NATIONAL INSTITUTE FOR CONGESTION REDUCTION

FINAL REPORT
FEBRUARY 2024

Use of Crowdsourced Data for Access Management Evaluation

Debbie Albert, Kartikeya Jha, Mahin Ramezani,
Max Steadman, Brianne Glover, Luke Albert

National Institute for Congestion Reduction
University of South Florida
Center for Urban Transportation Research | University of South Florida

NICR
NATIONAL INSTITUTE FOR CONGESTION REDUCTION

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

Use of Crowdsourced Data for Access Management Evaluation

Prepared by

Debbie Albert, Kartikeya Jha, Mahin Ramenzani, Max Steadman, Brianne Glover, Luke Albert

Texas A&M Transportation Institute

Prepared for

National Institute for Congestion Reduction

University of South Florida

Center for Urban Transportation Research

4202 E. Fowler Avenue, ENG030, Tampa, FL 33620-5375

nicr@usf.edu

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle	5. Report Date
Use of Crowdsourced Data for Access Management Evaluation	February 2024
	6. Performing Organization Code
7. Author(s) Debbie Albert, Kartikeya Jha, Mahin Ramenzani, Max Steadman, Brianne Glover, Luke Albert	8. Performing Organization Report No.
9. Performing Organization Name and Address Texas A&M Transportation Institute 3135 TAMU College Station, TX 77843	10. Work Unit No. (TRAIS)
	11. Contract or Grant No. 69A3551947136, 79075-00 SUB B
12. Sponsoring Organization Name and Address U.S. Department of Transportation University Transportation Centers 1200 New Jersey Avenue, SE Washington, DC 20590 United States National Institute for Congestion Reduction 4202 E. Fowler Avenue Tampa, FL 33620-5375 United States	13. Type of Report and Period Covered Final Report, May 1, 2022 – August, 2023
	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

The impacts of access management projects have been evaluated, documented, and debated by practitioners and adjacent property owners for decades. This research aims to evaluate the potential of using new-age data sources such as crowdsourced connected vehicle (CV) data, crowdsourced consumer spending pattern data, and traditional sources for sales tax and vehicle crash data in quantifying the effects of access management projects on arterial roadways. Based on observations and analyses highlighted in this study, it can be concluded that crowdsourced data provide valuable information on travel patterns along access managed corridors. The same information, if sought through manual data collection, can be expensive, time consuming, potentially biased, and relatively unscalable. An evaluation of crowdsourced data, such as Wejo, revealed that different elements of these datasets can be useful for a different set of audiences. Current limitations of new-age crowdsourced data sources as well as traditional data sources used for evaluating access management projects are also noted. The insights into traffic patterns and driving behavior gained using CV data can be used to inform transportation planning and policy-making efforts, as well as to improve road safety and reduce traffic congestion. Additionally, this project demonstrates the power of advanced analytical techniques, such as geospatial analysis, in extracting insights from large datasets and applying that knowledge for engineering, planning, public outreach, and business community engagement activities.

17. Key Words	18. Distribution Statement
Probe data, access management, connected vehicle data, safety, vehicle trajectory	

19. Security Classification (of this report)	20. Security Classification (of this page)	21. No. of Pages	22. Price
Unclassified.	Unclassified.	42	

Table of Contents

Figures

Tables

Executive Summary

Implementing access management projects and rationalizing their viability can be challenging because of seemingly conflicting demands of engineers, planners, business owners, the public, and elected representatives. The impacts of access management projects have been evaluated, documented, and debated by practitioners and adjacent property owners for decades. There have been efforts to evaluate the efficacy of access management projects, which have included the use of manual traffic data collection exercises, self-reported surveys, simulation, and evaluation of crash data many years after the project has been implemented. To date, comprehensive evaluations of the effects of these projects have shown positive safety and mobility benefits without causing negative economic outcomes. However, these studies are labor intensive and therefore are not completed as part of a regular assessment of access management programs, and they are limited to specific areas. As such, traditional data sources typically used for such evaluation are either impractically laborious, have a short span of data coverage, may suffer from stated preference biases, or are feasible only after a significant time has passed since the project was implemented. This also leads to deliberations on the applicability of the "national" results to local projects. State, county, and local traffic engineers need tools to help them convey the effects of access management projects based on local, timely data. Passive crowdsourced data has the potential to readily provide this information.

The objective of this research was to evaluate the potential of using new-age data sources such as crowdsourced connected vehicle (CV) data, crowdsourced consumer spending pattern data, and traditional sources for sales tax and crash data in quantifying the effects of access management projects on arterial roadways. Based on observations and analyses highlighted in this study, it can be concluded that crowdsourced data provide valuable information on travel patterns along access managed corridors. The same information, if sought through manual data collection, will be expensive, time consuming, and relatively unscalable. As discussed in detail using the example analysis of collected sales tax data, the available public information doesn't seem to be granular enough to derive conclusive insights on the impacts of access management. A few key observations regarding the crowdsourced data sources used in this study include the following:

- An evaluation of crowdsourced data, such as Wejo, revealed that different elements of this dataset can be useful for different sets of audience. For example, detailed information on vehicle journeys can be useful for businesses, property owners, and elected officials, while elements such as vehicle trajectories (paths), turning rates, and events like hard braking and acceleration can be useful information for engineers and planners.
- Using data from the same or a single vehicle manufacturer (OEM) for both the before and after periods can be comparatively more useful in providing insights on events like hard braking and evaluating the accompanying safety impacts.
- With the current form of the connected vehicle data, it is infeasible to associate movements with events. For example, it is difficult to tell if a certain movement type has more or fewer hard braking events than another movement type.
- Similarly, consumer spending data, such as the Spend data from SafeGraph, can be useful in comparing the same business or same type of business over time, but there is no consistent trend that is applicable across all kinds of businesses on a corridor. There can be several other factors at play—many, if not all of which are dynamic in nature—that can be impractical to capture with only a single or a few variables of analysis.

Limitations of traditional data sources used for evaluating access management projects are also noted. Most notably, using aggregated sales tax data from a state comptroller's office can be challenging to derive insights from mainly attributable to noted factors, some of which can be remedied, while others not to the same extent. Externalities can creep in because the before and after time windows might have a different composition of businesses in the reported dataset, inconsistent or different reporting frequencies of different businesses, and local economic conditions or peculiarities that might have a much larger impact on business activity than just the changes in access management around a business, among others.

Overall, this project provides insights into traffic patterns and driving behavior using CV data. The data can be used to inform transportation planning and policy-making efforts, as well as to improve road safety and reduce traffic congestion. Additionally, this project demonstrates the power of advanced analytical techniques, such as geospatial analysis, in extracting insights from large datasets and applying that knowledge for engineering, planning, public outreach, and business community engagement activities. A few ideas for potential future evaluation using these datasets are also discussed to enhance the scope and applicability of this research.

Chapter 1. Introduction

Implementation of access management projects has historically caused conflict between practitioners who support the safety and mobility benefits and business/property owners who are concerned about the potential negative economic impacts of reduced access to adjacent properties. Elected officials who are responsible for supporting or approving project funding find themselves in a delicate balancing act in the debate between safety, economy, and mobility. Past evaluations measuring the safety, mobility, and economic effects of access management projects have been limited to manual traffic data collection, self-reported surveys, simulation, and evaluation of crash data many years after the project has been implemented. In addition, evaluation of sales tax receipts before and after implementation of the project has been limited.

The purpose of this project was to test the use of innovative data sources (Wejo connected vehicle data and SafeGraph Spend data) combined with traditional data sources (sales tax data and publicly available crash data) to evaluate the safety, mobility, and economic effects of access management projects by comparing trips into and out of adjacent development before and after project implementation. Wejo connected vehicle data provides a rich data source with information about vehicle movements and vehicle events (e.g., hard braking, acceleration, etc.). The goal in evaluating the data was to determine if the percentages of vehicles turning in and out of adjacent businesses changed after a median was constructed, and whether rates of occurrence of events such as hard braking and hard acceleration changed. As a complementary outcome, this can also be reflective of the potential of such data sources in evaluating the safety, mobility, and economic effects of access management implementations. SafeGraph Spend data were used in an economic analysis to assess relationships between local sales tax and Spend data to determine the effect of access management on local businesses.

Chapter 2. Background

As defined by the Federal Highway Administration (FHWA), *access management* is "the proactive management of vehicular access points to land parcels adjacent to all types of roadways. Good access management promotes safe and efficient use of the transportation network" [1]. Examples of access management techniques include:

- Access spacing
- Driveway spacing
- Safe turning lanes
- Median treatments
- Right-of-way management

The conceptual functional hierarchy of corridor types and access is shown in Figure 1.

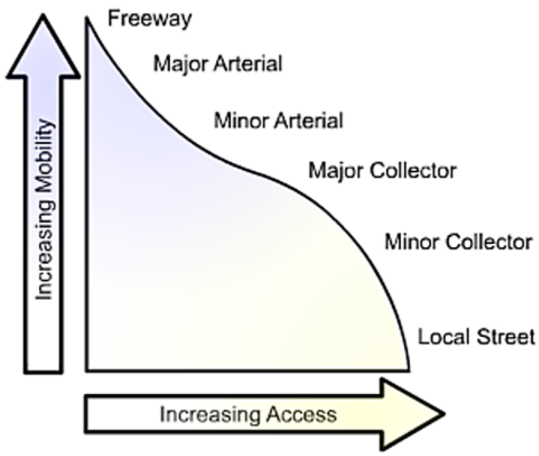
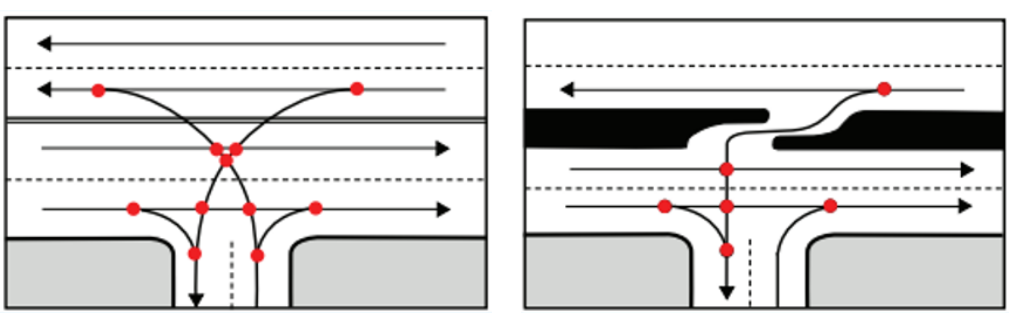


Figure 1. Conceptual roadway functional hierarchy [1].

A significant amount of research has been conducted on the safety and mobility benefits of access management. Enhanced safety is a high priority for everyone. Research points to reducing the number of conflict points as an effective way to improve safety. Various access management techniques help achieve that reduction. In the median example shown in Figure 2, there are 11 conflict points in the graphic on the left (full access - direct left turns out) versus 6 on the right (directional access).



Unlimited Access *Directional Access*

Figure 2. Reduction in number of potential conflict points because of access management.

Restricting turning movements can have a tremendous effect on safety. Converting a two-way left turn lane to a raised median reduces yearly crashes by 27% while converting a two-way left turn lane to a raised median reduces yearly pedestrian crashes by 46% [2]. Converting a two-way left turn lane to a raised median resulted in a 31.5% reduction in pedestrian crashes at unmarked crossings [3].

- Two out of three intersection related crashes involve left turns
- One out of three severe injuries involve left turns
- Left turns out are the most serious problem causing the most serious injuries
- Left turns in are the second most serious issue

Figure 3 illustrates the effects of increased access on safety. As the number of access points increases, there is an exponential increase in the number of crashes, thus reducing roadway safety.

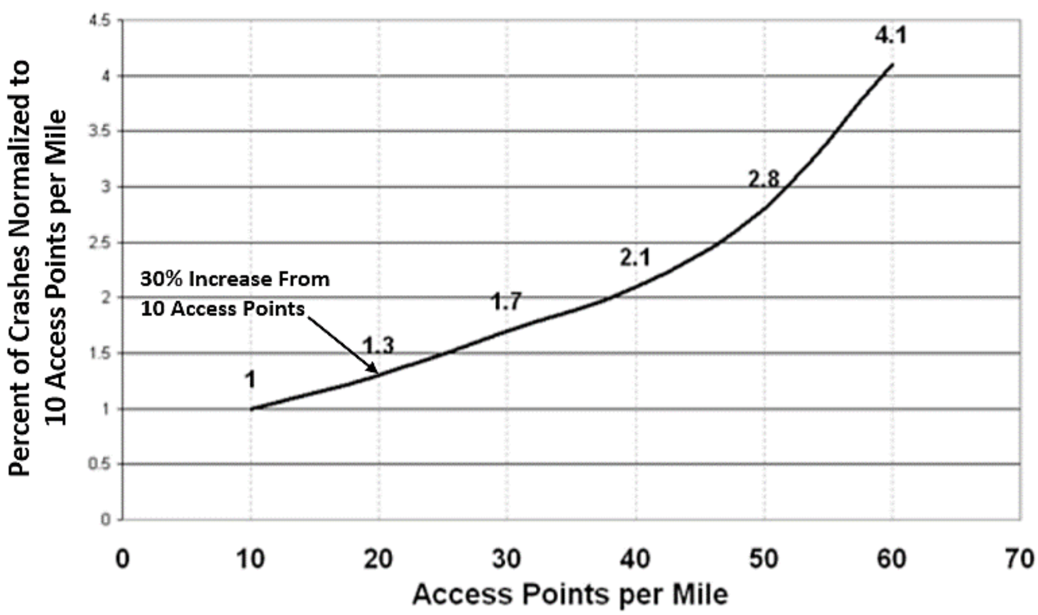


Figure 3. Composite crash rate indices [4].

In addition, Tables 1 and 2 demonstrate the relationship between mobility and access. In each case, mobility decreases as access points increase. Keeping people moving is another goal of access management techniques. For the customers, less congestion and more reliable travel times mean people reserve their time for what they want to do—family, shop, eat out, fun—in addition to travel.

Table 1. Travel Time and Signal Density [4]

Percentage Increase in Travel Time as Signal Density Increases	
Signals per Mile	**Percent Increase in Travel Time (Compared with 2 Signals per Mile)**
2	0
3	9
4	16
5	23
6	29
7	34
8	39

Table 2. Access Points and Free Flow Speed [5]

Access Points and Free-Flow Speed	
Access Points per Mile	**Reduction in Free Flow Speed, mph**
0	0.0
10	2.5
20	5.0
30	7.5
40 or more	10.0

Finally, efficiencies gained via less slowdowns and more reliable travel times to destinations mean consumers have more time to shop and they may even be more likely to enjoy getting to the store. A land use and transportation study found that a 30% reduction in travel speed can lead to a loss of nearly 50% of the market area for a business [6]. Therefore, if average speeds are 45 mph and they drop to just below 30 mph, businesses could expect to draw customers from an area 45% the original size (shown in Figure 4). More efficient access and fewer slowdowns mean more predictable travel times for customers and fewer shipment/delivery disruptions. Traffic backups play a huge role in shipping costs and reliability—which is critical considering the current time economy. Previous studies in Texas and other states have looked at measuring business impacts through surveys, and more recently by examining sales tax and property value information [7]. Through this project, researchers aimed to explore other crowdsourced sales data sources to assess opportunities for using them to conduct access management project evaluations.

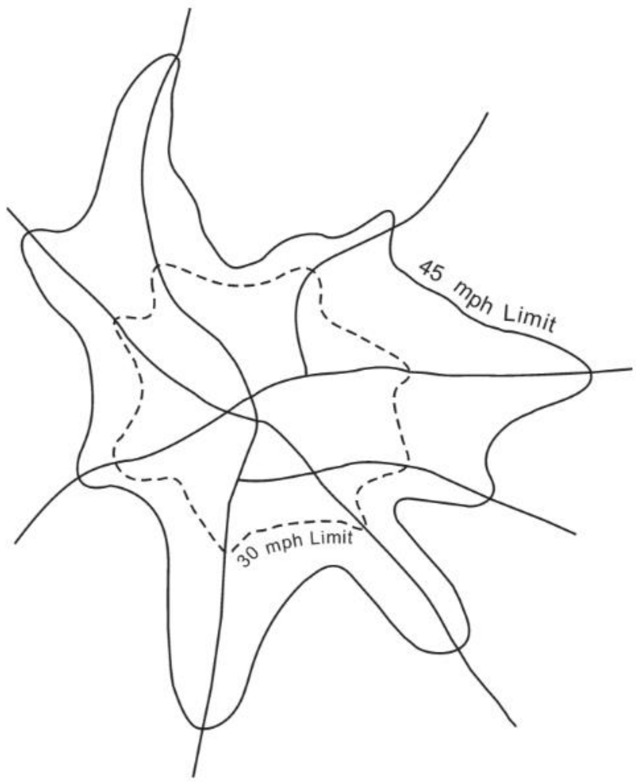


Figure 4. Reduction in market area with reduced travel speeds [6].

While comprehensive evaluations support the need for access management to improve safety and mobility without causing negative economic impacts, these studies require years of data for “after project” analyses, are laborious, and are not completed as part of a regular comprehensive assessment of access management programs. Therefore, they are limited to specific areas in which the study was completed, leading to deliberations on the applicability of “national” results to local projects. State, county, and local traffic engineers need tools to help them convey the effects of access management projects based on local, timely data. Passive crowdsourced data has the potential to supplement a closer analysis by providing this information.

Chapter 3. Data Sources

Researchers completed different analyses to explore the data sources as part of this project mostly using data from August through November of 2019 as the before condition and August through November of 2021 as the after condition. These data consist of the following sources:

- Wejo journeys
- Wejo events including key on, key off, hard accelerations, and hard braking
- Wejo paths and turning movements
- Crash data – TxDOT CRIS data
- Sales tax data – State Comptroller data for analysis window
- SafeGraph Spend data

Four of the six (italicized in the list above) involve new-age crowdsourced data types, while the other two involve traditional data sources. Wejo data (covering the first three bullets) are described in detail in the next section. The Spend data from SafeGraph look at debit and credit card transactions aggregated monthly to individual places in the United States going back to January 2019. When looking at the crash data, researchers calculated crash rates for each corridor using five years of collision data for the "before" window from the TxDOT Crash Reporting Information System (CRIS); the data for the longest duration accessible at the time of this analysis was used for the "after" condition. Sales tax data were collected from the state comptroller's office for the four-month before period and four-month after period for businesses along the study corridors.

Crowdsourced Passive Data

Connected Vehicle Travel Data

Wejo connected vehicle (CV) data include information on vehicle movements and driving events for newer passenger vehicles. The vehicle movements consist of geographic coordinates for every three-second interval, speed, time stamp, heading, year, make, model, and journey IDs. Driving events include geographic coordinates, driver seat belt status, event log for hard braking, hard acceleration, speeding above 80 mph, and ignition on/off events. Together, these very large datasets (multiple terabytes) require special computing techniques to process, analyze, and visualize. Figure 5 is a visualization of Wejo CV journey counts and hard braking rates along US 285 near CR 435 and CR 454 in the TxDOT Odessa district.

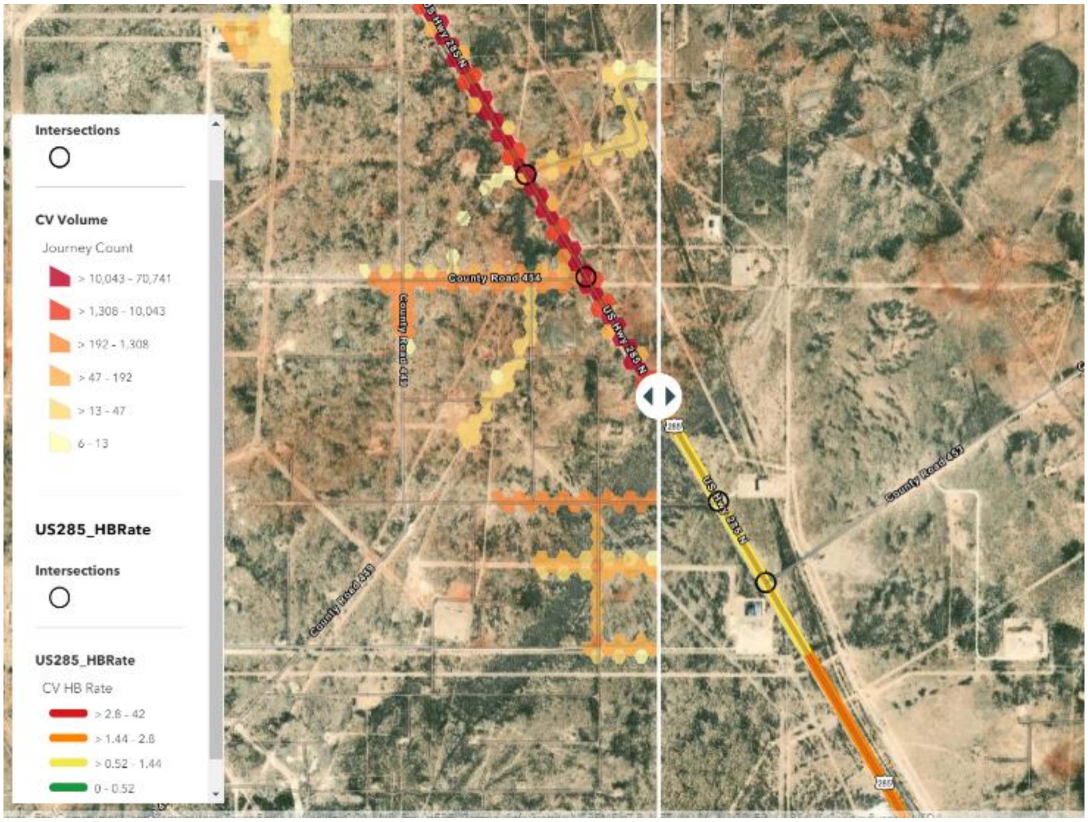


Figure 5. US 285 total connected vehicles (left) and hard braking rates (right).

It is supposed that certain driver events, such as hard braking, could serve as a surrogate measure of safety to potentially identify high-crash-risk locations before crashes occur. Currently, crash prediction models include only road characteristics and exposure, which does not account for drivers' behavioral influences. Researchers are currently developing crash prediction models with CV factors to explore whether a statistical and spatial relationship exists between CV events and historical crashes.

SafeGragh Spend Data

SafeGraph's Spend data include information about spending patterns by looking at debit and credit card transactions aggregated monthly for individual places in the United States going back to January 2019 [8]. According to SafeGraph, this dataset is ideal for:

- Competitor Analysis (e.g., In which markets are store sales growth outperforming competitors?)
- Site Selection (e.g., What co-tenants drive sales volume for individual stores within a brand?)
- Impact Measurement (e.g., Where did sales go up the most for a new product launch?)
- Any other use-cases where dynamic spend insights at individual locations are paramount

Organizations can use SafeGraph data to help make better decisions for growth. This proof-of-concept study explored relationships between the Spend data, traffic, and sales tax data to ascertain if any surrogate measures can be used to estimate the economic effects of access management projects.

Sales Tax Data

Sales tax data were requested from the Texas State Comptroller for the study locations. The comptroller's office provided a list of all taxpayers within the requested limits. The taxpayers—in other words businesses—were selected to include in the analysis and the comptroller provided an aggregated monthly total.

Vehicle Crash (Safety) Data

Vehicle safety data were collected from the Texas Department of Transportation (TxDOT) Crash Reporting Information System database. TxDOT collects crash reports from Texas law enforcement agencies for crashes occurring on public roadways and the state highway system. The retention schedule for crash reports and data is 10 years plus the current year. The database includes information about the date, time, severity, and manner of collision.

Chapter 4. Study Corridors

Wejo and SafeGraph are new, innovative data sources with information available since 2019. Therefore, study corridors were selected where the “before” and “after” conditions for the project could be evaluated between 2019 and 2021. In addition, it was reasonable to exclude 2020 data in this analysis because of the unique travel patterns due to the COVID-19 pandemic. The following two corridors were selected for this proof-of-concept evaluation.

RM 620 – Lohman’s Crossing to Bella Montagna Circle – Lakeway, TX

The RM 620 project corridor is fronted by many typical commercial businesses, such as grocery stores, banks, gas stations, and fast-food restaurants. Businesses are scattered along the entire stretch of roadway, but the retail and commercial development is concentrated more heavily on the west or south side of the project. Prior to 2020, this corridor had two through lanes in each direction with a two-way left turn lane. Access to adjacent properties was through eight street intersections and 52 private driveways. Of the 52 private driveways, 4 included a pork chop island that restricted access to right-in and right-out movements only. RM 620 is a high-volume arterial street where concrete block raised medians were installed on a 1.4-mile-long section of roadway in late summer/early fall 2020. A raised concrete median was installed on this commercial corridor in 2020 to restrict left turn movements. Access points that allow left turns along the project corridor upon completion of the project are shown in Figure 6. Following completion of the project between Lohman’s Crossing and Bella Montagna Circle, the corridor included three full access median openings and six three-quarters access (right-in/right-out/left-in) points.

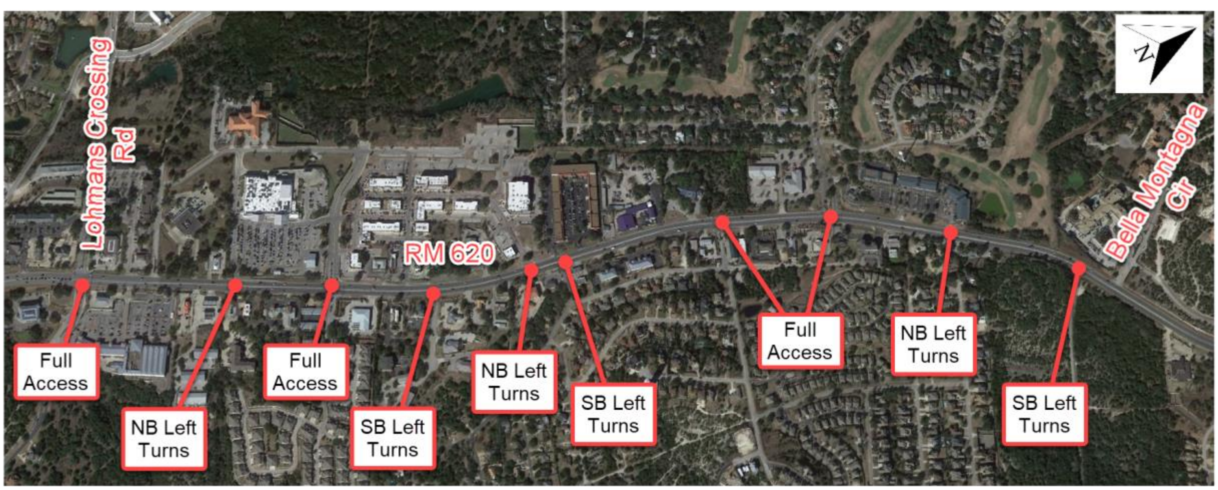


Figure 6. RM 620 – Lohman’s Crossing to Bella Montagna Circle left turn access points.

George Bush Drive (FM 2347) – Wellborn Road (FM 2154) to Houston Street – College Station, TX

The FM 2347, or George Bush Drive, project corridor is fronted on the south side by a gas station and other miscellaneous commercial businesses. Prior to 2020, this corridor was two through lanes in each direction with a two-way left turn lane, with access provided by three street intersections and five private driveways. The George Bush Drive project included installation of raised medians in February and March 2021. A raised concrete median was installed on this 0.25-mile-long commercial corridor in 2020 to restrict left turn movements. One three-quarters access (right-in/right-out/left-in) point that allowed left turns remained after the median was installed and is shown in Figure 7. Although the construction project covered a larger area to the east, this analysis focused on a 0.3-mile segment where there is some commercial development.

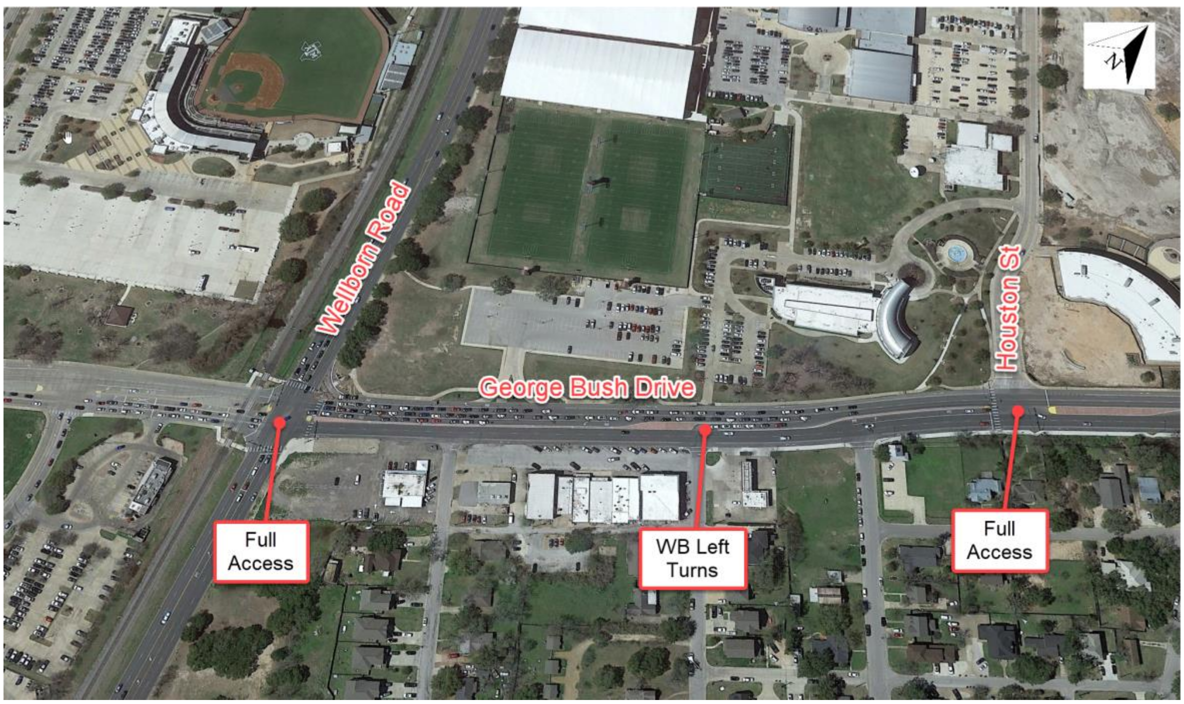


Figure 7. George Bush Drive – Wellborn Road to Houston Street left turn access points.

Chapter 5. Data Analysis and Results

To lay some background for the following data analysis exercise, Table 3 shows a list and schedule of special events that took place in College Station during 2019 and 2021, and which could influence the observed traffic patterns on George Bush Drive during the analysis time window. The event schedule for both the “before” and “after” time windows is included.

Table 3. Events and Activities in College Station during the Analysis Period

Date(s)	Event
August 9–10, 2019	Summer 2019 Commencement
August 17–21, 2019	Move-in Day/Howdy Week
August 26, 2019	First Day Fall 2019 Semester
August 29, 2019	Aggie Football v Texas State – Thursday Night Football
September 14, 2019	Aggie Football v Lamar
September 21, 2019	Aggie Football v Auburn
September 27, 2019	Aggie Ring Day (see details below)
October 12, 2019	Aggie Football v Alabama
October 26, 2019	Aggie Football v Mississippi State
November 1, 2019	Aggie Football v Univ of Texas at San Antonio
November 8, 2019	Aggie Ring Day
November 16, 2019	Aggie Football v South Carolina
November 27–29, 2019	Thanksgiving Break
August 13, 2021	Summer 2021 Commencement
August 26–29, 2021	Move-in Day/Howdy Week
August 30, 2021	First Day Fall 2021 Semester
September 4, 2021	Aggie Football v Kent State
September 18, 2021	Aggie Football v New Mexico
September 24, 2021	Aggie Ring Day
October 2, 2021	Aggie Football v Mississippi State
October 9, 2021	Aggie Football v Alabama
October 23, 2021	Aggie Football v South Carolina
November 6, 2021	Aggie Football v Auburn
November 12, 2021	Aggie Ring Day
November 20, 2021	Aggie Football v Prairie View A&M
November 24–26, 2021	Thanksgiving Break

These special events can have an observable effect on the traffic volumes and patterns, which is discussed in the next section. For example, more than 30,000 Texas A&M University students, friends, and family members attended the Aggie Ring Day festivities at the Clayton W. Williams, Jr. Alumni Center on Friday, September 27, 2019. According to estimates from The Association of Former Students, 4,300 Aggies received rings with the help of more than 320 volunteers and Association staff. Such events can increase both the volume of traffic and the associated travel-related activities such as shopping, recreational activities, and eating out, among others.

Wejo Connected Vehicle Travel Data

For this study, researchers utilized Wejo's connected vehicle (CV) data to track traffic volume, journey paths, and turning movements at two locations in Texas: George Bush Drive and RM 620, as outlined earlier. The analysis used eight months of Wejo data from August to November 2019, collected by one original equipment manufacturer (OEM), and from August to November 2021, collected by four OEMs. Wejo's connected vehicle data provide a wealth of information about vehicle movements, including location, time, speed, heading, and other attributes. Moreover, the data also contain events such as hard braking and hard acceleration, which can be used to gain insights into driving behavior. Figure 8 illustrates data captured in Texas during a one-hour period on November 5, 2021, containing approximately 120 million data points and providing insights into the movement of vehicles across the state. The data points are collected every three seconds and they cover both urban and non-urban areas.

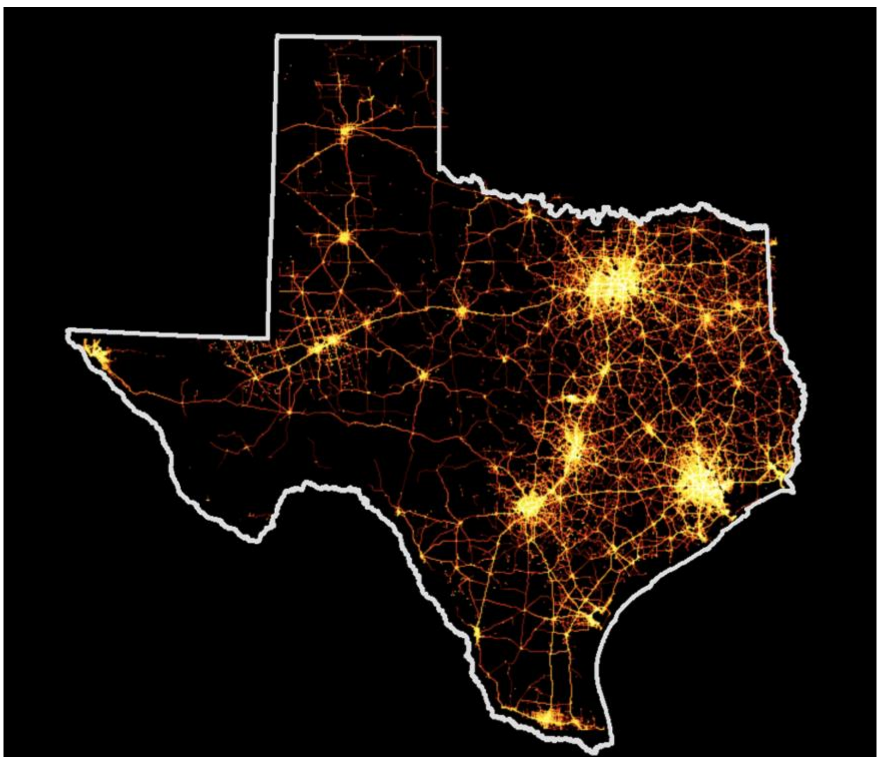


Figure 8. Wejo connected vehicle data for one day in Texas.

To analyze these data at a more granular level, the project focused on two study corridors in Texas, George Bush Drive and RM 620. The analysis was carried out using ArcGIS, a geographic information system, to create polygons and perform a geospatial join to filter the data for those areas. For this, researchers used Microsoft Azure cloud services as a storage and processing solution for the connected vehicle data. The GPS data were spatially linked to roadway polygons using Databricks, a cloud-based platform that leverages the Mosaic extension of the Apache Spark framework. This approach allowed the researchers to spatially index the GPS data using Uber's H3 hexagonal grid system. The analysis of data from George Bush Drive revealed 563,283 journeys, 53,991 hard accelerations, and 19,794 hard braking events for the eight months of analysis. Meanwhile, data from RM 620 showed 949,631 journeys, 157,621 hard accelerations, and 84,876 hard braking events. Wejo CV data also allowed the project team to track journey paths. Specifically, the analysis tracked journeys of vehicles that visit a property, and journeys with ignition on/off events, as shown in Figure 9. Midpoint journeys are where the vehicle visits a property/business but does not stop (i.e., there are no ignition on/off events), but there are three or more data points within the property/business area.

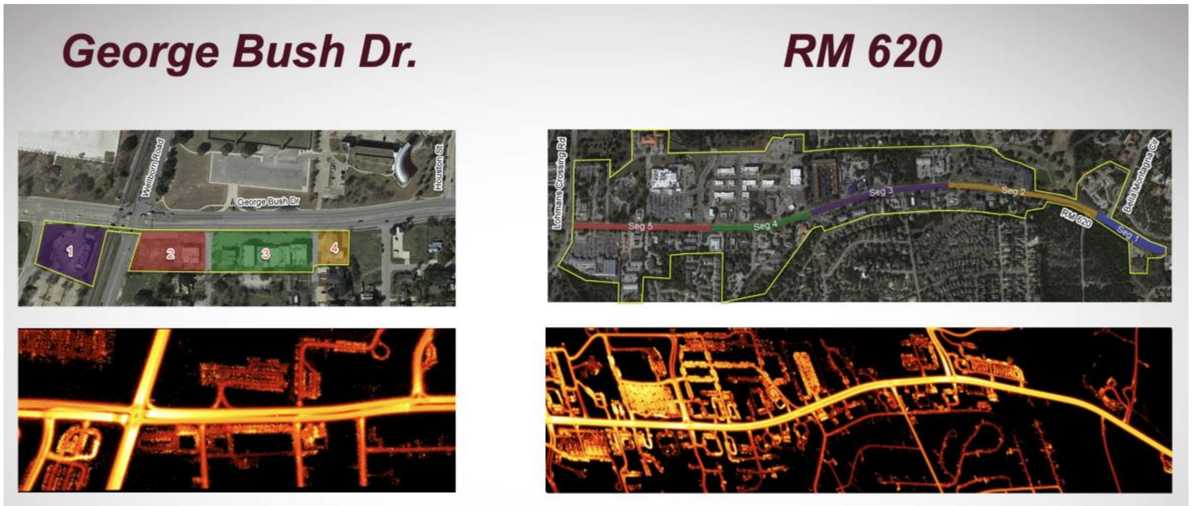


Figure 9. Journey paths traced using Wejo data.

In addition to the analysis of traffic volume and journey paths, researchers also examined turning movements. This was accomplished by utilizing location, time, and heading attributes from the Wejo data. Turning movements typically occur within a short time frame, usually a few seconds, at a single crossover location. For left and right turns, the vehicle's heading change should be around 90 degrees, while for U-turns, it should be approximately 180 degrees. Figure 10 shows a left turn performed during a journey on George Bush Drive. The data points in the exhibit were collected at three-second intervals, and the colors of the points indicate speed, with red indicating slower speeds and green indicating faster speeds. This visualization provides a clear representation of the vehicle's movement during the left turn maneuver and allows for a more detailed analysis of the vehicle's behavior during the turn.

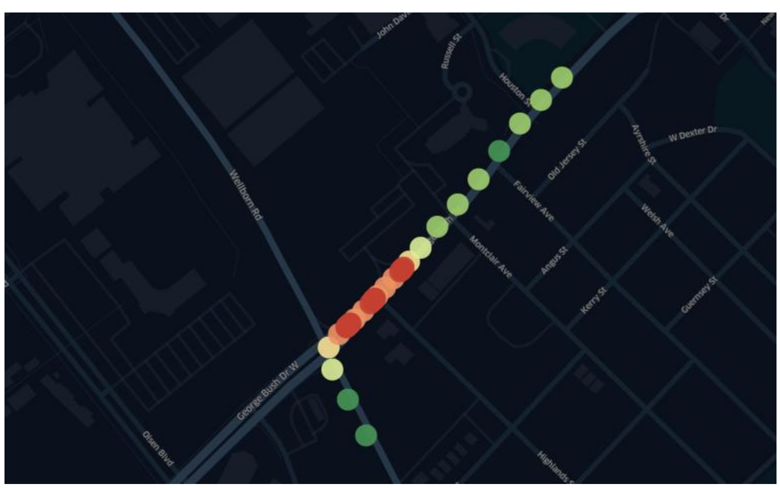


Figure 10. Example of left turn maneuver.

Extending the discussion in an earlier section on the effect of special events on local travel volumes and patterns, the following figures show the number of vehicle journeys captured in the Wejo database during the analysis time windows for years 2019 and 2021. The journey information is available by day. This is helpful for several reasons:

- One can do a quick test for expected pattern of traffic volumes and see if the journey frequencies follow anticipated levels. For example, weekend numbers are lower than Monday through Friday.
- Also, one can filter and see the effects of events or activities that might drive the journey numbers higher. George Bush Drive borders the largest university in Texas, so it might be anticipated that football gamedays have higher recorded journeys. Summer and Thanksgiving Break are quite a bit lower than typical weekdays.
- Finally, it is easy to compare time periods to see if there is anything that might affect other business metrics like sales or the number of customers. For example, September and November each had two games both in 2019 and 2021. However, in 2019, there were two football events in October compared to three in 2021. This could drastically affect monthly economic data given the size and scale of the events.

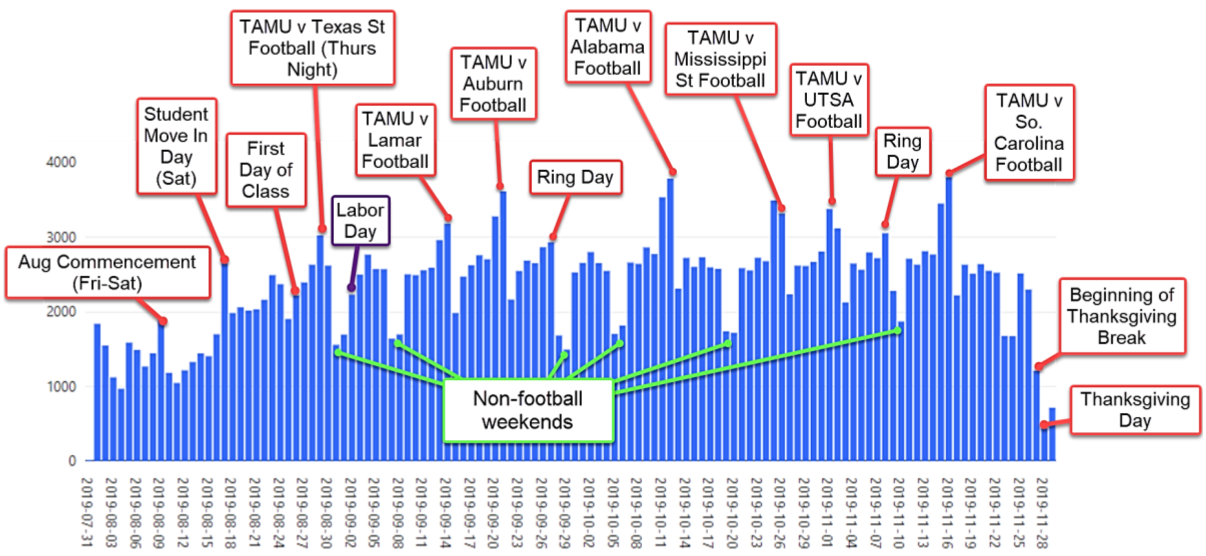


Figure 11. 2019 George Bush Drive Wejo journeys.

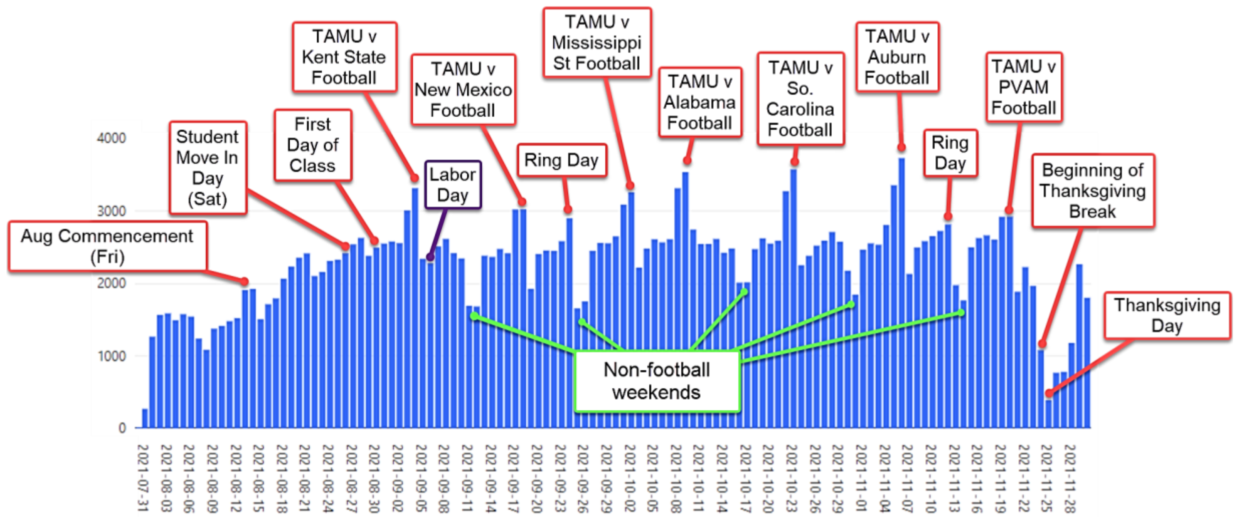


Figure 12. 2021 George Bush Drive Wejo journeys.

Figure 13 shows the number of journeys recorded for the George Bush Drive study area and the percentage of hard braking and hard accelerations observed. Again, the 2019 data represent the before condition and are the darker of the colors. The after period and the lighter color is 2021.

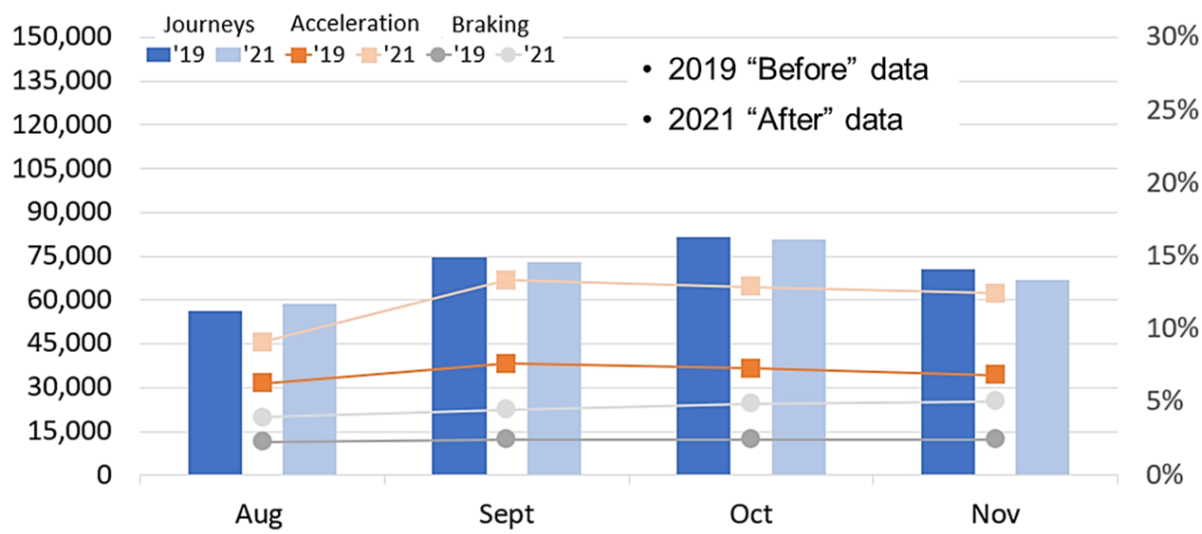


Figure 13. George Bush Drive Wejo journeys and hard acceleration/braking events.

Figure 14 and Figure 15 illustrate the previous graph in more detail. Figure 14 shows the number of Wejo journeys observed during the two four-month analysis periods. As can be seen on this graph, there was minimal change (within +/- 5%) from 2019 to 2021. This is somewhat unexpected because the 2021 data included more manufactures, so one would expect more sample data points. The negligible change could be an after effect of COVID-19 and the overall average daily traffic still being lower than 2019.

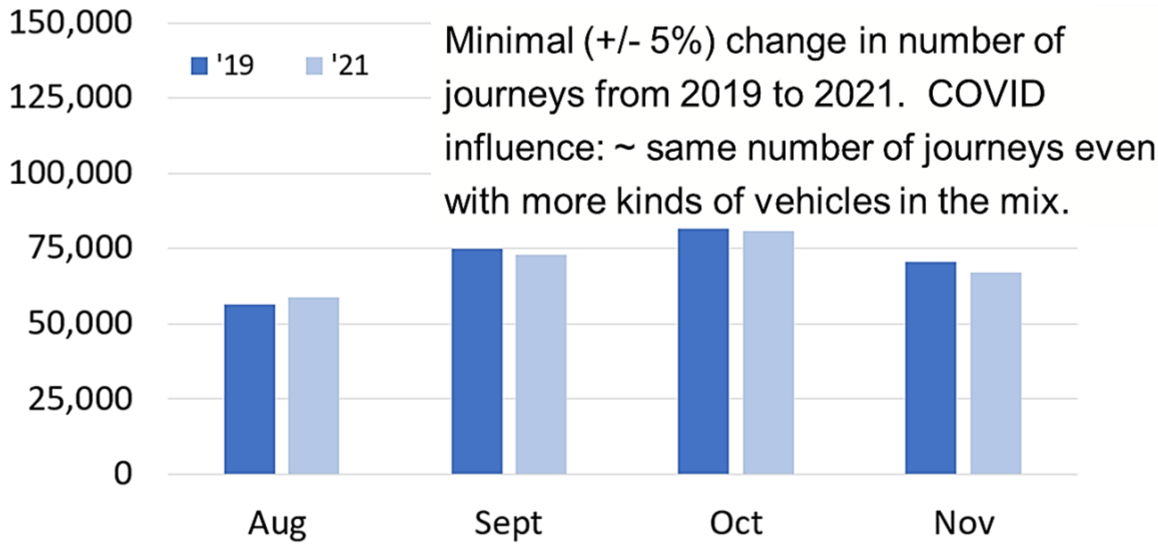


Figure 14. George Bush Drive Wejo journeys.

The graph in Figure 15 includes hard acceleration and braking event ratios based on the number of events reported relative to the total journeys in the study area. In each month there was between a 50% and 100% increase in the ratio from 2019 to 2021. This observed increase is likely due to manufacturers having different thresholds for these events—the same acceleration or braking rate may be considered an event in 2021, but not 2019. The hope was to be able to use this information as a stand-in or a leading indicator for safety effects. However, after digging into the data, researchers found that different vehicle manufacturers classify hard braking and hard acceleration at different threshold values. The 2019 Wejo dataset included only one manufacturer while there were four in 2021. The dataset does not include the threshold or the actual value for these events; it only indicates that an event occurred. Therefore, it was impractical at this point to assess if there was truly an increase or decrease in the event ratios. Further, there are ongoing investigations into the relationship between hard braking and crash frequency and severity. If the relationships are established, this data could be used to provide early safety results, but only if data from a single manufacturer is analyzed or a common threshold is used by all manufacturers to define events.

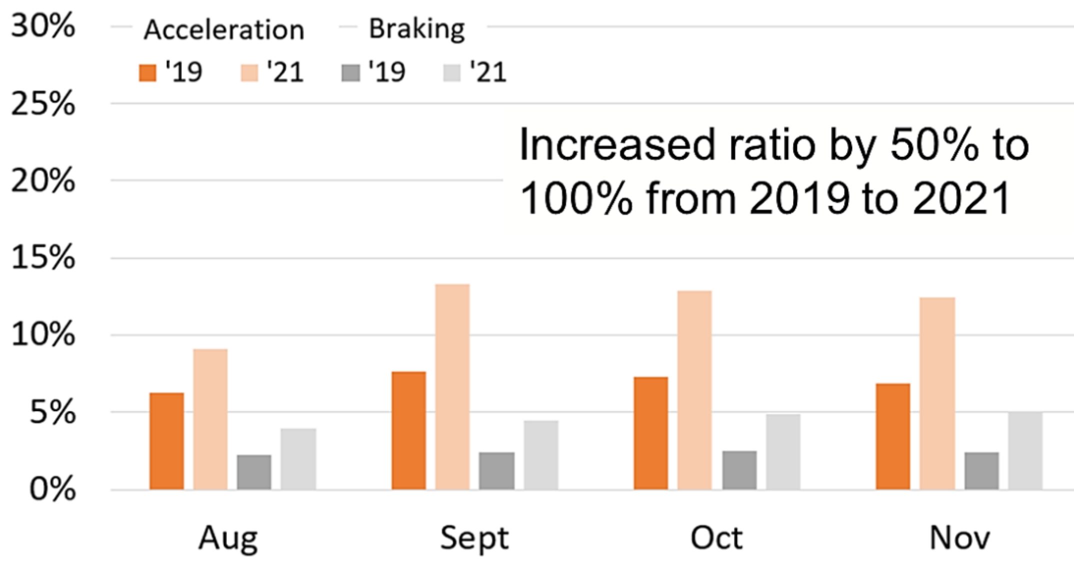


Figure 15. George Bush Drive Wejo hard accelerations and braking ratios.

The same data were evaluated for RM 620, as shown in Figure 16. The observations are similar to those at the George Bush Drive location, but here the number of journeys in the study area was down by 10–12% depending on the month, and the hard acceleration and braking ratios increased in the range of 52% to 116%.

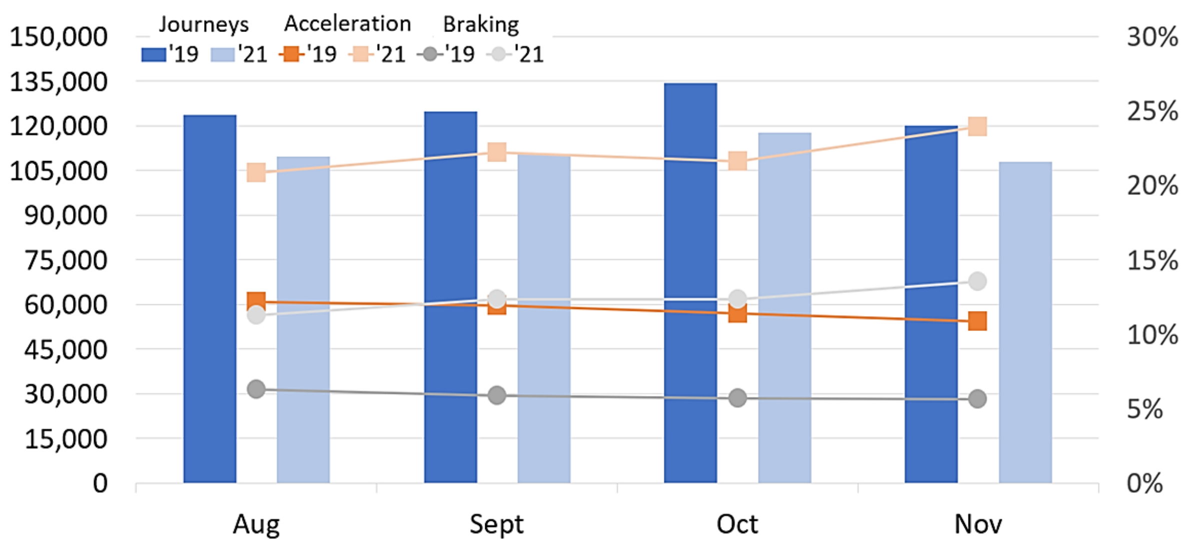


Figure 16. RM 620 Wejo number of journeys and hard acceleration and braking ratios.

To get a more detailed look at how crowdsourced connected vehicle data such as Wejo's can help in understanding patterns, the next few figures go through the journey type and turning movement results using examples from four different property areas along George Bush Drive to provide a flavor of the type of information available and how it applies to different scenarios. These four different areas are highlighted in Figure 17.

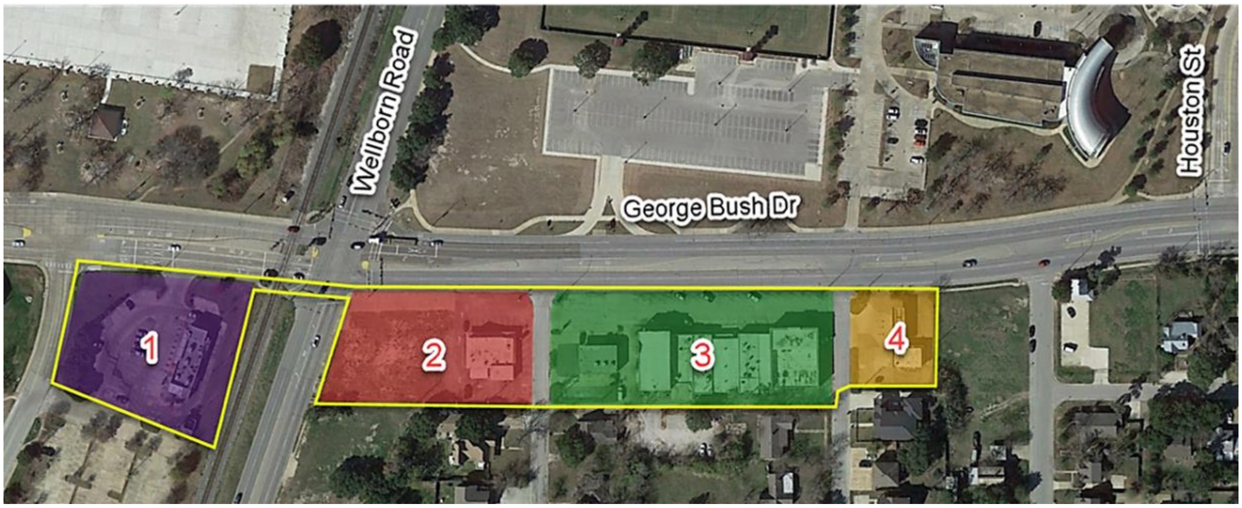


Figure 17. George Bush Drive – analysis properties.

In Figure 18, the rates shown in yellow indicate the frequency of the respective journey type for that property compared to all journeys on the main roadway. For this property (property 1), there is no change in median access. So, this was used as a control point. Median construction was on the other side (east side) of Wellborn Road. This is a fast-food restaurant land use. Midpoint journeys are four to six times the ignition on and ignition off journeys, which makes sense given the type of business and people using the drive-through windows.

The slight increase in midpoint journeys could be attributed to how people received service at fast food restaurants post COVID-19—there were more pick-ups and drive-throughs than dine-ins. The increase might be more than what is apparent because the on- and off-type events went down, so the delta is closer to 0.6%.

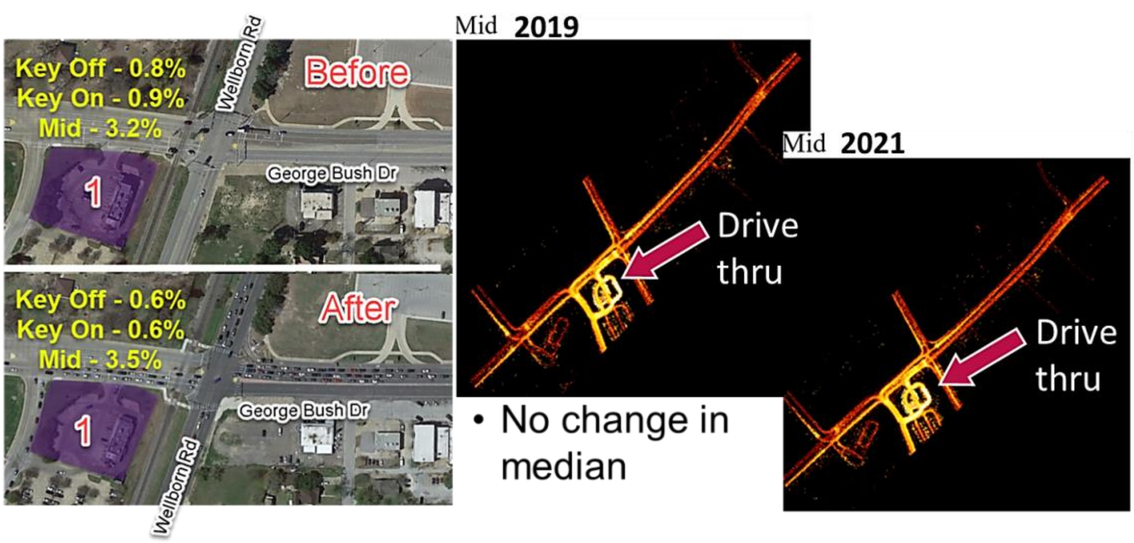


Figure 18. George Bush Drive – property 1 journey type rates.

Property 2 is a retail land use and is considered a destination business. Key-off journeys and turning ratios are shown on the left in Figure 19—this is the number of turning vehicles compared to the overall number of vehicles going in the same direction on the street. Note that in some cases the driveways are closely spaced or not clearly defined, so for those cases the turning movements have been combined (for example, two right turns from George Bush Drive and the rights and lefts from Highlands Street).

- Key-off events increased from 0.8% of journeys on George Bush Drive to 1.4% (a 75% increase).
- There was a change in access: left turns to and from Highlands Street were removed.
- Vehicles using that left turn access appear to have selected alternative access. These alterative paths are clearly shown in the journey heat maps on the right as well.
    - Option 1: Made a U-turn at Wellborn Road (U-turn increased from 0% to 0.2%) and then a right turn into the business (right turns increased from 0.9% to 2%).
    - Option 2: Made a left turn in advance of the property at Montclair Ave—a street to the east—and then traveled through the parcel to the east of property 2. This can be seen by a darker line south of and parallel to George Bush Drive on the "after" heat map.
    - Option 3: Come from the south. This is reflected in the left turning ratio from Highlands Street, which nearly doubled (increased from 2.7% to 5.3%).

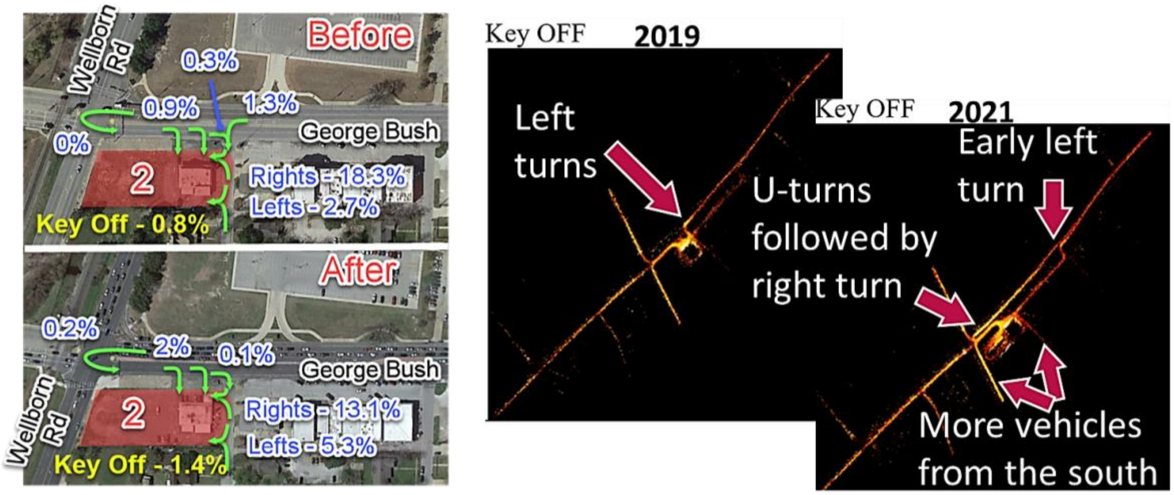


Figure 19. George Bush Drive – property 2 key-off journeys and inbound turning movement rates.

Figure 20 illustrates the data when looking at a parcel that has multiple businesses, including retail, restaurant, and office land uses.

- Four out of the seven businesses are the same in the before and after time windows.
- Overall journeys to and from the property decreased by nearly 1%, from 3.3% to 2.4%. It is unclear if the change in journeys was a result of the medians or if the change in business types and their trip generation caused the reduction.
- Based on the heat maps shown on the right, there are fewer journeys to and from the area west of Wellborn Road and there are fewer journeys recorded going westbound in front of the businesses. This is likely a result of fewer people turning left into the parking lots or onto Highlands Street.

- With the only opportunity for a left turn being on the east side of property 3 at Montclair Avenue, the westbound left turn at this location increased from 2.8% to 5.4%.

Figure 20. George Bush Drive – property 3 all journeys and inbound turning movement rates.

The fourth and final property analyzed along George Bush is shown in Figure 21. For this example, researchers looked at the rates for ignition or key-on events. This property is a gas station, which is generally associated with pass-by trips. Trips back to the residential area southwest of the project corridor disappeared. As shown on the right, there was a grocery store with gas pumps constructed during the project timeline, which likely contributed to the change in turning movement rates. For example, even though there were no left turns allowed out of the site in the "after" condition, the right turns decreased from 1.3% to 1.2%. Similarly, turn movements along Montclair Avenue decreased from 9% to 1.7% and from 7% to 5.7%.

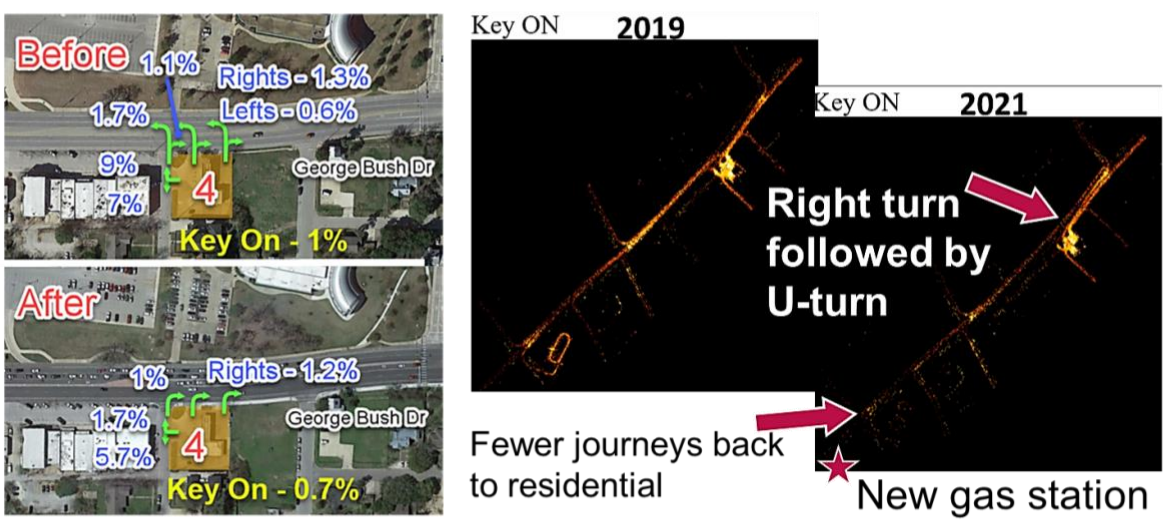


Figure 21. George Bush Drive – property 4 key-on journeys and outbound turning movement rates.

The second study corridor (RM 620) was divided into five segments, as shown in Figure 22, and turning movement rates were generated for each of them. One individual property on Segment 5 is discussed for illustrative purposes in Figure 23.

Figure 22. Study segments along the RM 620 corridor.

This property is a grocery store and did not have any changes in allowable inbound turning movements. When comparing a turn percentage for a movement in the "before" exhibit to the same location in the "after" exhibit, turning movement rates increased or stayed the same along RM 620, but decreased along Main Street and Medical Drive. For example, the eastbound left turn movements from RM 620 increased from 6.9% to 9.1% and 9.4% to 11.6%. Similarly, the right turns at the western entrance remained the same at 11% and increased from 5% to 6.1% at Main Street. It is challenging to conclude from this data alone if the additional vehicles turning from RM 620 were cutting through or going to other businesses. These changes in turning rates are good for engineers and planners to know to help better design future projects (i.e., median openings and turn bay lengths), but when looking at business impacts, it would be better to have the key-on, key-off, and mid journey data to know if there was an increase or decrease in customers.

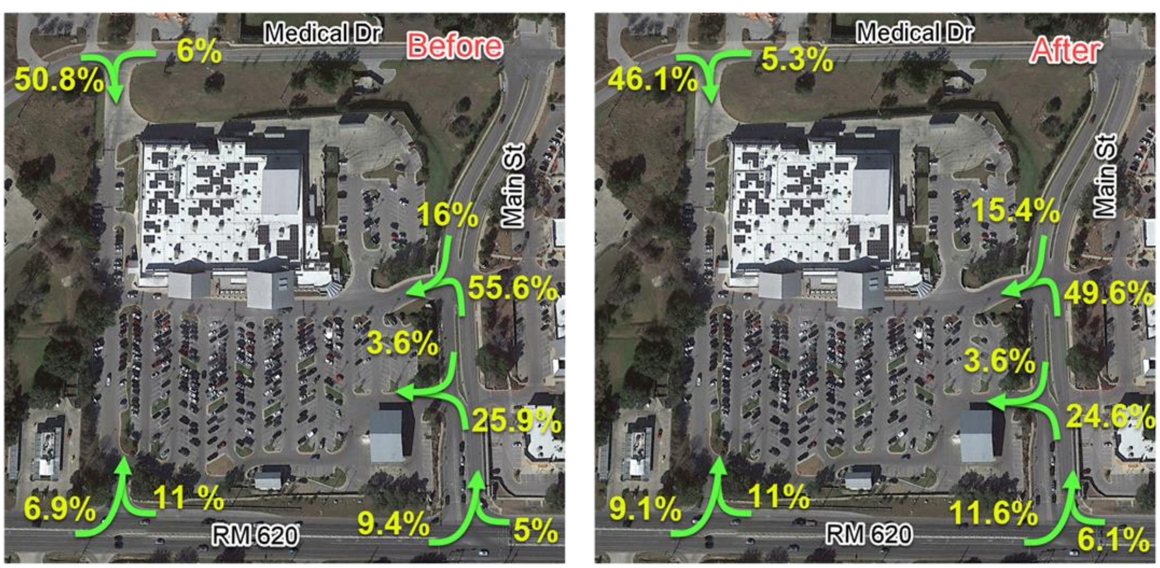


Figure 23. RM 620 – turning movement rates.

SafeGraph Spend Data

SafeGraph's Spend dataset contains anonymized contactless payment (debit and credit card transactions as well as Apple Pay and Apple Cash) data aggregated to individual places in the United States. This dataset helps reveal insights on how spend behavior is changing over time at specific points of interest, including the average transaction size, spend in-person versus online, spend by customer demographics such as income, and spend by customer loyalty.

As mentioned earlier, SafeGraph states this dataset is ideal for:

- Competitor Analysis (e.g., In which markets are store sales growth outperforming competitors?)
- Site Selection (e.g., What co-tenants drive sales volume for individual stores within a brand?)
- Impact Measurement (e.g., Where did sales go up the most for our new product launch?)
- Any other use-cases where dynamic spend insights at individual locations are paramount.

The Spend data are aggregated at a monthly time interval and delivered on the 20th day after each month's end. Historic Spend data are available back to January 2019. The dataset includes information on the type of business or service, location, monthly totals for total customer spending for that business, number of customers, median spending per transaction, median spending per customer, month-on-month changes in these metrics, and some other auxiliary information.

The following exhibits present data for a few example businesses on the two study corridors. Data for only those businesses were analyzed that were present in both the "before" and "after" time windows (years 2019 and 2021) at the same locations. Therefore, businesses that unfortunately closed or moved because of COVID-19 or otherwise in 2020 were not included for analysis. Figure 24 shows the total number of transactions for different types of businesses during the four-month analysis period for the "before" (2019) and "after" (2021) time windows on the George Bush Drive corridor. Figure 25 presents these data for the RM 620 corridor businesses.

As indicated earlier in this section, the Spend data from SafeGraph include information on total monthly sales of businesses as well, along with other metrics. However, information on total sales is not discussed here because it can be affected on a yearly basis by a multitude of factors including inflation, pricing strategy of businesses, product mix, and so on, which are more difficult to find information on at such a granular level and account for in this study. Moreover, businesses may care more about the effect of access management initiatives on the number of customers/transactions and footfalls at their locations. Therefore, information on the total number of transactions for different business types in the before and after period is presented in Figures 27 and 28.

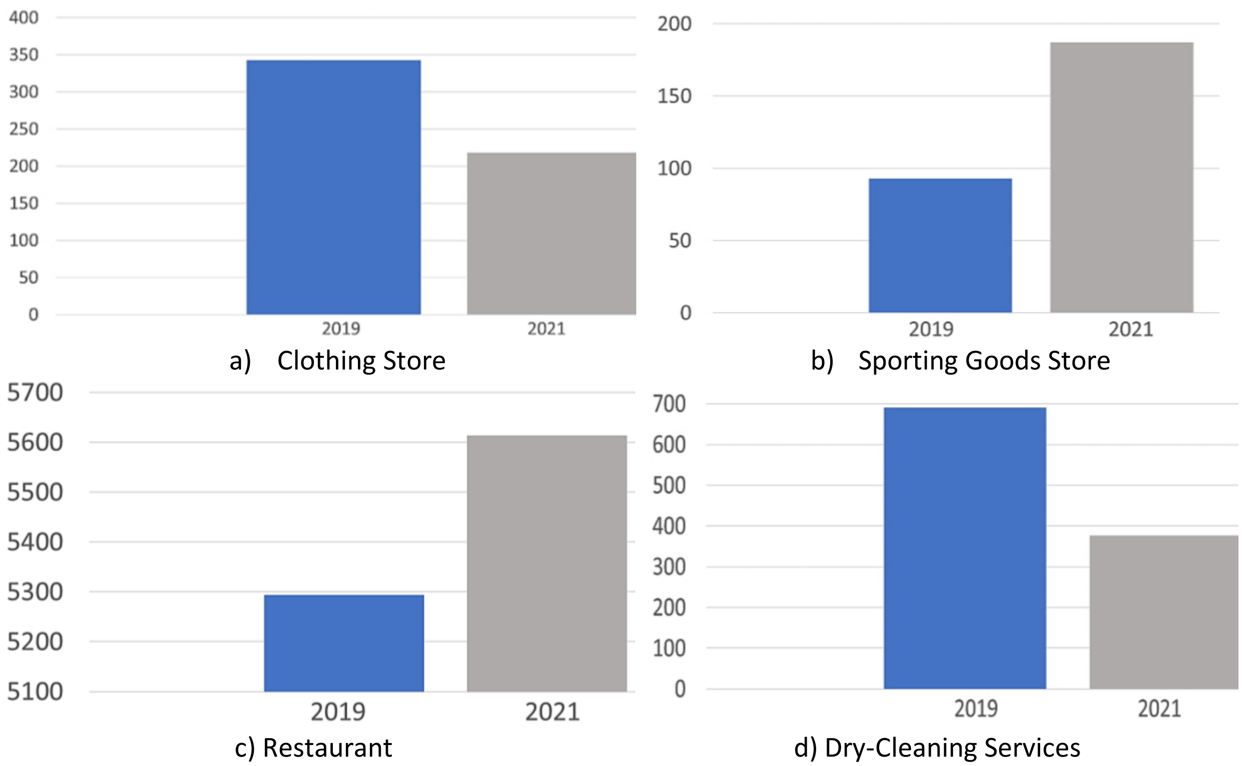
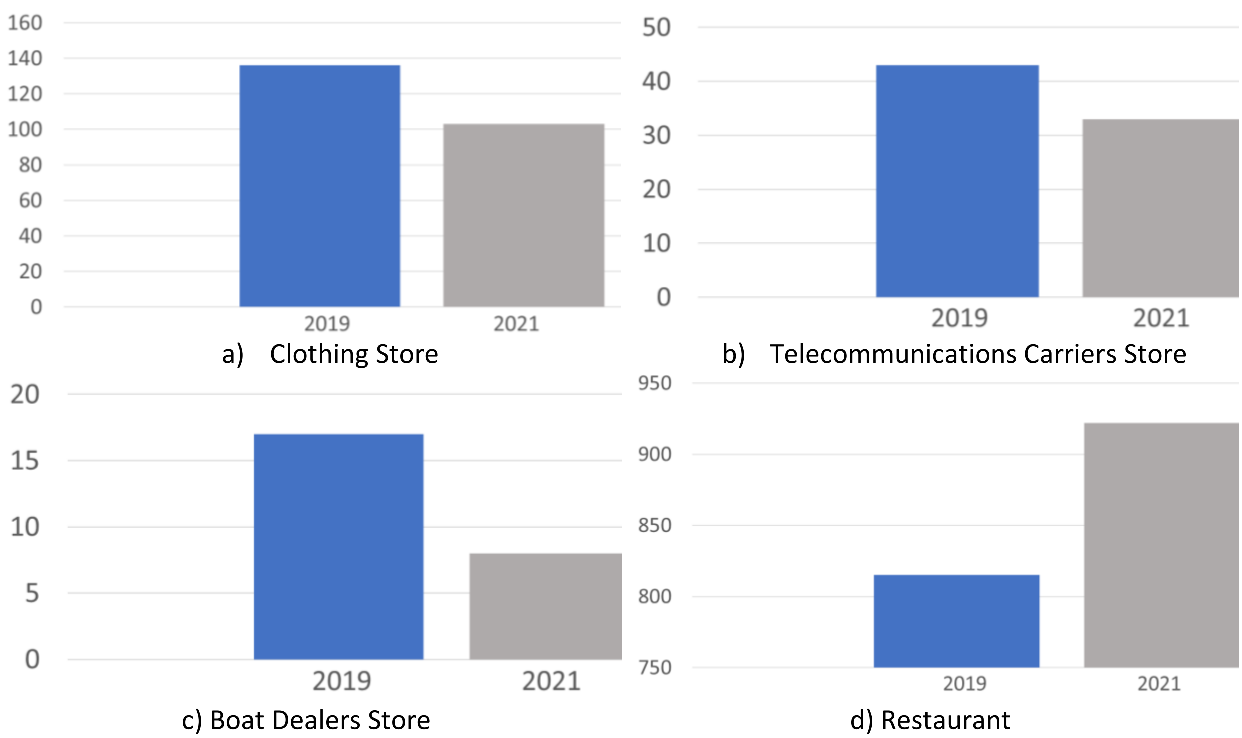
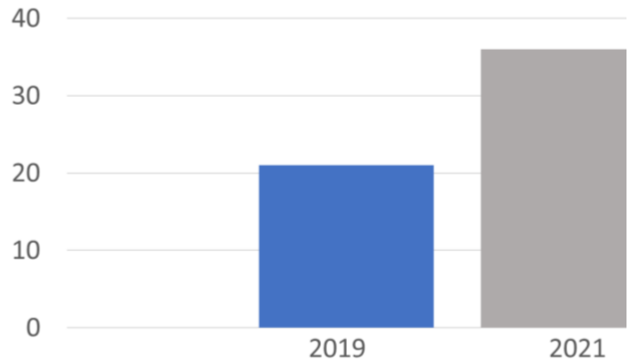
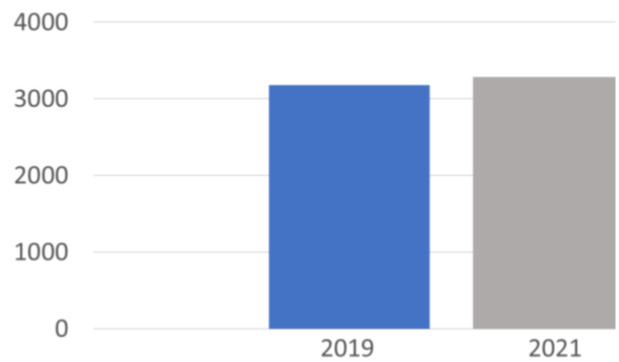


Figure 24. Number of transactions for businesses on George Bush Drive.

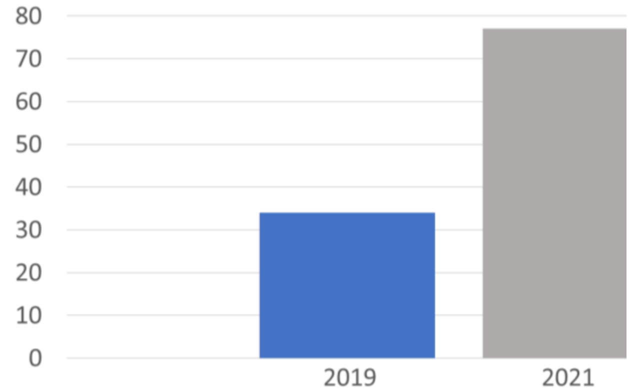




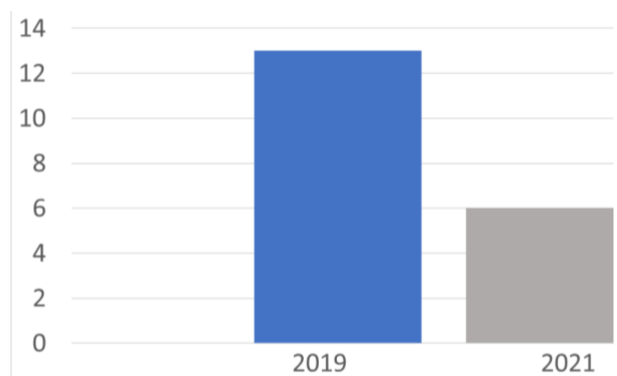
e) Fitness Service



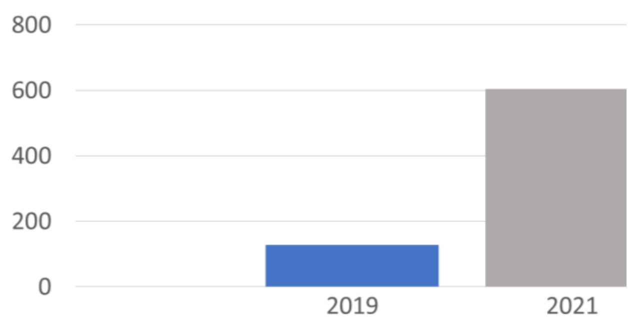
f) Health & Personal Care Store



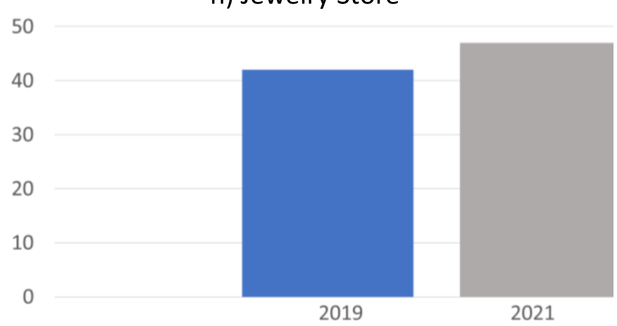
g) Personal Care Store



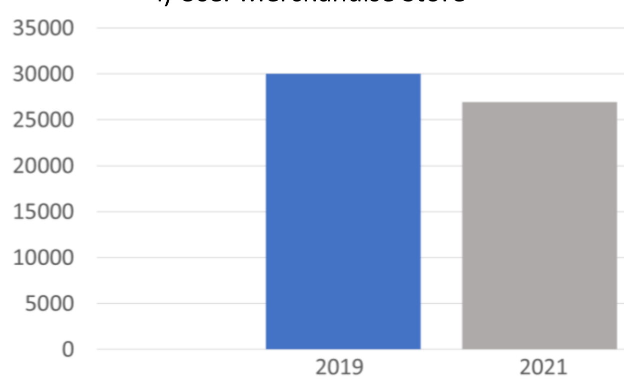
h) Jewelry Store



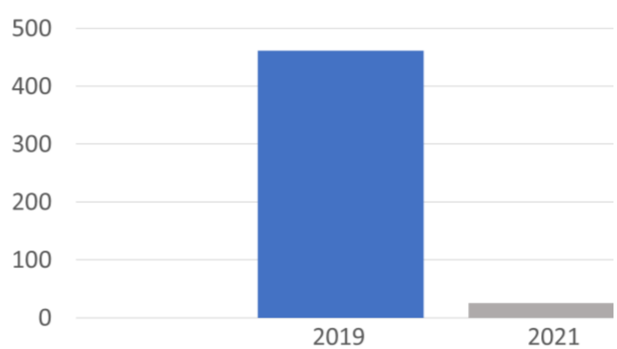
i) User Merchandise Store



j) Hotel



k) Grocery Store



l) Bowling Center

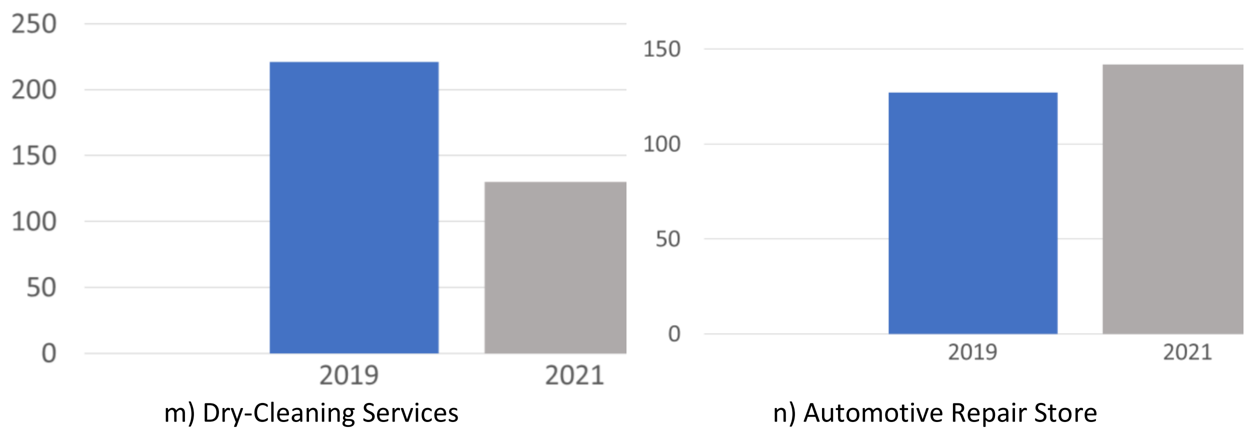


Figure 25. Number of transactions for businesses on RM 620.

The Spend dataset can be useful for comparing the same business or same type of business over time, but as illustrated in the two exhibits above, there is a lack of consistent trend that is applicable across all kinds of businesses on a corridor. The number of transactions or footfalls has increased for some types of businesses, remained about the same for some, and decreased for a few—indicating a lack of consistent growth or decline in business activity. Interestingly, in business categories that appeared on both George Bush Drive and RM 620, the transaction patterns were the same from before to after; restaurant transactions went up, while clothing stores and dry-cleaning services went down. Worth noting is that access management is just one of the factors that can potentially affect business activity, footfalls, or sales. There are several other factors at play, many of which are dynamic in nature, perhaps more important to business success than access alone, and difficult to capture with just a few variables of analysis.

Additionally, it is important to understand when looking at the data that the transaction information does not include all patrons of a business as the data panel includes only certain banks, which can change over time as contracts with third-party affiliates change and does not include cash pay customers. Contactless payment users can also choose to opt in or out of data sharing policies with their financial institution, thereby altering the sample size of potential data points. These fluctuations could lead to a change in customers or spending numbers regardless of a change in property access.

Sales Tax Data

Sales tax data requested from the Texas State Comptroller for the following location boundaries is shown in Table 4.

Table 4. Location Details for Sales Tax Data Collection

City	Zip	Street	From (block)	To (block)
College Station	77840	George Bush Drive	100	600
Lakeway	78734	Ranch Road 620S	600	2300

The comptroller's office provided a list of all taxpayers within the requested limits. Researchers selected the taxpayers to include in the analysis and the comptroller provided aggregated monthly totals shown in Figures 26 and 27. The comptroller's office reports data based on the allocation month. To accurately reflect the month the taxes were collected, the reporting period was shifted forward two months and then listed the taxpayers included in the total each month (Figures 28 and 29). This is where it was observed that several of the taxpayers disappeared from the before to after months. This was more noticeable in the George Bush Drive projects as there were far less taxpayers in total compared to the RM 620 corridor.

Worth noting is that in 2021 there was one more home football game in October compared to 2019. This could have impacted the sales tax collections for that month.

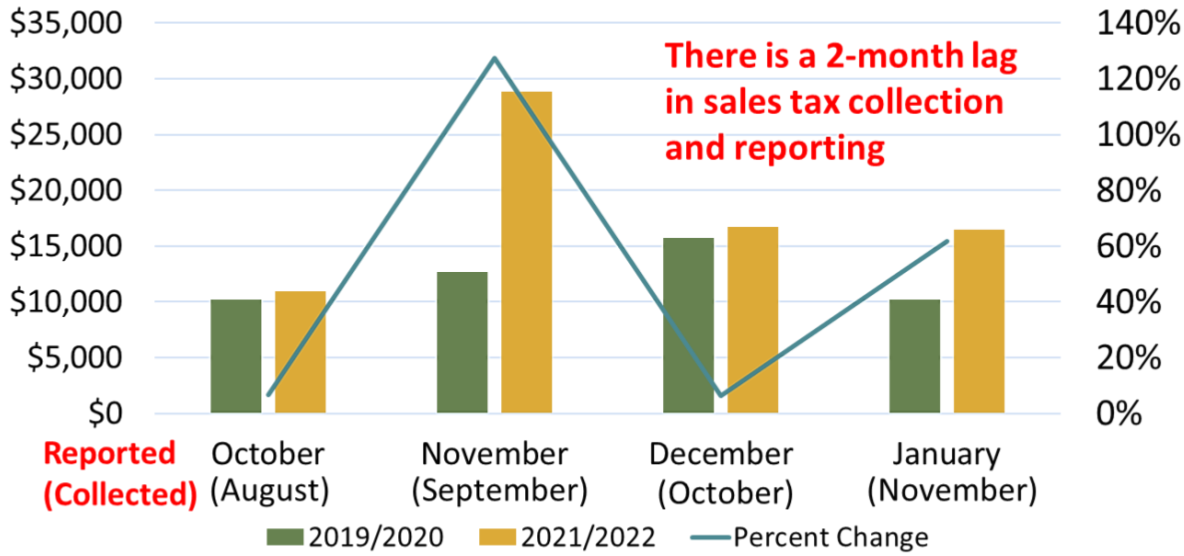


Figure 26. Sales tax data for George Bush Drive corridor.

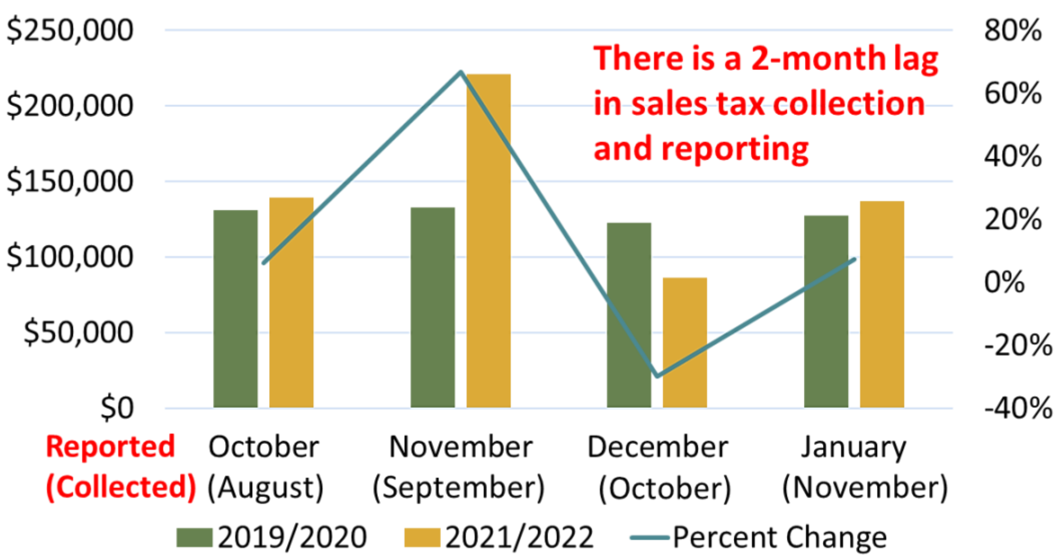


Figure 27. Sales tax data for RM 620 corridor.

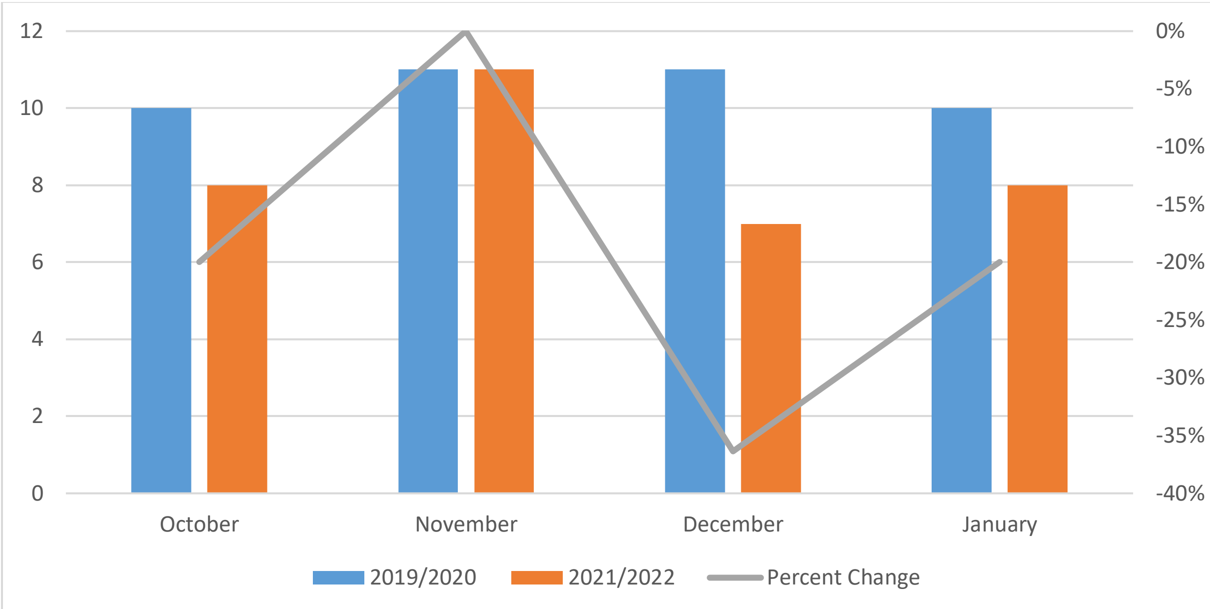


Figure 28. Number of businesses reporting sales tax data on George Bush Drive corridor.

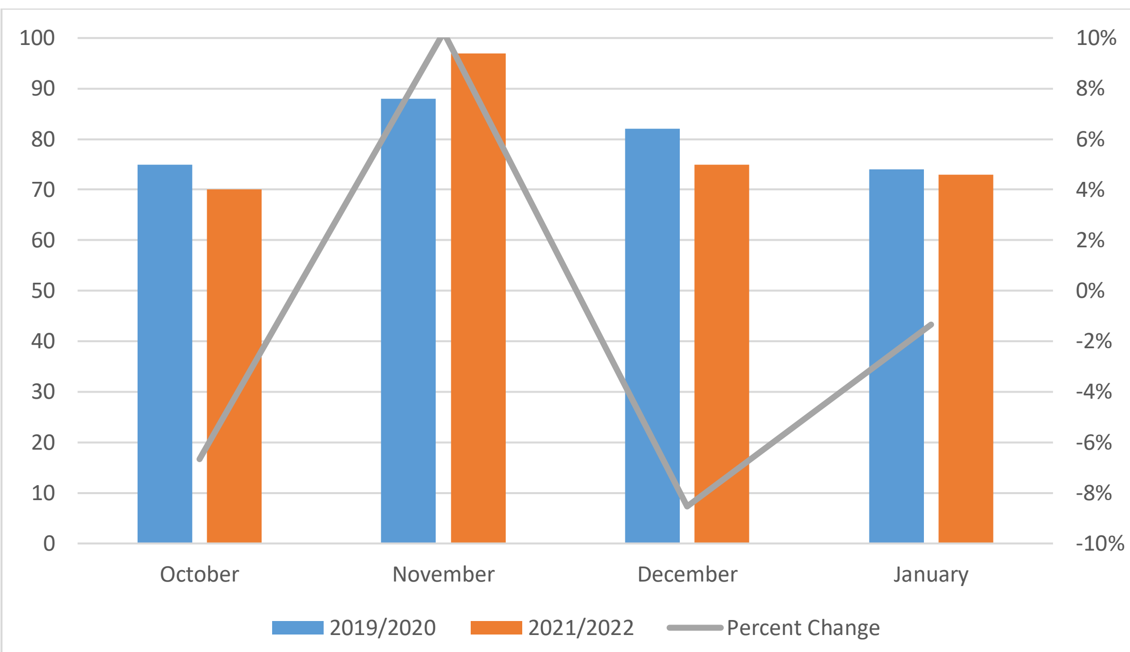


Figure 29. Number of businesses reporting sales tax data on RM 620 corridor.

The sales tax data showed increases in sales tax revenues in seven out of eight observations, with percent changes from year to year ranging from an increase of 127% to a decrease of 30%. These changes are too large to be caused by the change in access and are likely caused by other issues in the data, making the data poorly suited for this task for the following reasons:

1. The first issue with the data is that collection occurred during the pandemic when many businesses failed. During this period, 2 out of 12 businesses on George Bush closed and 2 new businesses opened. This included the closing of two restaurants and the opening of a grocery store and attached express center. These types of businesses produce dramatically different sales tax revenues. On RM 620, 19 out of 105 businesses closed, with 18 new businesses opening. The types of businesses that both closed and opened varied from restaurants and dry cleaners to pool and plumbing services. This makes it difficult to compare the time periods since the two periods don't share the same businesses. This could be mitigated by only comparing businesses that appear in the data for both years, however the dollar amounts are reported as an aggregate number to maintain anonymity. It may be possible in the future to take the analysis a step further and make a second request to the comptroller's office, requesting only the condensed list of taxpayers. These results may vary depending on the type of businesses excluded.
2. The second issue with the data is that businesses have different tax reporting schedules. Some report the revenues monthly, others bimonthly, others quarterly, while some report monthly but in an inconsistent manner. This makes comparing monthly changes difficult since a business may not report tax revenues in the same month each year. This could be mitigated by removing businesses that report inconsistently and or by expanding the period covered by the data. One possible solution could be to survey the businesses as to their sales tax reporting behavior and patterns. This would allow for a more tailored analysis.
3. The third issue with the data is inflation. Inflation as measured by CPI was about 8.3% from August 2019 to November 2021. Several months of the data showed a 6% to 7% increase in revenues, which would be a decrease in revenues if inflation was considered. The data could be adjusted for inflation, but it may not be appropriate to adjust revenues for local businesses using the national CPI.
4. The final issue with the data is that it is generally erratic and impacted by many outside factors, while the impact of reduced access likely has a relatively small impact. Much of the erratic nature can be explained by the previous points. However, even if those were accounted for, it may still be impossible to show any impact using the sales tax data since broader economic conditions could impact sales tax revenue much more than access. To show an impact, it would be necessary to control for general economic conditions using sales tax data from a broader area.
5. The limited number of observations also presents a problem. Revenue from a single month could be heavily impacted by one-time events. For example, a few days of bad weather in one month could have a large impact on revenues. This could be mitigated by having more observations in the data or increasing the period being covered by the data. If we were to examine a year of data before and a year of data after, there may be a way to account for some of the seasonality and different reporting time frames. If all these issues were accounted for, then this type of data could potentially be used to show an impact.

These issues highlight and underscore the challenges with assessing the economic effects of access management projects and why it is desirable to find an alternative to sales tax data.

Safety (Crash) Data

Crash data were also analyzed at the two study locations using the TxDOT Crash Reporting Information System (CRIS) database. As seen in Figure 30, there is a decrease in crash rates at the RM 620 location but an increase at the George Bush Drive location. Researchers assessed that a larger sample size was needed, thereby a longer "after" duration of observation to arrive at a meaningful finding from the crash data.

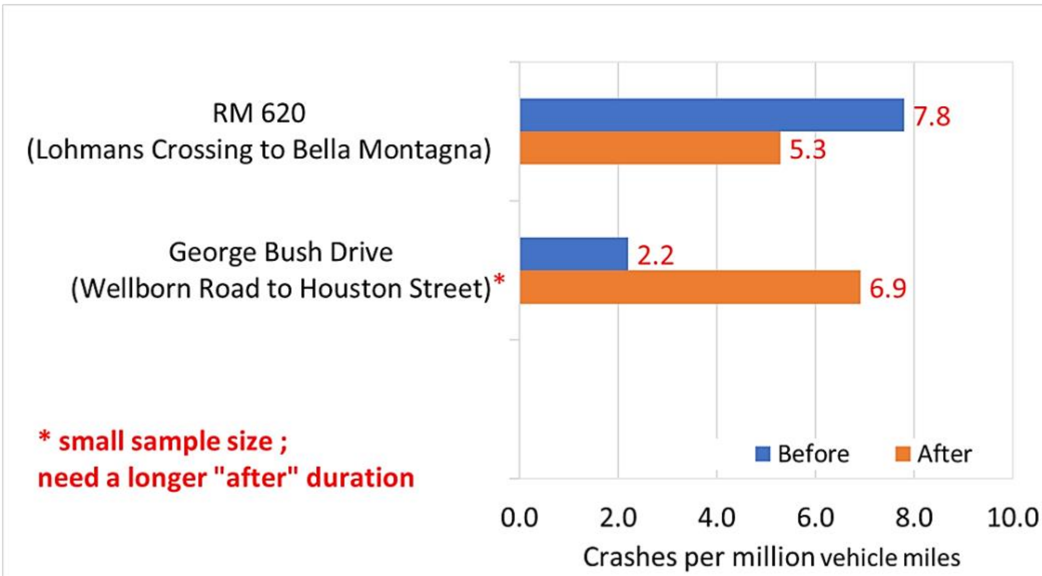
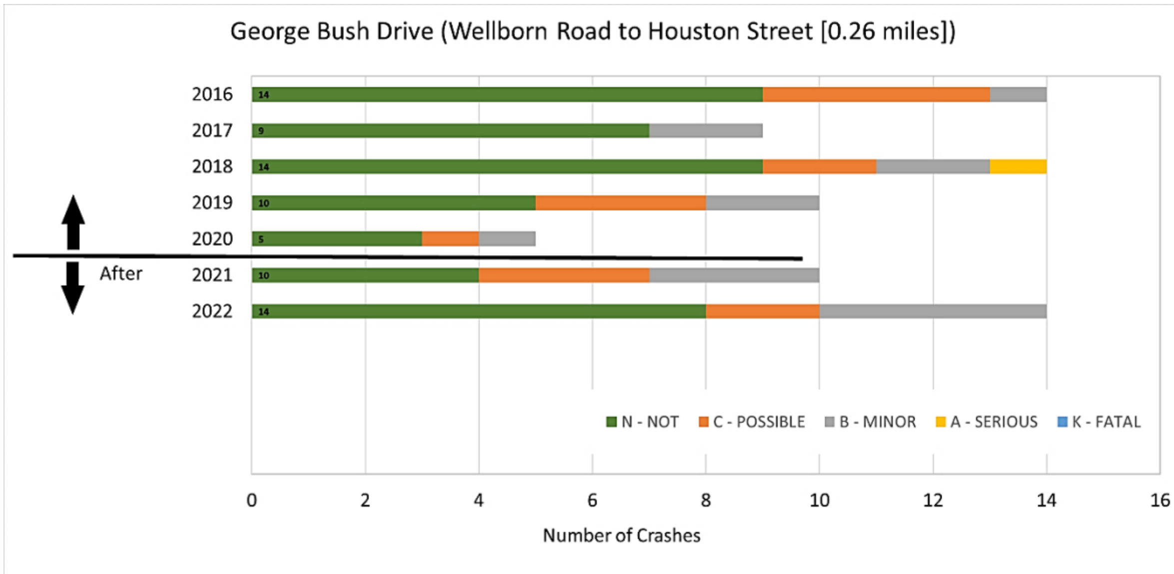


Figure 30. Crash data for the two study locations.

There might also be an overrepresentation of crashes occurring at the Wellborn Road and George Bush Drive intersection. Figure 31 presents data on the number crashes on the two study corridors broken down by the severity of the crashes.



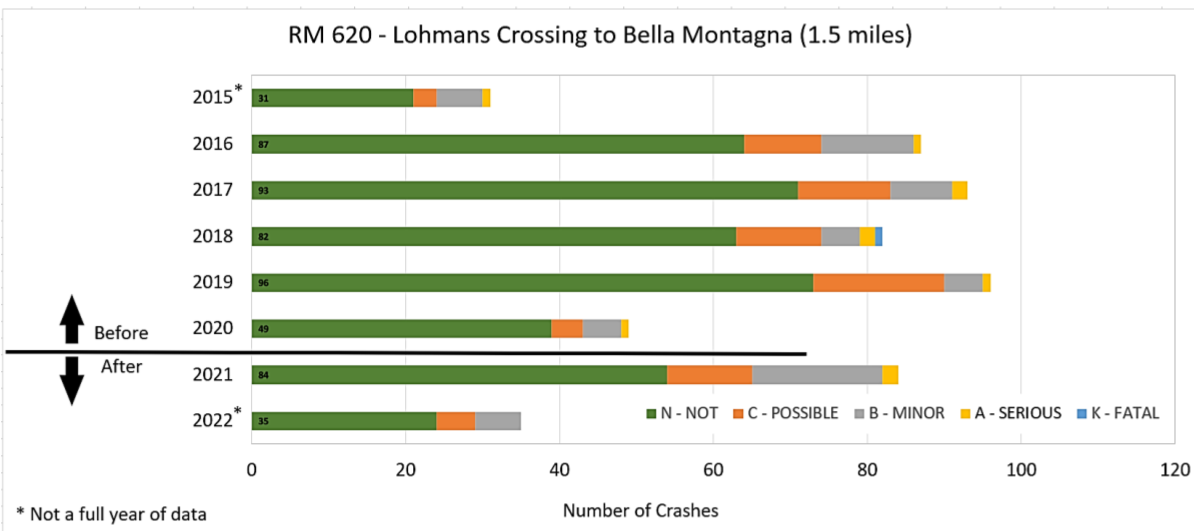


Figure 31. Crash severity data for the two study locations.

Chapter 6. Summary and Discussion

Implementing and justifying the viability of access management projects can be challenging for seemingly conflicting demands of engineers, planners, business owners, the public, and elected representatives. There have been efforts in the past to evaluate the efficacy of access management projects, which have included the use of manual traffic data collection exercises, self-reported surveys, simulation, and evaluation of crash data many years after the project has been implemented. As such, traditional data sources typically used for such evaluation are either impractically labor intensive, have a short span of data coverage, may suffer from stated preference biases, or are feasible only after a significant time has passed since the project was implemented.

The objective of this research was to evaluate the potential of using new-age data sources such as crowdsourced connected vehicle data, crowdsourced spending pattern data, and traditional sources for sales tax and crash data in quantifying the effects of access management projects on arterial roadways. Based on observations and analyses highlighted in this study, it can be concluded that crowdsourced data provide valuable information on travel patterns along access managed corridors. The same information, if sought through manual data collection, will be expensive, time consuming, and relatively unscalable. As discussed in detail using the example analysis of collected sales tax data, the available public information doesn't seem to be granular enough to derive conclusive insights on the impacts of access management. A couple of other observations as far as the crowdsourced connected vehicle (Wejo) data include the following:

- Using the same or a single vehicle manufacturer (OEM) for both the before and after periods can be comparatively more useful in providing insights on events like hard braking and evaluating the accompanying safety impacts.
- It was also observed that it is infeasible to associate movements with events using the current form of data. For example, it is hard to tell if a certain movement type has more or fewer hard braking events than another movement type.
- Similar to other large datasets, crowdsourced CV data require processing prior to use by engineers and planners.

Similarly, the Spend data from SafeGraph can be useful for comparing the same business or same type of business over time, but as discussed earlier in the report, there is no consistent trend that is applicable across all kinds of businesses on a corridor. The number of transactions or footfalls increased for some types of businesses, remained about the same for some, and decreased for a few—indicating a lack of consistent growth or decline in business activity. Worth noting is that access management is just one of the factors that can potentially affect business activity, footfalls, or sales. There are several other factors at play—many if not all of which are dynamic in nature—that can be impractical to capture with only a few variables of analysis.

However, having known these limitations, this evaluation of crowdsourced data such as Wejo revealed that different elements of this dataset can be useful for different sets of audience. For example, detailed information on vehicle journeys can be useful for businesses, property owners, and elected officials, while elements such as vehicle trajectories (paths), turning rates, and events like hard braking and acceleration can be useful information for engineers and planners.

Limitations of traditional data sources used for evaluating access management projects were also noted. Most notably, using aggregated sales tax data from the state comptroller's office can be challenging to derive insights from mainly attributable to noted factors, some of which can be remedied, while others not to the same extent. Externalities can creep in because the before and after time windows might have a different composition of businesses in the reported dataset, inconsistent or different reporting frequencies of different businesses, and local economic conditions or peculiarities that might have a much larger impact on business activity than just the changes in access management around a business, among others.

Overall, this project provides insights into traffic patterns and driving behavior using CV data. The data can be used to inform transportation planning and policy-making efforts, as well as to improve road safety and reduce traffic congestion. Additionally, this project demonstrates the power of advanced analytical techniques, such as geospatial analysis, in extracting insights from large datasets and applying that knowledge for engineering, planning, public outreach, and business community engagement activities.

The following are a few ideas for potential future evaluation using these datasets to enhance the scope or applicability of this research:

- Research is underway to investigate the correlation between hard braking events and safety or crash outcomes. Surrogate measures of safety can be useful in predicting performance and/or identifying possible remedies to improve roadway safety.
- Because crowdsourced data capture a sample of the overall traffic, a factor to scale up to determine the number of daily trips can be useful, particularly to road designers.
- Another possible evaluation is where one could look at hotspot identification using data for events such as hard braking and hard acceleration.
- Lastly, the correlation between turning movements and consumer spending data can be revealing in terms of traveler movements to and from businesses.

References

1. FHWA Office of Operations. (2021). What is Access Management? Federal Highway Administration. Accessed September 2022. https://ops.fhwa.dot.gov/access_mgmt/what_is_accsmgmt.htm
2. Bowman, B. L., Vecellio, R. L., & Maio, J. (1995). Vehicle and Pedestrian Accident Models for Median Locations. *Journal of Transportation Engineering, 121*(6).
3. Zegeer, C., Srinivasan, R., Lan, B., Carter, D., Smith, S., Sundstrom, C., Thirsk, N. J… Van Houten, R. (2017). *Development of Crash Modification Factors for Uncontrolled Pedestrian Crossing Treatments*, National Cooperative Highway Research Program, NCHRP Research Report 841. Washington, DC: The National Academies Press.
4. Gluck, J., Levinson, H. S., & Stover, V. G. (1999). *NCHRP Report 420: Impacts of Access Management Techniques*, National Cooperative Highway Research Program, Transportation Research Board. Washington, DC: The National Academies Press.
5. *Highway Capacity Manual*. (1994). Transportation Research Board. Washington, DC: The National Academies Press.
6. Stover, V. G., & Koepke, F. J. (2002). *Transportation and Land Development, 2nd Edition.* Washington, DC: Institute of Transportation Engineers.
7. Eisele, W. E., & Frawley, W. E. (1999). *A Methodology for Determining Economic Impacts of Raised Medians: Data Analysis on Additional Case Studies*, Research Report 3904-3, Texas Transportation Institute, College Station, Texas.
8. SafeGraph. Spend. Accessed March 2023. https://docs.safegraph.com/docs/spend

The National Institute for Congestion Reduction (NICR) will emerge as a national leader in providing multimodal congestion reduction strategies through real-world deployments that leverage advances in technology, big data science and innovative transportation options to optimize the efficiency and reliability of the transportation system for all users. Our efficient and effective delivery of an integrated research, education, workforce development and technology transfer program will be a model for the nation.

www.nicr.usf.edu