Using Web-Sourced Data to Estimate Demand for Climate Amenities

Bryan Parthum

Camp Resources

Asheville, NC

Tuesday, August 13th, 2019

ILLINOIS
Agricultural &
Consumer Economics
COLLEGE OF AGRICULTURAL, CONSUMER & ENVIRONMENTAL SCIENCES

BDEEP
Big Data in Environmental Economics and Policy Research Group

"Web scraping, or scraping, is a computer software technique of extracting information from the internet, usually transforming unstructured data on the web into structured data that can be stored and analyzed in a central database."

- distilnetworks.com

Bryan Parthum
Parthum2@illinois.edu

Bots
(1,000 RA's)

- Crawlers
  - From a given starting point, navigate (and record) every forward url on the page
    (RA's: create list of relevant literature)

- Scrapers
  - Collect information from a page
    (RA's: summarize literature and main findings)

- Interactive
  - Everything from navigating a page to put information into cells, to having a 'conversation' with a human.
    (RA's: write the first draft of your paper - without co-authorship)

Bryan Parthum
Parthum2@illinois.edu

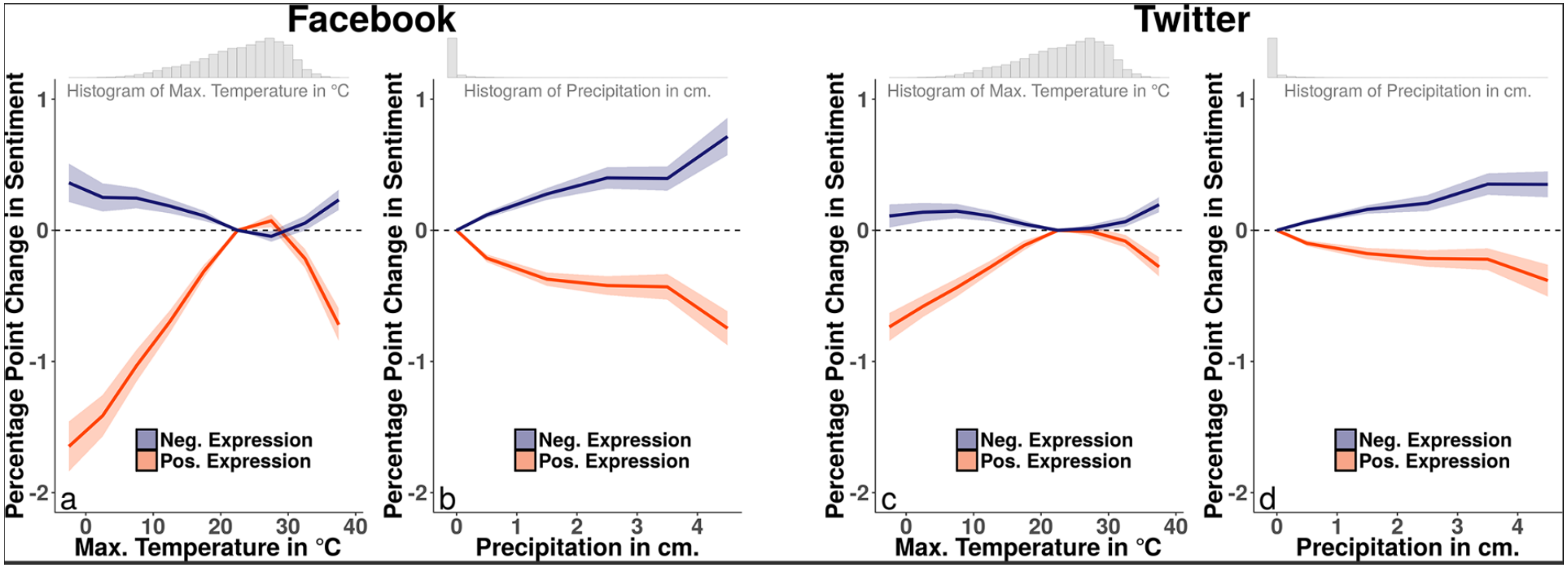
Rules

- The web is public domain (kind of, for now)
    - Legal precedent is mixed
        - (eBay v. Bidder's Edge; Intel v. Hamidi; AP v. Meltwater; LinkedIn v. hiQ)
    - Intellectual property, proper citing, etc. etc. still apply
    - If you're curious about IRB approval, then you should get IRB approval

- Site owners (or providers) can shut down excess traffic
    - Bots create lots of traffic (~43% of internet traffic!?)
    - Keep this in mind when designing a bot. Space out timing of requests, etc.

Bryan Parthum
Parthum2@illinois.edu

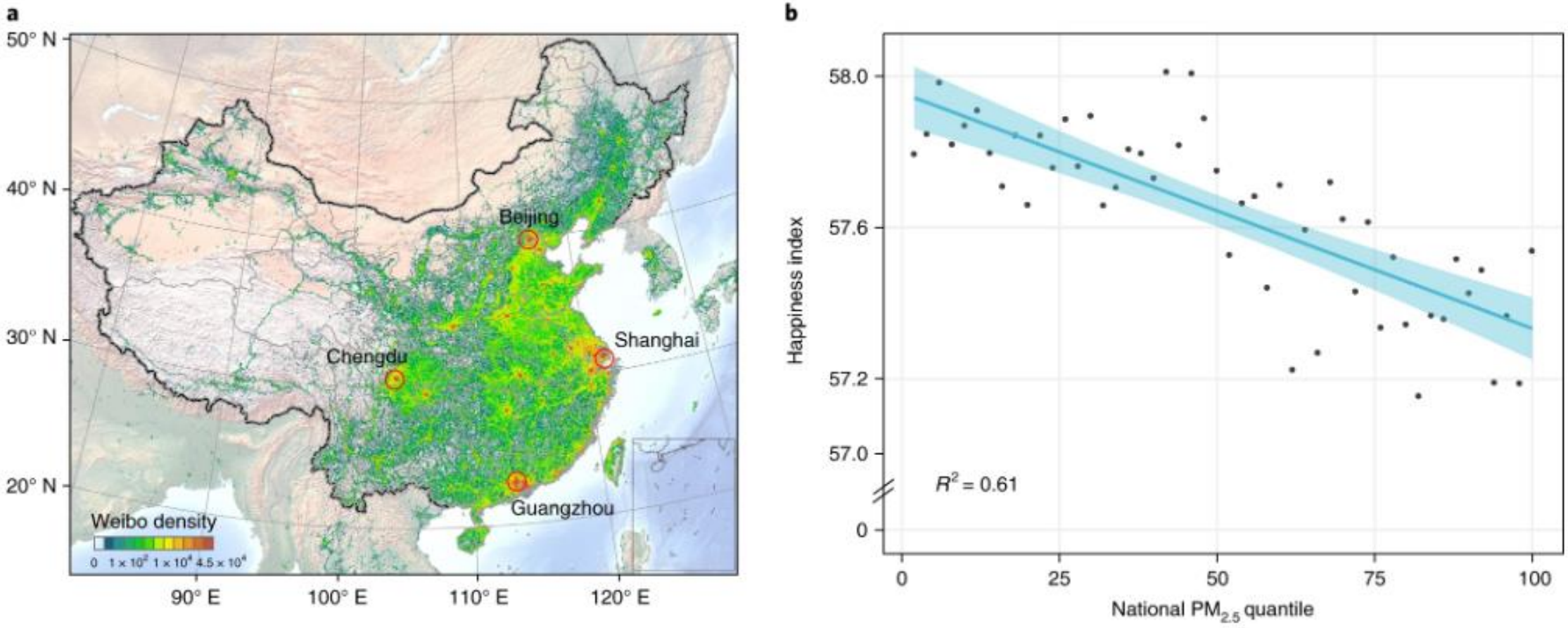
Application: API (application programming interface)



- Baylis et al. *Weather impacts expressed sentiment.* 2018. PloS One.
- Kramer, Guillory, Hancock. *Experimental evidence of massive-scale emotional contagion through social networks*. 2014. PNAS.

Bryan Parthum
Parthum2@illinois.edu

Application: API (application programming interface)



- Zheng et al. *Air pollution lowers Chinese urbanites' expressed happiness on social media.* 2019. Nature Human Behavior.

Bryan Parthum
Parthum2@illinois.edu

Application: API (application programming interface)

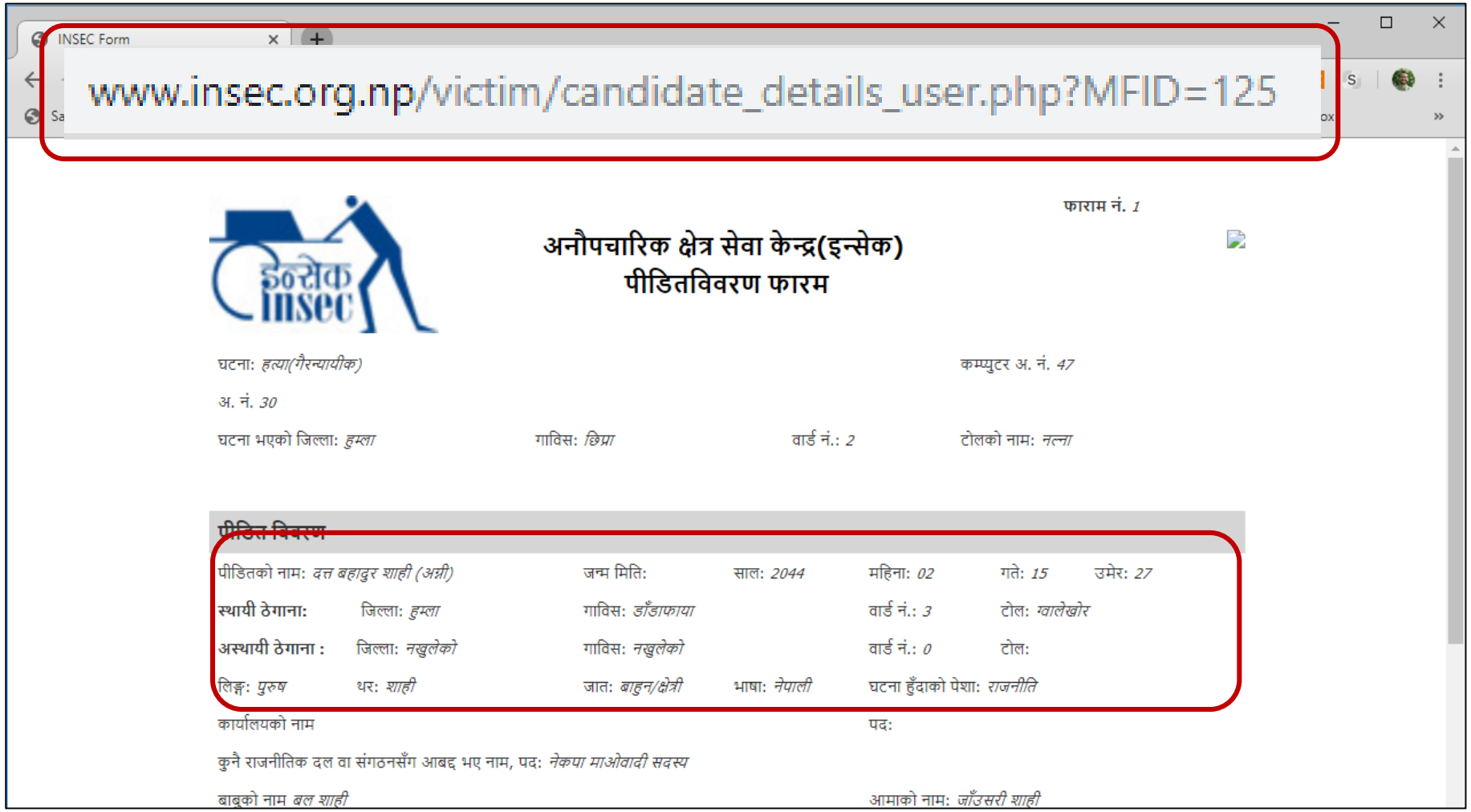
```
@author: parthum2@illinois.edu
"""

## GET TWEETS

import tweepy
import csv

ckey    = 'ApZ6PJFWxmX3C3w35leJREmIV'
csecret = 'KxrmUr1T0hOzv3keXbv6D3u179DV'
atoken  = '3401879230tsLuSzgda0ZfJ2l1WYBg0'
asecret = 'VIVs8KTr2diklsh07XdkWP132eB1Jg4SU'

auth = tweepy.OAuthHandler(ckey, csecret)
auth.set_access_token (atoken, asecret)

api = tweepy.API(auth, wait_on_rate_limit=True,wait_on_rate_limit_notify=True)

if (not api):print ("Can't Authenticate")sys.exit(-1)

# this is what we're searching for
searchQuery = 'Camp Resources'
latlongs = '38.9072,-77.0369,20mi'      # location of the tweets
language = 'en'                         # language of tweets
searchtype = 'recent'                   # type of tweet search, could be recent, popular, or mixed
entitiesin = 'false'                    # include or not some entities such as videos, images, urls in the tweets
maxTweets = 10000                       # Some arbitrary large number
tweetsPerQry = 1000                     # this is the max the API permits
fName = 'Camp_Resources_tweets.csv'     # We'll store the tweets in a csv file.
```


Bryan Parthum
Parthum2@illinois.edu

#CampResources2019

Bryan Parthum
Parthum2@illinois.edu

Application: Scraping



- Phadera, L. *Unfortunate Moms and Unfortunate Children: Impact of the Nepali Civil War on Women's Stature and Intergenerational Health.* 2019. World Bank Group. (scraper built by Mateus Souza)

Bryan Parthum
Parthum2@illinois.edu

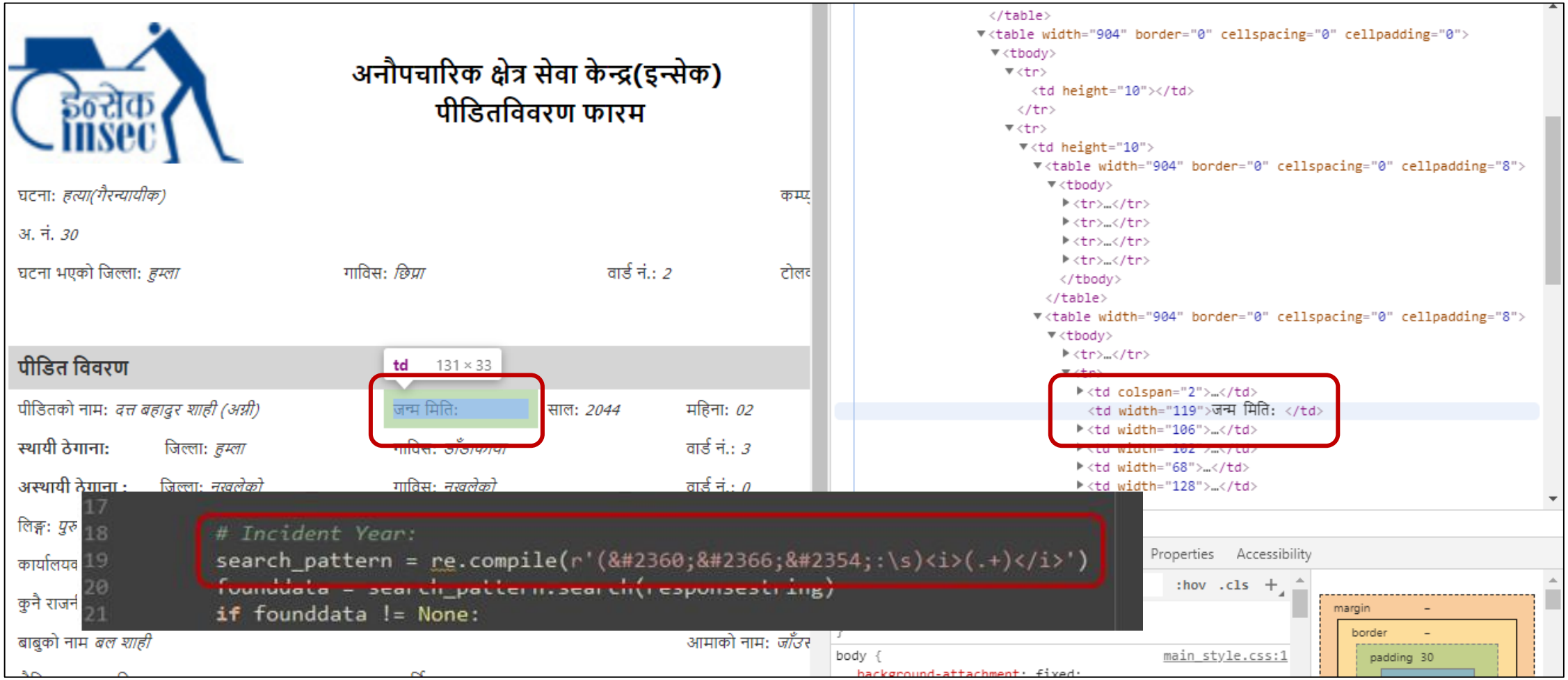
Application: Scraping

www.insec.org.np/victim/candidate_details_user.php?MFID=125

```
fullfile = [['mfid', 'incident_year', 'computer_id', 'incident_district', 'incident_vdc', 'incident_ward',

for person_id in range(0, 10802):
    current_record = []

    # Requesting the data looping over several person IDs
    url = "http://www.insec.org.np/victim/candidate_details_user.php?MFID=" + str(person_id)
    response = requests.request("GET", url)

    response.encoding = "UTF-8"
    responsestring = response.text
    print(person_id)
    if len(responsestring)>20000:

        # MFID:
        current_record.append(str(person_id))

        # Incident Year:
        search_pattern = re.compile(r'(&#2360;&#2366;&#2354;:\s)<i>(.+)</i>')
        founddata = search_pattern.search(responsestring)
        if founddata != None:
            wanteddata = founddata.groups()[1]
            current_record.append(wanteddata.strip())
        else: current_record.append('NA')
```

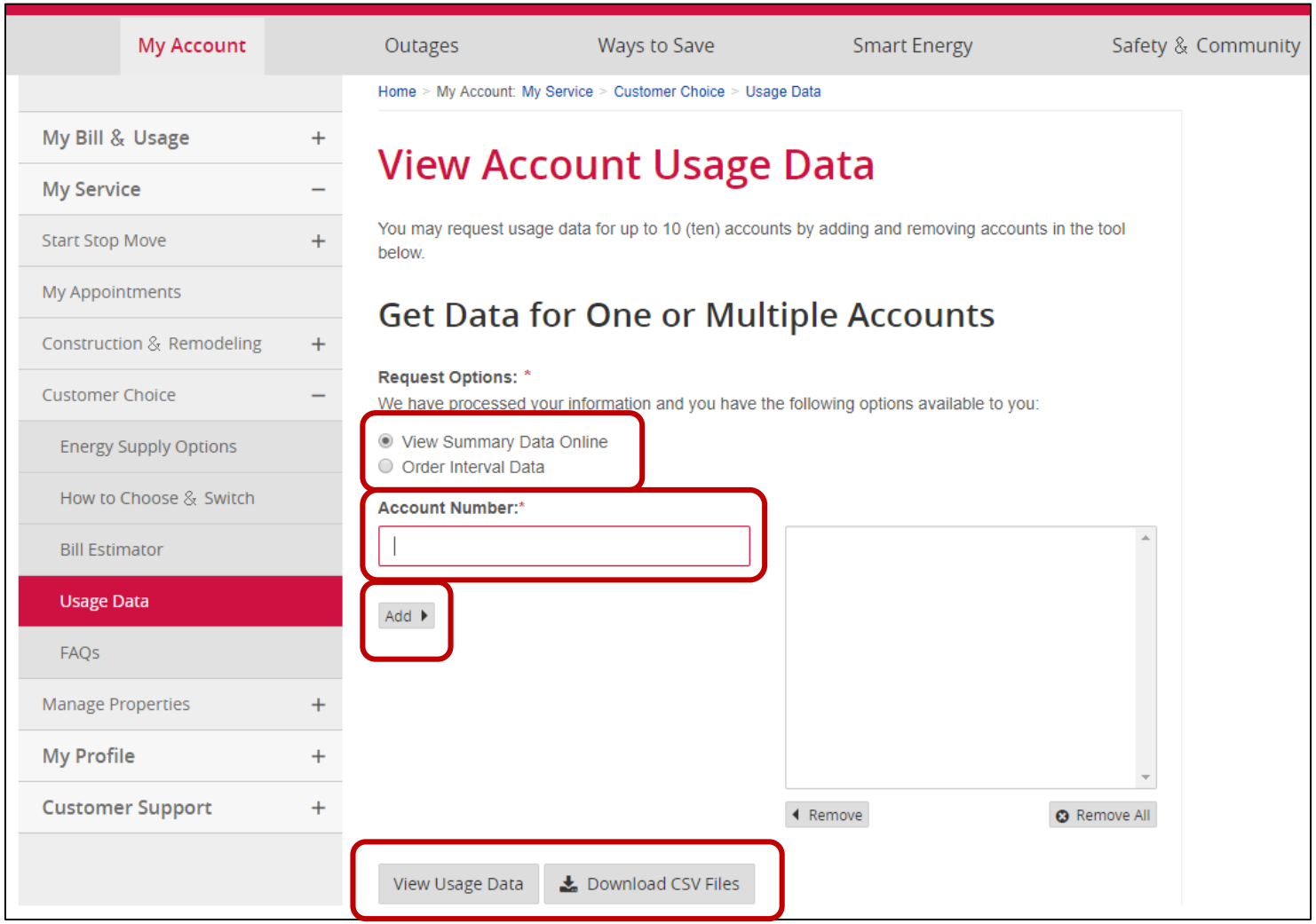

Bryan Parthum
Parthum2@illinois.edu

Application: Scraping



Bryan Parthum
Parthum2@illinois.edu

Application: Active Scraping

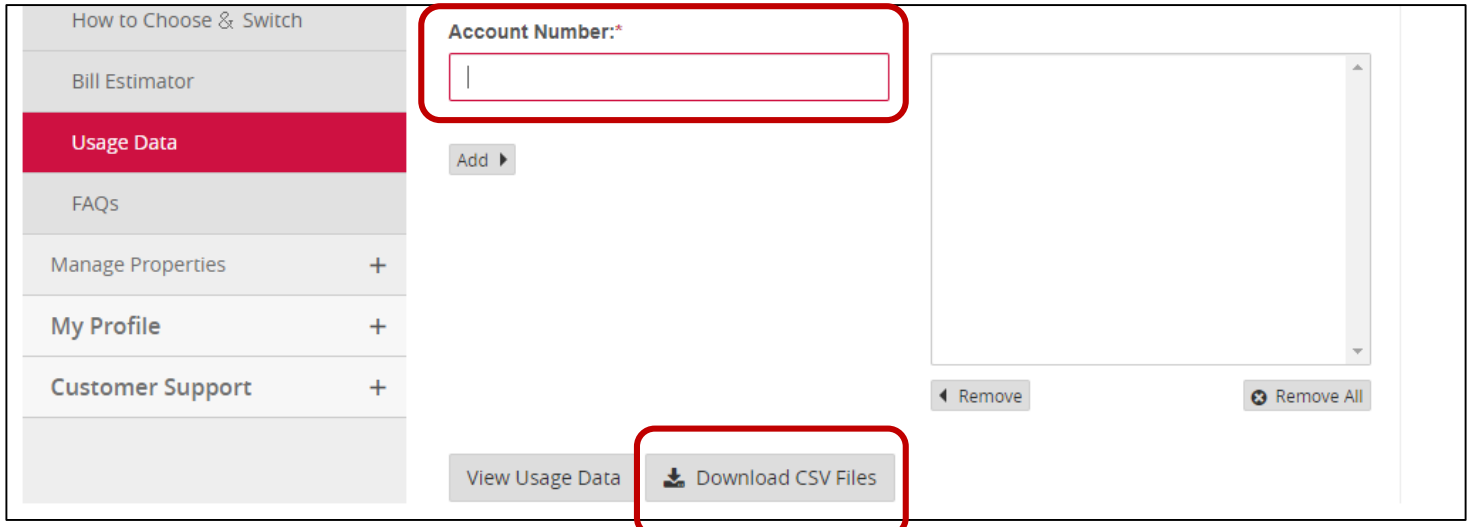


- Myers et al. *Decomposing the Wedge: Evidence from a Home Weatherization Program.*

Bryan Parthum
Parthum2@illinois.edu

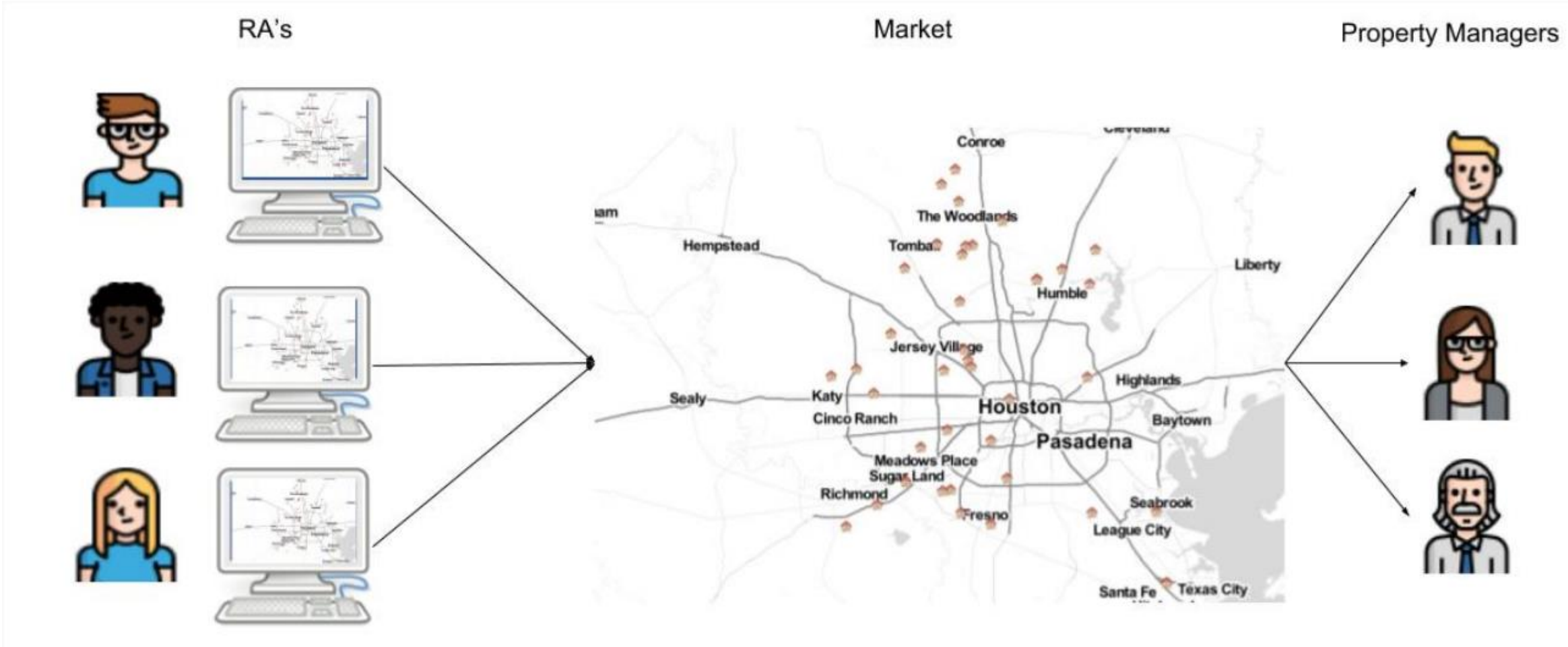
Application: Active Scraping

```
        try:
            textField = driver.find_element_by_id("AccountNumber")
            textField.send_keys(accnNum)
            addButton = driver.find_element_by_id("ctl00_ctl00_ContentPlaceHolderMain_FeaturedContentZone_ctl00_ctl00_Addbtn")
            addButton.click()
        except:
            print("Add Account Fail")
            driver.quit()
        try:
            downloadFile = driver.find_element_by_id("ctl00_ctl00_ContentPlaceHolderMain_FeaturedContentZone_ctl00_ctl00_Downloadbtn")
            downloadFile.click()
        except:
            print("Download File Fail")
            driver.quit()
        time.sleep(1.0)
        driver.quit()
```



Bryan Parthum
Parthum2@illinois.edu

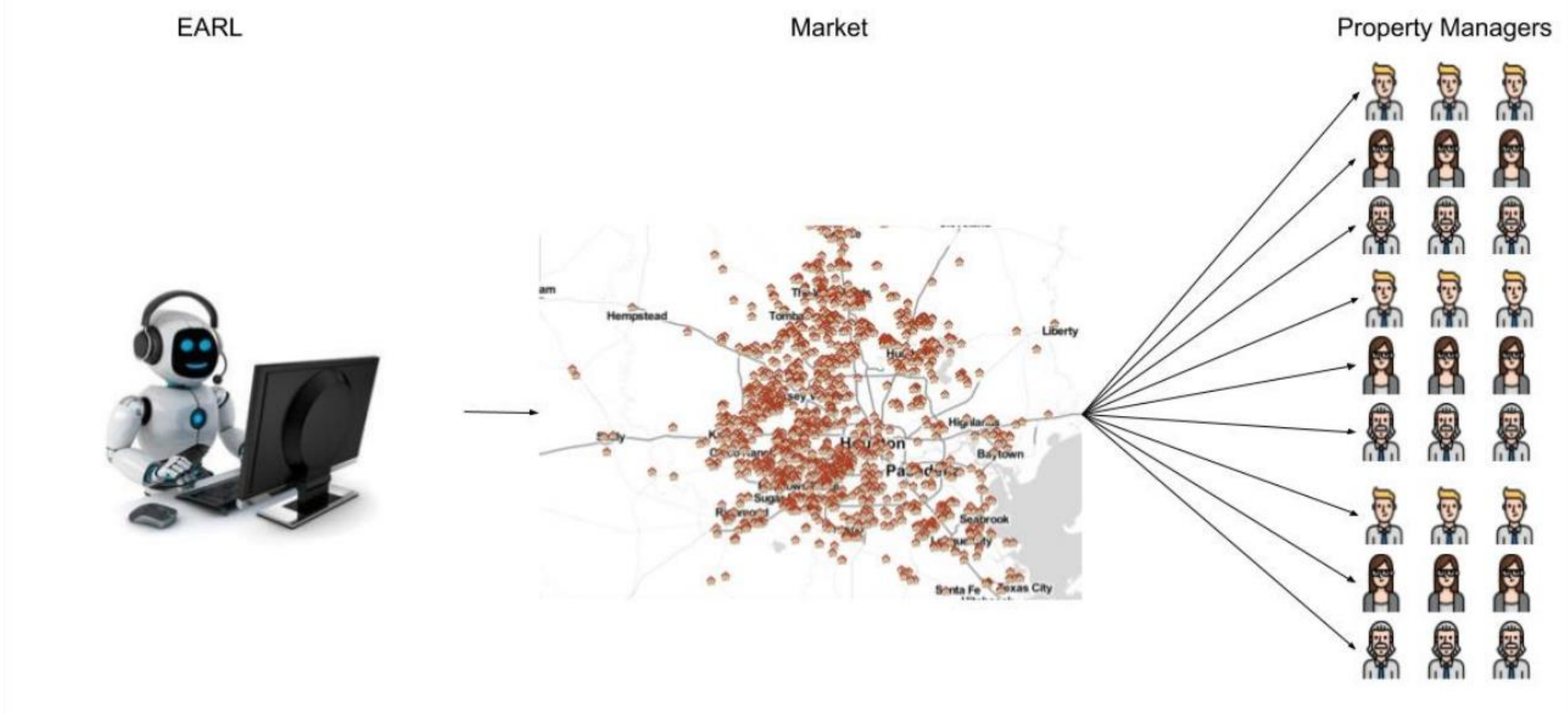
Application: Interactive



- Christensen and Timmins. *The Welfare Effect of Racial Discrimination in the US Rental Market for Housing.*

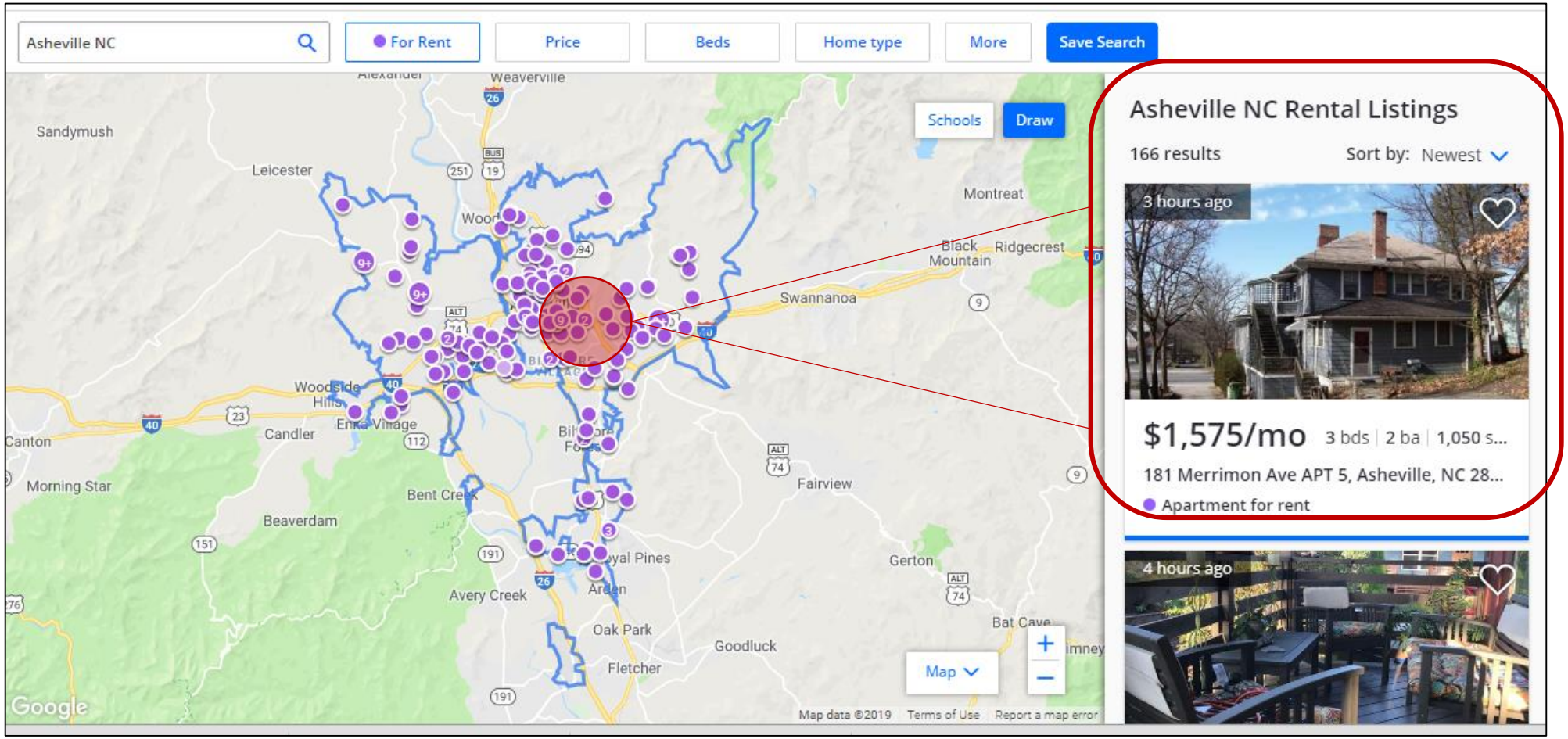
Bryan Parthum
Parthum2@illinois.edu

Application: Interactive



- Christensen and Timmins. *The Welfare Effect of Racial Discrimination in the US Rental Market for Housing.*

Application: Interactive



Bryan Parthum
Parthum2@illinois.edu

Application: Interactive

$1,575/mo 3 bd | 2 ba | 1,050 sqft

181 Merrimon Ave APT 5, Asheville, NC 28801

Apartment for rent

Request to apply | Request a tour

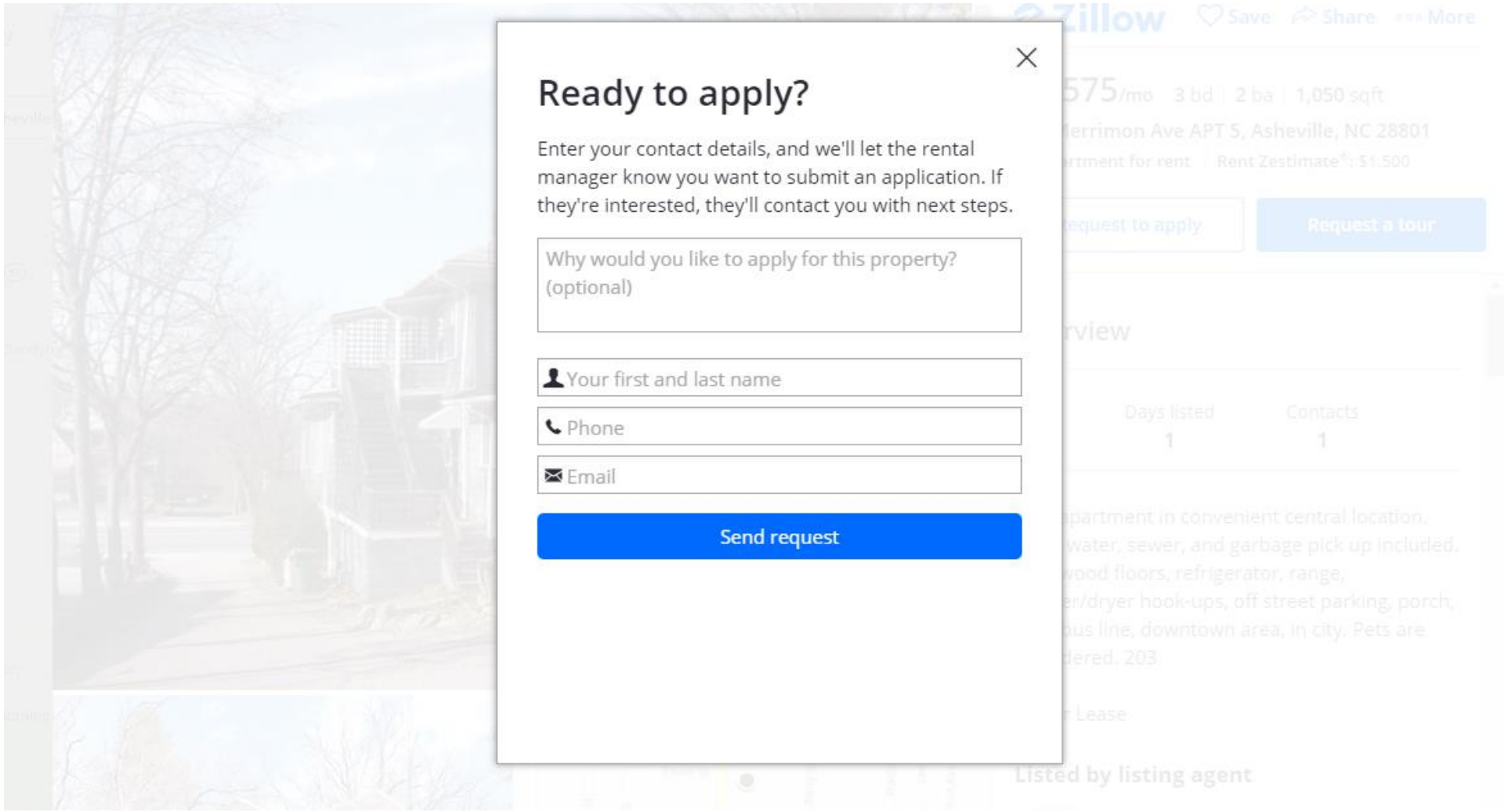
Overview

Days listed	Contacts
1	1

Nice apartment in convenient central location. Heat, water, sewer, and garbage pick up included. Hardwood floors, refrigerator, range, washer/dryer hook-ups, off street parking, porch, near bus line, downtown area, in city. Pets are considered. 203

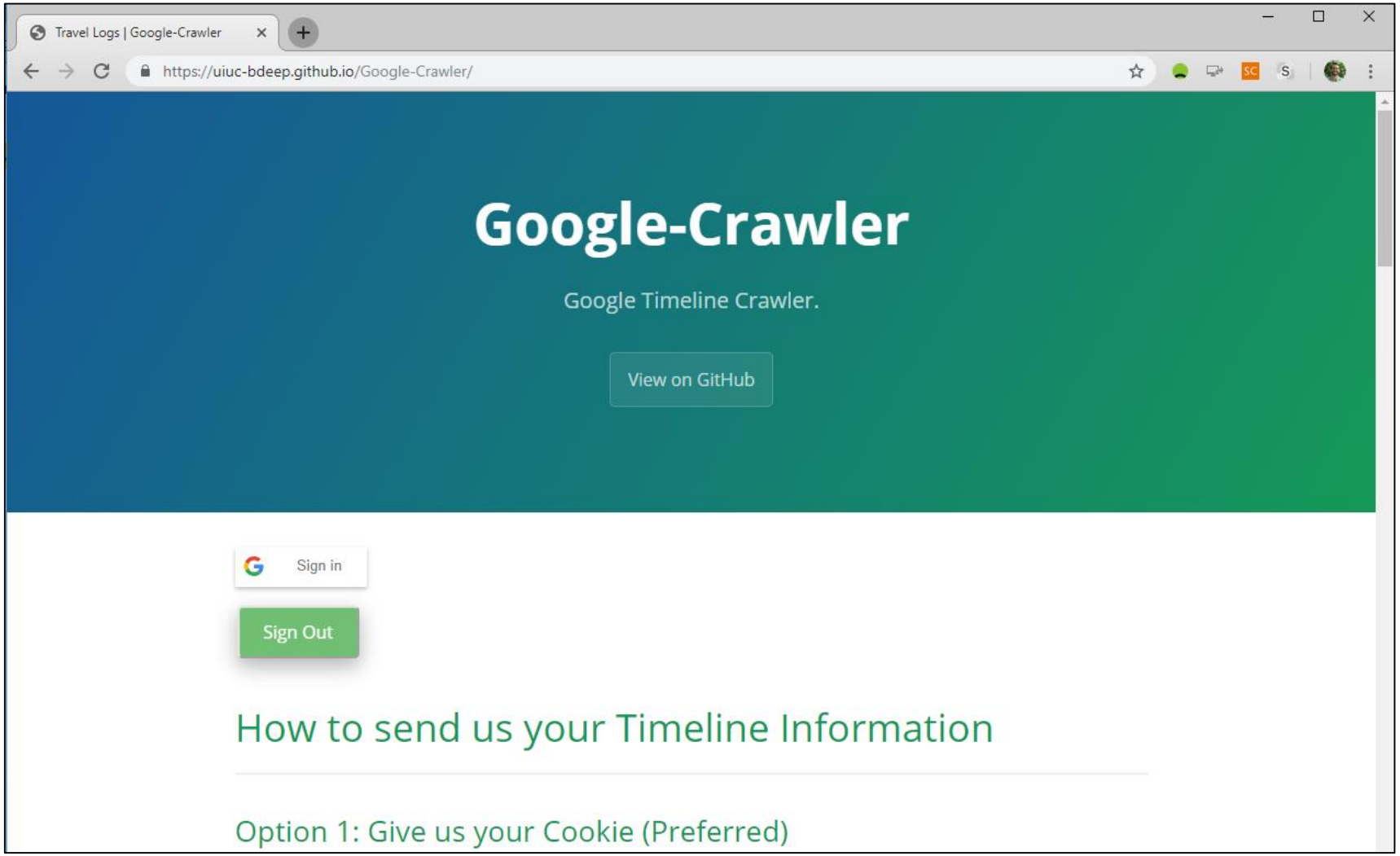
Bryan Parthum
Parthum2@illinois.edu

Application: Interactive



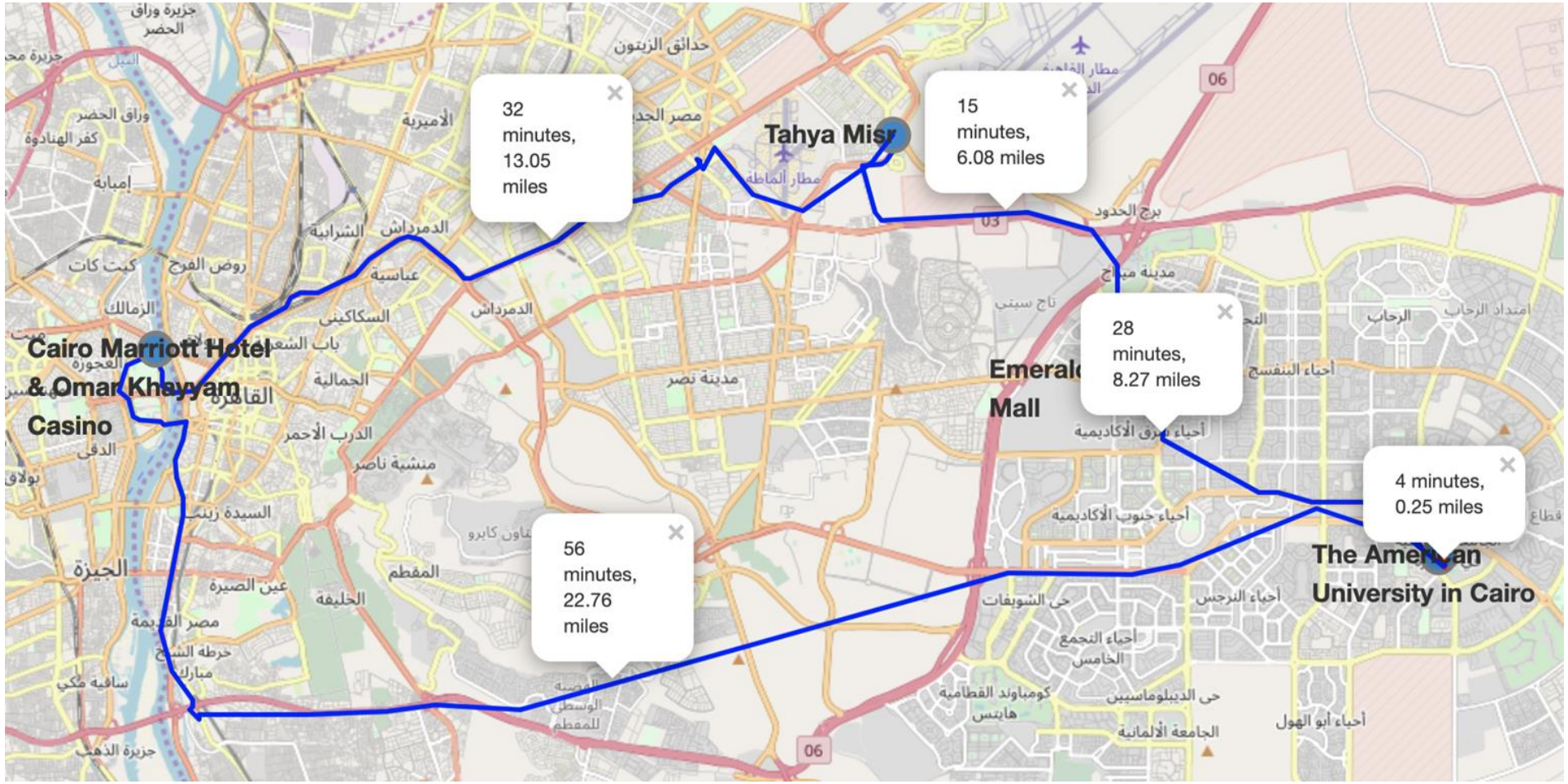
Bryan Parthum
Parthum2@illinois.edu

Application: Tracking



- Christensen and Osman. *The Demand for Mobility in a Transit-Constrained City: A Field Experiment with Uber in Cairo.*

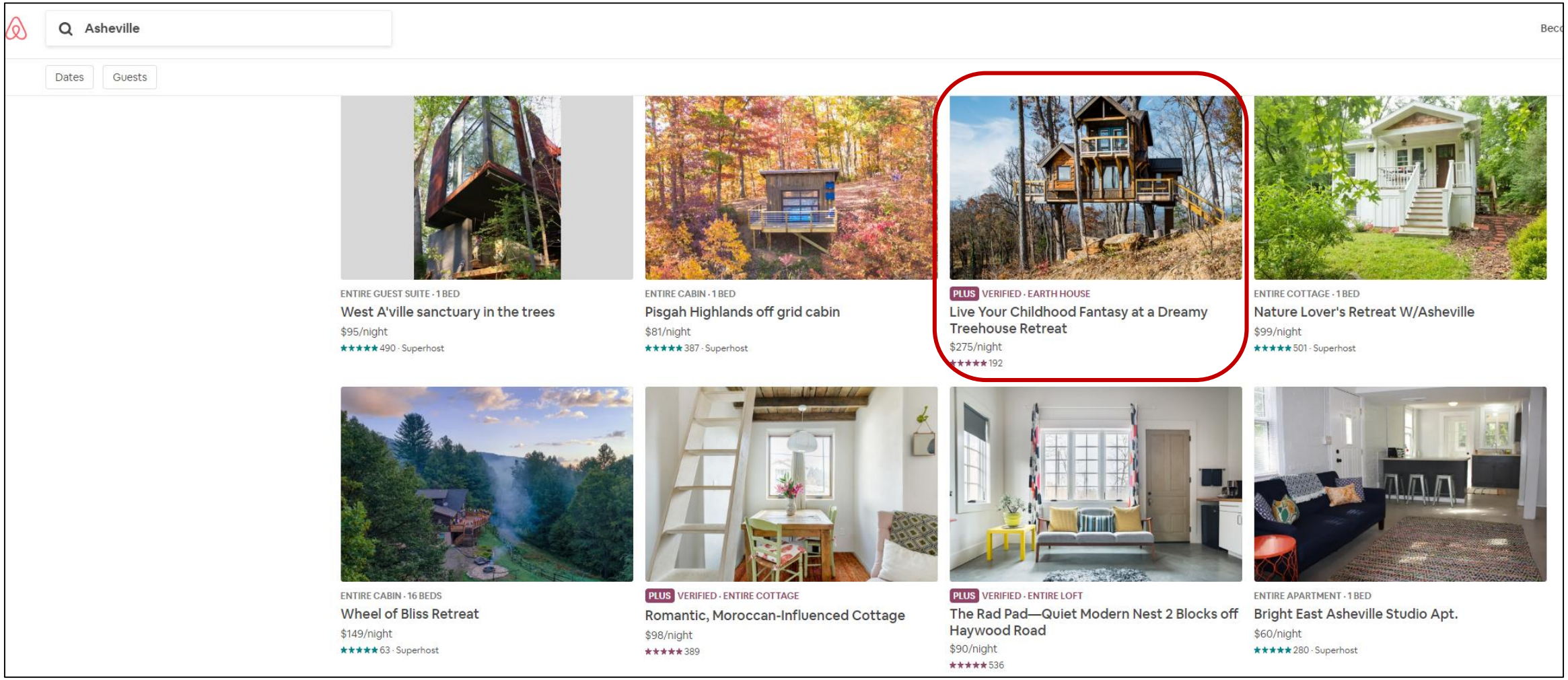
Application: Tracking



- Christensen and Osman. *The Demand for Mobility in a Transit-Constrained City: A Field Experiment with Uber in Cairo.*

Bryan Parthum
Parthum2@illinois.edu

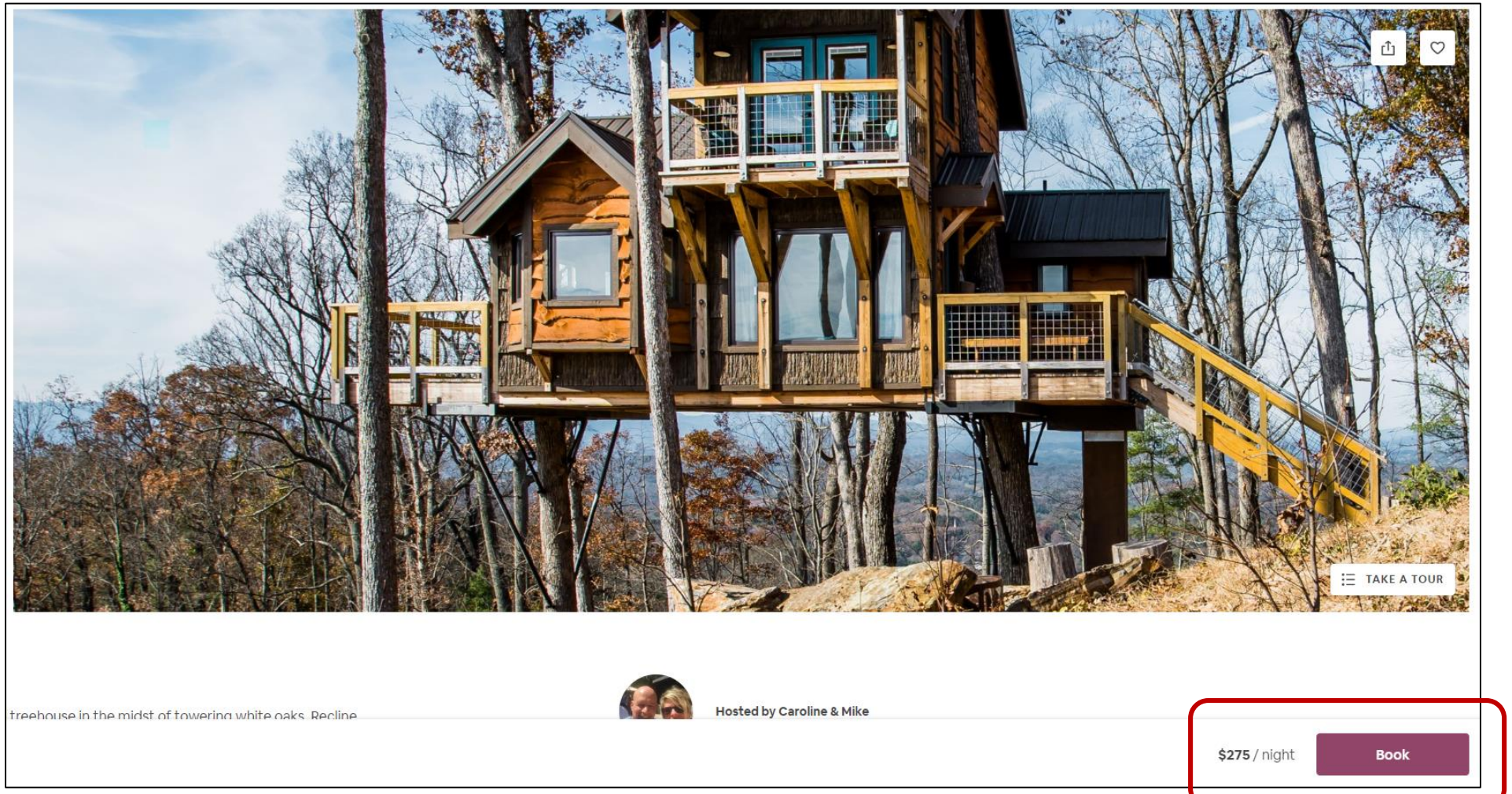
Application: Dynamic Scraping



- Parthum and Christensen. *The Price of Powder: Evidence on the Demand for Snow from Short Term Property Rentals.* Scraping done by *Airdna.co*

Bryan Parthum
Parthum2@illinois.edu

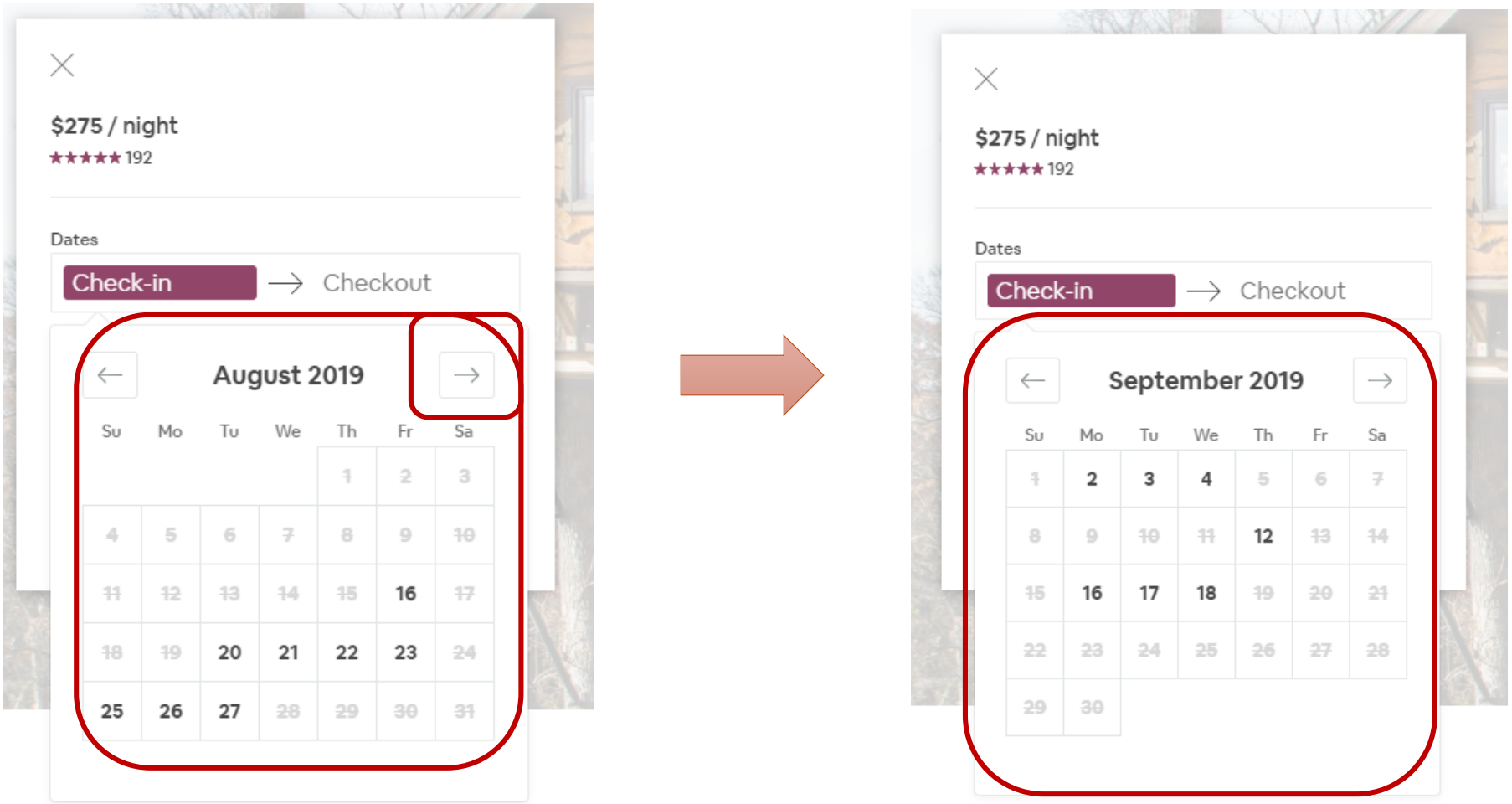
Application: Dynamic Scraping



- Parthum and Christensen. *The Price of Powder: Evidence on the Demand for Snow from Short Term Property Rentals.*

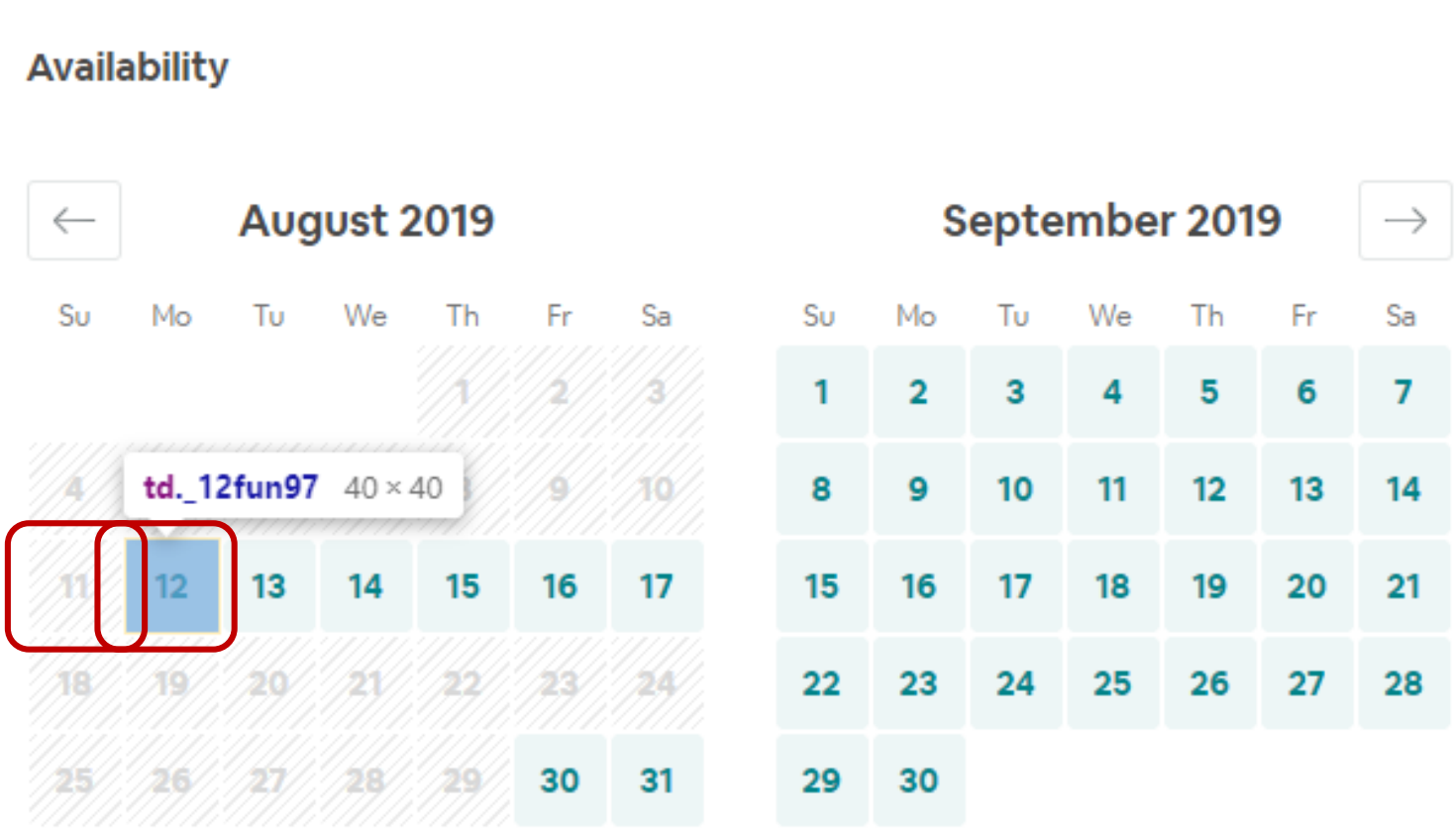
Bryan Parthum
Parthum2@illinois.edu

Application: Dynamic Scraping



- Parthum and Christensen. *The Price of Powder: Evidence on the Demand for Snow from Short Term Property Rentals.*

Application: Dynamic Scraping



- Parthum and Christensen. *The Price of Powder: Evidence on the Demand for Snow from Short Term Property Rentals.*

Bryan Parthum
Parthum2@illinois.edu

Application: Dynamic Scraping

https://www.onthesnow.com/british-columbia/whistler-blackcomb/historical-snowfall.html?y=2018&q=top

RESORTS REGIONS EXPLORE SPRING/SUMMER

Whistler Blackcomb Historical Snowfall CLOSED
https://www.whistlerblackcomb.com
1-888-583-8058

Compare ★★★★☆ 184 Resort Reviews Favorites

Related Regions: British Columbia, Canada, Vancouver, North America, Ski Canada

MOUNTAIN CONDITIONS PLAN EXPLORE

WHISTLER BLACKCOMB SNOW HISTORY

How much snow did Whistler Blackcomb get this ski season and over the last several seasons? Check out the ski resort historical snowfall using the graph, Whistler Blackcomb snow stats—such as Total Snowfall, Max Base Depth, Biggest Snowfall, Total Snowfall Days and Average Base Depth—or the calendar below that shows daily snowfall data for the current ski season. To see how much snow Whistler Blackcomb got last ski season or any ski season dating back to 2009/2010, click the corresponding tab. Click **Compare** at the top left to see a season-over-season comparison of Whistler Blackcomb snowfall. In addition to snow history, you can also view the mountain's Base Depth by selecting it from the dropdown.

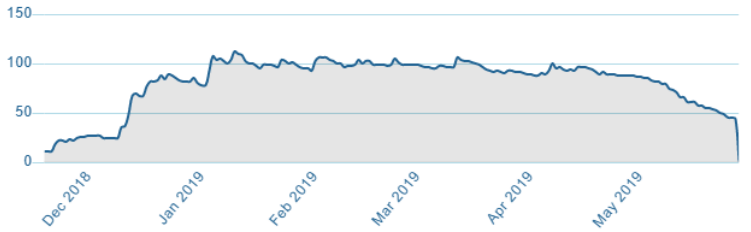
SNOWFALL HISTORY

Advertisement

SNOWFALL HISTORY

Compare | 2018/2019 | 2017/2018 | 2016/2017
2015/2016 | 2014/2015 | 2013/2014 | 2012/2013
2011/2012 | 2010/2011 | 2009/2010

SNOW DEPTH



Total Snowfall: 373"	Max Base Depth: 106"	Max Summit Depth: 112"
Total Snowfall Days: 82	Average Base Depth: 44"	Average Summit Depth: 82"
Biggest Snowfall: 29"		

The historical calendar and graph show all snowfall reported or sourced directly from the ski resort and include all historical snow data available, regardless of whether the ski resort was open or not.

Nov 2018						
Mon	Tue	Wed	Thu	Fri	Sat	Sun
19	20	21	22	23	24	25
11	10	10	18	22	22	20
26	27	28	29	30		
23	22	24	25	25		

Dec 2018						
Mon	Tue	Wed	Thu	Fri	Sat	Sun
					1	2
					27	27
3	4	5	6	7	8	9
26	26	24	24	24	24	24
10	11	12	13	14	15	16
35	36	48	67	69	67	67

Jan 2019						
Mon	Tue	Wed	Thu	Fri	Sat	Sun
	1	2	3	4	5	6
	78	78	93	107	104	105
7	8	9	10	11	12	13
102	100	104	112	110	108	102
14	15	16	17	18	19	20
100	100	97	95	98	98	98

Bryan Parthum
Parthum2@illinois.edu

Application: Dynamic Scraping

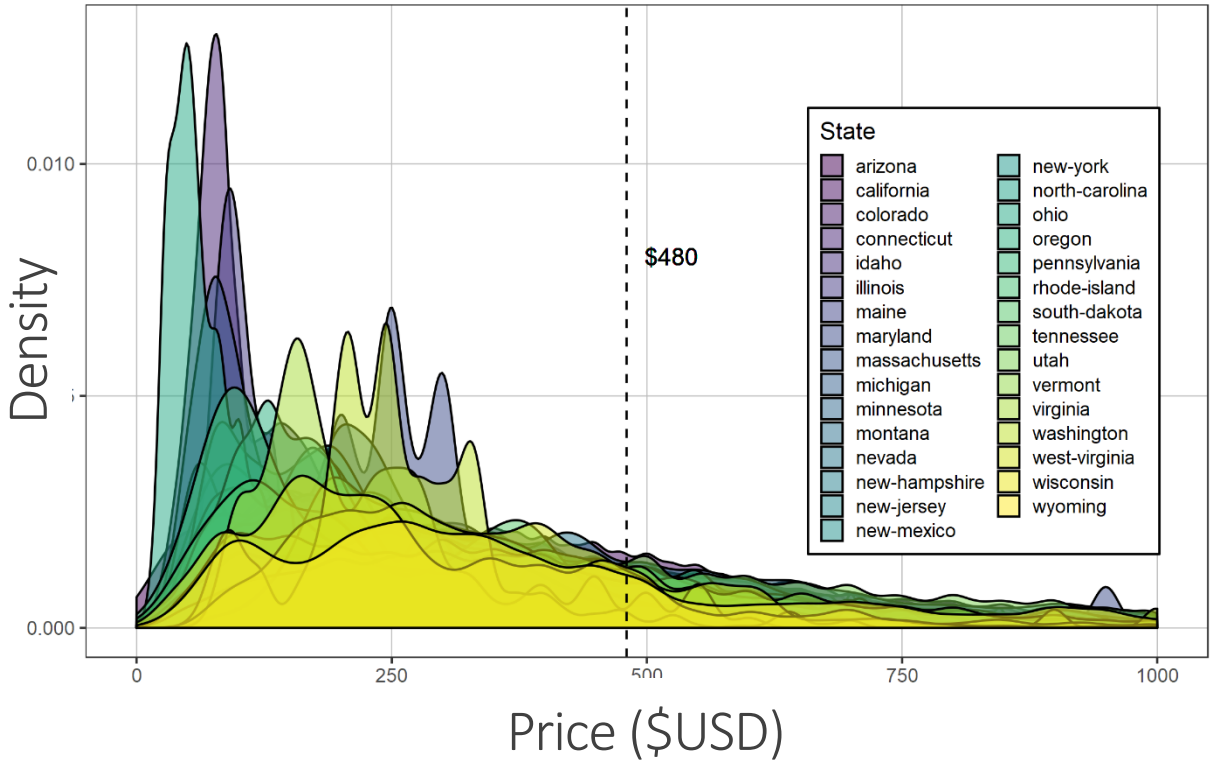
https://www.onthesnow.com/british-columbia/whistler-blackcomb/historical-snowfall.html?y=2018&q=top

```
11 def grabData(region, resort):
12     print(resort)
13     ## SCRAPES DATA FROM ON THE SNOW FOR THE RESORT AND REGION FROM 2009 to 2017
14     ## OUTPUTS DATA FORMATTED ONTO A CSV FILE
15     i = 2004
16     year = 2018
17     while i < year:
18         try:
19             page = requests.get("http://www.onthesnow.com/" + region + "/" + resort + '/historical-snowfall.html?&y=" + str(i) + "&v=list")
20             soup = BeautifulSoup(page.content, "lxml")
21 
22             headers = [c.get_text() for c in soup.find('tr').find_all('td')[0:4]]
23             data = [[cell.get_text(strip=True) for cell in row.find_all('td')[0:4]]
24                 for row in soup.find_all("tr")]
25 
26             for entry in data:
27                 masterCSV = open("snowdata.csv", 'a')
28                 masterWriter = csv.writer(masterCSV, delimiter=",")
29 
30                 date = str(entry).replace('[', '')
31                 date = date.replace(']', '')
32                 date = date.replace("'", '')
33                 date = date.replace(',', '')
34                 final = date.replace('u', '@')
35 
```

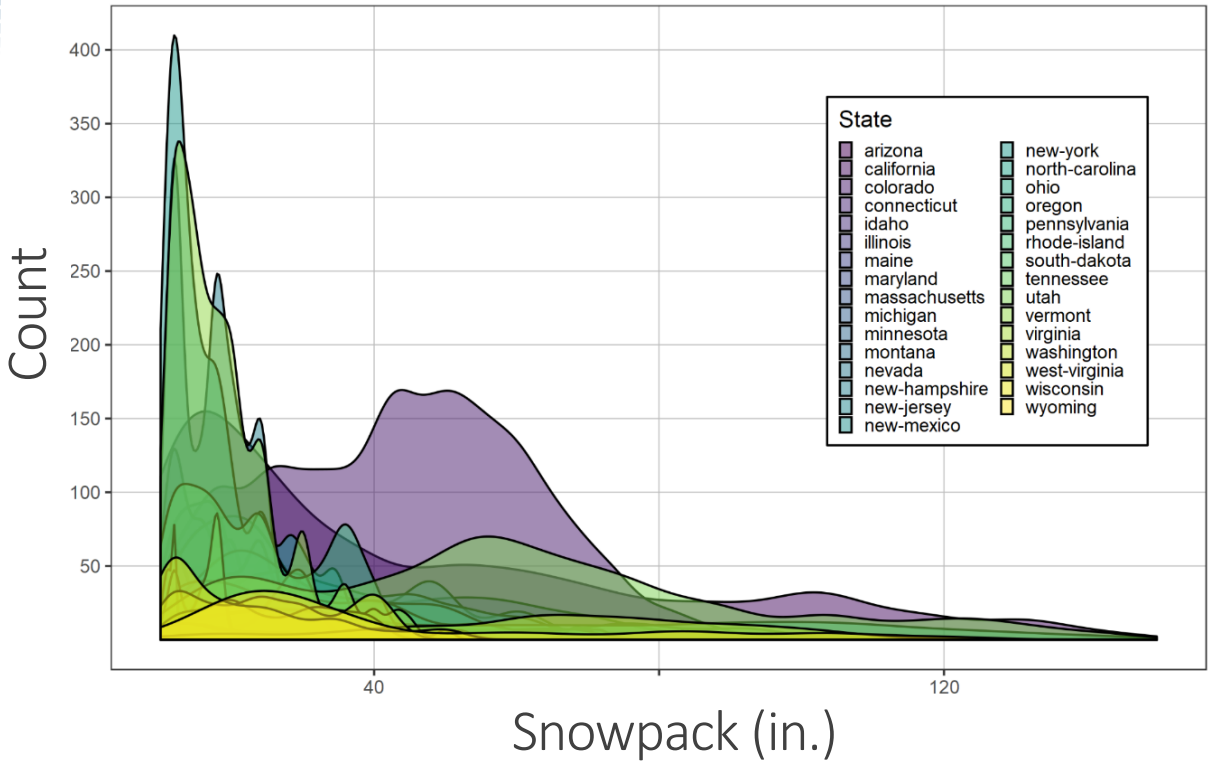

Bryan Parthum
Parthum2@illinois.edu

Application: Dynamic Scraping

Price of Short Term Rental (daily)



Resort Snowpack (daily)



Bryan Parthum
Parthum2@illinois.edu

Resources

- Introduction to Python
  - Python
    - Great resources for complete beginners. Experience with a keyboard and mouse is helpful.

- Text editors
  - Spyder
    - This one gives you the familiar comfort of a Stata or Rstudio environment, which is nice for getting started. (“where’s my .do file?”)
  - Sublime
    - CS kids seem to like this one

- Some places where you can find help for building your own bot
  - Hackernoon – Beautifulsoup to JSON
  - Codeburst – Beautifulsoup to .csv

Bryan Parthum
Parthum2@illinois.edu

References

- Baylis et al. Weather impacts expressed sentiment. 2018. PloS One.
- Kramer, Guillory, Hancock. Experimental evidence of massive-scale emotional contagion through social networks. 2014. PNAS.
- Zheng et al. Air pollution lowers Chinese urbanites' expressed happiness on social media. 2019. Nature Human Behavior.
- Phadera, L. Unfortunate Moms and Unfortunate Children: Impact of the Nepali Civil War on Women's Stature and Intergenerational Health. 2019. World Bank Group. (scraper built by Mateus Souza)
- Myers, Erica, Peter Christensen, Paul Francisco, Mateus Souza. Decomposing the Wedge: Evidence from a Home Weatherization Program.
- Christensen and Timmins. The Welfare Effect of Racial Discrimination in the US Rental Market for Housing.
- Christensen and Osman. The Demand for Mobility in a Transit-Constrained City: A Field Experiment with Uber in Cairo.
- Parthum and Christensen. The Price of Powder: Evidence on the Demand for Snow from Short Term Property Rentals.
- Big Data in Environmental Economics and Policy Research Group (BDEEP)

Bryan Parthum
Parthum2@illinois.edu

The Price of Powder: Evidence on the demand for snow from short term property rentals

Bryan Parthum Peter Christensen

Camp Resources

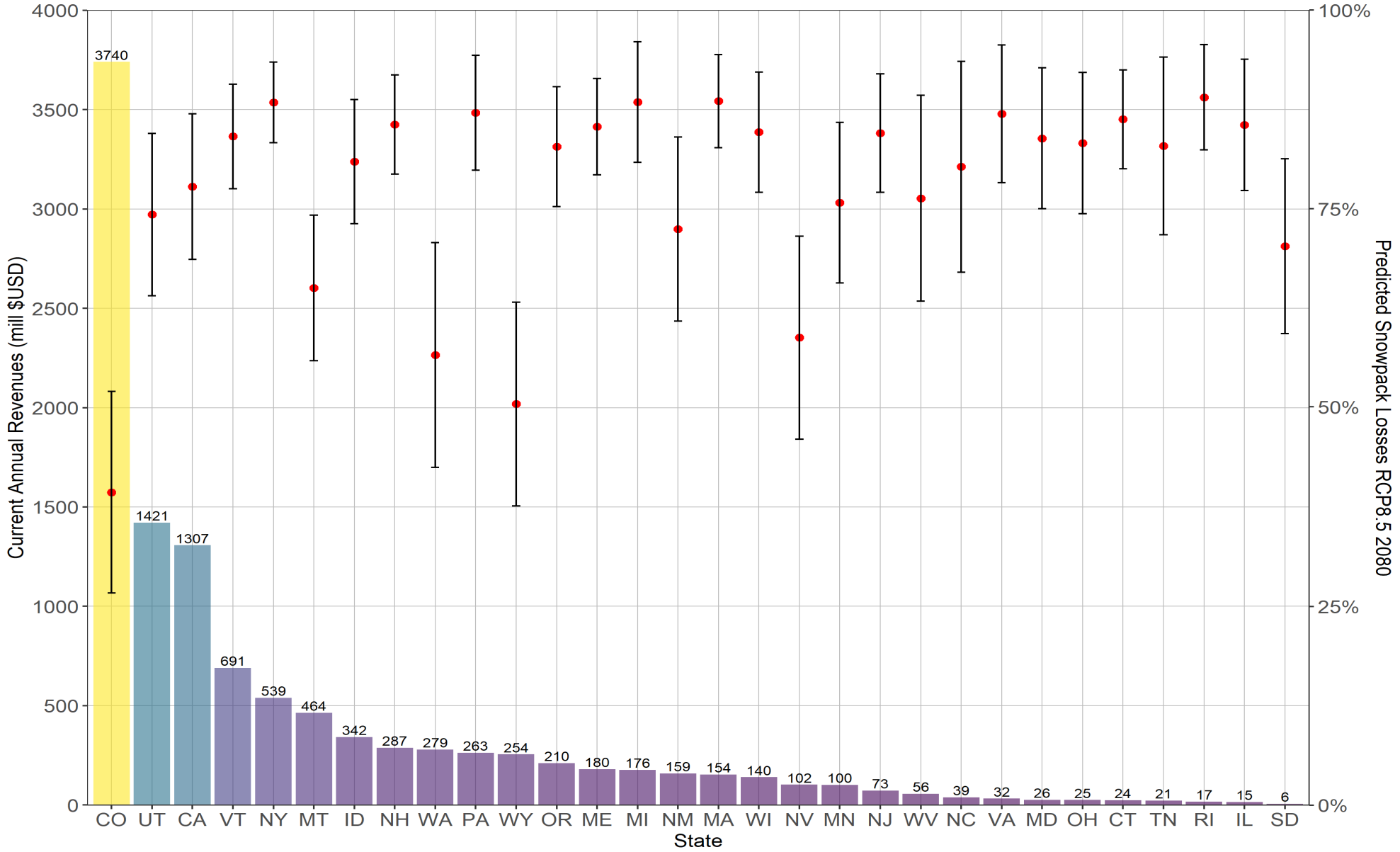
Asheville, NC

Tuesday, August 13th, 2019

ILLINOIS
Agricultural &
Consumer Economics
COLLEGE OF AGRICULTURAL, CONSUMER & ENVIRONMENTAL SCIENCES

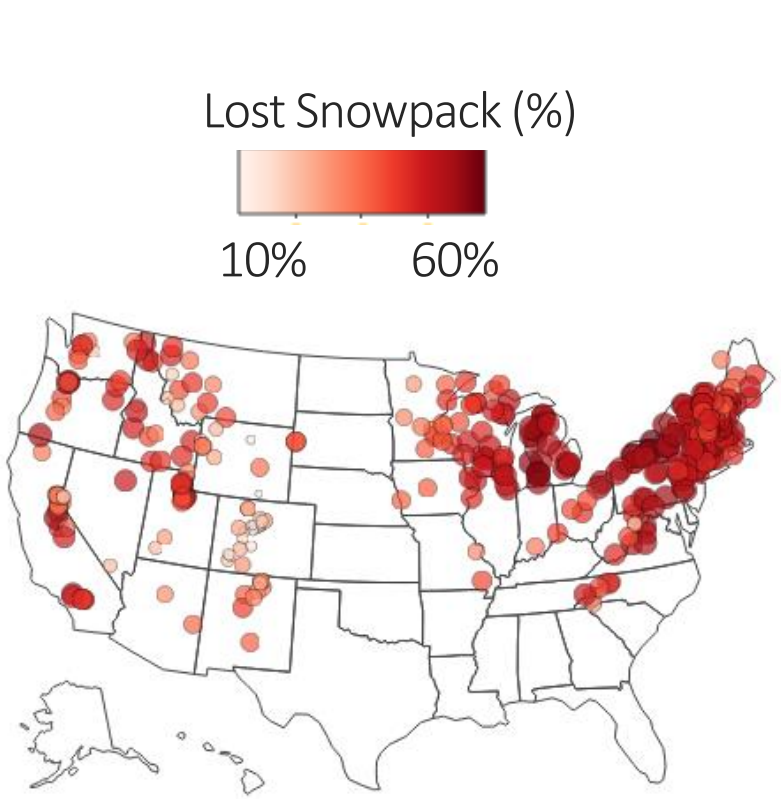
BDEEP
Big Data in Environmental Economics and Policy Research Group

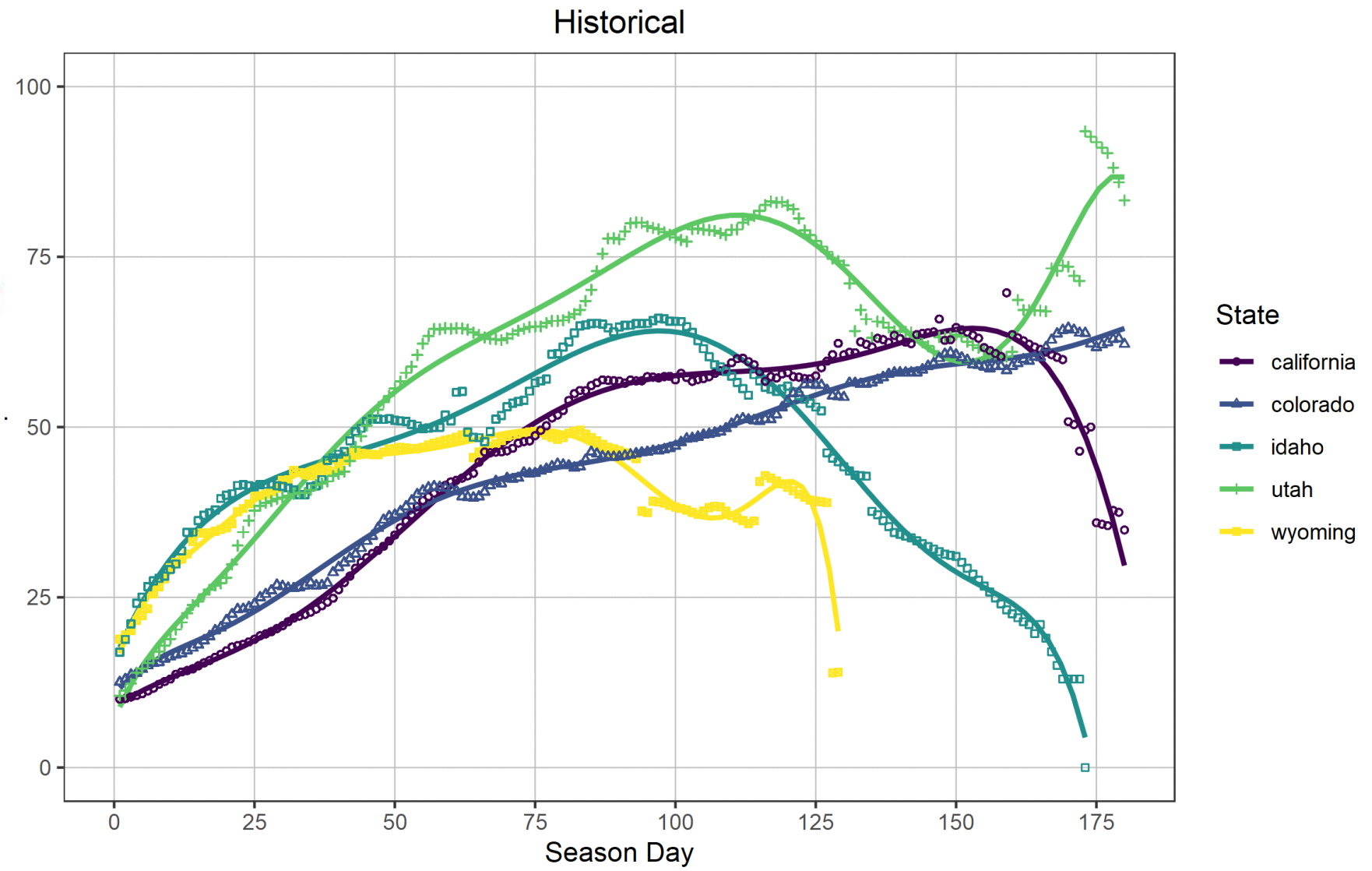
Motivation: Wintertime recreation is greatly threatened by climate change



Bryan Parthum
Parthum2@illinois.edu

Large Spatial and Temporal Variation





Bryan Parthum
Parthum2@illinois.edu

Research Questions

What is the willingness to pay (accept) for snowpack in ski resort markets?

Positive demand for snow → We estimate an average marginal utility of *snowpack* of 0.8. At the means, this implies a WTP of $5 for one inch of resort *snowpack*.

Markets are heterogeneous → We find marginal utilities are spatially heterogeneous, and vary at the state-level between 0.07 and 2.4.

What are the welfare implications of climate change for these consumers?

Climate change is bad → We estimate current consumer surplus to be $32 billion per ski season from *snowpack* alone. Damages under two future scenarios (RCP4.5 and RCP8.5) are estimated to be between $7 billion to $14 billion per ski season, or 30% to 60% of current levels.

Bryan Parthum
Parthum2@illinois.edu

Previous Work

Marginal Willingness to Pay for snow

- Morey (1981/84) kicks off the ski literature estimating share equations for Colorado resorts. Elasticities for annual *snowfall* are mixed in sign.
- Englin and Moeltner (2003/04) use survey data from the Tahoe region in California, include demand estimates relating to weekly *snowfall,* and estimate elasticities around 0.2.

Climate change and recreational demand

- Loomis and Crespi (1999), Mendelsohn and Markowski (1999) both estimate reductions in number of trips (52% and 39%) to estimate effects of climate change on skiing. Explicitly omit marginal values of *snowpack* and simply estimate effect of shorter seasons on number of trips (~$1.4b per year).
- Dundas and von Haefen (2019) discuss the importance of incorporating substitution into recreation demand models (recreational fishing)

Methods

- von Haefen and Domanski (2017): Estimating recreation demand with large choice sets
- Barry, Levinsohn, and Pakes (1995): Market share inversion to estimate own and cross elasticities

Bryan Parthum
Parthum2@illinois.edu

Data

Period

- August 2014 – May 2017
- Daily observations

Short term property rentals

- Price paid, date of stay and reservation
- 44k properties, ~7 million observations

Ski resorts

- 247 ski resorts across 30 states
- Daily *snowpack*, *snowfall*, *temperature*

Ski tourist information

- National Ski Area Association
- Trip costs (lift ticket, flights, etc.)
- Some skier demographics

Bryan Parthum
Parthum2@illinois.edu

Random Utility

$$V_{ij} = \lambda\, trip\ cost_{ij} + \beta_k \ln(snowpack)_{rj} \cdot D_k + \delta'_{rt} + \Gamma'_i + \psi_s + \gamma_w + \epsilon_{ij}$$

$$trip\ cost_{it} = rental + lift\ ticket + (\$0.33 * miles\ to\ metro)$$

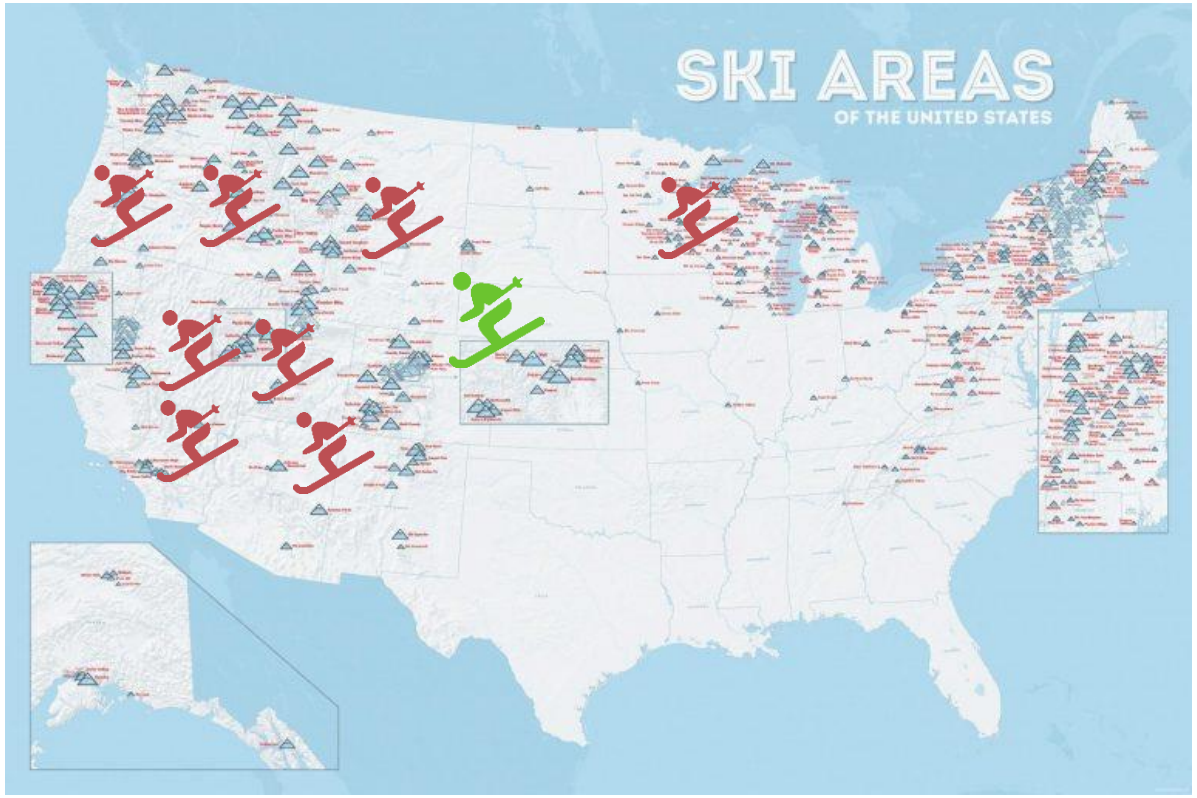
$D =$ destination state k
$\delta =$ vector of new *snowfall* bins;
controls for temperature (cubic);
holiday week (0,1)
$\Gamma =$ vector of prop characteristics (beds, baths, etc.)
$\psi =$ state and season FE
$\gamma =$ day-of-week (Sun.-Sat.) FE

$i =$ individual j $=$ alternative $r =$ resort;
$s =$ season $w =$ weekday $m =$ month-of-sample

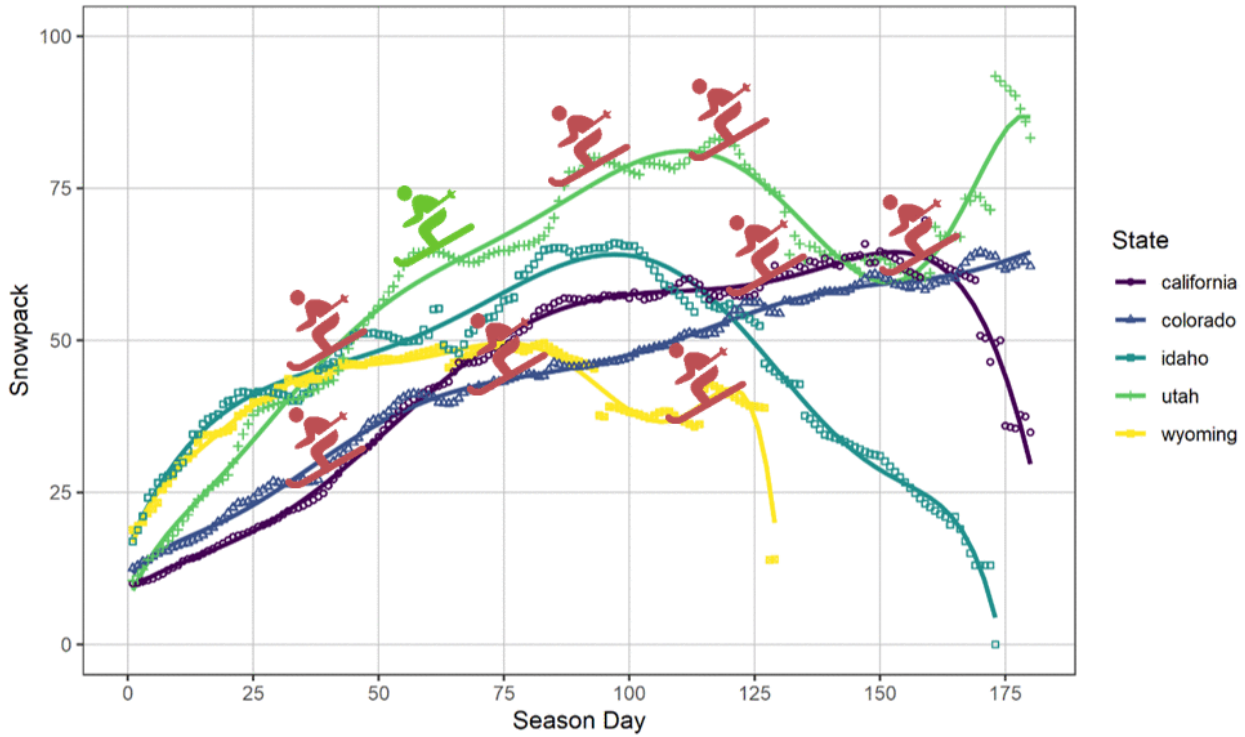
Bryan Parthum
Parthum2@illinois.edu

Choice Sets

Substitution across space



Substitution across space and time



= Observed choice V_{ij}

= Unobserved alternative $V_{i(-j)}$

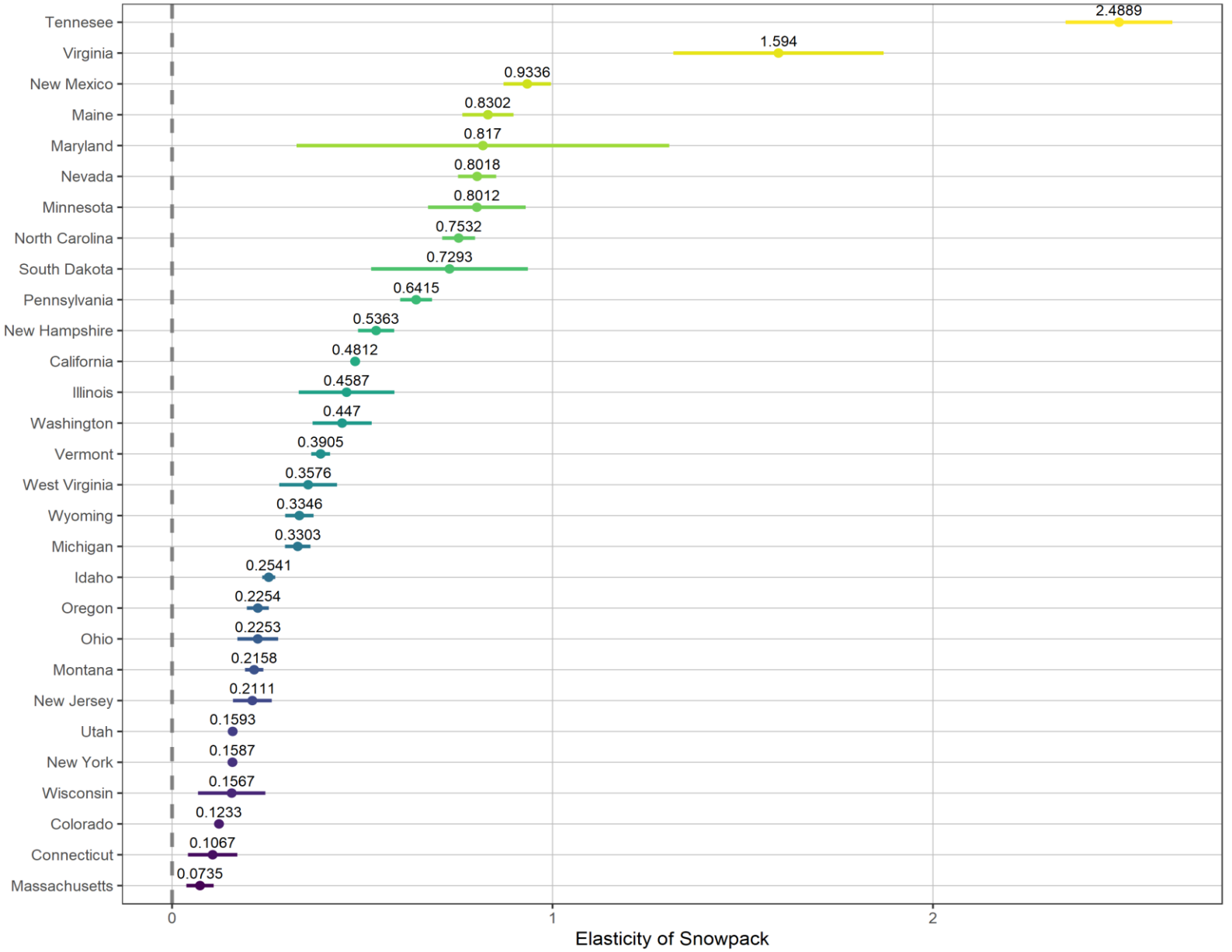
Bryan Parthum
Parthum2@illinois.edu

Results

Fixed Parameter Logit *(speedglm)*	(1) Home Price	(2) Trip Cost
Cost	-0.005***	-0.027***
	(0.002)	(0.008)
New snow 1"-3"	0.057***	0.059***
	(0.007)	(0.007)
New snow 3"-6"	0.047***	0.050***
	(0.008)	(0.009)
New snow 6"-9"	0.045***	0.043***
	(0.011)	(0.012)
New snow 9"-12"	0.050***	0.047***
	(0.018)	(0.012)
New snow 12"-15"	0.053***	0.056***
	(0.020)	(0.018)
New snow 15"+	0.027*	0.031**
	(0.016)	(0.015)
Observations	7,479,280	7,479,280
State FE	Y	Y
Season FE	Y	Y
Holiday Week FE	Y	Y
Clu. SE: NSAA Region	Y	Y
AIC	2,857,695	2,856,895
Standard errors in parentheses		*p<0.1; **p<0.05; ***p<0.01

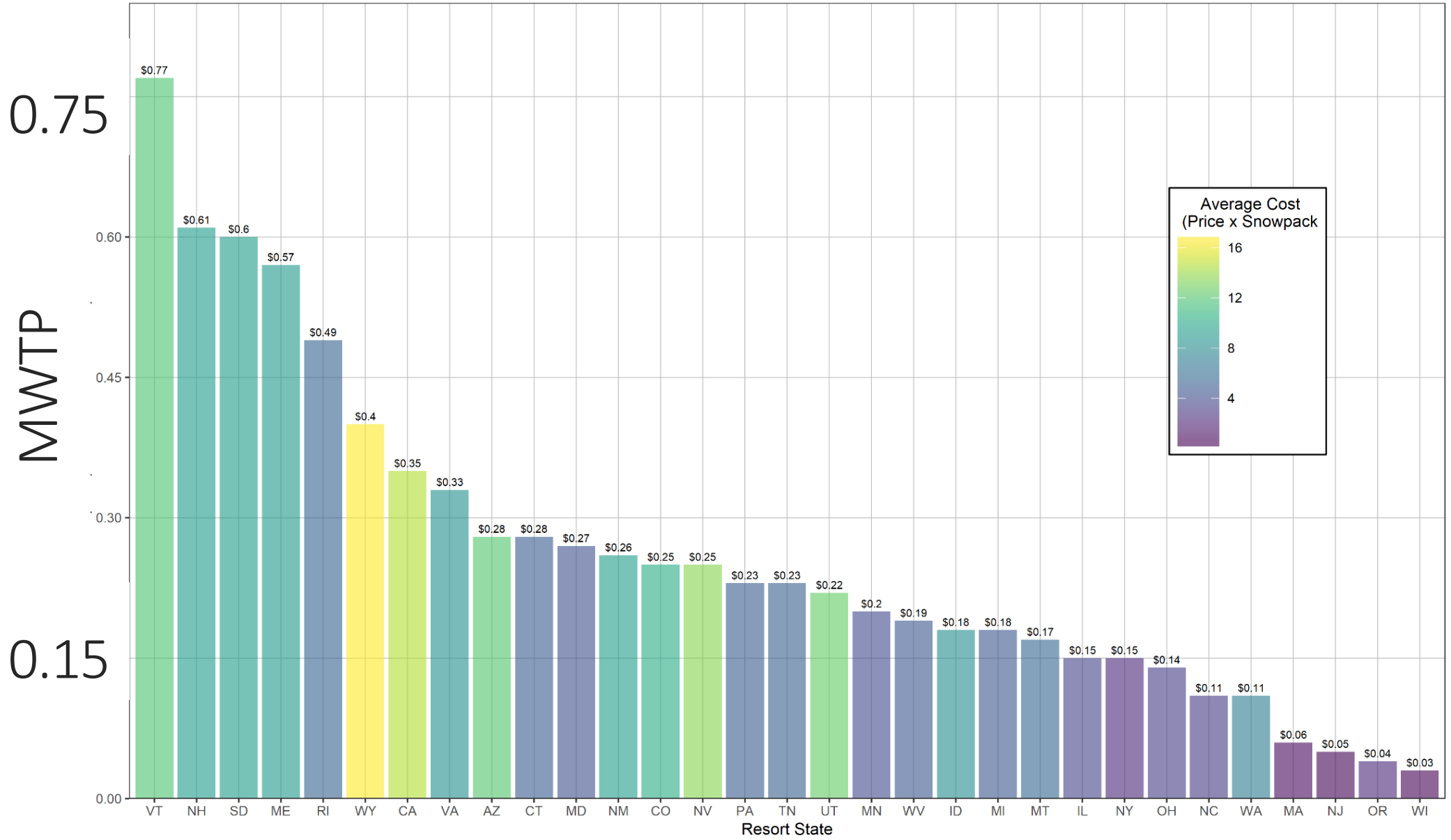
Bryan Parthum
Parthum2@illinois.edu

Marginal Utility of Snowpack



Bryan Parthum
Parthum2@illinois.edu

Marginal Willingness to Pay for 1 inch of Snowpack



Bryan Parthum
Parthum2@illinois.edu

Estimating Damages

Climate Data

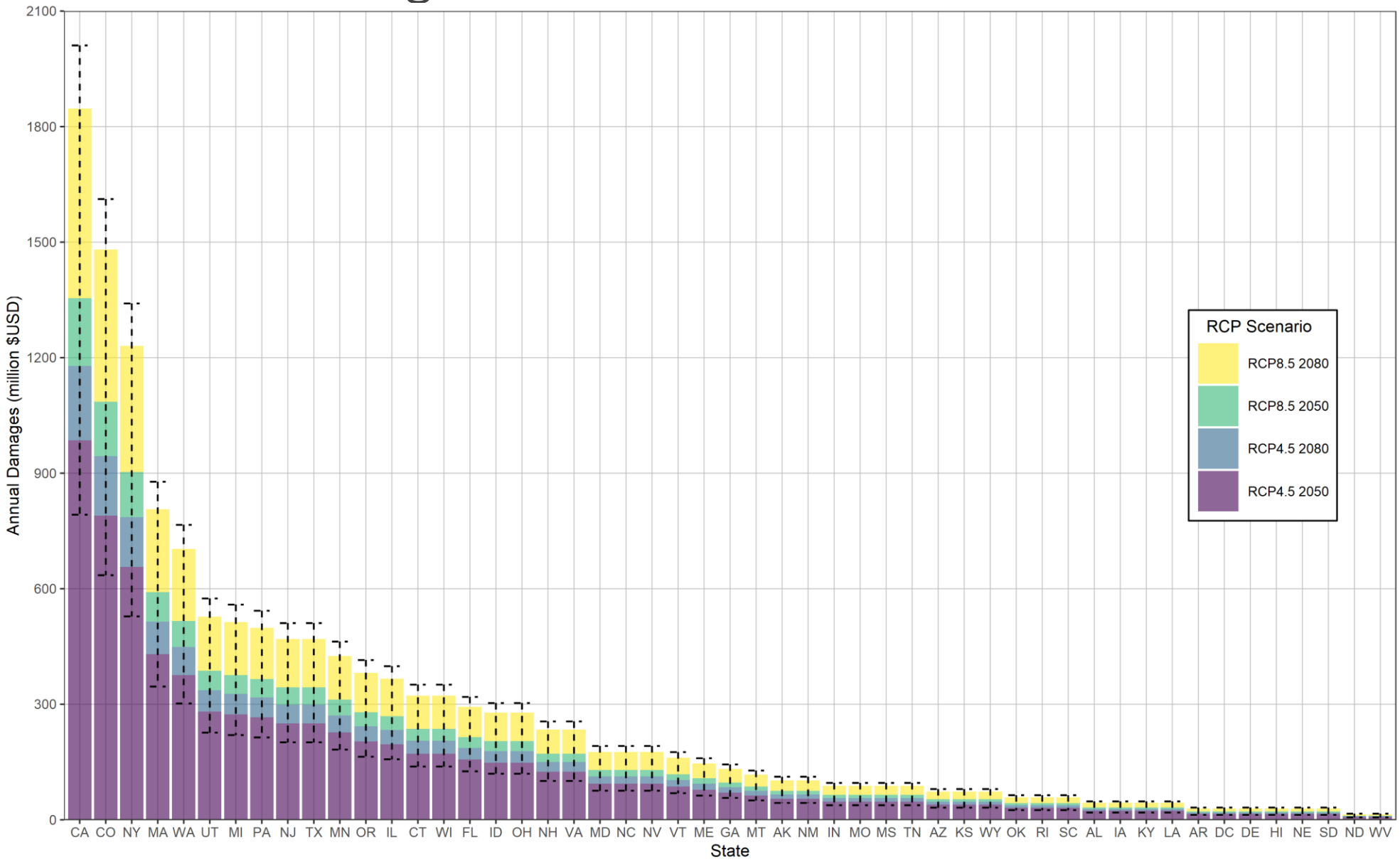
- Climate Change
  - Spatially downscaled estimates from USBR
  - Snow-water equivalent (*snowpack*)
  - 29 climate models
  - Historical: 1950-2015
  - Mid-century: 2040-2060
  - Late-century: 2070-2099
  - RCP4.5 and RCP8.5

Skier Demographics

- Ski tourist statistics
  - National Ski Area Association
  - Skier place of residence/destination
  - Visitation
  - Trip costs

Bryan Parthum
Parthum2@illinois.edu

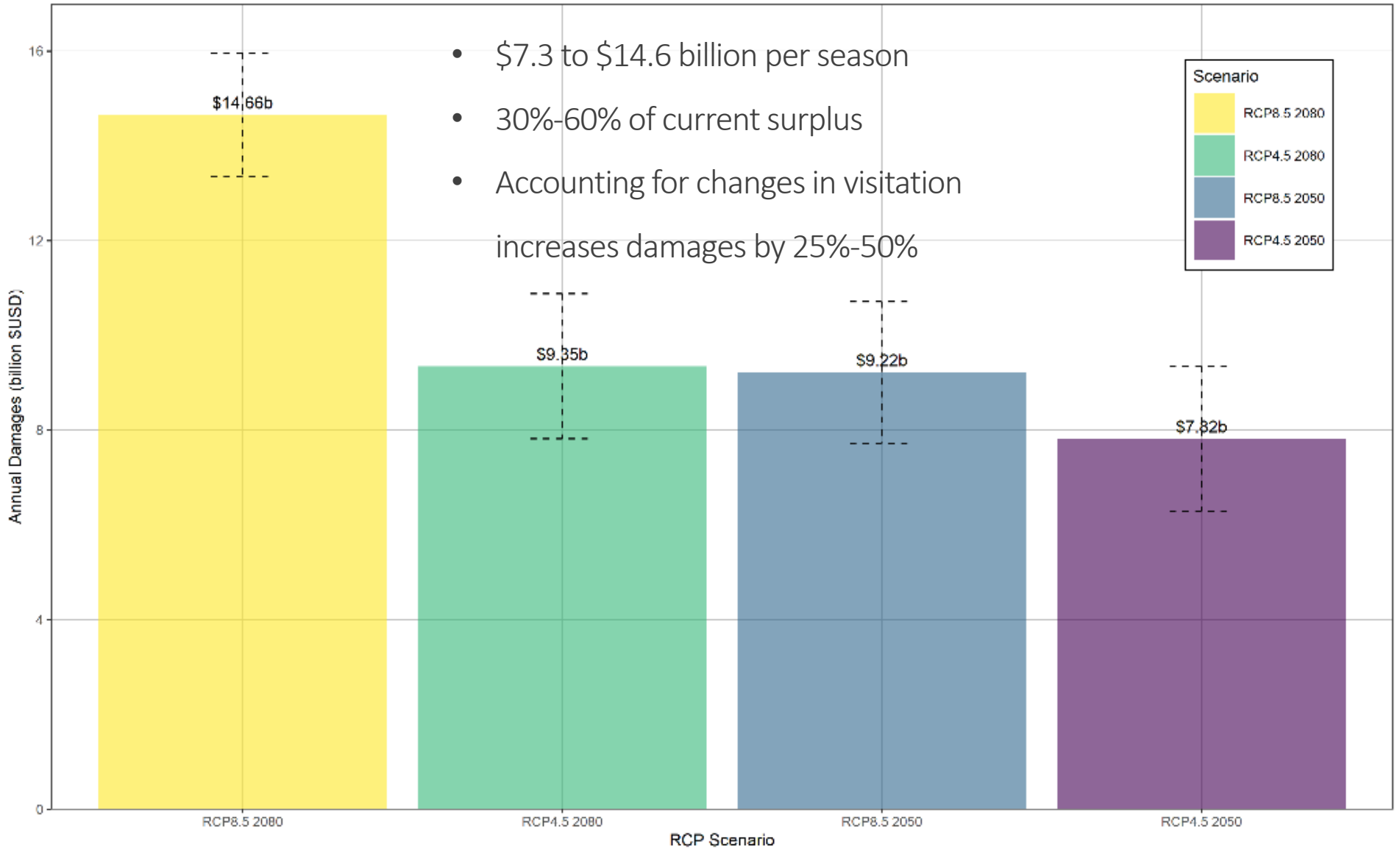
Damages Under future Climate Scenarios



Bryan Parthum
Parthum2@illinois.edu

Damages Under future Climate Scenarios

- $7.3 to $14.6 billion per season
- 30%-60% of current surplus
- Accounting for changes in visitation increases damages by 25%-50%



Bryan Parthum
Parthum2@illinois.edu

Market Share Transformation (BLP, 1995)

- $\Pr(j|x,\theta) = s_{jt} = \dfrac{\exp(\delta_{jt})}{1+\Sigma \exp(\delta_{kt})}$
- $\delta_{jt} = \lambda p_{jt} + X'_{jt}\beta + \xi_{ij}$
- $\xi_{ij} = log s_{jt} - \log s_{0t} - (-\alpha p_{jt} + X'_{jt}\beta)$

Bryan Parthum
Parthum2@illinois.edu

Market Share Transformation (BLP, 1995)

$$V_{ijt} = \lambda p_{jt} + X'_{jt}\beta + \xi_{ij}$$

$$-\left(\sum_{l=1}^{d} \pi_{1l} D_{il} + \sigma_1 \upsilon_{i1}\right) p_{jt}$$

$$+\sum_{k=1}^{K}\left(\sum_{l=1}^{d} \pi_{(k+1)l} D_{il} + \sigma_{(k+1)} \upsilon_{i(k+1)}\right) X'_{jk} + \epsilon_{ijt}$$

Bryan Parthum
Parthum2@illinois.edu

Results of BLP

	Dependent variable: ξ_{ij}	
	(1)	(2)
	Mean β	Std. Dev.
Trip Cost	-0.0116**	-0.047***
	(0.005)	(0.014)
log(Snowpack)	0.835***	0.443*
	(0.178)	(0.257)
Observations	1998	
State FE	Y	
Season FE	Y	
Holiday Week FE	Y	
Clu. SE: NSAA Region	Y	
Wald	32.47***	
Standard errors in parentheses	*p<0.1; **p<0.05; ***p<0.01	

Estimated using BLPestimatoR in R

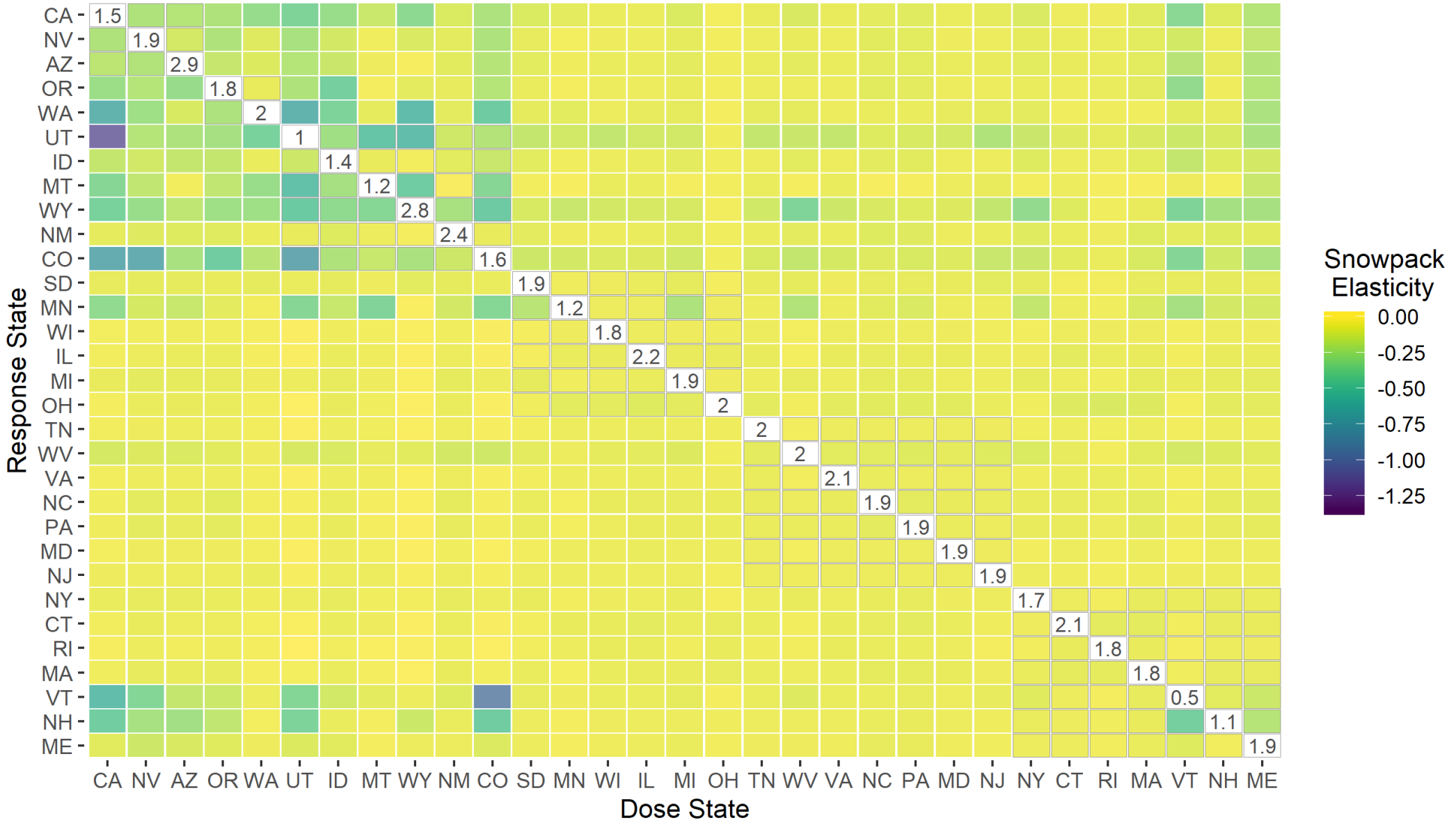
Bryan Parthum
Parthum2@illinois.edu

Elasticities of Snowpack

$$\eta_{jkt} = \frac{\partial s_{jt}}{\partial x_{kt}} \frac{x_{kt}}{s_{jt}} \begin{cases} \beta x_{jt}(1 - s_{jt}) & \text{if } j = k \\ \beta x_{kt} s_{kt} & \text{if } j \neq k \end{cases}$$

Bryan Parthum
Parthum2@illinois.edu

Own and Cross-snowpack Elasticities (state)



Bryan Parthum
Parthum2@illinois.edu

Thank You

www.BryanParthum.com
EconTwitter: @bryanparthum

ILLINOIS
Agricultural &
Consumer Economics
COLLEGE OF AGRICULTURAL, CONSUMER
& ENVIRONMENTAL SCIENCES

BDEEP
Big Data in Environmental
Economics and Policy
Research Group

Bryan Parthum
Parthum2@illinois.edu