > 55 years	-0.16	-0.13
	(0.15)	(0.16)
Child Asthma* Age > 55 years	-0.21	-0.46*
	(0.25)	(0.24)
Child Cognition* Age > 55 years	-0.28	-0.64***
	(0.22)	(0.23)
Alzheimer's* Age > 55 years	0.03	-0.04
	(0.23)	(0.23)
AP Invisibility* Age > 55 years	-0.06	0.10
	(0.14)	(0.18)
Controls		
Age > 55 years	0.49***	0.47**
	(0.19)	(0.21)
Female	0.19**	0.25***
	(0.08)	(0.09)
Asthma	0.53***	0.66***
	(0.10)	(0.11)
Children	0.60***	0.34***
	(0.09)	(0.11)
Child Asthma	0.08	0.07
	(0.16)	(0.18)
Knowledge of AQ	0.33***	0.39***
	(0.09)	(0.11)
Constant	4.94***	4.40***
	(0.07)	(0.08)
Observations	5,010	5,010
Adjusted R-squared	0.14	0.06
F	39.48	17.09

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In summary, we found that the exercise and AP invisibility message categories were the most effective. The framing had little effect within most categories except for the exercise category where the combined positive and negative framing was most effective. Children-based messages were more effective among parents/guardians even if the children did not have asthma and the child-asthma based message was effective among adults with asthma irrespective of whether they had children or not. In contrast, the Alzheimer's message was not effective among the elderly and the exercise message was not more effective among those who exercise outdoors frequently. The correlation matrix (Table 21 in the appendix) lends evidence for the relationship between the variables included in the regression. Demographics (age, gender) are correlated with the four dependent variables as are exercise frequency, asthma and income. Race and education are not strongly correlated though. Knowing the AQI and PM2.5 are moderately correlated (0.327) and thus only one of them is retained in the regression models to avoid collinearity.

4.2 App Data and Email experiment

4.2.1 Summary Statistics for App Intake Survey

App users encountered an intake survey when they first downloaded the app before they were able to access the AQ. While they were not required to fill out the survey, they received prompts requesting them to fill out the survey. As a result, the response rate was close to a 100%. Results from the intake survey are listed in Table 11 for 2741 users. While the app was downloaded over 3,000 times, users outside the US were dropped from the study. Researchers and beta testers were also dropped from the study.

App users differed from the general population with regards to their health conditions (Table 11).

Incidence of asthma among app users and among their children/guardians was much higher than US and CA averages; 15.4 % for adults compared to 7.4% for the US and 8.7 % for CA and for children 18.7 % compared to 8.6 % for US averages, more than double the national average. 14.1 % of the users had heart disease compared to the US average of 10.2 %. There was a big selection bias towards those with health conditions associated with poor AQ (Table 9 and Table 11).

Table 9: Prevalence of health conditions* aggravated by air pollution among app users and their children

	App Users (%)	**Children (%)**
At least 1 health condition	55.1	55.5
More than 1 health condition	13.3	13.8
No health condition	44.9	44.5

*Health conditions – asthma, outdoor allergies, lung disease, heart disease and other

4.2.2 Summary Statistics for App Usage

The sample size for the study was 2741. App users were predominantly iPhone users (75%) (Table 10). Since its launch, the app was opened 66000+ times (Table 10) and AQ information was accessed 164,000+ times (Table 12). Health info was also accessed very frequently (Table 12).

Table 10: Total App downloads and sessions

	N
Number of app downloads (U.S. only)	2741
iPhone users	2105
Android users	636
Total App Visits (since launch)	66,220
Daily Average	~107
Weekly Average	~753

Table 11: Summary Statistics for App Intake Survey (N=2741)

	App Users						US Census[1]/ CDC[2]	CA Census[1]/ CDC[2]
	N	%	M	SD	Min	Max	M	M
Female	1226	44.7	0.447	0.497	0	1	0.513	0.508
Age (>18 yrs.)	2741	-	43.0	15.5	18-24	≥ 65		
Health Conditions								
Heart Disease	385	14.1	0.141	0.345	0	1	0.102	
Lung Disease	102	3.72	0.037	0.189	0	1		
Asthma	421	15.4	0.154	0.361	0	1	0.076	0.078
Allergies	909	33.2	0.332	0.471	0	1		
Other Health Conditions	121	4.41	0.441	0.205	0	1		
Children (<18 yrs.) living in HH	959	35.0	0.350	0.477	0	1	0.250	
Children								
Heart Disease	113	11.8	0.118	0.323	0	1		
Lung Disease	18	1.88	0.019	0.136	0	1		
Asthma	179	18.7	0.187	0.390	0	1	0.084	
Allergies	337	35.1	0.351	0.478	0	1		
Other Health Conditions	32	3.34	0.334	0.180	0	1		
Exercise Frequency			4.04	1.43	1	6		
Have Knowledge of AQI	265	9.67	0.097	0.296	0	1		
Have Knowledge of $PM_{2.5}$[a]	810	38.7	0.387	0.487	0	1		

[1]US and CA Census 2015 data for age and sex; 2010 for children living in household (US Census Bureau, 2010, 2015a)
[2]Centers for Disease Control and Prevention (CDC) 2014 data for heart disease and asthma (CDC, 2014b. 2014c)
[a]Responses weren’t recorded correctly for this question for some users (n=647)

Table 12: App usage summary

App tabs	Information Content	Total Views or Searches
Air Quality	Changes hourly	164,196
Health	Static	87,547
Toxic Release Inventory	Changes based on current location and zip code	23,286
Prizes	Changes daily based on response to behavioral questions	12,328
Learn More	Static	4,594

4.2.3 App Engagement Before and After Email Experiment

We ran regressions to identify the effect of the AP message categories, message framing and other control factors affecting engagement with the app (Refer to Table 13 for complete list of variables). Heteroscedasticity checks were added but standard errors were not clustered. The total number of users considered in the experiment are 2641; there were some users that had not reported an email and only a phone number so they were removed from the experiment.

Table 13: Variables used in the regressions for exploring app engagement

Category	Name and Type	Description
Dependent Variable	open_app	Number of app visits per person per week
Treatment Variable	email_group (dummy)	12 dummies corresponding to the 12 message groups and 1 control group that received no email
	email_cat (dummy)	6 dummies corresponding to the 6 message categories and 1 control group that received no emails
	prepost (dummy)	Time dummy accounting for period before and after the email experiment
	openemail (dummy)	Dummy indicating whether an app user opened the email
Prior engagement controls	notif (dummy)	Dummy to control for the effect of weekly notifications. 0 for disabled always, 1 for enabled always, 2 for switching between enabled/disabled and 3 for android/no data for iPhone.
	wks_inactive5 (10,15 or 20) (dummy)	Dummies to control for user's engagement with the app prior to the experiment. Dummies for this variable represent whether the user has been inactive for periods longer than 5 (10, 15 or 20 weeks) since downloading the app
Demographic and health	age (continuous)	Age of the app user (values range from 1 to 6; see Table 24 in the appendix for categories)

Category	Name and Type	Description
controls		
	gender (dummy)	Gender of the app user
	exercise (continuous)	Exercise frequency of the app user (values range from 1 to 6; see Table 24 in the appendix for categories)
	aqi (dummy)	Knowledge of aqi; also correlated with knowledge of $PM_{2.5}$
	Children (dummy)	Accounting for whether the app user has children (<18 yrs and living in household)
	user_asthma (and other health conditions) (dummy)	Accounting for the user's health conditions aggravated by air pollution
	child_asthma (and other health conditions) (dummy)	Accounting for the health conditions of the user's children (<18 years and living in household) aggravated by air pollution
Time controls	week (dummy)	Week dummies to control for seasonality. Number of dummies correspond to the number of weeks since the user first downloaded the app

The first set of regressions (Table 14) are based on the difference-in-differences model presented in equation 2. Weekly app visits (open_app) is the dependent variable and the different columns split the sample based on the level of activity of users. Column 1 is for all users. Columns 2, 3 and 4 are for users whose maximum period of inactivity is 5,10 or 15 weeks. These time periods were determined by analyzing Google Analytics data for overall app activity which indicated that by about 12-15 weeks, most users had disengaged from the app. The control group for these regressions is the group that received no email messages. The treatment group is split by message categories (a total of 6 categories including the baseline message). Week dummies are included in the model but not in the table.

When looking at overall engagement, results indicate that the email messages are effective among those who have been inactive for up to a period of 15 weeks. There was not too much of a change in engagement for users who have been consistently active with the app (period of inactivity < 5 weeks, N=118) but the email messages were mostly effective at increasing engagement for those who have been inactive for longer periods of time i.e between 5-15 weeks.

While the email messages were effective at increasing engagement among participants who had been inactive between 5-15 weeks (N=290 out of 2641), there was no significant difference among users for those had been inactive for 15 weeks or longer (N=2233 out of 2641 users). There is no significant difference in engagement among those who have been inactive longer than 20 weeks or so (regression not shown there); it can also be assumed that these users have deleted the app. With the current level of technology, we are unable to record which users have deleted the app so users who have been inactive for a long time appear to be the same as those who have deleted the app.

When considering the engagement for different email messages, there is not much difference observed among the different categories. It appears that all message categories, except for the child cognition category among some groups, were effective at increasing engagement. This may suggest that users are responding to the emails as a reminder to check the app rather than the content of the email. This is also in contrast to the MTurk survey results that showed the exercise and AP invisibility categories to be the most effective. This is not an equivalent comparison though since the control groups are different; the control group in the survey is the baseline message respondents whereas in the app it is the group that received no message. That is investigated in a later set of regressions.

One confounding factor is that with most smartphones it is possible to read the content of the email without actually opening the email, making it impossible to entirely disentangle the effect of receiving an email and reading it. In order to test the potential additional effect of opening the email, these regressions include treatment terms for those who received the email not those who opened it. A dummy was added for those who opened the email and its coefficient was 0.08 (significant at the 1 % level) indicating that those who opened the email were more likely to check the app versus those who received it but did not open it. Because of the aforementioned issue, this effect might be underestimated.

Table 14: App engagement before and after the email experiment split by levels of user activity. Treatment groups are the 6 message categories in reference to the control group that received no email.

Study Characteristic	**(1) Check App All users (N=2641)**	**(2) Check App Inactive 5 weeks (N=2523)**	**(3) Check App Inactive 10 weeks (N=2362)**	**(4) Check App Inactive 15 weeks (N=2233)**
Treatment Groups				
Baseline	0.37***	0.11***	0.12***	0.12***
	(0.11)	(0.03)	(0.04)	(0.04)
Exercise	0.34***	0.07***	0.06***	0.09***
	(0.11)	(0.02)	(0.02)	(0.02)
Child Asthma	0.36***	0.04	0.05**	0.05**
	(0.11)	(0.03)	(0.02)	(0.02)
Child Cognition	0.29***	-0.03	-0.02	-0.02
	(0.11)	(0.03)	(0.02)	(0.02)
Alzheimer's	0.32***	0.06**	0.08***	0.09***
	(0.10)	(0.03)	(0.03)	(0.03)
AP Invisibility	0.38***	0.11***	0.09***	0.12***
	(0.11)	(0.03)	(0.03)	(0.03)
Email Groups				
Baseline	-0.01	-0.05**	-0.07***	-0.07***
	(0.03)	(0.02)	(0.02)	(0.02)
Exercise	0.02	-0.05***	-0.05***	-0.07***
	(0.03)	(0.02)	(0.02)	(0.02)
Child Asthma	0.03	0.03	-0.04**	-0.06***
	(0.03)	(0.02)	(0.02)	(0.02)
Child Cognition	0.04	0.06***	0.03*	0.03
	(0.03)	(0.02)	(0.02)	(0.02)
Alzheimer's	0.03	-0.01	-0.04**	-0.05***
	(0.03)	(0.02)	(0.02)	(0.02)
AP Invisibility	-0.06**	-0.07***	-0.07***	-0.10***
	(0.03)	(0.02)	(0.02)	(0.02)
Before/After Dummy	-5.34***	-4.86***	-4.78***	-4.70***
	(0.23)	(0.20)	(0.19)	(0.19)
Prior Engagement Controls				

Study Characteristic	(1) Check App All users (N=2641)	(2) Check App Inactive 5 weeks (N=2523)	(3) Check App Inactive 10 weeks (N=2362)	(4) Check App Inactive 15 weeks (N=2233)
Weeks of Inactivity (5 weeks)	5.03***	-	-	-
	(0.18)	-	-	-
Notification (Enabled Always)	0.03***	0.03***	0.00	0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Notification (Alternating between Disabled/Enabled)	0.50***	0.49***	0.31***	0.20***
	(0.04)	(0.04)	(0.03)	(0.02)
Notification (Status unknown for some devices)	-0.18***	-0.14***	-0.14***	-0.13***
	(0.01)	(0.01)	(0.01)	(0.01)
Health and Demographic Controls				
Age	-0.01	0.01	0.02**	0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Female	0.06***	0.12***	0.10***	0.08***
	(0.01)	(0.01)	(0.01)	(0.01)
Exercise	-0.02***	0.01**	0.01***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)
Heart Disease	0.19***	0.18***	0.17***	0.18***
	(0.02)	(0.02)	(0.02)	(0.02)
Lung Disease	0.10***	0.11***	-0.03	-0.00
	(0.03)	(0.03)	(0.02)	(0.02)
Asthma	0.02	0.06***	0.08***	0.06***
	(0.02)	(0.02)	(0.01)	(0.01)
Allergies	0.04***	0.07***	0.08***	0.07***
	(0.01)	(0.01)	(0.01)	(0.01)
Other health conditions	0.72***	0.39***	0.38***	0.38***
	(0.07)	(0.04)	(0.04)	(0.04)
Children	-0.01	0.01	0.02**	0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Children Heart Disease	0.12***	0.10***	0.06**	0.05**
	(0.03)	(0.03)	(0.02)	(0.02)
Children Lung Disease	-0.44***	-0.28***	-0.15***	-0.15***
	(0.04)	(0.03)	(0.03)	(0.03)
Children Asthma	0.33***	0.05**	-0.02	-0.00
	(0.04)	(0.02)	(0.02)	(0.02)
Children Allergies	0.05**	-0.03**	-0.06***	-0.04***
	(0.02)	(0.01)	(0.01)	(0.01)
Children Other health conditions	-0.13**	0.01	-0.02	-0.03

Study Characteristic	**(1) Check App All users (N=2641)**	**(2) Check App Inactive 5 weeks (N=2523)**	**(3) Check App Inactive 10 weeks (N=2362)**	**(4) Check App Inactive 15 weeks (N=2233)**
	(0.05)	(0.04)	(0.04)	(0.04)
Knowledge of AQ	0.05***	0.05***	0.07***	0.02
	(0.02)	(0.02)	(0.02)	(0.01)
Constant	4.86***	4.69***	4.65***	4.61***
	(0.21)	(0.20)	(0.19)	(0.19)
Observations	163,086	160,160	156,253	151,590
Adjusted R-squared	0.14	0.07	0.08	0.08
F	37.07	35.61	39.81	39

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The trends for most of the health and demographic control variables were similar to the MTurk Survey and as expected. In contrast to the survey, those who were older were less likely to check the message but the comparison might not be a good one because the average age of the app users was higher than those who took the survey. Those with health conditions aggravated by air pollution, or with children with those health conditions, were much more likely to check the app compared to those without those health conditions. Women were 6% more likely to check the app compared to men on a weekly basis.

In the second set of regressions, the message framing was considered as the treatment variable with all the same specifications as the regressions in Table 14. Only the coefficients for the message framing are presented in Table 15 for the purposes of brevity. All other coefficients were similar. There was no significant difference among the framing types, similar to the survey results. Although in the survey the combined exercise message was more effective than either the positive or negative version along, no such difference was observed in the app email experiment. Similarly, hypothesis A did not hold for the app experiment either; neither positively or negatively framed were more effective than the other at influencing engagement.

Table 15: Coefficients for message framing in the app email experiment

Treatment Groups	
Baseline	0.37***
Exercise Positive	0.41***
Exercise Negative	0.25*
Exercise Combined	0.36***
Child Asthma Positive	0.33***
Child Asthma Negative	0.38***
Child Cognition Positive	0.21*
Child Cognition Negative	0.35***
Alzheimer's Positive	0.29***
Alzheimer's Negative	0.34***
Invisibility Positive	0.40***
Invisibility Negative	0.36***
Total N	**2641**

In order to be consistent while comparing the results of the survey and app, the next set of regressions were run with the baseline message as the control group (Table *16*). Compared to the baseline message, none of the message categories were more effective; the child cognition had the reverse effect and actually decreased engagement relative to the baseline. This lends further support to the observation that app users might simply be responding to the email as a reminder to check the app, rather than responding to the content of the message.

Table 16: Comparing the effectiveness of message categories in the MTurk Survey and App email experiment

Treatment Groups	MTurk Survey	App experiment
Baseline	(Control)	(Control)
Exercise	0.40***	NS
Child Asthma	NS	NS
Child Cognition	NS	-0.09*
Alzheimer's	NS	NS

AP Invisibility	0.34***	NS
Total N	**835**	**2376**

4.2.4 Feedback Survey Results

All users with an email (N=2641, 99 users only submitted phone numbers) were solicited for the feedback survey, conducted a few weeks after the email experiment. Despite a low response rate of about 4% (N=99), the data collected was crucial in understanding the usability of the app. There was a strong selection bias for respondents with a high engagement; over 70% of the respondents checked the app at least once a week and 18% checked it daily. The first step was to understand the app user's experience with the app – their comprehensibility, relevance and learning associated with the information presented in the app and their ability to protect their health based on that information (Figure 1).

Learning was assessed by comparing responses from the intake survey and feedback survey for knowledge of the AQI and the AQI range (Table 17). There was a big increase in learning, however it is difficult to compare because of the bias in the feedback survey among those who were more engaged with this information.

Table 17: Pre/post learning of AQI among app users

	Intake Survey (N=2741)	Feedback Survey (N=99)
Knowledge of AQI		
Yes	9.7%	70.1%
No	90.3%	29.6%
Knowledge of AQI range[a]		
Yes	9.4%	97.1%
No	90.6%	2.9%

[a]N=69 based on those who responded yes to knowledge of AQI

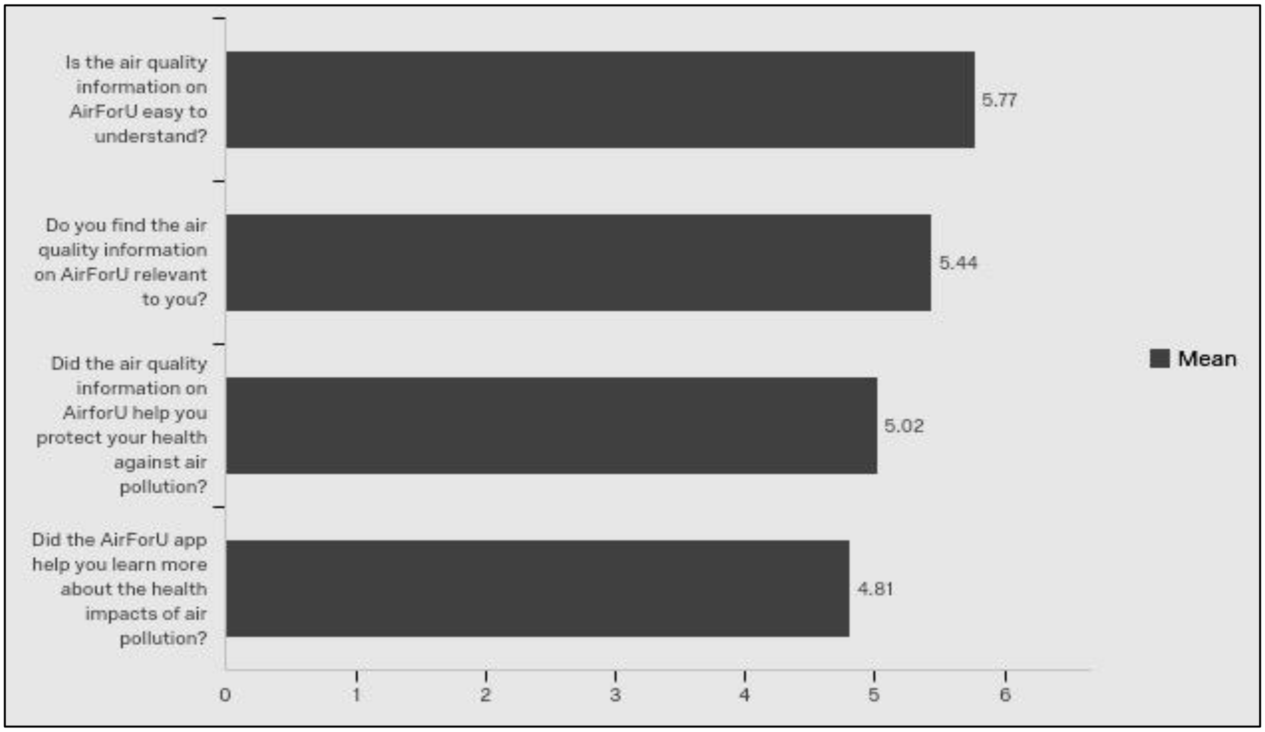

Figure 1: App users' responses for their experience with the AirForU app*(N=99)*. Numbers *at the end of the bars correspond* to means on a 7 point Likert scale (1= Strongly Disagree; 7== Strongly Disagree)

The primary purpose of AQI monitoring and reporting programs is the change in behavior to reduce health risk associated with air pollution. We made a list of all the health protective behaviors that app users could adopt and measured how many people adopted them (Table *18*). The most common action and one of the most effective (based on EPA guidelines) is not exercising outdoors during high air pollution. Those who checked the app more frequently were more likely to adopt these behaviors.

Table 18: Adoption of health protecting behaviors based on the information provided in the AirForU app as measured in the feedback survey (N=99)

Health Protective Behavior	**%**	**Number**
Talk to your healthcare provider about issues associated with poor air quality	5.4	14
Close windows during poor air quality episodes	20.2	52
Wear a breathing mask	4.4	11
Clean or change filters in your air conditioner more frequently	12.4	32
Missed school or work	1.6	4
Use an air filter/purifier	14.0	36
Change your outdoor exercise schedule	21.7	56
Plan for potential asthma attacks	5.4	14
Use your air conditioner more frequently	12.0	31
Other	3.2	8
Total	100 %	258

5 Conclusion

Through the MTurk survey, we learned that people are more likely to check AQ in response to AP messages that don't induce a strong fear appeal. When considering all groups, messages based on exercise and general invisibility of AP were the most effective. These were the least threatening messages. This might indicate that in the case of air pollution, resorting to fear appeal messages might not be effective.

In addition, positively framed messages were not more effective than negatively framed messages in this study. A combined positive and negative mixed message presenting a problem and then also providing a resolution to that message (exercise combined message) was the most effective message framing among all messages. Studies testing the effect of combined message framing are not as common as valence framing studies but combined framing has been more successful in some instances (Treiber, 1986; Wilson et al., 1990).

However, the results of the survey were in alignment with previous literature on targeted messages. Women, parents/guardians, those with asthma, those with a better knowledge of AQ are more likely to check AQ and thus potentially engage in protecting behaviors. Messages targeted at certain groups were more effective among those groups for the most part and had a higher potential for increasing effectiveness with AQ information.

Engagement for app users was widely varied. Users with health conditions or with children with health conditions were more engaged with the app in general, as were women compared to men. These results indicate that groups that are most impacted by AP are engaging with this information. Yet, despite the fact that the group of app users had a high proportion of sensitive groups, engagement tended to be short-lived for most users. Within 10-15 weeks most users (>2000 users) had disengaged with the app. There was a small group of highly motived individuals (~100 users) who remained actively engaged with the app from the time they first downloaded it. This group also reported the adoption of health protective behaviors as a result of the information provided in the app. Less engaged individuals may have adopted these behaviors as well but because they did not partake in the exit survey this is mere speculation.

By replicating (as much as possible) the MTurk survey among the app users and by delivering the same AP messages via email and measuring engagement with the app, we were able to compare the results of a survey and field experiment. Survey results were quite different from the field experiment. Engagement after receiving the emails was highly dependent on levels of engagement prior to receiving the emails. There was little effectiveness of messaging for users with a previously high engagement, while those who were less engaged became more engaged over time with messaging. However, there was little difference in engagement based on the content of the email – the message category or the framing type. Emails are an effective way of re-engaging some users and should be sent at "appropriately timed" intervals as users seems to forget to use the app. It is much harder to re-engage users who have been inactive for long periods of time (20 weeks or longer).

Another limitation of this study was the timing of the app experiment. The experiment was conducted over a year and half after the app's launch, which also coincided with much of the recruiting effort and a large portion of the total downloads. At the time of the experiment, many users had disengaged from the app and perhaps even deleted it from their phones.

While one of the limitations of this study is its external validity to the general population, it is not necessarily a disadvantage. There was a self-selection bias among app users; those affected by AQ were more likely to download and use the app. These individuals are more likely to contribute to the health burden and thus it is more important that they engage with AP information and adopt health protecting behaviors compared to the general population. Besides, vulnerable groups constitute a large part of the population anyway; millions of people fall into these groups.

Even though a majority of the app users were disproportionately affected by air pollution, they tended to lose interest in AQ information over time. One of the reasons could be that often AQ information didn't change much for long periods of time. While notifications (sent through the app) and emails were effective at re-engaging app some of the app users, they were not effective for those who had been inactive for long periods. Timely interventions might be necessary to keep users engaged over time. Different modes of reminders (emails, notifications) may be effective for different groups. Survey results indicate that targeted messages have potential but more research would have to be done to understand the effect of targeted messages.

Environmental information programs have great potential at increasing awareness of environmental pollution and encouraging the adoption of health protective behaviors especially among those that are most impacted. Ultimately, this would lead to a lower health burden. One big challenge is to keep people motivated in engaging with the information over time. Personalized information and timely reminders may play an important role in influencing engagement and improving public health protection.

Appendix 1 – MTurk Survey Additional Information and Results

A MTurk Survey Air Pollution Message Categories and Framing Type

1) Baseline
   a. Protect your health. Check your local air quality on AirForU today before engaging in outdoor activities!
2) Exercise (Positive, Negative and Mixed)
   a. Do you know that exercising outdoors when air pollution is low is beneficial to your health? Protect your health. Check your local air quality on AirForU today before engaging in outdoor activities!
   b. Do you know that exercising outdoors when air pollution is high can harm your health? Protect your health. Check your local air quality on AirForU today before engaging in outdoor activities!
   c. Do you know that while exercising is beneficial for your health, exercising outdoors when air pollution is high can harm your health? Protect your health. Check your local air quality on AirForU today before engaging in outdoor activities!
3) Child Asthma (Positive and Negative)
   a. Do you know that high air pollution can cause or worsen childhood asthma? Avoiding air pollution can reduce this risk. Check your local air quality on AirForU today before engaging in outdoor activities with your children!
   b. Do you know that high air pollution can cause or worsen childhood asthma? Exposure to air pollution can increase this risk. Check your local air quality on AirForU today before engaging in outdoor activities!
4) Child Cognition (Positive and Negative)
   a. Do you know that air pollution slows cognition in children by affecting their brain development? Avoiding air pollution can reduce this risk. Check your local air quality on AirForU today before engaging in outdoor activities!
   b. Do you know that air pollution slows cognition in children by affecting their brain development? Exposure to air pollution can increase this risk. Check your local air quality on AirForU today before engaging in outdoor activities with your children!
5) Alzheimer’s (Positive and Negative)
   a. Do you know that air pollution is linked to Alzheimer's disease? Avoiding air pollution can reduce this risk. Check your local air quality on AirForU today before engaging in outdoor activities!
   b. Do you know that air pollution is linked to Alzheimer's disease? Exposure to air pollution can increase this risk. Check your local air quality on AirForU today before engaging in outdoor activities!
6) AP Invisibility (Positive and Negative)
   a. Do you know that harmful air pollution is often invisibility to the naked eye? Avoiding air pollution can reduce your risk of harmful health effects. Check your local air quality on AirForU today before engaging in outdoor activities!
   b. Do you know that harmful air pollution is often invisibility to the naked eye? Exposure to air pollution can increase your risk of harmful health effects. Check your local air quality on AirForU today before engaging in outdoor activities!

B Additional Summary Statistics for MTurk Survey

Table 19: MTurk Survey: Summary Statistics for Demographics

	N	%	M	SD	Min	Max
Duration of survey (minutes)	835	-	7.43	7.53	1.08	73.7
Gender	835		0.430	0.495	0	1
Male (value=0)	476	57.0				
Female (value =1)	359	43.0				
Age	835					
18-24 years	103	12.3				
25-34 years	384	46.0				
35-44 years	195	23.3				
45-54 years	98	11.7				
55-64 years	40	4.80				
65 years or older	15	1.80				
Have Asthma	96	11.5	0.115	0.319	0	1
Children (<18 years) living in household	246	29.5	0.295	0.456	0	1
Children (<18 years) with asthma	58	23.6	0.236	0.425	0	1
Have Knowledge of AQI	151	18.1	0.181	0.385	0	1
Have Knowledge of PM2.5	336	40.3	0.403	0.491		
Air quality after 2 pm (Incorrect)	13	1.56				
Particulate matter with diameter **less**	336	40.3				
Performance measurement	43	5.16				
Powdered metallics with diameter	10	1.20				
I don't know	432	51.8				
Frequency of outdoor exercise[1]	430					
Once a year or less	46	10.7				
Several times a year	44	10.2				
A few times a month	91	21.2				
1-2 times a week	114	26.5				
3-4 times a week	89	20.7				
5 or more times a week	46	10.7				
Annual household income[1]	430					
Less than $24,999	99	23.0				
$25,000 to $49,999	134	31.2				
$50,000 to $74,999	106	25.6				
$75,000 to $99,999	28	6.51				
$100,000 to $149,999	53	12.3				
More than $150,000	10	2.33				
Highest level of education[1]	430					
Less than high school	5	1.2				
High school degree of equivalent	61	14.2				
Some college but no degree	99	23.0				
Associate or technical degree	53	12.3				
Bachelor's degree	165	38.4				

Graduate degree/professional	47	10.9
Race/Ethnicity[1]	430	
American Indian or Alaska Native	2	0.47
Asian	33	7.67
Black or African American	33	7.67
Native Hawaiian or Pacific Islander	0	0.00
Hispanic/Latino	26	6.05
White/Caucasian	312	72.6
Other or Mixed	24	5.58

Table 20: MTurk Survey: Summary Statistics for Air Pollution Messages (N=835)

	N	M	SD	Min	Max
Baseline Statement					
Comprehensibility	835	6.2	0.97	1	7
Realism	835	5.5	1.3	1	7
Relevance	835	5.4	1.4	1	7
Check AQ	835	4.8	1.6	1	7
Exercise Negative					
Comprehensibility	277	6.2	1.0	1	7
Realism	277	5.5	1.3	1	7
Relevance	277	5.4	1.4	1	7
Check AQ	277	5.2	1.6	1	7
Exercise Positive					
Comprehensibility	278	6.3	0.9	1	7
Realism	278	5.8	1.2	1	7
Relevance	278	5.3	1.6	1	7
Check AQ	278	5.2	1.6	1	7
Exercise Mixed					
Comprehensibility	280	6.2	1.0	1	7
Realism	280	5.9	1.2	1	7
Relevance	280	5.5	1.4	1	7
Check AQ	280	5.5	1.7	1	7
Child Asthma Negative					
Comprehensibility	421	6.3	0.9	1	7
Realism	421	5.9	1.1	1	7
Relevance	421	5.6	1.3	1	7
Check AQ	421	4.8	1.4	1	7
Child Asthma Positive					
Comprehensibility	414	6.3	0.9	2	7
Realism	414	6.1	1.0	2	7
Relevance	414	4.3	2.0	1	7
Check AQ	414	4.7	1.8	1	7
Child Cognition Negative					
Comprehensibility	418	6.4	0.8	2	7
Realism	418	6.1	1.0	1	7
Relevance	418	4.2	2.1	1	7
Check AQ	418	4.8	1.8	1	7
Child Cognition Positive					

	N	M	SD	Min	Max
Comprehensibility	417	6.1	1.0	1	7
Realism	417	5.4	1.4	1	7
Relevance	417	4.4	1.9	1	7
Check AQ	417	4.8	1.7	1	7
Alzheimer's Negative					
Comprehensibility	416	6.2	0.9	1	7
Realism	416	5.5	1.3	1	7
Relevance	416	4.2	2.1	1	7
Check AQ	416	4.9	1.8	1	7
Alzheimer's Positive					
Comprehensibility	419	6.1	1.0	2	7
Realism	419	4.9	1.6	1	7
Relevance	419	4.6	1.7	1	7
Check AQ	419	4.7	1.7	1	7
Invisibility Negative					
Comprehensibility	434	6.1	1.0	2	7
Realism	434	5.1	1.5	2	7
Relevance	434	4.8	1.7	1	7
Check AQ	434	5.2	1.8	1	7
Invisibility Positive					
Comprehensibility	401	6.2	0.9	1	7
Realism	401	5.9	1.0	1	7
Relevance	401	5.5	1.2	1	7
Check AQ	401	5.2	1.4	1	7

Table 21: Correlation Matrix for MTurk Survey Questions

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
1	1															
2	0.989	1														
3	0.017	0.021	1													
4	-0.066	-0.058	0.466	1												
5	-0.045	-0.051	0.696	0.462	1											
6	-0.024	-0.018	0.243	0.532	0.259	1										
7	0.000	0.000	0.120	0.119	0.102	0.107	1									
8	0.000	0.000	0.032	0.062	0.053	0.109	0.090	1								
9	0.000	0.002	0.151	0.095	0.110	0.074	0.089	-0.062	1							
10	0.000	0.001	0.088	0.063	0.158	0.075	0.167	0.030	0.059	1						
11	0.000	-0.002	0.061	0.016	0.045	0.020	0.040	0.107	0.045	-0.032	1					
12	0.000	0.001	0.057	0.015	0.006	0.017	-0.097	0.001	-0.111	-0.029	0.327	1				
13	0.000	0.000	-0.094	0.013	-0.035	0.038	-0.041	0.171	0.075	-0.040	0.061	0.033	1			
14	0.000	0.002	0.004	0.019	0.026	-0.019	0.031	-0.015	-0.037	0.068	0.044	0.080	0.184	1		
15	0.000	-0.001	0.077	0.040	0.117	0.033	0.064	-0.004	0.068	0.239	0.017	0.033	0.002	0.311	1	
16	0.000	0.006	0.061	0.020	0.107	0.013	-0.083	0.103	-0.021	-0.022	0.101	0.150	0.096	0.089	0.182	1

1 – Message Category; 2 – Message Framing; 3 – Check AQ; 4 – Realism; 5 – Relevance; 6 – Comprehension; 7 – Gender; 8 – Age; 9 – Asthma; 10 – Children; 11 – Knowledge of AQ; 12 – Knowledge of PM; 13 – Non-white ethnicity/race; 14 – Education; 15 – Income; 16 – Exercise Frequency

C Additional Regressions for MTurk Survey

Table 22: Comprehension and Realism Regressions for question categories relative to the baseline category (N=835)

Study Characteristic	(1) Comprehensibility	(2) Realism
Category Treatment		
Exercise	0.04	0.36***
	(0.05)	(0.06)
Child asthma	0.12***	0.55***
	(0.04)	(0.06)
Child Cognition	-0.04	-0.05
	(0.05)	(0.06)
Alzheimer's	-0.11**	-0.53***
	(0.05)	(0.07)
AP Invisibility	0.03	0.38***
	(0.05)	(0.06)
Controls		
Age > 55 years	0.08*	0.21***
	(0.05)	(0.06)
Female	0.18***	0.25***
	(0.03)	(0.04)
Asthma	0.07*	0.23***
	(0.04)	(0.05)
Children	0.11***	0.14***
	(0.03)	(0.04)
Knowledge of AQ	0.18***	0.23***
	(0.03)	(0.04)
Constant	6.06***	5.29***
	(0.04)	(0.05)
Observations	5,010	5,010
Adjusted R-squared	0.02	0.10
F	13.55	55.43

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 23: Regression results for message framing relative to the baseline (N=430)

Study Characteristic	**(1) Comprehensibility**	**(2) Realism**	**(3) Relevance**	**(4) Check AQ**
Question Treatment				
Exercise negative	0.07	0.25**	-0.07	0.29**
	(0.07)	(0.11)	(0.12)	(0.13)
Exercise positive	-0.11	0.37***	0.18	0.29**
	(0.08)	(0.10)	(0.12)	(0.12)

Study Characteristic	(1) Comprehensibility	(2) Realism	(3) Relevance	(4) Check AQ
Exercise mixed	0.07	0.35***	0.13	0.61***
	(0.07)	(0.09)	(0.11)	(0.12)
Child asthma positive	0.04	0.47***	-1.11***	-0.16
	(0.05)	(0.07)	(0.13)	(0.11)
Child asthma negative	0.13**	0.50***	-1.17***	0.09
	(0.05)	(0.08)	(0.14)	(0.12)
Child cognition positive	-0.14**	-0.20**	-1.05***	-0.08
	(0.07)	(0.10)	(0.13)	(0.13)
Child cognition negative	0.03	-0.02	-1.12***	0.01
	(0.05)	(0.09)	(0.14)	(0.12)
Alzheimer's positive	-0.21***	-0.73***	-0.92***	-0.12
	(0.06)	(0.12)	(0.12)	(0.12)
Alzheimer's negative	-0.02	-0.41***	-0.47***	0.18
	(0.06)	(0.11)	(0.11)	(0.11)
AP Invisibility positive	0.03	0.30***	0.17**	0.37***
	(0.05)	(0.08)	(0.09)	(0.10)
AP Invisibility negative	-0.02	0.32***	0.09	0.30***
	(0.06)	(0.08)	(0.09)	(0.10)
Controls				
Age > 55 years	0.17	0.23	0.42*	0.41
	(0.12)	(0.14)	(0.22)	(0.26)
Frequent Outdoor Exercise	0.04	0.07	0.26**	0.12
	(0.09)	(0.09)	(0.11)	(0.13)
College education	0.03	0.08	0.00	0.05
	(0.08)	(0.08)	(0.11)	(0.13)
Above median income	-0.07	-0.01	0.16	0.04
	(0.08)	(0.09)	(0.11)	(0.13)
Non-white	0.00	0.07	0.27**	0.40***
	(0.08)	(0.09)	(0.12)	(0.14)
Female	0.16**	0.26***	0.23**	0.30**
	(0.07)	(0.08)	(0.11)	(0.13)
Asthma	0.22**	0.39***	0.56***	0.76***
	(0.10)	(0.12)	(0.15)	(0.16)
Children	0.22***	0.22**	0.55***	0.27*
	(0.08)	(0.10)	(0.13)	(0.15)
Children with asthma	-0.33*	-0.40*	-0.02	-0.05
	(0.19)	(0.21)	(0.22)	(0.28)
Knowledge of AQ	0.02	0.00	0.13	0.20
	(0.11)	(0.12)	(0.15)	(0.17)
Constant	6.06***	5.27***	4.77***	4.23***
	(0.10)	(0.11)	(0.13)	(0.15)
Observations	2,580	2,580	2,580	2,580
Adjusted R-squared	0.03	0.10	0.15	0.06
F	3.057	13.69	15.55	7.587

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix 2 – AirForU App

A App Intake Survey – Additional Summary Statistics

Table 24: App Intake Survey Summary Statistics (N=2740)

	N	%
Age		
18-24 years	357	13.02
25-30 years	387	14.12
31-50 years	1144	41.77
51-64 years	531	19.37
65 years or older	321	11.71
Frequency of Outdoor exercise		
Once a year or less	163	5.95
Several times a year	269	9.82
A few times a month	491	17.93
1-2 times a week	656	23.95
3-4 times a week	686	25.05
5 or more times a week	474	17.31
Knowledge of $PM_{2.5}$		
Air quality after 2 pm	24	1.15
Particulate matter with a diameter less than 2.5 □m	810	38.68
Performance measurements standards for air quality	45	2.15
Powdered metallics with a diameter less than 2.5 □m	33	1.58
I don't know	1182	56.45

B App Email Experiment Summary Statistics

Delivery and total open rates for email experiment (Figure 2, Figure 3 and Figure 4) for 5 consecutive emails using data from MailChimp.

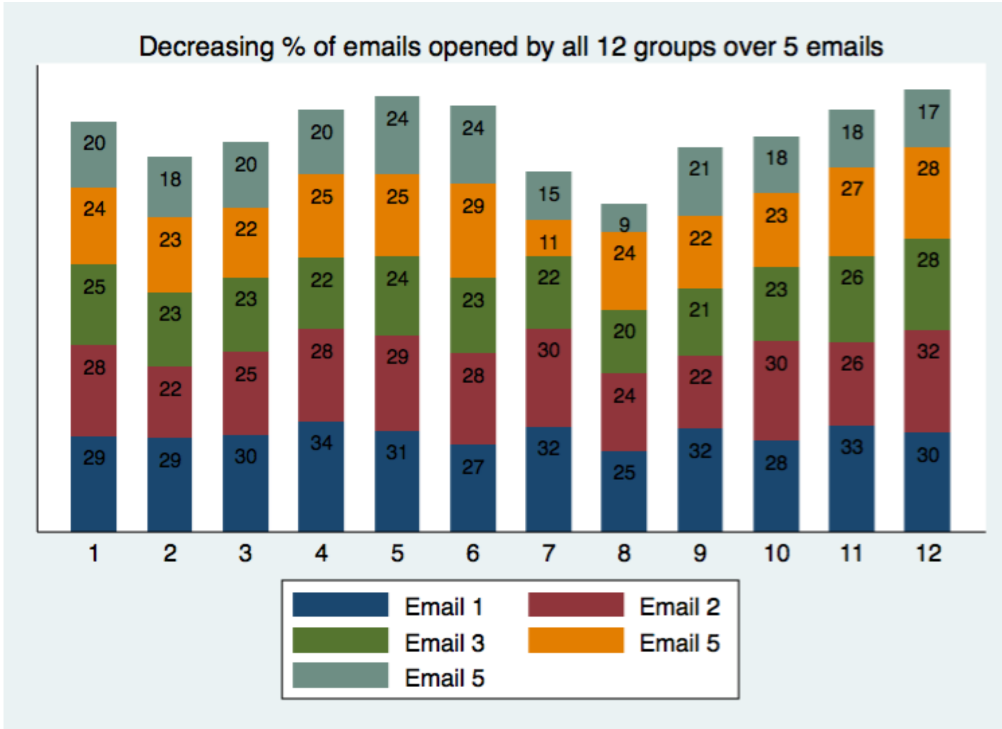

Figure 2: Decreasing % of emails opened by all groups over 5 consecutive emails

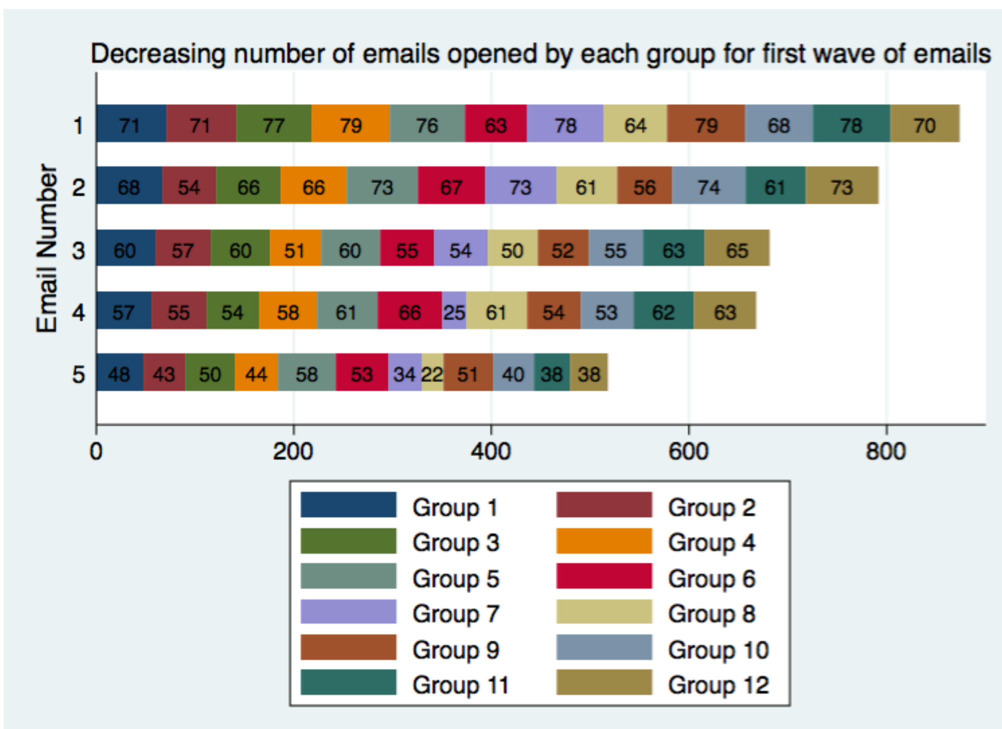

Figure 3: Decreasing number of emails opened by all 12 groups over 5 consecutive weeks

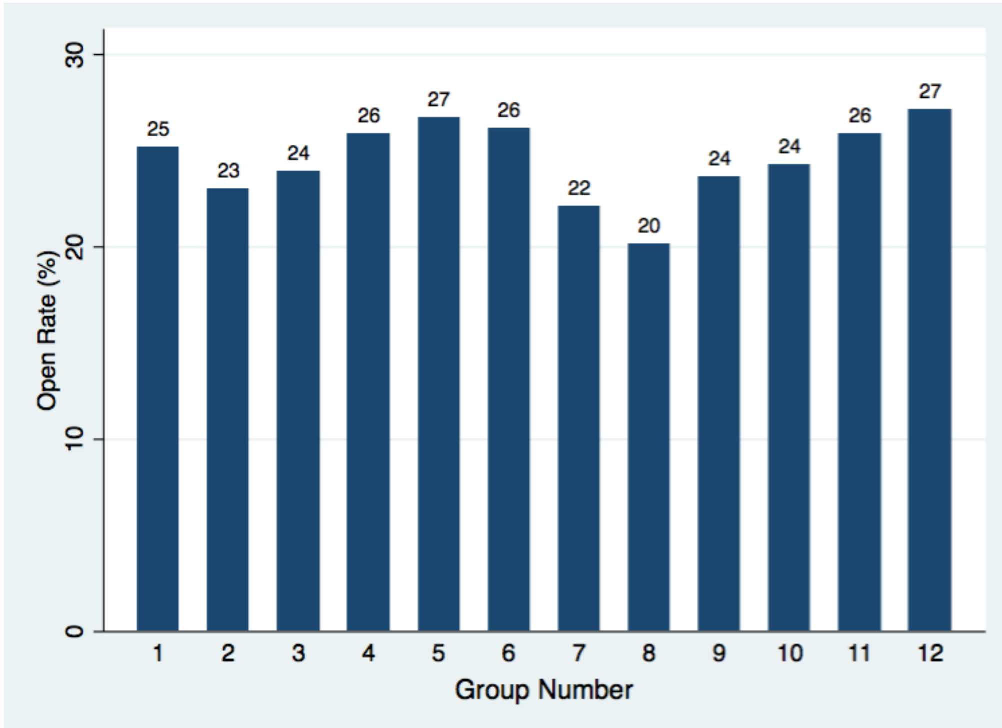

Figure 4: Average email open rate (%) for all 12 groups for 5 a total of 5 emails

Appendix 3 – Complete MTurk Survey

(Page numbers below correspond to the display screens for the survey question)

(Page 1)
Thank you for your participation in the AirForU air quality study. You must be 18 years or older to participate in this study.

As a study participant, you are entitled to know the following:
Your participation is voluntary.
Your answers are confidential.
There are no right or wrong answers.

For any questions regarding this survey or the research being conducted, please contact the researchers at engage@ioes.ucla.edu or at the address below:
Magali Delmas
Professor
UCLA Institute of the Environment and Sustainability
LaKretz Hall, Suite 300 Phone: 310-825-9310

I have read the above information and by clicking on the next page I agree to participate in the study.

(Page 2)
About the Survey

AirForU is an air quality app, developed by researchers at UCLA, that provides app users with information about real-time air quality conditions for cities throughout the US.

In this survey, we are testing which messages encourage people to check the air quality on AirForU. This information can be used by people to take steps to protect their health against air pollution.

In the following pages, we will show you statements about the effects of air pollution. Please rate these statements on their comprehensibility, realism, relevance, and effectiveness.

Thank you for your help!

(Page 3)
For the purposes of this study:

Outdoor activities include activities such as fast walking, running, running, biking, sports and intensive gardening conducted outdoors. These activities increase breathing rates and thus can be harmful when air pollution levels are high, even for healthy people.

(Respondents receive 6 questions (from a total of 12) one from each of the 6 categories. Each category contains between 1-3 questions. The order of categories is random and questions within each category are randomly assigned)

(Page 4 – Category 1: Baseline)

Protect your health. Check your local air quality on AirForU today before engaging in outdoor activities!

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
I understand this message	❍	❍	❍	❍	❍	❍	❍
The impacts described are realistic	❍	❍	❍	❍	❍	❍	❍
This message is relevant to me	❍	❍	❍	❍	❍	❍	❍
After reading this message, I would check air quality	❍	❍	❍	❍	❍	❍	❍

(Page 5 – Category 2: Exercise – Only one of the following statements)

Do you know that exercising outdoors when air pollution is high can harm your health? Protect your health. Check your local air quality on AirForU today before engaging in outdoor activities!
OR
Do you know that exercising outdoors when air pollution is low is beneficial to your health? Protect your health. Check your local air quality on AirForU today before engaging in outdoor activities!
OR
Do you know that while exercising is beneficial for your health, exercising outdoors when air pollution is high can harm your health? Protect your health. Check your local air quality on AirForU today before engaging in outdoor activities!

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
I understand this message	❍	❍	❍	❍	❍	❍	❍
The impacts described are realistic	❍	❍	❍	❍	❍	❍	❍
This message is relevant to me	❍	❍	❍	❍	❍	❍	❍
After reading this message, I would check air quality	❍	❍	❍	❍	❍	❍	❍

(Page 6 – Category 3: Child asthma – Only one of the following statements)

Do you know that high air pollution can cause or worsen childhood asthma? Avoiding air pollution can reduce this risk. Check your local air quality on AirForU today before engaging in outdoor activities with your children!
OR
Do you know that high air pollution can cause or worsen childhood asthma? Exposure to air pollution can increase this risk. Check your local air quality on AirForU today before engaging in outdoor activities!

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
I understand this message	❍	❍	❍	❍	❍	❍	❍
The impacts described are realistic	❍	❍	❍	❍	❍	❍	❍
This message is relevant to me	❍	❍	❍	❍	❍	❍	❍
After reading this message, I would check air quality	❍	❍	❍	❍	❍	❍	❍

(Page 7 – Category 4: Child cognition – Only one of the following statements)

Do you know that air pollution slows cognition in children by affecting their brain development? Avoiding air pollution can reduce this risk. Check your local air quality on AirForU today before engaging in outdoor activities!
OR
Do you know that air pollution slows cognition in children by affecting their brain development? Exposure to air pollution can increase this risk. Check your local air quality on AirForU today before engaging in outdoor activities with your children!

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
I understand this message	❍	❍	❍	❍	❍	❍	❍
The impacts described are realistic	❍	❍	❍	❍	❍	❍	❍
This message is relevant to me	❍	❍	❍	❍	❍	❍	❍
After reading this message, I would check air quality	❍	❍	❍	❍	❍	❍	❍

(Page 8 – Category 5: Alzheimer's– Only one of the following statements)

Do you know that air pollution is linked to Alzheimer's disease? Avoiding air pollution can reduce this risk. Check your local air quality on AirForU today before engaging in outdoor activities!
OR
Do you know that air pollution is linked to Alzheimer's disease? Exposure to air pollution can increase this risk. Check your local air quality on AirForU today before engaging in outdoor activities!

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
I understand this message	❍	❍	❍	❍	❍	❍	❍
The impacts described are realistic	❍	❍	❍	❍	❍	❍	❍
This message is relevant to me	❍	❍	❍	❍	❍	❍	❍
After reading this message, I would check air quality	❍	❍	❍	❍	❍	❍	❍

(Page 9 – Category 6: Perception of air quality – Only one of the following statements)

Do you know that harmful air pollution is often invisibility to the naked eye? Avoiding air pollution can reduce your risk of harmful health effects. Check your local air quality on AirForU today before engaging in outdoor activities!
OR
Do you know that harmful air pollution is often invisibility to the naked eye? Exposure to air pollution can increase your risk of harmful health effects. Check your local air quality on AirForU today before engaging in outdoor activities!

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
I understand this message	❍	❍	❍	❍	❍	❍	❍
The impacts described are realistic	❍	❍	❍	❍	❍	❍	❍
This message is relevant to me	❍	❍	❍	❍	❍	❍	❍
After reading this message, I would check air quality	❍	❍	❍	❍	❍	❍	❍

(Page 10; respondents only see the 6 choices they received in the 6 questions above and these choices appear randomly)

Which of the messages that you viewed earlier would be the most persuasive in encouraging you to check air quality? (Rank them in order of persuasiveness with 1 being the most persuasive and 6 being the least persuasive. Please assign only one ranking for each option.)
______ Protect your health. Check your local air quality on AirForU today before engaging in outdoor activities!
______ Do you know that exercising outdoors when air pollution is high can harm your health? Protect your health. Check your local air quality on AirForU today before engaging in outdoor activities!
______ Do you know that exercising outdoors when air pollution is low is beneficial to your health? Protect your health. Check your local air quality on AirForU today before engaging in outdoor activities!
______ Do you know that while exercising is beneficial for your health, exercising outdoors when air pollution is high can harm your health? Protect your health. Check your local air quality on AirForU today before engaging in outdoor activities!
______ Do you know that high air pollution can cause or worsen childhood asthma? Avoiding air pollution can reduce this risk. Check your local air quality on AirForU today before engaging in outdoor activities!
______ Do you know that high air pollution can cause or worsen childhood asthma? Exposure to air pollution can increase this risk. Check your local air quality on AirForU today before engaging in outdoor activities!
______ Do you know that air pollution slows cognition in children by affecting their brain development? Avoiding air pollution can reduce this risk. Check your local air quality on AirForU today before engaging in outdoor activities!
______ Do you know that air pollution slows cognition in children by affecting their brain development? Exposure to air pollution can increase this risk. Check your local air quality on AirForU today before engaging in outdoor activities!
______ Do you know that air pollution is linked to Alzheimer's disease? Avoiding air pollution can reduce this risk. Check your local air quality on AirForU today before engaging in outdoor activities!
______ Do you know that air pollution is linked to Alzheimer's disease? Exposure to air pollution can increase this risk. Check your local air quality on AirForU today before engaging in outdoor activities!
______ Do you know that harmful air pollution is often invisibility to the naked eye? Avoiding air pollution can reduce your risk of harmful health effects. Check your local air quality on AirForU today before engaging in outdoor activities!
______ Do you know that harmful air pollution is often invisibility to the naked eye? Exposure to air pollution can increase your risk of harmful health effects. Check your local air quality on AirForU today before engaging in outdoor activities!

If you were to write your own message to encourage people to check air quality on AirForU what would it be?

(Page 11)

Please provide some background information about yourself.

What is your gender?
- ❍ Male
- ❍ Female

What is your age?
- ❍ 18-24 years
- ❍ 25-34 years
- ❍ 35-44 years
- ❍ 45-54 years
- ❍ 55-64 years
- ❍ 65 years or older

Do you have asthma?
- ❍ Yes
- ❍ No

Do you have any children under 18 living in your household?
- ❍ Yes
- ❍ No

(This question only appears to those who responded Yes to the prior question)

Do any of the children (under 18) living in your house have asthma?
- ❍ Yes
- ❍ No

What is PM2.5?
- ❍ Air quality after 2 pm
- ❍ Particulate matter with a diameter less than 2.5 micrometers
- ❍ Performance measurements standards for air quality equipment
- ❍ Powdered metallics with a diameter less than 2.5 micrometers
- ❍ I don't know

Do you know what the Air Quality Index (AQI) is? If yes, please explain.
- ❍ Yes ____________________
- ❍ No

Approximately, how often do you exercise outdoors?
❍ Once a year or less
❍ Several times a year
❍ A few times a month
❍ 1-2 times a week
❍ 3-4 times a week
❍ 5 or more times a week

What category best describes your annual household income?
❍ Less than $ 24,999
❍ $ 25,000 to $ 49,999
❍ $ 50,000 to $ 74,999
❍ $ 75,000 to $ 99,999
❍ $ 100,000 to $ 149,999
❍ More than $ 150,000

What is the highest level of education you have completed?
❍ Less than high school
❍ High school degree of equivalent (e.g. GED)
❍ Some college but no degree
❍ Associate or technical degree
❍ Bachelor's degree
❍ Graduate degree/professional

What is your race/ethnicity? You may select more than one.
❑ American Indian or Alaska Native
❑ Asian
❑ Black or African American
❑ Native Hawaiian or Other Pacific Islander
❑ Hispanic/Latino
❑ White/Caucasian
❑ Other

(Page 12)

Thank you for participating. Your response has been recorded.

Your survey code is:
[Randomly generated number]

To receive payment for participating, click "Accept HIT" in the Mechanical Turk window, enter this survey code, then click "Submit."

References

AAFA, & NFC. (2005). *Ethnic disparities in the burden and treatment of asthma*. *Washington, DC: Asthma and Allergy Foundation of …*. Retrieved from http://www.aafa.org/media/Ethnic-Disparities-Burden-Treatment-Asthma-Report.pdf

Abrahamse, W., Steg, L., Vlek, C., & Rothengatter, T. (2007). The effect of tailored information, goal setting, and tailored feedback on household energy use, energy-related behaviors, and behavioral antecedents. *Journal of Environmental Psychology*, *27*(4), 265–276. http://doi.org/10.1016/j.jenvp.2007.08.002

ALA. (2016). *State of the Air 2016*. Retrieved from http://www.lung.org/assets/documents/healthy-air/state-of-the-air/sota-2016-full.pdf

Asensio, O. I., & Delmas, M. A. (2015). Nonprice incentives and energy conservation. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(6), E510–E515. http://doi.org/10.1073/pnas.1401880112

Attfield, S., Kazai, G., & Lalmas, M. (2011). Towards a science of user engagement (Position Paper). *WSDM Workshop on User Modelling for Web Applications*. Retrieved from http://www.dcs.gla.ac.uk/~mounia/Papers/engagement.pdf

Banks, S. M., Salovey, P., Greener, S., Rothman, A. J., Moyer, A., Beauvais, J., & Epel, E. (1995). The Effects of Message Framing on Mammography Utilization. *Health Psychology*, *14*(2), 178–184.

Beaumont, R., Hamilton, R. S., Machin, N., Perks, J., & Williams, I. D. (1999). Social awareness of air quality information. *The Science of the Total Environment*, *235*, 319–329.

Bennett, G. G., & Glasgow, R. E. (2009). The Delivery of Public Health Interventions via the Internet: Actualizing Their Potential. *Annual Review of Public Health*, *30*(1), 273–292. http://doi.org/10.1146/annurev.publhealth.031308.100235

Bergquist, S. P., Pal, J. S., Trott, W., Brown, A., Wang, G., & Luce, S. L. (2012). *Climate Change*

Implications for Ballona Wetlands Restoration Climate Change Implications for.

Bickerstaff, K., & Walker, G. (1999). Clearing the smog? Public responses to air-quality information. *Local Environment*, *4*(3), 279–294. http://doi.org/10.1080/13549839908725600

Bickerstaff, K., & Walker, G. (2001). Public understandings of air pollution: The "localisation" of environmental risk. *Global Environmental Change*, *11*(2), 133–145. http://doi.org/10.1016/S0959-3780(00)00063-7

Block, L. G., & Keller, P. A. (1995). When to Accentuate the Negative: The Effects of Perceived Efficacy and Message Framing on Intentions to Perform a Health-Related Behavior. *Journal of Marketing Research*, *32*(2), 192–203.

Block, M. L., & Calderón-Garcidueñas, L. (2009). Air pollution: mechanisms of neuroinflammation and CNS disease. *Trends in Neurosciences, 32*(9), 506–516. http://doi.org/10.1016/j.tins.2009.05.009

Boulos, M. N. K., Wheeler, S., Tavares, C., & Jones, R. (2011). How smartphones are changing the face of mobile and participatory healthcare: an overview, with example from eCAALYX. *Biomedical Engineering Online*, *10*, 24. http://doi.org/10.1186/1475-925X-10-24

Bresnahan, B. W., Dickie, M., & Gerking, S. (1997). Averting Behavior and Urban Air Pollution. *Land Economics, 73*(3), 340–357. Retrieved from http://www.jstor.org/stable/3147172

Brook, R. D., Franklin, B., Cascio, W., Hong, Y., Howard, G., Lipsett, M., … Tager, I. (2004). Air Pollution and Cardiovascular Disease A Statement for Healthcare Professionals From the Expert Panel on Population and Prevention Science of the American Heart Association. *Circulation*, *109,* 2655–671. http://doi.org/10.1161/01.CIR.0000128587.30041.C8

Brookmeyer, R., Gray, S., & Kawas, C. (1998). Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset. *American Journal of Public Health.*, *88*(9), 1337–1342. http://doi.org/10.2105/AJPH.88.9.1337

Brunekreef, B., & Holgate, S. T. (2002). Air pollution and health. *Lancet, 360*(9341), 1233–1242. http://doi.org/10.1016/S0140-6736(02)11274-8

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk. *Perspectives on*

Psychological Science, *6*(1), 3–5. http://doi.org/10.1177/1745691610393980

Bush, J., Moffatt, S., & Dunn, C. E. (2001). Keeping the public informed? Public negotiation of air quality information. *Public Understanding of Science*, *10*(2001), 213–229. http://doi.org/10.1088/0963-6625/10/2/304

Calderón-Garcidueñas. (2002). Safety Evaluation — Environmental Air Pollution and Brain Damage, *30*(3), 373–389.

Calderón-Garcidueñas, L., Solt, A. C., Henríquez-Roldán, C., Torres-Jardón, R., Nuse, B., Herritt, L., … Reed, W. (2008). Long-term air pollution exposure is associated with neuroinflammation, an altered innate immune response, disruption of the blood-brain barrier, ultrafine particulate deposition, and accumulation of amyloid beta-42 and alpha-synuclein in children and youn. *Toxicologic Pathology*, *36*(2), 289–310. http://doi.org/10.1177/0192623307313011

CDC. (2014). Center for Disease Control and Prvention: Air Pollution. Retrieved July 9, 2017, from https://www.cdc.gov/climateandhealth/effects/air_pollution.htm

Centers For Disease Control and Prevention. (2015). National Center for Health Statistics. Centers for Disease Control and Prevention Asthma. Retrieved July 14, 2017, from https://www.cdc.gov/nchs/fastats/asthma.htm

Colls, J. (2002). *Air Pollution Measurement, Modelling and Mitigation* (2nd ed.). Routledge. http://doi.org/10.4324/9780203476024.ch1

Curtis, L., Rea, W., Smith-Willis, P., Fenyves, E., & Pan, Y. (2006). Adverse health effects of outdoor air pollutants. *Environment International*, *32*(6), 815–30. http://doi.org/10.1016/j.envint.2006.03.012

D'amato, G., Liccardi, G., D'amato, M., & Cazzola, M. (2001). The role of outdoor air pollution and climatic changes on the rising trends in respiratory allergy. *Respiratory Medicine*, *95*(7), 606–611. http://doi.org/10.1053/rmed.2001.1112

Davis, J. J. (1995). The effects of message framing on response to environmental communications. *Journalism & Mass Communication Quarterly*, *72*(2), 285–299. http://doi.org/10.1177/107769909507200203

Delmas, M. A., Shimshack, J. P., & Montes-Sancho, M. (2010). Information Disclosure Policies : Evidence from the Electricity Industry UC Santa Barbara. *Economic Inquiry*, *48*(2), 483–498. http://doi.org/10.1111/j.1465-7295.2009.00227.x

Detweiler, J. B., Bedell, B. T., Salovey, P., Pronin, E., & Rothman, A. J. (1999). Message Framing and Sunscreen Use: Gain-Framed Messages Motivate Beach-Goers. *Health Psychology*, *18*(2), 189–196.

Fagan, J. C. (2014). The suitability of web analytics key performance indicators in the academic library environment. *Journal of Academic Librarianship*, *40*(1), 25–34. http://doi.org/10.1016/j.acalib.2013.06.005

Fjeldsoe, B. S., Marshall, A. L., & Miller, Y. D. (2009). Behavior Change Interventions Delivered by Mobile Telephone Short-Message Service. *American Journal of Preventive Medicine*, *36*(2), 165–173. http://doi.org/10.1016/j.amepre.2008.09.040

Freire, C., Ramos, R., Puertas, R., Lopez-Espinosa, M.-J., Julvez, J., Aguilera, I., … Olea, N. (2010). Association of traffic-related air pollution with cognitive development in children. *Journal of Epidemiology and Community Health*, *64*(3), 223–228. http://doi.org/10.1136/jech.2008.084574

Gehring, U., Wijga, A. H., Brauer, M., Fischer, P., De Jongste, J. C., Kerkhof, M., … Brunekreef, B. (2010). Traffic-related air pollution and the development of asthma and allergies during the first 8 years of life. *American Journal of Respiratory and Critical Care Medicine*, *181*(6), 596–603. http://doi.org/10.1164/rccm.200906-0858OC

Giles, L. V., & Koehle, M. S. (2014). The health effects of exercising in air pollution. *Sports Medicine*, *44*(2), 223–249. http://doi.org/10.1007/s40279-013-0108-z

Graff Zivin, J., & Neidell, M. (2009). Days of haze: Environmental information disclosure and intertemporal avoidance behavior. *Journal of Environmental Economics and Management*, *58*(2), 119–128. http://doi.org/10.1016/j.jeem.2009.03.001

Weuve, J., Puett, R.C., Schwartz, J., Yanoksy, J.D., Laden, F., Grodstein, F. (2012). Exposure to particulate air pollution and cognitive decline in older women. *Archives of Internal Medicine*, *172*(3), 219–227. http://doi.org/10.1001/archinternmed.2011.683

Jacob, D. J., & Winner, D. A. (2008). Effect of climate change on air quality. *Atmospheric Environment*, *43*(6), 51–63. http://doi.org/10.1016/j.atmosenv.2008.09.051

Kahneman, D., & Tversky, A. (1979). Prospect Theory : An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–292.

Kailas, A., Chong, C. C., & Watanabe, F. (2010). From mobile phones to personal wellness dashboards. *IEEE Pulse*, *1*(1), 57–63. http://doi.org/10.1109/MPUL.2010.937244

Kampa, M., & Castanas, E. (2008). Human health effects of air pollution. *Environmental Pollution*, *151*(2), 362–367. http://doi.org/10.1016/j.envpol.2007.06.012

Konar, S., & Cohen, M. (1997). Information As Regulation : The Effect of Community Right to Know Laws on Toxic Emissions *. *Journal of Environmental Economics and Management*, *32*, 109–124.

Krewski, D., Burnett, R. T., Goldberg, M. S., Hoover, K., Siemiatycki, J., Abrahamowicz, M., & White, W. H. (2004). Validation of the Harvard Six Cities Study of Particulate Air Pollution and Mortality. *N Engl J Med*, *350*(2), 198–199. http://doi.org/http://dx.doi.org/10.1056/NEJM200401083500225

Krewski, D., Burnett, R. T., Goldberg, M. S., Hoover, K., Siemiatycki, J., Jarret, M., … Institute), (Health Effects. (2000). Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of particulate air pollution and mortality. *Particle Epidemiology Reanalysis Project*.

Krishna, S., Austin Boren, S., & Balas, E. A. (2009). Healthcare via Cell Phones: A Systematic Review. *Telemedicine and E-Health*, *15*(3), 231–240. http://doi.org/10.1089/t mj.2008.0099

Lalmas, M., O'Brien, H., & Yom-Tov, E. (2014). *Measuring User Engagement Synthesis Lectures on Information Concepts, Retrieval, and Services*. (G. Marchionini, Ed.)*Synthesis Lectures on Information Concepts, Retrieval, and Services*. Morgan & Claypool Publishers. http://doi.org/10.2200/S00605ED1V01Y201410ICR038

Laumbach, R. J., & Kipen, H. M. (2012). Respiratory health effects of air pollution: update on biomass smoke and traffic pollution. *The Journal of Allergy and Clinical Immunology*, *129*(1), 3-11–3. http://doi.org/10.1016/j.jaci.2011.11.021

Lehmann, J., Lalmas, M., Yom-Tov, E., & Dupret, G. (2012). Models of user engagement. *20th*

Conference on User Modeling, Adaptation, and Personalization, 164–175. http://doi.org/10.1007/978-3-642-31454-4_14

Levesque, S., Surace, M. J., McDonald, J., & Block, M. L. (2011). Air pollution & the brain: Subchronic diesel exhaust exposure causes neuroinflammation and elevates early markers of neurodegenerative disease. *Journal of Neuroinflammation*, *8*, 105. http://doi.org/10.1186/1742-2094-8-105

Levin, I. P. (1987). Associative effects of information framing. *Bulletin of the Psychonomic Society*, *25*(2), 85–86. http://doi.org/10.3758/BF03330291

Lustria, M. L., Noar, S. M., Cortese, J., Van Stee, S. K., Glueckauf, R. L., & Lee, J. (2013). A meta-analysis of web-delivered tailored health behavior change interventions. *Journal of Health Communication*, *18*(9), 1039–69. http://doi.org/10.1080/10810730.2013.768727

Maheswaran, D., & Meyers-levy, J. (1990). The Influence of Message Framing and Issue Involvement. *Journal of Marketing Research*, *27*(3), 361–367.

Mannino, D. M., Homa, D. M., Akinbami, L. J., Moorman, J. E., Gwynn, C., & Redd, S. C. (2002). *Surveillance for Asthma - United States, 1980-1999*. *Morbidity and Mortality Weekly Report* (Vol. 51). Retrieved from http://origin.glb.cdc.gov/mmWR/preview/mmwrhtml/ss5101a1.htm

Mansfield, C., Reed Johnson, F., & Van Houtven, G. (2006). The missing piece: Valuing averting behavior for children's ozone exposures. *Resource and Energy Economics*, *28*(3), 215–228. http://doi.org/10.1016/j.reseneeco.2006.02.002

McConnell, R., Islam, T., Shankardass, K., Jerrett, M., Lurmann, F., Gilliland, F., … Berhane, K. (2010). Childhood incident asthma and traffic-related air pollution at home and school. *Environmental Health Perspectives*, *118*(7), 1021–1026. http://doi.org/10.1289/ehp.0901232

Meyerowitz, B. E., & Chaiken, S. (1987). The effect of message framing on breast self-examination attitudes, intentions, and behavior. *Journal of Personality and Social Psychology*, *52*(3), 500–510. http://doi.org/10.1037/0022-3514.52.3.500

Mickley, L. J., Jacob, D. J., Field, B. D., & Rind, D. (2004). Effects of future climate change on regional air pollution episodes in the United States. *Geophysical Research Letters*, *31*(24), 1–4.

http://doi.org/10.1029/2004GL021216

Moulton, P. V., & Yang, W. (2012). Air pollution, oxidative stress, and alzheimer's disease. *Journal of Environmental and Public Health*, *2012*(x). http://doi.org/10.1155/2012/472751

Neidell, M. (2006). Public Information and Avoidance Behavior: Do People Respond to Smog Alerts? *Center of Statistical and Environmental Science Technical Report*, (24).

Neidell, M. J. (2004). Air pollution, health, and socio-economic status: the effect of outdoor air quality on childhood asthma. *Journal of Health Economics*, *23*(6), 1209–36. http://doi.org/10.1016/j.jhealeco.2004.05.002

Noar, S. M., Benac, C. N., & Harris, M. S. (2007). Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychological Bulletin*, *133*(4), 673–693. http://doi.org/10.1037/0033-2909.133.4.673

Ozdalga, E., Ozdalga, A., & Ahuja, N. (2012). The Smartphone in Medicine: A Review of Current and Potential Use Among Physicians and Students. *Journal of Medical Internet Research*, *14*(5). http://doi.org/10.2196/jmir.1994

Pew Research Center. (2017). *Mobile Fact Sheet*. Retrieved from http://www.pewinternet.org/fact-sheet/mobile/

Pope, C. A., & Dockery, D. W. (2006). Health Effects of Fine Particulate Air Pollution: Lines that Connect. *Journal of the Air & Waste Management Association*, *56*(6), 709–742. http://doi.org/10.1080/10473289.2006.10464485

Rodden, K., Hutchinson, H., & Fu, X. (2010). Measuring the User Experience on a Large Scale : User-Centered Metrics for Web Applications. *SIGCHI Conference on Human Factors in Computing Systems*, 2395–2398. http://doi.org/10.1145/1753326.1753687

Ruggieri, M., & Plaia, A. (2012). An aggregate AQI: Comparing different standardizations and introducing a variability index. *The Science of the Total Environment*, *420*, 263–72. http://doi.org/10.1016/j.scitotenv.2011.09.019

Semenza, J. C., Wilson, D. J., Parra, J., Bontempo, B. D., Hart, M., Sailor, D. J., & George, L. a. (2008).

Public perception and behavior change in relationship to hot weather and air pollution. *Environmental Research*, *107*(3), 401–11. http://doi.org/10.1016/j.envres.2008.03.005

Sly, R. (1988). Mortality from asthma, 1979-1984. *The Journal of Allergy and Clinical Immunology*, *82*, 705–717.

Statista. (2017). Global market share held by the leading smartphone operating systems from 1st quarter 2011 to 2nd quarter 2016. *Statista*. Retrieved from http://www.statista.com/statistics/236035/market-share-of-global-smartphone-os-shipments-by-mobile-operating-system-per-quarter/

Suglia, S. F., Gryparis, A., Wright, R. O., Schwartz, J., & Wright, R. J. (2008). Association of black carbon with cognition among children in a prospective birth cohort study. *American Journal of Epidemiology*, *167*(3), 280–286. http://doi.org/10.1093/aje/kwm308

Terry, M. (2010). Medical Apps for Smartphones. *Telemedicine Journal and E-Health*, *16*(1), 17–22. http://doi.org/Doi 10.1089/Tmj.2010.9999

Treiber, F. A. (1986). A comparison of the positive and negative consequences approaches upon car restraint usage. *Journal of Pediatric Psychology*, *11*(1), 15–24. http://doi.org/10.1093/jpepsy/11.1.15

Tversky, A., & Kahneman, D. (1985). The Framing of Decisions and the Pscyhology of Choice. In V. T. Covello, J. L. Mumpower, P. J. M. Stallen, & V. R. R. Uppuluri (Eds.), *Evnvironmental Impact Assessment, Technology Assessment, and Risk Analysis*. http://doi.org/10.1007/978-3-642-70634-9

UN Environment Programme. (2016). *Annual Report 2016: Engaging People To Protect The Planet*.

UN WHO. (2014). 7 million premature deaths annually linked to air pollution. Retrieved from http://www.who.int/mediacentre/news/releases/2014/air-pollution/en/

US Census Bureau. (2015a). Age and Sex 2015 American Community Survey 1-Year Estimates. Retrieved July 9, 2017, from https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_15_1YR_S0101&prodType=table

US Census Bureau. (2015b). Selected Characteristics of Households by Total Money Income. Retrieved July 9, 2017, from https://www.census.gov/data/tables/time-series/demo/income-poverty/cps-hinc/hinc-01.html

US Dept. of Health and Human Services. (2012). *National surveillance of asthma: United States, 2001-2010. Vital & health statistics. Series 3, Analytical and epidemiological studies / [U.S. Dept. of Health and Human Services, Public Health Service, National Center for Health Statistics]*. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24252609

US EPA. (2006). *Guidelines for the Reporting of Daily Air Quality – the Air Quality Index (AQI)*.

US EPA. (2014). *A Guide to Air Quality and Your Health*.

US EPA. (2017a). Current Nonattainment Counties for All Criteria Pollutants. Retrieved September 25, 2017, from https://www3.epa.gov/airquality/greenbook/ancl.html

US EPA. (2017b). EnviroAtlas Benefit Category: Biodiversity Conservation. Retrieved September 25, 2017, from https://www.epa.gov/enviroatlas/enviroatlas-benefit-category-biodiversity-conservation

US EPA. (2017c). The Impaired Waters and TMDLs Program in your EPA Region, State or Tribal Land. Retrieved September 25, 2017, from https://www.epa.gov/tmdl/impaired-waters-and-tmdls-program-your-epa-region-state-or-tribal-land

US EPA. (2017d). Toxics Release Inventory (TRI) Program. Retrieved July 9, 2017, from https://www.epa.gov/toxics-release-inventory-tri-program

Wakefield, S. E. L., Elliott, S. J., Cole, D. C., & Eyles, J. D. (2001). Environmental risk and (re)action: Air quality, health, and civic involvement in an urban industrial neighbourhood. *Health and Place*, *7*(3), 163–177. http://doi.org/10.1016/S1353-8292(01)00006-5

Weiss, K. B., & Wagener, D. K. (1990). Changing Patterns of Asthma Mortality. Identifying Target Populations at High Risk. *The Journal of the American Medical Association*, *264*, 1683–1687.

Wilson, D. K., Wallston, K. A., & King, J. E. (1990). Effects of contract framing, motivation to quit, and self-efficacy on smoking reduction. *Journal of Applied Social Psychology*, *20*(7, Pt 2), 531–547. http://doi.org/10.1111/j.1559-1816.1990.tb00426.x

Witte, K., & Allen, M. (2000). A meta-analysis of fear appeals: implications for effective public health campaigns. *Health Education & Behavior : The Official Publication of the Society for Public Health Education*, *27*(5), 591–615. http://doi.org/10.1177/109019810002700506

Yearley, S. (2006). Bridging the science-policy divide in urban air-quality management: Evaluating ways to make models more robust through public engagement. *Environment and Planning C: Government and Policy*, *24*(5), 701–714. http://doi.org/10.1068/c0610j

III. Constructing Meaningful Environmental Indices: A Nonparametric Frontier Approach

1 Introduction

Environmental information disclosure has been touted as a powerful tool that can augment traditional command and control regulation and influence positive environmental performance through public pressure (Konar and Cohen, 1997). A well-known example is the Toxic Release Inventory (TRI), an initiative from the U.S. Environmental Protection Agency (EPA), which imposes mandatory disclosure requirements on large industrial facilities to release information on toxic emissions (Koehler and Spengler, 2007). The success of the TRI hinges on public pressure imposed on plants with poor environmental performance, which provides an incentive for plants to adopt stronger environmental measures and improve their environmental performance (Khanna and Anton, 2002) or, alternatively, positive public recognition for the best performers. Many users of TRI data tend to evaluate environmental performance based on a single metric such as total releases of toxic chemical emissions, while ignoring other potentially relevant dimensions of economic performance such as revenue generated, or employment data, which are crucial aspects of businesses. We argue that a composite environmental index (CEI) must consider environmental performance in conjunction with other measures of corporate performance to identify the "best" plants and practices, those that achieve both environmental and economic success.

Besides the variables used, the CEI must exhibit other properties such as "meaningfulness" and standardization. The terminology of a “meaningful index” originated from the influential study by Ebert and Welsch (2004) that characterized classes of environmental indices. As a fundamental scientific rule, ‘meaningfulness’ implies that the comparison of environmental performance across time or space based on CEIs must be free of ambiguity (Welsch, 2005). However, when variables with different scale properties, for example, tons of air pollutants (ratio-scale) and temperature measured on the Celsius scale (interval-scale), are combined, it is difficult to aggregate them into a CEI in a meaningful way (Böhringer and Jochem, 2007). In addition, since Tyteca (1996, 1997) scholars have been advocating for the development of a standardized aggregate index between zero and one in order to allow for a proper comparison of environmental performance between firms.[1] Hence it is important to construct a meaningful and standardized CEI that is capable of handling issues of mixed measurability of underlying variables (i.e. both ratio-scale and interval-scale variables are involved) and data irregularity (e.g. the existence of multiple zero entries).

In the existing literature, methods for measuring environmental performance for firms may be broadly classified into two groups. The first group aims to measure environmental performance from the perspective of productive efficiency, which involves classifying underlying variables into inputs and outputs and specifying an environmental production technology for modeling the joint production of good and bad outputs. Within this first group, there are two strategies for the measurement of environmental performance of firms or plants. One strategy is to calculate an adjusted measure of efficiency or productivity whereby a firm or plant is credited for

[1] The study by Tyteca (1996) provided an excellent review of the existing methods for measuring environmental performance of firms, which ranges from simple indicators reflecting only one aspect of the impact of activities to more sophisticated ones reflecting the overall impact on the environment.

simultaneously increasing good output production and reducing bad output production. Pittman (1983) conducted one of the earliest studies initiating this strategy by incorporating pollutants into productivity measurement. Subsequently, the seminal study by Färe et al. (1989) laid an elegant theoretical foundation for using nonparametric frontier methodology to evaluate productive efficiency with undesirable outputs. The framework developed by Färe et al. (1989) has been adopted by a large number of studies, which have focused on not only firms and plants (e.g. Boyd and McClelland, 1999; Färe et al., 1997; Färe et al., 2010; Khanna and Kumar, 2011) but also countries and regions (e.g. Zhou et al., 2010; Hoang and Coelli, 2011; Picazo-Tadeo et al., 2014). The other strategy involves constructing a formal environmental performance index (EPI) by using Shephard or Malmquist distance functions (Färe et al., 2004, 2006, 2010). Its advantage lies in the fact that the resulting EPI holds some desirable index number properties. Both of these strategies can be implemented by utilizing data envelopment analysis (DEA) models.

The second group attempts to aggregate multiple environmental variables into a CEI for performance evaluation and comparison. It allows the use of diverse variables in accordance with the environmental theme being studied. Ebert and Welsch (2004) showed that a geometric mean can lead to a meaningful index when the underlying variables are ratio-scale and strictly positive. Zhou et al. (2006) developed an information loss criterion to assess alternative aggregation rules for constructing CEIs. Munda and Nardo (2009) highlighted the usefulness of non-compensatory aggregation approach. While many previous studies focused on data aggregation, several scholars have examined other important issues such as weighting (e.g. Decancq and Lugo, 2013) and normalization (e.g. Zhou and Ang, 2009; Pollesch and Dale, 2016).

This paper contributes to the existing body of studies on CEIs in the following aspects. First, as an extension to the important work by Ebert and Welsch (2004), we classify 'meaningfulness' into ordinal and cardinal meaningfulness and argue the importance of constructing a cardinally meaningful CEI. Second, in the spirit of the influential studies by Färe et al. (1996, 2004, 2006, 2010) and Tyteca (1996, 1997), we advocate the use of a nonparametric frontier approach for constructing a standardized CEI that simultaneously satisfies cardinal meaningfulness. This approach can also address data irregularity issues that appear in TRI data such as a large number of zeros, which pose challenges in the application of the theoretically meaningful aggregation rules such as geometric mean. Third, we apply this methodology to TRI data for evaluating facility-level environmental performance in Los Angeles County, which provides perspectives on how facilities may improve their environmental performance.

The remainder of this paper is organized as follows. In Section 2, we describe the concept of a meaningful environmental index that is composed of two categories, namely "ordinal meaningfulness" and "cardinal meaningfulness." In Section 3, we present the nonparametric frontier approach and show its desirable theoretical properties as compared to arithmetic aggregation. In Section 4, we describe our empirical study using the nonparametric frontier methodology to evaluate the environmental performance of different facilities from three industrial sectors in Los Angeles County. Concluding remarks are presented in Section 5.

2 Meaningful Composite Environmental Indices

The concept of "meaningfulness" given by Ebert and Welsch (2004) is built upon the invariance of preference orderings with respect to the measurement units of underlying variables. In addition to orderings, the relative performance gaps of CEI values may also carry valuable information for performance comparison and improvement. As such, we classify 'meaningfulness' into 'ordinal meaningfulness' and 'cardinal meaningfulness' and use a simple example to illustrate the importance of cardinal meaningfulness in this section.

2.1 Definitions

Let $V_k = (v_{k1}, \cdots, v_{kn})$ denote a vector of n underlying environmental variables for entity k ($k = 1, \cdots, K$). Our task is to construct a CEI for each entity based on the n variables. One usage of constructing the CEI is to provide a ranking of different entities in environmental performance, which can be characterized by the preference ordering $\succeq$ defined on $\Re^n$. Thus a CEI can be represented by a mapping function $I : \Re^n \to R$ that satisfies

$$V_k \succeq V_l \Leftrightarrow I(V_k) \geq I(V_l) \quad \forall k, l \in \{1, \cdots, K\} \tag{1}$$

Note that the measurement units of the underlying n variables may be changed, which can be represented by a transformation function $F = (f_1, \cdots, f_n)$ such that

$$F : (v_{k1}, \cdots, v_{kn}) \to (f_1(v_{k1}), \cdots, f_n(v_{kn})) \tag{2}$$

As described in Ebert and Welsch (2004) and Welsch (2005), an admissible transformation involves expansion as well as translation, i.e. $f_i(v_{ki}) = \alpha_i v_{ki} + \beta_i, \alpha_i > 0$. With reference to CEIs, the orderings of different entities are expected to be invariant with respect to any admissible transformation of underlying variables (Ebert and Welsch, 2004; Welsch, 2005), i.e.

$$V_k \succeq V_l \Leftrightarrow F(V_k) \succeq F(V_l) \quad \forall k,l \in \{1,\cdots,K\} \tag{3}$$

Definition 1 (*Ordinal meaningfulness*). *I* is an ordinally meaningful index if it satisfies

$$I(V_k) \geq I(V_l) \Leftrightarrow I(F(V_k)) \geq I(F(V_l)) \quad \forall k,l \in \{1,\cdots,K\} \tag{4}$$

It should be pointed out that Ebert and Welsch (2004) and Welsch (2005) termed *I* satisfying Eq. (4) as a meaningful index while we refer to it as an ordinally meaningful index in this paper. Ebert and Welsch (2004) showed that geometric aggregation would yield an ordinally meaningful CEI when the underlying variables are ratio-scale noncomparable. Despite the importance of ordinal meaningfulness, as discussed by Böhringer and Jochem (2007), many popular CEIs for sustainability did not take it into account and were therefore misleading with respect to policy practice.

Acknowledging the importance of ordinal meaningfulness, we argue that it is valuable for a CEI to preserve a relative performance gaps between entities. This may be illustrated by the case of a city-level air pollutant index derived from several air pollutants. If the index values are respectively 150, 140 and 50, it says that the last city shows the best while the first city shows the worst air pollution level. When the index values become 80, 60 and 50, the same message is transmitted regarding their orderings in air pollution level. Beyond it, we observe that the performance gaps between the first two cities and the last one become smaller. It implies that the index values also carry valuable information through their relative performance gaps between entities. Thus we have

Definition 2 (*Cardinal meaningfulness*). *I* is a cardinally meaningful index if it satisfies

$$I(V_k) = \alpha I(F(V_k)) \quad \forall k \in \{1, \cdots, K\} \tag{5}$$

where α is a positive constant.

Eq. (5) says that a cardinally meaningful CEI preserves the relative performance gaps between entities for any admissible transformation of underlying variables. $\alpha = 1$ implies that the CEI values will not change, which is not necessary for satisfying cardinal meaningfulness but still desirable as the resulting CEI looks more standardized. Obviously, cardinal meaningfulness represents a stronger requirement than ordinal meaningfulness. That is to say,

Proposition 1. *A cardinally meaningful index must be an ordinally meaningful index; not vice versa*.

Once a cardinally meaningful CEI is theoretically defined, the next task is to identify a way to compute its values. While adequate weighting, normalization and aggregation of underlying variables are often regarded as pre-requisites for the practice, Böhringer and Jochem (2007) pointed out that there are no unambiguous rules for data weighting and normalization as they often imply value judgements.[2] Regarding data aggregation, Welsch (2005) and Böhringer and Jochem (2007) described several meaningful aggregation methods dependent on the scale and comparability characteristics of underlying variables. An important finding of their studies is that the arithmetic mean is meaningful for variables satisfying interval scale and full comparability.

[2] Normalization is the process of transforming the different measurement units of underlying variables into a common unit or dimensionless. The recent study by Pollesch and Dale (2016) provides a comprehensive discussion on alternative normalization methods in the context of sustainability assessment.

While normalization can help achieve comparability and may internally be linked to meaningful aggregation, in this paper we only focus on the aggregation of underlying variables without explicitly discussing the normalization scheme.

2.2 An illustrative example

We use a simple example to illustrate the issue of data irregularity occurring in the TRI database and explain why arithmetic and geometric means aggregation rules are inappropriate for the application. Table 25 shows the data on two environmental variables for four selected facilities in the Chemicals industry in the Los Angeles County (Delmas and Kohli, 2014). Clearly, one data irregularity is that there exist multiple zero entries.

Table 25: Data for the illustrative example

Facility	**Total toxic releases (Pounds)**	**Toxicity of on-site releases (Pounds-toxicity)**
A	0	0
B	2000	6000
C	200	22000
D	180	0

As shown in Table 25, facility A obviously has the best environmental performance since it represents the best practice for the total toxic releases and the toxicity of the releases. If an arithmetic mean is applied to evaluate the performance for facilities B and C, their CEI values are respectively (2000+6000)/2=4000 and (200+22000)/2=11000 respectively. However, if the measurement unit of the second variable is changed to "thousand pounds-toxicity", the CEI values of B and C computed by arithmetic mean will become (2000+6)/2=1003 and (200+22)/2=111. The preference orderings of the two facilities are reversed, which verifies the conclusion drawn by Ebert and Welsch (2004) that an arithmetic mean cannot yield an ordinally meaningful CEI without normalization. Note that the two variables are ratio-scale noncomparable so that a geometric mean would yield a meaningful CEI for facilities B and C as shown by Ebert and Welsch (2004). Based on geometric means, their CEI values are respectively $(2000\times6000)^{0.5}=3464$ and $(200\times22000)^{0.5}=2098$. Nevertheless, if the entire dataset is considered, facilities A and D have at least one variable equal to zero. This violates the condition given by Ebert and Welsch (2004) that the observations of underlying variables are strictly positive. If the geometric mean is applied, the CEI values for facilities A and D are equal to zero indicating that the aggregation rule is not zero robust. While only ratio-scale variables are considered in this example, both ratio-scale and interval-scale variables might be involved in the application (i.e the mixed measurability of underlying variables). In this circumstance, even if all the observations are strictly positive, the geometric mean aggregation rule will not yield a meaningful index (Ebert and Welsch, 2004; Welsch, 2005).

3 Methods

In this section, we introduce a nonparametric frontier approach, DEA, for constructing a cardinally meaningful and standardized CEI. This approach can easily address the issues of data irregularity that frequently appear in TRI data as well as the mixed measurability of the underlying variables.

3.1 DEA model

As a nonparametric frontier methodology, DEA employs linear programming to identify the best practice frontier and evaluate the relative performance of each entity based on the observations of inputs and outputs for a group of comparable entities (Coelli et al., 2005). Since the seminal study by Färe et al. (1989) and the influential work by Tyteca (1996), DEA has been widely applied to the measurement of environmental performance or pollutant-adjusted efficiency/productivity of different entities. Examples of such studies include Färe et al. (1996, 2004, 2006, 2007, 2010), Tyteca (1997), Boyd and McClelland (1999), Zhou et al. (2010), Hoang and Nguyen (2013) and Picazo-Tadeo et al. (2014).

The conventional use of DEA for environmental performance measurement starts from a differentiation between good and bad outputs as well as the specification of a production technology for modeling their joint production. In this line of research, Färe et al. (1989) has laid an elegant theoretical foundation, which makes the nonparametric frontier methodology popular for performance measurement with bad outputs such as pollutants. In constructing CEIs, however, there may not exist a productive relationship between underlying variables (e.g. the case of the air pollutant index). Nevertheless, the variables may also be divided into "inputs" and "outputs" which respectively satisfy the properties of "the smaller the better" and "the larger the better" from the perspective of performance improvement. To differentiate between inputs and outputs, we replace the vector $V_k = (v_{k1}, \cdots, v_{kn})$ given in Section 2 by $V_k = (X_k, Y_k) = (x_{k1}, \cdots, x_{km}, y_{k1}, \cdots, y_{ks})$ where X_k and Y_k are respectively the input and output vectors. Using all the observations, we can construct a quasi- reference technology as follows:[3]

[3] The term "quasi-reference technology" implies that it looks like a reference technology externally but the productive relationship between inputs and outputs may not exist. Since the choice of inputs and outputs for constructing a CEI is dependent on the environmental theme concerned, the commonly used inputs such as capital, labor and energy may not be included in the construction of CEIs. Actually, Färe et al. (2006, 2010) also excluded such inputs in developing an environmental performance

$$
\begin{aligned}
\mathrm{S}=\{(X,Y):&\sum_{k=1}^{K}x_{ki}z_k\le x_i,\quad i=1,\cdots,m\\
&\sum_{k=1}^{K}y_{kr}z_k\ge y_r,\quad r=1,\cdots,s\\
&\sum_{k=1}^{K}z_k=1\\
&z_k\ge 0,\ k=1,\cdots,K\ \}
\end{aligned}
\tag{6}
$$

With Eq. (6) as the constraint, we can formulate the following range adjusted DEA model (Aida et al., 1998; Cooper et al., 1999):

$$
\begin{aligned}
\max\ &\frac{1}{m+s}\left(\sum_{i=1}^{m}\frac{s_i^-}{R_i^-}+\sum_{r=1}^{s}\frac{s_r^+}{R_r^+}\right)\\
\text{s.t.}\ &\sum_{k=1}^{K}x_{ki}z_k+s_i^-=x_{oi},\quad i=1,\cdots,m\\
&\sum_{k=1}^{K}y_{kr}z_k-s_r^+=y_{or},\quad r=1,\cdots,s\\
&\sum_{k=1}^{K}z_k=1\\
&z_k\ge 0,s_i^-\ge 0,s_r^+\ge 0
\end{aligned}
\tag{7}
$$

where x_{oi} and y_{or} respectively denote the *i*-th input and *r*-th output for entity o ($o\in\{1,\cdots,K\}$);

R_i^- and R_r^+ denote the ranges for input *i* and output *r*, which are respectively defined as

$R_i^-=\max\{x_{ki},k=1,\cdots,K\}-\min\{x_{ki},k=1,\cdots,K\}$ and

$R_r^+=\max\{y_{kr},k=1,\cdots,K\}-\min\{y_{kr},k=1,\cdots,K\}$.

Eq. (7) belongs to the family of additive DEA models.[4] Its objective function, often referred to as

index that has an advantage of crediting a producer for adopting processes generating more good output per unit of bad output produced. A difference is that Färe et al. (2006, 2010) classified outputs into good and bad outputs while we treat bad outputs as inputs since they both follow the property of "the smaller the better" (Hailu and Veeman, 2001).

[4] As a slacks-based DEA model, Eq. (7) has a close relationship with the non-radial directional distance function (DDF) that has

range adjusted inefficiency measure, represents the average of slacks-based inefficiency measure for entity o. The constraints determine the best practice frontier from which the maximally potential reduction and expansion in inputs and outputs are identified. When the range for a variable is zero, it indicates that all the entities have the same value for the variable so that the variable may be excluded in environmental performance evaluation. In that circumstance, the relevant component in the objective function of Eq. (7) and the corresponding constraint need to be removed. The last constraint $\sum_{k=1}^{K} z_k = 1$ as a convexity condition guarantees that the measurement units of ratio-scale variables will not change the optimal solution (Cooper et al., 1999). For any admissible transformation of the original variables, i.e. $f_i(v_{ki}) = \alpha_i v_{ki} + \beta_i$, the focus on slacks (or gaps) accommodates the shift parameter (α_i) and the scaling factor (β_i) is handled by means of range adjustment.

Eq. (7) as a simple linear programming model can be easily solved by any linear programming software package. Once the optimal solution to Eq. (7) is derived, we can define a CEI as

$$CEI(V_o) = CEI(X_o, Y_o) = 1 - \frac{1}{m+s}\left(\sum_{i=1}^{m} \frac{s_i^{*-}}{R_i^-} + \sum_{r=1}^{s} \frac{s_r^{*+}}{R_r^+}\right) \quad \forall o \in \{1, \cdots, K\} \qquad (8)$$

where * denotes the corresponding optimal slack variable. It can be shown that a C*EI* derived from Eq. (8) satisfies the following properties (Cooper et al., 1999):

P1. $0 \le CEI(V_o) \le 1$.

P2. $CEI(V_o) = 1 \Leftrightarrow$ Entity *o* is located on the best practice frontier; $CEI(V_o) < 1 \Leftrightarrow$ Entity *o*

gained much popularity in efficiency and productivity analysis (Chambers et al., 1996; Zhou et al., 2012). In environmental economics, DDF has been widely used to assess environmental performance and the impact of environmental regulation. See, for example, Boyd and McClelland (1999), Hoang and Coelli (2011) and Picazo-Tadeo et al. (2014).

is not located on the best practice frontier and can be improved in certain dimensions.

P3. $CEI(V_o)$ is invariant to the measurement units of inputs and outputs.

P4. $CEI(V_o)$ is strongly monotonic.

P5. $CEI(V_o)$ is translation invariant.

P1 indicates that Eq. (8) yields a standardized index lying between zero and one, and a larger index value is linked to better environmental performance. P2 implies that the entities that do not play a role in constructing the best practice frontier have index values less than unity. From Eq. (7), we can easily identify the entities forming the best practice frontier as those associated with nonzero z_k. P3 indicates that the index is invariant with the measurement units of ratio-scale variables. The implication of P4 is that a reduction in any input or an increase in any output leads to an increase in the index value. P5 means that additions and subtractions of constants by any variables will not affect the index value, which is particularly useful when some interval-scale indicators are involved in constructing CEIs.[5] Combining P3 to P5, we have

Proposition 2. *The CEIs derived from Eqs. (7) and (8) are cardinally meaningful, i.e.* $CEI(V_k) = CEI(F(V_k)) \quad \forall k \in \{1, \cdots, K\}$.

Proposition 3. If $X_k \le X_l$, $Y_k \ge Y_l$ and there is at least one i (s) such that $x_{ki} < x_{li}$ ($y_{ks} > y_{ls}$), then $CEI(V_k) > CEI(V_l)$.

[5] It should be pointed out that the range adjusted DEA model, i.e. Eq. (7), is not the only choice for generating a CEI satisfying P1 to P5. For example, the bounded adjusted DEA model proposed by Cooper et al. (2011) may also be used to construct a cardinally meaningful CEI, which might be worth further investigating in future research.

Propositions 2 implies that the CEIs derived from Eqs. (7) and (8) can easily handle data irregularity issues such as multiple zero entries and mixed measurability of the underlying variables. The property of cardinal meaningfulness also facilitates the computation of CEIs when other data irregularity issues exist. In the case of the TRI dataset, the range of values for certain variables can be rather large, which may pose challenges in solving linear programming models due to computer rounding errors. Owing to the properties of unit and translation invariance of Eq. (7), we may rescale the variables to remove the zero values and force the variables to be comparable. Despite the advantages, a concern (and possible weakness) of the use of DEA to construct CEIs is that multiple entities may have index values of unity preventing them from being compared with each other. This, however, also indicates that each of these entities has its particular strengths in certain dimensions, allowing them to serve as benchmarks for similar entities.

3.2 Linkage with arithmetic mean aggregation

The derivation of CEIs by Eqs. (7) and (8) requires solving a series of linear programming models. However, when there is a "super-entity" dominating all the other entities in all dimensions, we may directly derive the CEIs without solving linear programming models. In this circumstance, other entities will automatically identify the "super-entity" as their benchmark, and the optimal slack in a variable for other entities will be equal to their distances from the "super-entity". Mathematically, the CEI can be derived by

$$CEI(V_k) = 1 - \frac{1}{m+s}\left(\sum_{i=1}^{m} \frac{x_{ki} - \min_k\{x_{ki}\}}{R_i^-} + \sum_{r=1}^{s} \frac{\max_k\{y_{kr}\} - y_{kr}}{R_r^+} \right)$$

$$= \sum_{i=1}^{m} \frac{1}{m+s} \left(\frac{\max_k\{x_{ki}\} - x_{ki}}{\max_k\{x_{ki}\} - \min_k\{x_{ki}\}} \right) + \sum_{r=1}^{s} \frac{1}{m+s} \left(\frac{y_{kr} - \min_k\{y_{ki}\}}{\max_k\{y_{kr}\} - \min_k\{x_{kr}\}} \right) \qquad (9)$$

Eq. (9) is a weighted sum of the normalized variables for all the entities for which a linear min-max normalization scheme is adopted. It suggests that the weighted sum method could lead to a meaningful index when there exists a “super-entity”. However, in reality a “super-entity” is unlikely to exist when multi-dimensional environmental performance is concerned. One may imagine that a “super-entity” could be artificially generated by taking the highest values for the outputs and the lowest values for the inputs. Indeed, this practice makes the use of Eq. (9) feasible, which simplifies the computation of CEIs. However, as Munda and Nardo (2009) discussed, the weighted sum aggregation rule assumes full compensability between different variables, implying that the variables are completely substitutable with each other. Since different dimensions cannot be fully substituted with each other, the assumption might not be appropriate for scientifically assessing environmental performance (Munda and Nardo, 2009). In the range adjusted DEA model, the substitution between the optimal slacks for different variables is allowed when there exist multiple optimal solutions. However, this kind of substitution indicates that an entity may have multiple choices to reach the best practice frontier. Different optimal solutions only imply different pathways while the ultimate goal is common – improving environmental performance!

4 Empirical study

4.1 Background

The public availability of the TRI database allows stakeholders as well as researchers to make comparisons of environmental performance across and between firms/plants over time for different purposes (Khanna et al., 1998). Prior studies using the TRI database have specified a variety of environmental indicators, but there is no consensus on which indicator represents an ideal proxy for the measurement of environmental performance (Toffel and Marshall, 2004). The environmental indicators used include aggregate toxic releases (Bui and Kapon, 2012), toxic releases weighted by toxicity factors (Cole et al., 2013), on-site toxic releases and off-site transfers (Khanna and Damon, 1999), the ratio of toxic releases to net sales (Konar and Cohen, 1997), and the toxic releases adjusted by distance (Hanna, 2007). Often, various indicators have been separately used, while it has been argued that the measurement of environmental performance based on TRI data needs to consider not only toxic releases but also other indicators such as revenue and toxicity factors (Gerde and Logsdon, 2001).

Scholars have shown that providing facility-level specific information allows greater transparency and can influence pollution abatement positively (Konar and Cohen, 1997). If environmental performance is evaluated only at the firm level, rather than the facility level, some facilities of a parent company performing above or below average might not be recognized depending on the environmental regulations of the state in which they are located. Despite the availability of this facility specific information, only a few studies, such as Färe et al. (2010) and Bui and Kapon (2012), assessed facility level environmental performance. Specifically, the interesting study by Färe et al. (2010) used Malmquist quantity index and DEA to develop a formal environmental performance index for assessing the performance of coal-fired power plants in releasing toxic chemicals. In this section, we shall employ the nonparametric methodology described in Section 3 to construct a meaningful and standardized CEI for assessing the facility-level environmental performance in toxic releases in Los Angeles County. The empirical analysis not only demonstrates the robustness of the CEI but also shows how the CEI can provide perspectives on the improvement of facility-level environmental performance.

4.2 Data description

Our analysis is based on 150 facilities from three major industries - Primary Metals, Fabricated Metals and Chemicals, which respectively have 29, 54 and 67 facilities and as a whole accounted for 59% of the total toxic releases reported to the TRI database in Los Angeles County for 2012 (Delmas and Kohli, 2014). As pointed out by Delmas and Blass (2010), it is inappropriate to compare the environmental performance of firms or plants from different industries due to their different operating characteristics. As such, the facilities are evaluated and compared with those in the same industry. Since the sample size varies across different industries, our analysis may also shed some insights on how the construction of the CEI is affected by sample size.

Our CEI is derived from four variables, which represent a facility's effort in generating revenue while simultaneously preventing toxics releases into the natural environment. The first variable is the *Quantity of Total Toxic Releases* (QTTR), which includes both on-site and off-site releases to the environment but excludes the toxic releases arising from catastrophic and extreme events. The second is the *Toxicity of Total On-site Toxic Releases* into the atmosphere (TTTR), which is the sum of chemical-specific toxic releases weighted by their corresponding toxicity factors. TTTR accounts for the varying toxicity of chemicals releases and is valuable in measuring the local health-related impacts of different facilities, which cannot be captured by QTTR alone. The third variable is referred to as the *Percentage of Waste Managed through Recycling, Energy*

Recovery and Treatment (PWM), which is the ratio of waste managed through recycling, energy recovery and treatment to the total waste including released and managed waste. The forth variable is the *Gross Revenue* (GR), which is a financial indicator that highlights each facility's ability in generating revenue given a certain amount of toxic releases. Of the four variables, QTTR and TTTR are used as inputs while PWM and GR are used as outputs, following the scheme that QTTR and TTTR are "the smaller the better" and PWM and GR are "the larger the better." Table 26 lists the summary statistics of the four variables by industry. A detailed description of these and additional TRI variables as well as data sources can be found in Delmas and Kohli (2014).

Table 26: Summary statistics of four variables by industry

Industry		**Quantity of Total Releases**	**Toxicity of Total On-site Toxic Releases**	**Percentage of Waste**	**Gross Revenue**
Chemicals	Mean	5500	427707	61	79.64
(67 facilities)	Std.	10697	1706483	44	213.21
	Min	0	0	0	0.09
	Max	55170	13616910	100	1310.00
Primary	Mean	109161	55936001	58	44.65
(29 facilities)	Std.	533393	292429498	48	76.79
	Min	0	0	0	1.14
	Max	2823311	1548000000	100	319.44
Fabricated	Mean	21847	293951	79	44.65
(54 facilities)	Std.	85782	1117620	36	62.78
	Min	0	0	0	0.15
	Max	496159	6475100	100	377.91

As shown in Table 26, the minimum values of QTTR, TTTR and PWM are zero for all the three industries and the ranges for the first two variables are extremely large. This may give rise to certain computational problems due to computer rounding errors if DEA models are directly solved. Fortunately, the range adjusted DEA models given by Eq. (7) are not affected by any linear transformations of the variables, which facilitate the calculation of our CEIs. For comparison purposes, we also compute the CEI values by using the weighted sum aggregation rule cum min-max linear normalization, i.e. Eq. (9). While a geometric aggregation can lead to a meaningful CEI given the fact that four ratio-scale variables are aggregated, we do not use the aggregation rule due to the existence of multiple zeros in the dataset.

4.3 Main results and discussions

Table 27 shows the summary statistics of the CEI values calculated from the nonparametric frontier approach, i.e. Eqs. (7)-(8), and the weighted sum aggregation rule cum min-max linear normalization, i.e. Eq. (9). In terms of variance, the two sets of CEIs are quite close to each other for the sectors of Primary Metals and Fabricated Metals, while for the Chemicals sector the CEI obtained from the nonparametric frontier approach showed a slightly larger variance than that from the weighted sum aggregation rule. In addition, the ranges of CEI values from the two methods are also very close to each other.

Table 27: Summary statistics of the CEIs by two aggregation methods

Industry		**Nonparametric frontier approach**	**Weighted sum aggregation cum min-max normalization**
Chemicals	Mean	0.75	0.63
	Std. Dev.	0.18	0.13
	Median	0.72	0.68
	Min	0.22	0.22
	Max	1.00	0.99
Primary Metals	Mean	0.84	0.66
	Std. Dev.	0.20	0.18
	Median	0.96	0.75
	Min	0.27	0.26
	Max	1.00	0.98
Fabricated Metals	Mean	0.83	0.70
	Std. Dev.	0.12	0.12
	Median	0.88	0.75
	Min	0.52	0.40
	Max	1.00	0.91

Four hypotheses are proposed and tested to investigate whether there exist significant differences in the CEIs computed by the two aggregation methods and for the three industries when only the nonparametric frontier approach is employed. The proposed null hypotheses are described as follows:

(1) The choice between the nonparametric frontier approach and the weighted sum aggregation does not affect the CEIs;

(2) The chemicals sector has the same environmental performance as primary metals sector in toxic releases;

(3) The primary metals sector has the same environmental performance as fabricated metals sector in toxic releases;

(4) The chemicals sector has the same environmental performance as the fabricated metals sector.

Since the two sets of CEI values as well as the differences derived do not follow a normal distribution, we employ the commonly used Wilcoxon-Mann–Whitney rank-sum-test to test the four hypotheses. To test hypothesis (1), the CEI values for the three sectors are separately used, which leads to three sets of testing results. For testing hypotheses (2) to (4), we only use the CEI values computed by the nonparametric frontier approach.

Table *28* shows the results of the hypothesis tests. It can be observed that the three null hypotheses for comparing two different aggregation methods are all rejected at the 0.01 level of significance implying that the nonparametric frontier approach yields larger CEI values than the weighted sum aggregation rule. In addition, the CEI results obtained from the nonparametric frontier approach suggest that the primary metals industry might show better environmental performance than the chemicals industry. However, there is no statistical evidence for rejecting the last two hypotheses at the 0.01 level of significance, which indicates the differences between the chemicals and fabricated metals industries as well as between the primary metals and fabricated metals industries in environmental performance are not significant.[6]

Table 28: Summary of hypothesis test results

Null hypothesis	**Mann-Whitney U**	***p*-value**
H_{01a}: Mean($CEI_{Ch\text{-}RAM}$)=Mean($CEI_{Ch\text{-}WS}$)	5225	0.0018
H_{01b}: Mean($CEI_{PM\text{-}RAM}$)=Mean($CEI_{PM\text{-}WS}$)	1062	0.0013
H_{01c}: Mean($CEI_{FM\text{-}RAM}$)=Mean($CEI_{FM\text{-}WS}$)	3839	0.0000
H_{02}: Mean(CEI_{Ch})=Mean(CEI_{PM})	1752	0.0059
H_{03}: Mean(CEI_{PM})=Mean(CEI_{FM})	1451	0.0263
H_{04}: Mean(CEI_{Ch})=Mean(CEI_{FM})	3650	0.0638

We also investigate the correlation between the CEIs derived from the two different aggregation rules. The Pearson and Spearman rank correlation coefficients obtained are shown in Table *29*. There exist significant positive correlations between the CEI values derived from the two alternative aggregation rules. In particular, we find that both the Pearson and Spearman rank correlation coefficients of the two sets of CEIs for Fabricated Metals sector are as high as 0.998, which might be an indication of the robustness of the rankings to the choice of aggregation rule for different facilities in this sector.

Further comparisons between the two aggregation rules in deriving CEIs may be conducted by looking through facility-level CEI values. When the nonparametric frontier approach is used,

[6] It is worth pointing out that the environmental performance comparisons between different groups of facilities should be performed with caution when the CEIs are constructed by the nonparametric frontier approach. Since the CEI values are relative and not absolute ones, another possible reason for the between-group differences might be that the facilities in an industry are closer – on average – to the frontier for the industry compared to the facilities in another industry.

more than one facility will achieve a CEI value of unity unless there exists a super-facility dominating all the other entities in all the dimensions. While the weighted sum aggregation usually has higher discriminating power, the CEI values derived from the method could sometimes be misleading since the full compensability between all the variables is implicitly assumed. To examine this point, in Table *30* we summarize the CEI values from the weighted sum aggregation as well as the original data for the facilities with a CEI value of unity using the nonparametric frontier approach.

Table 29: Correlation coefficients of CEIs derived from two aggregation models

Correlation type	Chemicals	Primary Metals	Fabricated Metals
Pearson	0.782*	0.913*	0.998*
Spearman	0.791*	0.902*	0.998*

*Correlated at the 1% significance level

Table 30: CEIs from weighted sum aggregation for the facilities with unity values using the nonparametric frontier approach

Industry	Facility	Quantity of Total Releases	Toxicity of Total On-site Toxic Releases	Percentage of Waste Managed (%)	Gross Revenue	CEI (Weighted sum)
Chemicals	No. 27	1085	488	99	1310	0.99
	No. 60	1	1760	0	93	0.52
	No. 63	0	0	100	91	0.77
Primary	No. 7	128	17600	100	54	0.79
	No.12	21	27896	67	129	0.77
	No.13	868	295434	93	319	0.98
	No.16	541	565800	87	266	0.92
	No. 23	0	0	100	49	0.79
	No. 24	5	37460	0	86	0.57
	No. 29	15823	73472	99	102	0.83
Fabricated	No. 10	16200	1592210	93	378	0.91
	No. 54	0	0	100	178	0.87

It can be observed from Table 30 that three facilities from the chemical industry and two facilities from the fabricated industries are located on the best practice frontiers. For the primary metals industry, however, seven facilities constitute its best practice frontier. The variation should mainly be attributed to the differences in the numbers of facilities in the three sectors. While the chemicals and fabricated metals industries respectively consist of 67 and 54 facilities, the primary metals industry has only 29 facilities. Although the higher discriminating power of weighted sum aggregation rule in constructing CEIs is insensitive to sample size, the meaningfulness of the CEIs based on this rule is questionable. For example, in the chemicals sector facility no. 63 had no toxic releases, which should be an indication of better environmental performance. However, since its gross revenue is substantially less than that of facility no. 27, the CEI value of facility no. 63 is much smaller than that of facility no. 27 although the latter produced 1085 pounds of toxic releases. The same cases occur for the facility no. 23 in the primary metals sector and the facility no. 54 in the fabricated metals sectors. In evaluating the facility-level environmental performance, it is logical to reward the facilities producing higher revenues with the same impact of toxic releases by giving them higher CEI scores. But it may not be reasonable to give better evaluation of a facility with relatively higher revenue and toxic releases than another facility without any toxic releases. Compared to the weighted sum aggregation rule, the nonparametric frontier approach treats the facilities with distinctive strengths in various dimensions indifferently by giving them the same CEI values of unity, which seems to be more reasonable for an appropriate environmental performance assessment.

As described in Section 3, an additional strength of the nonparametric frontier approach is that it may help each facility identify its benchmark (groups) as well as the directions of potential improvement in different dimensions. Take the variable QTTR as an example. We can use Eq. (7) to compute the optimal slack in QTTR for each facility within the three sectors, which indicates that facility's potential reduction in QTTR to reach the level of its benchmark. Figure *5* shows the sectoral potential reduction in QTTR as a percentage of the actual QTTR for each of the three sectors. It is observed that almost 90% of QTTR for Chemicals sector could be reduced if all the facilities reach the levels of their benchmarks. For primal metals and fabricated Metals sectors, the potential reductions in QTTR are more than 95% of the actual quantities! This result might be explained by the fact that there are some facilities with poor environmental performance and very high QTTR. For example, facility no. 21 in the primary metals sector could reduce 2,822,677 pounds of toxic releases with reference to a convex combination of facility no. 13 and no. 23, which accounts for 97% of potential reductions in QTTR for the whole sector. While it shows a huge potential in reducing toxic releases, this might be unrealistic due to the scale discrepancy between facility no. 21 and its benchmark group. Nevertheless, it at least offers a direction along which facility no. 21 may improve its environmental performance through managerial efforts.

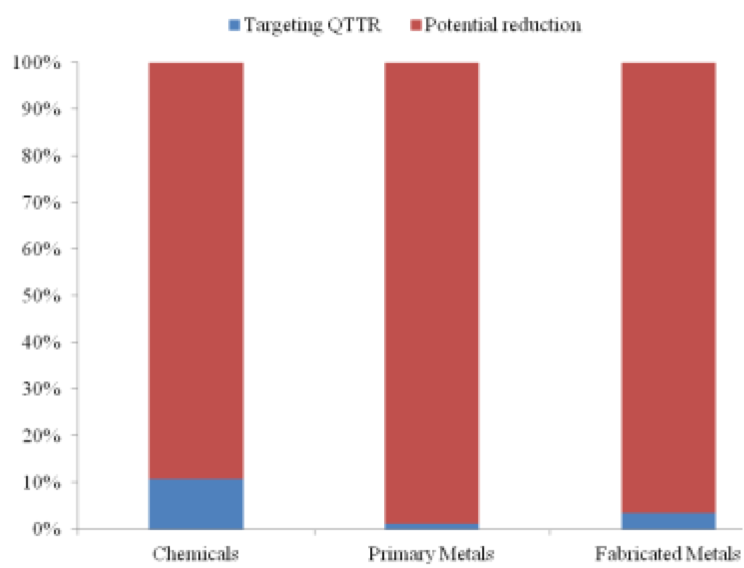

Figure 5: Percentages of potential reductions in QTTR for three industries

Table 31 summarizes the appearance frequencies of the facilities forming the best practice frontiers of the three industries. It can be seen that facility no. 27 in the Chemicals, no. 23 in the Primary Metals and no. 54 in the Fabricated Metals are most frequently identified as the benchmarks. Referring to Table 30, we find that the latter two are indeed the best performers in toxic releases with reasonable revenues. In terms of facility no. 27 in the chemical industry, 99% of those toxic releases were properly handled. Meanwhile, this facility generates a gross revenue that is over ten times the revenue of the other two facilities forming the best practice frontier of the Chemicals industry. In view of these features, it is not surprising that facility no. 27 has been identified as a benchmark most frequently. On the contrary, although facility no. 60 in the chemicals industry and nos. 12 & 24 in the primary metals industry also have CEI values of unity, they are not used to evaluate any other facilities except themselves. It implies that the three facilities, which did not perform well in managing toxic releases, cannot be dominated by any convex combination of other facilities, and therefore should not be set as the benchmarks in environmental performance assessment.

Table 31: Appearance frequencies of the facilities forming the best practice frontiers

Industry	**Facility no.**	**Frequency of appearance in the best practice frontier**
Chemicals	**#27**	**53**
	#60	*1*
	#63	35
Primary Metals	#7	5
	#12	*1*
	#13	13
	#16	9
	#23	**19**
	#24	*1*
	#29	5
Fabricated Metals	#10	30
	#54	**52**

4.4 Sensitivity analysis

Our CEI is derived through solving a series of linear programming models and may be affected by some uncertainty factors. First, the best practice frontier, formed by the existing facilities, is an estimate of the "true" frontier, which might be subject to uncertainty arising from the sampling variation of the obtained frontier. Although the uncertainty can be handled by using bootstrap methods for assessing the sampling variation (Simlar and Wilson, 2015), it should be noted that the best practice frontier in the context of constructing CEIs is somewhat different from the production frontier in efficiency and productivity analysis. One main usage of our CEI is to conduct cross-sectional comparison or monitor the environmental performance over time. Such an application context allows us to use the observations from all the comparable entities at different time points to form the best practice frontier and construct CEIs. As such, the uncertainty due to the sampling variation of the obtained frontier will not be studied in this paper.

Data accuracy is another important source of uncertainty. In the case of TRI data, as pointed out by Toffel and Marshall (2004), the uncertainty in data accuracy made the development of environmental performance metrics difficult. While the sensitivity of the nonparametric frontier approach with respect to data perturbation could be theoretically examined, in this paper we conduct a sensitivity analysis of the CEIs by artificially changing the observations in a random way. It is assumed that the data errors for all the observations are within $\pm 10\%$ of the data observed. By generating random numbers within [-10%, 10%], we create 50 datasets for each of the three industries based on which 50 sets of CEIs can be derived. Using the 50 sets of CEIs, we first compute the average CEI values as well the corresponding standard derivations for each of three industries. Figure 6 shows the box plots of the averages and standard deviations of CEI values by industry.

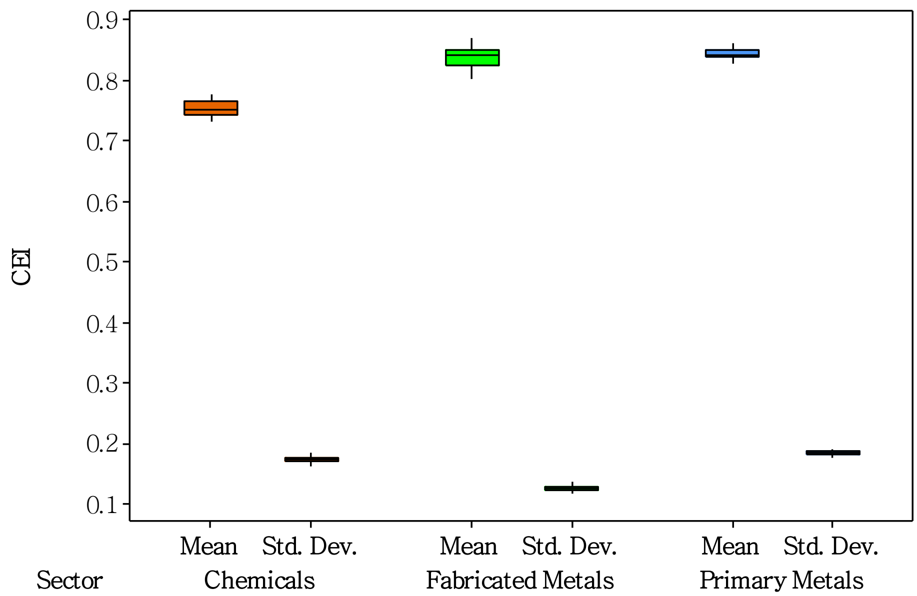

Figure 6: Boxplots of the average and standard deviation values of CEIs by industry

It can be observed from . Figure 6 that the impact of data uncertainty in the specified range on the average CEI value of each industry is relatively weak, especially for the case of the primary metals sector. Meanwhile, data variation has very little impact on the standard derivation of the sectoral average CEI values. It suggests that the sectoral average CEI values are quite insensitive to the uncertainty in data accuracy (within the specified range). To investigate whether the dispersion of CEI values for all the facilities in each of the three industries varies significantly when the uncertainty in data accuracy is considered, we show the box plots of the average CEI values from the simulated data for all the facilities in each of the three industries in Figure 7. For comparison purposes, the box plots of the CEI values derived from the original dataset are also provided. It can be seen from Figure 7. that there are few changes in the median and ranges of CEI values for the chemical and fabricated metals industries. However, the change in the median of CEI values for the primary metals industry seems to be slightly larger, which could be due to the relatively small number of facilities in this industry. It might be an indication that CEI values are insensitive to the uncertainty in data accuracy when the sample size is relatively large.

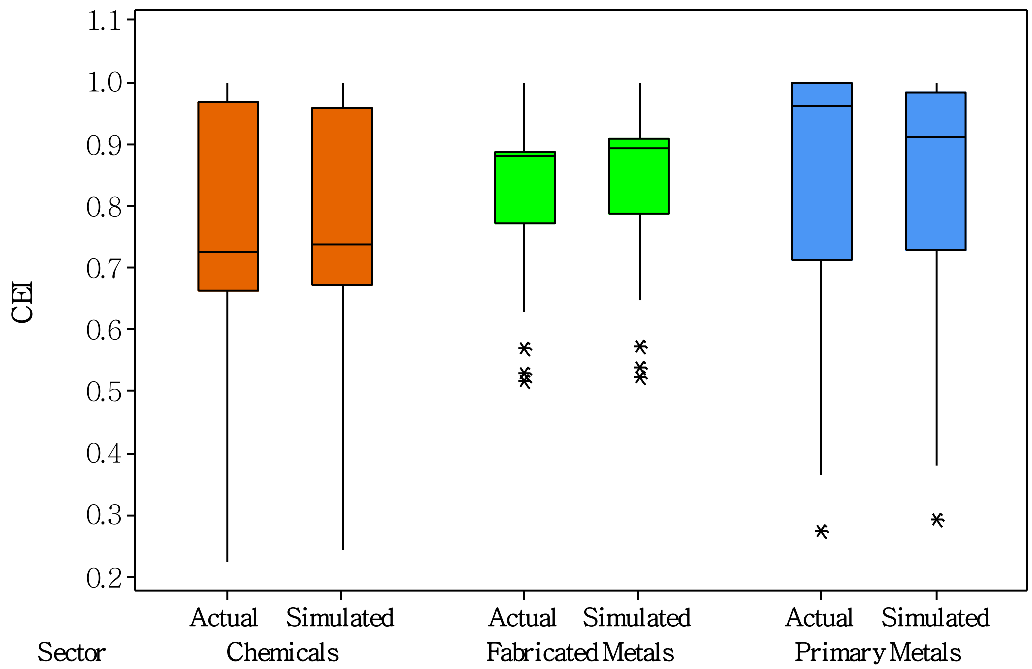

Figure 7: Boxplots of the CEI values for all facilities derived from both actual and simulated data

As the uncertainty in data accuracy has a relatively larger impact on the CEI values of facilities in the primary metals industry, it is worthwhile looking through the types of facilities that would be more easily affected in their CEI values. Figure 8 shows the CEI value of each facility in the primary metals industry as well as the corresponding average CEI value derived from the simulated data. It is found that most of the facilities have small changes in CEI when data variation exists. In particular, five facilities, i.e. nos. 12, 13, 16, 23 and 24, are always located at the best practice frontier which might be an indication of the robustness of the best practice frontier with respect to the uncertainty in data accuracy. However, several facilities such as nos. 8, 11 and 20 show relatively larger gaps between the actual CEI value and the CEI value from simulated data.

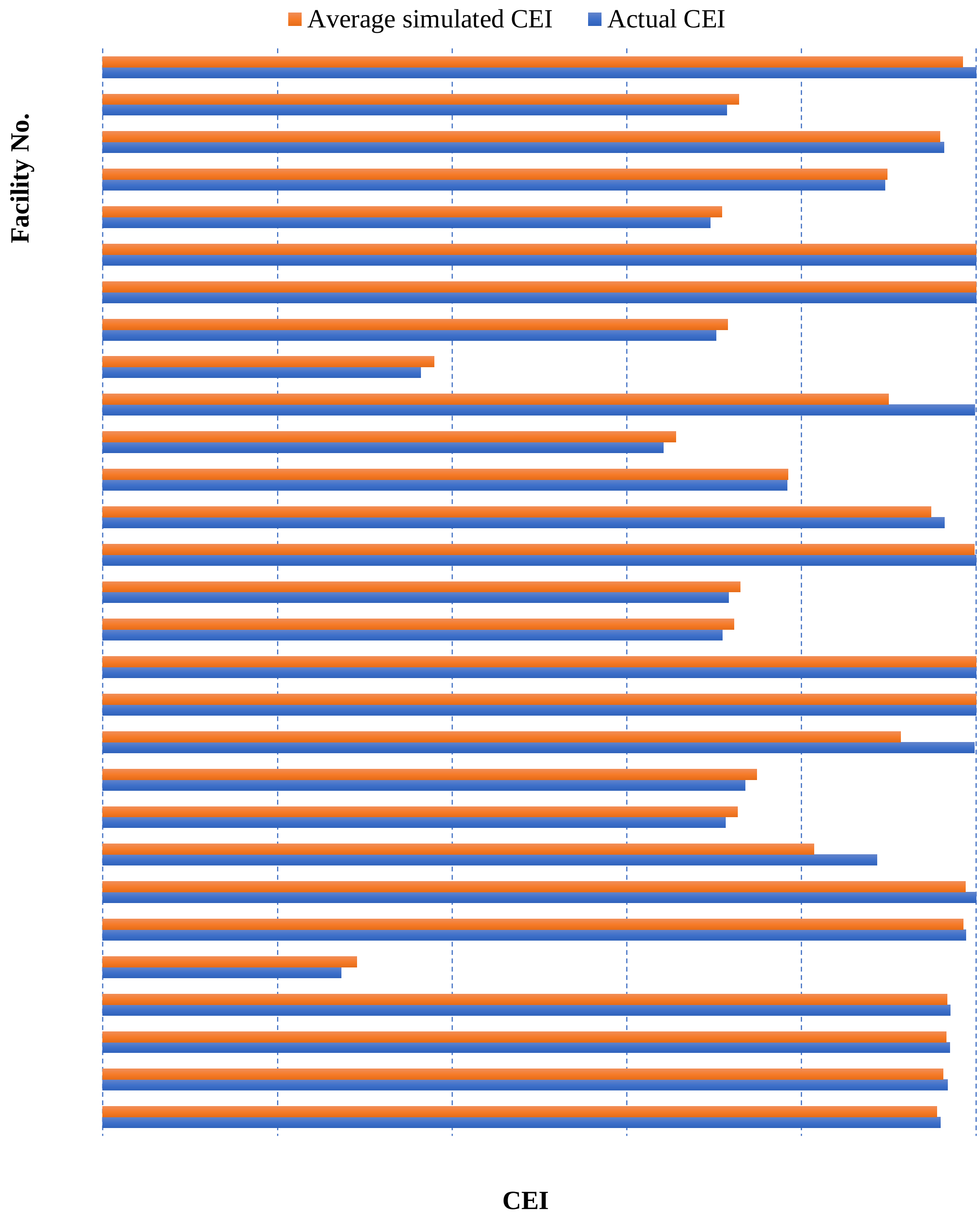

Figure 8: Comparison between CEIs from original and simulated datasets for the facilities in the Primary Metals industry

5 Conclusions

This paper argues that it is important to construct a cardinally meaningful and standardized CEI for the measurement of environmental performance. A CEI is said to be cardinally meaningful if its values are invariant with respect to the changes in the measurement units of underlying variables. This concept is particularly important when the cardinality characteristics of CEIs are concerned. The commonly used aggregation methods, e.g. arithmetic and geometric aggregation methods, cannot yield a cardinally meaningful CEI when mixed measurability of underlying variables is involved. We propose to use a nonparametric frontier approach, i.e. range adjusted DEA model, to construct a cardinally meaningful CEI, which can easily handle the issues of mixed measurability of underlying variables and data irregularity such as the existence of multiple zeros.

We apply the nonparametric frontier approach to constructing a CEI for evaluating the facility-level environmental performance of toxic releases in three industries (i.e. chemical, primary metals and fabricated metals) in Los Angeles County based on the latest TRI data. At the industry level, we find that the primary metals industry shows better environmental performance than the chemical industry while other pairwise comparisons do not show statistically significant differences. In addition, we summarize the benchmark facilities in every industry as well as their appearance frequency in forming best practice frontiers, which represent targets for other facilities to improve their environmental management practices. Finally, we investigate whether the uncertainty in data accuracy has a significant effect on the CEI results. Our results show that the distributions of the CEI values change very little when confronted with data errors of 10%, which might be an indication of the robustness of our CEI in evaluating facility-level environmental performance.

References

Aida, K., Cooper, W.W. Pastor, J.T., Sueyoshi, T., 1998. Evaluating water supply services in Japan with RAM: A range-adjusted measure of inefficiency. Omega 26, 207-232.

Böhringer, C., Jochem, P.E.P., 2007. Measuring the immeasurable – A survey of sustainability indices. Ecological Economics 63, 1-8.

Boyd, G.A., McClelland, J.D., 1999. The impact of environmental constraints on productivity improvement in integrated paper plants. Journal of Environmental Economics and Management 38, 121-142.

Bui, L.T.M., Kapon, S., 2012. The impact of voluntary programs on polluting behavior: Evidence from pollution prevention programs and toxic releases. Journal of Environmental Economics and Management 64, 31-44.

Chambers, R.G., Chung, Y., Färe, R., 1996. Benefit and distance functions. Journal of Economic Theory 70, 407-419.

Coelli, T.J., Rao, D.S.P., O'Donnell, C.J., Battese, G.E., 2005. An Introduction to Efficiency and Productivity Analysis (2nd ed.). Springer.

Cooper, W.W., Park, K.S., Pastor, J.T., 1999. RAM: A range adjusted measure of inefficiency for use with additive models, and relations to other models and measures in DEA. Journal of Productivity Analysis 11, 5-42.

Cooper, W.W., Pastor, J.T., Borras, F., Aparicio, J., Pastor, D., 2011. BAM: A bounded adjusted measure of efficiency for use with bounded additive models. Journal of Productivity Analysis 35, 85-94.

Cole, M.A., Elliott, R.J.R., Khemmarat, K., 2013. Local exposure to toxic releases: examining the role of ethnic fractionalization and polarization. Ecological Economics 93, 249-259.

Delmas, M.A., Blass, V.D., 2010. Measuring corporate environmental performance: The trade-offs of sustainability rating. Business Strategy and the Environment 19, 245-260.

Delmas, M.A., Kohli, A., 2014. UCLA-US EPA TRI university challenge: Assessing and communicating the environmental impact of TRI facilities in Los Angeles County. Project report, UCLA Institute of the Environment and Sustainability.

Ebert, U., Welsch, H., 2004. Meaningful environmental indices: A social choice approach. Journal of Environmental Economics and Management 47, 270-283.

Färe, R., Grosskopf, S., Hernandez-Sancho, F., 2004. Environmental performance: An index number approach. Resource and Energy Economics 26, 343-352.

Färe, R., Grosskopf, S., Lovell, C.A.K., Pasurka, C., 1989. Multilateral productivity comparisons when some outputs are undesirable: A nonparametric approach. The Review of Economics and Statistics 71, 90-98.

Färe, R., Grosskopf, S., Pasurka, C., 2006. Social responsibility: U.S. power plants 1985-1998. Journal of Productivity Analysis 26, 259-267.

Färe, R., Grosskopf, S., Pasurka, C., 2007. Environmental production functions and environmental directional distance functions. Energy 32, 1055-1066.

Färe, R., Grosskopf, S., Pasurka, C.A., 2010. Toxic releases: An environmental performance index for coal-fired power plants. Energy Economics 32, 158-165.

Färe, R., Grosskopf, S., Tyteca, D., 1996. An activity analysis model of the environmental performance of firms – application to fossil-fuel-fired electric utilities. Ecological Economics 18, 161-175.

Gerde, V.W., Logsdon, J.H., 2001. Measuring environmental performance: Use of the toxics releases inventory (TRI) and other US environmental databases. Business Strategy and the Environment 10, 269-285.

Hailu, A., Veeman, T.S., 2001. Non-parametric productivity analysis with undesirable outputs: An application to the Canadian pulp and paper industry. American Journal of Agricultural Economics 83, 605-616.

Hanna, B.G., 2007. House values, incomes, and industrial pollution. Journal of Environmental Economics and Management 54, 100-112.

Hoang, V.N., Coelli, T., 2011. Measurement of agricultural total factor productivity growth incorporating environmental factors: A nutrients balance approach. Journal of Environmental Economics and Management 62, 462-474.

Khanna, M., Damon, L.A., 1999. EPA's voluntary 33/50 program: Impact on toxic releases and economic performance of firms. Journal of Environmental Economics and Management 37, 1-25.

Khanna, M., Quimio, W.R.H., Bojilova, D., 1998. Toxics release information: A policy tool for environmental protection. Journal of Environmental Economics and Management 36, 243-266.

Koehler, D.A., Spengler, J.D., 2007. The toxic release inventory: Fact or fiction? A case study of the primary aluminum industry. Journal of Environmental Management 85, 296-307.

Konar, S., Cohen, M.A., 1997. Information as regulation: The effect of community right to know laws on toxic emissions. Journal of Environmental Economics and Management 32, 109-124.

Munda, G., Nardo, M., 2009. Noncompensatory/nonlinear composite indicators for ranking countries: A defensible setting. Applied Economics 41, 1513-1523.

Picazo-Tadeo, A.J., Castillo-Giménez, J., Beltrán-Esteve, M., 2014. An intertemporal approach to measuring environmental performance with directional distance functions: Greenhouse gas emissions in the European Union. Ecological Economics 100, 173-182.

Pittman, R.W., 1983. Multilateral productivity comparisons with undesirable outputs. Economics Journal 93, 883-891.

Pollesch, N.L., Dale, V.H., 2016. Normalization in sustainability assessment: Methods and implications. Ecological Economics 130, 195-208.

Simar, L., Wilson, P.W., 2015. Statistical approaches for non-parametric frontier models: A guided tour. International Statistical Review 83, 77-110.

Toffel, M.W., Marshall, J.D., 2004. Improving environmental performance assessment: A comparative analysis of weighting methods used to evaluate chemical release inventories. Journal of Industrial Ecology 8, 143-172.

Tyteca, D., 1996. On the measurement of the environmental performance of firms – A literature review and a productive efficiency perspective. Journal of Environmental Management 46, 281-308.

Tyteca, D., 1997. Linear programming models for the measurement of environmental performance of firms – Concepts and empirical analysis. Journal of Productivity Analysis 8, 183-197.

Welsch, H., 2005. Constructing meaningful sustainability indices. In: C. Böhringer, A. Lange (eds.), Applied Research in Environmental Economics, Physica-Verlag, Heidelberg, pp. 7-22.

Zhou, P., Ang, B.W., 2009. Comparing MCDA aggregation methods in constructing composite indicators using the Shannon-Spearman measure. Social Indicators Research 94, 83-96.

Zhou, P., Ang, BW., Han, J.Y., 2010. Total factor carbon emission performance: A Malmquist index analysis. Energy Economics 32, 194-201.

Zhou, P., Ang, B.W., Poh, K.L., 2006. Comparing aggregating methods for constructing the environmental index: An objective measure. Ecological Economics 59, 305-311.

Zhou, P., Ang, B.W., Wang, H., 2012. Energy and CO_2 emission performance in electricity generation: A non-radial directional distance function approach. European Journal of Operational Research 221, 625-635.

IV. Conclusion

1 Introduction

There has been a rise of environmental information programs in recent years. These programs have been implemented to increase public awareness of the environment, associate health risks and ultimately encourage the adoption of healthier behaviors and environmentally-friendly practices. As the public demands transparency and information about pollutants in the environment and their health impact, these programs also serve to increase trust between the government and the public.

A lot of resources are used to implement these programs; by collecting data and making it easily accessible to a large number of people. Effective communication of these programs is a critical factor that determines their success. Some important communication questions for discussion are:

- Is the information easily accessible and understandable?
- Is the information reaching the intended audience i.e. the people most impacted?
- Is it resulting in the adoption of healthier behaviors or environmentally friendly practices?

I explored some of these questions in my dissertation by investigating the use of data from the AirNow program and the Toxic Release Inventory (TRI) program.

2 Air Quality Information Programs

Through the AirNow program, established in the late 90's, real-time air quality (AQ) information is available for most cities in the US. This information is available through a variety of modes such as their website and mobile applications (apps). There is very little information about who uses this information and their response to it. I worked with my team of app developers to develop our own AQ app, AirForU, with an inbuilt-research study. The app was available for free for iPhones and Android devices. The app used AQ data from the AirNow program.

The goal of this research study was to understand engagement with AQ information. Over a period of a year and a half, I collected data on over 2700 app users and studied their engagement with this information. Since downloading the app was voluntary, there was a self-selection bias. The good news is that the bias leaned towards vulnerable groups that are more impacted by air pollution (AP). These groups are interested in this information and want to take steps to protect their health. Compared to CA and US averages, the % of app users who were suffering from health conditions aggravated by air pollution (such as asthmatics, heart and lung disease patients) was much higher. The same was true of app users who had children affected by some of these health conditions.

Over time (approximately 12 weeks), most user apps tended to become disengaged with the app. The next step was to see if we could influence engagement by sending AP messages via email to the app users. Using the AP health literature and message framing literature 12 AP messages were developed. There were 6 message categories and within each category there was a positively and negatively framed message. A survey was conducted to test hypothetical engagement. In the survey, the exercise and AP invisibility category were the most effective messages. Messages targeted at certain groups were more effective among those groups.

After a year and a half since the app's launch, an email experiment was conducted where app users were split into random groups and treated with the same AP messages as the survey. Their engaged with the app was recorded after the experiment. The results were quite different from the survey. Engagement after the emails were sent out was strongly dependent on prior levels of engagement; there was little change for those who were already highly engaged, the group of users that were moderately engaged tended to become more engaged and the group of users that were least engaged did not respond to the emails. There was very little difference in response to the content of the email, rather users responded to the emails as a reminder to check AQ. An exit survey indicated the adoption of health protective behaviors such as avoiding outdoor exercise, increase air conditioning use or the use of filters but these changes were limited to a small group of app users.

While the results of the research study were useful, there remains a lot to be done to better understand engagement with AQ information. An important avenue for research is to develop strategies to keep people more engaged. Some strategies to increase awareness and engagement that need more research are:

- Periodic reminders; the length of time might be different for different people
- Reminders via email, text or app alerts
- Targeted or tailored messages

- The use of prizes or raffles to encourage people to remain motivated and take action steps to protect their health

The number of strategies that could be tested in the present study was limited by the number of app users. Testing strategies is limited by the number of app users. The applicability of the results was also limited by geography; most AirForU users were in Southern California. UCLA health, located in this region, was a major partner in the recruiting effort. The app user base can be expanded in number and geography through the following steps

- Incorporating social media features that allow app users to share information with their friends and family easily through the app and send them invitations to download the app
- Partner with health networks throughout the nation to inform more people, particularly sensitive groups

The public can play an important role in improving reporting of AP (Yearley, 2006) and eventually in reducing AP particularly in urban areas (Bickerstaff & Walker, 2001). The public can provide information on local sources and improve the monitoring and reporting of AQ information (Yearley, 2006). This can be accomplished by establishing a 2-way interface in the app that allows users to be more active e.g. upload pictures of polluted areas. The public can also contact local policy-makers through the app. One of hindering factors in enabling people to act on the issue of AP is that they don’t know whom to contact or how to contact them (Wakefield, Elliott, Cole, & Eyles, 2001). Eventually the public can be motivated to take not just health protective steps but also environmentally friendly actions that reduce AP; a large percentage of people have indicated an interest in learning about the steps that they can take to reduce air pollution (Beaumont et al., 1999).

3 Toxic Release Inventory

This dissertation research also focused on improving the accessibility of data in the TRI program. One of the obstacles that hindered the success of the TRI was the complexity of the data that made it inaccessible to the public and the facilities reporting to the TRI. The use of a range-adjusted data envelopment analysis (DEA) model was explored in simplifying TRI data and it proved to be successful in developing an environmental and economic performance index. this was one of first studies that measured environmental and economic performance at the facility level, rather than the company level.

While this was an important step towards improving the accessibility of this data, it need to be more widely available to facilities and the public. Facilities can be trained in using this a tool to monitor and improve their performance. Currently this was only done for one year of TRI data and was restricted to facilities in the Los Angeles region. These methods can be used to performance analyses for several years and be used towards identifying trends. This information can be used by the public to contact local policymakers about facilities in their neighborhoods this generating publicity to influence positive environmental change.

4 Conclusion

All of the above steps can improve environmental information programs. More effective communication strategies will allow agencies to reach a larger number of people, potentially avoid part of the health burden associated with poor AQ and toxic releases, and even reduce AP and toxic chemical releases by promoting environmentally friendly behaviors. The implications of this research are beyond the AirNow and TRI programs and can easily be extended to other areas of health and environmental protection. Understanding dynamic responses, avoidance behavior patterns and the adoption of environmentally friendly behaviors can shed light towards the development of public policies that focus on improving health and environmental outcomes. Information policies, which are often low-cost, have the potential to become a major driver of behavioral change when implemented well.

References

Beaumont, R., Hamilton, R. S., Machin, N., Perks, J., & Williams, I. D. (1999). Social awareness of air quality information. *The Science of the Total Environment*, *235*, 319–329.

Bickerstaff, K., & Walker, G. (2001). Public understandings of air pollution: The “localisation” of environmental risk. *Global Environmental Change*, *11*(2), 133–145. http://doi.org/10.1016/S0959-3780(00)00063-7

Wakefield, S. E. L., Elliott, S. J., Cole, D. C., & Eyles, J. D. (2001). Environmental risk and (re)action: Air quality, health, and civic involvement in an urban industrial neighbourhood. *Health and Place*, *7*(3), 163–177. http://doi.org/10.1016/S1353-8292(01)00006-5

Yearley, S. (2006). Bridging the science-policy divide in urban air-quality management: Evaluating ways to make models more robust through public engagement. *Environment and Planning C: Government and Policy*, *24*(5), 701–714. http://doi.org/10.1068/c0610j