MEMORANDUM

To: City of Hendersonville, NC

From: Alta Planning + Design

Date: 02/11/2026

Re: Hendersonville, NC Downtown Streets Modernization Project BUILD Grant: Benefit-Cost Analysis Summary Memo

Benefit-Cost Analysis for Hendersonville, NC Downtown Streets Modernization Project BUILD Grant Application: All Components

Introduction

This Benefit-Cost Analysis (BCA) includes the benefits and costs for the proposed project that would be fully constructed if the BUILD grant is awarded. The analysis period was 23 years (3 years of planning, engineering and construction and 20 years of operation) and assumes a useful service life of 30 years for the project. All costs and benefits are presented in 2024 dollars.

Each component of the project was assessed in separate BCA following the principles documented in the USDOT Benefit-Cost Analysis Guidance for Discretionary Grant Programs as of December 2025.

The following categories of benefits were considered in the BCA:

- **Safety:** The expected reduction in collisions and associated costs.
- **Environmental Sustainability:** Includes reductions in the following pollutants that impact air quality, NO_X SO_2, and $PM_{2.5}$.
- **Quality of Life:** The expected reduction in mortality rates due to increased physical activity from new users of the project.
- **Economic Competitiveness:** Includes savings in household transportation costs and traffic congestion costs.
- **State of Good Repair:** Includes reductions in roadway maintenance costs.
- **Maintenance costs (dis-benefit):** Covers the ongoing costs of upkeep to the proposed project.

Results Summary

Table 1 displays the total benefits by category included in the BCAs. The capital costs included in all the BCAs are $12.6 million. This BCA estimates the project compared to the no-build scenario over a 23-year evaluation (2027-2049) and at a 7.0 percent real discount rate will have a net present value of **$4.3 million** and a benefit-cost ratio of **1.34:1.0.** This is summarized in Table 2.[1]

Table 1. Total Undiscounted Benefits over 20 years of Operation

CATEGORY	1. South Main Street Improvements	2. 7th Avenue Improvements	BUILD Grant All Components
	Monetary Value (2024 dollars)	Monetary Value (2024 dollars)	Monetary Value (2024 dollars)
Safety Benefits	$5,980,000	$28,600,000	$34,580,000
Environmental Sustainability	$5,900	$1,000	$6,900
Quality of Life	$7,758,000	$800,000	$8,558,000
Economic Competitiveness	$132,200	$13,600	$145,800
State of Good Repair	$20,600	$2,100	$22,700
Maintenance Costs (Dis-benefit)	($1,000,000)	($100,000)	($1,100,000)
Residual Value	$4,267,000	$3,405,000	$7,672,000
TOTAL BENEFITS (UNDISCOUNTED)	$16,461,000	$31,254,000	$47,715,000

[1] A 7.0% discount rate was used for all benefits and costs.

Table 2. Benefit-Cost Analysis Summary

CATEGORY	1. South Main Street Improvements	2. 7th Avenue Improvements	BUILD Grant All Components
	Monetary Value (2024 dollars)	Monetary Value (2024 dollars)	Monetary Value (2024 dollars)
Net Discounted Benefits	$5,434,000	$11,420,000	$16,854,000
Net Discounted Capital Costs	($8,172,000)	($4,430,000)	($12,602,000)
Net Present Value	($2,737,000)	$6,991,000	$4,254,000
Benefit - Cost Ratio	0.66:1	2.58:1	1.34:1

Table 3. Summarized GHG and Criteria Pollutants

METRIC TONS OF GHGS AND CRITERIA POLLUTANTS	1. South Main Street Improvements	2. 7th Avenue Improvements	BUILD Grant All Components
PARTICULATE MATTER 2.5 (PM 2.5) REDUCED	0.023	0.017	0.040
NITROUS OXIDES (NOx) REDUCED	2.361	1.790	4.151
SULFUR OXIDES (SOx) REDUCED	0.022	0.017	0.039

Additional details can be found in the individual BCA Memos and spreadsheets: South Main Street and 7th Avenue, included in the grant application package.

MEMORANDUM

To: City of Hendersonville, NC

From: Alta Planning + Design

Date: 02/11/2026

Re: BUILD Benefit-Cost Analysis Technical Memo: Component 1: South Main Street Complete Street Improvements

---

Benefit-Cost Analysis for Hendersonville, NC Downtown Streets Modernization Project BUILD Grant Application (Component 1: South Main Complete Street Improvements)

Executive Summary

This Benefit-Cost Analysis (BCA) includes the benefits and costs for the proposed project that would be fully constructed if the BUILD grant is awarded. The analysis period was 23 years (3 years of planning, engineering and construction and 20 years of operation) and assumes a useful service life of 30 years for the project. All costs and benefits are presented in 2024 dollars.

The following categories of benefits were considered in the BCA:

- **Safety:** The expected reduction in collisions and associated costs.
- **Environmental Sustainability:** Includes reductions in the following pollutants that impact air quality, NO_X SO_2, and $PM_{2.5}$.
- **Quality of Life:** The expected reduction in mortality rates due to increased physical activity from new users of the project.
- **Economic Competitiveness:** Includes savings in household transportation costs and traffic congestion costs.
- **State of Good Repair:** Includes reductions in roadway maintenance costs.
- **Maintenance costs (dis-benefit):** Covers the ongoing costs of upkeep to the proposed project.

Results Summary

Table 1 displays the total benefits by category included in the BCA. The capital costs included in the BCA are $10.7 million. This BCA estimates the project compared to the no-build scenario over a 23-year evaluation (2027-2049) and at a 7.0 percent real discount rate will have a net present value of **-$2,737,000** and a benefit-cost ratio of **0.66: 1.0.** This is summarized in Table 2.[1]

[1] A 7.0% discount rate was used for all benefits and costs.

Table 1:Total Undiscounted Benefits over 20 years of Operation

CATEGORY	MONETARY VALUE (In 2024 dollars)
Safety Benefits	$ 5,980,000
Environmental Sustainability	$ 5,900
Quality of Life	$ 7,758,000
Economic Competitiveness	$ 132,200
State of Good Repair	$ 20,600
Maintenance Costs (Dis-benefit)	$ (1,000,000)
Residual Value	$3,566,667
TOTAL BENEFITS (UNDISCOUNTED)	$ 16,461,000

Table 2: Benefit-Cost Analysis Summary

CATEGORY	DISCOUNTED [2] VALUE (in 2023 dollars)
Net Discounted Benefits	$ 5,434,000
Net Discounted Capital Costs	$ (8,172,000)
Net Present Value	$ (2,737,000)
Benefit - Cost Ratio	0.66:1

[2] A 7.0% discount rate was used for all benefits and costs.

Background

The benefit-cost analysis (BCA) for this project follows the principles documented in the USDOT Benefit-Cost Analysis Guidance for Discretionary Grant Programs as of December 2025 (hereafter referred to as “USDOT BCA Guidance”) and uses the recommended parameter values where applicable. The BCA includes the benefits and costs for the project that would be fully constructed if the BUILD grant is awarded. The analysis period was 23 years and assumes a useful service life of 30 years for the project. All costs and benefits are presented in 2024 base year dollars. Benefits and cost streams were discounted using a 7.0% per year discount rate. This memo contains a detailed explanation of the BCA methodology and the parameter values that were used.

Approach to Benefits and Study Area

This BCA approach expands on the methods suggested by the National Cooperative Highway Research Program (NCHRP) Report 552: Guidelines for Analysis of Investments in Bicycle Facilities by incorporating detailed local demographic information and using new data and research that has become available since Guidelines for Analysis was published in 2006.

While construction of the project will benefit all residents of and visitors to the region, those living within three miles (about a 15-minute bike ride) and one-half mile (about a 10-minute walk) of the project will have the most convenient access and will gain the most from its completion. Accordingly, this BCA focuses on the bicycling benefits attributed to residents living within three miles of the project and on the walking benefits attributed to residents living within one-half mile project. There are several benefit categories that benefit the region more widely (reduced roadway maintenance, healthcare costs), but these ranges are used to constrain this analysis to the main beneficiaries.

Benefits were primarily calculated by comparing walking and biking activity (including crashes) under the Baseline to a Build scenario in which the project has been implemented. The baseline and build scenarios encompass an identical geography (census block groups within three (3) miles of the project). **The benefits included in the Net Present Value and Benefit-Cost Ratio calculations are the net difference between the two scenarios.** The proposed improvements and expected benefits are summarized in Table 3.

Table 3: Summary Matrix

Baseline	Build Scenario	Type of Impacts
No current connection to nearby regional Ecusta Trail being installed south of Hendersonville, NC (fully funded, partially constructed)	Construction of street improvements on South Main Street to enhance connectivity of downtown Hendersonville, NC and through other portions of the BUILD project to near-term regional Ecusta Trail and improving connectivity throughout downtown.	Increase in walking within 0.5 miles of the study area and biking within 3 miles of the study area due to increased feelings of safety and comfort with addition of sidewalk, lighting, signal upgrades with leading pedestrian intervals, a separated bike lane/shared-use path, and crossing improvements on South Main Street. Reduced mortality benefits, reduced collisions, reduced roadway maintenance, reduced traffic congestion, and reduced household transportation costs within 3 miles of the study area.

Capital Costs

Refer to the main application for a detailed breakdown of project costs. The capital cost schedule is shown in Table 4**.** This schedule includes design, engineering, permitting, contracting and installation. The official cost estimate for this project was established with a construction sub-total of January 2026, so this was deflated to the baseline year of 2024 using the December Consumer Price Index of 2.6%.

Table 4: Project Construction Schedule and Cost

Design/Construction Year	Anticipated Cost
2027	$2,675,000
2028	$5,350,000
2029	$2,675,000
Total Capital Costs	$10,700,000

The estimated total annual maintenance costs are $50,000 per year (undiscounted) and they were included as a disbenefit in the benefit-cost ratio. The basis for this annual cost stems from conversations with the client around current maintenance activities along North Main Street.

Useful Life

The expected useful life of the street improvements is 30 years. The window of analysis used was 23 years. A residual value of $3,566,667 (undiscounted) was claimed as a benefit in the final year of the analysis period, assuming linear depreciation.

Demand

Next, the analysis estimated the expected number of biking and walking trips that would occur due to the street improvements. The primary inputs to the demand analysis were existing counts of pedestrians throughout several years in downtown Hendersonville, NC, see Table 5.

Table 5: Pedestrian Counts, City of Hendersonville, NC

Counter (Location)	Month	Year	Peak 24-Hr Daily Users
North Main Street between 3rd Ave and 4th Ave	May	2016	2,592
North Main Street at 5th, by Mast General Store	February	2022	1,960
Downtown Hendersonville, Main Street, between 5th Avenue and 6th Avenue, East side of the road	June	2016	2,047
South Main Street between 1st Street and Allen Street	June	2016	1,406

The pedestrian demand estimate stems from counts directly north of where the project will connect and improve conditions for pedestrians. The average peak 24-hour users for established Complete Street portions on North Main Street is 2,200 pedestrians. Further south, just one block north of the project corridor the peak 24-hour users was approximately 30% lower, 1,406 pedestrians.

On South Main Street, where the Festival Street project will be located, land uses and transportation facilities shift to be more auto-centric. The difference in visitation activity between complete street environments and existing auto-centric environments was approximately 30%, according to Placer.ai data from January 1, 2023 - September 1, 2024. This is similar to the difference in average pedestrian counts between complete street environments and auto-centric environments on North Main Street. Extending this pattern of activity linearly to the project on South Main Street, the project team estimates an existing baseline daily pedestrian activity of 980 users on South Main Street.

The project team assumed a 13% increase in pedestrian activity after implementation of the project based on review of recent estimates developed for the Caltrans Active Transportation Benefit-Cost Tool (Fitch et al. 2022)[3] and detailed results from the following research studies: Broach et al. (2012)[4], Broach and Dill (2015, 2016, 2017)[5], McNeil et al. (2015)[6], and Sevtsuk et al. (2021)[7].

Since there was no bike counter data available in the corridor, the project team estimated the number of existing and new daily bike trips along the corridor by extrapolating from the pedestrian counts. The team used the ratio of walking to bike commuting within 3 miles of the project – from American Community Survey (ACS) 2022 5-year primary commute mode data at the block group level – to estimate the number of existing bike trips at the project location (275 trips). Then, using this baseline estimate of existing bike trips, the team applied a percentage increase (31%) based on the degree of improvement in the experience of bicyclists along the corridor once the project is constructed (following a method developed for NCHRP 08-149[8]) to estimate the number of new bike trips on the corridor (85).

---

[3] Fitch, D.T., S. Kamalapuram, M. Favetti, and S.L. Handy. 2022. Caltrans Active Transportation Benefit-Cost Tool. Technical Documentation, Institute for Transportation Studies, University of California, Davis, Version 0.1.0. Last updated July 31, 2022. <
https://activetravelbenefits.ucdavis.edu/Caltrans%20ATP%20BC%20Tool%20Technical%20Documentation%20Final%20Draft%20(v2).pdf>.

[4] Broach, J., J. Dill, and J. Gliebe. 2012. "Where Do Cyclists Ride? A Route Choice Model Developed with Revealed Preference GPS Data." Transportation Research Part A: Policy and Practice 46(10), 1730–1740.

[5] Broach, J., and J. Dill. 2015. Pedestrian Route Choice Model Estimated from Revealed Preference GPS Data. In Transportation Research Board 94th Annual Meeting. Washington, DC.; Broach, J., and J. Dill. 2016. "Using Predicted Bicyclist and Pedestrian Route Choice to Enhance Mode Choice Models." Transportation Research Record 2564(1), 52–59. Broach, J., and J. Dill. 2017. Bridging the Gap: Using Network Connectivity and Quality Measures to Predict Bicycle Commuting. Paper presented at the 96th Annual Meeting of the Transportation Research Board.

[6] McNeil, N., C.M. Monsere, and J. Dill. 2015. "Influence of Bike Lane Buffer Types on Perceived Comfort and Safety of Bicyclists and Potential Bicyclists." Transportation Research Record 2520(1), 132–142.

[7] Sevtsuk, A., R. Basu, X. Li, and R. Kalvo. 2021. "A Big Data Approach to Understanding Pedestrian Route Choice Preferences: Evidence from San Francisco." Travel Behaviour and Society 25, 41–51. https://doi.org/10.1016/j.tbs.2021.05.010.

[8] Transportation Research Board. (2024). *Estimating benefits of closing gaps in active transportation networks (NCHRP 08-149)*. The National Academies of Sciences, Engineering, and Medicine. https://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=5086

Therefore, we expect 213 new daily users (85 bicyclists and 128 pedestrians) along South Main Street / Festival Street after implementation of the proposed improvements, as shown in Table 6.

Table 6. Demand Estimate

Project Name	Length (Mi)	Estimated Daily Average of Bike Trips	Estimated Daily Average of Pedestrian Trips	Average Daily Users
Festival Street	0.3	85	128	213

Benefits

The various benefits expected to result from implementation of the project are described in this section.

Walking and Biking Activity

The BCA estimated current levels of walking and biking within the project area using American Community Survey (ACS) 2022 5-year data. Table 7 displays the existing commute to work mode share for people within walking and biking distance of the proposed project. Population and demographic forecasts from Land of Sky Regional Council at the Transportation Analysis Zone (TAZ) level were used to estimate population growth in the study area over the analysis period. Actual population data from 2015 and forecasts for 2045 were collected and were interpolated for each intermediate year in the analysis.

Table 7: Means of Transportation to Work of People Living in the Study Area (2022 American Community Survey)

Corridor	Employed Population	Drove Alone	Carpool	Public Transit	Bicycled	Walked	Worked from Home	Other
Walkshed (within 0.5 miles)	4,613	3,190	775	48	0	95	459	46
Bikeshed (within 3 miles)	22,029	17,368	2,008	50	55	141	2,160	247

The means of transportation to work data was converted to daily estimates and extrapolated to annual trip volumes and broken into different trip types (i.e. commute, school, college, and utilitarian) using the existing travel patterns (Table 8) and data from the National Household Transportation Survey. The annual extrapolations account for the expected number of trips per week by trip type (i.e., commute, school, and college trips are expected five out of seven days a week, and other trip types are expected to occur seven days a week).

Table 8: Trip Purpose Multiplier[9]

	Bike	Walk
Utilitarian Trip Multiplier	5.33	8.77

[9] Travel Day Person Trips (in millions), NHTSA 2017 https://nhts.ornl.gov/

Increase in Walking and Biking Activity

The Baseline scenario assumes that the walking and biking mode share will remain constant and that trips will increase annually with expected population growth. In the Build scenario, the demand estimates for the proposed project were added to the existing walking and biking activity starting in 2030 (the expected opening year). The demand estimates were escalated by the expected population growth factor each year.

Decrease in Motor Vehicle Trips

Some of the estimated annual bicycle and pedestrian trips within the proposed project area are expected to replace motor vehicle trips. Calibrated to modal shift factors reported in literature[10], a univariate regression model estimates the motor vehicle trip replacement factor based on the percentage of trips less than four miles that terminate in census block groups within three miles (approximate bicycling distance) of the proposed facility. Trip distance data is provided by Replica for a typical travel in the Hendersonville region on a Thursday in Spring 2024[11]. More details on Replica are included in Appendix A. The motor vehicle trip replacement factor for the proposed project is 0.09. Additional details on the methodology are included in Appendix B.

To estimate the number of vehicle-miles that might be replaced by bicycling and walking trips, Table 9 shows the average trip distance of bicycling and walking trips by trip purpose. The number of vehicle miles reduced due to bicycle and pedestrian trips was calculated by multiplying the number of biking or walking trips by the trip replacement and trip distance factors. The estimated reduction in vehicle miles traveled is shown in Table 16.

Table 9: Average Trip Distance (miles)

	Bike	Walk
Commute Trips[12]	2.47	0.72
College Trips[13]	1.31	0.43
K-12 School Trips[14]	1.36	0.69
Utilitarian Trips[15]	2.28	0.83

[10] Volker et al (2019). Quantifying Reductions in Vehicle Miles Traveled from New Bike Paths, Lanes, and Cycle Tracks
[11] Replica Places (2019). https://replicahq.com/
[12] Federal Highway Administration. (2009). 2009 National Household Travel Survey, U.S. Department of Transportation, Washington, DC. Available online: https://nhts.ornl.gov.
[13] Ibid.
[14] Safe Routes National Center for Safe Routes to School, Trends in Walking and Bicycling to School from 2007 to 2013 (2015).
[15] Federal Highway Administration. (2009). 2009 National Household Travel Survey, U.S. Department of Transportation, Washington, DC. Available online: https://nhts.ornl.gov.

Environmental Sustainability Benefits

For every vehicle-mile reduced, there is an assumed decrease in greenhouse gases and criteria pollutants. Table 10 lists the reduction in greenhouse gases and criteria pollutants by vehicle-mile traveled. The cost to mitigate or clean-up those pollutants was calculated using the monetary values provided by the USDOT BCA Guidance Table A-6. Emission types not listed in that table were not included in the analysis. The estimated annual emission reduction benefits are shown in Table 17.

Table 10: Environmental Sustainability Multipliers

Pollutant	Value (metric tons/VMT)
Particulate Matter 2.5 ($PM_{2.5}$) [16]	0.000000008
Nitrous Oxides (NOx) [17]	0.0000008
Sulfur Oxides (SO_2) [18]	0.000000008

Quality of Life Benefits

More people bicycling and walking can help encourage an increase in physical activity levels, increased cardiovascular health, and other positive outcomes for users. The benefits from reduced mortality were calculated using the recommended values provided in the USDOT BCA Guidance (Table A-13) and the national distribution of age ranges and travel patterns. These benefits were only applied to the estimated number of walking and biking trips induced by the project (see Demand section). Table 11 displays the multipliers that were used. The estimated annual mortality benefits are shown in Table 18.

Table 11: Mortality Reduction Multipliers

Mortality Reduction Benefits of Induced Active Transportation	Value
Walking Value per Induced Trip	$8.36

---

[16] The Safer Affordable Fuel-Efficient Vehicles Rule for MY2021-MY2026 Passenger Cars, BUILD Guidance 2020, Table A-7 and Light Trucks Preliminary Regulatory Impact Analysis (October 2018) https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/ld_cafe_co2_nhtsa_2127-al76_epa_pria_181016.pdf

[17] The Safer Affordable Fuel-Efficient Vehicles Rule for MY2021-MY2026 Passenger Cars, BUILD Guidance 2020, Table A-7 and Light Trucks Preliminary Regulatory Impact Analysis (October 2018) https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/ld_cafe_co2_nhtsa_2127-al76_epa_pria_181016.pdf

[18] The Safer Affordable Fuel-Efficient Vehicles Rule for MY2021-MY2026 Passenger Cars, BUILD Guidance 2020, Table A-7 and Light Trucks Preliminary Regulatory Impact Analysis (October 2018) https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/ld_cafe_co2_nhtsa_2127-al76_epa_pria_181016.pdf

Cycling Value per Induced Trips	$7.45
Walking Age Proportion (20-74 years old)	68%
Cycling Age Proportion (20-64 years old)	59%
Trips induced from non-active modes	89%

Economic Competitiveness Benefits

For every vehicle-mile reduced, there is a reduction in household transportation costs and congestion costs. Table 12 displays the multipliers use to calculate economic competitiveness benefit. The estimated annual economic competitiveness benefits are shown in Table 19.

Table 12: Economic Competitiveness Multipliers

Type of Savings	Value
Household Transportation Cost Savings	$0.54 per VMT[19]
Congestion Cost Savings	$0.08 per VMT[20,21]

[19] Our Driving Costs, AAA (2016).
[20] Crashes vs. Congestion: What's the Cost to Society? AAA (2011). https://exchange.aaa.com/wp-content/uploads/2012/07/AAA-Crashes-vs-Congestion-2011.pdf
[21] Crashes vs. Congestion: What's the Cost to Society? AAA (2011). https://exchange.aaa.com/wp-content/uploads/2012/07/AAA-Crashes-vs-Congestion-2011.pdf

Safety Benefits

The proposed project would decrease conflicts between people walking and biking with motor vehicles. Collision data covering a nearly seven-year period (January 1, 2019-November 30, 2025) was provided by the North Carolina Department of Transportation. Collisions under consideration were located along the project corridor and included all types of collisions due to the various benefits of the proposed improvements (Table 13).

A total of 36 crashes were included in the analysis, including one fatality, a pedestrian who was struck while crossing at an intersection. Different Crash Reduction Factors were applied to the selected crashes for each segment of the project and the benefits were monetized using the values provided in the USDOT BCA Guidance Appendix A, Table A-1 for KABCO Level data.

Each crash was assigned to a single countermeasure. The countermeasures selected were:

- Corridor-specific traffic calming (Caswell St to Allen St) for injury crashes (CM ID 587, CRF 0.18); non-injury crashes (CM ID 590, CRF 0.06)
- Leading pedestrian interval (intersections with Caswell St and Barnwell St) for one pedestrian crash (CM ID 8893, CRF 0.13); injury crashes (CM ID 8892, CRF 0.14)
- Upgrading signal heads from 8-inch to 12-inch signal heads (intersections with Caswell St and Barnwell St) for angle crashes (CM ID 10117, CRF 0.42); non-angle crashes (CM ID 4776, CRF 0.03)
- Pedestrian countdown signals (intersections with Caswell St and Barnwell St) for rear-end crashes (CM ID 10117, CRF 0.13)

The estimated annual safety benefits are shown in Table 20.

Table 13: Summary of Collisions in project corridor, January 1, 2019-November 30, 2025, NCDOT

CMF	Total Collisions	Fatal (K)	Incapacitating (A)	Serious (B)	Possible (C)	PDO (O)	Injured – Unknown Severity (U)
Corridor-specific Traffic Calming - Injury	2	0	0	0	2	0	0
Corridor-specific Traffic Calming - No Injury	4	0	0	0	0	4	0
Leading Pedestrian Interval - Pedestrian	1	1	0	0	0	0	0
Leading Pedestrian Interval - Injury	1	0	0	0	0	0	1
Upgrading Signal Heads - Angle	17	0	0	0	3	14	0
Upgrading Signal Heads - Non-angle	4	0	0	0	0	4	0
Pedestrian Countdown Signals	7	0	0	0	1	6	0
Total	36	1	0	0	6	28	1

State-of-good Repair Benefits

Table 14 shows the estimated roadway maintenance cost savings associated with a reduction in vehicle-miles traveled.

Table 14: State of Good Repair Multiplier

	Value (metric tons/VMT)
Roadway Maintenance Cost Savings	$0.10 per VMT[22]

[22] Kitamura, R., Zhao, H., and Gubby, A. R. Development of a Pavement Maintenance Cost Allocation Model. Institute of Transportation Studies, University of California, Davis. https://trid.trb.org/view.aspx?id=261768

Results

Table 15 through Table 24 display the results of the benefit-cost analysis for each year of the analysis period. This BCA estimates the project compared to the no-build scenario over a 23-year evaluation (2027-2049) and at a 7.0 percent real discount rate will have a net present value of **-$2,737,000** and a benefit-cost ratio of **0.66: 1.0.**

Table 15: Estimated Annual Bicycle and Walk Trips

Year	Baseline	Build Scenario	Additional Trips
2027	923,700	923,700	-
2028	924,200	924,200	-
2029	924,600	924,600	-
2030	925,100	998,400	73,300
2031	925,500	999,900	74,400
2032	926,000	1,001,500	75,500
2033	926,400	1,003,000	76,600
2034	926,900	1,004,500	77,600
2035	927,400	1,006,100	78,700
2036	927,800	1,007,700	79,900
2037	928,300	1,009,300	81,000
2038	928,700	1,010,900	82,200
2039	929,200	1,012,500	83,300
2040	929,600	1,014,100	84,500
2041	930,100	1,015,800	85,700
2042	930,500	1,017,500	87,000
2043	931,000	1,019,100	88,100
2044	931,500	1,020,900	89,400
2045	931,900	1,022,600	90,700
2046	932,400	1,024,300	91,900
2047	932,800	1,026,100	93,300
2048	933,300	1,027,900	94,600
2049	933,700	1,029,700	96,000
Total Additional Trips:		1,683,700	

Table 16: Estimated Annual Vehicle Miles Reduced

Year	Baseline	Build Scenario	Additional Vehicle Miles Reduced
2027	101,000	101,000	-
2028	101,100	101,100	-
2029	101,100	101,100	-
2030	101,200	110,400	9,200
2031	101,200	110,600	9,400
2032	101,300	110,700	9,400
2033	101,300	110,900	9,600
2034	101,400	111,100	9,700
2035	101,400	111,300	9,900
2036	101,500	111,500	10,000
2037	101,500	111,700	10,200
2038	101,600	111,900	10,300
2039	101,600	112,100	10,500
2040	101,700	112,300	10,600
2041	101,700	112,500	10,800
2042	101,800	112,700	10,900
2043	101,800	112,900	11,100
2044	101,900	113,100	11,200
2045	101,900	113,300	11,400
2046	102,000	113,500	11,500
2047	102,000	113,700	11,700
2048	102,100	113,900	11,800
2049	102,100	114,100	12,000
Total Additional Vehicle Miles Reduced:		211,200	

Table 17: Estimated Annual Environmental Sustainability Benefits (Undiscounted)

Year	Baseline	Build Scenario	Benefits
2027	$-	$-	$-
2028	$-	$-	$-
2029	$-	$-	$-
2030	$3,000	$3,200	$200
2031	$3,000	$3,200	$200
2032	$3,000	$3,300	$300
2033	$3,000	$3,300	$300
2034	$3,000	$3,300	$300
2035	$3,000	$3,300	$300
2036	$3,000	$3,300	$300
2037	$3,000	$3,300	$300
2038	$3,000	$3,300	$300
2039	$3,000	$3,300	$300
2040	$3,000	$3,300	$300
2041	$3,000	$3,300	$300
2042	$3,000	$3,300	$300
2043	$3,000	$3,300	$300
2044	$3,000	$3,300	$300
2045	$3,000	$3,300	$300
2046	$3,000	$3,300	$300
2047	$3,000	$3,300	$300
2048	$3,000	$3,300	$300
2049	$3,000	$3,400	$400
Total Benefits:			$5,900

Table 18: Estimated Annual Quality of Life Benefits (Undiscounted)

Year	Baseline	Build Scenario	Benefits
2027	$-	$-	$-
2028	$-	$-	$-
2029	$-	$-	$-
2030	$4,392,000	$4,730,000	$338,000
2031	$4,394,000	$4,737,000	$343,000
2032	$4,396,000	$4,744,000	$348,000
2033	$4,398,000	$4,751,000	$353,000
2034	$4,400,000	$4,758,000	$358,000
2035	$4,403,000	$4,765,000	$362,000
2036	$4,405,000	$4,773,000	$368,000
2037	$4,407,000	$4,780,000	$373,000
2038	$4,409,000	$4,788,000	$379,000
2039	$4,411,000	$4,795,000	$384,000
2040	$4,413,000	$4,803,000	$390,000
2041	$4,416,000	$4,811,000	$395,000
2042	$4,418,000	$4,818,000	$400,000
2043	$4,420,000	$4,826,000	$406,000
2044	$4,422,000	$4,834,000	$412,000
2045	$4,424,000	$4,842,000	$418,000
2046	$4,426,000	$4,850,000	$424,000
2047	$4,429,000	$4,858,000	$429,000
2048	$4,431,000	$4,867,000	$436,000
2049	$4,433,000	$4,875,000	$442,000
Total Benefits:			$7,758,000

Table 19: Estimated Annual Economic Competitiveness Benefits (Undiscounted)

Year	Baseline	Build Scenario	Benefits
2027	$-	$-	$-
2028	$-	$-	$-
2029	$-	$-	$-
2030	$63,500	$69,300	$5,800
2031	$63,600	$69,400	$5,800
2032	$63,600	$69,500	$5,900
2033	$63,600	$69,600	$6,000
2034	$63,700	$69,800	$6,100
2035	$63,700	$69,900	$6,200
2036	$63,700	$70,000	$6,300
2037	$63,800	$70,100	$6,300
2038	$63,800	$70,200	$6,400
2039	$63,800	$70,400	$6,600
2040	$63,900	$70,500	$6,600
2041	$63,900	$70,600	$6,700
2042	$63,900	$70,700	$6,800
2043	$63,900	$70,900	$7,000
2044	$64,000	$71,000	$7,000
2045	$64,000	$71,100	$7,100
2046	$64,000	$71,300	$7,300
2047	$64,100	$71,400	$7,300
2048	$64,100	$71,500	$7,400
2049	$64,100	$71,700	$7,600
Total Benefits:			$132,200

Table 20: Estimated Annual Safety Benefits (Undiscounted)

Year	Baseline	Build Scenario	Benefits
2027	$-	$-	$-
2028	$-	$-	$-
2029	$-	$-	$-
2030	$-	$299,000	$299,000
2031	$-	$299,000	$299,000
2032	$-	$299,000	$299,000
2033	$-	$299,000	$299,000
2034	$-	$299,000	$299,000
2035	$-	$299,000	$299,000
2036	$-	$299,000	$299,000
2037	$-	$299,000	$299,000
2038	$-	$299,000	$299,000
2039	$-	$299,000	$299,000
2040	$-	$299,000	$299,000
2041	$-	$299,000	$299,000
2042	$-	$299,000	$299,000
2043	$-	$299,000	$299,000
2044	$-	$299,000	$299,000
2045	$-	$299,000	$299,000
2046	$-	$299,000	$299,000
2047	$-	$299,000	$299,000
2048	$-	$299,000	$299,000
2049	$-	$299,000	$299,000
Total Benefits:			$5,980,000

Table 21: Estimated Annual State of Good Repair Benefits (Undiscounted)

Year	Baseline	Build Scenario	Benefits
2027	$-	$-	$-
2028	$-	$-	$-
2029	$-	$-	$-
2030	$9,900	$10,800	$900
2031	$9,900	$10,800	$900
2032	$9,900	$10,800	$900
2033	$9,900	$10,900	$1,000
2034	$9,900	$10,900	$1,000
2035	$9,900	$10,900	$1,000
2036	$9,900	$10,900	$1,000
2037	$9,900	$10,900	$1,000
2038	$9,900	$10,900	$1,000
2039	$9,900	$11,000	$1,100
2040	$9,900	$11,000	$1,100
2041	$10,000	$11,000	$1,000
2042	$10,000	$11,000	$1,000
2043	$10,000	$11,000	$1,000
2044	$10,000	$11,100	$1,100
2045	$10,000	$11,100	$1,100
2046	$10,000	$11,100	$1,100
2047	$10,000	$11,100	$1,100
2048	$10,000	$11,100	$1,100
2049	$10,000	$11,200	$1,200
Total Benefits:			$20,600

Table 22: Estimated Annual Maintenance Disbenefits (Undiscounted)

Year	Baseline	Build Scenario	Benefits
2027			
2028			
2029			
2030	$-	$(50,000)	$(50,000)
2031	$-	$(50,000)	$(50,000)
2032	$-	$(50,000)	$(50,000)
2033	$-	$(50,000)	$(50,000)
2034	$-	$(50,000)	$(50,000)
2035	$-	$(50,000)	$(50,000)
2036	$-	$(50,000)	$(50,000)
2037	$-	$(50,000)	$(50,000)
2038	$-	$(50,000)	$(50,000)
2039	$-	$(50,000)	$(50,000)
2040	$-	$(50,000)	$(50,000)
2041	$-	$(50,000)	$(50,000)
2042	$-	$(50,000)	$(50,000)
2043	$-	$(50,000)	$(50,000)
2044	$-	$(50,000)	$(50,000)
2045	$-	$(50,000)	$(50,000)
2046	$-	$(50,000)	$(50,000)
2047	$-	$(50,000)	$(50,000)
2048	$-	$(50,000)	$(50,000)
2049	$-	$(50,000)	$(50,000)
Total Benefits:			$(1,000,000)

Table 23: Estimated Annual Benefits (Undiscounted)

Year	Baseline	Build Scenario	Benefits
2027	$-	$-	$-
2028	$-	$-	$-
2029	$-	$-	$-
2030	$4,468,000	$5,062,000	$594,000
2031	$4,470,000	$5,069,000	$599,000
2032	$4,473,000	$5,077,000	$604,000
2033	$4,475,000	$5,084,000	$609,000
2034	$4,477,000	$5,091,000	$614,000
2035	$4,479,000	$5,098,000	$619,000
2036	$4,481,000	$5,105,000	$624,000
2037	$4,484,000	$5,114,000	$630,000
2038	$4,486,000	$5,121,000	$635,000
2039	$4,488,000	$5,129,000	$641,000
2040	$4,490,000	$5,136,000	$646,000
2041	$4,492,000	$5,144,000	$652,000
2042	$4,495,000	$5,153,000	$658,000
2043	$4,497,000	$5,160,000	$663,000
2044	$4,499,000	$5,168,000	$669,000
2045	$4,501,000	$5,176,000	$675,000
2046	$4,503,000	$5,184,000	$681,000
2047	$4,506,000	$5,193,000	$687,000
2048	$4,508,000	$5,202,000	$694,000
2049	$4,510,000	$8,777,000	$4,267,000
Total Benefits:			$16,461,000

Table 24: Estimated Discounted Net Costs and Benefits

Year	Discounted Costs	Discounted Benefits	Net Cumulative Discounted Costs and Benefits
2027	$(2,184,000)	$-	$(2,184,000)
2028	$(4,081,000)	$-	$(6,265,000)
2029	$(1,907,000)	$-	$(8,172,000)
2030	$-	$396,000	$(7,777,000)
2031	$-	$373,000	$(7,404,000)
2032	$-	$351,000	$(7,052,000)
2033	$-	$331,000	$(6,721,000)
2034	$-	$312,000	$(6,409,000)
2035	$-	$294,000	$(6,115,000)
2036	$-	$277,000	$(5,838,000)
2037	$-	$261,000	$(5,576,000)
2038	$-	$246,000	$(5,330,000)
2039	$-	$232,000	$(5,098,000)
2040	$-	$219,000	$(4,879,000)
2041	$-	$206,000	$(4,673,000)
2042	$-	$195,000	$(4,478,000)
2043	$-	$183,000	$(4,295,000)
2044	$-	$173,000	$(4,122,000)
2045	$-	$163,000	$(3,959,000)
2046	$-	$154,000	$(3,805,000)
2047	$-	$145,000	$(3,660,000)
2048	$-	$137,000	$(3,523,000)
2049	$-	$786,000	$(2,737,000)
Total Net Discounted Costs: $ 8,172,000		Total Discounted Net Benefits: $5,434,000	Net Present Value: $(2,737,000)
Benefit-Cost Ratio: 0.66:1			

MEMORANDUM

rTo: City of Hendersonville, NC

From: Alta Planning + Design

Date: 02/11/2026

Re: BUILD Benefit-Cost Analysis Technical Memo: Component 2: 7th Avenue Complete Street Improvements

---

Benefit-Cost Analysis for Hendersonville, NC Downtown Streets Modernization Project BUILD Grant Application: Component 2: 7th Avenue Complete Street Improvements

Executive Summary

This Benefit-Cost Analysis (BCA) includes the benefits and costs for the proposed project that would be fully constructed if the BUILD grant is awarded. The analysis period was 23 years (3 years of planning, engineering and construction and 20 years of operation) and assumes a useful service life of 30 years for the project. All costs and benefits are presented in 2024 dollars.

The following categories of benefits were considered in the BCA:

- **Safety:** The expected reduction in collisions and associated costs.
- **Environmental Sustainability:** Includes reductions in the following pollutants that impact air quality, NO_X SO_2, and $PM_{2.5}$.
- **Quality of Life:** The expected reduction in mortality rates due to increased physical activity from new users of the project.
- **Economic Competitiveness:** Includes savings in household transportation costs and traffic congestion costs.
- **State of Good Repair:** Includes reductions in roadway maintenance costs.
- **Maintenance costs (dis-benefit):** Covers the ongoing costs of upkeep to the proposed project.

Results Summary

Table 1 displays the total benefits by category included in the BCA. The capital costs included in the BCA are $5.8 million. This BCA estimates the project compared to the no-build scenario over a 23-year evaluation (2027-2049) and at a 7.0 percent real discount rate will have a net present value of **$7 million** and a benefit-cost ratio of **2.58:1.0.** This is summarized in Table 2.[1]

[1] A 7.0% discount rate was used for all benefits and costs.

Table 1:Total Undiscounted Benefits over 20 years of Operation

CATEGORY	MONETARY VALUE (In 2024 dollars)
Safety Benefits	$ 28,600,000
Environmental Sustainability	$ 1,000
Quality of Life	$ 800,000
Economic Competitiveness	$ 13,600
State of Good Repair	$ 2,100
Maintenance Costs (Dis-benefit)	$ (100,000)
Residual Value	$ 1,933,333
TOTAL BENEFITS (UNDISCOUNTED)	$ 31,254,000

Table 2: Benefit-Cost Analysis Summary

CATEGORY	DISCOUNTED [2] VALUE (in 2023 dollars)
Net Discounted Benefits	$ 11,420,000
Net Discounted Capital Costs	$ (4,430,000)
Net Present Value	$ 6,991,000
Benefit - Cost Ratio	2.58:1

[2] A 7.0% discount rate was used for all benefits and costs.

Background

The benefit-cost analysis (BCA) for this project follows the principles documented in the USDOT Benefit-Cost Analysis Guidance for Discretionary Grant Programs as of December 2025 (hereafter referred to as "USDOT BCA Guidance") and uses the recommended parameter values where applicable. The BCA includes the benefits and costs for the project that would be fully constructed if the BUILD grant is awarded. The analysis period was 23 years and assumes a useful service life of 30 years for the project. All costs and benefits are presented in 2024 base year dollars. Benefits and cost streams were discounted using a 7.0% per year discount rate. This memo contains a detailed explanation of the BCA methodology and the parameter values that were used.

Approach to Benefits and Study Area

This BCA approach expands on the methods suggested by the National Cooperative Highway Research Program (NCHRP) Report 552: Guidelines for Analysis of Investments in Bicycle Facilities by incorporating detailed local demographic information and using new data and research that has become available since Guidelines for Analysis was published in 2006.

While construction of the project will benefit all residents of and visitors to the region, those living within three miles (about a 15-minute bike ride) and one-half mile (about a 10-minute walk) of the project will have the most convenient access and will gain the most from its completion. Accordingly, this BCA focuses on the bicycling benefits attributed to residents living within three miles of the project and on the walking benefits attributed to residents living within one-half mile project. There are several benefit categories that benefit the region more widely (reduced roadway maintenance, healthcare costs), but these ranges are used to constrain this analysis to the main beneficiaries.

Benefits were primarily calculated by comparing walking and biking activity (including crashes) under the Baseline to a Build scenario in which the project has been implemented. The baseline and build scenarios encompass an identical geography (census block groups within three (3) miles of the project). **The benefits included in the Net Present Value and Benefit-Cost Ratio calculations are the net difference between the two scenarios.** The proposed improvements and expected benefits are summarized in Table 3.

Table 3: Summary Matrix

Baseline	Build Scenario	Type of Impacts
No current connection to nearby regional Ecusta Trail being installed south of Hendersonville, NC (fully funded, partially constructed)	Construction of street improvements on 7th Avenue to enhance connectivity of the existing Oklawaha Greenway to downtown Hendersonville, NC and through other portions of the BUILD project to near-term regional Ecusta Trail.	Increase in walking within 0.5 miles of the study area and biking within 3 miles of the study area due to increased feelings of safety and comfort with addition of sidewalk, lighting, curb extensions and crossing improvements on 7th Avenue. Reduced mortality benefits, reduced collisions, reduced roadway maintenance, reduced traffic congestion, and reduced household transportation costs within 3 miles of the study area.

Capital Costs

Refer to the main application for a detailed breakdown of project costs. The capital cost schedule is shown in Table 4**.** This schedule includes design, engineering, permitting, contracting and installation. The official cost estimate for this project was established with a construction sub-total of 2024, so this was deflated to the baseline year of 2023 using the November 2024 Consumer Price Index of 2.7%.

Table 4: Project Construction Schedule and Cost

Design/Construction Year	Anticipated Cost
2027	$1,450,000
2028	$2,900,000
2029	$1,450,000
Total Capital Costs	$5,800,000

The estimated total annual maintenance costs are $5,000 per year (undiscounted) and they were included as a disbenefit in the benefit-cost ratio. The basis for this annual cost stems from conversations with the client around current maintenance activities along the roadway.

Useful Life

The expected useful life of the street improvements is 30 years. The window of analysis used was 23 years. A residual value of $1,933,333 (undiscounted) was claimed as a benefit in the final year of the analysis period, assuming linear depreciation.

Demand

Next, the analysis estimated the expected number of biking and walking trips that would occur due to the street improvements. The primary inputs to the demand analysis were existing counts of pedestrians at a location on the existing roadway, see Table 5.

Table 5: Pedestrians on 7th Avenue, City of Hendersonville, NC

Trail Counter (Location)	Month	Year	Peak 24-Hr Daily Users
North of 7th Avenue East, near the parking lot, on a signpost, on the eastern edge of William H King Memorial Park.	August	2024	317

This demand estimate stems from counts along 7th Avenue near where the project will connect and improve. According to this count, 317 trips on the nearby greenway in a 24-hour period. Therefore, we assume a similar baseline for the project corridor of 317 pedestrian trips each day. Examining current land uses, and based on knowledge of the study area, the project team assumed a baseline bicycle estimate of 10% of the estimated baseline number of pedestrian trips, or 32 bicycle trips per day.

Our team is a aware of recent estimates developed for the Caltrans Active Transportation Benefit-Cost Tool (Fitch et al. 2022)[3] and detailed results from the following research studies: Broach et al. (2012)[4], Broach and Dill (2015, 2016, 2017)[5], McNeil et al. (2015)[6], and Sevtsuk et al. (2021)[7]; all of which indicate that pedestrian use could be expected to increase by 4% percent and bicycle use by 29% after implementation of the proposed improvements. Therefore, we expect 22 new daily trips along 7th Avenue in Hendersonville, NC: 9 bicycle trips (32*0.29 = 9) and 13 pedestrian trips (317*0.04 = 13) as shown in Table 6.

Table 6. Demand Estimate

Project Name	Length (Mi)	Estimated Daily Average of Bike Trips	Estimated Daily Average of Pedestrian Trips	Average Daily Users
7th Avenue Streetscape Improvements	0.4	9	13	22
Total Estimate: 22 daily additional users				

[3] Fitch, D.T., S. Kamalapuram, M. Favetti, and S.L. Handy. 2022. Caltrans Active Transportation Benefit-Cost Tool. Technical Documentation, Institute for Transportation Studies, University of California, Davis, Version 0.1.0. Last updated July 31, 2022. <
https://activetravelbenefits.ucdavis.edu/Caltrans%20ATP%20BC%20Tool%20Technical%20Documentation%20Final%20Draft%20(v2).pdf>.

[4] Broach, J., J. Dill, and J. Gliebe. 2012. "Where Do Cyclists Ride? A Route Choice Model Developed with Revealed Preference GPS Data." Transportation Research Part A: Policy and Practice 46(10), 1730–1740.

[5] Broach, J., and J. Dill. 2015. Pedestrian Route Choice Model Estimated from Revealed Preference GPS Data. In Transportation Research Board 94th Annual Meeting. Washington, DC.; Broach, J., and J. Dill. 2016. "Using Predicted Bicyclist and Pedestrian Route Choice to Enhance Mode Choice Models." Transportation Research Record 2564(1), 52–59. Broach, J., and J. Dill. 2017. Bridging the Gap: Using Network Connectivity and Quality Measures to Predict Bicycle Commuting. Paper presented at the 96th Annual Meeting of the Transportation Research Board.

[6] McNeil, N., C.M. Monsere, and J. Dill. 2015. "Influence of Bike Lane Buffer Types on Perceived Comfort and Safety of Bicyclists and Potential Bicyclists." Transportation Research Record 2520(1), 132–142.

[7] Sevtsuk, A., R. Basu, X. Li, and R. Kalvo. 2021. "A Big Data Approach to Understanding Pedestrian Route Choice Preferences: Evidence from San Francisco." Travel Behaviour and Society 25, 41–51. https://doi.org/10.1016/j.tbs.2021.05.010.

Benefits

The various benefits expected to result from implementation of the project are described in this section.

Walking and Biking Activity

The BCA estimated current levels of walking and biking within the project area using American Community Survey (ACS) 2022 5-year data. Table 7 displays the existing commute to work mode share for people within walking and biking distance of the proposed project. Population and demographic forecasts from Land of Sky Regional Council at the Transportation Analysis Zone (TAZ) level were used to estimate population growth in the study area over the analysis period. Actual population data from 2015 and forecasts for 2045 were collected and were interpolated for each intermediate year in the analysis.

Table 7: Means of Transportation to Work of People Living in the Study Area (2021 American Community Survey)

Corridor	Employed Population	Drove Alone	Carpool	Public Transit	Bicycled	Walked	Worked from Home	Other
Walkshed (within 0.5 mile)	3,618	2,708	432	48	0	58	347	25
Bikeshed (within 3 miles)	23,446	18,405	2,272	50	55	225	1,965	472

The means of transportation to work data was converted to daily estimates and extrapolated to annual trip volumes and broken into different trip types (i.e. commute, school, college, and utilitarian) using the existing travel patterns (Table 8) and data from the National Household Transportation Survey. The annual extrapolations account for the expected number of trips per week by trip type (i.e., commute, school, and college trips are expected five out of seven days a week, and other trip types are expected to occur seven days a week).

Table 8: Trip Purpose Multiplier[8]

	Bike	Walk
Utilitarian Trip Multiplier	5.33	8.77

[8] Travel Day Person Trips (in millions), NHTSA 2017 https://nhts.ornl.gov/

Increase in Walking and Biking Activity

The Baseline scenario assumes that the walking and biking mode share will remain constant and that trips will increase annually with expected population growth. In the Build scenario, the demand estimates for the proposed project were added to the existing walking and biking activity starting in 2029 (the expected opening year). The demand estimates were escalated by the expected population growth factor each year.

Decrease in Motor Vehicle Trips

Some of the estimated annual bicycle and pedestrian trips within the proposed project area are expected to replace motor vehicle trips. Calibrated to modal shift factors reported in literature[9], a univariate regression model estimates the motor vehicle trip replacement factor based on the percentage of trips less than four miles that terminate in census block groups within three miles (approximate bicycling distance) of the proposed facility. Trip distance data is provided by Replica for a typical travel in the Hendersonville region on a Thursday in Spring 2024[10]. More details on Replica are included in Appendix A. The motor vehicle trip replacement factor for the proposed project is 0.09. Additional details on the methodology are included in Appendix B.

To estimate the number of vehicle-miles that might be replaced by bicycling and walking trips, Table 9 shows the average trip distance of bicycling and walking trips by trip purpose. The number of vehicle miles reduced due to bicycle and pedestrian trips was calculated by multiplying the number of biking or walking trips by the trip replacement and trip distance factors. The estimated reduction in vehicle miles traveled is shown in

[9] Volker et al (2019). Quantifying Reductions in Vehicle Miles Traveled from New Bike Paths, Lanes, and Cycle Tracks
[10] Replica Places (2019). https://replicahq.com/

Table 16.

Table 9: Average Trip Distance (miles)

	Bike	Walk
Commute Trips[11]	2.47	0.72
College Trips[12]	1.31	0.43
K-12 School Trips[13]	1.36	0.69
Utilitarian Trips[14]	2.28	0.83

Environmental Sustainability Benefits

For every vehicle-mile reduced, there is an assumed decrease in greenhouse gases and criteria pollutants. Table 10 lists the reduction in greenhouse gases and criteria pollutants by vehicle-mile traveled. The cost to mitigate or clean-up those pollutants was calculated using the monetary values provided by the USDOT BCA Guidance Table A-6. Emission types not listed in that table were not included in the analysis. The estimated annual emission reduction benefits are shown in Table 17.

Table 10: Environmental Sustainability Multipliers

Pollutant	Value (metric tons/VMT)
Particulate Matter 2.5 ($PM_{2.5}$) [15]	0.000000008
Nitrous Oxides (NOx) [16]	0.0000008

[11] Federal Highway Administration. (2009). 2009 National Household Travel Survey, U.S. Department of Transportation, Washington, DC. Available online: https://nhts.ornl.gov.
[12] Ibid.
[13] Safe Routes National Center for Safe Routes to School, Trends in Walking and Bicycling to School from 2007 to 2013 (2015).
[14] Federal Highway Administration. (2009). 2009 National Household Travel Survey, U.S. Department of Transportation, Washington, DC. Available online: https://nhts.ornl.gov.
[15] The Safer Affordable Fuel-Efficient Vehicles Rule for MY2021-MY2026 Passenger Cars, BUILD Guidance 2020, Table A-7 and Light Trucks Preliminary Regulatory Impact Analysis (October 2018) https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/ld_cafe_co2_nhtsa_2127-al76_epa_pria_181016.pdf
[16] The Safer Affordable Fuel-Efficient Vehicles Rule for MY2021-MY2026 Passenger Cars, BUILD Guidance 2020, Table A-7 and Light Trucks Preliminary Regulatory Impact Analysis (October 2018) https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/ld_cafe_co2_nhtsa_2127-al76_epa_pria_181016.pdf

Sulfur Oxides (SO_2)[17]	0.000000008

Quality of Life Benefits

More people bicycling and walking can help encourage an increase in physical activity levels, increased cardiovascular health, and other positive outcomes for users. The benefits from reduced mortality were calculated using the recommended values provided in the USDOT BCA Guidance (Table A-13) and the national distribution of age ranges and travel patterns. These benefits were only applied to the estimated number of walking and biking trips induced by the project (see Demand section). Table 11 displays the multipliers that were used. The estimated annual mortality benefits are shown in Table 18.

Table 11: Mortality Reduction Multipliers

Mortality Reduction Benefits of Induced Active Transportation	Value
Walking Value per Induced Trip	$8.36
Cycling Value per Induced Trips	$7.45
Walking Age Proportion (20-74 years old)	68%
Cycling Age Proportion (20-64 years old)	59%
Trips induced from non-active modes	89%

Economic Competitiveness Benefits

For every vehicle-mile reduced, there is a reduction in household transportation costs and congestion costs. Table 12 displays the multipliers use to calculate economic competitiveness benefit. The estimated annual economic competitiveness benefits are shown in Table 19.

Table 12: Economic Competitiveness Multipliers

Type of Savings	Value

[17] The Safer Affordable Fuel-Efficient Vehicles Rule for MY2021-MY2026 Passenger Cars, BUILD Guidance 2020, Table A-7 and Light Trucks Preliminary Regulatory Impact Analysis (October 2018) https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/ld_cafe_co2_nhtsa_2127-al76_epa_pria_181016.pdf

Household Transportation Cost Savings	$0.54 per VMT[18]
Congestion Cost Savings	$0.08 per VMT[19,20]

[18] Our Driving Costs, AAA (2016).
[19] Crashes vs. Congestion: What's the Cost to Society? AAA (2011). https://exchange.aaa.com/wp-content/uploads/2012/07/AAA-Crashes-vs-Congestion-2011.pdf
[20] Crashes vs. Congestion: What's the Cost to Society? AAA (2011). https://exchange.aaa.com/wp-content/uploads/2012/07/AAA-Crashes-vs-Congestion-2011.pdf

Safety Benefits

The proposed project would decrease conflicts between people walking and biking with motor vehicles. Collision data covering a five-year period (January 1, 2019-December 31, 2025) was provided by the North Carolina Department of Transportation. Collisions under consideration were located along the project corridor and included all types of collisions due to the various benefits of the proposed improvements (Table 13).

A total of 25 crashes were included in the analysis, including two fatalities: a crash into a fixed object at westbound approach to the bridge over Mud Creek and a crash where a driver struck and killed a bicyclist at the intersection of 7th Avenue and Oklawaha Greenway. Different Crash Reduction Factors were applied to the selected crashes for each segment of the project and the benefits were monetized using the values provided in the USDOT BCA Guidance Appendix A, Table A-1 for KABCO Level data. Each crash was assigned to a single countermeasure. The countermeasures selected were:

- Corridor-specific traffic calming (railroad tracks to Cherry St) for injury crashes (CM ID 587, CRF 0.18); non-injury crashes (CM ID 590, CRF 0.06)
- Pedestrian countdown signals (intersection with Ashe St) for rear-end crashes (CM ID 10117, CRF 0.13)
- Systemic signing and marking improvements (intersection with Cherry St) for injury crashes (CM ID 8893, CRF 0.186); non-injury crashes (CM ID 8892, CRF 0.121)
- Install bike lanes with lane width reduction (Cherry St to bridge over Mud Creek) for all crashes (CM ID 10741, CRF 0.266)
- Improve street lighting illuminance and uniformity (in and around Oklawaha Greenway and approaches to bridge over Mud Creek) for all crashes and especially the fatality, a crash into a fixed object at westbound approach to the bridge over Mud Creek (CM ID 11027; CRF 0.38)

The estimated annual safety benefits are shown in Table 20.

Table 13: Summary of Collisions in project corridor, January 1, 2019- December 31, 2025, NCDOT

CMF	Total Collisions	Fatal (K)	Incapacitating (A)	Serious (B)	Possible (C)	PDO (O)	Injured – Unknown Severity (U)
Install Bike Lanes with Lane Reduction	8	1	0	3	0	0	4
Corridor-specific Traffic Calming – Injury	1	0	0	0	1	0	0
Corridor-specific Traffic Calming – No Injury	4	0	0	0	0	4	0
Systemic Signing and Marking – Injury	2	0	0	0	2	0	0
Systemic Signing and Marking – No Injury	6	0	0	0	0	6	0
Pedestrian Countdown Signals	2	0	0	0	0	2	0
Improve Street Lighting Illuminance and Uniformity	2	1	0	0	0	0	1
Total	25	2	0	3	3	12	5

State-of-good Repair Benefits

Table 14 shows the estimated roadway maintenance cost savings associated with a reduction in vehicle-miles traveled.

Table 14: State of Good Repair Multiplier

	Value (metric tons/VMT)
Roadway Maintenance Cost Savings	$0.10 per VMT[21]

[21] Kitamura, R., Zhao, H., and Gubby, A. R. Development of a Pavement Maintenance Cost Allocation Model. Institute of Transportation Studies, University of California, Davis. https://trid.trb.org/view.aspx?id=261768

Results

Table 15 through Table 24 display the results of the benefit-cost analysis for each year of the analysis period. This BCA estimates the project compared to the no-build scenario over a 23-year evaluation (2027-2049) and at a 7.0 percent real discount rate will have a net present value of **$7 million** and a benefit-cost ratio of **2.58:1.0.**

Table 15: Estimated Annual Bicycle and Walk Trips

Year	Baseline	Build Scenario	Additional Trips
2027	664,200	664,200	-
2028	664,500	664,500	-
2029	664,900	664,900	-
2030	665,200	672,700	7,500
2031	665,500	673,200	7,700
2032	665,800	673,600	7,800
2033	666,200	674,100	7,900
2034	666,500	674,500	8,000
2035	666,800	674,900	8,100
2036	667,200	675,400	8,200
2037	667,500	675,800	8,300
2038	667,800	676,300	8,500
2039	668,100	676,700	8,600
2040	668,500	677,200	8,700
2041	668,800	677,600	8,800
2042	669,100	678,100	9,000
2043	669,500	678,500	9,000
2044	669,800	679,000	9,200
2045	670,100	679,400	9,300
2046	670,400	679,900	9,500
2047	670,800	680,400	9,600
2048	671,100	680,800	9,700
2049	671,400	681,300	9,900
Total Additional Trips:		173,300	

Table 16: Estimated Annual Vehicle Miles Reduced

Year	Baseline	Build Scenario	Additional Vehicle Miles Reduced
2027	81,900	81,900	-
2028	81,900	81,900	-
2029	81,900	81,900	-
2030	82,000	82,900	900
2031	82,000	83,000	1,000
2032	82,100	83,000	900
2033	82,100	83,100	1,000
2034	82,100	83,200	1,100
2035	82,200	83,200	1,000
2036	82,200	83,300	1,100
2037	82,300	83,300	1,000
2038	82,300	83,400	1,100
2039	82,300	83,400	1,100
2040	82,400	83,500	1,100
2041	82,400	83,500	1,100
2042	82,500	83,600	1,100
2043	82,500	83,700	1,200
2044	82,500	83,700	1,200
2045	82,600	83,800	1,200
2046	82,600	83,800	1,200
2047	82,700	83,900	1,200
2048	82,700	83,900	1,200
2049	82,700	84,000	1,300
Total Additional Vehicle Miles Reduced:		22,000	

Table 17: Estimated Annual Environmental Sustainability Benefits (Undiscounted)

Year	Baseline	Build Scenario	Benefits
2027	$-	$-	$-
2028	$-	$-	$-
2029	$-	$-	$-
2030	$2,400	$2,400	$-
2031	$2,400	$2,400	$-
2032	$2,400	$2,400	$-
2033	$2,400	$2,400	$-
2034	$2,400	$2,400	$-
2035	$2,400	$2,400	$-
2036	$2,400	$2,400	$-
2037	$2,400	$2,400	$-
2038	$2,400	$2,400	$-
2039	$2,400	$2,400	$-
2040	$2,400	$2,500	$100
2041	$2,400	$2,500	$100
2042	$2,400	$2,500	$100
2043	$2,400	$2,500	$100
2044	$2,400	$2,500	$100
2045	$2,400	$2,500	$100
2046	$2,400	$2,500	$100
2047	$2,400	$2,500	$100
2048	$2,400	$2,500	$100
2049	$2,400	$2,500	$100
Total Benefits:			$1,000

Table 18: Estimated Annual Quality of Life Benefits (Undiscounted)

Year	Baseline	Build Scenario	Benefits
2027	$-	$-	$-
2028	$-	$-	$-
2029	$-	$-	$-
2030	$3,077,000	$3,111,000	$34,000
2031	$3,078,000	$3,114,000	$36,000
2032	$3,080,000	$3,116,000	$36,000
2033	$3,081,000	$3,118,000	$37,000
2034	$3,083,000	$3,120,000	$37,000
2035	$3,084,000	$3,122,000	$38,000
2036	$3,086,000	$3,124,000	$38,000
2037	$3,087,000	$3,126,000	$39,000
2038	$3,089,000	$3,128,000	$39,000
2039	$3,090,000	$3,130,000	$40,000
2040	$3,092,000	$3,132,000	$40,000
2041	$3,093,000	$3,134,000	$41,000
2042	$3,095,000	$3,136,000	$41,000
2043	$3,096,000	$3,138,000	$42,000
2044	$3,098,000	$3,140,000	$42,000
2045	$3,100,000	$3,142,000	$42,000
2046	$3,101,000	$3,145,000	$44,000
2047	$3,103,000	$3,147,000	$44,000
2048	$3,104,000	$3,149,000	$45,000
2049	3,106,000	3,151,000	$45,000
Total Benefits:			$800,000

Table 19: Estimated Annual Economic Competitiveness Benefits (Undiscounted)

Year	Baseline	Build Scenario	Benefits
2027	$-	$-	$-
2028	$-	$-	$-
2029	$-	$-	$-
2030	$51,500	$52,100	$600
2031	$51,500	$52,100	$600
2032	$51,500	$52,100	$600
2033	$51,600	$52,200	$600
2034	$51,600	$52,200	$600
2035	$51,600	$52,200	$600
2036	$51,600	$52,300	$700
2037	$51,700	$52,300	$600
2038	$51,700	$52,400	$700
2039	$51,700	$52,400	$700
2040	$51,700	$52,400	$700
2041	$51,800	$52,500	$700
2042	$51,800	$52,500	$700
2043	$51,800	$52,500	$700
2044	$51,800	$52,600	$800
2045	$51,900	$52,600	$700
2046	$51,900	$52,600	$700
2047	$51,900	$52,700	$800
2048	$51,900	$52,700	$800
2049	$52,000	$52,700	$700
Total Benefits:			$13,600

Table 20: Estimated Annual Safety Benefits (Undiscounted)

Year	Baseline	Build Scenario	Benefits
2027	$-	$-	$-
2028	$-	$-	$-
2029	$-	$-	$-
2030	$-	$1,430,000	$1,430,000
2031	$-	$1,430,000	$1,430,000
2032	$-	$1,430,000	$1,430,000
2033	$-	$1,430,000	$1,430,000
2034	$-	$1,430,000	$1,430,000
2035	$-	$1,430,000	$1,430,000
2036	$-	$1,430,000	$1,430,000
2037	$-	$1,430,000	$1,430,000
2038	$-	$1,430,000	$1,430,000
2039	$-	$1,430,000	$1,430,000
2040	$-	$1,430,000	$1,430,000
2041	$-	$1,430,000	$1,430,000
2042	$-	$1,430,000	$1,430,000
2043	$-	$1,430,000	$1,430,000
2044	$-	$1,430,000	$1,430,000
2045	$-	$1,430,000	$1,430,000
2046	$-	$1,430,000	$1,430,000
2047	$-	$1,430,000	$1,430,000
2048	$-	$1,430,000	$1,430,000
2049	$-	$1,430,000	$1,430,000
Total Benefits:			$ 28,600,000

Table 21: Estimated Annual State of Good Repair Benefits (Undiscounted)

Year	Baseline	Build Scenario	Benefits
2027	$-	$-	$-
2028	$-	$-	$-
2029	$-	$-	$-
2030	$8,000	$8,100	$100
2031	$8,000	$8,100	$100
2032	$8,000	$8,100	$100
2033	$8,000	$8,100	$100
2034	$8,000	$8,100	$100
2035	$8,000	$8,100	$100
2036	$8,000	$8,100	$100
2037	$8,000	$8,200	$200
2038	$8,100	$8,200	$100
2039	$8,100	$8,200	$100
2040	$8,100	$8,200	$100
2041	$8,100	$8,200	$100
2042	$8,100	$8,200	$100
2043	$8,100	$8,200	$100
2044	$8,100	$8,200	$100
2045	$8,100	$8,200	$100
2046	$8,100	$8,200	$100
2047	$8,100	$8,200	$100
2048	$8,100	$8,200	$100
2049	$8,100	$8,200	$100
Total Benefits:			$2,100

Table 22: Estimated Annual Maintenance Disbenefits (Undiscounted)

Year	Baseline	Build Scenario	Benefits
2027	$-	$-	$-
2028	$-	$-	$-
2029	$-	$-	$-
2030	$-	$(5,000)	$(5,000)
2031	$-	$(5,000)	$(5,000)
2032	$-	$(5,000)	$(5,000)
2033	$-	$(5,000)	$(5,000)
2034	$-	$(5,000)	$(5,000)
2035	$-	$(5,000)	$(5,000)
2036	$-	$(5,000)	$(5,000)
2037	$-	$(5,000)	$(5,000)
2038	$-	$(5,000)	$(5,000)
2039	$-	$(5,000)	$(5,000)
2040	$-	$(5,000)	$(5,000)
2041	$-	$(5,000)	$(5,000)
2042	$-	$(5,000)	$(5,000)
2043	$-	$(5,000)	$(5,000)
2044	$-	$(5,000)	$(5,000)
2045	$-	$(5,000)	$(5,000)
2046	$-	$(5,000)	$(5,000)
2047	$-	$(5,000)	$(5,000)
2048	$-	$(5,000)	$(5,000)
2049	$-	$(5,000)	$(5,000)
Total Benefits:			$(100,000)

Table 23: Estimated Annual Benefits (Undiscounted)

Year	Baseline	Build Scenario	Benefits
2027	$-	$-	$-
2028	$-	$-	$-
2029	$-	$-	$-
2030	$3,139,000	$4,600,000	$1,461,000
2031	$3,140,000	$4,601,000	$1,461,000
2032	$3,142,000	$4,604,000	$1,462,000
2033	$3,143,000	$4,605,000	$1,462,000
2034	$3,145,000	$4,608,000	$1,463,000
2035	$3,146,000	$4,609,000	$1,463,000
2036	$3,148,000	$4,612,000	$1,464,000
2037	$3,150,000	$4,615,000	$1,465,000
2038	$3,151,000	$4,616,000	$1,465,000
2039	$3,153,000	$4,619,000	$1,466,000
2040	$3,154,000	$4,620,000	$1,466,000
2041	$3,156,000	$4,623,000	$1,467,000
2042	$3,157,000	$4,624,000	$1,467,000
2043	$3,159,000	$4,627,000	$1,468,000
2044	$3,160,000	$4,629,000	$1,469,000
2045	$3,162,000	$4,631,000	$1,469,000
2046	$3,163,000	$4,633,000	$1,470,000
2047	$3,165,000	$4,635,000	$1,470,000
2048	$3,167,000	$4,638,000	$1,471,000
2049	$3,168,000	$6,573,000	$3,405,000
Total Benefits:			$31,254,000

Table 24: Estimated Discounted Net Costs and Benefits

Year	Discounted Costs	Discounted Benefits	Net Cumulative Discounted Costs and Benefits
2027	$(1,184,000)	$-	$(1,184,000)
2028	$(2,212,000)	$-	$(3,396,000)
2029	$(1,034,000)	$-	$(4,430,000)
2030	$-	$973,000	$(3,456,000)
2031	$-	$910,000	$(2,546,000)
2032	$-	$851,000	$(1,696,000)
2033	$-	$795,000	$(900,000)
2034	$-	$744,000	$(156,000)
2035	$-	$695,000	$539,000
2036	$-	$650,000	$1,189,000
2037	$-	$608,000	$1,797,000
2038	$-	$568,000	$2,365,000
2039	$-	$531,000	$2,896,000
2040	$-	$497,000	$3,393,000
2041	$-	$464,000	$3,857,000
2042	$-	$434,000	$4,291,000
2043	$-	$406,000	$4,697,000
2044	$-	$380,000	$5,077,000
2045	$-	$355,000	$5,432,000
2046	$-	$332,000	$5,763,000
2047	$-	$310,000	$6,073,000
2048	$-	$290,000	$6,364,000
2049	$-	$627,000	$6,991,000
Total Net Discounted Costs: $ 4,430,000		Total Discounted Net Benefits: $11,420,000	Net Present Value: $6,991,000
Benefit-Cost Ratio: 2.58:1			

Appendix A:
Technical Documentation - Replica Methodology

Seasonal Mobility Model Methodology Extended (Places)

Document Purpose

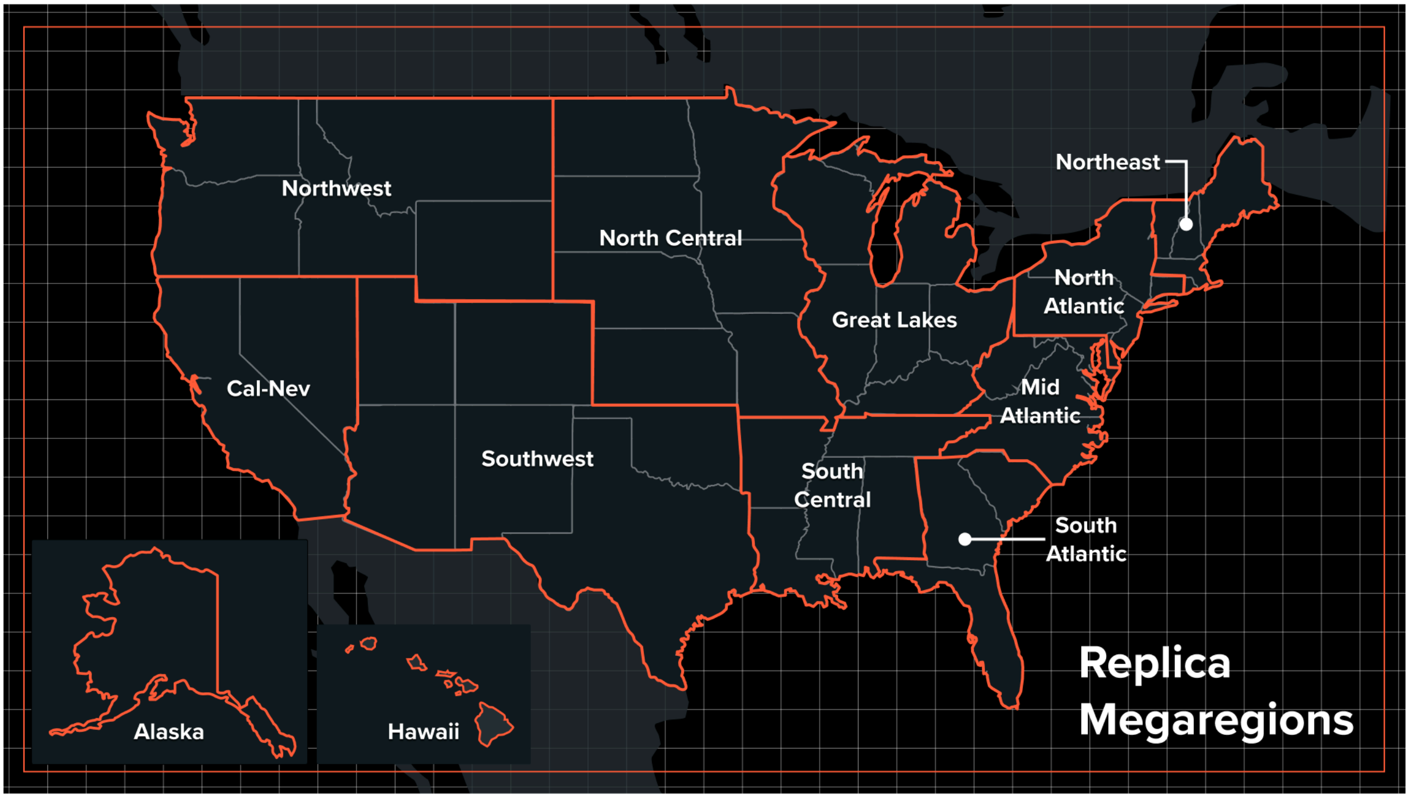
This document is written to provide a detailed explanation of the methodology used to create Replica's seasonal mobility model (Places) and weekly mobility model (Trends). At Replica, we understand that data is valuable only to the extent that it is trusted to inform analysis and decision-making. To that end, this document provides an overview of Replica's data sources, data processing methods, statistical inference systems, and data outputs, in order to help our customers evaluate the quality and accuracy of our models, and assess privacy and data security implications.

Introduction

Replica's seasonal mobility model (Places) is a high-fidelity activity-based travel model with network-link level granularity. Each model is a synthetically generated representation of the activities and movements of residents, visitors, and commercial vehicles on a typical weekday and typical weekend day for a given location and season.

The output of each seasonal mobility model is a complete, disaggregate trip and population table. Each completed model also includes a Quality Report, which shows how the outputs of the model compare to ground truth data. The report facilitates comparisons and validations between Replica's modeled outputs and a customer's observed counts.

Replica's seasonal mobility models are delivered at "megaregion" scale, most of which cover geographies that include between 10 million and 50 million residents.



Megaregion Boundaries

Data Sources

Replica's seasonal mobility model (Places) utilizes a diverse set of public and private data sources. This composite approach mitigates risk by minimizing sampling bias, creating resiliency against data quality issues, and protecting against data source disruption, while enabling Replica to deliver detailed modeled outputs. Building component models from different data sources independently also creates additional privacy protecting measures, as it enables Replica to abstract out potentially identifying details of any individual before combining these models into our aggregate outputs.

These data sources fall into five broad categories: 1) Mobile Location Data, 2) Consumer / Resident Data, 3) Built Environment Data, 4) Economic Activity Data, 5) Ground Truth Data.

This section provides a brief overview of each category of data. How each data source integrates into the data processing pipeline is addressed in later sections.

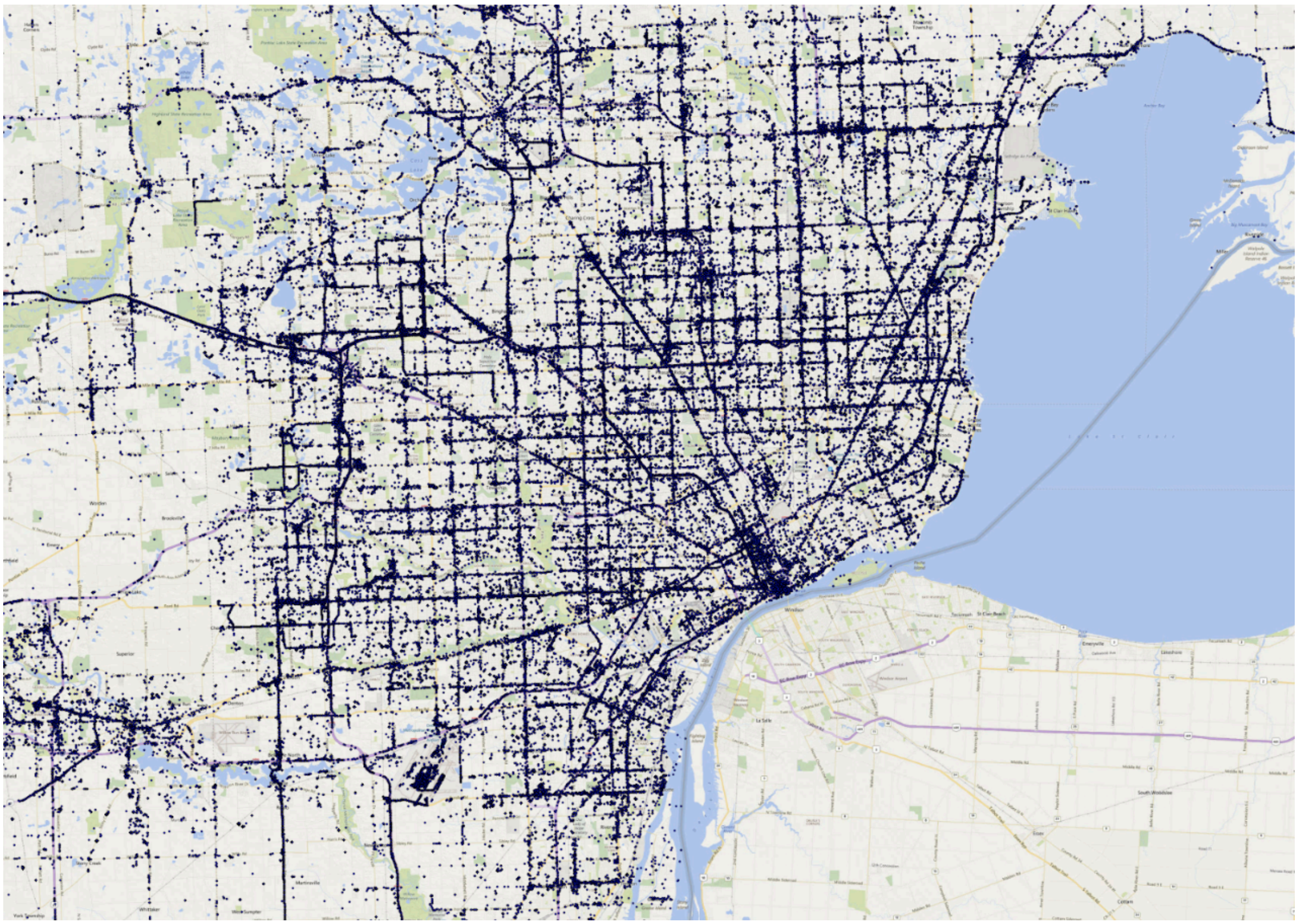
Mobile Location Data

Replica incorporates three types of mobile location data into its pipeline: 1) location-based services data, 2) vehicle in-dash GPS data, and 3) point-of-interest aggregates. Previous versions of Replica's model also included cellular networks data as another source of mobile location data.

Location-based Services (LBS) Data

As people move around with their phones in the real world, they use mobile apps that rely on their location, which is determined via GPS. Users opt in to sharing their location when using these apps. The user's location as identified at a specific point in time is commonly referred to as a location trace. For users that opt in to sharing their location, mobile app publishers then aggregate these location traces, keeping the localization accuracy set at the device by the user. The mobile app publisher then de-identifies the traces and licenses them to data aggregators (e.g., Cuebiq, Safegraph, Gravy Analytics). Using the mobile advertising ID, or "ADID", uniquely generated for each device, aggregators join data from multiple app publishers into a single-time sequenced data stream for each device. (Note: Users can reset their ADID at any time).

A single data aggregator usually sees a sample size equal to approximately 10% of the US population. Because location traces are dependent on a device using a specific application, and recent Android and iOS operating systems are more and more restrictive in background location data collection, the space-time coverage of this type of data is often volatile. Replica sees data from more than 30 million unique devices each month.



Visual representation of raw location data collected from several thousand random devices for a period of 1 hour in the Detroit region, across all data types.

Vehicle in-dash GPS Data

Car manufacturers often integrate both GPS and cellular connectivity hardware into their vehicles for real-time navigation, safety, and other services available through in-dash infotainment systems. This generates data on vehicle speeds and locations, which can be geo-matched to a particular road segment (telematics data), and transferred to centralized data storage and processing systems that monitor real-time congestion.

The primary purpose of these systems is to provide better routing services to consumers and/or manage operations of specialized vehicle fleets (e.g., freight, shuttles, taxi vehicles). This telematics data is also licensed by third parties, which is governed by EULAs. Replica processes in-dash GPS data from over 3 billion trips each month.

Point-of Interest Aggregates

Wireless technology can be used to detect and count an aggregate number of mobile devices present in a given venue (e.g., a park or a shopping mall). Aggregators of this point-of-interest (POI) information provide a total count of devices in their sample at each location, providing a signal to estimate the popularity of each venue. This signal enables Replica to generate the relative occupancy weights at different points and areas of interest for its own model. Replica processes POI data from over 4 million locations each month.

Consumer / Resident Data

Replica incorporates both public and private sources of resident and consumer data as the basis for creating a synthetic population. The majority of this data is sourced from the US Census Bureau. Replica utilizes the following US Census datasets:

- 5-Year American Community Survey (ACS)
- 1- and 5-Year Public Use Microdata Sample (PUMS)
- Longitudinal Employer-Household Dynamics (LEHD)
- Census Transportation Planning Products (CTPP)
- National Center for Education Statistics
- US Department of Education

Built Environment Data

Replica ingests and utilizes a number of sources for transportation network, land use, and real estate data. These include:

- OpenStreetMap (OSM)
- General Transit Feed Specification (GTFS) data. Replica updates its nationwide database of GTFS feeds twice each year. Currently, the database includes feeds for 300 agencies and 11,000 routes. If customers have GTFS data for additional agencies, Replica can add those agencies to its database.
- A number of proprietary sources including providers of land use categorization data, parcel data, building footprint data, building floor area data, POI data (e.g., restaurants, stadiums, theaters), parking data, and hotel inventory data. Parcel-level data is collected for roughly two-thirds of the land area of the United States.

Economic Activity Data

Financial data processing companies collect and store the aggregate number and characteristics of transactions (such as dollars spent or the time of transaction) made at different categories of vendors within a given geography in a given period of time. This payments ecosystem data broadly has 3 sources: (1) Merchant Acquiring Data, which ties transaction data to the venue where the transaction occurred; (2) Bank Issuing Data; and (3) Payments Networks, which tie transaction data to the card making the transaction. Replica currently ingests 2 of those 3 sources of economic activity data.

Ground Truth Data

Replica uses a number of different sources of ground truth, or observed counts, to calibrate its models. This data is “held back” during the model creation phase and then used to calibrate initial outputs. Specific types of ground truth include, but are not limited to:

- Auto / Traffic Counts, both in aggregate and for specific vehicle types (Private Auto, Freight)
- Transit Ridership Counts
- TNC/Taxi Counts

Ground truth is both sourced by Replica directly, and optionally provided by Replica’s customers. Each of Replica’s seasonal mobility models can be sufficiently calibrated exclusively with Replica-sourced ground truth. The inclusion of customer-provided ground truth is neither a requirement nor a dependency to delivering a seasonal mobility model. Nor does the absence of customer-provided ground truth meaningfully impair the quality of a modeled season.

However, certain customers with robust sets of ground truth data elect to include that data in the model calibration process to further calibrate the model, and to better compare Replica’s outputs against their own counts.

Core Data Product Creation

The first step in the Replica seasonal mobility model (Places) development process is the transformation of source data into Core Data Products (CDPs). If the Data Sources described above are the “raw materials” to Replica’s development process, the four CDPs described below are the essential building blocks of Replica's model. A common characteristic of each CDP is that they produce nationwide outputs.

Population CDP

The Population CDP is a nationwide synthetic population that is statistically similar to the census population in person-level attributes, household composition, and in aggregate, at each census-level geography. These synthetic people and households are assigned to housing units, work locations, and school locations, which represent long-term choices that do not vary during the one-week simulated time period of an individual seasonal mobility model. Replica’s nationwide synthetic population is updated seasonally, using the following process.

1. Creation of Households and Persons

The first step in creating the Population CDP is to train two Bayesian networks from the US Census's Public Use Microdata (PUMS) sample files — one network representing households, and another representing person attributes.

The purpose of the Bayesian network approach is to model a diversity of persons and households with attributes that reflect conditional probability distributions of the attributes specific to each given geographic area (e.g., a Public Use Micro Area [PUMA] in the US). In addition to geographical partitioning, the models are stratified by household size and person types: different household models are trained within each PUMA for households of a given size, and different person models are trained for a given role of a person in a household (e.g., employed adult, a child, a relative, etc).

An example of the Household and Person Bayesian network models structures.

The structure of the network represents conditional dependencies between the attributes. For example, for each household-sized group within a microdata sample area, for a particular household of a given size belonging to a given household-income group, the Bayesian network includes a link that represents the probabilities of owning different numbers of vehicles. Sample structures of the models are shown in the figure above.

When data is available, the set of standard attributes and the structure of the Bayesian network model can be extended, learned from data, and customized to include other variables of interest as well as to integrate auxiliary and proprietary data sources required for a particular use case.

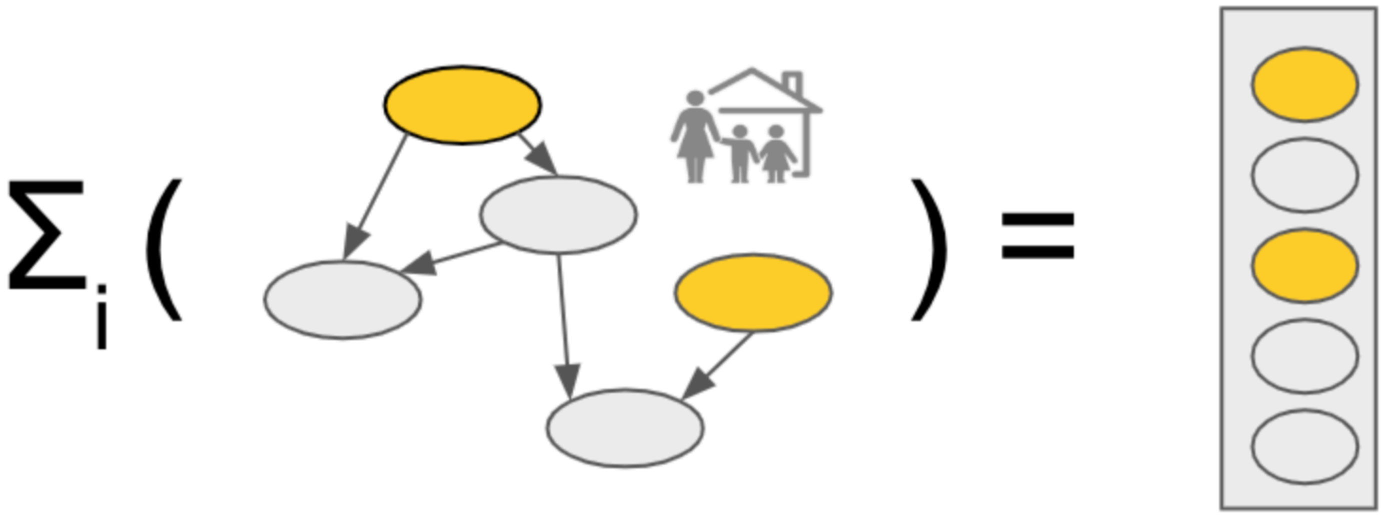
2. Matching Controls in Aggregate

In the second step of building the synthetic population, the system allocates microdata-based household profiles to census areas, (typically block groups), in order to match the known total census aggregate numbers (i.e. "controls"), using a convex optimization method. The output of the optimization method is a set of weights which, for every census area, describe how many times a given set of controlled seed attributes of a given household and person's record should be used for a matching Bayesian network model to generate a synthetic household and persons.

Not every person or household attribute can be "controlled" for in this method, so only salient attributes are chosen. These attributes, which are used as a set of controlled variables, include: household size, household income group, the number of vehicles owned, employment status, industry of employment, commute mode, housing tenure, age group, sex, race, ethnicity, and school grade attending. Meta-controls at the PUMA level are included in the allocation algorithm's objective function.

Due to the internal inconsistencies in census aggregates (attributed to data suppression, binning thresholds, and non-sampling errors), limited accuracy of data for hard-to-reach populations, presence of significant special population groups in given areas (e.g., military establishments), and other reasons, the allocation algorithm parameters may require adjustments considering the robustness of the synthesis with respect to data versus the overall accuracy across different areas.

At the end of this step, households and persons are generated by passing each household's profile and their members through the Bayesian network developed in the first step. In addition to the controlled seed attributes, each person and household is assigned other attributes based on the relationships and probabilities of the Bayesian network models.



The purpose of the allocation step is to find a combination of the households (weights) to use as seeds in Bayesian network synthesis, such that the total set of synthesized households and persons in aggregate match the given "controls" for the main attributes of interest within each area.

3. Home, Work, and School Assignments

In the final step of the process, households are assigned to housing units within a census block group, and employed persons and students are assigned workplaces and schools, respectively.

Home, work and school assignments are made at the sub-census block group resolution for the purposes of modeling fidelity and are not exposed in Replica's user interface.

- *Housing.* Decennial census estimates of the total population and dwelling units available within each census block, as well as proprietary or customer-supplied land use, geocoded postal addresses, and building footprint data are used at this step to improve spatial fidelity and modeling realism. While it is possible to adjust the controls and update the synthetic population to reflect recent population changes not captured by census, or to reproduce future scenarios, a standard Replica release is generally based on the most recent 5-year ACS data.

- *Places of Work*. Each employed synthetic person is allocated first to a workplace at the census tract-level, then at the census block group-level, then to an individual office unit. Using a combination of the most recent Census Transportation Planning Product (CTPP) tabulation of the ACS data and location-based data used to generate weekly home-work assignments, we construct a home-work matrix representing the number of residents traveling between a home census tract to a work census tract based on their commute mode, industry of employment and income group. The reason for incorporating the home-work allocation into the home-work assignment is to see seasonal changes in work assignment since CTPP data is generated infrequently and the most recent CTPP data was from 2016. These flows are given as marginal distributions in CTPP, and an optimization algorithm is applied to reconstruct the conditional distribution of the workplace as a function of the employed person's attributes. Land use and LEHD employment data are then used to redistribute flows to census block groups within tracts, according to the availability of employment in any given industry sector. Once work block groups have been assigned to each employed person, the individual office unit assignment leverages parcel land use, building, points of interest, and industry information to identify the most likely work location for that person.
- *School Enrollment*. School-aged residents are assigned to a school location based on each school's enrollment, school district boundaries, and proximity to home. Students are assigned to schools that offer classes matching their ages. Grade levels and enrollment counts are sourced from the US National Center for Education Statistics for the relevant school year. Residents aged 16 (legal working age) to 18 are assigned to a secondary school until enrollment counts are met. The balance of these residents will assume their general employment class (e.g., not in the labor force, employed, or unemployed). Residents aged 18 to 23 and 23 to 34 are assigned to undergraduate and graduate institutions respectively using the same process, until total enrollment counts are met. Residents who are classified as employed and a student are considered to have part-time employment and may attend school and a workplace on the same day.

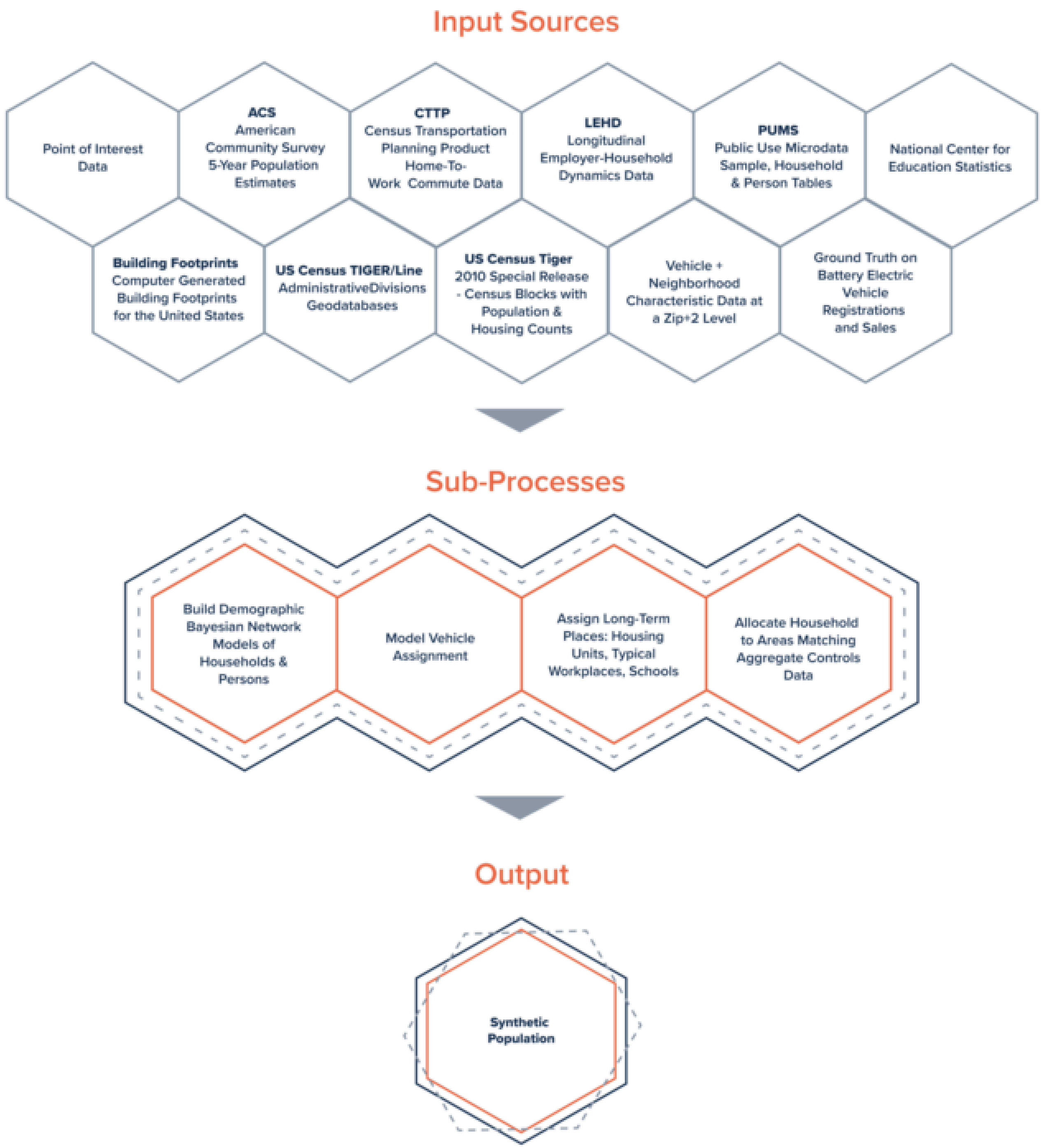
4. Vehicle Assignment Modeling

In the Spring 2023 season, Replica introduced vehicle assignments. Each private auto trip is now tied to one of two vehicle fuel types, Battery Electric Vehicle (BEV), or vehicle with another fuel type. Vehicle assignment is modeled as a four stage process:

- *Model the marginal attributes with respect to vehicle ownership for each household*. Data from a consumer marketing vendor is used to understand the temporal and spatial variability of different vehicle ownership attributes. This vendor collects data from a variety of sources including self-reported survey information, property data, and vehicle title/registration documents. We use two datasets at the zip+2 granularity: 1) a set of characteristics of the vehicles owned within the households of a given zip+2, and 2) sociodemographic characteristics of the individuals and households in the zip+2. We combine these datasets to produce a dataset at the zip+2 granularity that includes automotive and sociodemographic characteristics. For example the percentage of households that own a Battery Electric Vehicle (BEV), average household income for the zip+2, and percent of households by age distribution. We run a K-Nearest Neighbor algorithm taking into account the sociodemographic attributes of Age and Household Income using this combined zip+2-level dataset, and the previously synthesized households. As a result, every household in our synthesized population is assigned a set of average household characteristics, including the percent of households that own a BEV.

- *Model the target number of electric vehicles that exist in each zip code*. Our data vendor provides an estimate for the total number of vehicles, and the percent of households that own an electric vehicle in each zip code. However, the total population of vehicles may differ between our synthesized population and our data vendor. For states where we have publicly available data on BEV registrations or BEV sales, we normalize the zip code level vendor data to match state totals. We then produce a target number of BEVs for each zip code.
- *Generate the set of synthetic vehicles for each household.* Once all households have been assigned the most likely average vehicle ownership characteristics, and all zip codes have a target value for electric vehicles (N), we sample N households using a weighted sampling with replacement out of the set of households in that zip code. The weight used is the percentage of households that own a registered electric vehicle, a property that was assigned to each household in the first step. Thus, the sociodemographic characteristics of households that are more likely to own a BEV are captured.
- *Assign the vehicles in the household to individuals.* We assume that the same individual uses the same vehicle throughout the day. In cases where a household has more individuals above driving age than vehicles, the same vehicle can be assigned to multiple people.

After a vehicle is assigned to an individual, we assume that individual uses that vehicle for every private auto trip on the modeled day.



Built Environment CDP

The Built Environment CDP consists of a number of distinct data tables. The primary data output of the Built Environment CDP is a nationwide land use model that serves as the source of truth for all spatial data consumed by Replica's systems. It maintains three principal layers of disaggregate data: (1) parcels, or "lots", (2) buildings, and (3) points of interest. The union of these three layers with auxiliary spatial datasets such as schools and airports supports the home, work, and school assignments in the Population CDP, and the location choice model for activities in Replica's seasonal and weekly mobility models. It is responsible for creating the comprehensive set of locations an agent can travel to.

The land use of buildings and parcels is classified using direct input data from proprietary sources where available, and otherwise modeled using tract-level characteristics. Parcel-level data can come in a variety of forms, including spatially joined points of interest, zoning descriptions, and postal delivery codes. Tract-level characteristics consist of aggregate land use proportions that are further weighted using the relative, block-level concentration of housing units in Census data.

Replica's hierarchical land use categories are as follows:

Primary category (L1)	Sub-categories (L2)
residential	single_family multi_family
commercial	retail office non_retail_attraction
mixed_use	• combination of an other L2s
industrial	industrial
civic_institutional	healthcare education civic_institutional
transportation_utilities	transportation_utilities
agriculture	agriculture
open_space	open_space
other	other
unknown	unknown

Total square footage for each building is modeled using a composite of individual and aggregate features. Dwelling units are modeled on each residential or mixed-use parcel based on ACS-scaled Census counts. Points of interest are classified based on North American Industry Classification System (NAICS) industry code.

The Built Environment CDP also includes a number of distinct datasets that serve as "base layers," such as the OSM network and nationwide GTFS. All datasets are refreshed seasonally.

Economic Activity CDP

The Economic Activity CDP is a weekly estimate of total consumer spend across a number of categories, for every census tract in the country. An estimate is created for both (1) all "brick and mortar" spend, defined as the sum total of every transaction that occurs in-person within a given tract, regardless of where the purchaser lives, and (2) all on- and off-line resident spend, defined as the sum total of all transactions made by residents of a given tract, regardless of where that transaction occurs.

Consumer spend includes all transactions, including credit card, debit card, and cash transactions, that take place at a point of sale, such as at retail stores, supermarkets, restaurants, taxis, and bars. It also includes e-commerce transactions in these same categories. The data does not include all household expenditures; for example, rent, car payments, and healthcare spending is excluded. This most closely aligns Replica's consumer spend metric to the Census Bureau's Monthly Retail Trade Estimates. Transactions are categorized by the merchant’s NAICS code.

Both Brick and Mortar and Resident Spend for each census tract is generated for the 6 categories listed below. Due to source data limitations, an on- and off-line breakdown for Resident Spend are only generated for three of the six categories, which are marked with an asterisk.

1. Restaurants & Bars*
2. Retail*
3. Grocery Stores*
4. Airline, Hospitality & Car Rental
5. Entertainment & Recreation
6. Gas Stations, Parking, Taxis, & Tolls

Outputs are generated based on three types of source data: (1) Cardholder transaction data; (2) Merchant transaction data; (3) Point of Interest visitor patterns data. Outputs are calibrated against multiple sources of ground truth, including both Census and Bureau of Economic Analysis (BEA) datasets.

At a high level, the methodology consists of 3 steps: (i) aggregated spend by home location (both in-person and e-commerce); (ii) disaggregate spend by POI and aggregated spend by merchant location; (iii) updated in-person spend by home location using disaggregate estimator.

First, we use vendor data (which consists of transactions at merchants and for individual cardholders) to estimate the total spend (in-person + e-commerce) and use the cardholder data to determine what proportion of the total spend is e-commerce. According to census totals, we capture about 10-20% of total transactions (depending on the category), and to estimate aggregate totals, we scale up based on the population within each home tract. To correct for sample bias, we combine nationwide monthly census totals and state-wide annual totals from the BEA to a forecasted state-wide monthly estimator that forecasts ahead to the most recent month. We then scale up tract-level totals using state-wide bias correction factors based on the previous 12 months (to ensure the modeled spend is always forward-looking). Scaling factors typically vary from 1-3x depending on the state and category.

Second, we model POI-level spend using a combination of merchant-level transaction data and POI-level visit count data. Our POI-level model includes visit count data, transaction data, population data, and POI-specific features such as name, common brands, etc. We scale up POI-level spend to ensure consistency with aggregate totals generated in the previous step. Scaling factors are done at the PUMA level since PUMAs have relatively consistent populations.

Finally, to improve the spend-by-home location estimator (for in-person spend only) we use our POI-level spend estimator from the previous step in combination with POI-level modeling of home tract using POI-level visit count data.

Travel Activity CDP

The Travel Activity CDP is a nationwide set of travel behavior models — “Personas” — that describe and predict the movements of individual residents. These travel behavior models are trained using a combination of the mobile location data acquired by Replica as described earlier.

The purpose of creating personas from observed travel behaviors of real people is three-fold: preserve privacy of source data, provide explanation for travel behavior, and enable model sensitivity to urban design and policy changes. The personas are composed of three main underlying behavioral choice models: activity scheduling, destination location, and travel mode, which will be discussed in detail below. A combination of these models can be used to reproduce complete daily activity sequences (e.g., start the day at home, drive to work, walk to lunch at lunch break, complete work, drive to shop, shop, drive to home, stay at home) with each activity annotated with planned start and end times.

Staypoint Creation

The first step in the creation of the Travel Activity CDP is the transformation of traces to staypoints and travel periods. For each device identified by a non-persistent hash of device ID, individual location points (traces) are clustered with a space-time clustering method in order to identify recurrent locations where the device has been stationary over a period of time (a staypoint). The time gap between two consecutive staypoints is labeled as a period of missing data or, in the presence of consistent intermediate locations, a period of travel. The location of each stationary activity is assigned a ”staypoint ID” specific to that device. This process runs daily.

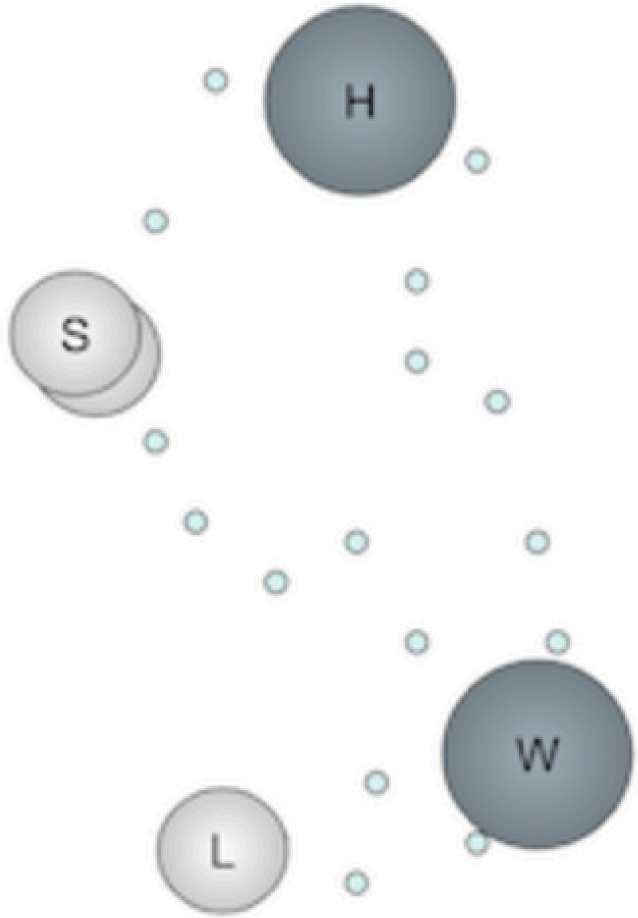
Staypoints are analyzed seasonally (Spring, for instance, includes March, April and May data). For each season, the set of staypoint IDs for a given device is referred to as the device user’s “habitat.” Repeated activity locations are annotated with a common staypoint ID, which is used to identify repeatedly visited locations.

A combination of land use attributes (generated from the Built Environment CDP) and an observed recurrent pattern of stay at the most frequented locations is used to identify home and work locations of a device owner. Home and work activities are identified in the sequence of device movements along with their typical start/end times and durations. Home and work (if observed) are the two most prominent activity types.

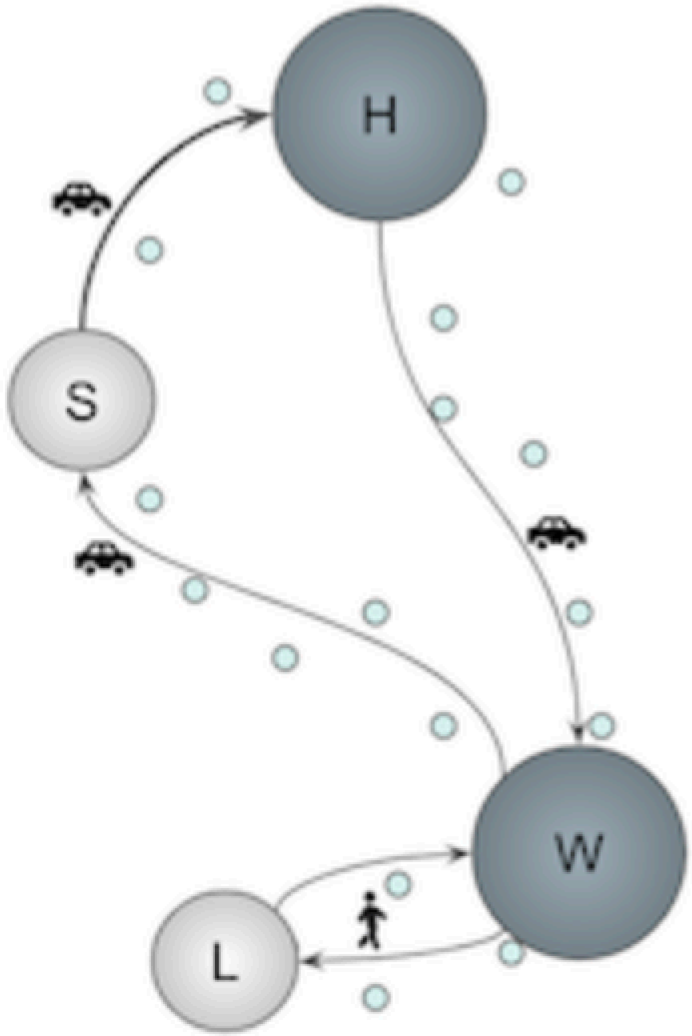
Following the assignment of home and work activities, a combination of land use, popularity scores for a set of venues at a given time, and the context of the stay time period within the sequence of the primary activities within the day are used to assign likely activity types to all secondary staypoints.

The figure below illustrates the three location data and main processing stages: 2) raw location data, b) staypoints detection and activity types inference, and c) travel periods segmentation and mode inference.

a) Raw location data



b) Staypoints detection and activity types inference



c) Travel periods segmentation and mode inference

Known Limitations in Staypoint Detection

During the staypoint creation process, insufficient spatial accuracy or a high dispersion of locations within individual clusters may lead Replica to deem the data not usable. At each stage in the process, data can be withdrawn from further processing in the pipeline depending on the quality and accuracy of the previous stage.

For instance, when significant temporal gaps are present and the coverage is not sufficient to identify secondary staypoints, the data from a device may only be useful to inform algorithms that require aggregate information of home and work flows.

It is worth noting that the start and end time of any activity (and the arrival and departure time for each travel period) cannot be observed with certainty in LBS and cellular data. A typical duration of a period of stay that can be detected from a typical mobile location data feed with continuous temporal sampling (1 location sample per minute) is a stay of over 5 minutes in duration (over 85% detection accuracy), with most stays of 15 minutes identified at 90% accuracy. Very short activities or local activities that happen within the range of localization accuracy can not be reliably detected. For example, it is often not possible to identify events such as buying a coffee at a drive-through location, particularly with no significant stationary periods (such as a long wait in line).

Generating Personas

Models of travel behavior ("personas") of adult residents are trained using mobile location data that was successfully processed into segmented sequences of staypoints and periods of travel as described above. Records from devices with insufficient temporal coverage or insufficient number of complete day records are not used for model training. A complete day is defined as a day with at least 14 hours of coverage (last observed time minus first observed time) and no gaps of more than three hours. Devices with fewer than 7 days of observed overnight stays during each season at identified home locations are also removed from the training data set.

Assigning attributes — such as age group, income, and employment status — to each persona is required for travel modeling (as well as the assignment of personas to the synthetic population). **Because Replica does not acquire nor handle any personally identified information (PII) associated with device owners, these attributes are inferred** using (1) the demographics of the residents and workers for the given home census block group; and (2) the socio-demographics of the workers with a given home-work commute pair detected for a persona as compared to that provided by the most recent CTPP. Commute mode detection and the presence of recurrent home-based trips by car gives a proxy for a vehicle availability attribute.

Weighting factors are computed based on the total number and the number of employed persona models residing within a given census block group. Iterative proportional fitting methods used in weighting traditional surveys are applied for the latter. These weighting factors are in turn used in the adjustments of home-to-work commute flows and persona matching.

Three models are trained for each device: (1) Activity Sequence Model (ASM), (2) Travel Destination Location Choice Model (LCM), and (3) Mode Choice Model (MCM). Each model is described below with respect to how the model is structured and trained on the observed data, including the privacy protection measures applied in the process.

(1) Activity Sequence Model (ASM)

The annotated "activity day sequences" (staypoints and travel periods) form the basis of the device persona's activity sequence model. ASM is a generative sequence learning model. When given historical data on the observed sequence of activities, it can produce the most likely sequence of activities of a person carrying the device on a given day of the week within the season. ASM is also used to fill in short unobserved gaps and predict the likely sequence of activities at the end of the day given its observed beginning. Three hour gaps in the observed data are imputed with the activities predicted by the ASM. This gap-filling functionality of the ASM includes access to a complete library of activities detected with high confidence and shared across all the devices observed in the region.

Without mitigations, activity sequences carry a privacy risk: A person could be re-identified if they have a rare or unique activity sequence that is reproduced in the synthetic data. Specifically, in the event a rare activity sequence is reconstituted in the modeled outputs (e.g., gym visit at 5 am every Wednesday followed by visiting a specific restaurant), it is possible to regenerate a part of a real person’s rare day itinerary. To mitigate this risk, we filter out rare day sequences from the training dataset.

Rare or unique sequences are identified by resampling it at regular time intervals within a 24-hour period beginning at 03:00 local time, and comparing across all observed activities. Figure 5.5 gives an example of this representation (using 30 minute resampling interval) for several days, where each row represents a day with the actual calendar date and day-of-week as metadata.

```
2018-01-09 2 | HHHHHHHHHHHHHHHHHHHH-WWWWWWWWWWWWWWWWWWW-HHHHHHH
2018-01-10 3 | HHHHHHHHHHHHHHHHHHHHH--WWWWWWWWWWWWWWWWWWhhHHHHHHH
2018-01-14 7 | HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHoHHHHHHHHH
2018-01-16 2 | HHHHHHHHHHHHHHHHHHHHHHWWWWWWWWWWWWWWWWWW-HHHHHHH
2018-01-17 3 | HHHHHHHHHHHHHHHHHHHHHH-WWWWWWWWWWWWWWWWWW-HHHHHHH
2018-01-24 3 | HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH-HHHHHHHHHHHHHH
2018-01-29 1 | HHHHHHHHHHHHH--WWWWWWWWWWWWWWWWWWWWWWWWW-HHHHHHHH
2018-02-02 5 | HHHHHHHHHHHHHHHHHHHHHHWWWWWWWWWWWWWWWWWWoo-HHHHHH
2018-02-05 1 | HHHHHHHHHHHHHHHHHHHHH-WWWWWWWWWWWWWWWWWWW-hhHHHHHH
2018-02-16 5 | HHHHHHHHHHHHHHHHHHHHH-WWWWWWWWWWWWWWWWWWWHHHHHHHHH
2018-02-17 6 | HHHHHH-o-hhhhhhhhhhhhhhhhhhhHHHHHHHHHHHHHHHHHH
2018-02-18 7 | HHHHHHHHHHHHHHHHHHHHHHHHHHHH-oo-hhhhhhhhHHHHHHHHH
2018-02-22 4 | HHHHHHHHHHHHHHHHHHHHH-WWWWWWWWWWWWWWWWWWWhh-HHHHHH
2018-03-01 4 | HHHHHHHHHHHHHHHHHHHHHHWWWWWWWWWWWWWWWWWWWoHHHHHHHH
2018-03-05 1 | HHHHHHHHHHHHHHHHHHHHH--WWWWWWWWWWWWWWWWWWoHHHHHHHH
2018-03-06 2 | HHHHHHHHHHHHHHHHHHHHH-WWWWWWWWWWWWWWWWWWWoo-HHHHHH
2018-03-15 4 | HHHHHHHHHHHHHHHHHHHHH--WWWWWWWWWWWWWWWWWWoHHHHHHHH
2018-03-19 1 | HHHHHHHHHHHHHHHHHHHHHHWWWWWWWWWWWWWWWWWWWhhhHHHHHH
2018-03-25 7 | HHHHHHHHHHHHHHHHHHHHHHHHH--oooooooooHHHHHHHHHHHHH
```

Example encoded day activity sequences for a single device/user (H:home, W:work, o:other, '-':in transit, h: secondary home).

Two day-sequences are considered similar if there is less than one hour shift in either direction and have less than two 30-minute time bins with different activity types. The figure below provides a working example. The “rare” day-sequences are then compared against each of the common day-sequences. If they are similar, the sequence is kept, otherwise it is removed from the dataset.

Common day-sequence:

```
1: HHHHHHHHHHHHHHHHHHH-WWWWWWWWWWWWWWWWWWWWW-HHHHHHHH
```

Day sequence to evaluate:

```
11: HHHHHHHHHHHHHHHHH-WWWWWWWWWooWWWWWWWWWW-HHHHHHHHHH
```

Evaluate: Is sequence 11 similar to sequence 1?

Step 1: Generate 5 permutations of sequence 11.

```
11(-2): HHHHHHHHHHHHHHH-WWWWWWWWWooWWWWWWWWWW-HHHHHHHHHHHH
11(-1): HHHHHHHHHHHHHHHH-WWWWWWWWWooWWWWWWWWWW-HHHHHHHHHHH
11( 0): HHHHHHHHHHHHHHHHH-WWWWWWWWWooWWWWWWWWWW-HHHHHHHHHH
11(+1): HHHHHHHHHHHHHHHHHH-WWWWWWWWWooWWWWWWWWWW-HHHHHHHHH
11(+2): HHHHHHHHHHHHHHHHHHH-WWWWWWWWWooWWWWWWWWWW-HHHHHHHH
```

Step 2: Compare each permutation of sequence 11 to sequence 1.

```
11(-2): HHHHHHHHHHHHHHH-WWWWWWWWWooWWWWWWWWWW-HHHHHHHHHHHH
     1: HHHHHHHHHHHHHHHHHHH-WWWWWWWWWWWWWWWWWWWWW-HHHHHHHH
```

Shift: -2, Edit distance: 12 → **Not similar**

```
11(-1): HHHHHHHHHHHHHHHH-WWWWWWWWWooWWWWWWWWWW-HHHHHHHHHHH
     1: HHHHHHHHHHHHHHHHHHH-WWWWWWWWWWWWWWWWWWWWW-HHHHHHHH
```

Shift: -1, Edit distance: 10 → **Not similar**

```
11( 0): HHHHHHHHHHHHHHHHH-WWWWWWWWWooWWWWWWWWWW-HHHHHHHHHH
     1: HHHHHHHHHHHHHHHHHHH-WWWWWWWWWWWWWWWWWWWWW-HHHHHHHH
```

Shift: 0, Edit distance: 8 → **Not similar**

```
11(+1): HHHHHHHHHHHHHHHHHH-WWWWWWWWWooWWWWWWWWWW-HHHHHHHHH
     1: HHHHHHHHHHHHHHHHHHH-WWWWWWWWWWWWWWWWWWWWW-HHHHHHHH
```

Shift: +1, Edit distance: 6 → **Not similar**

```
11(+2): HHHHHHHHHHHHHHHHHHH-WWWWWWWWWooWWWWWWWWWW-HHHHHHHH
     1: HHHHHHHHHHHHHHHHHHH-WWWWWWWWWWWWWWWWWWWWW-HHHHHHHH
```

Shift: +2, Edit distance: 2 → **Similar**

Working example of similarity between two day sequences.

At least 20 similar activity sequences have to be observed across all devices in order for a given sequence to be included into both a training set and the common library used for gap filling. Elimination of rare day sequences generally results in a lower trip rate than initially observed, an unavoidable consequence needed for privacy protection.

(2) Location Choice Model (LCM)

The location choice model is trained per device to learn and predict travel destination location choices for discretionary activities (i.e., not home/work/school) made by the device owner. The LCM operates at a disaggregate level, selecting individual privately-owned businesses (e.g, individual businesses, shops, services) and other venues (e.g., parks, places of historic interest, tourist attractions) as potential destinations. It is a model that, for a given set of alternative destinations represented as attributes of particular venues in the area, ranks them based on the observed choice made by the owner of the device, considering every other location ever visited by a person as an alternative. Every trip made for a discretionary activity observed becomes one training sample for the LCM for a given person. The LCM includes contextual variables such as distance to home, distance to work, a score describing the deviation from a regular home-work commute, the hour of the day, day of the week, allocated travel time, and duration of the next scheduled activity. For destinations located in dense commercial and urban areas when an exact venue visited by a device cannot be observed, the training sample is constructed from a combined set of attributes (using the same contextual variables and median values of accessibility variables).

(3) Mode Choice Model (MCM)

The mode choice model consists of two distinct components. The first is a mode inference model that assigns a likelihood score to each mode that could be used for every observed trip. It is a discriminative machine learning model with a strong prior based on geographical location of origin and destination, accessibility, and observed travel times and distances. It is often impossible to distinguish the mode from the observed attributes of a trip, particularly in dense urban areas with significant congestion, where a pattern of biking, taking a bus, riding in a taxi, or driving a private vehicle could have the same speed profile. The mode inference model is additionally conditioned on the context of the trip within a day (e.g., which tour the trip is a part of and what modes are not feasible), as well as the pattern of modes taken on recurrent trips for the same origin and destination, such as daily commute to work. Modes that are detected with high confidence are considered to be revealed preferences and serve as the inputs for the next stage.

The second component is a mode choice model that gives the likelihood of choosing a given mode, as a function of the alternatives and the inferred attributes of the decision maker. Given an origin and a destination with respective start and end travel times, routes are computed for each possible mode (private auto, walking, bicycling, public transit with different access/egress types, and on demand auto) and are included into a choice set for a given trip. Choice model structure is based on machine learning with a utility function specification that includes a standard set of attributes (travel time, cost, in-vehicle and out-of-vehicle times, walking duration, inferred age and household income, vehicle ownership) and an additive alternative specific constant (ASC).

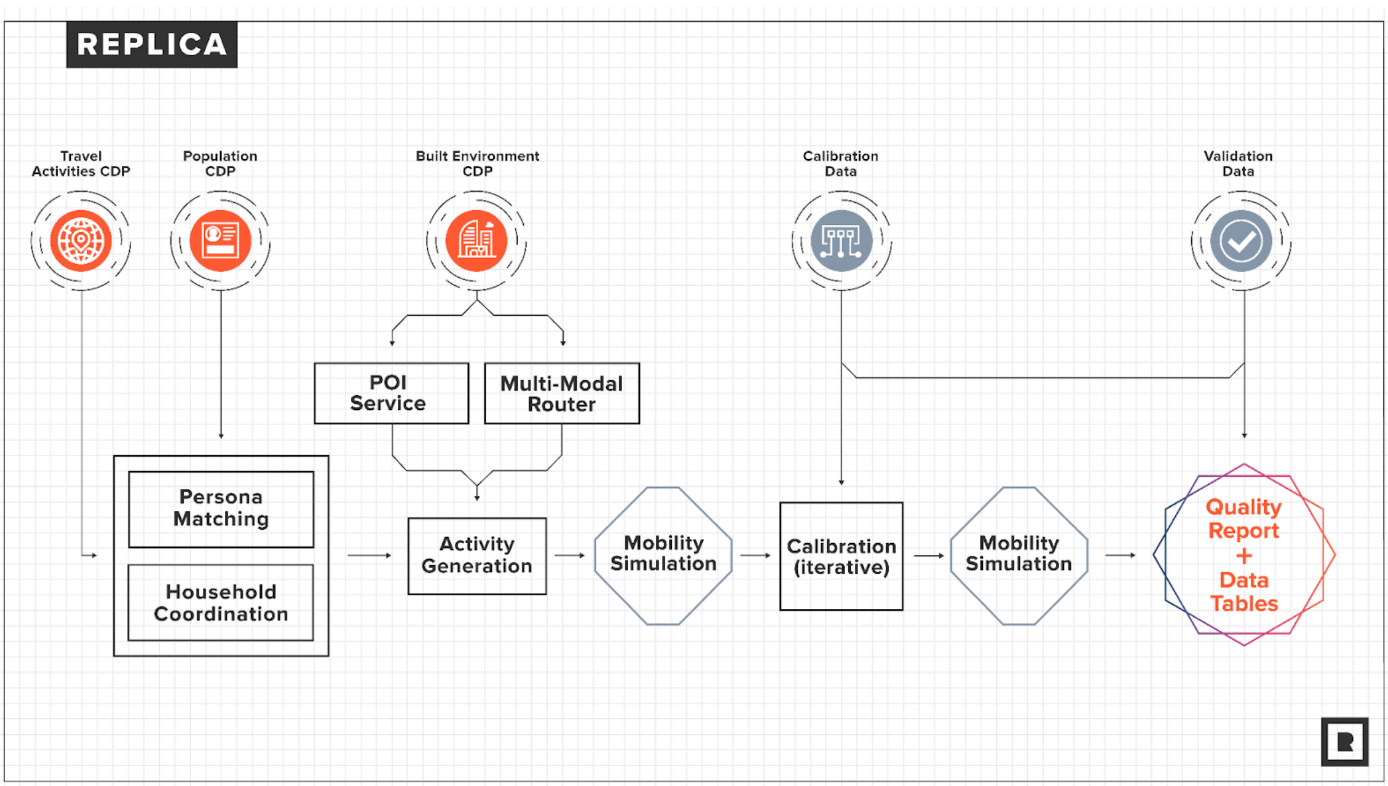
Special Cases in the Travel Activity CDP

Since Replica does not acquire mobile location data for minors, children and students are represented with a simplified model. Residents below the age of 5 do not receive a persona. All their travel is assumed to be represented by the travel of an accompanying adult from the household.

K-12 student personas are constructed synthetically, with the only daily activity being a trip to school. College student personas can represent students who are partly employed, and attend school and work on the same day. College student personas also have variability in class start and end times.

Model Creation

To create an individual seasonal mobility (Places) model, the CDPs are used to produce a simulation of a typical day of all movement in the modeled megaregion and season, leveraging the observed data in an accurate and privacy-preserving way. This synthetic representation is created using a 3-step process: (1) Persona Matching; (2) Activity Generation; (3) Mobility Simulation, each of which is described in detail below.



Model creation diagram

Following the 3-step process, the synthetic representation is calibrated and validated to observed travel metrics relevant to the given day of the modeled season. The calibration process is described in the later sections.

Persona Matching

The first step in the model creation process is to “motivate” a synthetic person to travel by matching that resident with a persona. For a given model, each synthetic person that “lives” within the megaregion boundaries is matched with a persona from the same geographic area and season — in essence, joining the Population and Travel Activity CDPs.

Residents aged 18 and older are matched with either working or non-working personas, depending on their employment status. A working synthetic resident is matched to a working persona that closely matches its home and work locations. Alternatively, a non-working resident is matched to a non-working persona that is found near its home location. This is done by weighting the personas using their inferred night-time locations as compared to the residential population (all persons aged over 7 years not living in group quarters) given by the ACS 5-year estimates.

Iterative proportional fitting methods are used for reweighting, following a methodology similar to those used in household travel surveys. Personas from areas with a lower device coverage are weighted higher than ones from areas with higher coverage. Candidate personas are then ranked by a weighted combination of the proximity of their home and work locations. The top 25-50 matches based on the proximity score (with exact limit randomly chosen each time) are returned. A persona is then chosen at random from the resulting set. School-aged residents (5-18) and college students are matched with a special student persona. Persona matching is performed sequentially for all members of the household.

Generally, for each specific megaregion and season, the expected ratio of adult residents to available personas is approximately 8:1. That is, if a megaregion has a synthetic population of 20 million adults, there would be approximately 2.5 million potential personas.

Visitors are generated by identifying personas that have activity in the modeled megaregion, but for whom the projected home location is outside the megaregion. Visitors are scaled using a composite of data sources, including hotel occupancy rates, and then included in both the Activity Generation and Mobility Simulation steps.

Activity Generation

Following the joining of synthetic persons and personas, the activity sequence model, location choice model, and mode choice models are each applied to generate activity. The activity sequence model is applied at the beginning of the simulated typical day, assuming the day starts at home for all residents. Location choice and mode choice models are sequentially applied for each tour and a trip of the simulated day.

Generation of Activity Sequence Model (ASM)

First, the ASM returns day sequences sampled from the trained model. ASM is applied sequentially for all the adult members of the household. For employed adults in the households it includes deterministic coordination of the morning commute in terms of departure times and escorting a student to school. School-aged students matched with a student persona travel to school only either in a carpool or “other” mode.

It is at this point in the model creation process that a work-from-home determination is made for each employed adult in the model. Whether or not an employed adult commutes to work or works from home on the modeled day is determined using a composite model that factors in industry of employment, home location, and the aggregate behavior of device personas that share the same home Census Public Use Microdata Area (PUMA) during the modeled season.

A typical trip rate produced by the standard version of an ASM is 3.7 trips (4.7 activities) per person per day for the nationwide sample of composite LBS/cellular data. This number varies geographically depending on the region, community type, and other factors. Region-specific trip rates are available in each megaregion’s Quality Report.

Generation of Location Choice Model (LCM)

Locations of primary activities (home, work and school) are assigned at the population synthesis stage and are not chosen within the simulated day.

For every discretionary trip, the LCM is queried with the synthetic person’s current (in simulation time) day-of-week, time-of-day, distance-from-home, distance-from-work, and commute deviation scores to rank accessible destinations for the planned activity type. The highest-ranked location is predicted by the LCM, but is not selected directly. Instead, it is used as a center of a spatial query to a POI service, which itself utilizes the Built Environment and Economic Activity CDPs. The POI service returns between 10 and 25 nearby venues of the matching category (exact count is randomly chosen). Every venue in the vicinity of a location returned as a candidate destination has an assigned popularity score computed from a combination of POI visits and credit transactions data, specific to the day of the week and the hour of the day when a visit takes place. Laplacian noise of variance, which is a function of a persona’s total number of observed visits in a day, is added to the raw place visit counts before this popularity score is computed. From this return set, a single venue is sampled based on the weighted popularity score (which may or may not be the original POI predicted by the LCM) and is assigned as a travel destination.

Generation of Mode Choice Model (MCM)

Following the assignment of a travel destination, the mode choice model is applied at the tour and trip level, generating an array of available route options between the current origin and the planned destination, based on the state of the transportation network at that time. The set of routes include an available set of transit options (utilizing all relevant agencies from Replica’s GTFS database) as well as multiple driving routes, with the travel times accounting for the expected congestion along each alternative.

Commute mode choice is governed by a separate model specifically trained on commute trips as described above. Tactical choices of modes for the trips within the tours are restricted based on the availability of a private vehicle (both car and bicycle) on the tour. It is assumed that a personal vehicle is accessible from any other destination accessed by walking from the last destination accessed by vehicle. Parking availability and capacity limits are not enforced in the current version of the model.

Mobility Simulation

Representation of all travel in the region on a typical day is derived with an agent-based simulation approach where every synthetic person in a household is engaged in travel to perform the activities at locations and times as predicted by their respective persona models. Interactions within the realized travel itineraries such as traffic congestion are accounted for in simulation runtime as described below.

Traffic Assignment

Routing algorithms in Replica use a combination of the observed routes from mobile location data and observed link-level speed collected from in-dash GPS, and a version of dynamic traffic assignment using link-level representation of congestion (traffic state is assumed to be the same along the link). Vehicle trips are assigned to the network incrementally for each vehicle, with subsequent drivers departing earlier or later than originally planned and/or rerouting in response to the developing congestion in the bottlenecks. Routing is performed on the OpenStreetMap network which is algorithmically verified for the consistency of attributes (number of lanes and speed limits) and is processed for ensuring routability.

Observed traffic volumes and routes are used in the network attributes verification process by ensuring the road types, number of lanes, speed limits, and turn restriction attributes are consistent with the traffic volumes observed on a typical day in a given season. No simplifications are applied to the network in terms of the categories of the roadways available for traffic, walking, and cycling.

Route choice and link performance functions are data-driven, and informed by the routes and travel speeds as observed in the in-dash GPS probe data. For instance, the true vehicular capacity of road segments is estimated replicating the Highway Capacity Manual speed/flow approach for every major road in the country that reaches capacity within the data available. These capacities are fed into the link performance functions and used to make re-routing decisions as the simulation progresses and congestion builds up.

The observed state of the network for a typical day is used for initial route planning, and the state of traffic on the network is stored for all the links and is updated in simulation time. The modes that interact in congested traffic are all vehicular modes (e.g., private autos, TNCs, and commercial vehicles) with the exception of public transit vehicles.

All drivers are assumed to be completely and equally informed of the state of the congestion and having access to the same routing service returning multiple distinct route alternatives per route choice request. Vehicle trips start and end at a road link adjacent to the destination, as opposed to, for example, a parking lot or a parking structure nearby. In contrast, real drivers tend to circle for parking, miss turns, and drive between different entrances of large destination areas such as the multi-lot parking areas at shopping plazas. In reality, drivers may take a non-optimal exit from the parking lot, make a U-turn, or any number of unusual actions that are not represented in routing vehicles in simulation.

Transit

Transit routing is based on the regional transit system represented by the set of the General Transit Feed Specification (GTFS) schedules valid during the simulated season in the modeled megaregion. The complete set of GTFS data used in a given season are available for download for reproducibility.

In simulation, transit stops can be accessed by walking, biking, carpooling and private auto over the same network as other motorized and non-motorized movements. Egress modes include walking, carpool, on-demand auto and private auto. Park and ride (P&R) behavior is implemented in the current version of Replica. Transit route options that include an ingress leg of private auto are considered when modeling inbound commute tours. These alternative routes are only added for those transit stations with available parking. The new alternatives are then included as part of the set of options returned by the router.

The transit router returns multiple options for every trip given the departure time window of 15 minutes from the planned departure time. Transit route choice is implemented as a machine learning-based model that includes walking distance, waiting time, number of transfers, in-vehicle travel time, and an estimated fare as attributes for every alternative. Additive constants representing idiosyncratic route preferences can be adjusted in calibration.

On-demand Auto (Taxi and TNC)

On-demand auto is one of the available modes in the option set. Mobile location data alone does not allow reliable identification of trips made by an on-demand auto (vs. for example a carpool). The choice of an on-demand option is usually a function of the limited option set of modes available to reach the observed destination on the tour/trip, with lack of access to a personal vehicle. It is preferable to use auxiliary ground truth data sources to represent travel by taxi and TNCs more accurately. On-demand auto trips with a passenger are a part of the traffic flow that contribute to congestion.

Non-motorized Travel

Non-motorized travel (walking and biking) is represented on the same street network as auto travel. Non-motorized travel restrictions are encoded as attributes of shared-mode streets, as well as dedicated sidewalks and bike paths. Travel speeds are assumed constant and do not vary by characteristics of travelers (e.g., age) or network (e.g., width of sidewalks), with travel time as the objective function in shortest path algorithms.

Origins and destinations of travel are set as the centroids of the building footprint traveled to/from, and the travel is considered started/completed from/at the travel link that is adjacent to this centroid. The accuracy of the representation of non-motorised travel is prone to the completeness and accuracy of the network, errors in the network attributes, as well as idiosyncratic route choices made by the travelers that might not be observable or representable due to the data limitations. At this time, Replica also only includes purposeful non-motorized travel in its model. Recreational trips (such as jogging for a workout, or a walk around a neighborhood or park that starts and ends at home) are not included.

Commercial (Freight) Travel

Commercial vehicle traffic is represented in the simulation and outputs of the model. Commercial fleets of medium and heavy vehicles are modeled by using a different statistical weighting methodology; while freight vehicles contribute to congestion, their travel is not subject to rerouting due to behavioral response to expected delays.

Pass-Through Traffic

Pass-through traffic is represented in the simulation and outputs of the model. Pass through (external-to-external) trips are a weighted sample of the devices that travel through the simulated region. The weighting methodology is similar to the one used in persona weighting and is aimed to represent the variability of the spatial coverage of the mobile location data sample. Carpool formation in the vehicles of pass-through traffic is not modeled and the occupancy rate is assumed fixed at 1.5 persons/vehicle. Routes of pass-through travel closely follow empirically observed routes (assuming they are the revealed route preferences in response to congestion). One reason Replica produces models at megaregion scale is to minimize the amount of pass-through traffic.

Model Calibration and Quality Reporting

After each individual simulation run, the modeled outputs are compared to aggregate control group data (i.e., observed counts, or “ground truth”) for quality and reporting purposes. This calibration process involves solving a set of large-scale optimization problems with an objective function defined as “fit to observed ground truth.” A careful balance is struck to ensure that the calibration algorithms do not overfit the modeled outputs to the calibration data, as both outliers and a certain level of noise is often present in every dataset.

Each completed seasonal mobility model includes an associated Quality Report that displays a comparison of modeled outputs to ground truth data, enabling users to compare model outputs to observed counts.

Customers have the option to supplement this process with their own ground truth data. However, it is important to note that the inclusion of customer-provided ground truth for calibration is neither a requirement nor a dependency to delivering a seasonal mobility model. Nor does the absence of customer-provided ground truth meaningfully impair the quality of a modeled season.

The Calibration Process

The calibration process compares modeled itineraries and the resulting travel volumes to observed counts, incorporating all ground truth available for the specific megaregion and season.

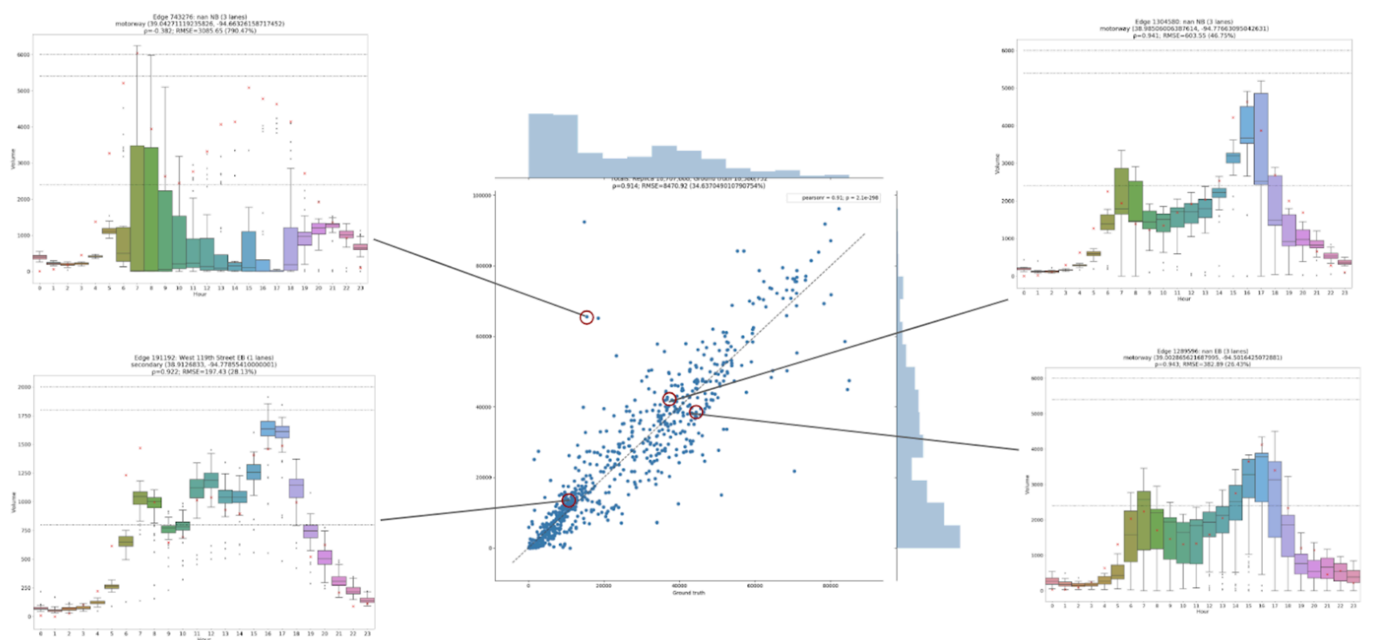
Traffic data, transit data, and non-motorized travel data can all be incorporated into the calibration process. Data collected from induction loop sensors, toll gantries, computer vision, turnstile and ticket sales, tap in/out systems, as well as manual counts, can be used.

Data Verification and Processing

The first step in the calibration process is to verify all data for consistency, identify the presence of gaps and outliers, and process the data to represent a typical day of the week in the modeled season. For example, average traffic volumes/ridership over all Wednesdays within a season are used to represent a typical Wednesday. Data outliers are identified and removed both algorithmically and manually in the cases of significant uncertainties. When day-specific information is not available, a typical weekday and weekend information is compiled in a similar manner.

Replica can process a wide range of ground truth. For traffic flow calibration, hourly volumes are preferred but not required. Zone-to-zone movements can be used for the calibration of the utility function for any given pair of planned trip origin and destination zones. For transit calibration, Replica aims to use the total number of boardings (ridership) per individual transit line. When little ground truth is available in a specific region, an estimation of trip volumes by a direct demand model trained on ground truth in other regions can be substituted.

Calibration algorithms are particularly sensitive to ground truth data quality. Every sample introduced into the calibration system is validated for the feasibility of its data. Faulty sensors, which might throw off the calibration accuracy, are identified and removed in a cascade of algorithmic data quality and feasibility filters. For example, in the figure below, there’s an outlier in the upper-left data point where the modeled count is significantly higher than the field count. The expanded upper-left diagram shows a significant day-to-day variance, which indicates this sensor is unreliable and should not be used.

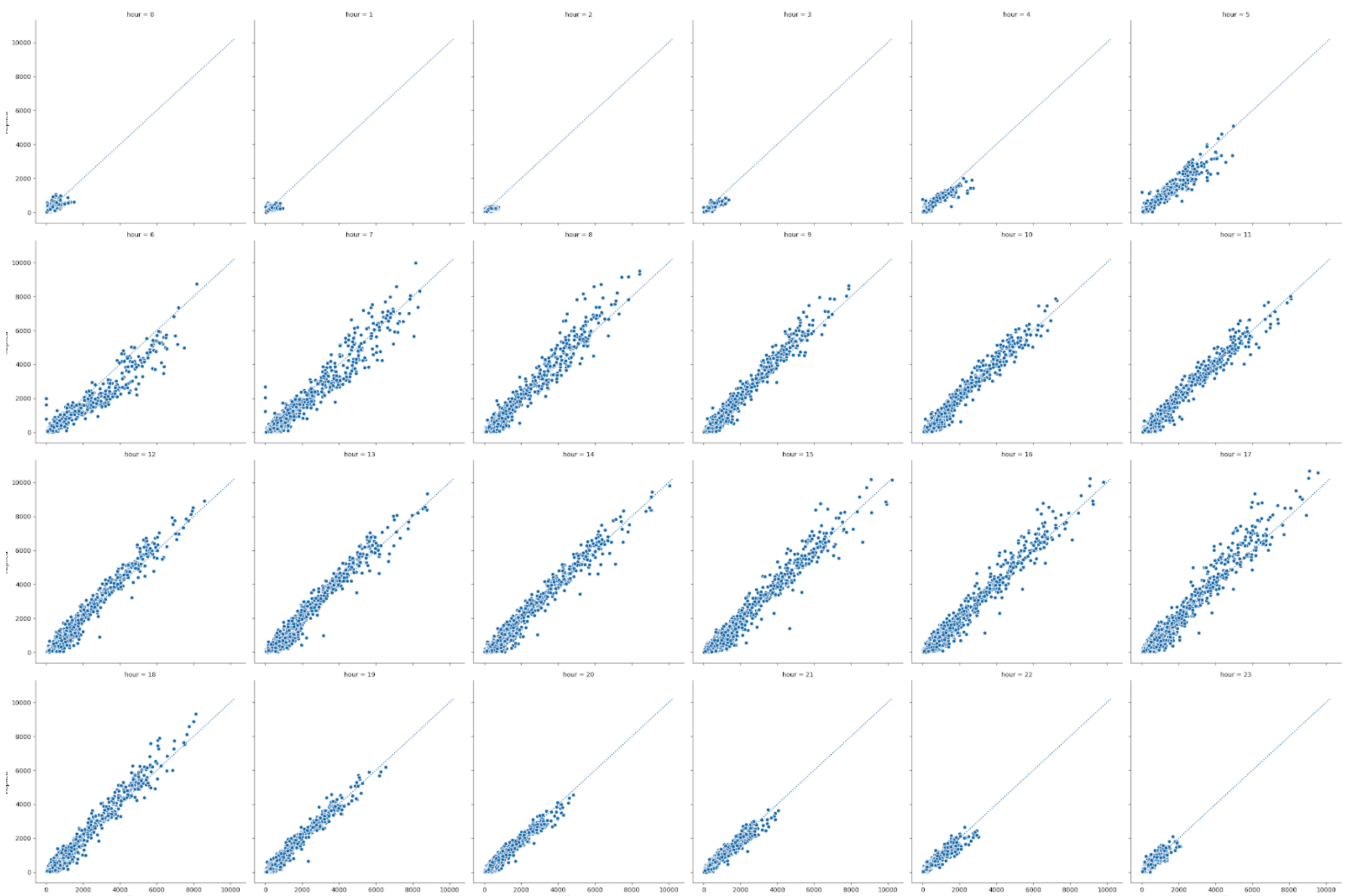


Example traffic count calibration scatter plot (center graphic) for an hour window. The X-axis is field count; the y-axis is modeled count. Inlet diagrams represent hourly volumes in a 24-hour period with an empirical variance and outliers presented as box plots at each hourly bin.

Running Calibration

Once all the faulty counts are removed from the calibration set, the calibration process is run. Choice parameters are adjusted in the calibration process depending on the ground truth data availability. They include route choice preference constants (i.e., tendency to choose a longer route with less experienced delay due to recurrent congestion), carpool choice preference, transit route choice parameters as a function of transit type and particular transit lines in an available transit option (i.e., light rail might be more attractive than a similar bus option), and TNC choice parameters (as a function of observed TNC zone-to-zone trip volumes).

In the calibration process, the algorithms are controlled to not overfit the calibration data. Despite the best attempts at validating ground truth data consistency and accuracy, both outliers and a certain level of noise are often present in every dataset. There is a delicate balance between the accurate fit to the ground truth data and the internal consistency of the simulation. There are often cases where the algorithms accept significant differences in certain counts if the modeled data represents a more consistent and balanced system state across other related metrics. These differences are typically resolved in favor of the model consistency and the observed data artifacts are disclosed to the customer.



Simulated hourly traffic volumes as compared to ground truth over a course of 24h period.

Quality Reporting

The final step in the seasonal mobility model deployment process is the production of a Quality Report, which both summarizes the outputs of the model and compares the modeled outputs to Replica-sourced and customer-provided ground truth. In the spirit of transparency and measurability, Quality Reports are always made available to Replica customers for each available season.

The Quality Report for each megaregion / season combination includes two types of information. The first is summary statistics, such as trip rate (both overall and by specific cohorts), mode and purpose breakdowns, and transit ridership counts. The second is a comparison of modeled outputs to ground truth. Root Mean Square Error (RMSE) is calculated, and plots and tables are provided in the Quality Report, for a number of key metrics including:

- Commute departures by county
- Auto volumes by hour and by road volume

- Transit ridership by transit mode, transit agency ridership, individual transit agencies, and by route
- Taxi/On-Demand volumes by origin tract

Quality Reports for each megaregion are available in our Help Center here. Example reports can be made available to prospects by request.

Data Outputs

The methodology detailed above outlines Replica's approach to running large-scale, computationally intensive simulations. It is this approach that enables Replica to deliver granular data outputs that match behavior in aggregate, but don't surface the actual movements (or compromise the privacy) of any one individual.

As such, the output of each simulation is a complete, disaggregate trip and population table for an average weekday and average weekend day in the modeled season (e.g., Fall 2021). The model represents a 24-hour period with second-by-second temporal resolution, and point-of-interest-level spatial resolution. Each row of data in the simulation output reflects a single trip, with characteristics about both the trip (e.g, origin, destination, mode, purpose, routing, duration) and trip taker (e.g., age, race/ethnicity, income, home location, work location).

The disaggregate nature of the data means that data can be filtered, and cross tabulated, by characteristics of either the trip or trip taker (or both). Specifically, the seasonal mobility model development process results in the generation of three specific output tables:

- People Table
- Trip Table
- Route Table

Data from these three tables can be joined using a combination of Trip ID, Person ID, and Route Segment ID.

Updated 2 months ago

← Seasonal Mobility Model Methodology Summary (Places)

Weekly Mobility Model Methodology (Trends) →

Appendix B: Modal Substitution Rate Methodology

MEMORANDUM – Confidential Business Information (CBI)

To: BCA Reviewers

From: Grace Young, Rohan Oprisko, Mike Sellinger, and David Wasserman, Alta Planning + Design

Date: April 1, 2022

Re: Modal Shift Model Notes

Modal Substitution Rates: Introduction

Modal substitution rates refer to the percentage of users of a facility who substituted one mode for another (Volker et al. 2019). These rates are often determined from survey instruments asking about alternative modes. When users substitute a carbon-free mode like biking for a carbon-intensive mode like driving, there is an associated emissions savings, proportional to the length of the trip. The following model provides a means for estimating the percentage of future facility users that will substitute a carbon-free mode in place of driving. This serves as a crucial step in identifying reductions in vehicle miles traveled and the emissions-saving benefits of the proposed facility.

Methodology

A series of univariate regression models were tested on peer-reviewed auto-to-bike substitution rates for projects in 10 cities around the United States. Six variables were collected at the city level and tested as inputs in a univariate regression model predicting the modal shift factor using an ordinary least squares regression from the statsmodels Python library. The variables are described in Table 1. The same variables were also tested in predicting the natural log of the modal shift percentage.

Data Review

Table 1. Peer-reviewed auto-to-bike modal shift factor and six demographic variables reported for the respective project cities[1]

City	Modal Shift (ratio)	Population Density (people per sq. mi.)	Median Income ($)	Travel Time to Work (min.)	% of Trips <4 Miles (ratio)	Active Mode Split (ratio)	Bike Mode Split (ratio)	Source
Los Angeles, CA	0.109	8,092	62,142	32	0.471	0.147	0.030	Matute et al. (2016)
Denver, CO	0.237	3,923	68,592	26	0.531	0.251	0.015	Piatkowski et al. (2015)
Boulder, CO	0.571	3,948	69,520	20	0.652	0.283	0.045	Piatkowski et al. (2015)
Littleton, CO[2]	0.724	3,215	76,105	26	0.512	0.254	0.060	Piatkowski et al. (2015)
Sacramento, CA	0.273	4,764	62,335	26	0.437	0.195	0.090	Piatkowski et al. (2015)

City	Modal Shift (ratio)	Population Density (people per sq. mi.)	Median Income ($)	Travel Time to Work (min.)	% of Trips <4 Miles (ratio)	Active Mode Split (ratio)	Bike Mode Split (ratio)	Source
Davis, CA	0.250	6,637	69,3709	23	0.636	0.220	0.095	Piatkowski et al. (2015)
Austin, TX	0.146	2,653	71,576	25	0.502	0.179	0.016	Monsere et al. (2014)
Chicago, IL	0.374	11,841	58,247	35	0.598	0.377	0.070	Monsere et al. (2014)
Portland, OR	0.202	4,375	71,005	27	0.538	0.267	0.027	Monsere et al. (2014)
San Francisco, CA	0.263	17,179	112,449	34	0.547	0.245	0.060	Monsere et al. (2014)
Washington, DC	0.202	9,856	86,420	31	0.564	0.311	0.018	Monsere et al. (2014)

Notes:

min. : minute

sq. mi. : square mile

1. Adapted from Volker et al. 2019.
2. Littleton, CO, was removed as an outlier in this modeling exercise for both final models.
3. All sources can be found in the Volker, J et. al (2019) paper specified in the references section.

Results

We found two acceptable models for contextual estimation of modal substitution rates given the available data: the examination of short trips (under 4 miles) and the active mode split model. Alta's preferred model is the examination of short trips due to its theoretical consistency with the idea that short trips are indicators that a higher proportion of vehicle trips can be converted to active modes given improved infrastructure and support. Alta uses the active mode split model depending on the available data sources on a given project or for sensitivity analysis to generate a conservative estimate.

Correlation and R-Squared

Table 2. Variable performance in correlation test and ordinary least squares univariate regression

Variable	Source	Correlation with Modal Shift	Correlation with ln (Modal Shift)	Adjusted R-Squared Predicting Modal Shift		Adjusted R-Squared Predicting ln (Modal Shift)	
				No Constant	With Constant	No Constant	With Constant
Population Density	Census	-0.21	-0.11	0.411	-0.063	0.663	-0.098

Variable	Source	Correlation with Modal Shift	Correlation with ln (Modal Shift)	Adjusted R-Squared Predicting Modal Shift		Adjusted R-Squared Predicting ln (Modal Shift)	
				No Constant	With Constant	No Constant	With Constant
Median Income	Census	-0.01	0.03	0.689	-0.111	0.813	-0.110
Travel Time to Work	Census	-0.32	-0.30	0.653	0.001	0.864	-0.014
Percent of Trips Under 4 Miles	Replica Places (2022)	0.31	0.41	0.744	-0.005	0.805	0.076
Active Mode Split (all trips)	Replica Places (2022)	0.39	0.53	0.763	0.057	0.709	0.200
Bike Mode Split	Replica Places (2022)	0.32	0.43	0.654	0.003	0.479	0.090

Note:

All values reported in this table are for models without the Littleton, CO outlier removed.

Linear Relationship Plots

Figure 1 and Figure 2 show the linear relationship between the log of modal shift and the percentage of trips less than 4 miles or active mode share, respectively. Littleton, CO, is identified as an outlier in both cases and thus removed for the final model development.

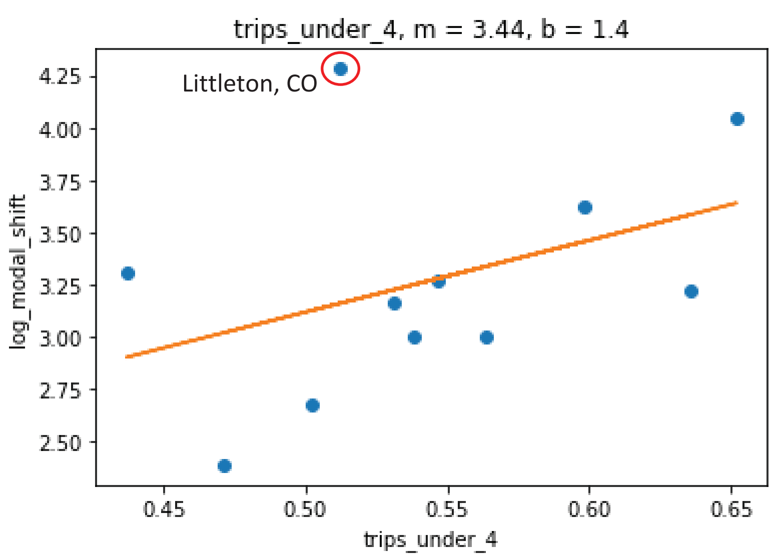


Figure 1. Modeled Relationships Between the Percentage of Short Trips and the Log of Modal Shift

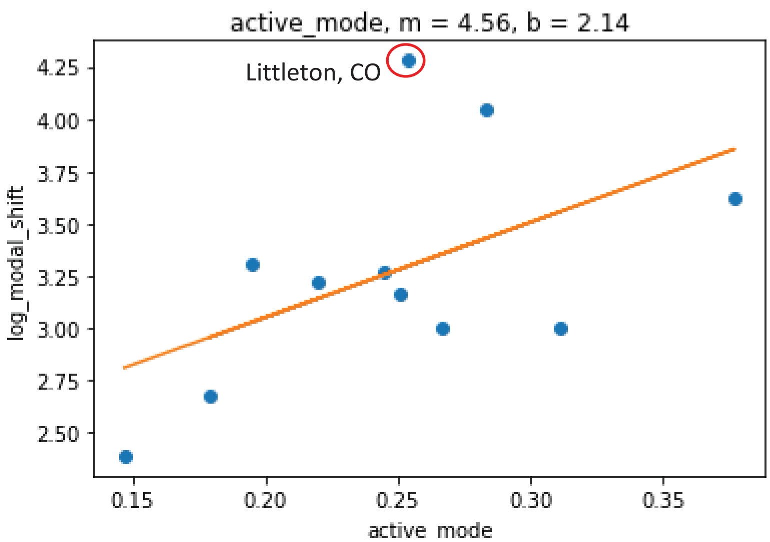


Figure 2. Modeled Relationships Between Active Mode Share and the Log of Modal Shift

Final Model Summaries

The two acceptable models are summarized in Table 3, along with the derived equations for applying each to a project-specific context.

Table 3. Model summaries for acceptable final models

Dependent Variable	*Log modal shift percentage*		Dependent Variable	*Log modal shift percentage*	
R-squared	*0.424*		R-squared	*0.414*	
Independent Variable	Coefficient	P-Value	Independent Variable	Coefficient	P-Value
Percent of trips under 4 miles	4.39	0.041	Active mode share	1.85	0.045
Constant	0.77	0.462	Constant	2.08	0.002
Equation			Equation		
ln(modal shift %) = 0.77 + 4.39*(% trips under 4 miles)			ln(modal shift %) = 2.08 + 1.85*(% active mode share)		

Discussion

These models enable a flexible and actionable approach to provide context-sensitive estimates of potential modal substitution rates given investments in multimodal infrastructure that are suitable for transportation planning practice. This approach aligns well with the understanding that compact, mixed-use locations with small urban footprints and high destination access encourage shorter trips and active travel (NASEM 2014). These models provide a decision-support tool to make informed and context-sensitive assessments of potential modal substitution rates given a project study boundary. Understanding how much reduction in vehicle miles traveled is possible given investments in active transportation is relevant to choosing a quick and responsive model.

However, there are limitations to this approach worth considering:

- While significant relationships were identified between these variables and modal substitution rates from literature, they are based on small sample sizes and depend on the removal of outliers.
- These models are not using any control variables. These univariate linear regression models are intended to enable quick determinations of possible modal substitution given a specific built context. While other variables such as population density or travel time to work were evaluated, they were not used as controls within the same model.
- Many other factors can influence rates of modal substitution beyond those identified here, and they warrant further study. It is highly complex result of localized intercept surveys, but their ranges from literature benefit from a context sensitive approach for analysis.

References

NASEM (National Academies of Sciences, Engineering, and Medicine). (2014). *Estimating Bicycling and Walking for Planning and Project Development: A Guidebook*. Washington, DC: The National Academies Press. https://doi.org/10.17226/22330

Volker, J., S. Handy, A. Kendall, and E. Barbour. (2019). *Quantifying Reductions in Vehicle Miles Traveled from New Bike Paths, Lanes, and Cycle Tracks: Summary Report.* California Air Resources Board (CARB). March 25, 2019.

Replica Places (2022). Replica Platform. Retrieved from https://replicahq.com/

CBI Rationale

These regression equations are the result of internal R&D at Alta and represent a data-driven approach to identifying realistic modal substitution rates given contextual information about a project area. Disclosure of these models before they can be further published in peer review research represents a disincentive for firms to advance research and development to advance context sensitive practice. This research was based on Alta Planning + Designs proprietary know-how and understanding of active transportation research and available data resources to inform them.