Lecture 6: Estimating Demand Functions with Travel Cost Models

Prof. Parthum
Environmental Economics
Econ 475

The Travel Cost Model

- In June of 1947, the National Parks Service, headed by Arthur Demaray, reached out to a group of economists to ask about the feasibility of estimating the economic value of the national park system.

The Travel Cost Model

- In June of 1947, the National Parks Service, headed by Arthur Demaray, reached out to a group of economists to ask about the feasibility of estimating the economic value of the national park system.

- In response they received 10 letters. One of these letters was from Harold Hotelling.

The Travel Cost Model

- In June of 1947, the National Parks Service, headed by Arthur Demaray, reached out to a group of economists to ask about the feasibility of estimating the economic value of the national park system.

- In response they received 10 letters. One of these letters was from Harold Hotelling.

The University of North Carolina
Institute of Statistics
Chapel Hill

DEPARTMENT OF MATHEMATICAL STATISTICS

June 18, 1947

Mr. Newton B. Drury, Director
National Park Service
Department of the Interior
Washington 25, D. C.

Dear Mr. Drury:

After a letter from Mr. A. E. Demaray, and a conference with Dr. Roy A. Pruitt of the National Park Service, I am convinced that it is possible to set up appropriate measures for evaluating, with a reasonable degree of accuracy, the service of national parks to the public.

The development of criteria for evaluating benefits to the public has been a long-term interest of mine. Following the example set a hundred years ago by the French engineer, Jules Dupuit, who wrote formulae for the benefits of roads, bridges, and canals, I have worked out more general formulae for benefits from wider and more complicated classes of public services.

These formulae, of course, involve coefficients which must, in each case, be determined by factual statistical studies. The development of such studies I believe to be possible through several modes of attack which Dr. Pruitt and I discussed. One of these, of whose feasibility I am confident, and which might be pursued further, is as follows:

Let concentric zones be defined around each park so that the cost of travel to the park from all points in one of these zones is approximately constant. The persons entering the park in a year, or a suitably chosen sample of them, are to be listed according to the zone from which they come. The fact that they come means that they service of the park is at least worth the cost, and this cost can probably be estimated with fair accuracy. If we assume that the benefits are the same no matter what the distance, we have, for those living near the park, a consumers' surplus consisting of the differences in transportation costs. The comparison of the cost of coming from a zone with the number of people who do come from it, together with a count of the population of the zone, enables us to plot one point for each zone on a demand curve for the service of the park. By a judicious process of fitting it should be possible to get a good enough approximation to this demand curve to provide, through integration, a measure of the consumers' surplus resulting from the availability of the park. It is this consumers surplus (calculated by the above process with deduction for the cost of operating the park) which measures the benefits to the public in the particular year. This, of course, might be capitalized to give a capital value for the park, or the annual measure of benefit might be compared directly with the estimated annual benefits on the hypothesis that the park area was used for some alternate purpose.

The Travel Cost Model

$$\mathrm{V_{ijt}} \quad = \quad \lambda cost_{ijt} \; + \quad \beta X \qquad + \quad \phi \quad + \quad \varepsilon_{ijt}$$

The Travel Cost Model

$$\mathrm{V_{ijt}} \quad = \quad \lambda cost_{ijt} \quad + \quad \beta X \quad + \quad \phi \quad + \quad \varepsilon_{ijt}$$

$$= \quad \$\$\$ \quad + \quad \quad + \quad \quad +$$

The Travel Cost Model

$$V_{ijt} = \lambda cost_{ijt} + \beta X + \phi + \varepsilon_{ijt}$$

= $$$ + + +

The Travel Cost Model

$$V_{ijt} = \lambda cost_{ijt} + \beta X + \phi + \varepsilon_{ijt}$$

= $$$ + + +

The Travel Cost Model

$$V_{\mathrm{ijt}} = \lambda cost_{ijt} + \beta X + \phi + \varepsilon_{ijt}$$

= $$$  +  + +

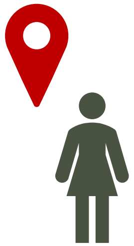
The Travel Cost Model

$$\mathrm{V_{ijt}} = \lambda cost_{ijt} + \beta X + \phi + \varepsilon_{ijt}$$

=  +  +  +

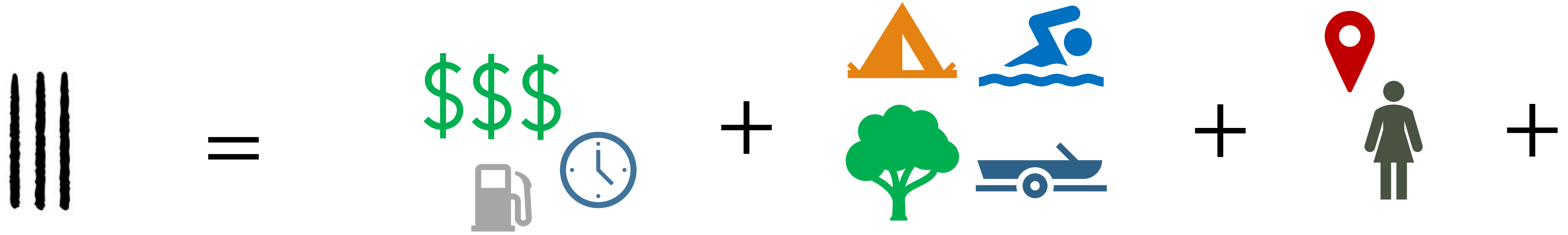
The Travel Cost Model

$$V_{ijt} = \lambda cost_{ijt} + \beta X + \phi + \varepsilon_{ijt}$$

= $$$  +  +  +

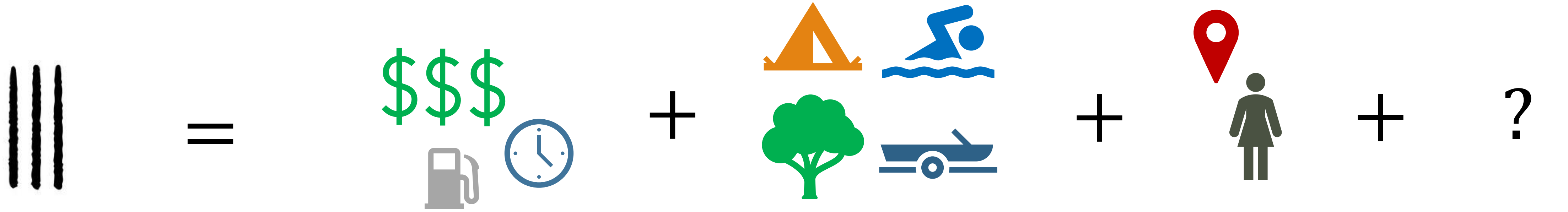
The Travel Cost Model

$$V_{ijt} = \lambda cost_{ijt} + \beta X + \phi + \varepsilon_{ijt}$$



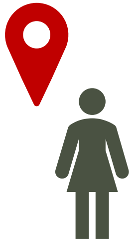
The Travel Cost Model

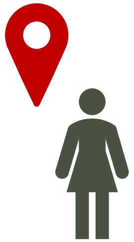
$$V_{ijt} = \lambda cost_{ijt} + \beta X + \phi + \varepsilon_{ijt}$$



The Travel Cost Model

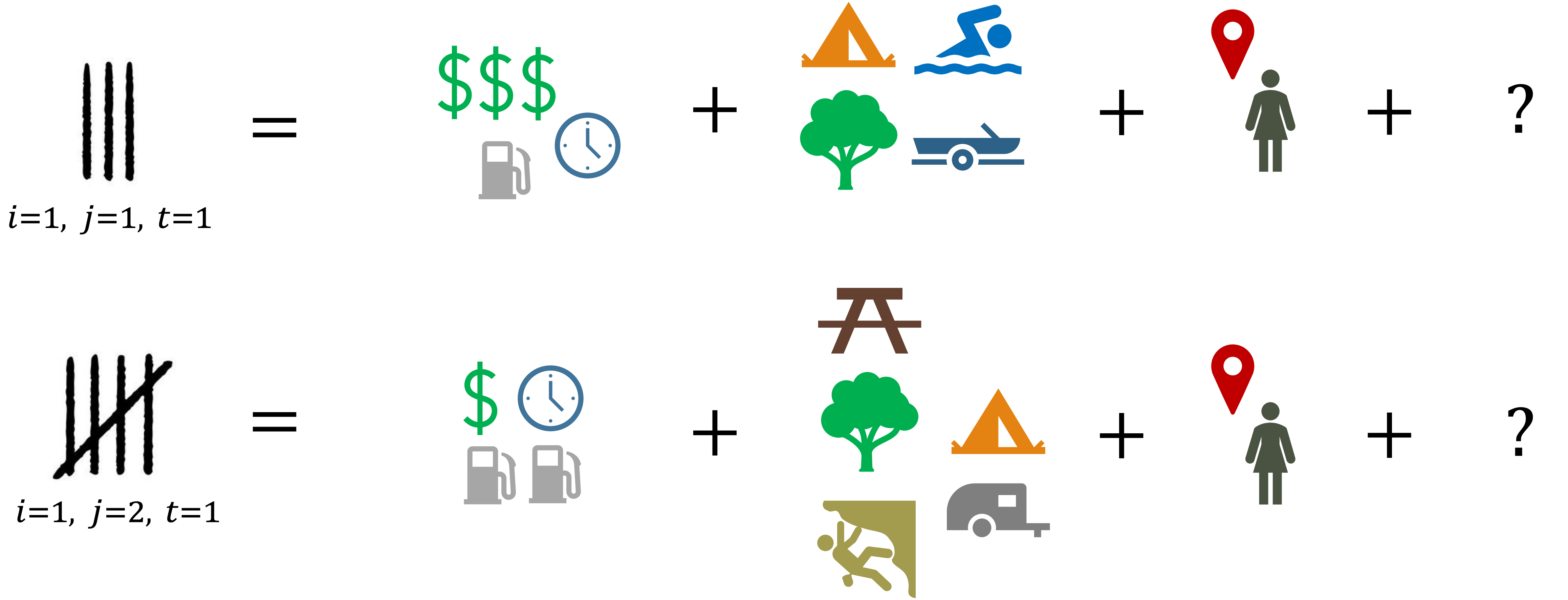
$$\mathrm{V}_{\mathrm{ijt}} = \lambda cost_{ijt} + \beta X + \phi + \varepsilon_{ijt}$$

 =  +  +  + ?

 =  +  +  + ?

The Travel Cost Model

$$V_{ijt} = \lambda cost_{ijt} + \beta X + \phi + \varepsilon_{ijt}$$



The Travel Cost Model

$$\widehat{V_{ijt}} = \lambda cost_{ijt} + \beta X + \phi + \varepsilon_{ijt}$$

$0.05 \times$ 

$+$

$1.5 \times$ 

$2.5 \times$ 

$1.8 \times$ 

$0.5 \times$ 

$0.7 \times$ 

$0.6 \times$ 

$0.3 \times$ 

$+$

$35.1 \times$ 

$15.2 \times$ 

The Travel Cost Model

Simple representation of consumer surplus if modeled linearly:

$$\widehat{Consumer\ Surplus}_{it} = \frac{\sum_{j=1}^{J} \beta_{ijt}}{\lambda}$$

However, typically modeled using logistic, Poisson, or negative binomial count data models:

$$\widehat{Consumer\ Surplus}_{it} = \frac{\log\left(\sum_{j=1}^{J} \mathrm{e}^{(V_{ijt})}\right)}{\lambda}$$

Valuing urban open space using the travel-cost method and the implications of measurement error (here)

By: Hanauer and Reid (2017)

- Research Questions:

1. What is the value of Taylor Mountain Regional Park in Santa Rosa, California?

2. Can using individual-specific travel cost estimates improve estimation?

Valuing urban open space using the travel-cost method and the implications of measurement error (here)

By: Hanauer and Reid (2017)

- Research Questions:

  1. What is the value of Taylor Mountain Regional Park in Santa Rosa, California?
     - Approximately $13.70 per trip, or $1.5mil each year

  2. Can using individual-specific travel cost estimates improve estimation?
     - They find evidence that using conventional methods could potentially understate the value of the regional park.

Valuing urban open space using the travel-cost method and the implications of measurement error (here)

By: Hanauer and Reid (2017)

- Data:

  1. 439 intercept surveys

     - Number of trips, costs, where they live, how they got there, if they paid for parking, income, demographics, etc.
     - Average number of self reported trips in the past year 32.8

Valuing urban open space using the travel-cost method and the implications of measurement error (here)

By: Hanauer and Reid (2017)

- Data:

1. 439 intercept surveys
    - Number of trips, costs, where they live, how they got there, if they paid for parking, income, demographics, etc.
    - Average number of self reported trips in the past year 32.8

2. How were costs calculated?

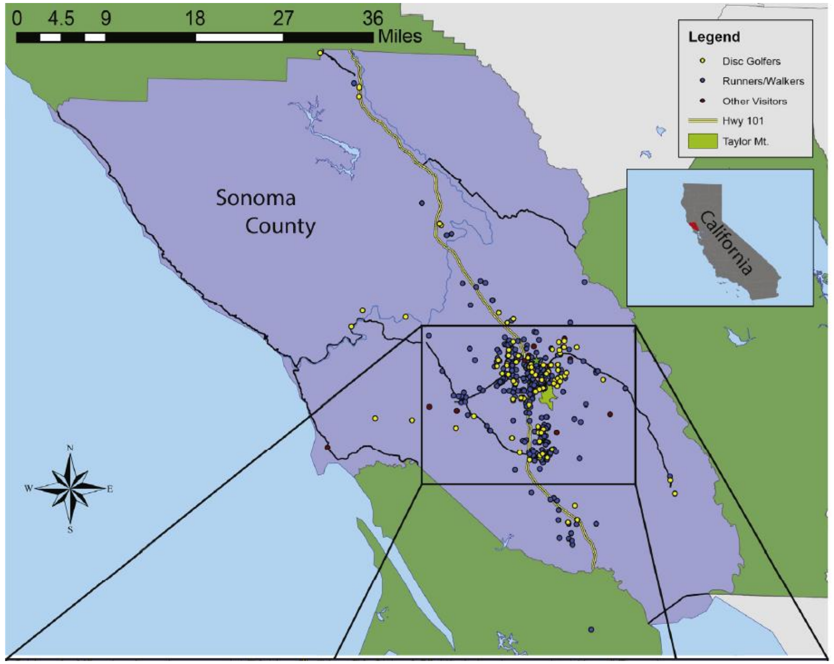
Valuing urban open space using the travel-cost method and the implications of measurement error (here)

By: Hanauer and Reid (2017)

- Data:

1. 439 intercept surveys
    - Number of trips, costs, home address, how they got there, if they paid for parking, activities they participate in, income, demographics, etc.
    - Average number of self reported trips in the past year 32.8
    - 72% white, $80k income

2. How were costs calculated?
    - $Trip\ Cost = \frac{1}{3} wage \times hours + vehicle\ costs + parking\ fees$
    - Average trip cost was $24



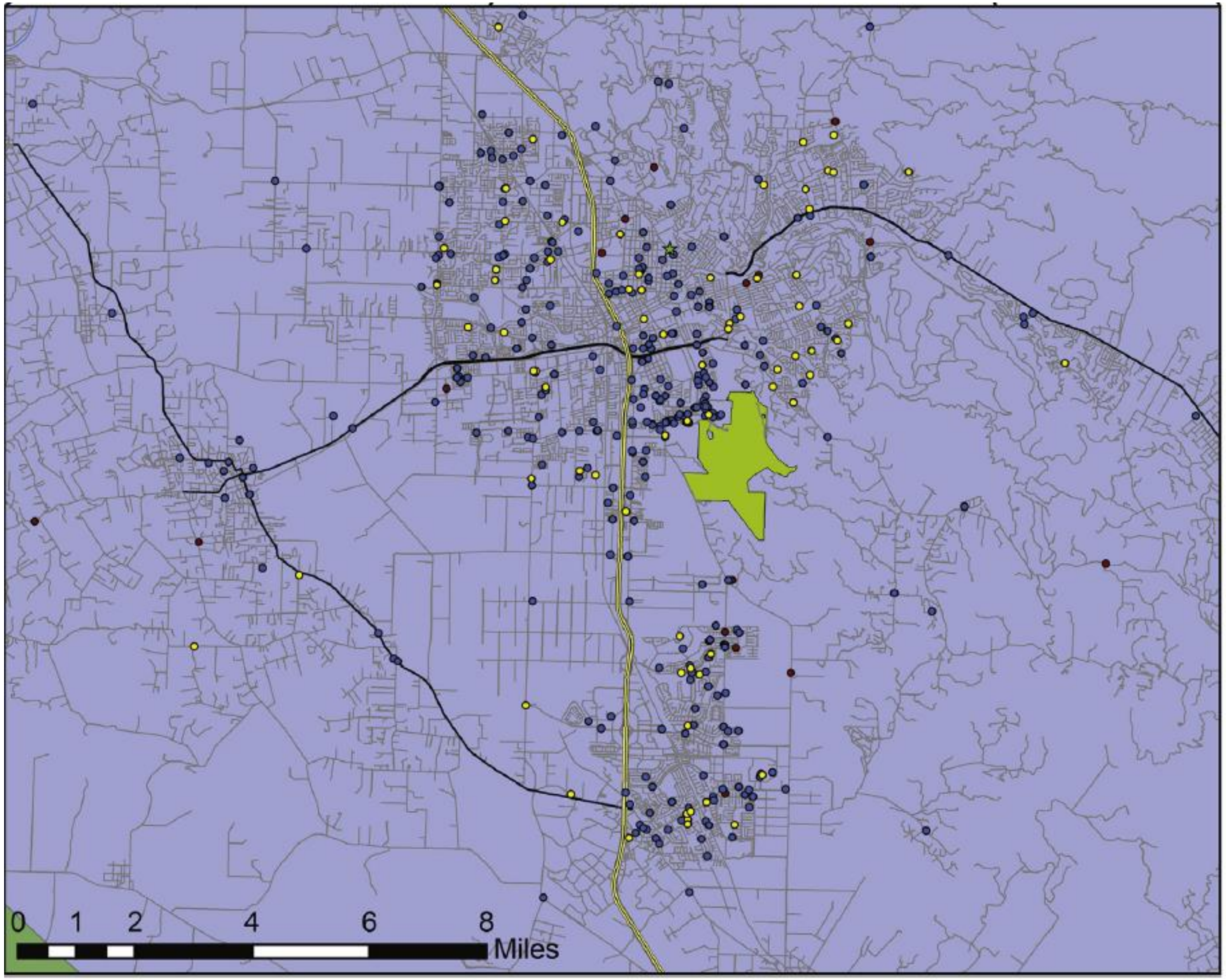


Fig. 1. Locations of Taylor Mountain Regional Park and households of visitors in the sample.

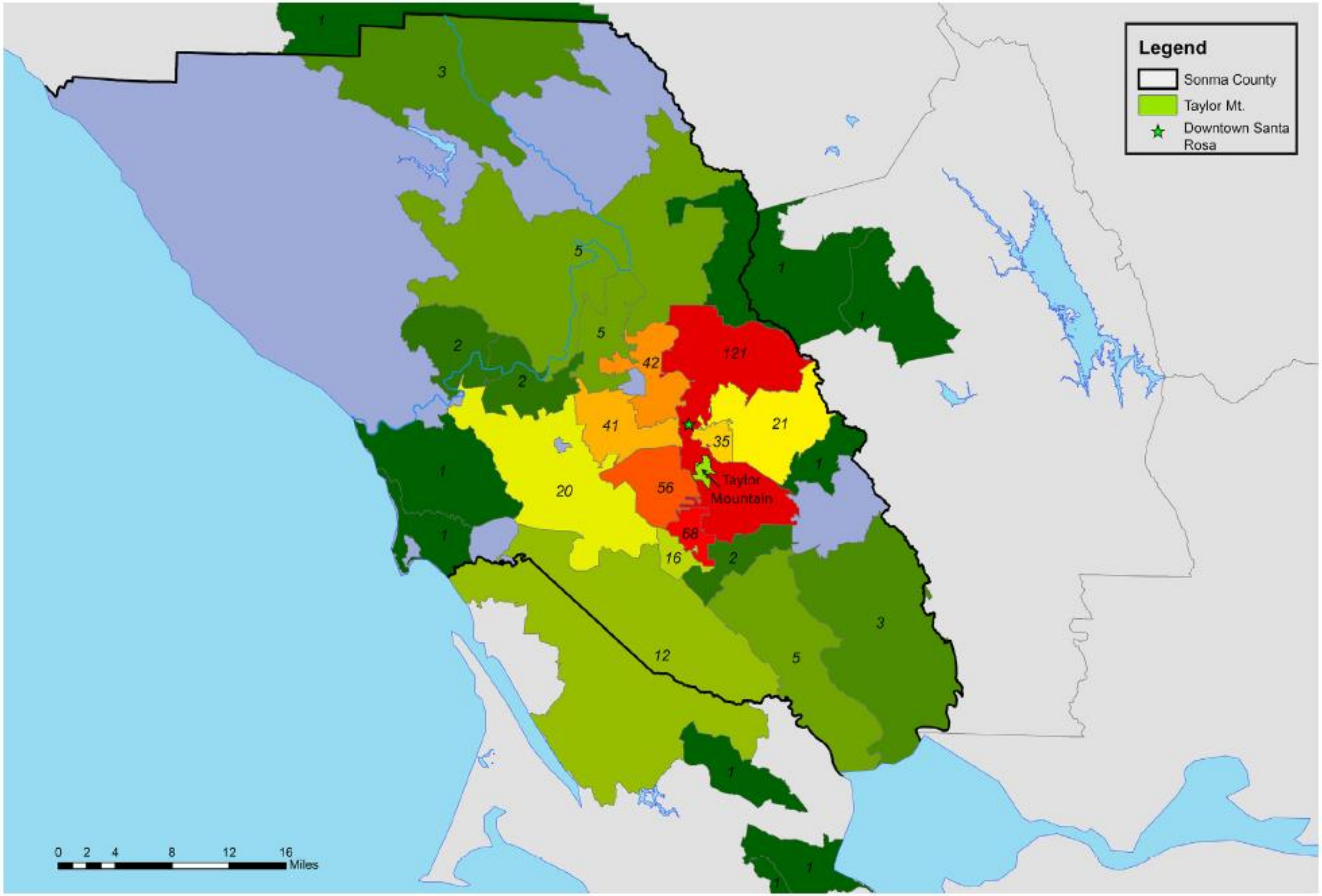


Fig. 2. Visitors' household location according to zip-code boundaries. The numbers represent the number of visitors within the sample that originated from the respective zip code.

$$V_{ijt} = \sum_{j=1}^{j=3} \lambda_j cost_{ijt} + \sum_{k=1}^{K} \beta_k x_{ikt}^{k} + \varepsilon_{ijt}$$

Table 2

Regression results from our primary measurement specification. Columns 1–3 present results from poisson regressions with various control variable specifications. Columns 4–6 present results from negative binomial regressions with various control variable specifications. All regressions account for zero truncation and endogenous stratification.

Dependent Variable: Number of Visits in Past Year						
Variables	Poisson			Truncated Negative Binomial		
	(1)	(2)	(3)	(4)	(5)	(6)
Travel Cost TM	−0.0329***	−0.0626***	−0.0630***	−0.0300***	−0.0697***	−0.0730***
	(0.000736)	(0.00117)	(0.00122)	(0.00227)	(0.00441)	(0.00463)
Travel Cost Annadel		0.0207***	0.0174***		0.0412***	0.0391***
		(0.00114)	(0.00119)		(0.00559)	(0.00585)
Travel Cost Crane Creek		−0.000513	0.00193***		0.00450	0.00515
		(0.000608)	(0.000637)		(0.00385)	(0.00371)
Income (1000s)		0.00481***	0.00419***		0.00168	0.00239**
		(0.000204)	(0.000214)		(0.00105)	(0.00121)
Constant	4.049***	3.661***	3.399***	−11.21***	−8.358	−10.52
	(0.0137)	(0.0170)	(0.0456)	(0.0683)	(26.94)	(76.89)
Additional Controls			Yes			Yes
Observations	439	439	439	439	439	439
AIC	24887.15	22152.69	20810.14	3801.91	3638.09	3588.79
Alpha				1.62***	1.41***	1.30***

Standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

$$V_{ijt} = \sum_{j=1}^{j=3} \lambda_j cost_{ijt} + \sum_{k=1}^{K} \beta_k x_{ikt}^{k} + \varepsilon_{ijt}$$

Table 2

Regression results from our primary measurement specification. Columns 1–3 present results from poisson regressions with various control variable specifications. Columns 4–6 present results from negative binomial regressions with various control variable specifications. All regressions account for zero truncation and endogenous stratification.

Dependent Variable: Number of Visits in Past Year						
Variables	Poisson			Truncated Negative Binomial		
	(1)	(2)	(3)	(4)	(5)	(6)
Travel Cost TM	−0.0329***	−0.0626***	−0.0630***	−0.0300***	−0.0697***	−0.0730***
	(0.000736)	(0.00117)	(0.00122)	(0.00227)	(0.00441)	(0.00463)
Travel Cost Annadel		0.0207***	0.0174***		0.0412***	0.0391***
		(0.00114)	(0.00119)		(0.00559)	(0.00585)
Travel Cost Crane Creek		−0.000513	0.00193***		0.00450	0.00515
		(0.000608)	(0.000637)		(0.00385)	(0.00371)
Income (1000s)		0.00481***	0.00419***		0.00168	0.00239**
		(0.000204)	(0.000214)		(0.00105)	(0.00121)
Constant	4.049***	3.661***	3.399***	−11.21***	−8.358	−10.52
	(0.0137)	(0.0170)	(0.0456)	(0.0683)	(26.94)	(76.89)
Additional Controls			Yes			Yes
Observations	439	439	439	439	439	439
AIC	24887.15	22152.69	20810.14	3801.91	3638.09	3588.79
Alpha				1.62***	1.41***	1.30***

Standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

$$V_{ijt} = \sum_{j=1}^{j=3} \lambda_j cost_{ijt} + \sum_{k=1}^{K} \beta_k x_{ikt}^{k} + \varepsilon_{ijt}$$

Table 2

Regression results from our primary measurement specification. Columns 1–3 present results from poisson regressions with various control variable specifications. Columns 4–6 present results from negative binomial regressions with various control variable specifications. All regressions account for zero truncation and endogenous stratification.

Dependent Variable: Number of Visits in Past Year						
Variables	Poisson			Truncated Negative Binomial		
	(1)	(2)	(3)	(4)	(5)	(6)
Travel Cost TM	−0.0329***	−0.0626***	−0.0630***	−0.0300***	−0.0697***	−0.0730***
	(0.000736)	(0.00117)	(0.00122)	(0.00227)	(0.00441)	(0.00463)
Travel Cost Annadel		0.0207***	0.0174***		0.0412***	0.0391***
		(0.00114)	(0.00119)		(0.00559)	(0.00585)
Travel Cost Crane Creek		−0.000513	0.00193***		0.00450	0.00515
		(0.000608)	(0.000637)		(0.00385)	(0.00371)
Income (1000s)		0.00481***	0.00419***		0.00168	0.00239**
		(0.000204)	(0.000214)		(0.00105)	(0.00121)
Constant	4.049***	3.661***	3.399***	−11.21***	−8.358	−10.52
	(0.0137)	(0.0170)	(0.0456)	(0.0683)	(26.94)	(76.89)
Additional Controls			Yes			Yes
Observations	439	439	439	439	439	439
AIC	24887.15	22152.69	20810.14	3801.91	3638.09	3588.79
Alpha				1.62***	1.41***	1.30***

Standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

$$V_{ijt} = \sum_{j=1}^{j=3} \lambda_j cost_{ijt} + \sum_{k=1}^{K} \beta_k x_{ikt}^{k} + \varepsilon_{ijt}$$

Table 2

Regression results from our primary measurement specification. Columns 1–3 present results from poisson regressions with various control variable specifications. Columns 4–6 present results from negative binomial regressions with various control variable specifications. All regressions account for zero truncation and endogenous stratification.

Dependent Variable: Number of Visits in Past Year						
Variables	Poisson			Truncated Negative Binomial		
	(1)	(2)	(3)	(4)	(5)	(6)
Travel Cost TM	−0.0329***	−0.0626***	−0.0630***	−0.0300***	−0.0697***	−0.0730***
	(0.000736)	(0.00117)	(0.00122)	(0.00227)	(0.00441)	(0.00463)
Travel Cost Annadel		0.0207***	0.0174***		0.0412***	0.0391***
		(0.00114)	(0.00119)		(0.00559)	(0.00585)
Travel Cost Crane Creek		−0.000513	0.00193***		0.00450	0.00515
		(0.000608)	(0.000637)		(0.00385)	(0.00371)
Income (1000s)		0.00481***	0.00419***		0.00168	0.00239**
		(0.000204)	(0.000214)		(0.00105)	(0.00121)
Constant	4.049***	3.661***	3.399***	−11.21***	−8.358	−10.52
	(0.0137)	(0.0170)	(0.0456)	(0.0683)	(26.94)	(76.89)
Additional Controls			Yes			Yes
Observations	439	439	439	439	439	439
AIC	24887.15	22152.69	20810.14	3801.91	3638.09	3588.79
Alpha				1.62***	1.41***	1.30***

Standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

$$V_{ijt} = \sum_{j=1}^{j=3} \lambda_j cost_{ijt} + \sum_{k=1}^{K} \beta_k x_{ikt}^k + \varepsilon_{ijt}$$

Table 2

Regression results from our primary measurement specification. Columns 1–3 present results from poisson regressions with various control variable specifications. Columns 4–6 present results from negative binomial regressions with various control variable specifications. All regressions account for zero truncation and endogenous stratification.

Dependent Variable: Number of Visits in Past Year						
Variables	Poisson			Truncated Negative Binomial		
	(1)	(2)	(3)	(4)	(5)	(6)
Travel Cost TM	−0.0329***	−0.0626***	−0.0630***	−0.0300***	−0.0697***	−0.0730***
	(0.000736)	(0.00117)	(0.00122)	(0.00227)	(0.00441)	(0.00463)
Travel Cost Annadel		0.0207***	0.0174***		0.0412***	0.0391***
		(0.00114)	(0.00119)		(0.00559)	(0.00585)
Travel Cost Crane Creek		−0.000513	0.00193***		0.00450	0.00515
		(0.000608)	(0.000637)		(0.00385)	(0.00371)
Income (1000s)		0.00481***	0.00419***		0.00168	0.00239**
		(0.000204)	(0.000214)		(0.00105)	(0.00121)
Constant	4.049***	3.661***	3.399***	−11.21***	−8.358	−10.52
	(0.0137)	(0.0170)	(0.0456)	(0.0683)	(26.94)	(76.89)
Additional Controls			Yes			Yes
Observations	439	439	439	439	439	439
AIC	24887.15	22152.69	20810.14	3801.91	3638.09	3588.79
Alpha				1.62***	1.41***	1.30***

Standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

$$V_{ijt} = \sum_{j=1}^{j=3} \lambda_j cost_{ijt} + \sum_{k=1}^{K} \beta_k x_{ikt}^k + \varepsilon_{ijt}$$

Table 2

Regression results from our primary measurement specification. Columns 1–3 present results from poisson regressions with various control variable specifications. Columns 4–6 present results from negative binomial regressions with various control variable specifications. All regressions account for zero truncation and endogenous stratification.

Dependent Variable: Number of Visits in Past Year						
Variables	Poisson			Truncated Negative Binomial		
	(1)	(2)	(3)	(4)	(5)	(6)
Travel Cost TM	−0.0329***	−0.0626***	−0.0630***	−0.0300***	−0.0697***	−0.0730***
	(0.000736)	(0.00117)	(0.00122)	(0.00227)	(0.00441)	(0.00463)
Travel Cost Annadel		0.0207***	0.0174***		0.0412***	0.0391***
		(0.00114)	(0.00119)		(0.00559)	(0.00585)
Travel Cost Crane Creek		−0.000513	0.00193***		0.00450	0.00515
		(0.000608)	(0.000637)		(0.00385)	(0.00371)
Income (1000s)		0.00481***	0.00419***		0.00168	0.00239**
		(0.000204)	(0.000214)		(0.00105)	(0.00121)
Constant	4.049***	3.661***	3.399***	−11.21***	−8.358	−10.52
	(0.0137)	(0.0170)	(0.0456)	(0.0683)	(26.94)	(76.89)
Additional Controls			Yes			Yes
Observations	439	439	439	439	439	439
AIC	24887.15	22152.69	20810.14	3801.91	3638.09	3588.79
Alpha				1.62***	1.41***	1.30***

Standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

$$\widehat{Consumer\ Surplus}_{it} = \frac{1}{\lambda_j}$$

Table 3
Consumer surplus calculations based on primary specification.

	Point Estimate	95% Confidence Interval	
		Lower	Upper
Mean CS_i/visit	$\boxed{\$13.70}$	$12.17	$15.65
Total CS/year	$1,480,782.42	$1,315,410.37	$1,691,550.73
Discount Rate	Present Value of Future Benefits		
1%	$148,078,242.44	$131,541,037.26	$169,155,072.56
3%	$49,359,414.15	$43,847,012.42	$56,385,024.19
5%	$29,615,648.49	$26,308,207.45	$33,831,014.51

$$\widehat{Consumer\ Surplus}_{it} = \frac{1}{\lambda_j}$$

Table 3
Consumer surplus calculations based on primary specification.

	Point Estimate	95% Confidence Interval	
		Lower	Upper
Mean CS_i/visit	\$13.70	\$12.17	\$15.65
Total CS/year	\$1,480,782.42	\$1,315,410.37	\$1,691,550.73
Discount Rate	Present Value of Future Benefits		
1%	\$148,078,242.44	\$131,541,037.26	\$169,155,072.56
3%	\$49,359,414.15	\$43,847,012.42	\$56,385,024.19
5%	\$29,615,648.49	\$26,308,207.45	\$33,831,014.51

$$\widehat{Consumer\ Surplus}_{it} = \frac{1}{\lambda_j}$$

Table 3
Consumer surplus calculations based on primary specification.

	Point Estimate	95% Confidence Interval	
		Lower	Upper
Mean CS_i/visit	\$13.70	\$12.17	\$15.65
Total CS/year	\$1,480,782.42	\$1,315,410.37	\$1,691,550.73
Discount Rate	Present Value of Future Benefits		
1%	\$148,078,242.44	\$131,541,037.26	\$169,155,072.56
3%	\$49,359,414.15	\$43,847,012.42	\$56,385,024.19
5%	\$29,615,648.49	\$26,308,207.45	\$33,831,014.51

- Does using zip-code level data to estimate travel costs bias the results?
    - The authors find that by failing to account for within-zip-code variation in travel costs the per trip consumer surplus and, thus, the overall benefits of Taylor Mountain Regional Park are underestimated compared to when more accurate individual-level costs are used.

Table 4

Comparison of primary results (bottom row) to results based on zip code point of origin. Each of the first six rows presents distribution of travel time estimates (columns 2–5), regression results (for the primary coefficient on travel cost to Taylor Mountain, column 7), estimated per-trip consumer surplus (column 8), and a comparison to the primary consumer surplus estimates (column 9), based on different assumptions on average travel speed for the linear zip code specification. The seventh row presents the same information for the route-based zip code specification.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Average MPH	Distribution of Travel Time Calculations (minutes)					Primary Coeff.	Mean CS_i/visit	Percent Diff.
	Mean	Median	SD	Min	Max			
55	5.67	3.56	5.02	1.34	45.4	−0.1364	$7.33	46.51%
45	6.93	4.35	6.14	1.64	55.49	−0.1349	$7.42	45.88%
35	8.91	5.59	7.89	2.1	71.34	−0.1321	$7.57	44.74%
25	12.47	7.83	11.05	2.94	99.88	−0.1262	$7.92	42.18%
15	20.79	13.04	18.42	4.9	166.47	−0.1104	$9.06	33.89%
10	31.18	19.57	27.63	7.36	249.7	−0.0910	$10.98	19.83%
Route-Based	28.94	17.58	30.9	10.8	122.3	−0.1154	$8.67	36.74%
Primary	17.71	14.75	9.75	3.37	56.58	−0.0730	$13.70	–

Next class

- Hedonics:
  - Two papers:
    1. Kuwayama et al. (2022)
    2. Christensen and Timmins (2021)

- Case Study due Sept 25th by 11:59pm
  - Estimate a travel cost or hedonic model.
  - Data and code (or at least most of it) will be provided.
  - Your job will be to submit a writeup of your results, explain the intuition, any potentially missing information/data that you would have like to have included in the analysis.
  - You will also submit your code.