APPENDIX: METHODOLOGY

Downloading the data

We project each dataset to North America Albers Equal Area Conic (ESRI:102008) before we analyze them so that distances and areas can be compared across the continental United States, Hawaii, and Alaska.

HIGHWAY PERFORMANCE MONITORING SYSTEM

The Federal Highway Administration publishes the Highway Performance Monitoring System (HPMS) annually. This product compiles submissions of public roadways from state agencies into a national dataset. Because states self-report data, reporting errors within and between states are possible.

We accessed HPMS Spatial All Sections - 2024 on February 9, 2026. To prevent downloads from failing, we downloaded data in tranches, only querying roads owned by the state department of transportation (where the ownership field is 1). We then merged all tranches into a national dataset.

CENSUS URBAN AREAS

We downloaded 2020 census urban area boundaries from the Census Bureau's TIGER/Line on April 14, 2026, dropping urban areas in U.S. territories. We then buffered these urban areas by 20 feet (6.096 meters). Many state-owned roads form the edges of urban areas, and a 20-foot buffer makes sure those are counted without inadvertently capturing others.

NATIONAL TRANSIT MAP

The National Transit Map, published by the Bureau of Transportation Statistics, is a national inventory of fixed-guideway and fixed-route transit. It covers light rail, heavy rail, subways, buses, ferries, cable trams, aerial lifts, funiculars, trolleybuses, and monorails. It excludes demand response service and paratransit.

The National Transit Map relies on General Transit Feed Specification data submitted by National Transit Database reporters, so reporting errors are possible. The Map's stops also have no information about service levels, frequencies, or headways. As a result, we cannot compare corridors by service quality, meaning that a stop with a bus arriving every hour is considered equal to a station with five-minute headways. This is an especially important topic for future research.

We downloaded all National Transit Map stops on April 13, 2026. At that point, the data were last updated on December 2, 2025. We only kept stops within the buffered urban areas.

OPENSTREETMAP

OpenStreetMap (OSM) is an open-source, volunteer-built map of the world. We downloaded OSM data for all five U.S. regions defined by Geofabrik between February 9 and 12, 2026, using the R package osmextract. Roads are stored in OSM's broader "lines" layer, so we queried only lines with a non-null highway tag. We kept roads tagged as trunk, primary, secondary, tertiary, unclassified, residential, living street, or pedestrian. We then merged the five regions into one national dataset.

We focus on those roads specifically because they make up the bulk of public streets that pedestrians and drivers can use to access adjacent land. As a result, our database excludes motorways and service roads, among others.

CENSUS BUREAU LONGITUDINAL EMPLOYER-HOUSEHOLD DYNAMICS ORIGIN-DESTINATION EMPLOYMENT STATISTICS

The Census Bureau's Longitudinal Employer-Household Dynamics Origin-Destination Employment Statistics (LODES) reports the number of jobs (workplace area characteristics) and the number of workers (residence area characteristics) in each census block.

The job counts within LODES data have known limitations. Studies have shown that jobs in rural areas may be undercounted, and jobs at firms with multiple locations may appear concentrated at headquarters.

We downloaded 2023 LODES 8.0 data on April 23, 2026, for all states except Michigan and Alaska. For those two states, 2023 workplace data were unavailable, so we used the most recent years: 2021 and 2016, respectively. We then joined LODES data to census block boundaries downloaded April 23, 2026.

REPLICA

Replica is a private data service that combines several inputs to model synthetic trips and land use. Because Replica's data are modeled and not observed, all measurements—including trips by mode—are proprietary estimates. For a more in-depth explanation of Replica data, see past Brookings research.

All Replica data are from the fall 2025 season, and trip data are for a typical weekday. We downloaded data on the following dates:

- Trip data for corridors from September 17 to 21, 2026
- Trip data for urban areas from September 11 to 14, 2026
- Land use data, by both parcel and building area, for urban areas and corridors on September 22 and 23, 2026

Replica does not model transit trips in many urban areas, so we exclude those urban areas in certain summaries. Those urban areas have a median population of about 15,000, and almost none have corridors with fixed guideway transit stops. Furthermore, trip estimates become less reliable as geometry sizes decrease. Throughout the report, we account for that by publishing findings for corridors longer than 1 mile.

MICROSOFT GLOBAL ML BUILDING FOOTPRINTS

Microsoft publishes a global dataset of buildings footprints based on imagery from Bing Maps and classified using machine learning. For many—but not all—buildings in the United States, that same algorithm estimates height as well. We discuss how we handle missing heights later in this document.

We downloaded the building data between June 30 and July 4, 2026.

FATALITY ANALYSIS REPORTING SYSTEM

The Fatality Analysis Reporting System (FARS) is a national census of all vehicle-related deaths on public roads, published by the National Highway Traffic Safety Administration.

We downloaded data for 2020 through 2024 between September 10 and 17, 2026. We used the fatalities field in the crash-level file to create five-year totals of deaths and mapped each death by its coordinates. The five-year total serves a dual purpose: It helps smooth what can be significant year-to-year fluctuations, and it helps account for unique pandemic trends in 2020 and 2021.

Deriving corridors

STEP 1: FINDING STATE-OWNED ROADS THAT CAN SERVE ADJACENT LAND

To find the universe of state-owned roads that could (transit aside) support TOD, we start with state-owned roads in HPMS and only keep roads with access to adjacent land. That means excluding interstates and other limited-access roadways that do not connect to nearby street networks and cannot access nearby development by design. We also exclude all bridges, tunnels, and causeways, which are physically separated from adjacent land. We also exclude ramps.

Though research has not found that TOD is impossible near stations in limited-access right-of-way, it routinely finds that TOD is more difficult to deliver in such locations and tends to underperform.

Table 1 lists the HPMS fields and values used. HPMS does not report every field for every road, so we note where that partial coverage exists. Lack of reporting for access control (a measure of connectivity with other roads and surrounding property) and structure type on lower functional classes is a limited issue, as most limited-access roads and major bridges are a higher functional class; we keep null values throughout so that roads with incomplete reporting aren't dropped.

TABLE 1

HPMS fields and their use within the Brookings TOD database

HPMS field	Reported for	Included	Excluded
Access control	All interstates, freeways and expressways, and other principal arterials Sample panel data for selected portions of minor arterials, major collectors, and urban minor collectors No rural minor collectors nor local roads	Partial access control (2) No access control (3) Null values	Full access control (1)
Facility type	All roadway segments and ramps	One-way roadway (1) Two-way roadway (2) Non-mainline facilities (5) Null values	Ramp (4) Non-inventory facilities (6) Planned and unbuilt facilities (3)
Structure type	All applicable roadway segments, except rural minor collectors and local roads	Null values	Bridge (1) Tunnel (2) Causeway (3)
Functional system	All roadway segments and ramps	Principal arterials (2, 3) Minor arterial (4) Major collector (5) Minor collector (6) Local (7) Null values	Interstates (1)

We calculate each segment's geodesic length in U.S. survey miles. That kind of measurement makes sure lengths stay accurate across the dataset, which stretches from Hawaii and Alaska to Florida and Maine.

From there, we intersect the HPMS data with the buffered urban areas, which tags each segment with its urban area for later summaries. Because the urban areas were buffered by 20 feet, some roadways are duplicated in the 40-foot overlap between abutting urban areas. This happens on 26.8 miles of roadways. After removing those duplicates, states own 113,375.8 miles of qualifying urban roadways.

STEP 2: FIND STATE-OWNED ROADS NEAR TRANSIT

We buffer every transit stop by a quarter-mile, then clip the roads from the previous step to those buffers. We split roadways into separate segments where gaps between stop buffers cut a single roadway into noncontinuous pieces.

We then used the HPMS route ID field to join contiguous segments of the same route. This is how we construct our "corridors." We then do a spatial join on each of these continuous segments with the transit stops, which lets us count the number of stops of each mode within a quarter-mile.

A few notes on this process:

- Overpasses and underpasses tend to be marked as bridges and tunnels. Though this is a methodological choice meant to isolate roadways that can access nearby development, their exclusion significantly increases the number of corridors.
- The quarter-mile radius captures many transit stops that do not directly serve the state road. These stops may instead serve parallel routes or nearby roads. This again is a methodological choice, as TOD is defined by a stop nearby, not on a specific road.

STEP 3: CREATING CORRIDORS

Finally, we buffer each continuous stretch of road from the step above by a quarter-mile. The resulting zones are "corridors," and that buffered area is what's analyzed in the section below. Corridor area is, specifically, the geodesic area of the corridor geometry in square miles. Overlapping corridors are common when roads are separated by a small gap or near other state roads; that can prove problematic for analysis of nominal counts, such as of roadway deaths, which we address below. The buffered radius also means that some parts of the zone can be up to a half-mile from the nearest transit stop.

Analyzing characteristics of corridors and urban areas

JOB AND POPULATION DENSITY

For urban areas, we compute job and population density by summing jobs and workers from LODES across each urban area's census blocks and normalizing by land area, as reported in Census Bureau urban area data.

For corridors, we intersect the census blocks joined to LODES data and assign each block's jobs and residents to the corridor, weighting by the share of a block's area that fills within the corridor. Corridor densities are those sums divided by the corridor's area.

This assumes a uniform distribution of jobs and population across each block. That may not be accurate in all blocks, particularly for jobs, which LODES tends to concentrate at a single address. This issue is mitigated by virtue of corridors being in more dense urban areas with smaller block sizes. It also means our estimates are relatively conservative, because jobs are more likely than housing to cluster directly along corridors.

INTERSECTION DENSITY

We use OSM roads to generate intersection densities by corridor and urban area. We first merge OSM roads by the OSM "name" field. We then intersect the roads with themselves, which generates points wherever OSM roads cross over one another. We delete duplicate points and use the resulting intersection totals to compute intersection density, normalizing by land area.

This approach can overcount intersections; roads that cross at separate grades, such as overpasses and bridges, may be included in intersection totals.

PERCENT DEVELOPED LAND AND FLOOR-AREA RATIO

For every building in the national Microsoft dataset, we systematically check whether its interior point is in each corridor and urban area. That process associates buildings and their attributes with the corridors and urban areas in which they sit. Since we report data as ratios and the buffered corridors can overlap, buildings that are within two corridors are counted fully for both corridors.

Percent developed land is the total building footprint area of all buildings in a geography, either a corridor or urban area, divided by its land area.

We calculate the floor-area ratio (FAR) of every geography by first estimating each building's floor area: multiplying its footprint by its number of floors (its height divided by 10 feet). We then divide the sum of floor area across all buildings in a geography by its land area to derive FAR.

A floor height of 10 feet is standard for residential development and some commercial real estate. In some cases, though, commercial real estate may have larger floor heights, meaning our analysis may overstate FAR. However, these estimates may very well be conservative, as the land area denominator includes undevelopable land such as the actual surface of the road forming the center of the corridor and sidewalks.

LAND USE AND THE ENTROPY INDEX

Replica estimates land use for corridors and urban areas by parcel and building area. The company derives land use from public and private data on parcel-level land use, enriched by point-of-interest data. We calculate land use shares and an entropy index for both, but we tend to focus on parcel area, because that's how travelers experience land use and reveals actual allocation of land, including undeveloped space that may be available for TOD.

TABLE 2

Replica land use categories

Land use category		
Residential single-family	Residential multifamily	Commercial retail
Commercial office	Commercial non-retail attraction	Mixed-use*
Industrial	Civic institutional	Health care
Education	Transportation utilities	Open space
Agriculture	Other	Unknown

*ONLY PUBLISHED BY PARCEL AREA, NOT BY BUILDING AREA.

We summarize land use mix for each corridor and urban area using the entropy index, which assigns a single number to describe the diversity of land use in a certain area. We use the following formula to derive the entropy index:

$$\text{Entropy Index} = -\frac{\sum_{i=1}^{k} P_i \ln\left(P_i\right)}{\ln(k)}$$

where p_i is the share of area in land use i, and k is the number of land use categories

An entropy index of 0 means the entire land area has one use, while 1 means it's split evenly across all uses. Replica notes where parcel and building area have unknown land use—though it's rare, as only 38 and eight urban areas out of 2,611 total have more than 10% of building and parcel area categorized as such, respectively. We ensure entropy indices are only calculated for known land uses by removing the unknown shares from above sums and total categories k.

TOTAL TRIPS AND TRIPS BY MODE

Replica models individual trips ending in both urban areas and corridors. We focus on trips ending in a corridor to measure its attractiveness as a destination, which TOD is meant to support.

We count all trips ending in each urban area and corridor, then normalize by square mile. Each trip has a mode associated via Replica data (Replica explains how it infers mode here, which is proprietary; they use existing modal estimates as inputs and validate their own estimates). Replica breaks those down by private automobile, automobile passenger, walking, biking, public transit, taxi, commercial, and other. We count trips by mode in each corridor and urban area, then calculate each mode's share of all trips—in the report, we tend to report trip data for corridors longer than 1 mile to ensure higher accuracy Replica estimates.

FATALITIES

We count roadway deaths between 2020 and 2024 in each urban area and corridor. To ensure corridor-specific counts are accurate, we count a death in overlapping corridor areas as in each corridor. For each corridor and urban area, we calculate the fatality rate by dividing number of deaths by total OSM road mileage.

We also count deaths anywhere in an urban area's corridor—essentially a de-duplicated version of the corridor count described above. To do that, we merge all corridors in an urban area into one feature. We use those totals to calculate what share of roadway deaths in an urban area occurred in corridors. Because corridors can, in rare circumstances, extend small distances beyond urban area boundaries, we only count deaths within the urban area.

As a point of comparison, we calculate the share of OSM road miles in an urban area that are within corridor areas. We say corridors' collective area is "disproportionately deadly" in an urban area if it has a higher share of that urban area's deaths than its roads.

Normalizing by road miles instead of vehicle miles does not account for differing traffic volumes between corridors and non-corridor roads. That's a methodological decision, though, as measuring fatalities per mile driven does not capture that design choices and TOD can reduce driving in a way that improves safety.

Comparing corridors and urban areas

To ensure that each corridor's multiple metrics were meaningfully contextualized, we compare each corridor metric to the same in its urban area. We say that a corridor outperforms its urban area when its value is higher (or lower, for roadway deaths).

Some of those comparisons have exceptions. For example, we only compare floor-area ratio when at least half of total building footprint in both a corridor and its urban area has height data. We also exclude urban areas where Replica does not model transit trips from transit mode share comparisons, and roadway death summaries exclude the 83 urban areas (all with fewer than 25,000 residents) with corridors that had zero roadway deaths between 2020 and 2024.

BROOKINGS

1775 Massachusetts Ave NW,
Washington, DC 20036
(202) 797-6000
www.brookings.edu