Leveraging behavioral data for public health: exploring sleep and nature exposure using mobile phones and social media

A Dissertation Presented

by

Kelsey Meegan Linnell

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
Specializing in Mathematical Science

August, 2022

Defense Date: May 16th, 2022
Dissertation Examination Committee:

Christopher M. Danforth, Ph.D., Advisor
Peter Sheridan Dodds, Ph.D., Advisor
Donna Rizzo, Ph.D., Chairperson
Jun Yu, Ph.D.
Cynthia J. Forehand, Ph.D., Dean of Graduate College

Abstract

Health surveillance and assessment are considered essential components of a functional public health system. The recent ubiquity of mobile devices and social media have created a wealth of behavioral data, and bring into existence new forms of population health monitoring. These new digital sources can provide direct and passive data for more detailed and nuanced health factors, and have expanded the human, spatial, and temporal scales at which these factors can be measured. In this project, I leverage digital trace data from tweets and mobile device location pings to explore population scale sleep loss, and nature exposure through park visitations in the United States. Both sleep and nature exposure are essential contributors to well-being, and have historically relied on either survey data or direct observation of individuals to measure. I begin by demonstrating the ability of Twitter data to passively reflect population-scale sleep loss at the state level. This is followed by an exploration of park visitation measured through mobile device GPS data. Changes in county-scale park visitation behavior at the onset of the COVID-19 pandemic are analyzed and comparisons are made using population density, employment sector, income, and voting records. In the final chapter I investigate the viability of predicting park visitation using demographic information from the surrounding neighborhood. I conclude with a brief discussion of the significance of measuring these behaviors, and the potential for health policy improvement.

For Clementine

Acknowledgements

This document exists as a result of the extreme generosity and daily support of my family, friends, and the members of Computational Story Lab. This space is not sufficient to encompass the many, many, people and opportunities for which I have gratitude, and without which this work would not have been possible.

In the Fall of 2018 I began to feel lost in academia, and by December several unfortunate events had taken place in my personal life. In the Spring of 2019 I was without hope, and had decided to abandon my pursuit of the PhD. During this time I encountered a period of identity fluidity, and commitment to pursuing happiness that led to the acquisition of two cows and a small herd of dairy goats.

In June of 2019 I was tending to my ruminants on a small farm in rural Maine, feeling conflicted about my 'failed' pursuit of the PhD, and contemplating what my future should be. Professor Christopher Danforth was at this moment the chair of the Mathematics Department at UVM, and though he was (and still is) a very busy individual, he had noticed my absence and reached out. That one phone call enabled this entire work, and proved pivotal for my life as a whole.

I am deeply grateful that Professor Danforth made the decision to support me as my advisor following that phone call. I had come to believe that I was deficient and incapable of engaging in research; Professor Danforth chose to believe that I only needed a different form of support and guidance. The risk he was willing to take by offering me a position in his lab gave me hope that my beliefs about myself were wrong. For years I was unable to believe directly in myself, but was able to make progress because I trusted Professor Danforth, and he believed in me.

At the time I began this journey as a part of the Computational Story Lab I

was already living in Maine, and as mentioned, had acquired livestock. It was clear that if I was to reinvest myself in pursuing a PhD it would be some sort of hybrid remote situation. I distinctly recall Professor Danforth saying ”Peter and I have never successfully advised someone who hasn’t been on campus very regularly before, but we are willing to try.” I am still in awe of the generosity that Professors Danforth and (Peter) Dodds showed me in this effort. Creating this work remotely has been difficult, but the ability to live with my husband and family, and to keep and care for my livestock, has been central to my well-being in the last three years.

I am also deeply grateful to Professors Dodds and Danforth for bringing Mikaela Fudolig into the lab. Mikaela has acted as a primary advisor for me in the last year and a half, and without her I would have made much less progress. Mikaela is enthusiastic, empathetic, practical, and driven. Working with her has made me a better scientist, and has made science more enjoyable. I am forever thankful to her for the time she has spent reading my papers, reviewing my code, meeting with me, and picking at my figures to make them at least presentable.

One of the final catalytic moments of this work came as a text received from Molly Rose Kelly-Gorham on Monday December 6 of 2021: "Is there any chance you want to do a power hour together each workday until Christmas?" Molly and I met through the IGERT Smart Grid program, and she was pursuing a PhD in Electrical Engineering. After the onset of the pandemic several of the grad students who had been in the IGERT program dispersed and began working remotely - when we felt like working together over a video call for a short burst we called it a ”power hour.” The pandemic started with many power hours, but they had dwindled to occasional or rare events within a few months. We agreed to try it. After Christmas break we decided to keep going. Almost every week day since I have woken up, milked and fed my animals, and then joined Molly to begin work for the day. I don't think I could have finished this without her.

To everyone who should be mentioned, but whose story I cannot fit - I hope that I show you enough love and appreciation everyday to make up for it.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Health surveillance and monitoring are essential components of a public health system. Without adequate data reflecting the health status of a population there can be no informed strategy or intervention development. The need for public health monitoring has been recognized for thousands of years, and historically has been most significant during epidemics and pandemics [3]. It is during these periods of high mortality and contagion that humans have developed methods for recording deaths, illnesses, contact tracing, and begun monitoring others for the purpose of quarantining them [3].

Though the collection of data on deaths and on contractions of communicable disease has been in existence for thousands of years, decisions to monitor public health in a more continuous manner, and in what might be considered more mundane regards, have only slowly developed over the last several hundred years. Regular reporting for the totality of deaths, for example, did not arise until the 1600s when the London Mortality Bills began a weekly publication [3, 4]. When considering the irregularity of reporting, and limited information collected, it is important to consider both the availability of health data- but also of the purpose of that data.

The maintenance of public health as a function of the government is, in itself, a rather recent development. It was not until the French Revolution that public health became a matter for the government, rather than individuals [3, 5]. As governments in the West assumed responsibility for public health over the next hundred years, and recognized a relationship between policy and health status of their populations, and the frequency of collection and richness of information reported increased [6]. Still, this data was concerned with communicable diseases - primarily cases, and deaths. It wasn't until the 1900s, when chronic disease became a greater threat to mortality than infectious disease, that the scope of public health monitoring expanded.

The expansion past the simple concern of infectious disease is incredibly recent. The establishment of a national public health agency in the United States took place in 1946, and this agency was known as the "Communicable Disease Center." The CDC did not engage in chronic disease surveillance until the 1970s, and it wasn't until the 80's that interest was extended beyond the presence of disease and into risk factors including behavior [7]. In 1984 the Center for Disease Control began to monitor and evaluate public health with respect to behavior via an annual telephone survey called the Behavioral Risk Factor Surveillance System (BRFSS) [8].

The organization and refinement of the BRFSS marked a maturation of the public health system by asserting human behavior to be an influential component of a public health worthy of surveillance. This assertion carried with it the implication that behavior was impacted by, and could possibly be mitigated using, policy. Despite clear interest in population behavior, monitoring and measuring behavior at the population scale remains a challenge.

In the last decade new opportunities for behavioral monitoring have arisen from the near ubiquity of mobile devices, the wide-spread use of wearables, and regular engagement with social media platforms. A large majority of Americans have a smartphone (85%) and use the social media platform Facebook (69%), and 30% are now using wearable health care devices like Fitbits and Apple Watches [9,10]. Interactions with technology, including simply wearing a fitbit, or carrying a cell phone, create digital traces which can be used to infer information about user behavior, and when aggregated can generate insights into entire populations.

The prevalence of this type of data makes monitoring at the population scale possible, and the regular and frequent use of technology allows for not only greater temporal scale, but finer granularity in data collection as well. Additionally, these data sources allow for both passive data-collection, and the consideration of more nuanced behavior than previously examined. By studying the language used on Twitter, for example, it has been possible for scholars to passively monitor the ambient happiness of nations on a daily basis for several years [11]. Language on Twitter has also been used to demonstrate caloric imbalances between exercise and eating that correspond to regional rates of obesity - a measure that would have previously required collection of medical records, and food and exercise diaries [12].

Chapter Two of this work interrogates the ability of Twitter data to reflect another aspect of behavioral health: sleep. Short sleep, less than six hours a night, is a well established risk factor for several severe chronic diseases such as diabetes, heart disease, and cancer [13–16]. Sleep is also considered an aspect of behavioral health which can be directly impacted by policy such as residential noise limits, and school start times [17]. Despite the clear opportunity to improve public health by developing strategies to address short sleep, efforts to do so are hindered, in part, by the lack of sufficient population sleep monitoring methods. A recent paper by Leypunskiy et al [18] explored a method of measuring population sleep by using time of tweet postings as a proxy for sleep and wake data, and successfully found that when aggregated to a population scale Twitter data was able to reveal a "window of sleep opportunity."

Chapter two builds on Leypubskiy et al's work by further examining time of tweet posting as an appropriate proxy for sleep data. Spring Forward, discontinued in 2023, was the annual practice of an artificial time-loss in the middle of the night for the purpose of aligning sleep-wake hours with daylight. The event corresponded to a single hour artificially lost from the night, and thus constituted a sleep-loss event of approximately one hour for most of the United States. Historically, Spring Forward was used as the basis for determining the effect of sleep loss on unpredictable phenomena including strokes, heart attacks, and traffic accidents – all of which increased dramatically following Spring Forward [19–22]. Knowing that Spring Forward was a serious population scale sleep loss event, with serious public health ramifications, we posited that a tool used to measure population sleep scale should be able to capture this event; Chapter Two investigates whether or not time of tweet posting does.

Similarly to sleep, nature exposure has been both recognized as instrumental to well-being, and dependent upon either direct observation or surveys to measure. Chapters Three and Four leverage location data (GPS coordinates along with time stamps) collected from mobile devices to measure nature exposure through park visitations. With 85% of Americans using smartphones, monitoring the location of these devices gives incredible insight into mobility in the United States [23].

During the COVID 19 pandemic interest in mobile device location data spiked, as this data became useful for monitoring the impact of social distancing measures by investigating the number and types of locations visited, and the time spent in those locations [24]. One of the many types of venues that was scrutinized with this data was parks [25]. Early mobility studies determined that park visitations increased at the onset of the pandemic, however these studies were regionally focused, and they did not account for the seasonality of park visitation [25, 26].

Chapter Three builds on the exploration of park visitation behavior at the onset of the pandemic by considering thousands of parks across the United States, and accounting for seasonality in visitation. Using daily visitation counts for 2019 as well as 2020 allows us to account for seasonality. Models were constructed for park visitation in each state, and for counties containing a park for which we had data. By applying the Bayesian Estimator for Abrupt Seasonal and Trend change (BEAST) [27], which fits the model and detects change points in a time series, we were able to identify abrupt changes in park visitation which occurred at the beginning of the pandemic, and were not due to seasonality of visitations. Identifying the regions for which these changes took place allowed us to further explore differences in the populations of those regions, including differences in income, employment-share in industrial sectors, and voteshare for the 2016 presidential election.

Chapter Four makes further use of the park visitation dataset by investigating the relationship between the demographic features of the neighborhood surrounding a park, and the visitation received by that park. The neighborhood surrounding the park is referred to as 'the walkshed,' and is considered the area which is within walking distance (usually 10 minutes) of the park; this is the area that a park's primary users are considered to live [28–30]. The residents of the walkshed can also be thought of as a component of the park environment; the walkshed is the area that will need to be crossed to enter the park, and the clothing, houses, cars, and language within the walkshed will all be part of a perceived experience of traveling to, and visiting, a park.

A recent study using social media in New York City demonstrated that parks which were surrounded by populations with greater proportions of racial and ethnic minorities were visited less [30]. A second study which observed the proportion of days that mobile devices visited a park found that racial and ethnic minorities had less access to parks in terms of visitation [31]. If the members of a walkshed are the primary users of the park, these studies together may imply that park visitation can be predicted by the demographic features of its walkshed. The first study may also imply that the appeal of a park's environment, both the quality of the park, and the environment created by the neighborhood surrounding it, may be related to the demographic features of the walkshed.

These possible implications are investigated in Chapter Four by testing how successfully the demographic features of a park's walkshed can be used to predict the visitation received by the park. For this inquiry data was restricted to visits made in 2019, prior to disruption by the pandemic, and to parks which are in suburban or urban areas, and unlikely to be tourist destinations. The data curation for the study was designed to examine parks which were most likely to be visited regularly and by a fairly constant population.

In Chapter Four, digital mobility data allows us to passively observe park visitation to 2,506 parks across the United States over the course of a year. Not only is the scale of this study feasible because of the type of data being used, but the data is allowing us a more detailed look at accessibility to parks by exploring actual visitation. Several scholars have posited that inequity may exist in park accessibility as a result of constraints on time, safety, or other non-tangible aspects of life [32–37]. Saxon's 2021 work [31] using mobile device data for measuring park visitations demonstrated that actual park usage was not adequately captured by models which assumed usage based on proximity. By looking directly at visitation, we are able to recognize these unmeasurable constraints, and consider how they may modify behavior in ways that become apparent at the system scale.

Each of the chapters in this work demonstrates the applicability of a large, passively collected, behavioral dataset to understanding aspects of public health. Utilizing measurements of sleep and nature exposure gathered through digital traces has allowed me to observe phenomena and relationships at temporal, geographic, and human scales that were previously not possible to observe. These observations serve not only as interesting archeological artifacts- evidence of past human behavior and full of culture and geospatial heterogeneity, but as baselines for understanding future health data that may be collected, as well as examples of methodology for monitoring critical aspects of population health.

1.1 Publications

Material from this dissertation has been published as:

Kelsey Linnell, Michael Arnold, Thayer Alshaabi, Thomas McAndrew, Jeanie Lim, Peter Sheridan Dodds, and Christopher M. Danforth. The sleep loss insult of spring daylight savings in the US is observable in twitter activity. *Journal of Big Data*, 8(1):121, 2021

Material from this dissertation has been submitted for publication in PLOS Global Public Health on February 7, 2022

Chapter 2

The sleep loss insult of Spring Daylight Savings in the US is observable in Twitter activity

2.1 Abstract

Sleep loss has been linked to heart disease, diabetes, cancer, and an increase in accidents, all of which are among the leading causes of death in the United States. Population-scale sleep studies have the potential to advance public health by helping to identify at-risk populations, changes in collective sleep patterns, and to inform policy change. Prior research suggests other kinds of health indicators such as depression and obesity can be estimated using social media activity. However, the inability to effectively measure collective sleep with publicly available data has limited large-scale academic studies. Here, we investigate the passive estimation of sleep loss through a

proxy analysis of Twitter activity profiles. We use "Spring Forward" events, which occur at the beginning of Daylight Savings Time in the United States, as a natural experimental condition to estimate spatial differences in sleep loss across the United States. On average, peak Twitter activity occurs 15 to 30 minutes later on the Sunday following Spring Forward. By Monday morning however, activity curves are realigned with the week before, suggesting that the window of sleep opportunity is compressed in Twitter data, revealing Spring Forward behavioral change.

2.2 Introduction

The American Academy of Sleep Medicine recommends adults sleep 7 or more hours per night [39]. However, studies show only 2/3 of adults sleep for this length of time consistently. In 2014, the Centers for Disease Control and Prevention's (CDC's) Behavioral Risk Factor Surveillance System suggested that between 28% and 44% of the adult population of each state received less than the recommended 7 hours of sleep [14]. Despite the scientific consensus that adequate sleep is essential to health, many adults are sleeping less than 7 hours a night on average—a state referred to as *short sleep*. Results from the most recent National Health Interview Survey determined that since 1985, the age-adjusted average sleep duration has decreased, and the percentage of adults who experience short sleep, on average, rose by 31% [40].

Because adequate sleep is necessary for optimal cognition, short sleep is adverse to productivity and learning, and reduces the human capacity to make effort- related choices such as whether to take precautionary safety measures [41–43]. Short sleep's impact on human cognition is harmful in the workplace, and poses a pronounced and distinct threat to public safety when operating a vehicle [44–47]. Short sleep is linked to increased risk of serious health conditions, including heart disease, obesity, diabetes, arthritis, depression, strokes, hypertension, and cancer [13–15], and a recent study found that disrupted sleep is also associated with DNA damage [48]. The link between sleep loss and cancer is so strong that the World Health Organization has classified night shift work as “probably carcinogenic to humans” [16]. Socio-economic status is positively correlated with quality of sleep [49–52]. Due to such detrimental effects, and high prevalence among the population, insufficient sleep accounts for between $280 and over $400 billion lost in the United States every year [53].

Accurately measuring short sleep in a large population is difficult, and there is often a trade-off between accuracy and the size of the study. Polysomnography—considered the most accurate way to measure sleep—can only measure an individual’s sleep patterns in a controlled laboratory setting [54, 55]. Large studies have relied on participants recording their own sleep, but suffer from reporting bias [14, 56, 57].

Wearable technology can measure short sleep at the population scale, and has the potential to measure short sleep accurately enough to study its association with adverse health risks [41,54,58,59]. One recent large sleep study enrolled 31,000 participants and used sleep data from wearable devices along with participant’s interactions with a web based search engine to compare sleep loss and performance [41]. The authors [41] showed that measurements of cognitive performance (including keystroke and click latency) vary over time, follow a circadian rhythm, and are related to the duration of participant’s sleep, results that closely mirrored those from laboratory settings and validated their methodology. Another study using wearables was able to analyze nine metrics of sleep, including social jetlag, duration, and variability for 69,650 individuals [59]. The authors' analysis of these metrics found gendered differences in sleep behaviors across the cohort [59].

While promising in the long run, present studies that use wearable devices have limitations. To infer from wearables that individuals are sleeping, data must first go through a pipeline of preprocessing, feature extraction and classification. The pipeline for processing sleep data is typically proprietary and dependent on the specific wearable used, and changes to how data is processed can impact results [60]. Moreover, validation studies have yet to explore the effectiveness of these devices across genders, ages, culture, and health [60].

Social media may be an alternative way to measure sleep disturbances in a large population, for example by studying the link between screen time and sleep [18, 61]. Researchers have found that Tweeting behavior can reveal "sleep-wake" behavior for individuals as well as cities [62, 63]. In particular, the correlation between sustained low activity on Twitter and sleep time as measured by conventional surveys has been validated against data collected from the CDC on sleep deprivation [18]. The relationship between time of onset of Twitter activity and wake time has been used to explore and demonstrate social jetlag - the discrepancy between weekend and weekday sleep behavior [18, 64]. Other work has shown evidence of an increase in a user's smart phone screen time as being associated with an increase in short sleep [61]. Other mental and physical characteristics have been measured from sociotechnical systems. Several instruments developed by members of our research group including the Hedonometer [11], which measures population sentiment through tweets, and the Lexicocalorimeter [12], which measures caloric balance at the state level, have demonstrated an ability to infer population-scale health metrics from Twitter data. Circadian rhythms in mood and cognitive processes have also been inferred from tweets [65,66]. Twitter data has also been used to identify users who experience sleep deprivation and study the ways their social media interactions differ from others [67].

In urban, industrialized societies where social timing is synced to clock time, Daylight Savings- a biannual sudden upset to clock time- creates behavioral stability across seasons [68, 69]. The onset of DST, Spring Forward, is associated with a one hour sleep disruption due to the disconnect between the "human clock" and the mechanical clock [70]. Past work has used Daylight Savings as a natural experiment to show that a one hour collective sleep loss event has large and quantifiable effects on health, safety, and the economy [19–22], with two striking findings being a one day increase in heart attacks by 24% and a loss of $31 billion on the NYSE, AMEX, and NASDAQ exchanges in the United States [19, 71].

We hypothesize here that sleep loss is measurable in behavioral patterns on Twitter, and changes in population-scale sleep patterns due to Spring Forward can be observed through changes in these behavioral patterns. In what follows we describe the process by which we used the local time of tweet posting to explore patterns in posting frequency relative to time of day, and how these patterns were affected by the clock shift known as Spring Forward. The data is described in detail, followed by the specific methodologies employed to analyze the patterns in the frequency of posting. Then, we visualize and describe the results before concluding with a discussion of limitation and implications.

2.3 Materials and methods

2.3.1 Data

We collected a 10% random sample of all public tweets—offered by Twitter's Decahose API—for Sundays and Mondays in the four weeks leading up to, the week of, and the four weeks following Spring Forward events during the years 2011-2014. Spring Forward is defined as the instantaneous clock adjustment from 2 a.m. to 3 a.m. on the second Sunday of March each year. We included tweets in the study if the user who created the tweet reported living in the U.S. in their bio, or if the tweet was geo-tagged to a GPS coordinate within the U.S. [72]. With these conditions, we ended up selecting approximately 7% of the messages in the Decahose random sample for analysis [73]. The sample was composed of 13.1 million tweets.

Twitter provided the time-zone from which each message was posted during the period from 2011 to 2014 (for privacy purposes, Twitter discontinued publication of time zone information in 2015). We used the time-zone to determine the local time of posting for each tweet. Tweets for which the time-zone was incompatible with the assigned location were discarded. This process enabled us to analyze the method on data for which the local time of posting is known. We binned tweets by 15 minute increments according to the local time of day they were posted.

2.3.2 Experimental setup

The Spring Forward event of Daylight Savings was used as a natural experiment in which the control is behavior prior to the event, and the experiment is behavior directly after the clock change and known sleep loss event. Change in Twitter posting behavior was observed in this experiment. To estimate behavioral change associated with Daylight Savings, we partitioned tweets into various groups, primarily a "Before Spring Forward" (BSF) group and a "Spring Forward" (SF) group. To establish a convenient 'control' pattern of behavior, all tweets posted on any of the four Sundays before the Spring Forward event were classified as "Before Spring Forward" tweets. We classified the 'experimental' set of tweets posted on the Sunday coincident with the Spring Forward event as "Spring Forward". The above classification created, for every year, a 4:1 matching of before to week of Spring Forward activity. We analyzed tweets posted 1-4 weeks following Spring Forward separately to quantify relaxation to the original behavior.

2.3.3 Analysis

We binned tweets by time in 15 minute intervals starting at the top of the hour, and normalized their frequencies by dividing by the total number of tweets posted on the corresponding day. In this way, we establish a discrete description of the posting volume over the course of a typical 24-hour period.

We averaged the Before Spring Forward tweets over the four Sundays, and the four years as follows:

$$T_{\mathrm{BSF}}(k) = (4 \times 4)^{-1} \sum_{Y=2011}^{2014} \sum_{\mathrm{S}=1}^{4} \frac{C_{YS}(k)}{C_{YS}},$$

where $C_{YS}(k)$ is the number of tweets in the k^{th} 15 minute interval of the S^{th} Sunday of year Y, C_{YS} is the total number of tweets posted on that Sunday and year, and

$T_{\mathrm{BSF}}(k)$ is the average fraction of tweets posted in the k^{th} 15 minute interval of a Sunday prior to Spring Forward,

We also normalized the Spring Forward tweets against daily activity:

$$T_{\mathrm{SF}}(k) = (4)^{-1} \sum_{Y=2011}^{2014} \frac{C_Y(k)}{C_Y}.$$

These averages enabled us to aggregate more data, building a more reliable pattern of daily activity, and decrease the susceptibility to daily variation. To reduce noise that could depend on our choice of bin size and spatial scale, we smoothed normalized tweet activity using Gaussian Process Regression (GPR) [74, 75]. We fit a GPR with a squared exponential kernel and characteristic length scale of 150 minutes (a total of 10 bins of size 15-minutes) to normalized tweets. We chose a characteristic length of 150 minutes for consistency with previous work [18]. Tikhonov regularization with an α penalty tuned manually to 0.1 was included when finding weights ω_k to prevent overfitting [75]. GPR yielded a smooth **behavioral curve**, $B(t)$, of the functional form:

$$B(t) = \sum_{k=1}^{96} \omega_k \exp\left[-\frac{1}{2} k \left(\frac{t}{150}, \frac{t_k}{150}\right)^2\right],$$

where ω_k is a weight determined by the regression process, k is the squared-exponential kernel (commonly called a radial basis), t is the time in minutes since midnight (00:00), and t_k is the k^{th} 15 minute interval of the day, i.e. t_5 corresponds to 75 minutes past midnight, or 1:15 a.m. The sum to 96 refers to the number of 15 minute intervals in a single 24 hour period.

We generated behavioral curves $B(t)$ for the BSF and SF groups by state, and for the U.S. in aggregate. To estimate behavioral change induced by a Spring Forward event, we calculate two quantities from the behavioral curves: (i) the time of peak activity and (ii) the time of the inflection point between the peak and trough. The inflection point is referred to as a 'twinflection' point, and represents a point of diminishing losses in Twitter activity for the night. Peak shift is defined as:

$$\arg\max_{t} \{B_{\mathrm{SF}}(t)\} - \arg\max_{t} \{B_{\mathrm{BSF}}(t)\}$$

and twinflection shift is defined as:

$$\arg\min_{t\in N} \{B'_{\mathrm{SF}}(t)\} - \arg\min_{t\in N} \{B'_{\mathrm{BSF}}(t)\},$$

where $N = \{t : \arg\max_t B(t) < t < \arg\min_t B(t)\}$. We were able to reliably measure peak activity and twinflection because behavioral curves exhibited a consistent diurnal wave structure: a rise in the evening corresponding to peak Twitter posting activity, followed by a trough during typical sleeping hours, and a plateau throughout the day. Contraction of the trough associated with sleeping hours is considered to be reflective of lost sleep opportunity, and may indicate sleep loss itself.

We measured the loss of sleep opportunity by calculating the peak and twinflection times for the four weeks Before Spring Forward and the week of Spring Forward itself. We then characterize differences between the BSF and SF measures for each state, and for the total U.S., as a proxy for sleep loss.

2.4 Results

Our overall finding is that peak Twitter activity occurs 15-30 minutes later on the Sunday evening immediately following Spring Forward for most states, with this shift varying among states. By Monday morning, activity is back to normal, suggesting that the window of sleep opportunity is visibly compressed in Twitter behavior.

In Fig 2.1, we plot $B(t)$ for the subset of posts containing the words ‘breakfast’, ‘lunch’, and ‘dinner’ for the period beginning 6 a.m. on Sunday and ending 9 p.m. on Monday, both before (solid) and the weeks of (dashed) Spring Forward events. These curves were constructed for states observing Eastern Time (top row) and Pacific Time (bottom row). These regions were chosen as they are the zones with the greatest spatial difference among zones with significant data density. Observing a shift in behavior for each assures us that these shifts are not limited to a particular geographic region of the country.

Meal-related language reveals a daily pattern of behavior in which peak volume occurs around the time that meal typically takes place. On an average Sunday, breakfast is most mentioned at 10:30 a.m., lunch at 1:15 p.m., and dinner at 6:45 p.m. in Eastern Time Zone states (see Fig 2.1). On the average Monday, breakfast mentions peak at 10:45 a.m., lunch peaks at 1:30 p.m., and dinner at 7:15 p.m. Breakfast and Lunch are mentioned more often on Sunday than on Monday.

There is essentially no discussion of meals during the period from 2 a.m.-6 a.m. These plots also exhibit a small forward shift in time following Spring Forward, suggesting that each meal was tweeted about, and probably eaten, later in the day on Sunday. The effect is greater on the East Coast, and disappears on both coasts by

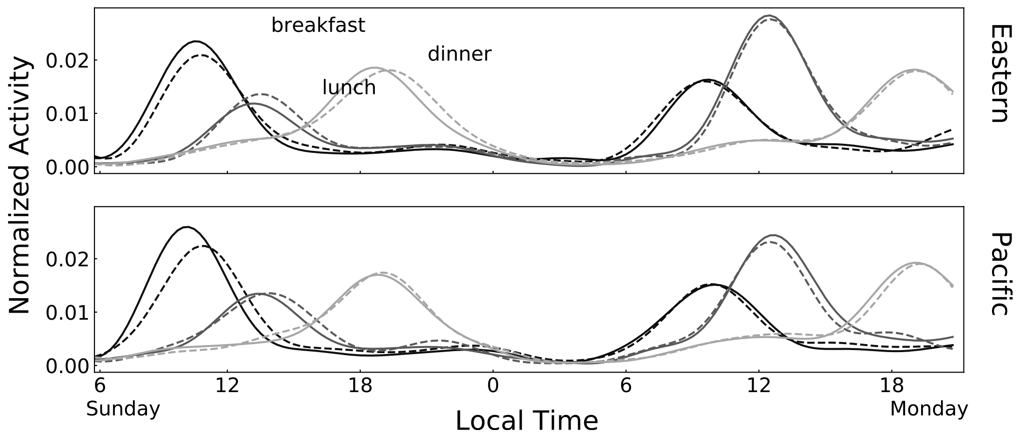

Figure 2.1: ***Diurnal collective attention to meals quantified by normalized usage of the words 'breakfast', 'lunch', and 'dinner' for states observing Eastern Time (top) and Pacific Time (bottom), for the weeks before (solid) and of (dashed) Spring Forward.*** *The x-axis represents the interval between 6 a.m. Sunday and 9 p.m. Monday local time. Counts for tweets containing each individual word were tallied in 15 minute increments, normalized by the total number of tweets mentioning that word, and smoothed using Gaussian Process Regression to create a "Normalized Activity" curve. Each day has a clear pattern for frequency of meal name appearance in tweets, with the peak for breakfast, lunch, and dinner occurring in the respective order of the meals themselves. For each of the meals, we observe a slight forward shift in the peak following Spring Forward, suggesting that meals are taking place later than usual on the corresponding Sunday. By Monday, the peak for each meal name appears to be aligned with the week before, with the exception of 'dinner' on the west coast, which is still a bit later.*

Monday.

Broadening from messages mentioning specific meals to all messages, daily activity plots of B_{BSF} and B_{SF} reveal a regular diurnal pattern of behavior that is consistently shifted forward in time the evening following Spring Forward events. Fig 2.2 shows this shift for the year 2013, but the results were similar for other years.

Panel (a) suggests overall activity across the U.S. peaks around 9 p.m. on Sundays before Spring Forward (red circles), and experiences a minimum around 5am. The peak shifts approximately 45 minutes later on the Sunday of Spring Forward (blue

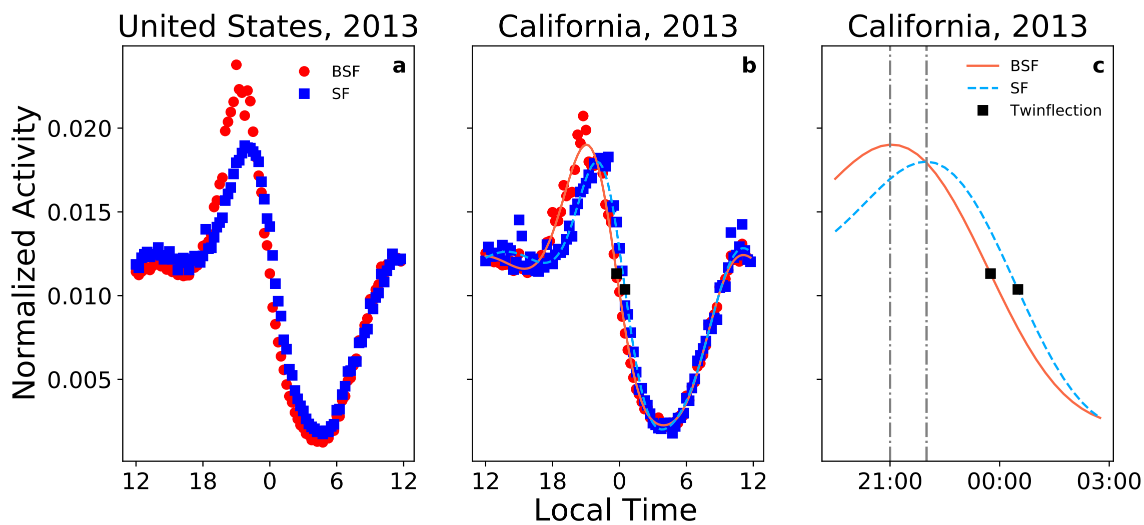

Figure 2.2: ***Twitter activity behavioral curves*** $B(t)$***.*** ***(a)*** *Normalized count of tweets posted from a location within the United States between 12 p.m. Sunday and 12 p.m. Monday before (red) and the week of (blue) the 2013 Spring Forward Event. The time recorded for the tweet is that local to the author. Though the pattern of behavior is preserved following Daylight Savings, peak activity is translated forward in time.* ***(b)*** *The same plot, with location of tweet origin restricted to the state of California. California is the state for which we have the most data, and therefore the most representative behavior profile after smoothing with Gaussian Process Regression (lines). We note that Fig 2.5 shows behavioral curves for all states.* ***(c)*** *The smoothed behavioral pattern for California during the hours of 9 p.m. to 3 a.m. Pacific Time. Activity peaks are denoted by vertical dashed lines, and twinflection points are marked by squares. To estimate the behavioral shift in time, we compute the distance along the temporal axis between these pairs of lines/points. California's BSF peak is one hour earlier than the SF peak.*

squares) before synchronizing again by early morning Monday. In panel (b) California is used as an illustrative example of these patterns existing at the state level, and the smooth behavioral pattern constructed using Gaussian Process Regression. The pattern is similar to that observed for the entire country, with the exception of a slightly reduced amplitude. Twinflection points are illustrated by black squares in panels (b) and (c).

Fig 2.2 demonstrates evidence that there is a shift in the peak time spent interacting with Twitter on Sunday evening following Spring Forward, relative to prior Sundays. Given the absence of a corresponding delay in interaction Monday morning, we infer a decrease in sleep opportunity experienced on Sunday night.

To explore the spatial distribution of the behavioral changes induced by Spring Forward, in Fig. 2.3 we map the time of peak Twitter activity on Sunday night for each state before (top) and the week of (bottom) Spring Forward, averaged across the years 2011-2014. On the Sundays leading up to Spring Forward (top), peak twitter activity occurs near either 10 p.m. for states on the East Coast, or 9:15 p.m., for most of the other states. The week of Spring Forward, nearly all states exhibit peak activity later in the night.

Looking at Texas as an individual example, before Spring Forward we see peak activity around 9:15 p.m. local time, and the week of Spring Forward it occurs at 10:15 p.m. local time. While Texas is one of the latest peaks observed on the evening following Spring Forward, several other states are up late as well including Oklahoma, Georgia, and Mississippi each peaking around 10:15 p.m.

In the Supplemental Information, we show maps estimating the time of peak activity for each of the individual 9 weeks centered on Spring Forward (see Supplemenatry Fig S1 online). There is some week-to-week variation, most notably in the second week prior to Spring Forward, which was the night of the Academy Awards for three of the four years. By four weeks after Spring Forward, the peak activity map has relaxed to roughly the same pattern as BSF.

The magnitude of the forward shift in behavior illustrated in Fig 2.3 is considered a proxy for the loss of sleep opportunity on the Sunday night following Spring Forward.

We used two distinct methods to estimate this magnitude, namely the peak shift

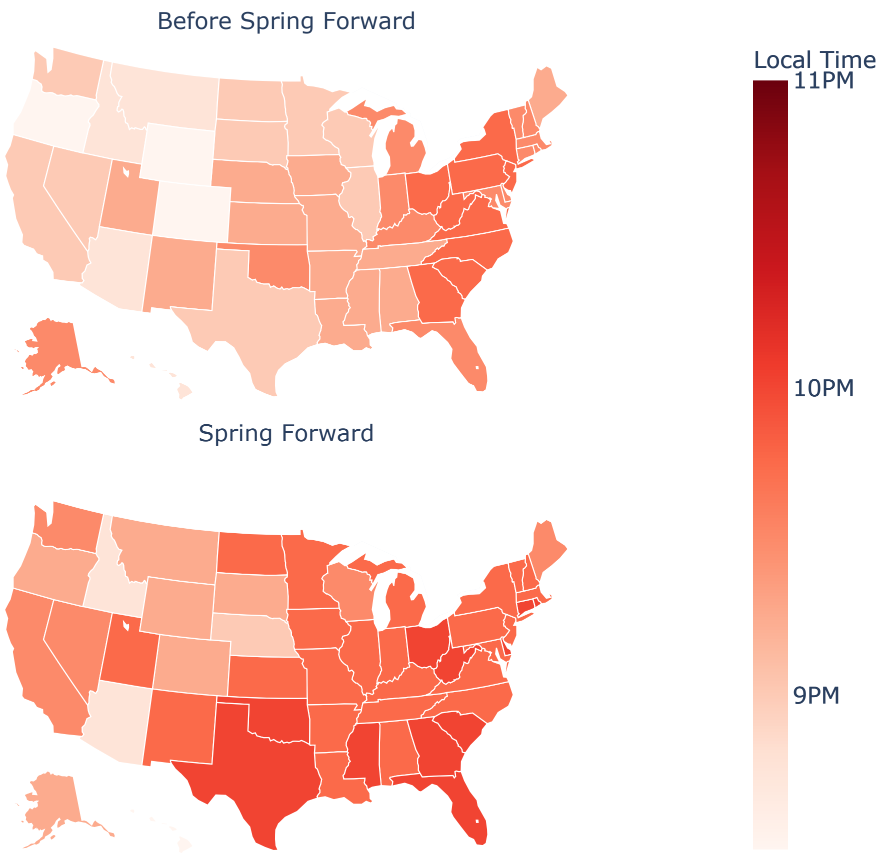

Figure 2.3: ***Time of peak Twitter activity on Sunday night for each state before (top) and the week of (bottom) Spring Forward for the four events observed between 2011 and 2014.*** *Before Spring Forward (BSF), the time of peak activity occurs around 10 p.m. most states in the Eastern Time Zone, and around 9:15- 9:30 p.m. for most of the other states. The week of Spring Forward (SF), peak Twitter activity occurs between 0 and 60 minutes later for each state, with the exception of Alaska, Nebraska, and Hawaii for which the peak occurred earlier. Texas has the latest peak at 10:15 p.m. local time, a shift of 60 minutes forward compared with prior Sundays. We note again that the BSF estimates are based on the aggregation of four Sundays prior to Spring Forward, while the SF estimates are based on the Sunday coincident with Spring Forward, and are therefore estimated using roughly 1/4 of the data. [1]*

and the twinflection shift. A comparison of the spatial estimates made using each method are shown in Fig 2.4.

Panel (a) illustrates the average shift in peak activity observed for 2011-2014 by computing the difference between the pair of maps in Fig 2.3 (bottom minus top). There is clear spatial variation in the shift in time on the night of Spring Forward, while most states exhibit a positive forward shift some exhibit none, and Alaska, Hawaii, and Nebraska show a negative shift. The peak in Twitter behavior for the east and west coasts occurred 0-30 minutes later Sunday night, while it occurred 30-60 minutes later for the central U.S. (Fig 2.4 panel a).

Fig 2.4 panel (b) estimates the change using twinflection, namely the change in concavity of the behavior activity curve from down to up. Every state except Hawaii, Alaska, and Wyoming exhibits a shift forward in time, and with similar spatial regularity. When measured with twinflection shift, Texas and Mississippi are seen to have the greatest temporal shift following Spring Forward. Texans were tweeting 105 minutes later than usual following a Spring Forward event. Most of the east and west coast states were measured as tweeting 15 to 30 minutes later (Fig 2.4 panel b). Both measures agreed on a positive shift for the country as a whole. However, the two measures yielded different results for the magnitude of these shifts, with twinflection shift generally estimating a more positive shift.

Fig 2.4 panels (c) and (d) illustrate the amount of data contributing to calculations for the behavioral curves, and the density of this data with respect to each state's population. Idaho, Alaska, Hawaii, Montana, Wyoming, North Dakota, South Dakota, and Vermont were the states offering the smallest amount of data, and subsequently have the highest potential for a poor behavioral curve model fit. Wyoming was unique in that in 2013 for the 24 hour observation window on the week of Spring Forward there were no tweets meeting inclusion requirements, making conclusions about this state particularly tenuous.

Though the amount of data available for California and Texas is much greater than the other states, when considering their large population size we find their twitter activity per capita to be similar to most other states. Based on our estimate of tweets per capita, we expect behavioral curves for most states to be more or less equally representative of their tweeting populations.

Looking at the diurnal cycle of Twitter activity for each individual state, we see remarkable consistency. Fig. 2.5 shows the 24 hour period spanning noon Sunday to noon Monday local time for the year 2012. Plots for the other 3 years exhibit similar behavior. Before Spring Forward (red), most states show a peak between 9:15 and 10:00 p.m., local time. The week of Spring Forward (blue), nearly all states have a peak after 9:30 p.m. While states differ slightly in the time of peak, and magnitude of shift in the peak, most exhibit a clear positive shift (see Supplementary Fig. S3 online). By Monday morning, nearly all curves have re-aligned. We also consistently observe higher peaks for the BSF curves which we believe to be driven by televised events such as the Oscars. The Sunday of Spring Forward does not have a regularly scheduled popular television event, and as a result the SF curves have lower amplitude.

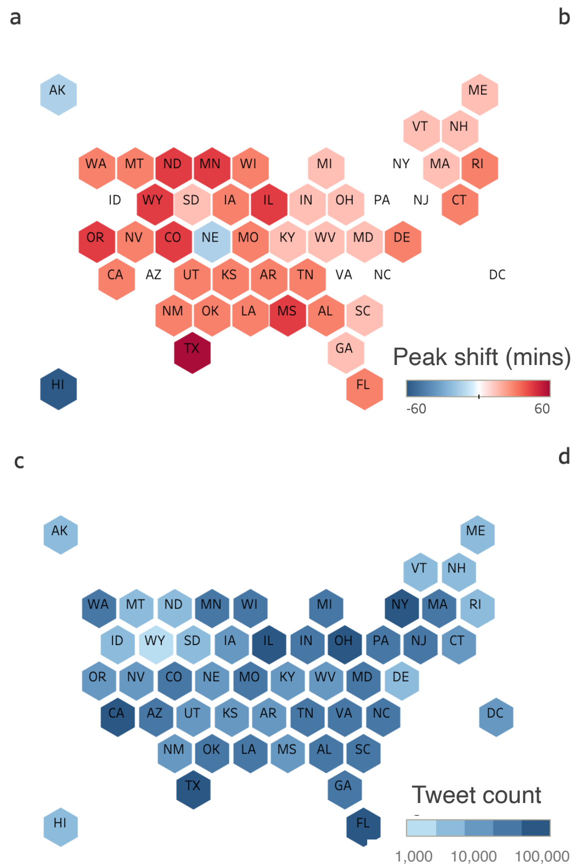

Figure 2.4: ***The magnitude of Twitter behavioral shift following a Spring Forward event, averaged for the four years from 2011 to 2014.*** *(a) Shift measured using behavioral curve peaks, the difference between the pair of maps in Figure 2.3 (bottom minus top). Texas is estimated to have experienced the greatest time shift. The effect of Spring Forward is more pronounced in the South, and center of the country. Alaska, Nebraska, and Hawaii have negative shifts.* ***(b)*** *The same map, but with measurements calculated using twinflection shift instead. The states most affected are Texas and Mississippi, where the shift was 105 and 75 minutes respectively. Hawaii and Alaska are estimated to have negative shifts (15, and 30 minutes respectively). Twinflection shift produces similar spatial results to peak shift, with greater shift estimates.* ***(c)*** *The number of tweets posted from each state in the period after Spring Forward. California and Texas both contributed over 200,000 tweets, while Alaska, Hawaii, Idaho, Wyoming, Montana, North Dakota, South Dakota, Wyoming, Delaware, New Hampshire, Maine and Vermont each produced less than 10,000 tweets.* ***(d)*** *The density of data used to establish the experimental pattern of behavior, as measured by tweets per capita. This measurement reflects the ability of the data to capture the behavior of the tweeting population of each state. While Idaho, Wyoming, Montana, Utah and South Dakota have relatively little data compared to their populations, the remaining states have similar data density, with somewhere between five and eleven tweets per thousand residents, with the exception of the District of Columbia which has 35. Note: both panels (c) and (d) use logarithmically spaced colorbars*

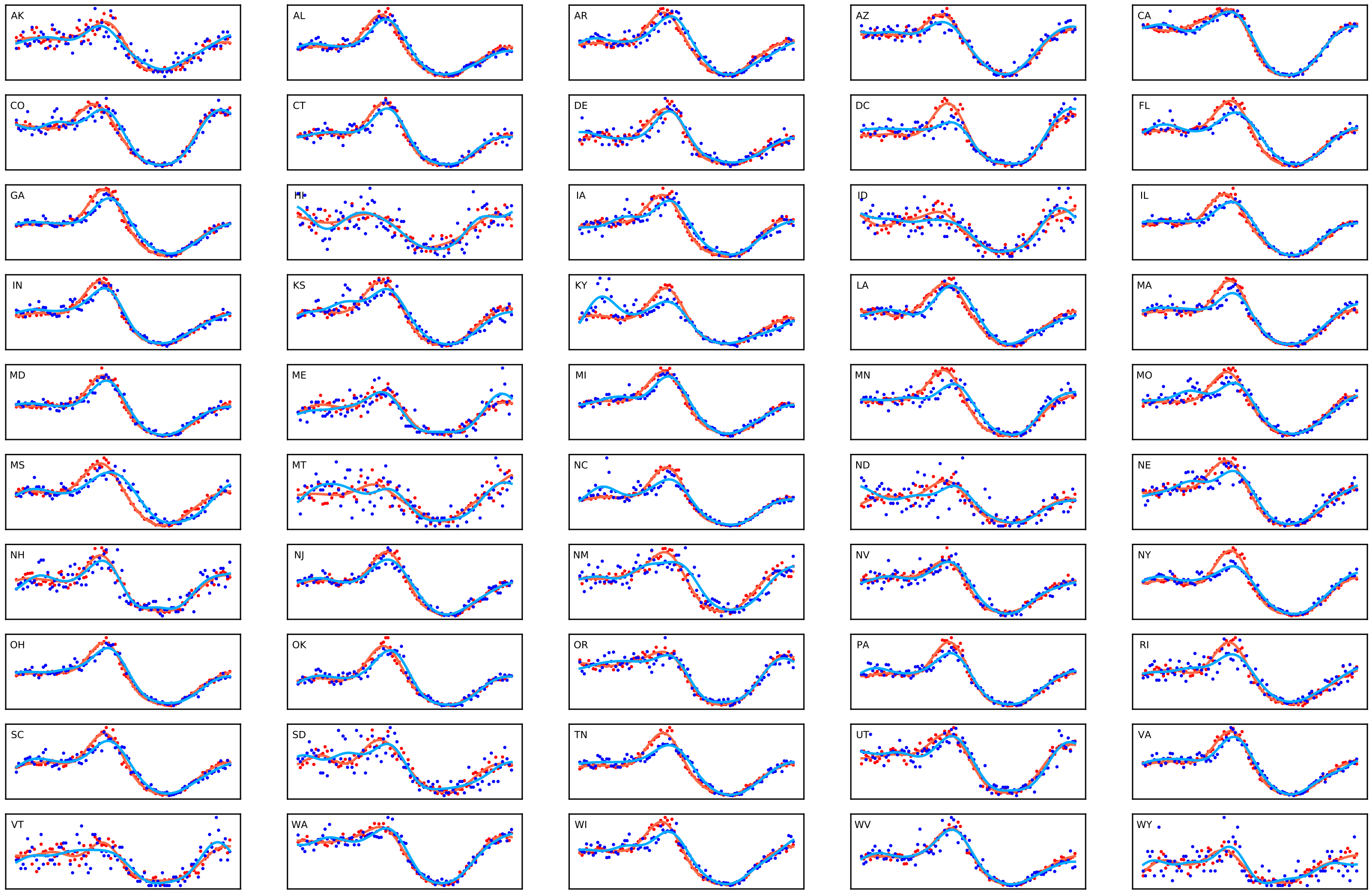

Figure 2.5: ***Normalized Twitter activity between 12 p.m. Sunday and 12 p.m. Monday prior to and following the 2012 Spring Forward event for each state.*** *Red indicates an aggregation of data from the specified period over four weeks before the Spring Forward Event. Blue indicates data from the single 24 hour period after Spring Forward has occurred. Dots are indicative of 'raw' data, while the corresponding curves demonstrate Gaussian smoothing. Texas exhibits the largest change following Spring Forward. Curves for nearly all states have aligned by Monday morning. The BSF peaks are slightly higher than the SF peaks in some states, largely due to televised events Before Spring Forward such as the Oscars. The Sunday of Spring Forward does not have a regularly scheduled popular television event, and as a result the SF curves have lower amplitude.*

Both the peak and twinflection demonstrate that it is possible to observe a measurable decrease in the amount of sleep opportunity people in the United States receive on average due to Spring Forward. They also both demonstrate uneven geographic distribution of the effect of Spring Forward, and therefore the ability to determine geographic disparity in sleep loss.

We also discovered that the Super Bowl occurred exactly 5 weeks prior to Spring Forward in each of the years studied. This annual event watched by over 100 million individuals in the U.S. caused peak Twitter activity to synchronize at roughly the same time nationally, around 9 p.m. Eastern, during the second half of the football game. The map in Fig 2.6 shows the time of peak activity for each state on Super Bowl Sunday, averaged over the years 2011 to 2014. The colormap is the same as the scale used for 2.3, with the additional cooler range brought in capture the time of peak relative to the usual times.

The map bears a remarkable resemblance to the timezone map, demonstrating a synchronization of collective attention across the country. Data from Super Bowl Sunday was not included in the Before Spring Forward data, as it does not accurately reflect the spatial distribution of typical posting behavior on a Sunday evening.

2.5 Discussion

Technically speaking, Spring Forward occurs very early Sunday morning, and the instantaneous clock adjustment from 2 a.m. to 3 a.m. is witnessed by very few waking individuals. In addition, we speculate that the majority of individuals do not set an alarm clock for Sunday morning. As a result, we expect that the hour lost

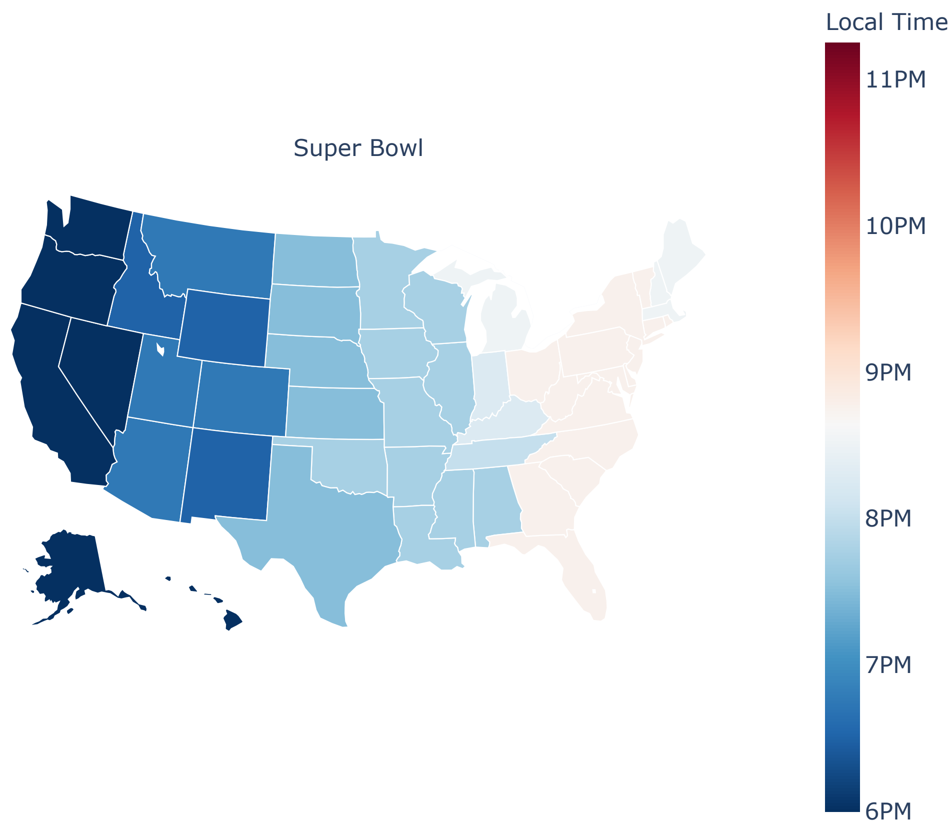

Figure 2.6: ***Peak activity time (local) for Super Bowl Sunday, 5 weeks prior to Spring Forward, averaged over the years 2011 to 2014.*** *Activity exhibits a clear resemblance to the U.S. timezone map, with a peak near 9 p.m. Eastern Time just following the halftime performance. The data suggests a national collective synchronization in attention. Green Bay Packers d. Pittsburgh Steelers (2011), New York Giants d. New England Patriots (2012), Baltimore Ravens d. San Francisco 49ers (2013), and Seattle Seahawks d. Denver Broncos (2014). Performers included The Black Eyed Peas, Usher, and Slash (2011), Madonna, LMFAO, Cirque du Soleil, Nicki Minaj, M.I.A., and Cee Lo Green (2012), Beyoncé, Destiny's Child (2013), and Bruno Mars, Red Hot Chili Peppers (2014). We note that the colormap here the same as the scale used for 2.3, with blue colors included to reflect the relatively early times of the peaks relative to the other weeks.*

to Spring Forward will be felt by our bodies most meaningfully on Monday morning. Indeed, we are likely to experience the Monday morning alarm as occurring an hour early, as Spring Forward shortens the time typically reserved for sleep opportunity Sunday night by one hour.

Considering the correlation between screen time and lack of sleep, the Sunday evening shift, and the corresponding Monday morning re-synchronization, we observe evidence that sleep opportunity is lost in some states on the evening of Spring Forward. By estimating the magnitude and spatial distribution of the shift in Twitter behavioral curves, we have approximated a lower bound on sleep loss at the state level.

Our pair of measurement methodologies have a Pearson correlation coefficient of 0.575, and a Spearman correlation coefficient of 0.467 (See Supplementary Fig S3 online). While they produced slightly different estimates of the magnitude of temporal shift in behavior, the resulting geographic profiles of sleep loss were similar. Both suggest that states along the coast are least affected by Spring Forward, while Texas and the states surrounding it to the North and East are the most affected.

Peak shift suggests the temporal shift in behavior due to Spring Forward generally less than the actual clock shift (1 hour). California, the state for which we have the most data and therefore the most representative behavior profile after smoothing, was found to have a peak shift of 30 minutes.

Considering the clock adjustment of exactly one hour, both measurements are plausibly directly representative of sleep lost, however the differing magnitudes of the measurements indicate that future work should clarify the relationship between these measurements and actual shifts. Twinflection measured similar shifts for most states, but for a few estimated larger effects. While California was measured as having the same 30 minute shift, Texas, the state for which we have the second most data, was estimated by twinflection to be delayed by an additional 45 minutes.

Twinflection measured a small forward shift for the state of Arizona, which does not observe DST. This could indicate that the twinflection method overestimates the behavioral shift. It is also possible that a shift in behavior could occur for residents of Arizona, as a result of their connections to those in neighboring states, and in their former timezone. In example, some residents likely work in bordering states, and are forced to observe DST, and some will likely engage in more online activity and discussion when their peers are present- those peers being initially established by a shared time of activity. This we believe to be an important distinction between Arizona and Hawaii, which also does not observe DST.

Hawaii is measured to have gained sleep opportunity by both accounts. Lacking the observation of DST, neighboring states, and other states in the same timezone, it is plausible that behavior in Hawaii would be unlike any other state, and be more independent of behaviors in other states. However, Hawaii's results should be considered tentative at best, given the sparsity of data available. This sparsity of data and relative independence from other states is shared with Alaska, the other state with a measured sleep opportunity gain by both measures. Caution should likewise be extended to measurements ascribed to South Dakota, North Dakota, Wyoming, Idaho, Montana, Vermont, New Hampshire, Rhode Island, Delaware, and Maine. These states have smaller populations, less population density, and lower volume of tweets. As a result, the behavioral curves associated with these states are less reliable.

Discrepancies in available data were determined to be largely accounted for by differences in population. Thus, we expect results for each state (exclusive of those mentioned earlier) to be comparably reliable in their representation of sleep loss for the state as a whole.

Incremental future work in this area could investigate state specific sleep loss related to Spring Forward events, which would allow further clarification of the relationship between the magnitude of behavioral shifts on Twitter and population sleep loss. Other directions might include looking at other sleep opportunity interruption events such as the end of Daylight Savings in November, where we are ostensibly given an additional hour of sleep opportunity. Our findings suggest that the sleep behavior associated with other annual events including New Year's Eve and Thanksgiving ought to be visible through tweets. This and other works would also benefit from exploration of the relationship between measurements of sleep opportunity as given by social media activity and actual sleep duration. More ambitiously, proxy data such as this could be verified by matching wearable measurements of sleep (e.g. Fitbit) with social media accounts.

Our study suffers from several limitations associated with our data source, we describe a few such examples here. The geographic location users provide in their Twitter bio is static and unlikely to be updated when traveling. As a result, user locations (time zone, state) inferred from this field will not always reflect their precise location. The GPS tagged messages included in our analysis will not suffer from this same uncertainty. Furthermore, the tweeting population of each state is likely to have complicated biases with respect to their representation of the general population [76].

Our dataset likely contains automated activity. Indeed, an entire ecology of algorithmic tweets evolved during the period in which we collected data for this study.

However, we expect the majority of this activity to be scheduled using software that updates local time automatically in response to Daylight Savings. As such, this ‘bot’ type activity should largely serve to reduce our estimate of the time shift exhibited by humans.

As we showed for the Super Bowl, live televised events (e.g. sports, awards shows) have the potential to be a forcing mechanism to synchronize our collective attention throughout the week, and especially on Sunday evenings. Indeed, many individuals take to Twitter as a second screen during such events to interact with other viewers. In addition, streaming services such as Netflix and HBO often release new episodes of popular shows on Sunday night to align with peak consumption opportunity. These cultural attractions exert a temporal organizing influence on our leisure behavior, and the Spring Forward disturbance translates this synchronization forward in time.

It is worth noting that early March is a rather dull time of year for popular professional sports in the United States. While the National Basketball Association and National Hockey League are finishing up their regular seasons, the National Football League is in its off-season and Major League Baseball beginning pre-season exercises. Arguably the most engaging live-televised sporting contests taking place in early March are the NCAA College Basketball Conference Championship games, with March Madness happening weeks after Spring Forward.

In 2014, the Academy Awards were hosted by Ellen DeGeneres on Sunday March 2. Her famous selfie tweet containing many famous actors was posted that evening, a message which held the record for most retweeted status update for several years [77]. The event happened the week before Spring Forward, and led to anomalous behavior compared with all other Sundays we looked at.

Since Spring Forward only occurs once per year, the specific language of the tweets is highly dependent on events occurring on that specific day. The variability in daily events and susceptibility of affect to these daily events makes study of the actual language in the tweets unreliable.

Finally, Twitter (and other social media companies) have access to much higher fidelity information regarding user activity than we have analyzed here. We are not able to analyze consumption activity on the site, e.g. when individual messages are interacted with via views, likes, or clicks. These forms of interaction with the Twitter ecosystem are likely to occur chronologically following the final posting of a message in the evening, and prior to the initial posting of a message in the morning. As a result, we expect our estimate of the sleep opportunity lost due to Spring Forward to be a lower bound.

2.6 Conclusion

Privacy preserving passive measurement of daily behavior has tremendous potential to transform population-scale human activity into public health insight. The present study leverages a natural experiment in sleep loss to identify behavioral adaptation from Twitter data. It demonstrates a proof-of-concept along the path to a far more ambitious goal: construction of an 'Insomniometer' capable of real-time estimation of large-scale sleep duration and quality. Which cities in the U.S. slept well last night? Which states are increasingly suffering from insomnia? Answers to questions like these are not available today, but could lead to better public health surveillance in the near future. For example, communities exhibiting disrupted sleep in a collective pattern may be in the early stages of the outbreak of the flu or some other virus. Current methodologies for answering these questions are not scalable, but social media, mobile devices, and wearable fitness trackers offer a new opportunity for improved monitoring of public health.

Chapter 3

Spatial changes in park visitation at the onset of the pandemic

3.1 abstract

The COVID-19 pandemic disrupted the mobility patterns of a majority of Americans beginning in March 2020. Despite the beneficial, socially distanced activity offered by outdoor recreation, confusing and contradictory public health messaging complicated access to natural spaces.

Working with a dataset comprising the locations of roughly 50 million distinct mobile devices in 2019 and 2020, we analyze weekly visitation patterns for 8,135 parks across the United States.

Using Bayesian inference, we identify regions that experienced a substantial change in visitation in the first few weeks of the pandemic.

We find that regions that did not exhibit a change were likely to have smaller populations, and to have voted more republican than democrat in the 2020 elections.

Our study contributes to a growing body of literature using passive observations to explore who benefits from access to nature.

3.2 Introduction

Parks are important public infrastructure that provide a venue for interaction with nature, socialization, and exercise. Park access and use has been found to offer both mental and physical health benefits [78–83]. Among the many benefits of exposure to nature are faster healing, decreased stress and increased ability to manage life's challenges [84, 85]. During the COVID–19 pandemic, access to parks may have been important for mitigating and managing the secondary impacts of the virus. Recent publications indicate that access to parks during the pandemic is important for a variety of reasons including providing a venue for exercise, increasing happiness, and improving social cohesion [86–89].

While park visitation may have provided significant support to personal and public health at the time, it is unclear whether park visitation changed, to what extent, and for whom in the United States. In March of 2020, stay at home orders were issued in most states, and many non-essential workplaces and public spaces were closed. Following these events, overall mobility decreased dramatically for most Americans, reaching a maximum reduction by 34 to 69% depending on the state [90, 91]. While Americans were visiting fewer locations in general, some research suggests that park visitation may not have been subject to this decline. An early study of parks on the West Coast determined changes in visitation at the onset of the pandemic to be primarily motivated by seasonal change, while a study of parks in New Jersey found that early pandemic visitation was higher than the baseline [25, 26]. Together these results indicate that visits to parks may have differed from other points of interest at the onset of the pandemic.

Preliminary examination of trends suggest that changes in park visitation were not universal. In the United States, partisanship, even at the regional level, is associated with behavioral differences. Researchers have found that Thanksgiving dinners were 30 to 50 minutes shorter when the guests and hosts resided in voting precincts that had been in opposition in 2016 [92]. Mobility studies of Americans during the pandemic have found differences along partisan lines as well. The American political system is largely dominated by two political parties: Democrats, and Republicans. This divide in political ideology has been found to be indicative of differing identities and behaviors. This is particularly true of COVID-19 policy response and preferences [93]. Republicans have been found to have lower vaccination rates, have a smaller decrease in mobility during the pandemic, and to be less compliant with non-pharmaceutical interventions [24, 90, 94–97]. Counties with more Republicans also had less severe mobility restrictions, and were less responsive to their governor's recommendation to stay home [90, 98]. Given these partisan differences in general mobility, we seek to determine whether changes in local park visitation at the onset of the pandemic also differed by partisanship, or whether park visitation uniquely transcended these differences.

Studies of park usage in March and April of 2020 have thus far relied on survey data, or have been geographically limited, and neglected to establish a baseline of seasonality of park usage [25, 26, 99, 100]. Here we utilize mobile device data from across the United States to explore abrupt non-seasonal changes in park visitation at the regional level. We use data from 2019 to discern seasonal visitation patterns, and employ a change-point detection algorithm to diagnose sudden changes in behavior at the onset of the pandemic. By classifying regions by whether or not an abrupt change in park visitation took place, we are able to discern whether or not these abrupt changes occurred along partisan lines. We conduct further comparisons across population, income, and share of employment by industry to provide insight into other factors that may have influenced whether or not an abrupt change occurred. In Section 3.3 we introduce the data used to classify regions, and make these comparisons. Section 3.4 then gives a detailed explanation of the data aggregation for each region, and the classification procedure and methods of comparison applied to the aggregated data. The results of the comparisons are described in Section 3.5, and are discussed in Section 3.6.

3.3 Data

To determine whether a partisan effect is observed in park visitation we used park visitation data from across the United States, and voting share data from the 2020 Presidential Election. Differences in regions with and without abrupt changes in visitation were further analyzed using population estimates, and income and employment share data from the US Census and the Bureau of Economic Affairs. Details for these data sources are provided below.

3.3.1 Park Visitation Data

Our park visitation dataset was acquired from UberMedia (now part of Near), and consists of daily visitation counts for non-commercial parks for each day of 2019 and 2020. There are 8,135 parks in the data set, including municipal, neighborhood, and city parks. National and State Parks were specifically excluded as predominantly travel destinations. Parks are located in each of the 50 states, and Washington DC. A total of 1,033 counties, roughly a third of all counties, contain at least one park from our dataset.

Daily visitation counts were determined using location data from mobile devices. Each unique device appearing within a park's bounds on a single day was counted as a visit. A device's location was reported when an individual used one of over 400 apps utilizing a GPS Software Development Kit (SDK) in partnership with UberMedia (90% of data by volume), or when a user interacted with an advertisement through real time bidding on one of over 250,000 apps (10% of data by volume). GPS location and an accompanying timestamp were determined from the device's operating system.

The number of devices reporting activity in at least one location in the US on a given day is referred to as the Daily Active Users (DAUs). This number refers to all locations, not simply parks. In 2019 and 2020 the monthly DAUs varied between 38 and 60 million, and represented roughly 10% of the adult population in the United States.

In mid December 2019 the set of SDK's in partnership with UberMedia was updated. This change in data collection corresponded to a large increase in observations throughout the US, and was not spatially uniform. Thus, while the raw 2019 and

2020 park visitation data are not directly comparable, we analyze their relationship where possible.

3.3.2 Voting and Economic Data

Voting data at the state and county levels from the 2020 election was retrieved from MIT Election Data and Science Lab, and is available at https://electionlab.mit.edu/data.

The BEA publishes data on employment by industry (using the North American Industry Classification System (NAICS)) for each county in table "CAEMP25N", which can be found at https://apps.bea.gov/regional/downloadzip.cfm . In this table “Farming" and “Forestry" are considered as separate, though they appear as one sector in the NAICS classification. For this study they were considered separately, as they appear in the table. County level population, income, and economic data from the 2019 American Community Survey were obtained from US census API.

3.4 Methods

3.4.1 Aggregation

Daily visits to a park were defined as the unique number of mobile devices reporting GPS coordinates found inside the park polygon bound on a day. This daily visit count was then normalized by the average Daily Active Users for the month in which it was found, approximating the percentage of devices observed in parks relative to all observed devices. The normalized visitation was then summed over each week in order to minimize noise. Weekly visitation was summed for parks contained in a

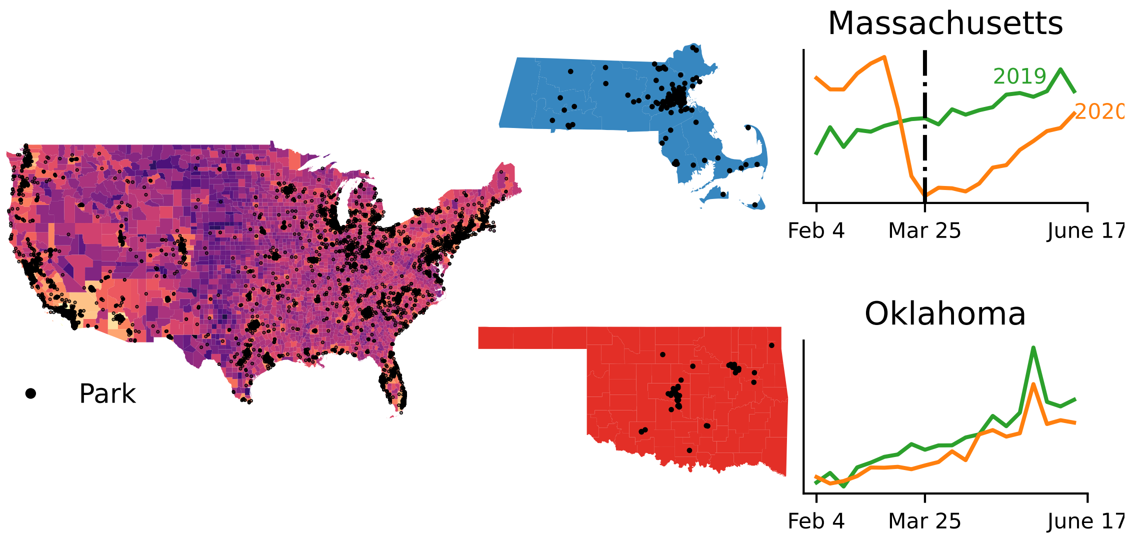

Figure 3.1: ***A heat map of the population of the contiguous United States in log scale overlaid with the locations of the parks used in the study, with each park demarcated with a black point.*** *Observation of this map indicates that the parks in this data set have an urban bias, and that the parks are roughly distributed according to population distribution in the United States. Population heat maps in the log scale are shown for Massachusetts (top) and Oklahoma (bottom), overlaid with park locations in black. The color scale chosen for each state represents the political party receiving the most votes in the 2020 Presidential election (Democrats for Massachusetts, and Republicans in Oklahoma). Normalized weekly park visitation for each state is plotted to the right. Visitation for 2019 is plotted in blue, while visitation for 2020 is plotted in orange. For Massachusetts there is a significant dip in visitation bottoming out the week of March 25, 2020, and the visitation plots for 2019 and 2020 diverge. For Oklahoma visitation does not drop off in March, and does not diverge from 2019 visitation patterns.*

county, or a state, and thus a time series of weekly park visitation between 2019 and 2020 was created for each county and state containing at least one park from our data set.

3.4.2 Change Point Detection

To determine whether a substantive change in visitation is observed in each time series, we use the Bayesian Estimator of Abrupt Change, Seasonality, and Trend (BEAST) [27]. This method decomposes a time series into a seasonal (harmonic) component, and trend (linear) component, and uses Bayesian Inference to fit a model which estimates the location of change points in either of the components. BEAST was chosen because the underlying model acknowledges the seasonal nature of most park visitation time series (more visits in summer). By specifying a 52 week season length, we were able to train the model to the annual cycle shape of the data.

Parametric methods applied without the seasonal decomposition are susceptible to under estimating

change points in these particular time series because of the combination of seasonality and the proximity in the series of the data collection change in December 2019 to the onset of the pandemic in March 2020. The initial event represents a sharp increase in visitation volume (roughly 150 pct), while the second appears, for most regions, as a sharp decline. When fit with a single model, these two features appear together as a change in variance, and a parametric model can be nicely fit using a single change point in December 2019.

By decomposing the time series and forcing a decoupling of the two events by specification of seasonal, length we make each event visible as a unique discontinuity in the linear component.

The December 2019 discontinuity could then be accommodated with a trend change point, which incorporates a discontinuity into the linear component. In this way the model was fit while accounting for seasonality, and the abrupt change in data volume.

Allowing a trend change point to be used as described above, the model was effectively limited to selecting a single trend change point, which enabled it to identify the most likely change point in the data. It is possible for the algorithm to detect no change point, reducing concern that one would be identified artificially.

Regions which had a change point occurring in between mid March and mid April 2020 were considered to have had an abrupt change in park visitation coinciding with the onset of the pandemic and social distancing measures. If a region was found not to have had a change point in this window, it can be assumed that either no change point was found in the time series, or any change occurring in the specified window was not as significant as a change at another time.

Changes induced by seasonality are in most cases more gradual than those that occur in the window of interest, and these changes are accounted for by the harmonic component of the model. The harmonic component is fit using both 2019 and 2020 data, which informs the model of the expected seasonal shape. Since these changes are accounted for in the model fitting, it is unlikely that change points identified in the window of interest are due only to seasonal variation. Because the length of the time series only included two seasons (park visitation demonstrates a yearly cycle), it was not pertinent to search for changes in the seasonal structure.

BEAST is less effective in identifying change points in time series with high variance. The recorded park visits in some of the counties were low enough that the behavior of only a few individuals could have large impacts on the time series itself. To ensure that BEAST was only considering counties for which there was enough data we used a mean normalized visitation threshold of $10^{-5.5}$ (this corresponds to about 120 visits per week in the month with the least DAUs) in 2020. A total of 322 counties did not meet this criteria and were excluded from further analysis. The remaining 711 counties that contain parks in our dataset met this criteria. The counties included in the analysis are roughly 21% of all the counties in the United States, and span all of the states. Details on the selection of the visitation threshold can be found in the Supplementary Materials (See Fig 7.1).

3.4.3 Comparison

Regions were binned according to whether a change point in mid March 2020 was identified or not, and comparisons of the populations of the regions in each category were made. Using data from the 2020 election, states and counties were assigned a percent of the population having voted either Republican (Trump and Pence) or Democrat (Biden and Harris) in the 2020 election. Counties were assigned personal incomes, and population counts using Census estimates from 2019. Voting records and census data were combined to determine the votes cast per capita for each county. Finally, a fraction of employment (employment share) for each industry in the North American Industry Classification System (NAICS) was assigned to each county in the study using data from the BEA. Counties which had no available employment data for an industry were exclcuded from the analysis of that particular industry. Counties with and without detected change points were compared across vote share, population, votes per capita, personal income, and industry employment using Kolomogorov-Smirnov two sample tests. This test was chosen for its ubiquity in the literature and ability to compare distributions with different sample sizes. The means of the distributions are compared using Welch's t-test, which also accommodates different sample sizes.

3.5 Results

3.5.1 Partisanship in Abrupt Changes

With the parameters discussed in the Methods section, BEAST found a change point in the window of interest for 21 states, while the remaining 29 states did not exhibit an abrupt change in visitation. Comparison of the 2020 presidential election results for states where visitation did and did not change abruptly is shown in the top row of Fig 3.2. The distributions across vote share for the two sets of states were not significantly different for either the Democratic or Republican parties (KS statistic=0.2, p=0.63 and KS statistic=0.2, p=0.63 respectively). The distributions for each party are neither similar, nor mirrored. The difference is accounted for by third party votes, most notably Libertarian votes.

Comparison of the distributions across percent voting Libertarian (which accounted for less than 3% of the vote in all states) indicates that the distributions were significantly different (KS statistic=0.44, p=0.015), where Libertarians had greater vote share in states without an abrupt change. Supplementary Fig 7.2 demonstrates the relative proportion of Democrat, Republican, and third party votes for each state and county. The state appearing as an outlier in the distributions, where Democrats had the highest vote share, and which did not have an abrupt change, is Washington DC, which was treated as a state for this study. Exclusion of Washington DC does not change the results.

Partitioning the data by county led to significantly different distributions across vote share (bottom row of Fig 3.2). When the BEAST classification procedure was applied to county level aggregations of visitation data, 123 of the 711 counties had abrupt visitation changes at the onset of the pandemic. The distribution across Democratic vote share for counties with abrupt changes is shifted to the right of the distribution for counties without- indicating that Democrats were more likely to have greater vote share in counties with abrupt changes. Kolomogorov Smirnov 2 sample results confirm that these distributions are significantly different (KS statistic=0.33, p=4.23e-10). Observation of the same distributions across Republican vote share reveals that Republicans were more likely to have greater vote share in counties without abrupt changes. KS 2 sample test results support that these distributions are also significantly different (KS statistic=0.33, p=2.35e-10).

Not only are the distributions significantly different, but across the vote share for each party they are translated across the x=0.5 line (drawn in red). This line represents the dividing point in the majority party support in a county. This reveals that Democrats were not only more likely to have greater vote share in counties with abrupt changes, they were more likely to hold a majority in those counties. Likewise, Republicans were more likely to hold a majority in counties without abrupt changes.

The distributions of the counties across vote share for the Democratic and Republican parties are not mirrored on account of votes going to third parties, meaning that

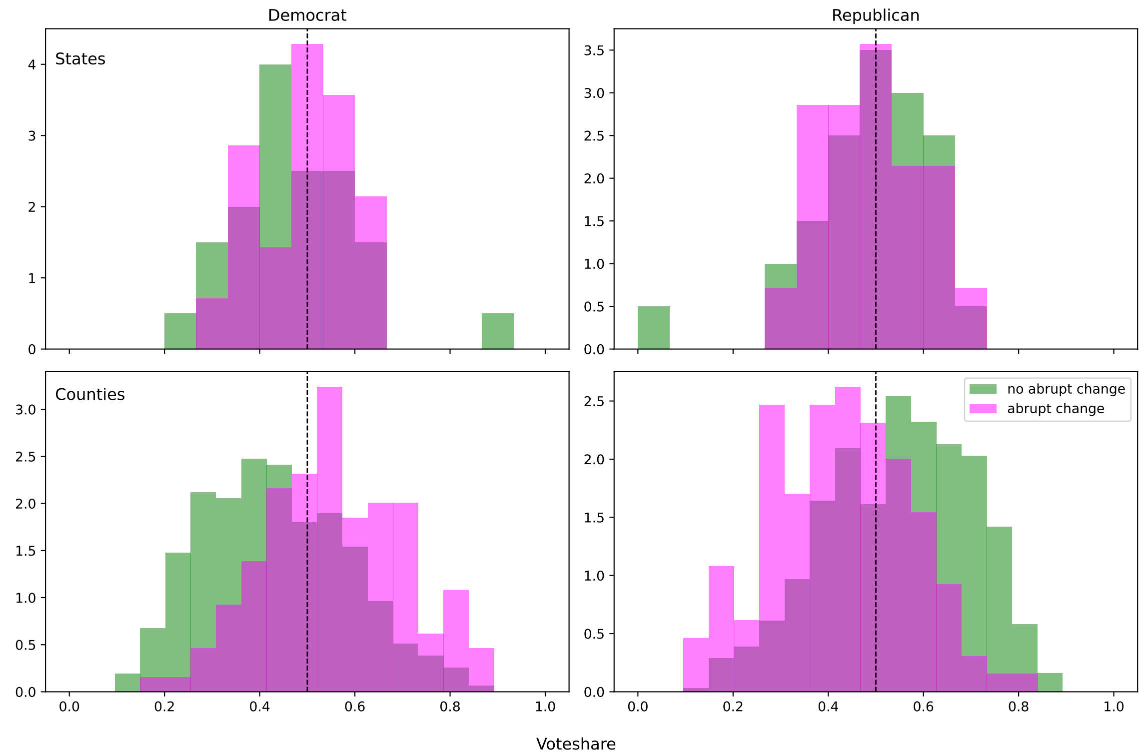

Figure 3.2: ***Top: Distributions of states where a pandemic response change point was detected (pink), and not (green), across proportion of votes cast for the Democratic (left) and Republican (right) parties in the 2020 Presidential Election.*** *The distributions are not significantly different across voting proportion for either party. The apparent outlier in each figure is Washington DC, which did not exhibit a change point, and for which more than 80% of the votes cast were for the Democratic party. Exclusion of DC from the analysis did not change the results.* ***Bottom: Distributions of counties with (pink) and without(green) detected change points across percent voting for the Democratic (left) and Republican (right) parties in the 2020 Presidential Election.*** *Distributions across the percent of votes cast for the Democratic party were determined to be significantly different by the Kolomogorov Smirnov 2 sample test (KS statistic=0.33, p=4.23e-10). The distribution of counties with a change point was shifted to the right of those without a change point, indicating counties with change points had greater proportions of votes for the Democratic party. The bulk of the mass of the two distributions lies on either side of 0.5, meaning that the majority of counties with a change point are majority Democrat counties. The distributions across percent voting for the Republican party are likewise significantly different (KS statistic=0.33, p=2.35e-10), and indicate counties without change points had greater proportions of votes for the Republican party, and were more likely to be a majority Republican county.*

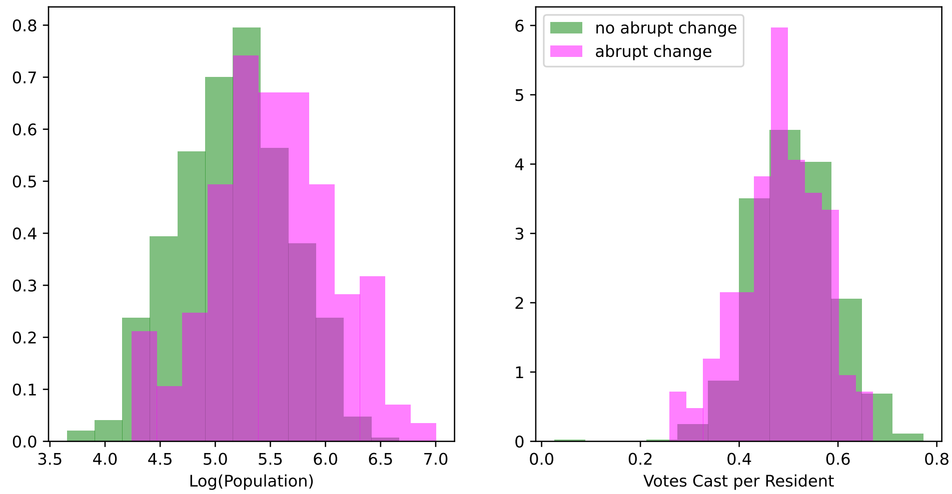

Figure 3.3: ***Distributions of counties with and without change points across the log base 10 of 2019 county population (left) and the votes cast per capita in the 2020 Presidential Election (right).*** *The distributions over population are roughly log normal, and visibly and significantly different (KS statistic=0.28, p=1.08e-07). The mean population of counties with change points was 331,131 people, which is more than twice the mean population of counties without change points, namely 144,544. The counties with the lowest populations were exclusively without change points, while the counties with the greatest populations were exclusively those with change points. The distributions across votes per capita are also visibly and significantly different (KS statistic=0.13, p=0.045) with the counties with change points having fewer votes per capita than counties without.*

"not Democrat" is not the same as “Republican”. Both distributions taken together support that there is a partisan divide between counties with and without abrupt changes in park visitation at the onset of the pandemic. This is further supported by no significant difference found in the distributions of the counties over percent voting Libertarian (KS statistic=0.13, p=0.087).

3.5.2 Population, Income, and Employment

Partitioning the data by county allowed further analysis using population, employment, and income data. Differences in distribution across population size, and votes cast per resident, for the counties with and without abrupt changes, are displayed in Fig 3.3. Counties with an abrupt change had more than twice the mean population of counties exhibiting no change, and fewer votes per resident than counties that did not. The distributions across each of these variables is significantly different (KS statistic=0.28, p=1.08e-07 for log 10 scale population, and KS statistic=0.13, p=0.045 for votes cast per resident).

The incomes of the counties were not significantly different (KS statistic=0.10, p=0.23), as seen in the distributions in Fig 3.4.

Counties were also compared on the basis of percent employment in each of the 20 NAICS sectors. The distributions of the counties with and without change points across percent of employment were significantly different ($p < 0.05$) for 14 of the sectors. This includes both Farming Employment, and Forestry, Fishing and Related Activities, which comprise a single sector in the NCAIS, but are considered separately here. Of the 14 sectors with significantly different distributions, Welch's T-Tests found only 10 had significantly different means. The distributions for these 10 sectors is shown for counties with and without abrupt changes in Figure 3.5. For each sector, the box plot to the left shows the distribution over the fraction of employment for counties with an abrupt change (pink), and the box plot to the right represents the same distribution for counties without an abrupt change (green).

For the 10 sectors with significantly different distributions and means, 5 had higher

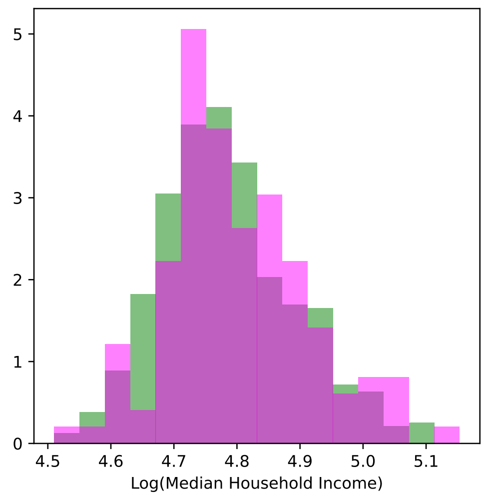

Figure 3.4: ***Distributions of counties with and without change points across the log base 10 of 2019 personal income as reported by the census.*** *The distributions are visually similar, and not significantly different (statistic=0.10, p-value = 0.23). There is not a statistically significant difference between the incomes of counties where abrupt park visitation changes occurred and those where it did not.*

mean employment share in counties with abrupt changes: Information, Finance and insurance, Professional, scientific, and technical services, Educational services, and Health care and social assistance. These sectors are primarily comprised of white collar workers, and with the exception of Health care and social assistance, require less onsite work. Farm employment, Mining, quarrying, and oil and gas extraction, Construction, Manufacturing, and Retail trade all had higher mean employment share in counties where abrupt changes in park visitation did not occur.

3.6 Discussion

At the state level, there was no significant difference in the partisanship of regions where an abrupt change in park visitation took place, and those where it had not. There was a significant difference in the vote share of Libertarians, with Libertarians having smaller vote share in states with an abrupt change. However, Libertarian voters account for less than 3 % of voters in each state, and are unlikely to be themselves pivotal in deciding overall park visitation behavior for a state. Thus, the practical significance of the difference in Libertarian vote share is doubtful. However, at the county level there is a clear divide in the partisanship of regions where park visitation did and did not undergo abrupt change. Counties with an abrupt change were more likely to be majority Democratic, while counties without a change point were more likely to be Republican. Taken together with the urban bias of the data set, it is possible that the state results are confounded by an over representation of urban park visits.

If abrupt park visitation changes were more associated with Democrat behavior, since urban areas have a Democratic bias, it is possible that the behavior of the urban park goers (who are more likely to be Democrats) may have overshadowed park going behavior in the rural parts of states. This possibility is made further plausible by the observation that the counties with a change point tend to be more populated. If park visitation changes are more likely in areas of greater population, and these areas are also over represented in the data, it stands to reason that aggregation to the state level may obscure behavior of the rural residents in the park visitation data.

Of course there is a second implication of these observations which is that whether or not park visitation exhibited an abrupt change is directly related to population density. If true, this relationship would explain why there is a disparity in population size for counties with and without abrupt changes, and why the counties with the lowest populations did not have abrupt changes, while the counties with the greatest populations did. In this case, differences in party affiliation of the respective areas is possibly unrelated, and only appears due to the confounding correlation between population density and party affiliation [101]. Since there is a connection between small populations and extreme partisanship as well, this would offer a potential explanation for why the span of the distribution across vote share for either party is greater for counties without abrupt changes.

Counties without abrupt changes in park visitation were more likely to have higher proportions of employment in Manufacturing, Construction, Mining, and Farming. Many of the workers in these sectors would have been considered "essential," and much of the work would be site specific. Meanwhile, counties with abrupt changes were more likely to have greater proportions of jobs in Information, Finance and insurance, Professional, scientific, and technical services, and Educational services;

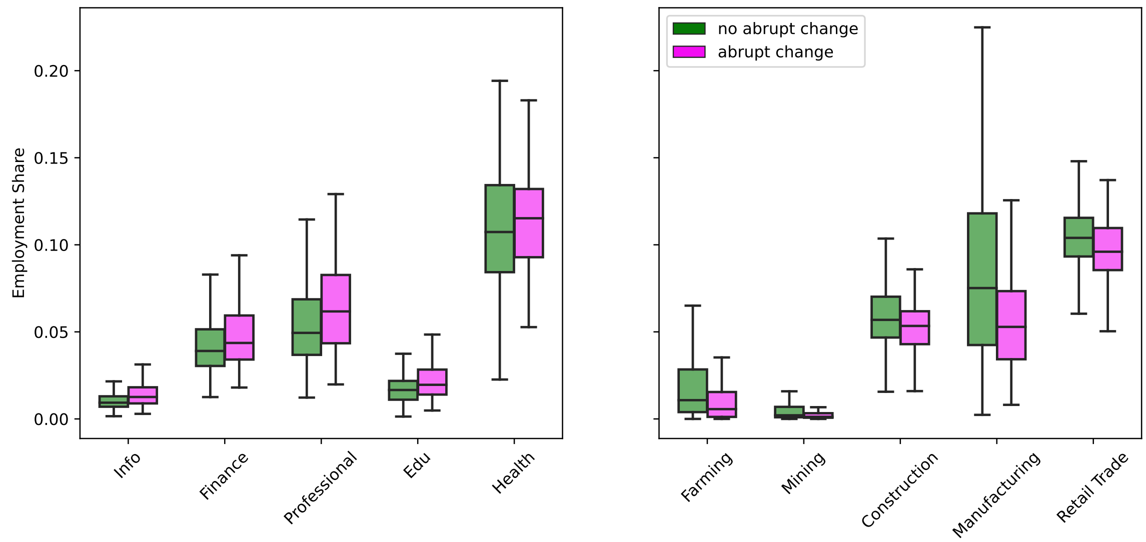

Figure 3.5: ***Box plots showing the distribution across employment share for counties with and without change points in the sectors where the distributions and their means were significantly different.*** *Sectors in the plot to the left were those where counties with a change point had significantly higher means ($p < 0.05$)), sectors in the plot to the right had significantly greater mean employment share in counties without change points. The distributions across employment share for counties with change points are shown in pink, while the distributions for counties without change points is shown in green. While the differences in mean and distribution for all shown sectors are significant, they are small.*

sectors where remote work would have been more widely adopted. It is curious that regions with greater proportions of remote workers, who may have had greater time and opportunity to visit parks at the time, were more likely to experience a drop-off in visits. The difference is interesting and suggests it is possible that reductions in employment related mobility impacted other mobility decisions, such as whether or not to visit parks.

However, while there are differences in employment share by sector, they are small, and their practical significance remains undetermined. The most striking differences found in this study were in population, and partisanship. Recent work [90,94,98] suggests that regions with higher Republican vote share exhibited less social distancing at the onset of the pandemic, were slower to adopt stay at home orders, and residents visited more points of interest than residents of regions with higher Democratic vote share, suggesting that overall mobility reduction was greater for Democratic counties than Republican ones. Insight from these new studies suggests that the lack of change in park visitation behavior among Republican regions simply reflects this partisan difference in mobility, and indicates that parks were not necessarily uniquely visited more or less relative to other points of interest.

3.7 Limitations and Future Directions

This study did not account for differences in local COVID-19 response policies. Incorporation of these differences would be necessary to understand how local governance impacted park access, and how willing residents were to defy local mobility restrictions for parks as opposed to other locations.

The spatial distribution of the parks in our data set roughly corresponds to the spatial distribution of the population, creating a substantial urban bias that we do not control for in this study. Weighing park visitation in such a way to allow for aggregation to the state level without over representing the urban parks would enable more revealing analysis at the state level, and additional insight into the demographic differences between counties with and without change points, with less influence from population density.

Augmenting the current data set with visitation data for more rural parks could also aid in these goals. Greater representation of rural parks would also allow a better investigation into population and park access as it relates specifically to population density and general nature accessibility.

Due to a change in collection methodology at the end of 2019, which led to a spatially non-uniform increase in total visitation counts, we were unable to directly compare 2019 and 2020 data. While there are visibly dramatic dips in behavior for some states and counties at the end of March 2020, it is not possible to clearly quantify how these changes deviate from expected behavior, nor how the magnitude of these changes compare across regions. Future work could investigate other park visitation data, and attempt to use it to normalize and perhaps compare visitation changes.

Comparison of visitation levels across years and regions, especially following the initial pandemic reaction, would be extremely helpful in determining whether or not there were differences in how park visitation was valued in different regions. This could also be achieved by comparing dips in park visitation to dips in visitation to other points of interest. In particular, it would be useful to understand how different areas, and different populations, weigh the benefits and risks of park usage in the pandemic, and how park usage diverted visitation to other destinations. Studies indicating which populations had access to parks, which may have been greatly beneficial during 2020, could be used to address potential social inequality, and reduce public health risk in the future.

Chapter 4

Park visitation and walkshed demographics in the United States

4.1 Abstract

A large and growing body of research demonstrates the value of local parks to mental and physical well-being. Recently, researchers have begun using passive digital data sources to investigate equity in usage; exactly who is benefiting from parks? Early studies suggest that park visitation differs according to demographic features, and that the demographic composition of a park's surrounding neighborhood may be related to the utilization a park receives. Employing a data set of park visitations generated by observations of roughly 50 million mobile devices in the US in 2019, we assess the ability of the demographic composition of a park's walkshed to predict its yearly visitation. Predictive models are constructed using Support Vector Regression, LASSO, Elastic Net, and Random Forests. Surprisingly, our results suggest that the demographic composition of a park's walkshed demonstrates little to no utility for

predicting visitation.

4.2 Introduction

The positive impact of park access, proximity, and use on well-being has been well established. Populations with park access have been found to be more active, have lower rates of obesity, and overall better cardiovascular health [81, 83, 102, 103]. As a form of nature exposure, park access also benefits cognition and mental well-being and is associated with lower stress and reduced rates of depression [80, 82, 104, 105]. Independent of time spent in the park, residential proximity is positively correlated with improved mental health [85].

Inequity in park resources can arise as a function of accessibility, or in qualitative differences in parks [106–112]. Studies have found that even when spatial access is equitable, income and race are linked to park quality. Parks associated with whiter and more affluent parks have more acreage, more tree canopy, and different amenities —including more playgrounds [108, 110, 111, 113–115].

Differences in park quality are significant because they may drive differences in visitation, which is the assumed mechanism by which parks offer health benefits. For example, while studies find correlations between residential proximity and lower rates of obesity, the underlying mechanism assumed to create this correlation is often that proximity to a park increases the likelihood of exercise in the park [116–119]. Similarly, studies on the impact of green space on mental health indicate that benefits are either accessed or increased by visitation to the space [80, 120, 121]. Park-based physical activity, for example, can mediate the relationship between park proximity and mental health benefits [36]. Thus, realized usage, or visitation, is an important variable when studying equity and the heath impact of parks [31]. Measuring park usage is particularly useful in equity studies because it can be used to quantify the effect of non-geographic barriers.

Non-geographic barriers to access—such as time constraints and safety —create inequity in the benefits that communities receive from parks. A much studied example of such a barrier is the perceived safety of parks [32–37]. Scholars have found that parks in higher income areas had fewer safety concerns than parks in low income areas and that perceptions of park safety are strongly tied to the odds of visiting a park [34, 37]. The impact of perceived safety as a barrier is visible in a 2020 study by Orstad, where mental health benefits were only associated with living in proximity to a park for residents who did not have concerns about park crime [36].

Classifying non-geographic barriers, and measuring their impact on communities remains an active area of research. Investigations with regards to these barriers requires measuring realized usage. Historically, park visits have been quantified using either surveys or by in-person observation of visitors [34, 35, 85, 103, 116, 117, 122–124]. These methods are limited both geographically and temporally, and can be expensive to implement.

Recently, these temporal and geographic limitations have been managed by using digital data sources such as social media and GPS data from mobile devices [30, 121, 125–127]. Digital data sources offer the ability to measure park usage at all times of the day, for prolonged periods of time, over a large geographic scale. By using Twitter and Flickr data, Hamstead et al. [30] were able to observe differences in visitation to all of the parks in New York City; their findings suggest visitation varies based on not only characteristics of the park, but on the demographic composition of the neighborhood surrounding the park. If park visitation varies with the demographic composition of its neighborhood for the US in general, it may be possible to predict park visitation using demographic data from its neighborhood.

Predicting which parks are being underutilized could inform decisions about which parks need more investment, better infrastructure, new programming, or could benefit from investigative studies into the barriers preventing full utilization. To explore whether park visitation can be predicted from demographic features, we use a novel dataset of daily park visitation counts obtained through observations of approximately 50 million mobile devices for 2,506 parks throughout the contiguous United States. In particular, we focus on the population residing within a ten minute walk of the park. This area is referred to in the literature as a "walkshed" and represents the residential area for which a park is considered "accessible" [28–30]

Historically, residents of the walkshed have been considered the primary users of the park, and walksheds have been conceptualized of as a fixed radius buffer around the park. Here we establish a park's walkshed using the convex hull of a pedestrian walking network within ten minutes walk of a park boundary. Using this method, the walkshed is limited to areas that have a walkable route to the park, and parks which cannot be reached by foot are excluded. Employing census data, we attribute demographic characteristics to the residents of a park's walkshed, and evaluate the ability of these characteristics to predict the visitations received by the park itself. Echoing the literature, we explore dimensions of race, income, educational attainment, gender, and age on park visitation in the United States.

4.3 Data

4.3.1 Park Visitation Data

Our data set consists of daily visitation records for 7,997 non-commercial parks in the contiguous United States for the year of 2019. Non-commercial parks are not operated for profit and include city, municipal, and neighborhood parks. National and state parks are excluded because they are typically larger and attract many tourists and infrequent visitors from a wide geographic area and therefore are not representative of the target relationship in this study.

The number of daily visitors to each park was estimated using the number of unique mobile devices that reported GPS data from within the park bounds on that day to the company UberMedia (since acquired by Near) [128].

UberMedia is a data vendor specializing in GPS data acquired from cellphones. UberMedia observes billions of devices world wide by collecting GPS coordinates with timestamps from a device's operating system when one of over 400 location collecting apps is used, or when the device is exposed to advertising through real-time bidding. In 2019, the average number of unique devices UberMedia observed each day (referred to as Daily Active Users or DAUs) varied between 44 and 60 million, and represented approximately 10% of the adult population in the United States [128].

4.3.2 Census and Geographic Data

Demographic data was retrieved from the 2019 American Community Survey 5 year data using the Census API. Census Tracts were chosen as the geographic unit as they are designed to be relatively homogeneous units. Demographic data gathered included race, ethnicity, age, median income, and educational attainment of residents over 25.

Demographic data was joined to park visitation data based on geographical overlap of census tracts with areas within walking distance of the park. In order to make these calculations, we used the geographic boundaries of the parks as given by UberMedia, the geographic boundaries of census tracts in 2019, and the network of walking routes surrounding each park obtained from the OpenStreetMaps API [129]. This process is described in section 4.4.1.

4.4 Methods

Exploring the predictive value of the demographic features of the surrounding neighborhood for a park's visitation required joining the data sets mentioned in 4.3, confirming the suitability of the resulting data for the study, and evaluating the predictive ability of several models. Each of these steps are detailed in the subsections below. Section 4.4.1 explains how demographic data from the US Census was joined to park visitation data using a geographic unit called a walkshed. This section also discusses how population data was aggregated for the walkshed. The inclusion criteria chosen for the study is delineated in section 4.4.2. The resulting data set is described in section 4.4.3. Compared to the US population, the walksheds included in our data have a large distribution of demographic features. Finally, section 4.4.4 explains the methodology by which models were chosen, fit, and evaluated.

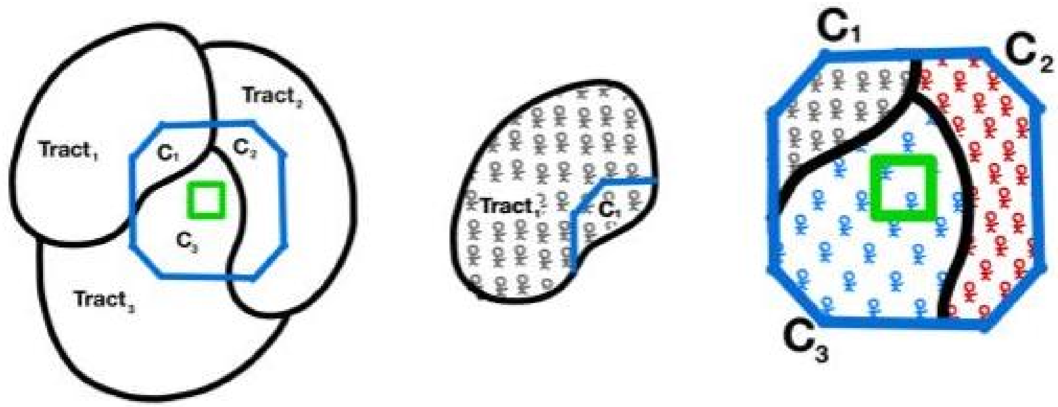

Figure 4.1: ***A conceptual model of the walkshed.*** *In the left most image the park is represented by a green square, and the walkshed is represented by the blue polygon surrounding the green square. The three black bordered shapes labeled "Tract" demonstrate how a walkshed could intersect multiple Census Tracts. The intersections of the walkshed with the census tracts are labeled C_i, where C indicates the region is a component of the walkshed, and i refers to the census tract that the particular component lies within. The middle image demonstrates the assumed uniform spatial distribution of a homogenous population within the tract, where the people associated with the component have the same features as the tract, and the population is proportional to the area of the census tract covered by the component. The final image is of the walkshed, with each of its components populated with respect to the census tract in which they lie. The total walkshed population is considered to be the aggregation of the populations of each component.*

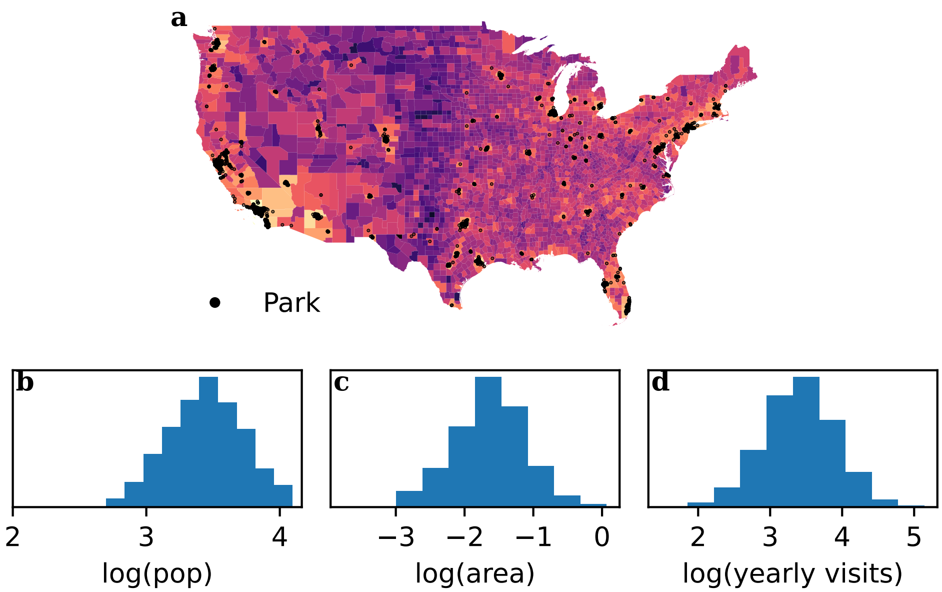

Figure 4.2: ***Heat map of the population of the contiguous United States overlaid with the locations of the parks used in the study, with each park demarcated by a black point, and distributions of park sizes and popularity below****. Histograms display the log normal distributions of parks in the dataset across walkshed population, park area, and yearly visits. The median park has 2785 people in its walkshed, covers 5.68 acres, and received 2528 visits in 2019.*

4.4.1 Walkshed Construction

For each of the parks in our dataset we constructed a walkshed and calculated an estimated walkshed population. The walkshed was defined to be the convex hull of the graph of the walking network obtained from OpenStreetMaps that represented a ten minute walk to the boundary of the park. This walkshed was then considered to be composed of the disjoint components lying in unique Census Tracts (see 4.3.2). Census Tracts are treated as homogenous populations uniformly distributed over a geographic area. Walksheds for each park in our study were created through a series of steps (Figure 4.1). The estimated number of people in a walkshed associated with park p, P^p_{walkshed} was computed as

$$P^p_{walkshed} = \sum_{i=0}^{n} P_{tract_i} \frac{A_i}{A_{tract}} \tag{4.1}$$

where P_{tract_i} is the estimated number of residents in the i^{th} census tract intersecting with the walkshed, A_i is the area of that intersection, and n is the total number of census tracts intersecting with the walkshed.

In addition to the number of people in the walkshed, we estimated the income, gender, age, educational attainment, ethnicity, and racial composition of the walkshed population. Taking the census tract as a homogeneous unit, the income of a person within the walkshed was calculated as the population-weighted average of the median incomes of the tracts intersecting that walkshed:

$$I_{walkshed} = \frac{1}{P_{walkshed}} \sum_{i=0}^{n} P_i I_i \tag{4.2}$$

where $I_{walkshed}$ is our estimate of the median income of a person in the walkshed and I_i is the median income of the census tract containing component i.

The average age of the residents was calculated similarly. The proportion of the walkshed belonging to a racial group, ethnicity, age range, sex, or having reached a given educational attainment was calculated as:

$$P_{walkshed,j} = \frac{1}{P_{walkshed}} \sum_{i=0}^{n} P_{i,j} \tag{4.3}$$

.

Where i refers to the component, and j refers to the classification of interest, for example $j \in \{race_1, race_2, \cdots race_J\}$.

4.4.2 Inclusion Criteria

A park was included in analysis if (i) the walkshed had an area greater than zero (i.e. the park could be accessed on foot) and (ii) demographic data was available from the US census for each census tract intersecting with the walkshed. Parks were omitted from the study if their walkshed included a component contained in a tract for which the US Census reported no data.

Because we assumed people in a census tract were distributed uniformly, we excluded a walkshed if it contained a census tract with less than one person per quarter-acre.

Parks with walksheds containing fewer than 500, or more than 12,500 people were also excluded. Parks with high population walksheds were considered large enough to attract tourists. Parks with fewer than 500 residents were considered too rural, with potentially inconsistent visitation. These parks accounted for 154 of the parks in the dataset. Of the initial 7,997 parks, 2,506 met our inclusion criteria. The primary reason for exclusion was insufficient population density.

4.4.3 Study Set Description

Our data set contained 2506 parks, which we observed to exhibit log-normal distributions in walkshed population, park area, and yearly visits (Figure 4.2). The average park in our data set was 0.023 km^2 (5.68 acres), contained 2,785 people in its walkshed, and received 2,528 visits each year.

The income associated with walksheds was also log-normally distributed. The average income associated with a park's walkshed was $76,155 per year, which is $7,542 more than the median income for the general population of the United States. Approximately 54% of communities associated with a walkshed had a lower proportion of individuals living below the Federal Poverty Level than the United States as a whole.

Figure 4.3 presents a summary of the distribution of walksheds with regards to race, ethnicity, educational attainment, and age. 55% and 46% of the parks in our dataset had a greater proportions of Asians and Hispanics than in the general population, while only 22% and 37% of parks had greater proportions of Black and White people than the general population. The distribution of parks over composition by gender was normal, with the average park having a similarly slightly female composition as the United States as a whole (51% and 49.8% respectively).

Comparing the educational attainment of walkshed populations to that of the general population, we see that roughly 60% of the walkshed populations had more highschool graduates, college graduates, and persons with advanced degrees than the general population. Overall, walkshed populations tended to have higher educational attainment than the general population.

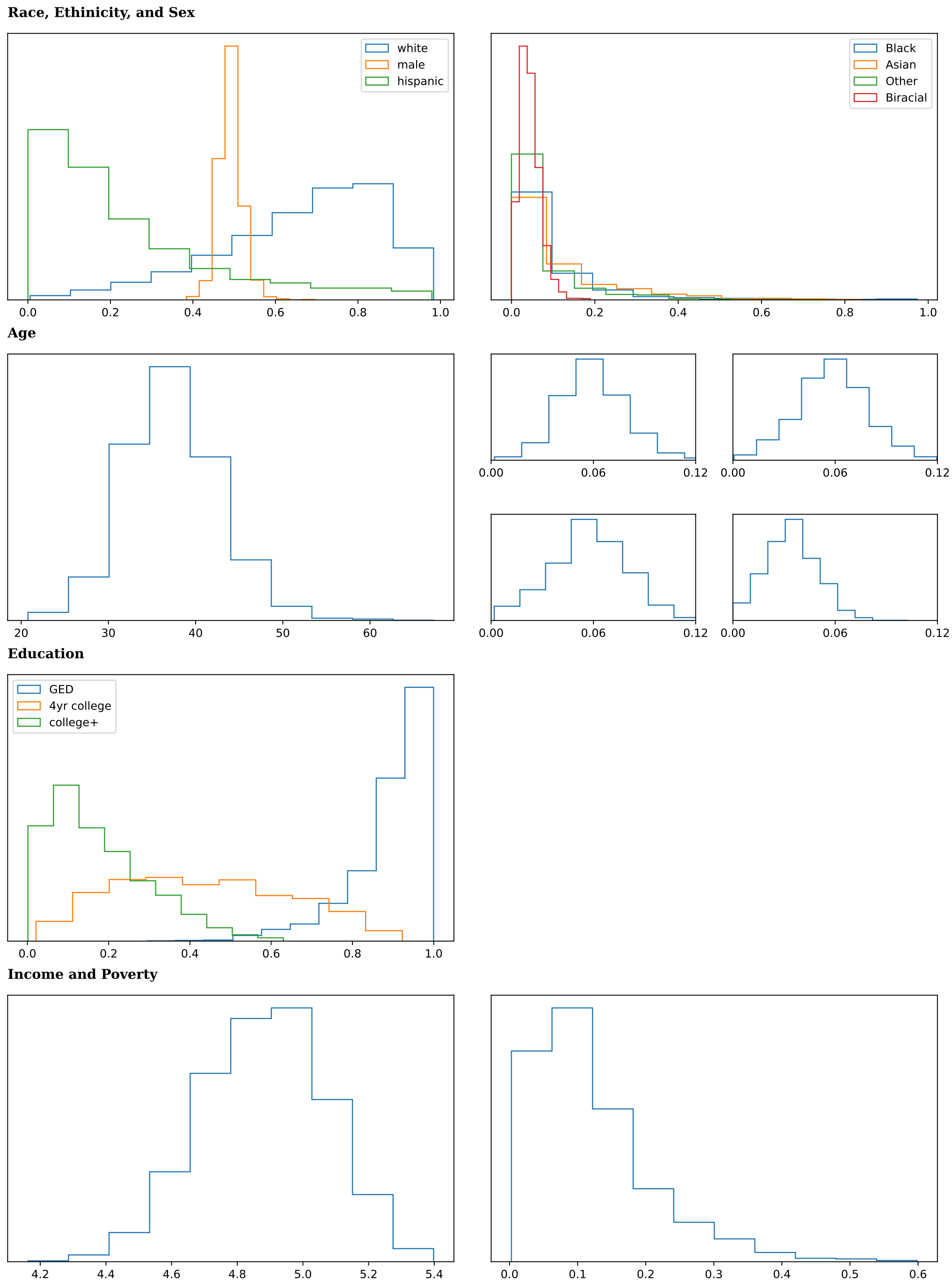

Figure 4.3: ***Distributions of the parks in the study set across race, ethnicity, sex, educational attainment, age and measures of wealth.*** *The first two rows display the distribution of parks in the study set across fraction of walkshed population in four racial categories (Black, Asian, Multiracial, and White), fraction identifying as Hispanic, and fraction male. The third row displays the distribution of parks in the study set across fraction of walkshed population over 25 having earned at least a high school diploma (or equivalent) (left), at least a Bachelor's degree(middle), and with more than a Bachelor's degree (right). The final two rows indicate the distribution of parks in the study set across average age -defined as the population weighted average of the medians for each component, and the fraction of walkshed residents falling into each of four child age ranges (under 5, 6-10, 11-14, and 15-17).*

The age composition of parks was also normally distributed. On average the parks in our data set had a similar median age, and similar proportions of children in each age group, as the general population.

4.4.4 Analysis

In our data set, park size and walkshed population are both positively correlated with yearly visitation on the log-log scale (Pearson correlation of 0.61 and 0.47 respectively, Spearman correlation of 0.59 and 0.47). Both of these factors effect the spatial accessibility of the park, but do not necessarily reflect the users of the park, or the demographic environment of the park. In order to focus on the latter features and the predictive value they have for park visitation, we normalize for the increased visitation that is likely only related to the increased population to which the park is available. To normalize we fit simple linear regressions for park area and walkshed population in the log-log space. The resulting slopes were used as exponents in the following normalization parameter:

$$n_p = \frac{1}{A_p^{0.58} * P_w^{0.84}} \tag{4.4}$$

where A_p is the area of the park in km^2, and P_w is the population of the walkshed. The exponents for population and area were determined using the slope of the line of best fit for each variable when plotted against yearly visits in log-log space (See Figure 4.4).

Normalized visitation was then used as the target value for a set of predictive models using the demographic features associated with the walkshed as inputs. Each

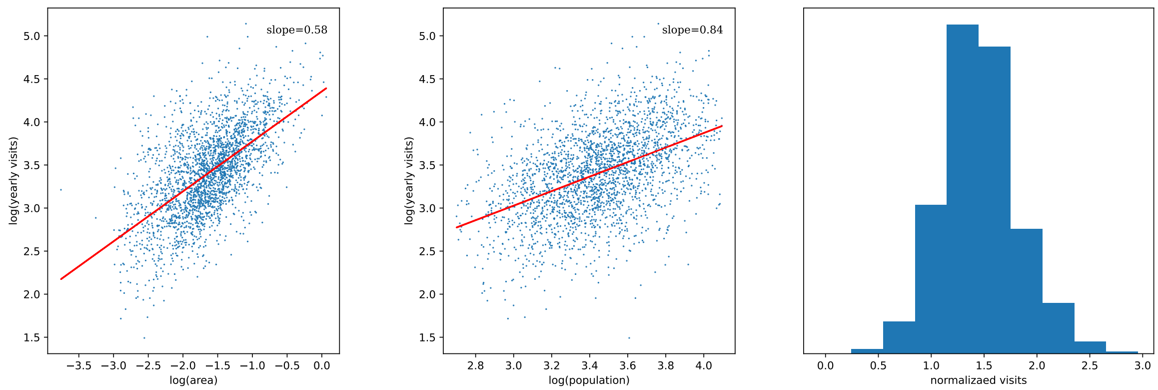

Figure 4.4: ***Log-log plots displaying the relationship between yearly visits and park area and walkshed population, and the distribution of normalized visitation.*** *In log space park area is positively correlated with yearly visits, as is population of the walkshed. The slope of the line of best fit is 0.58 for park area, and 0.84 for walkshed population. These slopes were used to determine the normalized visitation value, given by Equation 4.4. Visitation remains log-normally distributed after normalization.*

demographic feature was included as a separate input variable. The models tested were: Support Vector Machine Regression with linear, polynomial, and radial basis function kernels; Random Forest Regression; LASSO Regression; and Elastic Net Regression. These models were selected for the variety of functional forms offered, and for the ability of some to consider subsets of features. Random Forests were included to account for possible interaction effects between demographic features. Model fitting and analysis were performed in Python using Scikit-learn [2].

For each model, both input and target values were standardized by subtracting the mean and dividing by the sample standard deviation, prior to fitting. Model parameters were optimized using a grid search of the parameter space (see Table 4.1) over which models were evaluated using 5-fold cross validation. The best performing parameters and the model scores are presented in Table 4.1.

Model performances were compared against a null model: a constant function set to the mean of the normalized visitation. The null model asserts that all parks receive the same visitation From observing the distribution of parks across normalized visitation (Figure 4.4) we know that parks do not receive uniform normalized visitation, and therefore, that the null model is a poor predictor.

4.5 Model Performance

Model	Hyperparameters	Hyperparameter Grid	Tuned Hyperparameters	Mean Absolute Error
Constant (null)	-	-	-	0.368
SVR - linear kernel	C	$[10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}]$	10^{-4}	0.2806
SVR - rbf kernel	C	$[10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}]$,	0.99	0.2793
	γ	$[\frac{1}{n_{features}}, \frac{1}{(n_{features}*X.var()}]$	$\frac{1}{n_{features}}$	
SVR - polynomial kernel	C	$[10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}]$,	10^{-1}	0.2736
	γ	$[\frac{1}{n_{features}}, \frac{1}{(n_{features}*X.var()}]$	$\frac{1}{n_{features}}$	
	degree	[1,2,3,4,5]	3	
LASSO	α	$[10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}]$	10^{-4}	0.2759
Elastic Net	α	$[10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}]$,	10^{-4}	0.2758
	l1_ratio	[0.00, 1, 0.01]	0.01	
Random Forest	$n_{estimators}$	[200, 400]	200	0.2664
	$max_{features}$	$n_{features}, \sqrt{n_{features}}$	$\sqrt{n_{features}}$	
	max_{depth}	[10,30,50,70,90,110]	10	
	min samples to split	[2,5,10]	2	
	min samples at leaf	[1,2,4]	1	
	bootstrap	[True, False]	False	

Table 4.1: ***Results of training and testing predictive models for predicting normalized yearly park visitation using the demographic features associated with that park's walkshed.*** *Model types are presented along with the hyperparameters they use. The hyperparameters were tuned using a 5 fold cross-validation and a grid search over the parameter space. For each model the grid used for the grid search is reported followed by the tuned hyperparameters and the Mean Absolute Error of the tuned model. The notation* $n_{features}$ *refers to the number of independent variables, and* $X.var()$ *refers to the variance of the feature array. The l1_ratio is the ratio of the weights of the* $L1$ *and* $L2$ *penalties, such that* $l1\_ratio = 1$ *is the LASSO penalty. All hyperparameters refer to those in the scikit-learn library [2].*

Modeling the data using a constant function set to the mean of the normalized visitation yielded a Mean Absolute Error (MAE) of 0.2810. The random forest model was most successful of the regression models, achieving a MAE of 0.2664, or an 5.20% improvement over the null model.

Support Vector Regression (SVR) with a polynomial kernel was the second most effective model. This model resulted in a MAE of 0.2746, which is 2.33% less than the null model. SVR using radial basis function kernels had slightly worse performance, obtaining a MAE of 0.2793; only a 0.60% improvement over the null model. When paired with a linear kernel SVR performed comparably to the null model, with a MAE of 0.2810.

The other two linear models performed equally poorly. LASSO had a minimum MAE of 0.2759, and Elastic Net managed a very similar minimum MAE of 0.2758, indicating that they probably converged to very similar linear models. This is only an 1.81%, and 1.85% improvement upon the null model.

Given that the most successful model, the random forest, only achieved an improvement of 5.20% over the null model, we can conclude that demographic information did not significantly contribute to more accurate predictions of normalized yearly visits.

4.6 Discussion

Our approach investigated a variety of functional forms, employed methods for decreasing noise created by unimportant features (LASSO and Elastic Net), and incorporated the potential for interaction effects through sequential variable consideration in the Random Forest model. Additionally, the hyperparameters of each model were tuned. None of the predictive models performed substantially better than the null model. The null model provides no information about individual park usage, and has very weak predictive power. It also assumes that the demographic features considered in this study are irrelevant to visitation prediction, as it does not use them at all. Since a comprehensive set of models was tested, and none performed considerably better than the null model, we can infer that the demographic features have little predictive value for average yearly park visitation.

This result is perhaps unexpected in light of Hamstead et al's 2018 findings in New York City [30]. In contextualizing this discrepancy, it is important to note that our inclusion criteria intentionally focuses our research on community parks in suburban or urban residential areas. This effort decreased the influence of tourism on visitation by decreasing the number of parks that would be considered tourist destinations. In addition, exclusion of these parks naturally excludes potential demographic disparity in residential proximity to destination parks (i.e. homes near destination parks may be more expensive) as a confounding effect. Thus it is possible that our results differ because of the type of parks considered.

This study is limited by the inability to determine the residential proximity of visitors. A correlation was observed between the number of people in a park's walkshed and the number of visits received by a park. This suggests that the number of people a park "serves" is related to the number of people who visit it. In order to control for the effect of park size, and geographical accessibility, a normalization constant was applied to control for this relationship. In doing so, we made an assumption that a ten minute walk to the park encompassed the 'service area' of every park.

It is possible that depending on demographic features, some populations may travel farther to parks on average than others, but maintain similar visitation rates. If this were the case, the normalization constant for parks serving those populations would be too small, inflating the normalized visitation observed, and potentially obscuring important trends.

It is worth reiterating that these results are park-centered; results speak to the usage that a park receives, but do not reflect *who* uses the park. Our data does not provide the origin of the mobile device visiting the park there is no way to determine if visits are made by the people living in the walkshed, or by persons living farther away. Therefore, whether park usage differs for different *populations* is not addressed by this work.

Future work should include more detailed data on the origin of the devices observed in each park. This data would allow for a more accurately constructed walkshed. It would also allow further exploration into who visits parks, and which parks they visit. In addition, consideration should be given to park quality; this work did not incorporate features of the park itself into the predictive modeling. Park amenities, condition, and environment are all important contributors to visitor attraction. Since aspects of park quality could be confounding factors with walkshed demographic features, a logical extension of this work is to control for this confounding.

Further future analyses would benefit from consideration of park visitation relative to season and weather. In the current study yearly visitation is used, which obscures the difference in “visitation-season" length between parks; parks in Southern California may be used more days in a year relative a park in Boston, which creates a different story of population and visitation.

Chapter 5

Concluding Remarks

This work has demonstrated and explored the use of passively collected behavioral data for evaluating population health. Both sleep and nature exposure are pertinent to well-being, and have historically been difficult to measure at the population level. Because of the growing ubiquity of mobile devices, and interactions with technology, these aspects of health are now observable at the population scale complementing traditional survey-based methods. The scaling of the new monitoring tactics allows for a more extensive inquiry into the population, over greater periods of time, with finer granularity, and with the promise of a near real-time estimate.

While the data used in these projects has been illustrative of behavioral patterns directly related to health, it is severely limited in comparison to what exists. At the time of this writing, wearable devices are capable of capturing heart rate, taking EKGS, and measuring blood oxygen levels. Validating these features, ensuring their success across broad and diverse populations, and comparing results to standard measurement devices is ongoing. Still, for many applications the estimates provided by these devices may be enough.

In chapter two we built on work by Lepunskiy et al [18] which demonstrated the ability of Twitter data to measure population sleep duration. We were able to use Twitter data to observe the sleep loss event of Spring Forward, which is associated with enormous negative public health outcomes. If we are able to observe such meaningful events with what can be considered a rather abstract data source, there is enormous potential for using data which is more closely aligned with sleep itself. Scholars and industry professionals have been able to use data from mobile devices, as well as wearables, to infer sleep for individuals from accelerometers, device usage patterns, and ambient noise levels. Some apps offer even greater insight, not only into the duration of sleep, but the quality.

If this type of data could be aggregated and applied to the population scale it would create opportunities for natural experiments based on current policies - for example on the impact of elementary and high school start times on sleep health- and the ability to create targeted policies to improve the health of sleep disadvantaged communities. An example of such a targeted policy might be measures to reduce anthropogenic noise in a specific neighborhood, perhaps through provision of funds for sound barrier construction near a busy road way, to create a better sleep environment. Policy makers, and in turn society, would experience health improvements if we are able to harness our ability to measure and quantify the public health benefits of such policies.

Despite the relationship between sleep and health, and sleep and policy, sleep is often understood as a personal responsibility. By measuring sleep at the population scale we can reframe sleep as a public health matter, and establish the impact of societal norms, expectations, and pressures on sleep.

Nature exposure, in contrast to sleep, has the benefit of being considered - to some extent - as within the realm of public health. This is related in part to the clearly observable physical geography created through urban planning, which is a function of governments. There has been a consensus that greenery should be for everyone, reflected in both the literature and the 'greening' efforts of many cities, and while there are still ways in which the physical availability of green space falls short of equitable, there is less consideration to the role of policy in the actual usage of the green space.

In chapter three and four we investigate green space access through a proxy measure for usage. Looking at park visitation with respect to the pandemic in chapter three, we observed that some counties experienced a severe drop in park visitation at the onset of the pandemic, while others did not. We found that the counties that did and did not experience this drop differed in population density, income, employment sector, and most notably in voting results for the 2016 presidential election. However, when these results are contextualized within the literature regarding mobility at the onset of the pandemic, we find that park visitation was not necessarily special – mobility decreased (or failed to do so) in similar ways across all destinations at this time, and largely in alignment with messaging from local governments.

In chapter three we did not explore the impact of differing policies on park visitation. Such an investigation may have been well served by a case study with comparisons made across regions with measurably dissimilar COVID-19 response policies. Our interest, however, was in observing behavior for as much of the country as possible - and in broad differences in population. As the pandemic response was so fragmented, regionally and temporally, within the US, as well as inconsistent and unique depending not only on Federal, State, and County policy - but municipal as well - it was not practical to have a meaningful comparison of policy and park visitation behavior at the scale we wished.

Though we were unable to look at the impact of policy on park visitation, the pandemic provided a clear case in which park access through usage could be understood as a result of policy rather than simply personal choice. In some areas parks were closed entirely, the influence of which is as understandable as a physical barrier to access. Other policies, for example providing well-lit paths to access parks, may be less obvious in their influence, and so require the direct observation of park usage rates to discern.

In chapter four we address the existence of ‘unseen’ barriers by exploring visitation rates in relation to demographic features of the walkshed. The walkshed, or the area within a ten minute walk of a park, is thought to house the majority of a park’s visitors, but also to form part of the park’s environment. We sought to understand whether certain demographic features associated with a park through its walkshed population (namely sex, age, race, wealth, and educational attainment) could be used to predict the visitation a park receives. Our results conclude that these features have little predictive value for park visitation.

This result is somewhat surprising given that the features that make a park more desirable to visit (the amount of canopy, the type of infrastructure available, etc) are often associated with more affluent neighborhoods. It is also surprising given work which finds racial and ethnic minorities have less access to parks through visitation, and that visitation to parks in New York City is related to the ethnic and racial composition of the neighborhood surrounding a park. However, there are future studies that may shed light on this discrepancy.

In particular, our study inherently assumes that a standard and temporally constant walkshed is relevant to every park. The mobile device data used in this study only includes whether a device visited the park, it does not include where the device originated from, though that data does exist. If data based on the origin of the visiting devices was incorporated into this study, it would be possible to explore the distances travelled to reach each park, and to establish more relevant walksheds (or even drivesheds). Moreover, understanding who was using the parks might be more informative than simply who is living around the parks. Though we are not seeing a predictive relationship between the demographic features of the people immediately surrounding a park and visitation to the park, we have no way at the moment of looking at the demographic features of the actual visitors. Thus, we continue to rely at least partially on the assumption that those living in the walkshed are the primary users of the park.

Additionally, gathering a larger dataset, or for a different set of parks, would be informative. Studies have found that larger parks are associated with greater wealth, and our own work suggests that larger parks receive more visits. These observations together suggest that there should be a predictive relationship between wealth and visitation. However, in our data set there is no relationship between wealth and park area, which could be obscuring the relationship between wealth and visitation.

The data to vastly improve the studies presented in this work exists currently. There are two primary obstacles to use: privacy, and cost. Digital data has become a commodity in recent years, in which interest is substantial. Most of the data is collected by private firms, and is not freely available. Moreover, because of the commoditization and wide spread use of digital data, concerns have arisen regarding privacy. While the new technologies available for monitoring health could allow us to quite literally take the pulse of the entire country, whether this is ethical, or can be done in such a way that conserves privacy is as of yet unknown.

Chapter 6

Supplementary Material for Chapter Two

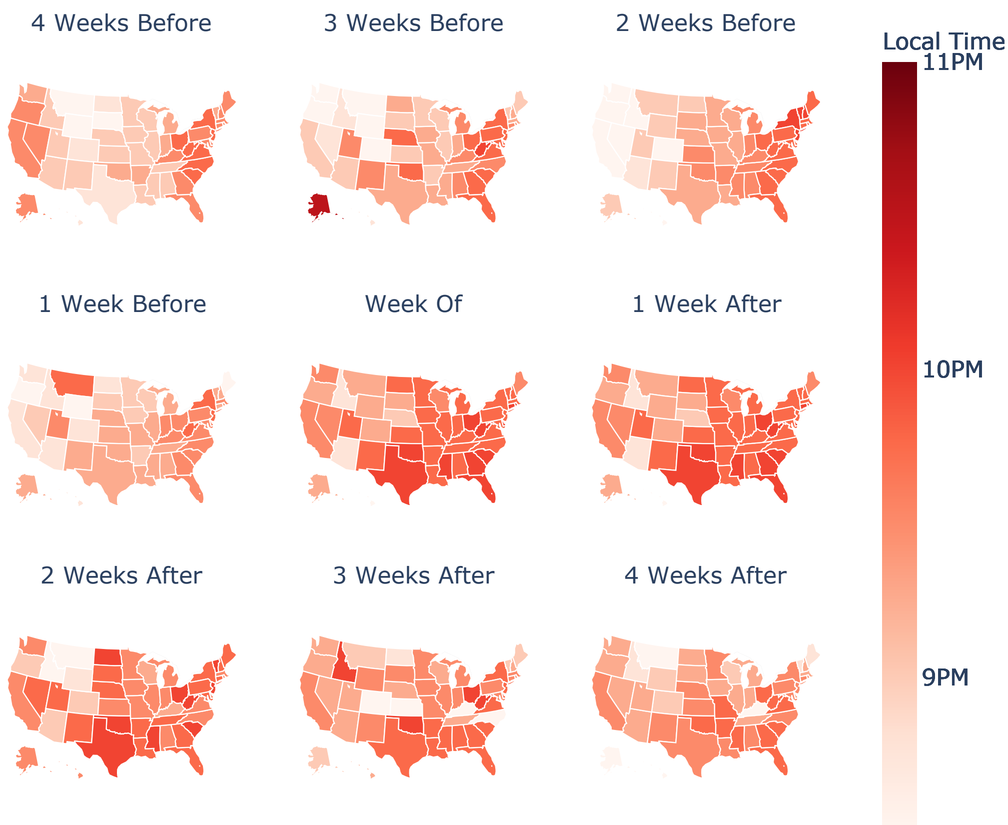

Figure 6.1: ***Peak activity time (local) for the Sunday of the four weeks prior to, the week of, and the four weeks following Spring Forward, aggregated from 2011 to 2014.*** *We have used the same colormap as for Fig. 2.3 in the main manuscript. States shown in white had a peak time that was 9 pm or earlier. From 2011 to 2013, the Academy Awards took place two weeks prior to Spring Forward, while in 2014 they took place one week prior. A clear discontinuity is visible between the "One Week Before" and "Week Of" maps.*

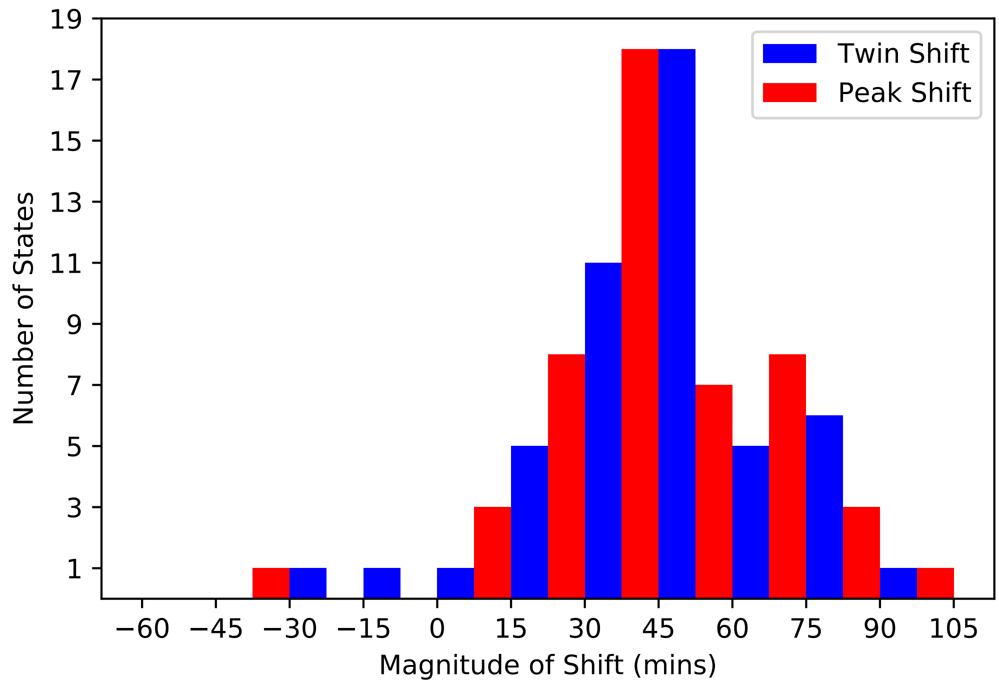

Figure 6.2: ***Histogram showing the Peak Shift and Twinflection Shift measured for each state in 2013.*** *The magnitude of the shift in minutes is on the x axis, and the height of each bar is the number of states with a shift of this magnitude. Blue bars represent Peak Shift, while red bars represent Twin Shift. Both measurements display a positive shift for most states. For Peak Shift the exceptions were the District of Columbia, having a -45 minute shift, and Hawaii having a -150 minute shift (not shown). For Twin Shift the exceptions were Alaska, with a -15 minute shift, and Hawaii with a -30 minute shift. Wyoming is not included in this figure as there were no tweets posted from Wyoming on the day following Spring Forward in 2013.*

State	count	State	count	State	log(TPC)	State	log(TPC)
AK	5265	CA	326963	AK	-2.14	DC	-1.45
AL	51955	TX	223565	AL	-1.97	LA	-1.93
AR	26229	NY	151345	AR	-2.05	AL	-1.97
AZ	40326	FL	134044	AZ	-2.21	MS	-1.99
CA	326963	IL	115887	CA	-2.07	IA	-2.00
CO	34070	OH	104102	CO	-2.18	OK	-2.00
CT	23189	GA	96759	CT	-2.19	GA	-2.01
DE	6079	PA	86693	DE	-2.18	MA	-2.03
DC	22628	NC	79102	DC	-1.45	OH	-2.04
FL	134044	MA	62707	FL	-2.16	WA	-2.05
GA	96759	WA	62114	GA	-2.01	IL	-2.05
HI	8143	MI	61833	HI	-2.23	TX	-2.05
IA	30965	VA	60125	IA	-2.00	AR	-2.05
ID	3962	TN	56028	ID	-2.61	NE	-2.07
IL	115887	LA	53912	IL	-2.05	CA	-2.07
IN	41224	AL	51955	IN	-2.20	TN	-2.07
KS	23894	NJ	49374	KS	-2.08	KS	-2.08
KY	29543	WI	45130	KY	-2.17	WV	-2.09
LA	53912	IN	41224	LA	-1.93	NC	-2.09
MA	62707	MD	40483	MA	-2.03	RI	-2.10
MD	40483	AZ	40326	MD	-2.16	SC	-2.10
ME	7448	MO	39990	ME	-2.25	WI	-2.10
MI	61833	OK	38231	MI	-2.20	NY	-2.11
MN	34862	SC	37326	MN	-2.19	VA	-2.13
MO	39990	MN	34862	MO	-2.18	AK	-2.14
MS	30418	CO	34070	MS	-1.99	FL	-2.16
MT	3708	IA	30965	MT	-2.43	MD	-2.16
NC	79102	MS	30418	NC	-2.09	NV	-2.16
ND	3469	KY	29543	ND	-2.30	PA	-2.17
NE	15951	AR	26229	NE	-2.07	KY	-2.17
NH	6787	OR	25003	NH	-2.29	MO	-2.18
NJ	49374	KS	23894	NJ	-2.25	DE	-2.18
NM	11319	CT	23189	NM	-2.27	CO	-2.18
NV	18905	DC	22628	NV	-2.16	MN	-2.19
NY	151345	NV	18905	NY	-2.11	CT	-2.19
OH	104102	NE	15951	OH	-2.04	OR	-2.19
OK	38231	WV	15061	OK	-2.00	IN	-2.20
OR	25003	UT	12271	OR	-2.19	MI	-2.20
PA	86693	NM	11319	PA	-2.17	AZ	-2.21
RI	8367	RI	8367	RI	-2.10	HI	-2.23
SC	37326	HI	8143	SC	-2.10	ME	-2.25
SD	3940	ME	7448	SD	-2.33	NJ	-2.25
TN	56028	NH	6787	TN	-2.07	NM	-2.27
UT	12271	DE	6079	UT	-2.37	VT	-2.27
VA	60125	AK	5265	VA	-2.13	NH	-2.29
VT	3358	ID	3962	VT	-2.27	ND	-2.30
WA	62114	SD	3940	WA	-2.05	SD	-2.33
WI	45130	MT	3708	WI	-2.10	UT	-2.37
WV	15061	ND	3469	WV	-2.09	MT	-2.43
WY	1545	VT	3358	WY	-2.57	WY	-2.57
TX	223565	WY	1545	TX	-2.05	ID	-2.61

Table 6.1: ***Tweet Counts.*** *Tweet count and tweets per capita ($\log_{10}$) sorted alphabetically and in order of volume for the four ASF Sundays observed in 2011-2014.*

State	BSF	State	BSF	State	SF	State	SF
AK	09:45	GA	10:00	AK	09:30	CT	10:15
AL	09:30	NC	10:00	AL	10:00	DE	10:15
AR	09:30	NJ	10:00	AR	10:00	FL	10:15
AZ	09:00	NY	10:00	AZ	09:00	GA	10:15
CA	09:15	OH	10:00	CA	09:45	MS	10:15
CO	08:45	PA	10:00	CO	09:30	OH	10:15
CT	09:45	SC	10:00	CT	10:15	OK	10:15
DE	09:45	VA	10:00	DE	10:15	RI	10:15
DC	09:45	WV	10:00	DC	09:45	SC	10:15
FL	09:45	AK	09:45	FL	10:15	TX	10:15
GA	10:00	CT	09:45	GA	10:15	WV	10:15
HI	09:00	DC	09:45	HI	08:00	AL	10:00
IA	09:30	DE	09:45	IA	10:00	AR	10:00
ID	09:00	FL	09:45	ID	09:00	IA	10:00
IL	09:15	IN	09:45	IL	10:00	IL	10:00
IN	09:45	KY	09:45	IN	10:00	IN	10:00
KS	09:30	MA	09:45	KS	10:00	KS	10:00
KY	09:45	MD	09:45	KY	10:00	KY	10:00
LA	09:30	MI	09:45	LA	10:00	LA	10:00
MA	09:45	NH	09:45	MA	10:00	MA	10:00
MD	09:45	OK	09:45	MD	10:00	MD	10:00
ME	09:30	RI	09:45	ME	09:45	MI	10:00
MI	09:45	VT	09:45	MI	10:00	MN	10:00
MN	09:15	AL	09:30	MN	10:00	MO	10:00
MO	09:30	AR	09:30	MO	10:00	NC	10:00
MS	09:30	IA	09:30	MS	10:15	ND	10:00
MT	09:00	KS	09:30	MT	09:30	NH	10:00
NC	10:00	LA	09:30	NC	10:00	NJ	10:00
ND	09:15	ME	09:30	ND	10:00	NM	10:00
NE	09:30	MO	09:30	NE	09:15	NY	10:00
NH	09:45	MS	09:30	NH	10:00	PA	10:00
NJ	10:00	NE	09:30	NJ	10:00	TN	10:00
NM	09:30	NM	09:30	NM	10:00	UT	10:00
NV	09:15	TN	09:30	NV	09:45	VA	10:00
NY	10:00	UT	09:30	NY	10:00	VT	10:00
OH	10:00	CA	09:15	OH	10:15	CA	09:45
OK	09:45	IL	09:15	OK	10:15	DC	09:45
OR	08:45	MN	09:15	OR	09:30	ME	09:45
PA	10:00	ND	09:15	PA	10:00	NV	09:45
RI	09:45	NV	09:15	RI	10:15	WA	09:45
SC	10:00	SD	09:15	SC	10:15	WI	09:45
SD	09:15	TX	09:15	SD	09:30	AK	09:30
TN	09:30	WA	09:15	TN	10:00	CO	09:30
UT	09:30	WI	09:15	UT	10:00	MT	09:30
VA	10:00	AZ	09:00	VA	10:00	OR	09:30
VT	09:45	HI	09:00	VT	10:00	SD	09:30
WA	09:15	ID	09:00	WA	09:45	WY	09:30
WI	09:15	MT	09:00	WI	09:45	NE	09:15
WV	10:00	CO	08:45	WV	10:15	AZ	09:00
WY	08:45	OR	08:45	WY	09:30	ID	09:00
TX	09:15	WY	08:45	TX	10:15	HI	08:00

Table 6.2: ***Time of Peak Twitter Activity by State.*** *Time of peak Twitter activity Before Spring Forward (BSF) and the week of Spring Forward (SF) for each state, listed alphabetically and by time of peak.*

State	Peak	State	Peak	State	Twin	State	Twin
AK	-15	TX	60	AK	-30	TX	105
AL	30	CO	45	AL	45	MS	75
AR	30	IL	45	AR	45	ID	60
AZ	0	MN	45	AZ	15	IL	60
CA	30	MS	45	CA	30	LA	60
CO	45	ND	45	CO	45	NM	60
CT	30	OR	45	CT	15	AL	45
DE	30	WY	45	DE	15	AR	45
DC	0	AL	30	DC	30	CO	45
FL	30	AR	30	FL	30	GA	45
GA	15	CA	30	GA	45	IA	45
HI	-60	CT	30	HI	-15	MN	45
IA	30	DE	30	IA	45	MO	45
ID	0	FL	30	ID	60	ND	45
IL	45	IA	30	IL	60	OK	45
IN	15	KS	30	IN	30	TN	45
KS	30	LA	30	KS	30	VT	45
KY	15	MO	30	KY	15	CA	30
LA	30	MT	30	LA	60	DC	30
MA	15	NM	30	MA	15	FL	30
MD	15	NV	30	MD	15	IN	30
ME	15	OK	30	ME	15	KS	30
MI	15	RI	30	MI	30	MI	30
MN	45	TN	30	MN	45	MT	30
MO	30	UT	30	MO	45	NC	30
MS	45	WA	30	MS	75	NE	30
MT	30	WI	30	MT	30	NV	30
NC	0	GA	15	NC	30	NY	30
ND	45	IN	15	ND	45	OR	30
NE	-15	KY	15	NE	30	RI	30
NH	15	MA	15	NH	15	SC	30
NJ	0	MD	15	NJ	15	SD	30
NM	30	ME	15	NM	60	VA	30
NV	30	MI	15	NV	30	WI	30
NY	0	NH	15	NY	30	AZ	15
OH	15	OH	15	OH	15	CT	15
OK	30	SC	15	OK	45	DE	15
OR	45	SD	15	OR	30	KY	15
PA	0	VT	15	PA	15	MA	15
RI	30	WV	15	RI	30	MD	15
SC	15	AZ	0	SC	30	ME	15
SD	15	DC	0	SD	30	NH	15
TN	30	ID	0	TN	45	NJ	15
UT	30	NC	0	UT	15	OH	15
VA	0	NJ	0	VA	30	PA	15
VT	15	NY	0	VT	45	UT	15
WA	30	PA	0	WA	15	WA	15
WI	30	VA	0	WI	30	WV	15
WV	15	AK	-15	WV	15	WY	0
WY	45	NE	-15	WY	0	HI	-15
TX	60	HI	-60	TX	105	AK	-30

Table 6.3: ***Spring Forward Time Shift (minutes) by State.*** *The temporal shift in (1) peak activity and (2) twinflection sorted alphabetically and by magnitude. Times reported are differences between columns in the preceding table, and reported in minutes.*

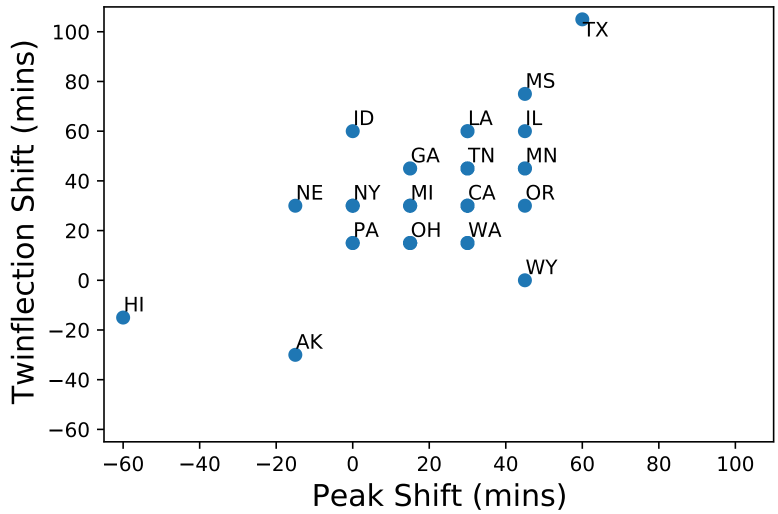

Figure 6.3: ***Correlation of Peak and Twinflection shift estimates.*** *Blue discs represent one or more states having that combination of ordered pair estimates (peak shift, twinflection shift). State abbreviations label each comparison. Given that there is overlap, we label each concurrent point with the state contributing the greatest number of tweets. Table 6.3 reports all states and shifts using each measure. The Pearson correlation of the two measures plotted here is 0.575, while the Spearman rank correlation is 0.467.*

Chapter 7

Supplementary Material for Chapter Three

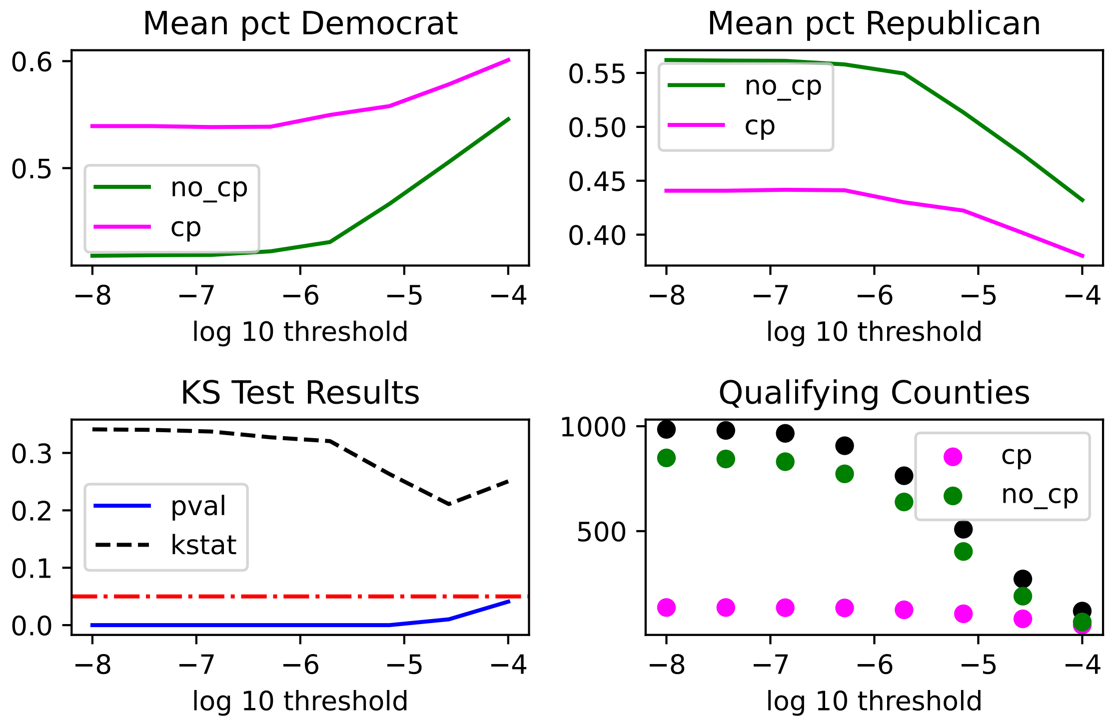

Figure 7.1: ***Plots of the effect of the mean visitation threshold on study results.*** *Top: The mean percent having voted Democrat(left) and Republican (right) in the 2020 Presidential election of the counties with and without change points as the threshold is increased at the log 10 scale. When the threshold is between -8 and -6 the gap in mean vote share between counties with and without abrupt park visitation changes is stable. As the threshold increases past -6 the gap begins to shrink, with the counties with abrupt changes becoming slightly more democrat, and the counties without abrupt changes becoming much more democrat, and both becoming less Republican. Bottom Left: The p-value (blue) and statistic(black dashed) results of the KS 2 sample test on the partisan differences in counties with and without abrupt changes as the threshold increases. The pvalue is stable until the threshold is greater than -5, when it begins to increase, but never crosses the p=0.05 significance threshold (red dashed). The k statistic remains stable until the threshold is increased past -6, when it decreases, but never falls below 0.2. Bottom Right: The number of counties (black) which meet inclusion criteria as the visitation threshold is increased. There is is rapid decline in counties included in the study beginning at a threshold of -6. Past a threshold of -5 fewer than half of all counties in our data set meet inclusion criteria, and at -4 there are almost none. The number of counties with an abrupt visitation change (pink) remains constant in the study until a threshold greater than -5, reflecting that these are among the counties with the greatest visitation. The number of counties without a change (green) declines almost in parallel to the total (black) counties, indicating that the threshold criteria eliminates these counties almost exclusively.*

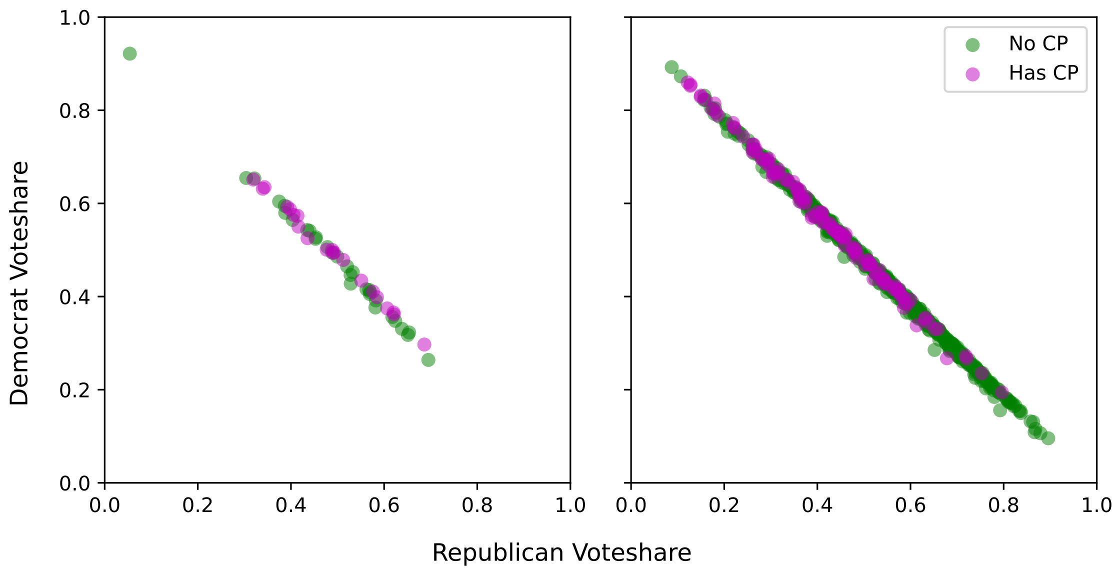

Figure 7.2: ***Scatter plots where each state and county is represented by a dot, the color of which corresponds to whether or not an abrupt chagne took place.*** *The location in the x-y plane is determined by the percent of votes for the Republican(x) and Democratic(y) candidates in the 2020 Presidential Election.*

Bibliography

[1] Plotly Technologies Inc. Collaborative data science. https://plot.ly, 2015.

[2] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[3] Bernard C. K. Choi. The past, present, and future of public health surveillance. *Scientifica*, 2012:875253, 2012.

[4] Niall Boyce. Bills of mortality: tracking disease in early modern london. *The Lancet*, 395(10231):1186–1187, 2020. Publisher: Elsevier.

[5] Ann Elizabeth Fowler La Berge. *Mission and Method: The Early Nineteenth-Century French Public Health Movement*. Cambridge University Press, 2002. Google-Books-ID: SwyGvuMLmiMC.

[6] A brief history of public health. https://sphweb.bumc.bu.edu/otlt/mph-modules/ph/publichealthhistory/publichealthhistory_print.html.

[7] Indicators for chronic disease surveillance. https://www.cdc.gov/mmwr/preview/mmwrhtml/rr5311a1.htm.

[8] BRFSS | CDC. https://www.cdc.gov/chronicdisease/resources/publications/factsheets/brfss.htm.

[9] John Gramlich. 10 facts about americans and facebook. https://www.pewresearch.org/fact-tank/2021/06/01/facts-about-americans-and-facebook/.

[10] Ranganathan Chandrasekaran, Vipanchi Katthula, and Evangelos Moustakas. Patterns of use and key predictors for the use of wearable health care devices by US adults: Insights from a national survey. *Journal of Medical Internet Research*, 22(10):e22443, 2020.

[11] Peter Sheridan Dodds, Kameron Decker Harris, Isabel M Kloumann, Catherine A Bliss, and Christopher M Danforth. Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PLOS ONE*, 6(12), 2011.

[12] Sharon E Alajajian, Jake Ryland Williams, Andrew J Reagan, Stephen C Alajajian, Morgan R Frank, Lewis Mitchell, Jacob Lahne, Christopher M Danforth, and Peter Sheridan Dodds. The Lexicocalorimeter: Gauging public health through caloric input and output on social media. *PLOS ONE*, 12(2), 2017.

[13] Goran Medic, Micheline Wille, and Michiel EH Hemels. Short-and long-term health consequences of sleep disruption. *Nature and Science of Sleep*, 9:151, 2017.

[14] Center for Disease Control. *Short Sleep Duration Among US Adults*, 5 2017.

[15] Michiaki Nagai, Satoshi Hoshide, and Kazuomi Kario. Sleep duration as a risk factor for cardiovascular disease-a review of the recent literature. *Current cardiology reviews*, 6(1):54–61, 2010.

[16] Maggie Fox. Shift work may cause cancer, world agency says. *Reuters*, 2007.

[17] Christopher M. Barnes and Christopher L. Drake. Prioritizing sleep health: Public health policy recommendations. *Perspectives on Psychological Science*, 10(6):733–737, 2015. Publisher: SAGE Publications Inc.

[18] Eugene Leypunskiy, Emre Kıcıman, Mili Shah, Olivia J Walch, Andrey Rzhetsky, Aaron R Dinner, and Michael J Rust. Geographically resolved rhythms in Twitter use reveal social pressures on daily activity patterns. *Current Biology*, 28(23):3763–3775, 2018.

[19] Amneet Sandhu, Milan Seth, and Hitinder S Gurm. Daylight Savings Time and myocardial infarction. *Open Heart*, 1(1):e000019, 2014.

[20] Jussi OT Sipilä, Jori O Ruuskanen, Päivi Rautava, and Ville Kytö. Changes in ischemic stroke occurrence following Daylight Saving Time transitions. *Sleep Medicine*, 27:20–24, 2016.

[21] Jason Varughese and Richard P Allen. Fatal accidents following changes in Daylight Savings Time: The American experience. *Sleep Medicine*, 2(1):31–36, 2001.

[22] José María Martín-Olalla. Traffic accident increase attributed to daylight saving time doubled after energy policy act. *Current Biology*, 30(7):R298–R300, Apr 2020.

[23] 1615 L. St NW, Suite 800 Washington, and DC 20036 USA202-419-4300 {\textbar} Main202-857-8562 {\textbar} Fax202-419-4372 {\textbar} Media Inquiries. Mobile fact sheet. https://www.pewresearch.org/internet/fact-sheet/mobile/.

[24] Samuel Engle, John Stromme, and Anson Zhou. Staying at home: Mobility effects of COVID-19. *SSRN Electronic Journal*, 2020.

[25] William L. Rice and Bing Pan. Understanding changes in park visitation during the COVID-19 pandemic: A spatial application of big data. *Wellbeing, Space and Society*, 2:100037, 2021.

[26] Zoe M. Volenec, Joel O. Abraham, Alexander D. Becker, and Andy P. Dobson. Public parks and the pandemic: How park usage has been affected by COVID-19 policies. *PLOS ONE*, 16(5):e0251799, 2021. Publisher: Public Library of Science.

[27] Kaiguang Zhao, Michael A. Wulder, Tongxi Hu, Ryan Bright, Qiusheng Wu, Haiming Qin, Yang Li, Elizabeth Toman, Bani Mallick, Xuesong Zhang, and Molly Brown. Detecting change-point, trend, and seasonality in satellite time series data to track abrupt changes and nonlinear dynamics: A bayesian ensemble algorithm. *Remote Sensing of Environment*, 232:111181, 2019.

[28] Keith K. Miyake, Andrew R. Maroko, Kristen L. Grady, Juliana A. Maantay, and Peter S. Arno. Not just a walk in the park: Methodological improvements for determining environmental justice implications of park access in new york city for the promotion of physical activity. *Cities and the Environment*, 3(1):1–17, 2010.

[29] B.A. Sandalack, F.G. Alaniz Uribe, A. Eshghzadeh Zanjani, A. Shiell, G.R. McCormack, and P.K. Doyle-Baker. Neighbourhood type and walkshed size. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, 6(3):236–255, 2013. Publisher: Routledge _eprint: https://doi.org/10.1080/17549175.2013.771694.

[30] Zoé A. Hamstead, David Fisher, Rositsa T. Ilieva, Spencer A. Wood, Timon McPhearson, and Peleg Kremer. Geolocated social media as a rapid indicator of park visitation and equitable park access. *Computers, Environment and Urban Systems*, 72:38–50, 2018-11-01.

[31] James Saxon. Empirical measures of park use in american cities, and the demographic biases of spatial models. *Geographical Analysis*, 53(4):665–685, 2021. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/gean.12265.

[32] Lisa Groshong, Sonja A. Wilhelm Stanis, Andrew T. Kaczynski, and J. Aaron Hipp. Attitudes about perceived park safety among residents in low-income and high minority kansas city, missouri, neighborhoods. *Environment and Behavior*, 52(6):639–665, 2020-07-01. Publisher: SAGE Publications Inc.

[33] Lisa Groshong, Sonja A. Wilhelm Stanis, Andrew T. Kaczynski, J. Aaron Hipp, and Gina M. Besenyi. Exploring attitudes, perceived norms, and personal agency: Insights into theory-based messages to encourage park-based physical activity in low-income urban neighborhoods. *Journal of Physical Activity & Health*, 14(2):108–116, 2017-02.

[34] Sandra C Lapham, Deborah A Cohen, Bing Han, Stephanie Williamson, Kelly R Evenson, Thomas L McKenzie, Amy Hillier, and Phillip Ward. How important is perception of safety to park use? a four-city survey. *Urban Studies*, 53(12):2624–2636, 2016. Publisher: Sage Publications, Ltd.

[35] Deborah A. Cohen, Bing Han, Kathryn P. Derose, Stephanie Williamson, Terry Marsh, Laura Raaen, and Thomas L. McKenzie. The paradox of parks in low-income areas: Park use and perceived threats. *Environment and Behavior*, 48(1):230–245, 2016-01-01. Publisher: SAGE Publications Inc.

[36] Stephanie L. Orstad, Kristin Szuhany, Kosuke Tamura, Lorna E. Thorpe, and Melanie Jay. Park proximity and use for physical activity among urban residents: Associations with mental health. *International Journal of Environmental Research and Public Health*, 17(13):4885, 2020-01. Number: 13 Publisher: Multidisciplinary Digital Publishing Institute.

[37] Amir A. Kamel, Paula B. Ford, and Andrew T. Kaczynski. Disparities in park availability, features, and characteristics by social determinants of health within a u.s.-mexico border urban area. *Preventive Medicine*, 69 Suppl 1:S111–113, 2014-12.

[38] Kelsey Linnell, Michael Arnold, Thayer Alshaabi, Thomas McAndrew, Jeanie Lim, Peter Sheridan Dodds, and Christopher M. Danforth. The sleep loss insult of spring daylight savings in the US is observable in twitter activity. *Journal of Big Data*, 8(1):121, 2021.

[39] Consensus Conference Panel, Nathaniel F Watson, M Safwan Badr, Gregory Belenky, Donald L Bliwise, Orfeu M Buxton, Daniel Buysse, David F Dinges,

James Gangwisch, Michael A Grandner, et al. Joint consensus statement of the american academy of sleep medicine and sleep research society on the recommended amount of sleep for a healthy adult: Methodology and discussion. *Sleep*, 38(8):1161–1183, 2015.

[40] Earl S Ford, Timothy J Cunningham, and Janet B Croft. Trends in self-reported sleep duration among us adults from 1985 to 2012. *Sleep*, 38(5):829–832, 2015.

[41] Tim Althoff, Eric Horvitz, Ryen W. White, and Jamie Zeitzer. Harnessing the web for population-scale physiological sensing: A case study of sleep and performance. *WWW '17 Proceedings of the 26th International Conference on World Wide Web*, pages 113–122, 2017.

[42] Giuseppe Curcio, Michele Ferrara, and Luigi De Gennaro. Sleep loss, learning capacity and academic performance. *Sleep Medicine Reviews*, 10(5):323–337, 2006.

[43] Mindy Engle-Friedman. The effects of sleep loss on capacity and effort. *Sleep Science*, 7(4):213–224, 2014.

[44] Mark R Rosekind, Kevin B Gregory, Melissa M Mallis, Summer L Brandt, Brian Seal, and Debra Lerner. The cost of poor sleep: Workplace productivity loss and associated costs. *Journal of Occupational and Environmental Medicine*, 52(1):91–98, 2010.

[45] Bonnie Dean, Daniel Aguilar, Colin Shapiro, William C Orr, Joshua A Isserman, Brian Calimlim, and Gregory A Rippon. Impaired health status, daily functioning, and work productivity in adults with excessive sleepiness. *Journal of Occupational and Environmental Medicine*, 52(2):144–149, 2010.

[46] JM Owens, TA Dingus, F Guo, Y Fang, M Perez, J McClafferty, and B Tefft. Prevalence of drowsy-driving crashes: Estimates from a large-scale naturalistic driving study. *AAA Foundation for Traffic Safety*, 2018.

[47] Clare Anderson, Suzanne Ftouni, Joseph M Ronda, Shantha MW Rajaratnam, Charles A Czeisler, and Steven W Lockley. Self-reported drowsiness and safety outcomes while driving after an extended duration work shift in trainee physicians. *Sleep*, 41(2):zsx195, 2018.

[48] V Cheung, VM Yuen, GTC Wong, and SW Choi. The effect of sleep deprivation and disruption on dna damage and health of doctors. *Anaesthesia*, 74(4):434–440, 2019.

[49] Nirav P Patel, Michael A Grandner, Dawei Xie, Charles C Branas, and Nalaka Gooneratne. "sleep disparity" in the population: Poor sleep quality is strongly associated with poverty and ethnicity. *BMC Public Health*, 10(1):475, 2010.

[50] Vijay Kumar Chattu, Soosanna Kumary Chattu, David Warren Spence, Md Dilshad Manzar, Deepa Burman, and Seithikurippu R Pandi-Perumal. Do disparities in sleep duration among racial and ethnic minorities contribute to differences in disease prevalence? *Journal of Racial and Ethnic Health Disparities*, 6(6):1053–1061, 2019.

[51] Megan E Ruiter, Jamie DeCoster, Lindsey Jacobs, and Kenneth L Lichstein. Normal sleep in African-Americans and Caucasian-Americans: A meta-analysis. *Sleep Medicine*, 12(3):209–214, 2011.

[52] David S Curtis, Thomas E Fuller-Rowell, Mona El-Sheikh, Mercedes R Carnethon, and Carol D Ryff. Habitual sleep as a contributor to racial differences in cardiometabolic risk. *Proceedings of the National Academy of Sciences*, 114(33):8889–8894, 2017.

[53] Marco Hafner, Martin Stepanek, Jirka Taylor, Wendy M Troxel, and Christian Van Stolk. Why sleep matters - the economic costs of insufficient sleep: A cross-country comparative analysis. *Rand Health Quarterly*, 6(4), 2017.

[54] Matt T Bianchi. Sleep devices: Wearables and nearables, informational and interventional, consumer and clinical. *Metabolism*, 84:99–108, 2018.

[55] Neil J Douglas, Stephen Thomas, and Mohammed A Jan. Clinical value of polysomnography. *The Lancet*, 339(8789):347–350, 1992.

[56] Allison G Harvey and Nicole KY Tang. (Mis) perception of sleep in insomnia: A puzzle and a resolution. *Psychological Bulletin*, 138(1):77, 2012.

[57] Diane S Lauderdale, Kristen L Knutson, Lijing L Yan, Kiang Liu, and Paul J Rathouz. Self-reported and measured sleep duration: How similar are they? *Epidemiology*, pages 838–845, 2008.

[58] Miguel Marino, Yi Li, Michael N Rueschman, John W Winkelman, JM Ellenbogen, Jo M Solet, Hilary Dulin, Lisa F Berkman, and Orfeu M Buxton. Measuring sleep: Accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography. *Sleep*, 36(11):1747–1755, 2013.

[59] Sigga Svala Jonasdottir, Kelton Minor, and Sune Lehmann. Gender differences in nighttime sleep patterns and variability across the adult lifespan: a global-scale wearables study. *Sleep*, 44, 2021.

[60] Sirinthip Roomkham, David Lovell, Joseph Cheung, and Dimitri Perrin. Promises and challenges in the use of consumer-grade devices for sleep monitoring. *IEEE Reviews in Biomedical Engineering*, 11:53–67, 2018.

[61] Matthew A Christensen, Laura Bettencourt, Leanne Kaye, Sai T Moturu, Kaylin T Nguyen, Jeffrey E Olgin, Mark J Pletcher, and Gregory M Marcus. Direct measurements of smartphone screen-time: Relationships with demographics and sleep. *PLOS ONE*, 11(11), 2016.

[62] Miguel Rios and Jimmy Lin. Visualizing the Pulse of World Cities on Twitter. *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*, page 4, 2013.

[63] Till Roenneberg. Twitter as a means to study temporal behaviour. *Current Biology*, 27(17):R830–R832, September 2017.

[64] Tatjana Scheffler and Christopher C M Kyba. Measuring Social Jetlag in Twitter Data. *Proceedings of the Tenth International AAAI Conference on Web and Social Media (ICWSM 2016)*, page 4, 2016.

[65] Fabon Dzogang, Stafford Lightman, and Nello Cristianini. Circadian mood variations in Twitter content. *Brain and Neuroscience Advances*, 1:2398212817744501, January 2017. Publisher: SAGE Publications Ltd STM.

[66] Fabon Dzogang, Stafford Lightman, and Nello Cristianini. Diurnal variations of psychometric indicators in Twitter content. *PLOS ONE*, 13(6):e0197002, June 2018.

[67] David J McIver, Jared B Hawkins, Rumi Chunara, Arnaub K Chatterjee, Aman Bhandari, Timothy P Fitzgerald, Sachin H Jain, and John S Brownstein. Characterizing sleep issues using twitter. *Journal of Medical Internet Research*, 17(6):e140, 2015.

[68] José María Martín-Olalla. the long term impact of daylight saving time regulations in daily life at several circles of latitude. *Scientific Reports*, 9(1):1–13, 2019.

[69] José María Martín-Olalla. Scandinavian bed and rise times in the age of enlightenment and in the 21st century show similarity, helped by daylight saving time. *Journal of Sleep Research*, Sep 2019.

[70] Thomas Kantermann, Myriam Juda, Martha Merrow, and Till Roenneberg. The Human Circadian Clock's Seasonal Adjustment Is Disrupted by Daylight Saving Time. *Current Biology*, 17(22):1996–2000, November 2007.

[71] Mark J Kamstra, Lisa A Kramer, and Maurice D Levi. Losing sleep at the market: The Daylight Saving anomaly. *American Economic Review*, 90(4):1005–1011, 2000.

[72] Tyler J Gray, Andrew J Reagan, Peter Sheridan Dodds, and Christopher M Danforth. English verb regularization in books and tweets. *PLOS ONE*, 13(12), 2018.

[73] Twitter. Developer application program interface (API). https://developer.twitter.com/en/docs/tweets/sample-realtime/overview/decahose, 2020.

[74] Carl Edward Rasmussen. Gaussian processes in machine learning. In *Summer School on Machine Learning*, pages 63–71. Springer, 2003.

[75] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12(Oct):2825–2830, 2011.

[76] Stefan Wojcik and Adam Hughes. Sizing up twitter users, 2019.

[77] Ellen DeGeneres. https://twitter.com/theellenshow/status/440322224407314432, March 2, 2014.

[78] Howard Frumkin, Gregory N Bratman, Sara Jo Breslow, Bobby Cochran, Peter H Kahn Jr, Joshua J Lawler, Phillip S Levin, Pooja S Tandon, Usha Varanasi, Kathleen L Wolf, et al. Nature contact and human health: A research agenda. *Environmental health perspectives*, 125(7):075001, 2017.

[79] Ariane L. Bedimo-Rung, Andrew J. Mowen, and Deborah A. Cohen. The significance of parks to physical activity and public health: A conceptual model. *American Journal of Preventive Medicine*, 28(2):159–168, 2005.

[80] Elizabeth Orsega-Smith, Andrew J. Mowen, Laura L. Payne, and Geoffrey Godbey. The interaction of stress and park use on psycho-physiological health in older adults. *Journal of Leisure Research*, 36(2):232–256, 2004. Publisher: Routledge _eprint: https://doi.org/10.1080/00222216.2004.11950021.

[81] Kyung Hee Lee, Jinmoo Heo, Roop Jayaraman, and Shay Dawson. Proximity to parks and natural areas as an environmental determinant to spatial disparities in obesity prevalence. *Applied Geography*, 112:102074, 2019.

[82] Ietza Bojorquez and Lina Ojeda-Revah. Urban public parks and mental health in adult women: Mediating and moderating factors. *International Journal of Social Psychiatry*, 64(7):637–646, 2018. Publisher: SAGE Publications Ltd.

[83] Aaron Reuben, George W. Rutherford, Jameze James, and Nooshin Razani. Association of neighborhood parks with child health in the united states. *Preventive Medicine*, 141:106265, 2020.

[84] Rita Berto. The role of nature in coping with psycho-physiological stress: a literature review on restorativeness. *Behavioral sciences*, 4(4):394–409, 2014.

[85] Roland Sturm and Deborah Cohen. Proximity to urban parks and mental health. *The journal of mental health policy and economics*, 17(1):19–24, 2014.

[86] Yingyi Cheng, Jinguang Zhang, Wei Wei, and Bing Zhao. Effects of urban parks on residents' expressed happiness before and during the COVID-19 pandemic. *Landscape and Urban Planning*, 212:104118, 2021.

[87] Shu Liu and Xinhao Wang. Reexamine the value of urban pocket parks under the impact of the COVID-19. *Urban Forestry & Urban Greening*, 64:127294, 2021.

[88] Sandy J. Slater, Richard W. Christiana, and Jeanette Gustat. Recommendations for keeping parks and green space accessible for mental and physical health during COVID-19 and other pandemics. *Preventing Chronic Disease*, 17:E59, 2020.

[89] Jing Xie, Shixian Luo, Katsunori Furuya, and Dajiang Sun. Urban parks as green buffers during the COVID-19 pandemic. *Sustainability*, 12(17):6751, 2020. Number: 17 Publisher: Multidisciplinary Digital Publishing Institute.

[90] D. Hsiehchen, M. Espinoza, and P. Slovic. Political partisanship and mobility restriction during the COVID-19 pandemic. *Public Health*, 187:111–114, 2020.

[91] Jean-Paul R. Soucy, Shelby L. Sturrock, Isha Berry, Duncan J. Westwood, Nick Daneman, Derek R. MacFadden, and Kevin A. Brown. Estimating effects of physical distancing on the covid-19 pandemic using an urban mobility index. *medRxiv*, 2020.

[92] M. Keith Chen and Ryne Rohla. The effect of partisanship and political advertising on close family ties. *Science*, 360(6392):1020–1024, 2018. _eprint: https://www.science.org/doi/pdf/10.1126/science.aaq1433.

[93] Shana Kushner Gadarian, Sara Wallace Goodman, and Thomas B Pepinsky. Partisanship, health behavior, and policy attitudes in the early stages of the covid-19 pandemic. *Plos one*, 16(4):e0249596, 2021.

[94] Hunt Allcott, Levi Boxell, Jacob Conway, Matthew Gentzkow, Michael Thaler, and David Yang. Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic. *Journal of Public Economics*, 191:104254, 2020.

[95] J. Clinton, J. Cohen, J. Lapinski, and M. Trussler. Partisan pandemic: How partisanship and public health concerns affect individuals’ social mobility during covid-19. *Science Advances*, 7(2):eabd7204, 2021.

[96] Anton Gollwitzer, Cameron Martel, William J. Brady, Philip Pärnamets, Isaac G. Freedman, Eric D. Knowles, and Jay J. Van Bavel. Partisan differences in physical distancing are linked to health outcomes during the COVID-19 pandemic. *Nature Human Behaviour*, 4(11):1186–1197, 2020. Bandiera_abtest: a Cg_type: Nature Research Journals Number: 11 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Human behaviour;Politics and international relations;Viral infection Subject_term_id: human-behaviour;politics-and-international-relations;viral-infection.

[97] Xinyuan Ye. Exploring the relationship between political partisanship and COVID-19 vaccination rate. *Journal of Public Health*, page fdab364, 2021.

[98] Guy Grossman, Soojong Kim, Jonah M. Rexer, and Harsha Thirumurthy. Political partisanship influences behavioral responses to governors' recommendations for COVID-19 prevention in the united states. *Proceedings of the National Academy of Sciences*, 117(39):24144–24153, 2020. Publisher: National Academy of Sciences Section: Social Sciences.

[99] Dehui (Christina) Geng, John Innes, Wanli Wu, and Guangyu Wang. Impacts of COVID-19 pandemic on urban park visitation: a global analysis. *Journal of Forestry Research*, 32(2):553–567, 2021.

[100] Bianca Lopez, Christopher Kennedy, and Timon McPhearson. Parks are critical urban infrastructure: Perception and use of urban green spaces in NYC during COVID-19. Publisher: Preprints, 2020.

[101] James G. Gimpel, Nathan Lovin, Bryant Moy, and Andrew Reeves. The urban–rural gulf in american political behavior. *Political Behavior*, 42(4):1343–1368, 2020.

[102] Gregory S Macfarlane, Nico Boyd, John E Taylor, and Kari Watkins. Modeling the impacts of park access on health outcomes: A utility-based accessibility approach. *Environment and Planning B: Urban Analytics and City Science*, 48(8):2289–2306, 2021-10-01. Publisher: SAGE Publications Ltd STM.

[103] James H. Stark, Kathryn Neckerman, Gina S. Lovasi, James Quinn, Christopher C. Weiss, Michael D. M. Bader, Kevin Konty, Tiffany G. Harris, and Andrew Rundle. The impact of neighborhood park access and quality on body mass index among adults in new york city. *Preventive Medicine*, 64:63–68, 2014-07-01.

[104] Jenny J. Roe, Catharine Ward Thompson, Peter A. Aspinall, Mark J. Brewer, Elizabeth I. Duff, David Miller, Richard Mitchell, and Angela Clow. Green space and stress: Evidence from cortisol measures in deprived urban communities. *International Journal of Environmental Research and Public Health*, 10(9):4086–4103, 2013-09. Number: 9 Publisher: Multidisciplinary Digital Publishing Institute.

[105] Gregory N. Bratman, J. Paul Hamilton, and Gretchen C. Daily. The impacts of nature experience on human cognitive function and mental health. *Annals of the New York Academy of Sciences*, 1249(1):118–136, 2012. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1749-6632.2011.06400.x.

[106] Dajun Dai. Racial/ethnic and socioeconomic disparities in urban green space accessibility: Where to intervene? *Landscape and Urban Planning*, 102(4):234–244, 2011-09-30.

[107] Alessandro Rigolon. A complex landscape of inequity in access to urban parks: A literature review. *Landscape and Urban Planning*, 153:160–169, 2016-09-01.

[108] Alessandro Rigolon, Matthew Browning, and Viniece Jennings. Inequities in the quality of urban park systems: An environmental justice investigation of cities in the united states. *Landscape and Urban Planning*, 178:156–169, 2018.

[109] Jennifer Wolch, John P. Wilson, and Jed Fehrenbach. Parks and park funding in los angeles: An equity-mapping analysis. *Urban Geography*, 26(1):4–35, 2005-02-01. Publisher: Routledge _eprint: https://doi.org/10.2747/0272-3638.26.1.4.

[110] Xiaolu Zhou and Jinki Kim. Social disparities in tree canopy and park accessibility: A case study of six cities in illinois using GIS and remote sensing. *Urban Forestry & Urban Greening*, 12(1):88–97, 2013-01-01.

[111] Alessandro Rigolon. Parks and young people: An environmental justice study of park proximity, acreage, and quality in denver, colorado. *Landscape Planning*, 165:73–83, 2017-09.

[112] Jennifer R. Wolch, Jason Byrne, and Joshua P. Newell. Urban green space, public health, and environmental justice: The challenge of making cities 'just green enough'. *Landscape and Urban Planning*, 125:234–244, 2014-05.

[113] Katherine B. Vaughan, Andrew T. Kaczynski, Sonja A. Wilhelm Stanis, Gina M. Besenyi, Ryan Bergstrom, and Katie M. Heinrich. Exploring the distribution of park availability, features, and quality across kansas city, missouri by income and race/ethnicity: an environmental justice investigation. *Annals of Behavioral Medicine*, 45:S28–S38, 2013-02-01.

[114] Dong-ah Choi, Keunhyun Park, and Alessandro Rigolon. From XS to XL urban nature: examining access to different types of green space using a 'just sustainabilities' framework. *Sustainability*, 12(17):6998, 2020. Publisher: Multidisciplinary Digital Publishing Institute.

[115] Dexter H. Locke, Billy Hall, J. Morgan Grove, Steward T. A. Pickett, Laura A. Ogden, Carissa Aoki, Christopher G. Boone, and Jarlath P. M. O'Neil-Dunne. Residential housing segregation and urban tree canopy in 37 US cities. *npj Urban Sustainability*, 1(1):15, 2021.

[116] Deborah A. Cohen, Thomas L. McKenzie, Amber Sehgal, Stephanie Williamson, Daniela Golinelli, and Nicole Lurie. Contribution of public parks to physical activity. *American Journal of Public Health*, 97(3):509–514, 2007-03-01. Publisher: American Public Health Association.

[117] Jing-Huei Huang, J. Aaron Hipp, Oriol Marquet, Claudia Alberico, Dustin Fry, Elizabeth Mazak, Gina S. Lovasi, Whitney R. Robinson, and Myron F. Floyd. Neighborhood characteristics associated with park use and park-based physical activity among children in low-income diverse neighborhoods in new york city. *Preventive Medicine*, 131:105948, 2020-02-01.

[118] Megan Knapp, Jeanette Gustat, Revonda Darensbourg, Leann Myers, and Carolyn Johnson. The relationships between park quality, park usage, and levels of physical activity in low-income, african american neighborhoods. *International Journal of Environmental Research and Public Health*, 16(1):85, 2019-01. Number: 1 Publisher: Multidisciplinary Digital Publishing Institute.

[119] Sharon E. Alajajian, Jake Ryland Williams, Andrew J. Reagan, Stephen C. Alajajian, Morgan R. Frank, Lewis Mitchell, Jacob Lahne, Christopher M. Danforth, and Peter Sheridan Dodds. The lexicocalorimeter: Gauging public health through caloric input and output on social media. *PLOS ONE*, 12(2):e0168893, 2017. Publisher: Public Library of Science.

[120] Hon Yuen and Gavin Jenkins. Changes in well-being immediately after urban park visit. *The American Journal of Occupational Therapy*, 74(4):7411515403p1, 2020-08-01.

[121] Aaron J. Schwartz, Peter Sheridan Dodds, Jarlath P. M. O'Neil-Dunne, Christopher M. Danforth, and Taylor H. Ricketts. Visitors to urban greenspace have higher sentiment and lower negativity on twitter. *People and Nature*, 1(4):476–485, 2019. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pan3.10045.

[122] Ching-Hua Ho, Vinod Sasidharan, William Elmendorf, Fern K. Willits, Alan Graefe, and Geoffrey Godbey. Gender and ethnic variations in urban park preferences, visitation, and perceived benefits. *Journal of Leisure Research*, 37(3):281–306, 2005-09-01. Publisher: Routledge _eprint: https://doi.org/10.1080/00222216.2005.11950054.

[123] Dong Wang, Gregory Brown, Yan Liu, and Iderlina Mateo-Babiano. A comparison of perceived and geographic access to predict urban park use. *Cities*, 42:85–96, 2015-02-01.

[124] Elise Rivera, Anna Timperio, Venurs HY Loh, Benedicte Deforche, and Jenny Veitch. Important park features for encouraging park visitation, physical activity and social interaction among adolescents: A conjoint analysis. *Health & Place*, 70:102617, 2021-07-01.

[125] Henrikki Tenkanen, Enrico Di Minin, Vuokko Heikinheimo, Anna Hausmann, Marna Herbst, Liisa Kajala, and Tuuli Toivonen. Instagram, flickr, or twitter: Assessing the usability of social media data for visitor monitoring in protected areas. *Scientific Reports*, 7(1):17615, 2017-12-14. Bandiera_abtest: a Cc_license_type: cc_by Cg_type: Nature Research Journals Number: 1 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Biogeography;Conservation biology Subject_term_id: biogeography;conservation.

[126] Marie L. Donahue, Bonnie L. Keeler, Spencer A. Wood, David M. Fisher, Zoé A. Hamstead, and Timon McPhearson. Using social media to understand drivers

of urban park visitation in the twin cities, MN. *Landscape and Urban Planning*, 175:1–10, 2018-07-01.

[127] Aaron J. Schwartz, Peter Sheridan Dodds, Jarlath P. M. O'Neil-Dunne, Taylor H. Ricketts, and Christopher M. Danforth. Gauging the happiness benefit of US urban parks through twitter. *PLOS ONE*, 17(3):e0261056, 2022. Publisher: Public Library of Science.

[128] Understanding mobile location data – UM. https://um.co/portfolio/understanding-mobile-location-data/. Accessed: 2022-03-11.

[129] OpenStreetMap contributors. Planet dump retrieved from https://planet.osm.org . https://www.openstreetmap.org, 2017.