ORIGIN-DESTINATION ANALYSIS OF TWO BIG DATA PROVIDERS:
REPLICA AND STREETLIGHT DATA

Zhuo Yang
TPB Transportation Data Analyst

Travel Forecasting Subcommittee
September 20, 2024

National Capital Region
Transportation Planning Board

Agenda Item 4

Background

- The TPB has traditionally relied on **local jurisdictions** for traffic volume and travel speed data. Recently, **big data** have become instrumental in enhancing travel monitoring to understand travel trends and performance.

- TPB currently uses big data obtained from CATT[1] lab to analyze congestion conditions in the TPB planning region. Big data has the potential to be **beneficial to numerous TPB projects.**

- When comparing analysis results from Replica and StreetLight, we hope to better understand (1) the **state of the practice** in applying big data in transportation planning and analysis and (2) **which big data products** have the potential to meet COG/TPB's programmatic needs.

1. CATT lab: University of Maryland Center For Advanced Transportation Technology

National Capital Region
Transportation Planning Board

Metric and Data Coverage

Replica:

Data: Trip counts,

Modes: all modes available

Temporal coverage: A typical Thursday and Saturday of spring and fall from 2019 to 2023

Geographic coverage: Entire TPB Modeled Area

StreetLight:

Data: StreetLight All Vehicles Volume (Vehicle Trips),
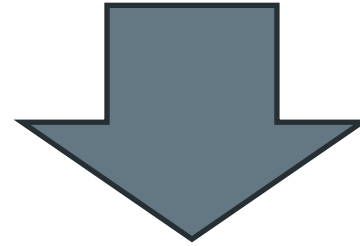
Modes: all modes available

Temporal coverage: 2018 – 2021 (Jan – Dec)

Geographic coverage: TPB Planning Area

Construct Best Apples-to-Apples Comparison

	Replica	StreetLight
Differences	• Only **Thursday** available for selected seasons • Person trips	• More temporal coverage • Vehicle trips

	Replica	StreetLight
Solution	**Analysis year:** 2019, Fall (Sep.- Nov.) **Time period:** entire Thursday **Selected Modes** to represent vehicle trips: "Private_auto", "on_demand_auto", and "Commercial"	**Analysis year:** 2019, Fall (Sep.- Nov.) **Time period:** entire Thursday **Mode:** All Vehicles LBS+

National Capital Region
Transportation Planning Board

Methods (Replica)

Platform Limitations: summaries are grouped by either O or D. No way to disaggregate to O-D pairs.

Solution:

1. Retrieve trip table for the TPB planning region. (requires 8 hours to retrieve data)

2. Spatial join trip table with TPB shapefile; Assign origin and destination jurisdictions according to origin and destination coordinates.

3. Group by origin and destination jurisdictions and sum up trip counts.

National Capital Region
Transportation Planning Board

Output (Replica)

Segment trip table by 1M
Total size: 14.7G
Total # of trips: 21M

Name	Date modified	Type	Size
2019-Q4_trips_in_tpb_1000000_2000000.csv	6/15/2024 10:34 AM	Microsoft Excel C...	742,215 KB
2019-Q4_trips_in_tpb_2000000_3000000.csv	6/15/2024 10:51 AM	Microsoft Excel C...	749,888 KB
2019-Q4_trips_in_tpb_3000000_4000000.csv	6/15/2024 11:07 AM	Microsoft Excel C...	763,698 KB
2019-Q4_trips_in_tpb_4000000_5000000.csv	6/15/2024 11:25 AM	Secrosoft Excel C...	755,702 KB
2019-Q4_trips_in_tpb_5000000_6000000.csv	6/15/2024 11:41 AM	Microsoft Excel C...	726,645 KB
2019-Q4_trips_in_tpb_6000000_7000000.csv	6/15/2024 11:58 AM	Microsoft Excel C...	753,119 KB
2019-Q4_trips_in_tpb_7000000_8000000.csv	6/14/2024 2:15 PM	Microsoft Excel C...	741,502 KB
2019-Q4_trips_in_tpb_8000000_9000000.csv	6/14/2024 2:32 PM	Microsoft Excel C...	720,698 KB
2019-Q4_trips_in_tpb_9000000_10000000.csv	6/14/2024 2:49 PM	Microsoft Excel C...	723,673 KB
2019-Q4_trips_in_tpb_10000000_11000000.csv	6/14/2024 3:06 PM	Microsoft Excel C...	743,585 KB
2019-Q4_trips_in_tpb_11000000_12000000.csv	6/14/2024 3:24 PM	Microsoft Excel C...	726,923 KB
2019-Q4_trips_in_tpb_12000000_13000000.csv	6/14/2024 3:40 PM	Microsoft Excel C...	733,423 KB
2019-Q4_trips_in_tpb_13000000_14000000.csv	6/14/2024 3:57 PM	Microsoft Excel C...	731,467 KB
2019-Q4_trips_in_tpb_14000000_15000000.csv	6/14/2024 4:15 PM	Microsoft Excel C...	775,791 KB
2019-Q4_trips_in_tpb_15000000_16000000.csv	6/14/2024 4:34 PM	Microsoft Excel C...	777,224 KB
2019-Q4_trips_in_tpb_16000000_17000000.csv	6/14/2024 4:51 PM	Microsoft Excel C...	782,356 KB
2019-Q4_trips_in_tpb_17000000_18000000.csv	6/14/2024 5:09 PM	Microsoft Excel C...	792,767 KB
2019-Q4_trips_in_tpb_18000000_19000000.csv	6/14/2024 5:26 PM	Microsoft Excel C...	798,128 KB
2019-Q4_trips_in_tpb_19000000_20000000.csv	6/14/2024 5:42 PM	Microsoft Excel C...	796,873 KB
2019-Q4_trips_in_tpb_20000000_21000000.csv	6/14/2024 5:50 PM	Microsoft Excel C...	366,566 KB

National Capital Region
Transportation Planning Board

Methods (StreetLight)

Submit OD analysis

Output (StreetLight)

- Summary tables
- Shapefiles

Name	Date modified	Type	Size
Analysis Details	6/17/2024 2:04 PM	File folder	
Shapefile	6/17/2024 2:04 PM	File folder	
Zone Activity	6/17/2024 2:04 PM	File folder	
1756059_OD_wi_tpbplanning_by_cnty2019_veh_od_all.csv	6/17/2024 2:04 PM	Microsoft Excel C...	443 KB
1756059_OD_wi_tpbplanning_by_cnty2019_veh_od_trip_all.csv	6/17/2024 2:04 PM	Microsoft Excel C...	1,206 KB

A	B	D	I	M	N	O
Data Periods	Mode of Travel	Origin Zone Name	Destination Zone Name	Day Type	Day Part	Average Daily O-D Traffic (StL Volume)
Sep 01, 2019 - Nov 30, 2019	All Vehicles LBS Plus - StL All Vehicles Volume	Alexandria	Alexandria	0: All Days	0: All Day	113,257
Sep 01, 2019 - Nov 30, 2019	All Vehicles LBS Plus - StL All Vehicles Volume	Alexandria	Alexandria	0: All Days	1: Early AM	2,817
Sep 01, 2019 - Nov 30, 2019	All Vehicles LBS Plus - StL All Vehicles Volume	Alexandria	Alexandria	0: All Days	2: Peak AM	22,373
Sep 01, 2019 - Nov 30, 2019	All Vehicles LBS Plus - StL All Vehicles Volume	Alexandria	Alexandria	0: All Days	3: Mid-Day	38,382
Sep 01, 2019 - Nov 30, 2019	All Vehicles LBS Plus - StL All Vehicles Volume	Alexandria	Alexandria	0: All Days	4: Peak PM	33,462
Sep 01, 2019 - Nov 30, 2019	All Vehicles LBS Plus - StL All Vehicles Volume	Alexandria	Alexandria	0: All Days	5: Late PM	16,509
Sep 01, 2019 - Nov 30, 2019	All Vehicles LBS Plus - StL All Vehicles Volume	Alexandria	Alexandria	1: Thursday	0: All Day	116,442
Sep 01, 2019 - Nov 30, 2019	All Vehicles LBS Plus - StL All Vehicles Volume	Alexandria	Alexandria	1: Thursday	1: Early AM	2,909
Sep 01, 2019 - Nov 30, 2019	All Vehicles LBS Plus - StL All Vehicles Volume	Alexandria	Alexandria	1: Thursday	2: Peak AM	25,476

National Capital Region
Transportation Planning Board

OD Matrix (Replica, Weekday Vehicle Trips, 2019)

	Arlington County	Charles County	City of Alexandria	District of Columbia	Fairfax County/Cities	Frederick County	Loudoun County	Montgomery County	Prince George's County	Prince William County/Cities	Total
Arlington County	142,702	1,536	24,056	41,883	90,639	732	5,707	12,545	13,600	10,492	343,892
Charles County	1,511	172,537	1,731	6,693	6,779	157	523	1,405	38,726	2,086	232,148
City of Alexandria	23,810	1,961	93,575	21,640	85,624	275	2,844	3,997	14,846	9,321	257,893
District of Columbia	35,577	4,986	18,012	470,680	66,894	2,842	9,305	97,426	155,394	15,488	876,604
Fairfax County/Cities	92,030	7,017	80,659	70,146	1,761,039	3,258	148,154	54,109	43,825	132,328	2,392,565
Frederick County	656	151	211	2,714	3,462	412,352	5,478	37,534	3,411	609	466,578
Loudoun County	5,096	571	2,553	9,326	141,465	5,546	554,121	11,913	4,101	25,355	760,047
Montgomery County	10,614	1,145	3,615	101,532	53,427	36,564	11,877	1,364,035	119,845	5,491	1,708,145
Prince George's County	12,661	40,197	14,424	160,956	41,915	3,348	4,400	120,703	1,140,264	7,748	1,546,616
Prince William County/Cities	8,174	2,138	8,234	10,459	131,771	560	25,293	5,099	7,608	690,625	889,961
	332,831	232,239	247,070	896,029	2,383,015	465,634	767,702	1,708,766	1,541,620	899,543	9,474,449

OD Matrix (Streetlight, Weekday Vehicle Trips, 2019)

	Arlington County	Charles County	City of Alexandria	District of Columbia	Fairfax County/Cities	Frederick County	Loudoun County	Montgomery County	Prince George's County	Prince William County/Cities	Total
Arlington County	211,778	2,210	30,276	68,885	99,774	970	8,690	13,692	15,600	14,453	466,328
Charles County	2,289	235,676	2,132	13,291	4,809	161	379	1,973	41,389	634	302,733
City of Alexandria	29,876	2,211	116,442	22,576	76,180	383	2,917	4,454	14,877	9,549	279,465
District of Columbia	70,755	13,318	23,385	843,379	83,451	3,235	12,850	109,033	180,842	17,669	1,357,917
Fairfax County/Cities	99,938	4,943	75,711	80,632	2,187,055	3,969	160,637	44,959	37,413	139,902	2,835,159
Frederick County	1,042	142	365	3,223	3,850	558,228	6,002	48,368	3,004	749	624,973
Loudoun County	8,495	370	2,860	12,273	157,698	6,162	797,544	9,677	3,844	25,138	1,024,061
Montgomery County	14,368	1,952	4,365	108,114	44,991	48,926	10,392	1,841,648	99,451	5,185	2,179,392
Prince George's County	15,756	42,944	13,811	177,858	36,518	3,135	4,046	99,451	1,230,519	6,055	1,630,093
Prince William County/Cities	13,792	699	9,132	17,318	136,428	713	24,698	5,130	5,968	957,679	1,171,557
Total	468,089	304,465	278,479	1,347,549	2,830,754	625,882	1,028,155	2,178,385	1,632,907	1,177,013	11,871,678

- Trip distributions across counties are similar.

- StreetLight estimates are larger than those of Replica, but the magnitudes are similar.

National Capital Region
Transportation Planning Board

TOP OD Comparison (Jurisdiction Internal)

	Replica	2019-Q4					Streetlight	2019-Q4		
Rank	origin_geo	dest_geo	count	Column1		Rank	origin_geo	dest_geo	avg_day	Column1
1	Fairfax County/Citie	Fairfax County/Cities	1,761,039	internal		1	Fairfax County/Cities	Fairfax County/Cities	2,187,055	internal
2	Montgomery County	Montgomery County	1,364,035	internal		2	Montgomery County	Montgomery County	1,841,648	internal
3	Prince George's Cou	Prince George's County	1,140,264	internal		3	Prince George's Coun	Prince George's Coun	1,230,519	internal
4	Prince William Cou	Prince William County/C	690,625	internal		4	Prince William Count	Prince William Count	957,679	internal
5	Loudoun County	Loudoun County	554,121	internal		5	District of Columbia	District of Columbia	843,379	internal
6	District of Columbia	District of Columbia	470,680	internal		6	Loudoun County	Loudoun County	797,544	internal
7	Frederick County	Frederick County	412,352	internal		7	Frederick County	Frederick County	558,228	internal
8	Charles County	Charles County	172,537	internal		8	Charles County	Charles County	235,676	internal
12	Arlington County	Arlington County	142,702	internal		9	Arlington County	Arlington County	211,778	internal
20	City of Alexandria	City of Alexandria	93,575	internal		16	City of Alexandria	City of Alexandria	116,442	internal

- The absolute number of trips shows large difference. Replica estimates are generally less than StreetLight.
- Six out of the top ten internal trips rank the same from the two data sources.
- Top OD jurisdictions are common among the top 8, though with slightly different rankings.

National Capital Region
Transportation Planning Board

TOP OD Comparison (Jurisdiction External)

Replica		2019-Q4		
Rank	origin_geo	dest_geo	count	Column1
9	Prince George's Cou	District of Columbia	160,956	external
10	District of Columbia	Prince George's County	155,394	external
11	Fairfax County/Citie	Loudoun County	148,154	external
13	Loudoun County	Fairfax County/Cities	141,465	external
14	Fairfax County/Citie	Prince William County/C	132,328	external
15	Prince William Cou	Fairfax County/Cities	131,771	external
16	Prince George's Cou	Montgomery County	120,703	external
17	Montgomery County	Prince George's County	119,845	external
18	Montgomery County	District of Columbia	101,532	external
19	District of Columbia	Montgomery County	97,426	external

Streetlight		2019-Q4		
Rank	origin_geo	dest_geo	avg_day	Column1
10	District of Columbia	Prince George's Coun	180,842	external
11	Prince George's Coun	District of Columbia	177,858	external
12	Fairfax County/Cities	Loudoun County	160,637	external
13	Loudoun County	Fairfax County/Cities	157,698	external
14	Fairfax County/Cities	Prince William Count	139,902	external
15	Prince William Count	Fairfax County/Cities	136,428	external
17	District of Columbia	Montgomery County	109,033	external
18	Montgomery County	District of Columbia	108,114	external
19	Fairfax County/Cities	Arlington County	99,938	external
20	Arlington County	Fairfax County/Cities	99,774	external

- Origin and destination counties appear in pairs.
- PG – DC, FC – LC, FC – PW, and DC – MC are top 5 round-trip OD pairs.
- Top jurisdictions are commonly ranked with both data sources.

National Capital Region
Transportation Planning Board

Key Takeaways

- Streetlight offers more temporal coverage while Replica covers a typical day in the spring and fall seasons.

- Replica offers more detailed access to its data and enables more analytical possibilities.

- Jurisdictional internal trips account for most of the total trip counts.

- For jurisdictional external trips, OD counties tend to rank in pairs.

- The top OD pairs include trips between Prince George's County and District of Columbia, Fairfax County and Loudoun County, Fairfax County and Prince William County, and District of Columbia and Montgomery County.

Insights

- **Coverage:** confirm the data coverage and granularity before starting the analysis, both temporally and geographically.

- **Accessibility:** request for more options of accessing the data to enable more possibilities.

- **Documentation/Consultation:** learn from documentations, trainings, and webinars. Consult vendors to set up analysis parameters correctly.

- **Efficiency:** decide either using platforms or develop programs with considerations of computing capacities and efficiencies.

National Capital Region
Transportation Planning Board

Next Step

- Select a third data set as the ground-truth data to validate the analysis results from big data vendors.

- Zoom into more granular geo unit (e.g., Tract/TAZ) and evaluate the results.

Zhuo Yang

Transportation Data Analyst

(202) 962-3370

zyang@mwcog.org

777 North Capitol Street NE, Suite 300

Washington, DC 20002

mwcog.org

National Capital Region
Transportation Planning Board