

THESIS

EXPLORING THE USE OF CREDIT/DEBIT CARD TRANSACTION DATA IN ESTIMATING
NATIONAL PARK VISITOR SPENDING: A MOUNT RUSHMORE CASE STUDY

Submitted by

Evan Stockmoe

Department of Agricultural and Resource Economics

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 2024

Master's Committee:

Advisor: Jesse Burkhardt

Jude Bayham

Tim Komarek

Copyright by Evan Stockmoe 2024

All Rights Reserved

ABSTRACT

EXPLORING THE USE OF CREDIT/DEBIT CARD TRANSACTION DATA IN ESTIMATING NATIONAL PARK VISITOR SPENDING: A MOUNT RUSHMORE CASE STUDY

Visitor spending refers to trip-related expenditures made by tourists. Estimates of visitor spending are needed for national park economic contribution studies, and are an essential component in evaluating tourism-related economic activity [Thomas et al., 2019, Wilton and Nickerson, 2006]. Traditional methods for estimating visitor spending rely on visitor surveys, which are costly and subject to multiple forms of survey bias [Stynes and White, 2006, Sinclair et al., 2023, Wilton and Nickerson, 2006]. Using Mount Rushmore National Memorial as a case study, I explore the use of granular credit and debit card transaction data to estimate visitor spending without the need for survey data.

I use Safegraph *Spend* as the source of credit and debit card transaction data. I gather transaction data at stores within 100 miles of Mount Rushmore for the period 2019-2023. Combining this data and Mount Rushmore visitation data from the National Park Service, I develop and compare multiple models that use fixed effect regressions to estimate average spending per visit in the Accommodation, Food Service, Retail, and Arts/Entertainment/Recreation sectors. Using results from the best performing model, for 2022, I estimate Mount Rushmore visitor spending to be \$41.0 million in the Accommodation sector, \$38.4 million in the Food Service sector, \$154.8 million in the Retail sector, and \$6.4 million in the Arts/Entertainment/Recreation sector. I compare these estimates with NPS survey-based Mount Rushmore visitor spending estimates. I find that model estimates for the Food Service and Retail sectors are statistically indifferent to NPS estimates, however, model estimates for the Accommodation and Arts/Entertainment/Recreation sectors are below NPS estimates.

I find several strengths and weaknesses in the credit/debit card transaction models. One strength is the use of observed spending data rather than stated spending data. Another strength is the representation of yearly visitor spending habits. A third strength is the ability to provide measures of estimate precision, like standard errors. One weakness is the inability to identify park-specific visitor spending when other nearby tourist attractions have similar visitation. Another weakness is the failure to account for visitor trip purpose. Additionally, I find that Safegraph *Spend* underrepresents spending in the Accommodation sector since vacation rental websites (like Airbnb) are not reflected in the data. Looking forward, further research should focus on methodological refinement and the integration of other data sources to improve visitor spending estimation using credit/debit card transaction data.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Jesse Burkhardt, and my committee member, Dr. Jude Bayham, for their input and guidance throughout the course of this thesis. I would also like to thank my friends, family, peers, and loved ones for their camaraderie and support.

TABLE OF CONTENTS

	ABSTRACT	ii
	ACKNOWLEDGEMENTS	iv
	LIST OF TABLES	vi
	LIST OF FIGURES	vii
Chapter 1	Introduction	1
1.1	Mount Rushmore National Memorial Background Information	4
Chapter 2	Data	6
2.1	SafeGraph	6
2.2	National Park Service Visitor Use Statistics	13
2.3	National Park Service Traffic Counter Data	14
2.4	Merging data	16
Chapter 3	Methods	18
3.1	Base Model	18
3.2	Binning Spending by Distance	19
3.3	Incorporating Other National Parks	22
3.4	Using Year-by-month Fixed Effects	23
3.5	Scaling	23
3.6	Estimating Mount Rushmore Visitor Spending for 2022	24
Chapter 4	Results and Discussion	26
4.1	Base Model	26
4.2	Binning Spending by Distance	29
4.3	Incorporating Other National Parks	33
4.4	Using Year-by-month Fixed Effect	35
4.5	Comparison to NPS Visitor Spending Estimates	37
Chapter 5	Conclusion	42

LIST OF TABLES

2.1	Correlation between Mount Rushmore visits, South Dakota cardholders included in <i>Spend</i> , and cardholders for rest of the US included in <i>Spend</i>	17
4.1	Base Model	26
4.2	Base Coefficients Scaled	28
4.3	Binned Model	30
4.4	Binned Coefficients Scaled	32
4.5	Base Model with Badlands and Wind Cave	34
4.6	Year-by-Month Fixed Effects Model	35
4.7	Year-by-month Fixed Effects Model Coefficients Scaled	36
4.8	NPS estimates and model estimates for 2022 Mount Rushmore visitor spending	38

LIST OF FIGURES

1.1	Mount Rushmore National Memorial	4
2.1	Distribution of POIs around Mount Rushmore	8
2.2	Safegraph <i>Spend</i> weekly consumer spending totals by sector	10
2.3	Monthly <i>Spend</i> sample from the United States and South Dakota	11
2.4	Weekly recreation visitors for Mount Rushmore, Badlands, and Wind Cave	16
3.1	Distribution of POIs around Mount Rushmore, binned by distance	20
4.1	Mount Rushmore visits coefficients across bins	31

Chapter 1

Introduction

Designated outdoor recreation areas, like national parks, are often key components within a regional economy. Groups who visit outdoor recreation parks spend money along the way (at restaurants, gas stations, hotels, etc.). The money spent by visitors related to their trip is referred to as visitor spending, and it generates economic activity within the local economy. An estimate of economic activity directly resulting from visitor spending is useful information for policymakers, park managers, and researchers. These estimates are also often used to justify park funding.

Performing an economic contribution analysis is a commonly used method to estimate the economic activity related to a particular industry, event, or attraction [Liu et al., 2022]. The National Park Service (NPS) uses this method to assess the economic importance of their national parks [Flyr and Koontz, 2023]. Economic contribution studies estimate the gross economic activity resulting from visitor spending, accounting for the direct, indirect, and induced spending effects¹. Performing a park economic contribution analysis requires two components: an estimate of annual visitor spending by industry and an Input-Output model of the park's local economy. In this research, I focus on the visitor spending estimation component, exploring a novel estimation approach using credit/debit card transaction data. I use Mount Rushmore National Memorial as a case study, estimating park visitor spending. I then compare model estimates with NPS estimates. Mount Rushmore National Memorial was recently surveyed by the NPS in 2022, and has newly published visitor spending estimates available in the 2022 National Park Visitor Spending Effects VSE Report [Flyr and Koontz, 2023].

To estimate visitor spending for their analysis, the NPS performs an on-site survey, called the Social Economic Monitoring (SEM) survey. Visitors report expenditures in the local economy

¹Importantly, economic contribution analysis should not be confused with economic impact studies or economic benefit studies. These estimate net impacts and social welfare effects, which are distinctly different from the gross effects estimated by economic contribution studies [Watson et al., 2007].

related to their visit to the park, as well as other information like group size and trip purpose [Thomas et al., 2019]. The survey data is used to make visitor profiles. Visitor profiles are then combined with annual park visitation estimates to generate an estimate of annual visitor spending². Visitor surveys are widely used when estimating visitor spending at parks. Survey methods are well-established in the literature and accepted as an estimation approach. However, there are limitations to the survey approach.

Traditional survey methods are prone to multiple forms of bias, like interviewer effects, recall bias, telescoping errors, self-selection bias, and non-response bias [Sinclair et al., 2023, Wilton and Nickerson, 2006]. Interviewer effects refer to the influence interviewer behavior can have on responses during survey administration. Recall bias refers to respondents misreporting trip details due to incorrect recollection of trip spending. For example, Howard et al. [1991] found that visitors were more likely to under-report visitor spending on surveys when trips last longer than 10 days. Telescoping errors are defined by Stynes and White [2006] as “when the subject reports expenses that are not supposed to be included, for example, reporting expenses outside the study region, beyond the time frame defined, or even for a different trip.” Non-response bias can occur as well, particularly when mail surveys are used [Stynes and White, 2006]. Finally, bias can be present because survey data only contains visitor information from a snapshot in time, which may not be an accurate representation of visitors throughout the entire year [Sun, 2005]. If surveys are not carefully designed and administered these forms of bias can result in inaccurate visitor data, leading to inaccurate visitor spending estimates.

Visitor surveys are also costly and time-intensive to perform [Wilton and Nickerson, 2006]. The cost of administering a visitor survey prevents many parks from generating economic contribution estimates. Parks that administer visitor surveys typically do so infrequently, which can lead to the use of outdated data in estimation. For example, while the NPS is responsible for generating economic contribution estimates for each of the 400+ units in the NPS system, they only perform visitor surveys at 24-30 units a year. Non-surveyed parks rely on survey data from a previous year

²For an in-depth description of the NPS methodology see Thomas et al. [2019].

or use a benefit transfer approach [Flyr and Koontz, 2023]. This can potentially result in outdated or inaccurate visitor spending estimates, particularly for parks that have never performed a visitor survey or have not updated their survey data for many years.

Mount Rushmore National Memorial is an example of this. In 2021, the NPS estimated total annual Mount Rushmore visitor spending at \$164 million [U.S. National Park Service, 2023]. The following year, the SEM survey of Mount Rushmore visitors was performed, updating trip characteristic data, spending data, and local area definition. Using the updated survey data in 2022, the NPS estimated \$386 million in total park visitor spending, a \$222 million increase [Flyr and Koontz, 2023]. This example demonstrates how outdated survey data can impact visitor spending estimation. The large change in estimation from 2021 to 2022 shows that accurate visitor spending estimation requires current visitor spending data. This is not always possible using visitor surveys due to the amount of resources they require.

The two large drawbacks of visitor survey methods are that they rely on stated data, which can be subject to various biases, and they are resource-intensive. The limitations of the visitor survey approach suggest that there are benefits from visitor spending estimation that is not reliant on visitor surveys and stated data. With the recent creation of a credit and debit card transaction dataset by United States-based analytics company SafeGraph, there is an opportunity to econometrically estimate annual park visitor spending without the need for survey data. This allows for the use of revealed behavior rather than stated behavior in visitor spending estimation. It also allows for the consideration of year-round visitor spending habits. Most importantly, it is a cheaper method of estimation.

In this research, I present multiple fixed-effect models that attempt to estimate Mount Rushmore visitor spending using credit/debit card transaction data as the source of visitor expenditure information. I then compare estimates from the best performing model with NPS Mount Rushmore visitor spending estimates. Along the way, I detail *Spend* data features, quirks, and limitations in its use in estimating park visitor spending. This research has two main contributions. First, it presents a novel visitor spending estimation method applied to Mount Rushmore National Memo-



Figure 1.1: Mount Rushmore National Memorial

rial. At the time of this writing, to my knowledge, there is currently no research in the literature that uses credit/debit card spending data to estimate visitor spending. Second, it provides insight into SafeGraph data regarding its use in outdoor recreation visitor spending research.

1.1 Mount Rushmore National Memorial Background Information

In this section, I present background information regarding Mount Rushmore National Memorial tourism. It informs the scope of the analysis, identifying the appropriate geographic area for the analysis as well as the economic sectors relevant to visitor spending. It also provides readers with context for the memorial and the surrounding tourism industry.

Mount Rushmore National Memorial is a sculpture of the faces of 4 historical presidents of the United States, carved on the side of a mountain. It is a national destination that attracts visitors from all over the country [Otak et al., 2023]. Viewing the memorial is the main activity available for park visitors [U.S. National Park Service, 2024]. Paved roads allow visitors to drive to memorial viewing spots. Mount Rushmore is located in the Black Hills region of South Dakota. This region sits along the western border of South Dakota and is a popular tourist destination. In 2022, the

Black Hills region saw 38.3% of all tourist spending in South Dakota [TourismEconomics, 2023]. In addition to Mount Rushmore, multiple other outdoor recreation attractions are located in this region, including Badlands National Park, Custer State Park, Wind Cave National Park, and Crazy Horse Memorial. The region also contains the second largest city in the state, Rapid City, South Dakota, and several smaller municipalities like Deadwood, South Dakota, and Spearfish, South Dakota. I use a 100-mile radius around Mount Rushmore to represent the Mount Rushmore local economy, as this captures the Black Hills region and the possible spending locations for visitors to Mount Rushmore. This definition of the local economy for Mount Rushmore is similar to the Mount Rushmore local economy definition used by the NPS. [Otak et al., 2023].

To inform which economic sectors I should examine when estimating Mount Rushmore visitor spending, I use the 2022 NPS Visitor Spending Effects report, which includes estimates of 2022 Mount Rushmore visitor spending. The NPS estimates Mount Rushmore visitor spending on hotels, campsites, gas, groceries, restaurants, retail, recreation, and local transportation [U.S. National Park Service, 2023]. I also use Tourism Economics' "Economic Impact of Tourism in South Dakota - County Results 2022" report, which provides 2022 visitor spending estimates for the Black Hills region. In the report, Tourism Economics estimates Black Hills visitor spending in the lodging, food and beverage, recreation, retail, and transportation industries. I estimate visitor spending in the Accommodation, Food Service, Retail, and Art/Entertainment/Recreation sectors, using the North American Industry Classification System (NAICS) to define sectors. These four sectors cover the same spending categories as the NPS and Tourism Economics reports. As defined by the North American Industry Classification System (NAICS), the Accommodation sector includes hotel and campsite expenditures, the Food Service sector includes restaurant expenditures, the Retail sector includes grocery, gas, and retail expenditures, and the Art/Entertainment/Recreation sector includes recreation expenditures. Local transportation is the only category used by the NPS that I do not include in the analysis. However, the NPS estimates that local transportation only makes up 0.4% of Mount Rushmore visitor spending, so omitting this category should not be too impactful [Flyr and Koontz, 2023].

Chapter 2

Data

Three data sources are used in this research: Safegraph, National Park Service Visitor Use Statistics, and National Park Service Traffic Data.

2.1 SafeGraph

The data processing company Safegraph provides a database of global points of interest (POIs), referred to as their *Places* dataset. A point of interest (POI) indicates a specific physical location, like a business, store, or landmark. As of January 2024, over 51 million POIs are included in *Places*. *Places* contains data on POI name, location, and commercial details. In addition to *Places*, Safegraph provides a dataset containing POI debit and credit card transaction data, designed to be used in conjunction with *Places*. It is referred to as their *Spend* dataset. In this research, I use *Places* to identify relevant businesses around Mount Rushmore National Memorial. Then, I use *Spend* to gather consumer spending data at these businesses for the period of 2019-2023.

Places covers multiple countries around the globe, however, POIs in the United States make up the majority of the dataset. *Places* is updated monthly to improve POI coverage and refine existing POI entries. For the United States, *Places* includes the majority of stores and businesses. Each quarter, Safegraph estimates the coverage rate of their US *Places* data by randomly selecting two zip codes and comparing POI data with other industry-leading POI databases like Google's. With their January 2024 release, Safegraph estimates an 82% coverage rate, meaning that of the zip codes examined, an estimated 82% of all POIs are included in *Places*. Previously estimated coverage rates have ranged from 72%-92%, depending on the zip codes being examined and the version of *Places* being assessed [SafeGraph, 2024]. For this research, I assume an 82% coverage rate of POIs in the Mount Rushmore area. This percentage is consistent with the current and previous coverage rate estimates.

From the *Places* dataset, I use four variables: “placekey”, “latitude”, “longitude”, and “naics_code”. “placekey” uniquely identifies each POI. “latitude” and “longitude” provide the location of each POI, which I use to filter *Places* to only include POIs within 100 miles of Mount Rushmore National Memorial. “naics_code” provides the North American Industry Classification System (NAICS) code of each POI, which I use to identify each POI’s economic sector. Based on each POI’s North American Industry Classification System (NAICS) code, I filter *Places* further, only including POIs in the Accommodation, Food Service, Retail Trade, and Arts/Entertainment/Recreation sectors. The retained POIs represent stores, businesses, and landmarks within 100 miles of Mount Rushmore that operate within the economic sectors relevant to Mount Rushmore visitor spending. I then merge these POIs with the *Spend* dataset to gather available consumer spending data for each POI.

Spend is a panel dataset, containing observations of consumer spending at individual POIs through time, beginning in 2019. The dataset is generated from anonymized debit and credit transactions sourced from thousands of financial institutions [SafeGraph, 2023a]. Each month, transactions are assigned to the appropriate POI. Transactions at each POI are then aggregated to generate metrics like POI daily spend, monthly spend, and monthly number of customers. The metrics reflect transactions made by the cardholders whose information is sourced from the financial institutions. Because the spending metrics are organized by POI, *Spend* shows the spatial distribution of cardholder spending across the United States. I use *Spend* to estimate how cardholder spending fluctuates around Mount Rushmore through time. New *Spend* data is continually released on a monthly basis. As of January 2024, over 11 million cardholders are represented in *Spend* [SafeGraph, 2023b].

Spend is organized by month by POI, meaning each row in the dataset represents one month of spending at one POI. From the *Spend* dataset, I use three variables: “spend_by_day”, “spend_date_range_start”, and “spend_date_range_end”. The “spend_by_day” variable represents POI daily in-person spending during a given month. It contains each daily spend total for the month, organized together in a string and presented as one observation. For every “spend_by_day”

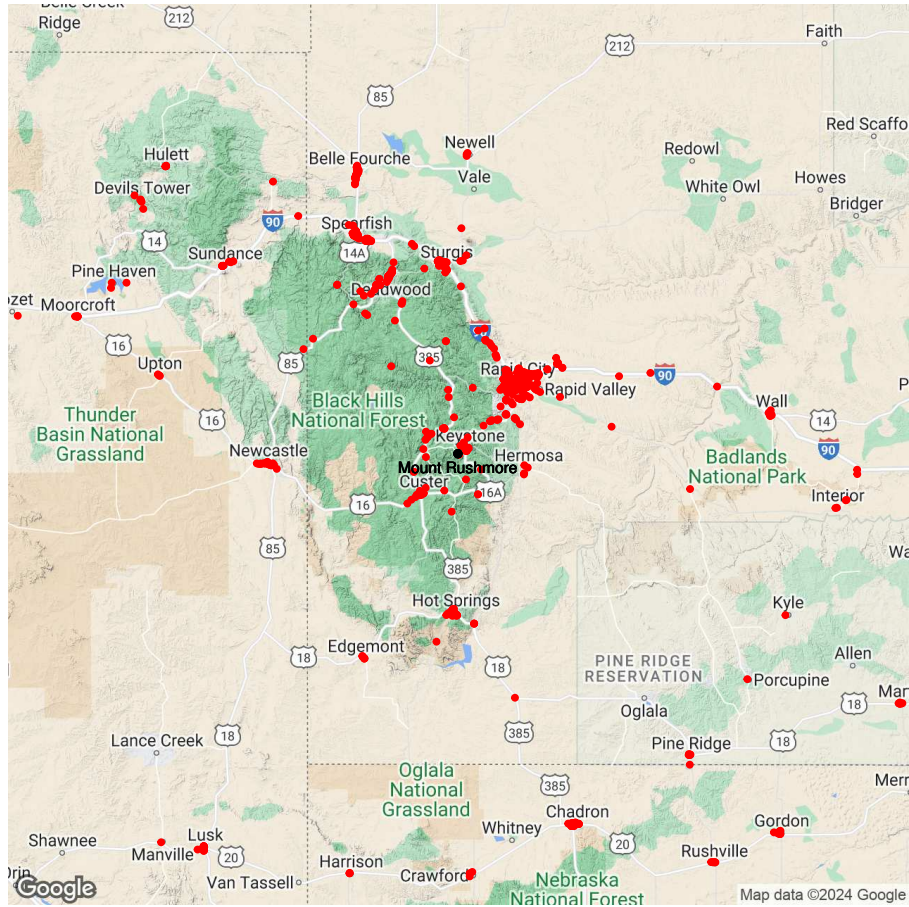


Figure 2.1: Distribution of POIs around Mount Rushmore

observation, I break up the string of daily spend totals and assign each to the appropriate date using the “spend_date_range_start” and “spend_date_range_end” variables. These variables indicate the starting date and ending date of the month represented by the “spend_by_day” observation. I separate the “spend_by_day” string, and attach the starting date of the month to the first daily spend total, and attach the ending date of the month to the last daily spend total. Once the starting and ending dates for the string have been established, I fill in the dates for the in-between daily spend totals. I do this for every “spend_by_day” observation, reorganizing the data into individual observations of daily spend by POI.

However, the daily spend totals are inaccurate for Saturdays, Sundays, and Mondays. Weekend spending is artificially low, and Monday spending is artificially high. This is because some card

transactions occurring on the weekend are not processed until Monday³. To prevent this from impacting the analysis, I aggregate POI daily spend by week. I use a Tuesday-Monday week to group Monday and the previous weekend together. I do this to ensure that Monday's inflated spending is associated with the previous weekend. The aggregation results in observations of POI weekly spend. I then sum the observations of POI weekly spend by economic sector, using each POI's North American Industry Classification System (NAICS) code to identify its' economic sector. The resulting dataset contains observations of weekly spend for the Accommodation, Food Service, Retail, and Arts/Entertainment/Recreation sectors. The sector weekly spend totals reflect spending made at stores and businesses within 100 miles of Mount Rushmore. I use these observations of sector weekly spend to estimate how consumer spending around Mount Rushmore fluctuates over time.

Figure 2.2 shows the weekly spending totals from the cardholders included in *Spend* for POIs within a 100-mile radius of Mount Rushmore in the Accommodation, Food Service, Retail, and Arts/Entertainment/Recreation sectors. The Retail sector contains the highest levels of spending. Following Retail, Food Service is the next largest sector, then Accommodation, then Arts/Entertainment/Recreation. Seasonal spending fluctuations are present, particularly for the Accommodation and Arts/Entertainment/Recreation sectors. These fluctuations are primarily driven by tourism. As detailed in the Mount Rushmore background section, this area of South Dakota sees lots of summer tourism, primarily for outdoor recreation activities. Tourism declines in the winter months, driven by cold weather and a lack of winter tourist activities. Less tourism-dominated sectors, like Food Service and Retail Trade, see some seasonal spending fluctuations, but overall have steadier consumer spending. Additionally, spending declines rather significantly at the end of 2022 for the Retail and Food Service sectors, and does not recover. This decline is likely due to fluctuations in underlining cardholder data used to generate *Spend*.

The number of cardholders represented in the *Spend* data fluctuates from month to month. This is because the number of credit and debit cards open and active at financial institutions fluctuates

³In this data, Mondays have the highest average spending, slightly less than double the next highest day.

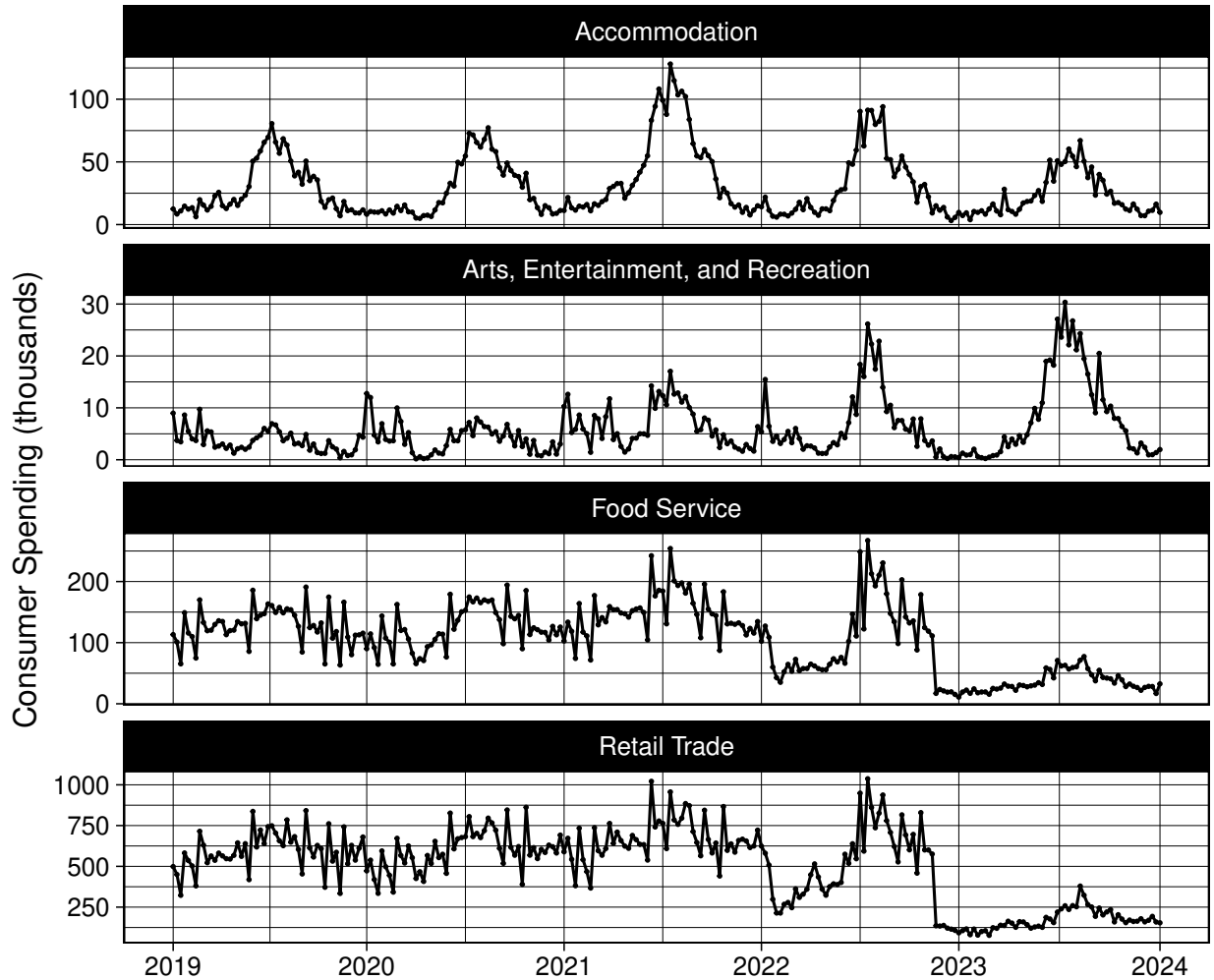


Figure 2.2: Safegraph *Spend* weekly consumer spending totals by sector

month to month. For example, if a new customer opens a credit card at one of the financial institutions used by *Spend*, the number of cardholders represented would increase by one. Or, if a customer closes a bank card, the number of cardholders represented would decrease by one. Cardholder fluctuations likely have some impact on the level of spending recorded. More customers represented could result in more observed spending, and fewer customers could result in less observed spending. To allow researchers to consider this potential impact on the *Spend* data, Safegraph publishes summary data about the card panel used each month to generate *Spend*. The summary data shows the number of unique cardholders represented in *Spend*. It also shows estimates of the number of unique cardholders who reside in each state. The observed sector spending

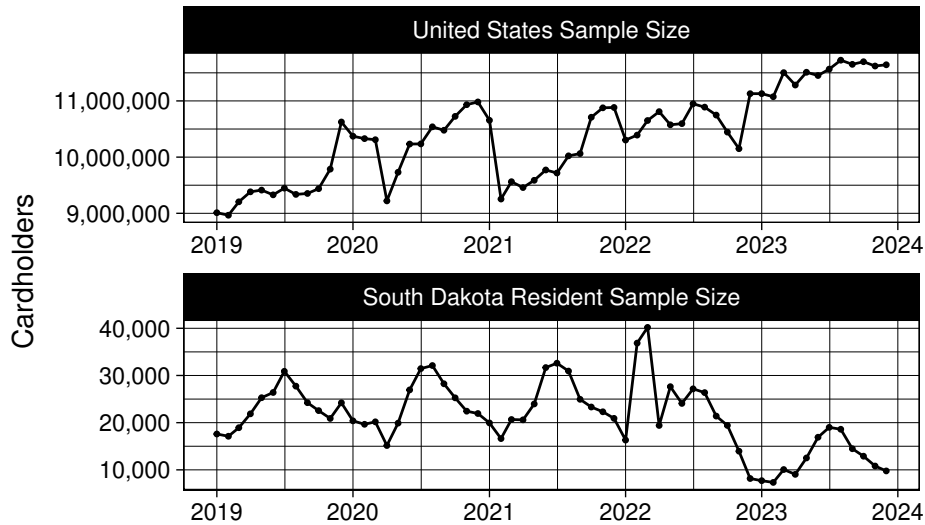


Figure 2.3: Monthly *Spend* sample from the United States and South Dakota

I use in this analysis reflects spending by both South Dakota cardholders and cardholders who live elsewhere in the United States. To be clear, if a cardholder represented in *Spend* resides in Colorado and travels to South Dakota and spends money, their spending is recorded in South Dakota but they are not counted as a South Dakota cardholder. However, if a cardholder represented in *Spend* resides in South Dakota and spends in South Dakota, their spending is recorded in South Dakota and they are counted as a South Dakota cardholder. Fluctuations in either of these metrics can impact observed spending. Figure 2.3 graphically displays the number of United States monthly cardholders included in *Spend*, as well as the number of monthly cardholders Safegraph estimates to be South Dakota residents.

The total number of United States cardholders represented in *Spend* follows an increasing trend over time. However, the number of cardholders who are South Dakota residents follows a seasonal trend. Additionally, it declines significantly at the end of 2022. The seasonal pattern can be explained by seasonal residents - those who live part-time in South Dakota in the summer months. But, it is unclear what causes the decline at the end of 2022⁴. Regardless of the cause, the decrease

⁴Safegraph has performed an analysis assessing potential geographic sampling bias in the *Spend* data. This analysis compared sample size changes to population changes for each state. It found that the average geographic bias per state was less than 1% [Ho, 2021]. Notably, this analysis only examined data from 2019-2021. Data from 2021-2023 have not been assessed and may contain higher rates of bias depending on the state.

in *Spend* cardholders residing in South Dakota aligns with the decrease in consumer spending observed in the Retail and Food Service sectors. It is likely that the reduction in South Dakota residents included in the panel caused the reduction in observed spending in these sectors. It is probable that the Retail and Food Service sectors have higher proportions of resident spending, and are more impacted by a decrease in South Dakota cardholders included in *Spend*. The Accommodation and Arts/Entertainment/Recreation sectors seem to be less impacted by the 2022 decrease in South Dakota cardholders included in *Spend*. These sectors may be more dominated by out-of-state visitor spending.

The number of cardholders included in *Spend*, and the residential distribution of those cardholders, can impact observed consumer spending levels. To mitigate the effects of the fluctuating number of *Spend* cardholders both in South Dakota and the rest of the United States, I use *Spend* cardholder metrics in the analysis. I use the monthly number of South Dakota cardholders represented in *Spend*. This metric strictly reflects the number of cardholders included in *Spend* that Safegraph estimates to be residents of South Dakota. I also use the monthly number of United States cardholders represented in *Spend* for the rest of the United States. To obtain this metric, I subtract the monthly South Dakota cardholders from the monthly United States cardholders. This metric strictly reflects the number of cardholders in the United States estimated to be living outside South Dakota. The observed consumer spending data from *Spend* that I use in this analysis reflect spending by both South Dakota cardholders and cardholders who live elsewhere in the United States. However, the impacts from cardholder fluctuation seem to be more significant for South Dakota cardholders and less significant for non-South Dakota cardholders. Because of this, I use separate metrics for South Dakota monthly cardholders and cardholders for the rest of the United States. By separating I can better account for their varying effects in the analysis.

I also explored the possibility that the observed consumer spending is impacted by underlying changes in the POIs represented in the data. For example, business NAICS classification could change, re-segmenting the POI out of one sector into another. If many businesses were reclassified, a sector may see a change in spending. In the *Spend* data I use in this analysis, there are no cases

of POIs being reclassified. Spending at each POI is attributed to the same sector for the duration of the period examined.

2.2 National Park Service Visitor Use Statistics

The National Park Service Visitor Use Statistics database contains park visitation estimates. Monthly visitation estimates are published for each NPS unit. For a few units, like Mount Rushmore, daily visitation estimates are published as well. Visitation estimation procedures vary across parks depending on specific park attributes. Generally speaking, traffic counters are used to source raw counts of visits. Raw counts are combined with other information like persons-per-vehicle to produce visitation estimates. Complete documentation on estimation procedures at each park can be found on the NPS Visitor Use Statistics database [U.S. National Park Service].

I use the Mount Rushmore “Daily Report” data from 2019-2023, which provides estimated daily visits to Mount Rushmore. The estimated daily visits represent both recreation and non-recreation visits. Recreation visits refer to those who enter the park on a recreation trip, while non-recreation visits refer to those who enter the park for non-leisure purposes, like commuters and those on commercial business. Visitor spending estimates for economic contribution analyses should only include spending by visitors who attend the park for recreation or leisure purposes [Thomas et al., 2019]. Non-recreation visitors are not tourists, and their spending should not be included in visitor spending estimates. Because the Mount Rushmore daily visit estimates represent both recreation and non-recreation visits, I remove non-recreation visits from the daily totals. To do this, I multiply daily visits by the Mount Rushmore recreation adjustment factors used by the NPS [U.S. National Park Service, 1995]. For each month, these adjustment factors estimate the proportion of visitors to Mount Rushmore that are recreation visitors. I remove estimated non-recreation visits by multiplying daily visits with the appropriate monthly recreation adjustment factor.

From the NPS Visitor Use Statistics database, I also use the “Recreation Visits by Month” dataset for both Badlands National Park and Wind Cave National Park. This report provides esti-

mates of monthly recreation visits to each park. Data from 2019-2023 is used. I use the monthly recreation visits estimates in combination with park traffic counter data to generate estimates of daily recreation visits for Badlands and Wind Cave National Parks (as discussed in the National Park Service Traffic Count Data). I generate estimates of daily recreation visits for Badlands National Park and Wind Cave National Park because daily visit estimates for these parks are not provided directly by the NPS like they are for Mount Rushmore.

2.3 National Park Service Traffic Counter Data

The NPS operates several traffic counters located in NPS units across the country⁵. The traffic counters report the number of vehicles traveling along a given road. I use daily traffic counts at Badlands National Park and Wind Cave National Park. These were the only national park units in the Mount Rushmore local area with traffic data. Jewel Cave National Memorial, Minuteman Missile National Historic Site, and Devils Tower National Memorial are located in the Mount Rushmore local area but do not have traffic data available. That being said, Badlands National Park and Wind Cave National Park are the two most visited national parks in the local area, after Mount Rushmore.

For Badlands National Park, daily vehicle counts from traffic monitor N2603N serve as the primary source of traffic data. I use traffic data for 2019-2023. In June 2023 and July 2023, data at this counter is unavailable, so for these months, I use substitute data from another nearby park traffic counter in its place. In the traffic data used, there are 19 missing daily traffic observations. 12 missing observations occur in December 2022. Because a large amount occur in this month, I drop the entire month and treat the observations as NAs. The daily visit estimation procedure I use is based on distributing total monthly recreation visits according to traffic frequency. Multiple missing days within a month result in skewed daily visit estimates. As such, months with more than a few missing observations must be dropped. The remaining 7 missing observations are

⁵NPS unit traffic data can be accessed at <https://geocounts.com/traffic/us/nps>.

scattered across months, with no more than 2 occurring in a given month. These individual missing observations are treated as NAs.

For Wind Cave National Park, daily vehicle counts from traffic station N2801E serve as the source of traffic count data. I use traffic data for 2019-2023. Data is missing for June 2020, July 2020, and January 2021 at this counter. Unfortunately, there is no substitute data available at other traffic counters within the park for these months. With no data available, I treat these months as NAs. Beyond the missing months, there are also 11 missing daily observations scattered across months. The individual missing observations are treated as NAs.

I use the following equation to generate estimates of daily recreation visits at Badlands National Park and Wind Cave National Park:

$$DRV = DTC/MTC * MRV$$

where *DRV* is daily recreation visits, *DTC* is daily traffic count, *MTC* is monthly traffic count, and *MRV* is monthly recreation visits. To estimate daily recreation visits, I convert daily traffic counts to proportions of monthly traffic. I then multiply each proportion by monthly recreational visitors. This distributes monthly recreation visits to days of the month, according to traffic frequency. High-traffic days are assigned more visits and low-traffic days are assigned fewer visits. Estimating daily visits in this fashion ensures that the daily recreation visit estimates are calibrated to the monthly NPS recreation visit estimates. When summed by month, the daily recreation visit estimates exactly match the monthly recreation visit estimates provided by the NPS.

I then sum the daily recreation visits estimates for Mount Rushmore National Memorial, Badlands National Park, and Wind Cave National Park by week. I use a Tuesday-Monday week to match the *Spend* data weekly aggregation. Figure 2.4 graphically displays weekly recreation visits to Mount Rushmore National Memorial, Badlands National Park, and Wind Cave National Park. Visitation is seasonal at all three parks. Mount Rushmore attracts the highest amount of visits. It attracts over 50,000 visits a week in the summer months, significantly surpassing Badlands National Park and Wind Cave National Park. Mount Rushmore has side-by-side visitation peaks each

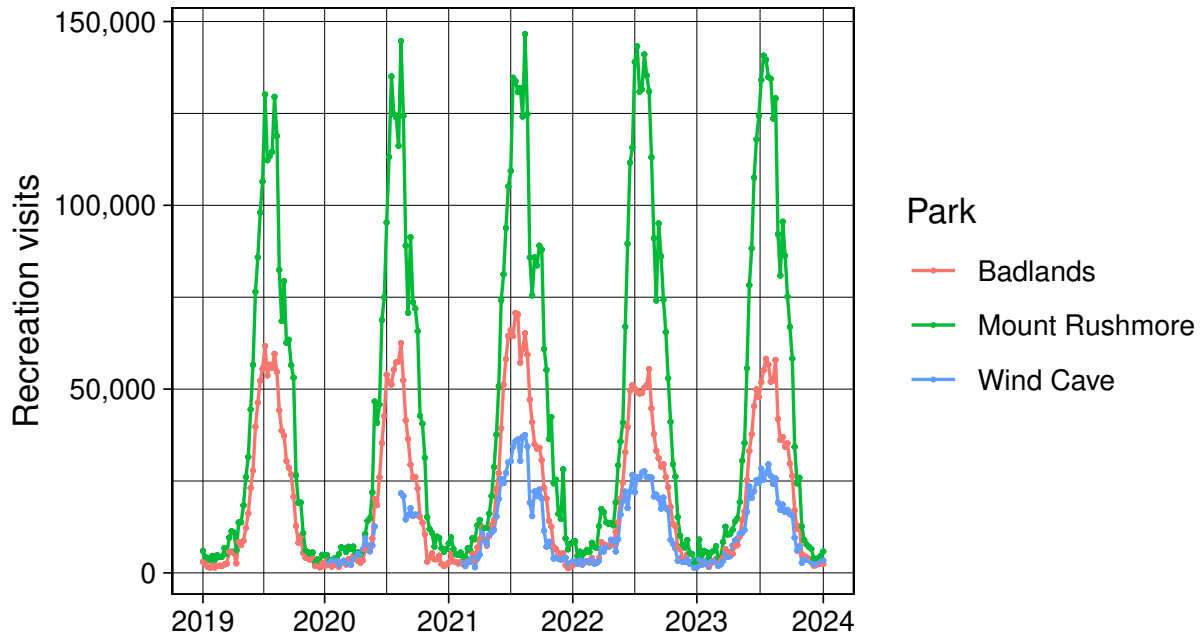


Figure 2.4: Weekly recreation visitors for Mount Rushmore, Badlands, and Wind Cave

summer. These peaks correspond with July 4 and the Sturgis Bike Rally. The Sturgis Bike Rally is a large annual motorcycle rally occurring during the first week of August each year in Sturgis, SD, located just under 50 miles from the park. There is also a visit bump occurring in September. This corresponds with Labor Day weekend. These yearly bumps show that there are structural visitation increases each year around certain holidays and annual events. In the winter months, visit rates at the three parks are more similar.

2.4 Merging data

I join the combined *Places* and *Spend* dataset with the monthly *Spend* cardholder metrics by date. I then join the National Park weekly visits data for Mount Rushmore National Memorial, Badlands National Park, and Wind Cave National Park by date. The final dataset contains weekly consumer spending at 1221 POIs across 4 sectors (Accommodation, Food Service, Retail Trade, and Arts/Entertainment/Recreation), the monthly *Spend* cardholders from South Dakota and the

rest of the United States, and recreation visits to Mount Rushmore National Memorial, Badlands National Park, and Wind Cave National Park. Data covers 2019-2023.

Table 2.1 shows the correlation between Mount Rushmore visits and the number of South Dakota cardholders represented in *Spend*, as well as the correlation between Mount Rushmore visits and the number of cardholders represented in *Spend* for the rest of the United States. The correlation between Mount Rushmore visits and South Dakota cardholders is moderate at 0.48. This correlation can be explained by seasonal trends. Mount Rushmore visitation is seasonal. Additionally, South Dakota has seasonal residents who live in the state only during the summer months. Both metrics tend to increase in the summer and decrease in the winter, and as such are moderately correlated. The correlation between Mount Rushmore visits and cardholders for the rest of the United States is weak at 0.05.

Table 2.1: Correlation between Mount Rushmore visits, South Dakota cardholders included in *Spend*, and cardholders for rest of the US included in *Spend*

	Mount Rushmore Visits
Mount Rushmore Visits	1.00
<i>Spend</i> South Dakota cardholders	0.48
<i>Spend</i> cardholders for the rest of the US	0.05

Chapter 3

Methods

To estimate visitor spending in the local economy from visitors to Mount Rushmore, I specify four fixed effect linear regression models. Each model is run for the Accommodation, Food Service, Retail Trade, and Arts/Entertainment/Recreation sectors. The models assume a linear relationship between Mount Rushmore visits and consumer spending. The coefficients on the Mount Rushmore visits variable reflect the average consumer spending correlated with each visit to Mount Rushmore.

3.1 Base Model

I run a regression for each economic sector to obtain a per-visit spending estimate by spending type. The sectors modeled are Accommodation, Food Service, Retail Trade, and Arts/Entertainment/Recreation. For each sector, I estimate the base model with the following specification:

$$y_t = \beta_{MR}x_{1t} + \delta x_{2t} + \psi x_{3t} + \gamma_m + \epsilon_t \quad (3.1)$$

where y_t is the total consumer spending in week t , x_{1t} is the total visits to Mount Rushmore in week t , x_{2t} is the number of South Dakota residents represented in the *Spend* data in week t , x_{3t} is the number of residents for the rest of the United States represented in the *Spend* data in week t , β_{MR} is the coefficient on the Mount Rushmore visits variable, δ is the coefficient on the South Dakota resident sample size variable, ψ is the coefficient on the rest of the United States sample size variable, γ_m is a month fixed effect, and ϵ_t is the error term in week t .

I include the monthly number of cardholders represented in *Spend* that live in both South Dakota and the rest of the United States in the model to capture changes in observed spending attributable to cardholder fluctuations. Ideally, there would be no fluctuation in the underlying sample size of cardholders, with a consistent sample through time. However, this is not the case.

I include the number of cardholders used to generate the *Spend* data as variables to account for the effects of the changing sample size. I include the number of cardholders from South Dakota as a separate variable from the total number of cardholders since the number of South Dakota cardholders represented appears to have a larger impact on spending in the Mount Rushmore area compared to the number of cardholders represented residing in other states. This is indicated by the large decline in Retail and Food Service spending aligning with the decline in South Dakota cardholders represented in *Spend*. δ and ψ theoretically capture the average impact on observed spending within the 100-mile radius around Mount Rushmore when the number of cardholders represented in *Spend* changes.

A month fixed effect is included to account for structural monthly changes in consumer spending seen every year. Western South Dakota's economy is seasonal, with structural spending increases in the summer months and decreases in the winter months. Including a month effect allows each month to have a different spending intercept. This better isolates the coefficient estimate on Mount Rushmore visits.

3.2 Binning Spending by Distance

Next, I adjust the base model to allow coefficients to change as POI distance from Mount Rushmore increases. I place POIs into 1 of 4 bins according to their distance from Mount Rushmore, and sum consumer spending within each bin. I then create dummy variables for each bin, and interact the dummy variables with the Mount Rushmore visits variable and both *Spend* sample size variables. The interaction allows coefficients to vary with each distance bin. I define the distance bins as follows:

- Bin 1: spending at POIs 0-25 miles from Mount Rushmore
- Bin 2 (d_2): spending at POIs 25-50 miles from Mount Rushmore
- Bin 3 (d_3): spending at POIs 50-75 miles from Mount Rushmore
- Bin 4 (d_4): spending at POIs 75-100 miles from Mount Rushmore

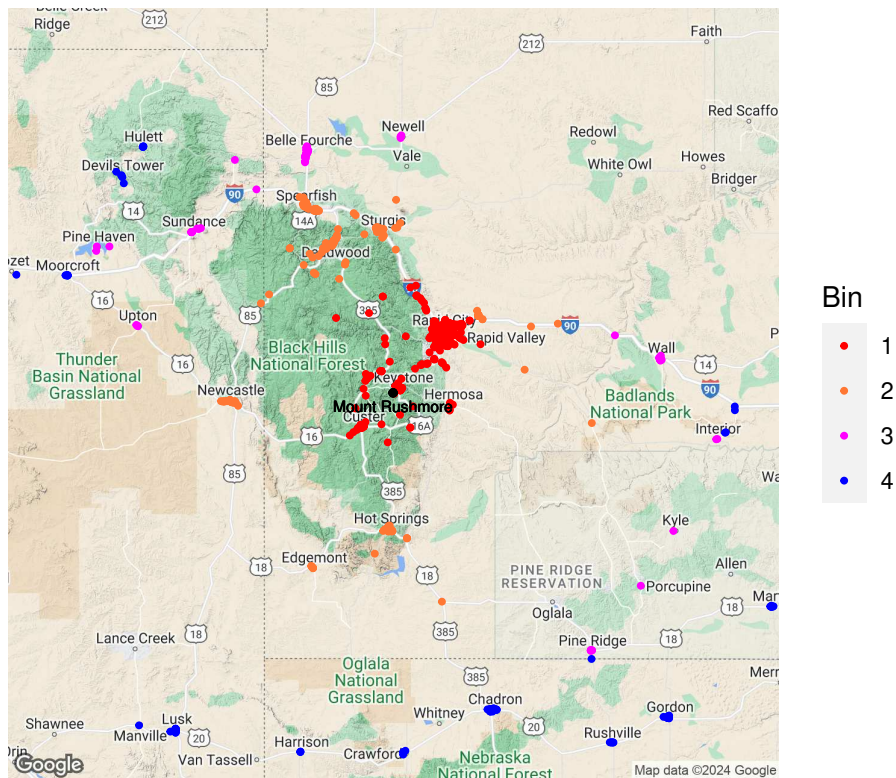


Figure 3.1: Distribution of POIs around Mount Rushmore, binned by distance

Figure 3.1 shows the distribution of binned POIs. Bin 1 includes Keystone, South Dakota, Custer, South Dakota, and Rapid City, South Dakota. Keystone and Custer are both gateway towns for tourists. Mount Rushmore is located in Keystone, and Crazy Horse Memorial, Custer State Park, and Wind Cave National Park are located near Custer. Rapid City is the largest city in western South Dakota and the second largest in the state. Bin 2 includes municipalities like Spearfish, South Dakota, Deadwood, South Dakota, and Hot Springs, South Dakota. While these towns are still tourist areas, they do not contain as many outdoor recreation attractions as bin 1. Bin 3 includes small rural towns and Badlands National Park. Bin 4 contains small rural towns, Badlands National Park, and Devils Tower National Memorial.

I run the following specification for each sector:

$$\begin{aligned}
y_{it} = & \beta_{MR,1}(d_1x_{1t}) + \beta_{MR,2}(d_2x_{1t}) + \beta_{MR,3}(d_3x_{1t}) + \beta_{MR,4}(d_4x_{1t}) + \\
& \delta_1(d_1x_{2t}) + \delta_2(d_2x_{2t}) + \delta_3(d_3x_{2t}) + \delta_4(d_4x_{2t}) + \\
& \psi_1(d_1x_{3t}) + \psi_2(d_2x_{3t}) + \psi_3(d_3x_{3t}) + \psi_4(d_4x_{3t}) + \\
& \gamma_{im} + \epsilon_{it}
\end{aligned} \tag{3.2}$$

where y_{it} is the total consumer spending in bin i in week t , x_{1t} is the total visits to Mount Rushmore in week t , x_{2t} is the number of South Dakota residents represented in the *Spend* data in week t , and x_{3t} is the number of residents for the rest of the United States represented in the *Spend* data in week t . d_2 , d_3 , and d_4 are dummy variables that indicate distance bin. $\beta_{MR,1}$ is the coefficient on the Mount Rushmore visits variable for bin 1. $\beta_{MR,2}$, $\beta_{MR,3}$, and $\beta_{MR,4}$ represent the coefficient change in $\beta_{MR,1}$ for bin 2, bin 3, and bin 4, respectively. δ_1 is the coefficient on the South Dakota resident sample size variable for bin 1. δ_2 , δ_3 , and δ_4 represent the coefficient change in δ_1 for bin 2, bin 3, and bin 4, respectively. ψ_1 is the coefficient on the rest of the United States sample size variable for bin 1. ψ_2 , ψ_3 , and ψ_4 represent the coefficient change in ψ_1 for bin 2, bin 3, and bin 4, respectively. γ_{im} is a combined month-bin fixed effect (meaning there is a fixed effect for each month for each bin). ϵ_{it} is the error term for bin i in week t .

This form removes the assumption that the variables have the same relationship to consumer spending across space. As distance from the park grows, coefficient estimates can now increase or decrease to reflect a changing relationship with spending. Intuitively, as distance from the park increases, I expect the coefficient on Mount Rushmore visits to decrease, since visitors are more likely to spend close to the park rather than far away. By allowing coefficients to change, a clearer picture of how visitor spending disperses through the economy is revealed.

I use a combined month-bin fixed effect to allow for month-fixed effects to vary by bin. The amount of consumer spending differs greatly across bins, with high spending in bin 1 and low spending in bins 3 and 4. Then, within each bin, consumer spending differs across months. Be-

cause month effects are different across bins, a combined month-bin effect is needed. Including this effect accounts for spending variation within each bin that is attributable to the month.

3.3 Incorporating Other National Parks

Next, I include Badlands National Park visits and Wind Cave National Park visits in the base model. This attempts to disentangle spending associated with Mount Rushmore visits from spending associated with Badlands visits and Wind Cave visits. However, visits to the three parks are highly correlated with each other. Pearson correlation coefficients are all over 0.94. The high correlation is because visits to the parks all occur during similar times of the year. Given their proximity, it is likely that these parks see lots of the same visitors as well. Because the correlation is so high, problematic multicollinearity is present in this specification. However, I still run the model to see how the Mount Rushmore visits coefficient changes in the presence of visits at other nearby parks (and multicollinearity). I run the following specification for each sector:

$$y_t = \beta_{MR}x_{1t} + \delta x_{2t} + \psi x_{3t} + \beta_{WC}x_{4t} + \beta_{BL}x_{5t} + \gamma_m + \epsilon_t \quad (3.3)$$

where y_t is the total consumer spending in week t , x_{1t} is the total visitors to Mount Rushmore in week t , x_{4t} is the total visitors to Wind Cave in week t , x_{5t} is the total visitors to Badlands in week t , β_{MR} is the coefficient on the Mount Rushmore visitation variable, β_{WC} is the coefficient on the Wind Cave visitation variable, and β_{BL} is the coefficient on the Badlands visitation variable. x_{2t} is the number of South Dakota residents represented in the *Spend* data in week t , x_{3t} is the number of residents for the rest of the United States represented in the *Spend* data in week t , δ is the coefficient on the South Dakota resident sample size variable, ψ is the coefficient on the rest of the United States sample size variable, γ_m is a month fixed effect, and ϵ_t is the error term in week t .

3.4 Using Year-by-month Fixed Effects

Lastly, I remove the *Spend* sample size variables from the base model and add a year-month combined fixed effect. This specification accounts for the consumer spending impacts from the fluctuating *Spend* sample size using a fixed effect, instead of including the sample size variables in the regression. The year-month combined fixed effect allows each month each year to have a different intercept. Since the number of *Spend* cardholders for South Dakota and for the rest of the United States vary monthly, but the consumer spending and Mount Rushmore visit variables vary by week, the year-month fixed effect accounts for any baseline change to monthly consumer spending that is driven by a change in monthly *Spend* cardholders. The fixed effect also captures the structural changes to consumer spending seen each month due to western South Dakota's seasonal economy. The use of a year-by-month fixed effect allows me to isolate and estimate the relationship between Mount Rushmore visits and consumer spending because the fixed effect accounts for monthly variations in consumer spending driven by both the changing *Spend* sample and the time of the year. I run the following specification for each sector.

$$y_t = \beta_{MR}x_{1t} + \gamma_{ym} + \epsilon_t \quad (3.4)$$

where y_t is the total consumer spending in week t , x_{1t} is the total visits to Mount Rushmore in week t , and γ_{ym} is a combined year-month fixed effect (meaning each month each year has a fixed effect). β_{MR} is the coefficient on the Mount Rushmore visitation variable, and ϵ_t is the error term in week t .

3.5 Scaling

Following the estimation of each model, I scale the Mount Rushmore visits to reflect total visitor spending. Because *Spend* measures only a sample of consumers at a sample of POIs, but park visitation data is for the entire population, the regression coefficients only reflect the correlation between visits and spending within the sample. They do not reflect spending made by

the entire population of visitors, or spending made at all POIs. Assuming a representative sample of cardholders and stores, I scale the park visitation coefficients in two ways.

First, I scale Mount Rushmore coefficients by the population sample rate. The population sample rate is equal to the average *Spend* sample size for the United States divided by the average United States population for 2019-2023. This calculation results in an estimated population sample rate of about 3.12% ($10390637/332691300 \approx 3.12\%$). I divide each coefficient by the sample rate so they reflect all consumer spending, not just sampled consumer spending. Second, I scale coefficients by POI sample rate. As outlined in section 2.1, in this research I assume an 82% POI coverage rate of Safegraph's *Places* dataset. Since *Spend* is based on the POIs identified in *Places*, I assume that *Spend* has the same POI coverage rate by extension. This means that spending at an estimated 82% of stores is captured in the data. Spending at an estimated 18% of stores is not captured, even if a cardholder that is represented in the *Spend* sample spends there. I divide each Mount Rushmore coefficient by 82% so they reflect all POIs, not just sampled POIs.

3.6 Estimating Mount Rushmore Visitor Spending for 2022

I use the scaled Mount Rushmore visits coefficients from the year-by-month fixed effects model to approximate annual Mount Rushmore visitor spending for 2022 in each sector. I then compare with NPS estimates of 2022 Mount Rushmore visitor spending. The following equation is used:

$$VS_{2022,s} = \omega_{MR} \times v_{2022} \quad (3.5)$$

$VS_{s,2022}$ represents 2022 visitor spending in sector s , ω_{MR} represents the scaled β_{MR} coefficient, and v_{2022} represents total recreation visits to Mount Rushmore in 2022. For each sector, ω_{MR} represents the estimated average spending correlated with each Mount Rushmore visit for the 100-mile radius around the park. I multiply ω_{MR} by the total recreation visits to Mount Rushmore for 2022 to calculate an estimate of the total Mount Rushmore visitor spending in 2022. I also

calculate 95% confidence intervals using the same equation, replacing ω with the scaled upper and lower bound estimates for each Mount Rushmore visits coefficient.

Chapter 4

Results and Discussion

The following section presents, interprets, and discusses the results for each of the four models described in the Methods section.

4.1 Base Model

For the base model, in all sectors there is a significant positive association between Mount Rushmore visits and consumer spending. Table 4.1 presents the base model results. Table 4.1 coefficient estimates are not scaled.

Table 4.1: Base Model

Dependent Variable:	Consumer Spending			
	Accommodation	Food Service	Retail	Arts/Ent/Rec
<i>Variables</i>				
Mount Rushmore visits	0.4521*** (0.0421)	0.5964*** (0.1362)	1.958** (0.6780)	0.1186*** (0.0310)
SD resident sample size	0.3107 (0.2259)	2.448* (1.306)	13.09** (5.819)	-0.0812 (0.0614)
Rest of US sample size	-0.0048*** (0.0014)	-0.0337*** (0.0076)	-0.1400*** (0.0319)	0.0015* (0.0008)
<i>Fixed-effects</i>				
Month	Yes	Yes	Yes	Yes
<i>Fit statistics</i>				
Observations	261	261	261	261
R ²	0.850	0.593	0.606	0.590
Adjusted R ²	0.841	0.570	0.584	0.567
F-test	20.8	5.34	5.65	5.28

Clustered (Month) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

The Mount Rushmore coefficients can be interpreted as follows: for consumers and stores present in the Spend panel, on average each additional visit to Mount Rushmore is associated with \$0.4624 spent in the Accommodation sector, \$0.6776 spent in the Food Service sector, \$2.390 spent in the Retail sector, and \$0.1159 spent in the Arts/Entertainment/Recreation sector. While these estimates are associated with Mount Rushmore visits, they are not directly attributable to Mount Rushmore. Some visitors to Mount Rushmore are on multi-purpose trips visiting other destinations in the local area as well. Additionally, since visits to other attractions are not included in this model, some spending that should be attributed to other attractions is captured by the Mount Rushmore coefficient. Because of this, the coefficient values on Mount Rushmore visits represent spending associated with the bundle of attractions in the area, including Mount Rushmore. Given that Mount Rushmore is the most prominent attraction (most prominent meaning it sees the highest visitation of all other state and national parks in the area), it is fair to assume that a majority of associated spending captured by the coefficients is connected to Mount Rushmore visitors either through a primary purpose or multi-purpose trip.

As mentioned in the Methods section, I include the number of cardholders represented in *Spend* from both South Dakota and the rest of the United States in the model as variables. The variables capture changes in spending attributable to fluctuations in cardholder representation, accounting for the effects of the changing sample size. I include the number of cardholders from South Dakota as a separate variable from the total number of cardholders since the number of South Dakota cardholders represented appears to have a larger impact on spending in the Mount Rushmore area compared to the number of cardholders represented residing in other states. Across sectors, the coefficients on the *Spend* cardholders for the rest of the United States variable are very small but significant. They provide little use in interpretation. They do not follow intuition, because logically I'd expect an increasing sample size to be correlated with higher spending. The coefficients on the *Spend* South Dakota cardholders variable are small and insignificant in the Accommodation and Arts/Entertainment/Recreation sectors, but are large and significant in the Food Service and Retail sectors. Given that the Food Service and Retail sectors tend to experience more resident spending,

it makes sense that they are more impacted by the number of South Dakota cardholders included in the *Spend* data.

Regarding model fit statistics, all sectors have significant F-stats, suggesting that visitor spending is an important source of sector spending. The Accommodation model has the highest significance with the largest F-test value and adjusted- R^2 value. This indicates that the Accommodations sector is most closely tied to tourism compared to the other sectors.

Table 4.2: Base Coefficients Scaled

Dependent Variable:	Consumer Spending			
	Accommodation	Food Service	Retail	Arts/Ent/Rec
Mount Rushmore visits	17.65 (14.04 - 21.27)	23.29 (11.58 - 34.99)	76.45 (18.18 - 134.71)	4.63 (1.97 - 7.29)

Scaled 95% confidence intervals in parentheses

Scaled, the model estimates, on average, each additional visit to Mount Rushmore is associated with \$17.65 spent in the Accommodation sector, \$23.29 spent in the Food Service sector, \$76.45 spent in the Retail sector, and \$4.63 spent in the Arts/Entertainment/Recreation sector. These results indicate that the average national park visitor to the Black Hills region spends the most in Retail, followed by Food Service, Accommodations, and Arts/Entertainment/Recreation. Retail sector spending covers a wide array of goods, including gas, groceries, souvenirs, and other retail purchases. This wide range is likely why the coefficient estimate is the largest. Additionally, the Black Hills region of South Dakota is popular for RV/road trip tourism. These types of visitors spend heavily on gas and groceries.

4.2 Binning Spending by Distance

In the base model, I estimate the average spending associated with Mount Rushmore visits across the entire 100-mile radius around the park. However, it is likely that visitor spending makes up a higher portion of spending in areas closer to the park and a lower portion further from the park. To assess this hypothesis, I segment consumer spending into 4 distance bins, with each bin covering a 25-mile distance from the park. I then create distance bin dummy variables and interact them with the Mount Rushmore visits variable and both *Spend* sample size variables.

When coefficients are allowed to change based on distance from the park, we see a clearer picture of where visitor spending occurs. For all sectors, the amount of spending correlated with Mount Rushmore visits is highest within 0-25 miles of the park. This is shown by the Mount Rushmore coefficient being the largest for bin 1 across all sectors. Consistent with expectations, coefficient estimates decrease for bin 2. As distance from the park grows, there is less spending correlated with visits. The Mount Rushmore coefficient estimates for bins 3 and 4 are very small, indicating that very small amounts of visitor spending occur 50-100 miles from Mount Rushmore. This is consistent with NPS SEM survey results, which found that the communities within 25 miles of the park saw the highest amount of spillover from Mount Rushmore visitors. 61% of park visitors stopped or stayed in Rapid City, 59% in Custer, and 56% in Keystone [Otak et al., 2023]. These municipalities are all located in bin 1.

For all sectors, the coefficients on Mount Rushmore visits are significant in each distance bin. This shows that the full 100-mile radius around the park is impacted by visitor spending. But, as noted in the previous paragraph, the areas further from the park see smaller amounts of visitor spending. For the Accommodations and Food Service sectors, the Mount Rushmore visits variable is significant at the $\alpha = 0.01$ level for all bins. In the Retail sector, the Mount Rushmore visits variable is significant at the $\alpha = 0.01$ level for distance bins 2, 3, and 4, however, bin 1 is less significant at the $\alpha = 0.05$ level. While the difference in significance from bin 1 to bins 2,3, and 4 is minimal, a possible explanation is that bin 1 contains the large urban area of Rapid City. It is

Table 4.3: Binned Model

Dependent Variable: Model:	Consumer Spending			
	Accommodation	Food Service	Retail	Arts/Ent/Rec
<i>Variables</i>				
Rushmore × Bin = 1	0.2978*** (0.0369)	0.3599*** (0.0893)	1.105** (0.4756)	0.0521*** (0.0189)
Rushmore × Bin = 2	0.1215*** (0.0094)	0.1785*** (0.0612)	0.5596*** (0.1814)	0.0377*** (0.0064)
Rushmore × Bin = 3	0.0198*** (0.0062)	0.0220*** (0.0026)	0.2139*** (0.0211)	0.0005** (0.0003)
Rushmore × Bin = 4	0.0129*** (0.0031)	0.0360*** (0.0038)	0.0794*** (0.0176)	0.0282*** (0.0075)
SD resid sample × Bin = 1	0.1785 (0.1441)	1.932* (1.009)	11.22** (4.682)	-0.1783** (0.0727)
SD resid sample × Bin = 2	0.0846 (0.0584)	0.4240** (0.2079)	1.273* (0.7223)	0.0921** (0.0374)
SD resid sample × Bin = 3	0.0413** (0.0195)	0.0188 (0.0171)	0.3393*** (0.1254)	-0.0001 (0.0010)
SD resid sample × Bin = 4	0.0063 (0.0063)	0.0732* (0.0374)	0.2573* (0.1361)	0.0051 (0.0100)
Rest of US sample × Bin = 1	-0.0045*** (0.0010)	-0.0289*** (0.0059)	-0.1233*** (0.0251)	0.0017*** (0.0005)
Rest of US sample × Bin = 2	-0.0003 (0.0005)	-0.0049*** (0.0012)	-0.0166*** (0.0044)	-0.0006* (0.0003)
Rest of US sample × Bin = 3	-0.0002 (0.0001)	-7.54×10^{-5} (0.0001)	0.0007 (0.0009)	-1.03×10^{-5} (1.06×10^{-5})
Rest of US sample × Bin = 4	8.72×10^{-5} (6.59×10^{-5})	0.0002 (0.0002)	-0.0008 (0.0011)	0.0004*** (0.0001)
<i>Fixed-effects</i>				
Month-Bin	Yes	Yes	Yes	Yes
<i>Fit statistics</i>				
Observations	1,044	1,044	1,044	1,044
R ²	0.896	0.873	0.897	0.604
Adjusted R ²	0.890	0.865	0.891	0.580
F-test	33.7	26.8	34.0	5.97

Clustered (Month-Bin) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

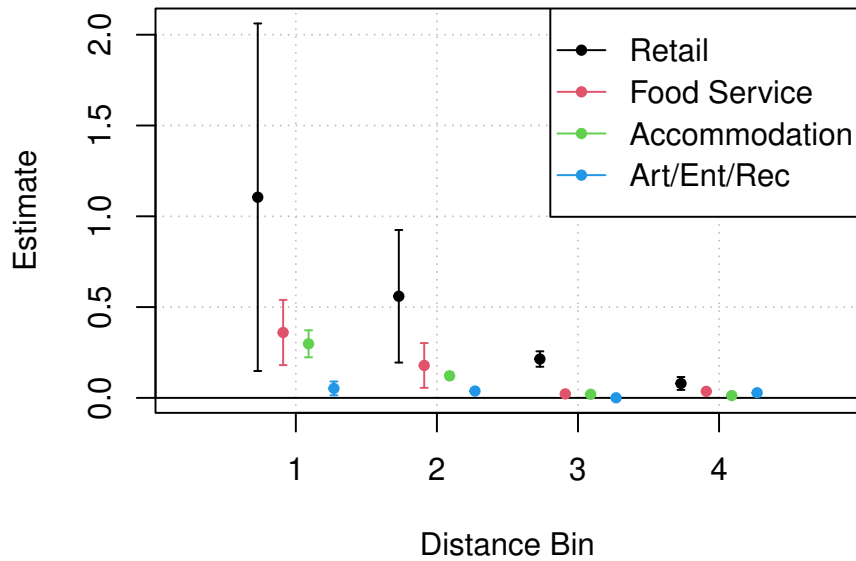


Figure 4.1: Mount Rushmore visits coefficients across bins

plausible that visitor spending makes up a smaller portion of Rapid City’s economy which results in the Mount Rushmore visits variable having less explanatory power.

Similar to the base model, the *Spend* cardholder coefficients are largely uninterpretable. I include them to account for observed changes in spending that are due to fluctuations in the underlying number of cardholders represented by *Spend*. I interact these variables with each distance bin to allow their impacts on spending to vary depending on the geographic group of POIs being analyzed. The amount of spending attributed to sample size changes is different for high-spending areas (0-25 miles around Mount Rushmore) than low-spending areas (50-100 miles around Mount Rushmore).

Scaled to include all visitor spending at all stores, the model estimates that for the Accommodation sector, on average, each additional visit to Mount Rushmore is associated with \$11.63 spent 0-25 miles from the park, \$4.75 spent 25-50 miles from the park, \$0.78 spent 50-75 miles from the park, and \$0.50 spent 75-100 miles from the park. For the Food Service sector, the model

Table 4.4: Binned Coefficients Scaled

Dependent Variable:	Consumer Spending			
	Accommodation	Food Service	Retail	Arts/Ent/Rec
Rushmore x Bin = 1 (0-25 mi)	11.63 (8.73 - 14.5)	14.05 (7.04 - 21.07)	43.14 (5.78 - 80.51)	2.04 (0.55 - 3.52)
Rushmore x Bin = 2 (25-50 mi)	4.75 (4.01 - 5.48)	6.97 (2.16 - 11.78)	21.85 (7.60 - 36.10)	1.47 (0.97 - 1.93)
Rushmore x Bin = 3 (50-75 mi)	0.78 (0.28 - 1.27)	0.86 (0.66 - 1.06)	8.35 (6.70 - 10.01)	0.02 (-0.03 - 0.04)
Rushmore x Bin = 4 (75-100 mi)	0.50 (0.26 - 0.75)	1.41 (1.11 - 1.70)	3.10 (1.71 - 4.49)	1.01 (0.39 - 1.75)

Scaled 95% confidence intervals in parentheses

estimates \$14.05 spent 0-25 miles from the park, \$6.97 spent 25-50 miles from the park, \$0.86 spent 50-75 miles from the park, and \$1.41 spent 75-100 miles from the park. For the Retail sector, estimates are \$43.14 spent 0-25 miles from the park, \$21.85 spent 25-50 miles from the park, \$8.35 spent 50-75 miles from the park, and \$3.10 spent 75-100 miles from the park. Finally, for the Arts/Entertainment/Recreation sector, the model estimates \$2.04 spent 0-25 miles from the park, \$1.47 spent 25-50 miles from the park, \$0.02 spent 50-75 miles from the park, and \$1.01 spent 75-100 miles from the park.

Compared to the base model, the fit statistics for the binned model are quite a bit more significant in the Accommodation, Food Service, and Retail sectors. Fit statistics are about the same for the Arts/Entertainment/Recreation sector. For 3 out of the 4 sectors, introducing spatial flexibility in coefficient estimation improves model fit and significance. The increased model performance demonstrates that Mount Rushmore visitor spending fluctuates by distance to the park.

4.3 Incorporating Other National Parks

When I include Badlands National Park visits and Wind Cave National Park visits in the base model, many of the resulting coefficient estimates are counter-intuitive. The Badlands coefficient estimates are significant and positive for the Accommodation sector, however, are insignificant and negative for the other sectors. The Wind cave coefficient estimates are insignificant for all sectors. The Mount Rushmore coefficient estimates are all significant, but, for the Accommodation sector, the coefficient is negative. Compared to the base model, Mount Rushmore coefficients change significantly. As described in the methods section, the model results cannot be considered accurate or legitimate due to the high correlation between Mount Rushmore visits, Badlands visits, and Wind Cave visits. The model results illustrate that Mount Rushmore some tourists are visiting multiple parks and destinations in the Black Hills region. The model results also demonstrate a large limitation to the methodology I use in this research to estimate park visitor spending. When visits to other nearby attractions are correlated with the park of interest, visitor spending correlated with each individual park cannot be identified.

Table 4.5: Base Model with Badlands and Wind Cave

Dependent Variable:	Consumer Spending			
Model:	Accommodation	Food Service	Retail	Arts/Ent/Rec
<i>Variables</i>				
Mount Rushmore visits	-0.1729* (0.0850)	0.8081** (0.2944)	3.751** (1.658)	0.1145** (0.0370)
Badlands visits	1.203** (0.3999)	-1.784 (1.329)	-8.585 (6.050)	-0.0738 (0.1152)
Wind Cave visits	0.7376 (0.5463)	1.135 (1.377)	4.099 (5.623)	-0.0383 (0.2267)
SD resident sample size	0.1044 (0.1402)	1.303 (1.098)	7.921 (4.905)	-0.1012 (0.0641)
Rest of US sample size	-0.0072*** (0.0018)	-0.0536*** (0.0088)	-0.2274*** (0.0385)	0.0012 (0.0007)
<i>Fixed-effects</i>				
Month	Yes	Yes	Yes	Yes
<i>Fit statistics</i>				
Observations	175	175	175	175
R ²	0.906	0.682	0.704	0.737
Adjusted R ²	0.897	0.650	0.674	0.711
F-test	21.27	4.712	5.23	6.17

Clustered (Month) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

4.4 Using Year-by-month Fixed Effect

In the base model, I account for the impacts on observed spending due to the fluctuating number of cardholders represented in *Spend* by including the South Dakota resident sample size and the sample size for the rest of the United States as variables in the regression. However, this may be an imperfect method of accounting for the sample size impacts. To assess its robustness, I run a model using a year-month combined fixed effect, and remove the *Spend* sample size variables. The year-month combined fixed effect allows each month each year to have a different intercept. Since the sample size of *Spend* cardholders varies by month, but the consumer spending and Mount Rushmore visit variables vary by week, the year-month fixed effect accounts for the baseline change to monthly consumer spending driven by the change in *Spend* sample size.

Table 4.6: Year-by-Month Fixed Effects Model

Dependent Variable: Model:	Consumer Spending			
	Accommodation	Food Service	Retail	Arts/Ent/Rec
<i>Variables</i>				
Mount Rushmore visits	0.4301*** (0.0545)	0.4028** (0.1739)	1.625** (0.6172)	0.0670*** (0.0134)
<i>Fixed-effects</i>				
Month-Year	Yes	Yes	Yes	Yes
<i>Fit statistics</i>				
Observations	261	261	261	261
R ²	0.952	0.808	0.836	0.861
Adjusted R ²	0.938	0.750	0.787	0.819
F-test	1,175.6	248.2	301.4	365.2

Clustered (Month-Year) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

The model results vary somewhat compared to the base model results. The Mount Rushmore coefficients are lower for every sector. The Mount Rushmore coefficient estimate is 4.5% lower in the Accommodation sector, 32.5% lower in the Food Service sector, 17% lower in the Retail sector,

and 43.5% lower in the Arts/Entertainment/Recreation sector. Accounting for the *Spend* sample fluctuations by including them as variables, as I did in the base model, results in higher estimates. When instead accounting for the *Spend* sample fluctuations using a month-year combined fixed effect, as I did in this model, estimates are lower. However, despite the difference in coefficient sizes, the significance of the Mount Rushmore coefficients are the same across models for every sector except Food Service. In the base model, the Food Service coefficient is significant at the $\alpha = 0.01$ level. In this model, the Food Service coefficient is slightly less significant at the $\alpha = 0.05$ level. This shows that both models find similar strengths in the relationship between Mount Rushmore visits and spending. Where they differ is in the size of the relationship.

In the base model, the coefficient estimates on the *Spend* sample size for the rest of the US are negative for all sectors except the Arts/Entertainment/Recreation sector. Because they are negative, they push the Mount Rushmore coefficients higher, since the Mount Rushmore coefficients compensate for the negative impact on model fitted values. The Mount Rushmore coefficients are not pushed higher when the *Spend* sample size for the rest of the US variable is not included in the model. This is most likely the reason for the difference in coefficient size between models.

Table 4.7: Year-by-month Fixed Effects Model Coefficients Scaled

Dependent Variable:	Consumer Spending			
	Accommodation	Food Service	Retail	Arts/Ent/Rec
Mount Rushmore visits	16.79 (12.53 - 21.05)	15.73 (2.14 - 29.31)	63.44 (15.22 - 111.66)	2.62 (1.57 - 3.66)

Scaled 95% confidence intervals in parentheses

Across sectors, the F-test values for the year-by-month fixed effect model are much more significant compared to the base model. The adjusted- R^2 values are larger as well. This indicates that the year-by-month fixed-effect model provides a better fit to the data than the base model. Com-

pared to the binned distance model, the year-by-month fixed effect model has higher adjusted- R^2 values for the Accommodation and Arts/Entertainment/Recreation sectors and lower adjusted- R^2 values for the Food Service and Retail sectors. The F-test values are larger across sectors for the year-by-month fixed effect model. The similarity in adjusted- R^2 values indicates that the binned distance model and the year-by-month fixed effect model have similar explanatory power, but the superiority of the F-test values in the year-by-month fixed effect model indicates it has higher statistical significance. Overall, using the model fit statistics, I find that the year-by-month fixed effect model performs best when compared to the base model and the binned distance model.

4.5 Comparison to NPS Visitor Spending Estimates

I use Mount Rushmore visits coefficient estimates from the year-by-month fixed effect model for my comparison to the NPS visitor spending estimates. The year-by-month fixed effect model has more explanatory power than the base model, higher significance than the binned distance model, and does not suffer from the multicollinearity issues seen in the multiple national parks model. I use the coefficient estimates from the year-by-month fixed effect model to calculate an estimate of total spending associated with Mount Rushmore visitors for 2022. I then compare these estimates with the survey-based NPS estimates of 2022 Mount Rushmore visitor spending⁶. If both estimates are accurate, then the model estimates should be somewhat higher than NPS estimates. This is because the NPS visitor spending estimates solely reflect spending attributable to Mount Rushmore, while the model estimates capture some visitor spending related to other attractions. That being said, when comparing estimates it should be noted that neither are the object truth. Both are estimates of the true value of Mount Rushmore visitor spending in 2022.

The NPS visitor spending sectors are slightly different than the sectors I use. To make estimates comparable, I combine NPS estimates for Hotels and Camping to match my Accommodation estimate. I also combine the NPS estimates for Retail, Gas, and Grocery to match the Retail estimate.

⁶NPS estimates for Mount Rushmore visitor spending, as well as for every NPS unit, are available at <https://www.nps.gov/subjects/socialscience/vse.htm>.

As mentioned in the Mount Rushmore Background section, there is one spending category included in the NPS estimate that is not included in the model estimate, local transportation. This is an unimpactful omission as the NPS estimates local transportation only makes up 0.4% of visitor spending at Mount Rushmore [Flyr and Koontz, 2023]. Table 4.8 shows the comparison between the *Spend*-based model and the NPS model, with model estimates in the first column and NPS estimates in the second column. For the *Spend*-based model, 95% confidence intervals are included in parenthesis underneath the estimates. The third column indicates whether or not the NPS estimate is included in the 95% confidence interval. It shows whether or not the *Spend* model estimate is statically different than the NPS estimate.

Table 4.8: NPS estimates and model estimates for 2022 Mount Rushmore visitor spending

	<i>Spend</i> Model (millions)	NPS Model (millions)	Captured?
Accommodation	\$41.0 (\$30.6 - \$51.4)	\$191.8	No
Food Service	\$38.4 (\$5.2 - \$71.5)	\$58.6	Yes
Retail	\$154.8 (\$37.2 - \$272.5)	\$100.4	Yes
Arts/Ent/Rec ⁷	\$6.4 (\$3.8 - \$8.9)	\$33.1	No
Transportation	-	\$1.5	-
Total	\$240.6 (\$76.8 -\$404.4)	\$385.4	Yes

95% CI shown in parenthesis

⁷Estimates in this sector are not directly comparable. The NPS estimate includes spending on equipment rentals and tour operators but the *Spend* model estimate does not.

In the Accommodation sector, the *Spend* model estimates \$41.0 million of visitor spending, while the NPS estimates \$191.8 million. The NPS estimate is not captured within the estimate's confidence interval. This indicates that the model underestimates accommodation spending. This is mainly for two reasons. First, online transactions are not included in the daily POI spending totals reported in *Spend*⁸. Second, online businesses like Airbnb and VRBO are not covered by *Spend*. This means that only in-person accommodation transactions are included in the data. Spending on hotel rooms paid for in advance and spending on vacation rental websites are omitted from the analysis. This results in a significant chunk of accommodation spending missing from the visitor spending estimation. Additionally, while *Spend* does cover campgrounds, it does not cover campsites that only take cash or check payments, like those operated by the US Forest Service.

In the Food Service sector, the model estimates \$38.4 million of visitor spending, while the NPS estimates \$58.6 million. The model estimate is less than the NPS estimate, however the NPS estimate is captured within the 95% confidence interval. The similarity suggests that the *Spend* model captures visitor spending in the Food Service sector appropriately. Food Service spending is well represented in the data since most food service purchases are made in-person rather than online.

In the Retail sector, the model estimates \$154.8 million of visitor spending, while the NPS estimates \$100.4 million. The model estimate is larger than the NPS estimate, however the NPS estimate is captured within the confidence interval. Similar to the Food Service sector, the similarity in estimates suggests the *Spend* model captures visitor spending in the Retail sector appropriately. While some retail purchases are made online, the majority are made in person. This is particularly true for the retail areas most relevant to visitor spending like groceries, gas, and souvenirs.

In the Arts/Entertainment/Recreation sector, the model estimates \$6.4 million of visitor spending, while the NPS estimates \$33.1 million. The NPS estimate is not captured within the 95% confidence interval. Importantly, these estimates are not directly comparable. The difference in estimates may be due to the different definition of the Arts/Entertainment/Recreation sector used

⁸*Spend* does provide POI online spending data but only at the monthly resolution.

by the NPS. The NPS estimate includes visitor spending on equipment rentals and tour operators. The NAICS definition of the Arts/Entertainment/Recreation sector I use in the *Spend* model does not include this spending. As such, less spending is included in the *Spend* estimate.

Totaling visitor spending across categories, the NPS estimates 2022 Mount Rushmore visitor spending to be \$385.4 million, while the *Spend* model estimates \$240.6 million. The NPS estimate is included in the 95% confidence interval, meaning that, on aggregate, the two estimations are not statically different from each other. However, when looking at estimated visitor spending by category, there are clear differences in estimates for the Accommodation sector and the Arts/Entertainment/Recreation sector.

The comparison between the NPS estimates and the *Spend* model estimates highlights several limitations of the model and the underlying data. With regard to the model, the presence of other tourist attractions in the area prevents accurate identification of Mount Rushmore visitor spending. Additionally, multi-purpose trips are not accounted for. Both these limitations mean model estimates are inflated, and more likely reflect visitor spending by general visitors to the area. With regard to the *Spend* data used, it is not representative of visitor spending in the Accommodation sector, as visitor spending on accommodation bookings made through vacation rental sites like Airbnb are not included in the data. Additionally, when using *Spend* data, impacts on observed spending from the fluctuating number of monthly cardholders represented in the data must be accounted for in the analysis.

Despite these limitations, the model does have several strengths. Using only observed spending data, it generates Mount Rushmore visitor spending estimates that are not subject to forms of survey bias like recall bias and telescoping bias. By using credit and debit card transaction data as the source of visitor expenditure information, the spending analyzed is guaranteed to have occurred within the identified time frame and within the identified geographic area. The model estimates reflect observed spending rather than remembered spending. Additionally, model estimates are up-to-date and reflective of visitor spending habits for the entire year. Survey-based methods only gather visitor spending data from a snapshot in time, and may not be representative

of all yearly visitors. Also, the model provides precision measurements like standard errors and confidence intervals. This gives the model estimates appropriate context useful to policymakers, park managers, and researchers. No measurement of estimate precision is provided through the survey-based methods. Lastly, the model does not require a visitor survey.

Chapter 5

Conclusion

In this research, I explore the feasibility and effectiveness of using credit and debit card transaction data to econometrically estimate national park visitor spending, using Mount Rushmore National Memorial as a case study. I build upon the vast amount of literature concerned with accurate visitor spending estimation, and introduce the use of granular transaction data as a source of visitor spending information. Using Safegraph data and Mount Rushmore visitation statistics, I develop multiple models using fixed effect regressions to estimate average spending per visit in the Accommodation, Food Service, Retail, and Arts/Entertainment/Recreation sectors. For the best performing model, in 2022, I estimate Mount Rushmore visitor spending to be \$41.0 million in the Accommodation sector, \$38.4 million in the Food Service sector, \$154.8 million in the Retail sector, and \$6.4 million in the Arts/Entertainment/Recreation sector. I then compare these estimates with NPS survey-based visitor spending estimates. The model estimates for the Food Service and Retail sectors are statistically indifferent to NPS estimates. However, model estimates for the Accommodation and Arts/Entertainment/Recreation sector are below NPS estimates.

The model has several strengths. Using only observed spending data, it generates Mount Rushmore visitor spending estimates that are not subject to forms of survey bias like recall bias and telescoping bias. By using credit and debit card transaction data as the source of visitor expenditure information, the spending analyzed is guaranteed to have occurred within the identified time frame and within the identified geographic area. The model reflects observed spending rather than remembered spending. Additionally, model estimates are up-to-date and reflective of visitor spending habits for the entire year. Survey-based methods only gather visitor spending data from a snapshot in time, and may not be representative of all yearly visitors. Also, the model provides precision measurements like standard errors and confidence intervals. This gives estimates appropriate context useful to policymakers, park managers, and researchers. Lastly, the model does not require a visitor survey.

The model also has several limitations. The presence of other tourist attractions in the area prevents accurate identification of Mount Rushmore visitor spending. Additionally, multi-purpose trips are not accounted for. Both these limitations mean model estimates are inflated, and more likely reflect visitor spending by general visitors to the area. With regard to the *Spend* data used, it is not representative of visitor spending in the Accommodation sector, as visitor spending on accommodation bookings made through vacation rental sites like Airbnb are not included in the data. Additionally, when using *Spend* data, impacts on observed spending from the fluctuating number of monthly cardholders represented in the data must be accounted for in the analysis.

Overall, visitor spending estimation using granular credit and debit card transaction data may present a possible alternative to traditional survey-based methods. This approach may be particularly beneficial for parks lacking the resources to conduct comprehensive visitor surveys. Looking forward, future research should focus on methodological refinement and the integration of other data sources to improve accuracy and applicability.

Bibliography

Dani T. Dagan and Emily J. Wilkins. What is “big data” and how should we use it? The role of large datasets, secondary data, and associated analysis techniques in outdoor recreation research. *Journal of Outdoor Recreation and Tourism*, page 100668, July 2023. ISSN 2213-0780. doi: 10.1016/j.jort.2023.100668. URL <https://www.sciencedirect.com/science/article/pii/S2213078023000658>.

Matthew Flynr and Lynne Koontz. 2022 national park visitor spending effects: Economic contributions to local communities, states, and the nation. Technical report, National Park Service, August 2023. URL <https://irma.nps.gov/DataStore/Reference/Profile/2299764>.

Jeff Ho. Is your spend data geographically representative?: Quantifying sampling bias in safegraph spend data, December 2021. URL <https://colab.research.google.com/drive/16eBT1dqnUK76GnBMgPMY5rJO4vBD2gdQ?usp=sharing>.

Dennis R. Howard, Samuel V. Lankford, and Mark E. Havitz. A measure for authenticating pleasure travel expenditures. *Journal of Travel Research*, 29(Spring):19–23, 1991.

Wassily W. Leontief. Quantitative Input and Output Relations in the Economic Systems of the United States. *The Review of Economics and Statistics*, 18(3):105–125, 1936. ISSN 0034-6535. doi: 10.2307/1927837. URL <https://www.jstor.org/stable/1927837>. Publisher: The MIT Press.

Yun Liang, Junjun Yin, Bing Pan, Michael Lin, and Guangqing Chi. Assessing the validity of SafeGraph data for visitor monitoring in Yellowstone National Park. *Travel and Tourism Research Association: Advancing Tourism Research Globally*, 2021.

Anyu Liu, Yoo Ri Kim, and Haiyan Song. Toward an accurate assessment of tourism economic impact: A systematic literature review. *Annals of Tourism Research Empirical Insights*, 3(2): 100054, November 2022. ISSN 2666-9579. doi: 10.1016/j.annale.2022.100054. URL <https://www.sciencedirect.com/science/article/pii/S2666957922000222>.

- James Mak. *Tourism and the Economy: Understanding the Economics of Tourism*. University of Hawai'i Press, 2004. ISBN 978-0-8248-2789-2. URL <https://www.jstor.org/stable/j.ctvvn7rt>.
- NPS. Visitation Numbers (U.S. National Park Service), February 2023. URL <https://www.nps.gov/aboutus/visitation-numbers.htm>.
- Inc. Otak, RRC Associates, and University of Montana. 2022 socioeconomic research of mount rushmore national memorial: Report on 2022 data collection. Technical Report NPS/MORU/NRR-2023/2546, National Park Service, U.S. Department of the Interior, Fort Collins, Colorado, 2023. URL <https://doi.org/10.36967/2299623>.
- Melville Saayman and Andrea Saayman. Estimating the Economic Contribution of Visitor Spending in the Kruger National Park to the Regional Economy. *Journal of Sustainable Tourism*, 14(1):67–81, January 2006. ISSN 0966-9582. doi: 10.1080/09669580608668592. URL <https://doi.org/10.1080/09669580608668592>. Publisher: Routledge
_eprint: <https://doi.org/10.1080/09669580608668592>.
- SafeGraph. Announcing safegraph spend, 2023a. URL <https://www.safegraph.com/events/announcing-safegraph-spend>. Accessed: 2024-06-26.
- SafeGraph. Spend | SafeGraph Docs, 2023b. URL <https://docs.safegraph.com/docs/spend>.
- SafeGraph. January 2024 release notes, 2024. URL <https://docs.safegraph.com/changelog/january-2024-release-notes>. Accessed: 2024-05-16.
- Michael Sinclair, Saeed Maadi, Qunshan Zhao, Jinhyun Hong, Andrea Ghermandi, and Nick Bailey. Assessing the socio-demographic representativeness of mobile phone application data. *Applied Geography*, 158:102997, September 2023. ISSN 0143-6228. doi: 10.1016/j.apgeog.2023.102997. URL <https://www.sciencedirect.com/science/article/pii/S0143622823001285>.
- Daniel J. Stynes and Eric M. White. Reflections on Measuring Recreation and Travel Spending. *Journal of Travel Research*, 45(1):8–16, August 2006. ISSN 0047-2875, 1552-6763. doi: 10.1177/0047287506288873. URL <http://journals.sagepub.com/doi/10.1177/0047287506288873>.

Y Sun. Survey biases from three national park service visitor survey projects. In *Proceedings of the Travel and Tourism Research Association 36th Annual Conference*, 2005.

Catherine Cullinane Thomas, Egan Cornachione, Lynne Koontz, and Christopher R. Keyes. National Park Service socioeconomic monitoring pilot survey: Visitor spending analysis. Technical Report 2019/1924, National Park Service, 2019. URL <https://pubs.usgs.gov/publication/70204070>.

TourismEconomics. Economic impact of tourism in south dakota – county results 2022. Technical report, South Dakota Office of Tourism, April 2023. Prepared for: South Dakota Office of Tourism.

U.S. Bureau of Economic Analysis. Population, January 1959. URL <https://fred.stlouisfed.org/series/POPTHM>. Publisher: FRED, Federal Reserve Bank of St. Louis.

U.S. National Park Service. VUStats FAQ - Social Science (U.S. National Park Service). URL <https://www.nps.gov/subjects/socialscience/statistics-faq.htm>.

U.S. National Park Service. *Mount Rushmore National Memorial: Public Use Counting and Reporting Instructions*, January 1995. Effective date: January 1, 1995.

U.S. National Park Service. Visitor spending effects - economic contributions of national park visitor spending, 2023. URL <https://www.nps.gov/subjects/socialscience/vse.htm>. Accessed: 2024-05-16.

U.S. National Park Service. Mount rushmore national memorial, 2024. URL <https://www.nps.gov/moru/index.htm>. Accessed: 2024-05-16.

Philip Watson, Dawn Thilmany, Josh Wilson, and Susam Winter. Determining Economic Contributions and Impacts: What is the difference and why do we care? *Journal of Regional Analysis and Policy*, 37:140–146, January 2007.

James J. Wilton and Norma Polovitz Nickerson. Collecting and Using Visitor Spending Data. *Journal of Travel Research*, 45(1):17–25, August 2006. ISSN 0047-2875. doi: 10.1177/0047287506288875. URL <https://doi.org/10.1177/0047287506288875>. Publisher: SAGE Publications Inc.

Deying Zhou, John F. Yanagida, Ujjayant Chakravorty, and PingSun Leung. Estimating economic impacts from tourism. *Annals of Tourism Research*, 24(1):76–89, January 1997. ISSN 0160-7383. doi: 10.1016/S0160-7383(96)00035-7. URL <https://www.sciencedirect.com/science/article/pii/S0160738396000357>.