

Historical activity-transition rules can misallocate travel demand: Nationwide evidence from US

Yilin Zhu, Haiyang Lin, Wenhan Wu, Yuji Yoshimura, Jiayu Pan, Rui Wang, Faan Chen

Electronic supplementary material

Supplementary methods are followed by supplementary results, figures and tables.

Supplementary Methods

S1 Trip-chain representation

Each traveling person-day was represented as an ordered activity-purpose sequence, for example [home, work, shopping/errands, home]. Trips were sorted by TDTRPNUM, and destination-purpose codes were mapped to the harmonized nine-category taxonomy. A home boundary was added only when the raw first-origin or final-destination purpose code identified home and the sequence did not already begin or end at home. The same codebook-based rule was applied across waves. For example, two recorded trips Home→Work and Work→Home were represented as [home, work, home]. If the first observed destination was already Home, no duplicate initial Home state was added. Home→Home adjacent pairs therefore represented successive Home states in the constructed sequence rather than an unconditional duplication of a loop endpoint. All retained trip records, including records flagged as loop trips in 2017, remained in the primary analysis. We distinguish observed trip count from purpose-sequence length because the latter may additionally include raw-code-supported home boundaries.

Formally, for person-day i , let

$$S_i = (x_{i1}, x_{i2}, \dots, x_{iT_i}), x_{it} \in \mathcal{X}, \quad (1)$$

where $\mathcal{X}$ denotes the nine-category purpose state space and T_i is the length of purpose sequence.

Let l_i denote the observed number of trips of the i th trip chain, which refers to chain length. This distinction mattered because l_i counted observed trips, whereas T_i referred to the boundary-augmented purpose sequence used for purpose-based sequence analysis.

Secondary mode representation

As a supplementary specificity analysis, we also represented each person-day using the observed trip-level mode sequence and the corresponding observed purpose-mode pairs. These representations were not used in the primary purpose-sequence, transfer, decomposition, motif or geographic analyses. At the trip level, let

$$m_i = (m_{i1}, m_{i2}, \dots, m_{il_i}), m_{it} \in \mathbf{M}, \quad (2)$$

where $\mathbf{M}$ denotes the harmonized mode state space. Let

$$c_i = ((x_{i1}^{obs}, m_{i1}), \dots, (x_{il_i}^{obs}, m_{il_i})) \quad (3)$$

denote the observed trip-level purpose mode sequence.

Boundary-derived home states carried no observed travel mode; mode measures therefore used only observed trip-level records. Each person-day inherited survey weight WTTRDFIN. The information-measure sample comprised 622,496 traveling person-days, each with at least one retained trip. The complete age-by-worker-by-settlement-by-income sample used for Markov, decomposition and destination-purpose counterfactual models retained 108,805 person-days in 2001, 214,130 in 2009 and 189,928 in 2017, corresponding to 77.3%, 81.6% and 86.6% of traveling person-days and 74.2%, 79.9% and 82.8% of their weighted totals, respectively (Supplementary Table 10). When statistics were defined on purpose tokens, adjacent transitions or complete sequences, person-day weights were propagated through the corresponding weighted counts.

S2 Information-theoretic measures

All entropy and mutual-information measures were estimated from weighted empirical counts. Equations (4)–(9) define counts and probabilities for all-position purpose tokens, adjacent purpose pairs and purpose triplets. We denote the all-position token distribution by $p_{tok}(x)$, the adjacent-pair joint distribution by $p_{pair}(a,b)$, and its origin and destination marginals by $p_O(a) = \sum_b p_{pair}(a,b)$ and $p_D(b) = \sum_a p_{pair}(a,b)$. The primary sequential entropy gap and the strict adjacent-pair mutual-information measure use these quantities but differ in their marginal support, as specified below.

All-position marginal purpose entropy H_0 (equation (10)) is computed from p_{tok} and summarizes diversity across all weighted purpose positions. First-order conditional entropy H_1 (equation (11)) is computed from p_{pair} and p_O ; second-order conditional entropy H_2 (equation (12)) conditions on adjacent purpose pairs. The primary sequential entropy gap G (equation (13)) is $H_0 - H_1$. Because H_0 includes all sequence positions while H_1 uses successor positions in adjacent pairs, G is a corpus-level entropy gap rather than strict pairwise mutual information. Strict adjacent-pair MI was therefore computed separately as $\sum_{a,b} p_{pair}(a,b) \log_2[(p_{pair}(a,b)) / (p_O(a)p_D(b))]$.

Secondary mode-based measures.

As a supplementary specificity analysis, mode entropy and purpose–mode coupling were computed from observed trip-level mode records and purpose–mode pairs rather than from boundary-augmented purpose sequences. Equations (14)–(20) define the corresponding weighted counts, probabilities, mutual information and normalized mutual information. The normalized measure was retained because absolute mutual information can remain stable when marginal uncertainty changes. These measures were not part of the primary activity-purpose sequence stability analysis.

Full-chain entropy was computed as the weighted entropy of complete purpose sequences (equation (21)). Uncertainty for weighted entropy, G and strict adjacent-pair MI estimates was quantified with 1,000 household-cluster bootstrap resamples within survey year. Cross-year significance used 1,000 permutation tests that shuffled survey-year labels across household clusters while preserving within-household dependence.

For purposes, the basic weighted counts were:

$$N(x) = \sum_i w_i \sum_{t=1}^{T_i} 1\{x_{it} = x\} \quad (4)$$

$$N(x, x') = \sum_i w_i \sum_{t=2}^{T_i} 1\{x_{i,t-1} = x, x_{it} = x'\} \quad (5)$$

$$N(x, x', x'') = \sum_i w_i \sum_{t=3}^{T_i} 1\{x_{i,t-2} = x, x_{i,t-1} = x', x_{it} = x''\} \quad (6)$$

The corresponding empirical probabilities were:

$$\hat{p}(x) = \frac{N(x)}{\sum_{a \in \chi} N(a)} \quad (7)$$

$$\hat{p}(x, x') = \frac{N(x, x')}{\sum_{a \in \chi} \sum_{b \in \chi} N(a, b)} \quad (8)$$

$$\hat{p}(x, x', x'') = \frac{N(x, x', x'')}{\sum_{a \in \chi} \sum_{b \in \chi} \sum_{c \in \chi} N(a, b, c)} \quad (9)$$

First, marginal purpose entropy was defined as

$$H_0 = H(X_t) = -\sum_{x \in \chi} \hat{p}(x) \log_2 \hat{p}(x) \quad (10)$$

where X denotes a randomly drawn purpose token from the weighted purpose distribution. H_0 summarizes purpose diversity.

First-order and second-order conditional entropy were defined as

$$H_1 = H(X_t | X_{t-1}) = -\sum_{a \in \chi} \sum_{b \in \chi} \hat{p}(a, b) \log_2 \frac{\hat{p}(a, b)}{\hat{p}(a)} \quad (11)$$

and

$$H_2 = H(X_t | X_{t-1}, X_{t-2}) = -\sum_{a \in \chi} \sum_{b \in \chi} \sum_{c \in \chi} \hat{p}(a, b, c) \log_2 \frac{\hat{p}(a, b, c)}{\hat{p}(a, b)} \quad (12)$$

respectively, where $\hat{p}(a) = \sum_{b \in \chi} \hat{p}(a, b)$, $\hat{p}(a, b) = \sum_{c \in \chi} \hat{p}(a, b, c)$.

The primary sequential entropy gap was defined as

$$G = H_0(X) - H(X_{t+1} | X_t) \quad (13)$$

For the companion pair-consistent measure, strict adjacent-pair mutual information was

$$I_{pair} = \sum_{a, b} p_{pair}(a, b) \log_2 [(p_{pair}(a, b)) / (p_O(a) p_D(b))] . \text{ The two statistics satisfy } G = I_{pair} + (H_0 - H_{destination}) ,$$

where $H_{destination}$ is the entropy of successor states in the adjacent-pair sample. The second term is a position-composition term. G , strict adjacent-pair MI and this term were estimated with the same survey weights and household-cluster bootstrap and permutation procedures.

Mode entropy and purpose-mode coupling were computed from the observed trip-level mode and purpose-mode pairs. The relevant weighted counts were:

$$N(m) = \sum_i w_i \sum_{t=1}^{l_i} 1\{m_{it} = m\} \quad (14)$$

$$N(m, m') = \sum_i w_i \sum_{t=2}^{l_i} 1\{m_{i,t-1} = m, m_{it} = m'\} \quad (15)$$

$$N(x, m) = \sum_i w_i \sum_{t=1}^{l_i} 1\{x_{it}^{obs} = x, m_{it} = m\} \quad (16)$$

From these counts, we computed

$$H(M) = -\sum_{m \in M} \hat{p}(m) \log_2 \hat{p}(m) \quad (17)$$

$$H(M_t | M_{t-1}) = -\sum_{a \in M} \sum_{b \in M} \hat{p}(a, b) \log_2 \frac{\hat{p}(a, b)}{\hat{p}(a)} \quad (18)$$

$$I(X; M) = \sum_{x \in M} \sum_{m \in M} \hat{p}(x, m) \log_2 \frac{\hat{p}(x, m)}{\hat{p}(x) \hat{p}(m)} \quad (19)$$

along with the normalized mutual information

$$NMI(X, M) = \frac{I(X; M)}{\min(H(X), H(M))} \quad (20)$$

For this purpose-mode calculation, $\hat{p}(x, m)$, $\hat{p}(x)$, and $\hat{p}(m)$ were all derived from the observed trip level purpose mode sequences $\mathcal{C}_i$, rather than from the boundary-augmented purpose sequences. The normalized measure was recorded because absolute mutual information can remain stable even when marginal uncertainty changes.

Next, full-chain entropy was calculated as the weighted entropy of the observed purpose sequence distribution within each year. Let $\mathcal{S}_y$ denote the set of distinct observed purpose sequences in year y . For sequence $s \in \mathcal{S}_y$, define $W_y(s) = \sum_{i: \text{year}_i = y} w_i 1\{s_i = s\}$,

$$\hat{p}_y(s) = \frac{W_y(s)}{\sum_{r \in \mathcal{S}_y} W_y(r)}.$$

Then chain entropy was

$$H_{chain, y} = - \sum_{s \in \mathcal{S}_y} \hat{p}_y(s) \log_2 \hat{p}_y(s) \quad (21)$$

S3 Demographic-conditioned Markov chain

To model sequence evolution while preserving observed heterogeneity, we estimated a demographic-conditioned first-order Markov chain for each survey year (equation (22)). Demographic cells were defined by age group (<25, 25–44, 45–64, 65+), worker status (0/1), survey-defined settlement status (urban/rural), and four survey-reported nominal-dollar income bands (under \$25,000, \$25,000–49,999, \$50,000–74,999, and \$75,000 or more), yielding 64 cells. The analysis pipeline did not apply an inflation adjustment to these reported income bands.

For year y , cell d , origin state a , and destination state b , weighted transition counts and origin totals were defined by equations (23) and (24). Transition probabilities were then estimated with Laplace smoothing (equation (25)), using $\alpha = 1$ to avoid zero-probability destinations when some within-cell transitions were sparse.

A weighted chain-length distribution was estimated separately for each demographic cell and survey year using equations (26) and (27). The first-order restriction was deliberate. Although descriptive entropy analyses included second-order conditional entropy on pooled sequences, second-order demographic-conditioned transition models were not fitted because doing so would have imposed substantially heavier support requirements within already partitioned cells. The counterfactual analysis therefore used the first-order model throughout.

The estimated demographic-conditioned first-order Markov chain for each survey year:

$$P_y(X_t | X_{t-1}, D) \quad (22)$$

where D denotes a demographic cell.

For year y , cell d , origin state a , and destination state b , the weighted transition count was

$$N_y(a, b | d) = \sum_{i: \text{year}_i = y, D_i = d} w_i \sum_{t=2}^{T_i} 1\{x_{i,t-1} = a, x_{it} = b\} \quad (23)$$

The corresponding origin total was

$$N_y(a | d) = \sum_{b \in \mathcal{X}} N_y(a, b | d) \quad (24)$$

With Laplace smoothing parameter $\alpha = 1$ and $|\mathcal{X}| = 9$, the fitted transition probability was

$$\hat{P}_y(X_t = b \mid X_{t-1} = a, D = d) = \frac{N_y(a, b \mid d) + \alpha}{N_y(a \mid d) + \alpha \mid \chi} \quad (25)$$

A weighted chain-length distribution was also estimated for each demographic cell and year. For observed trip count $L = l$,

$$W_y(l \mid d) = \sum_{i: \text{year}_i = y} w_i 1\{D_i = d, l_i = l\} \quad (26)$$

$$\hat{P}_y(L = l \mid D = d) = \frac{W_y(l \mid d)}{\sum_r W_y(r \mid d)} \quad (27)$$

S4 Cross-year transfer evaluation

Within the complete four-dimension sample, temporal portability was evaluated by training the demographic-conditioned Markov model on year A and scoring sequences from year B, producing a 3×3 matrix of within-wave and cross-wave perplexities. Equations (28)–(31) define transferred log-likelihood, weighted aggregate log-likelihood, weighted scored transitions and perplexity. Diagonal entries are fitted within-wave reference scores rather than held-out performance estimates; off-diagonal entries diagnose how transition structure estimated in one survey wave fits another.

For a test sequence $S_i = (x_{i1}, \dots, x_{iT_i})$ from year B with demographic cell $D_i = d$, the transferred log-likelihood under the year-A model was

$$l_{A \rightarrow B, i} = \sum_{t=2}^{T_i} \log \hat{P}_A(x_{it} \mid x_{i,t-1}, D_i = d) \quad (28)$$

Aggregating over all test person-days in year B, the weighted total log-likelihood was

$$L_{A \rightarrow B} = \sum_{i: \text{year}_i = B} w_i l_{A \rightarrow B, i} \quad (29)$$

and the corresponding weighted number of scored transitions was

$$N_{A \rightarrow B}^{tr} = \sum_{i: \text{year}_i = B} w_i (T_i - 1) \quad (30)$$

The reported perplexity was

$$\text{Perplexity}_{A \rightarrow B} = \exp\left(-\frac{L_{A \rightarrow B}}{N_{A \rightarrow B}^{tr}}\right) \quad (31)$$

Destination-specific portability counterfactual

To translate temporal portability into destination-purpose transition rates, the primary analysis excluded any observed adjacent pair if either endpoint belonged to the residual other/unaligned category. Only originally adjacent pairs with two retained endpoints were used. We then refitted the 2009 and 2017 demographic-conditioned transition matrices and retained origin-purpose exposure. Holding the observed 2017 demographic-cell-by-origin retained transition exposure and weighted person-day denominator fixed, we allocated exposure across the eight retained destination purposes using the refitted 2009 transition probabilities. Fixing origin exposures also fixed total transition volume across those destinations. For demographic cell, origin purpose and destination purpose, the projected weighted count was

$$\hat{N}_{dab}^{2009 \rightarrow 2017} = N_{da}^{2017} \hat{P}_{dab}^{2009}. \quad (32)$$

Observed and projected destination counts were summed across demographic cells and origin purposes, divided by total weighted 2017 person-days, and expressed per 100 weighted person-days. This construction preserves total retained transition volume by

design and changes only its allocation across destination purposes. In the 2017 NHTS, final trip weights equal final person weights multiplied by 365. The average-day national gap was therefore calculated as the gap per 100 weighted person-days multiplied by the annual weighted person-day total (75,579,689,913), then divided by 100 and 365. The full nine-category calculation, retaining the residual other/unaligned category, was repeated as a sensitivity specification. Uncertainty was estimated using 1,000 household-cluster bootstrap resamples, independently resampling households within the 2009 and 2017 waves and recomputing the fitted probabilities and target-year exposure in each replicate.

S5 Counterfactual generation and sequence-level decomposition

The main decomposition compared 2001 and 2017 and was also applied to 2001–2009 and 2009–2017. For source year s and target year e , the counterfactual combined source-year modeled sequence structure with target-year cell shares across the four measured demographic dimensions. The main specification paired 2001 transition structure with 2017 cell shares.

For each target-year person-day in demographic cell d , the simulation generated synthetic sequences in four steps. First, it drew an observed trip length from the source-year cell-specific trip-count distribution (equation (33)). Second, because the fitted model operated on boundary-augmented purpose sequences rather than raw trip counts alone, it drew an empirical padding difference and converted observed trip count to simulated purpose-sequence length (equations (34) and (35)). The empirical padding difference used the same codebook-based home-boundary rule described in Section S1. Third, it drew an initial activity from the source-year initial-state distribution (equations (36) and (37)). Fourth, it generated later states recursively from the fitted first-order transition model (equation (38)).

For each target-year person-day, we generated $K = 50$ Monte Carlo draws using random seed 20260321. Each counterfactual summary metric was averaged across draws. Total fitted change was partitioned into a measured cell-share component and a residual within-cell sequence component through equations (40)–(43). Same-year reconstruction diagnostics bound interpretation of the fitted components.

The decomposition was applied to mean chain length, purpose entropy measures including G , selected trip-purpose shares and the simple-commute motif (Home–Work–Home). The residual within-cell sequence component combines remaining changes in chain length, initial state and first-order transition structure within the broad cells and is interpreted as a reduced-form sequence component rather than a structural preference parameter. Same-year reconstruction recovered weighted mean chain length within 0.003 trips in each wave but retained residual differences for G and the simple-commute motif (Supplementary Table 11). The decomposition is therefore used as model-based fitted attribution across the four measured dimensions rather than exact reconstruction of all higher-order sequence features.

For each target-year person-day i in the target year e with demographic cell $D_i = d$, the simulation drew a synthetic chain in four steps. First, it drew an observed trip length

$$L_i^* \sim \hat{P}_s(L | D = d) \quad (33)$$

Second, because the fitted Markov model operated on purpose sequence rather than on observed trip count alone, the simulation drew the empirical padding difference

$$B_i^* \sim \hat{P}_s(B | D = d, L = L_i^*), \quad B = T - L \in \{0, 1\} \quad (34)$$

The simulated purpose sequence length was then

$$T_i^* = L_i^* + B_i^* \quad (35)$$

Third, the initial-purpose distribution was estimated directly from the source-year sequence:

$$\hat{P}_s(X_1 = x | D = d) = \frac{\sum_{i: \text{year}_i = s} w_i 1\{D_i = d, x_{i1} = x\}}{\sum_{i: \text{year}_i = s} w_i 1\{D_i = d\}} \quad (36)$$

The first simulated state was drawn as

$$X_{i1}^* \sim \hat{P}_s(X_1 | D = d) \quad (37)$$

Fourth, for $t = 2, \dots, T_i^*$, later states were drawn recursively from the fitted first-order model:

$$X_{it}^* \sim \hat{P}_s(\cdot | X_{i,t-1}^*, D = d) \quad (38)$$

For each target-year person-day, we compute $K = 50$ Monte Carlo draws. The point estimate of any counterfactual summary metric g was taken as the simulation average

$$\hat{g}(R_s(D_e)) = \frac{1}{K} \sum_{k=1}^K g(R_s^{(k)}(D_e)) \quad (39)$$

where $R_s^{(k)}(D_e)$ denotes the full set of simulated sequences in draw k .

For any metric g , total change was decomposed as

$$\Delta g = g(actual_e) - g(actual_s) \quad (40)$$

$$Composition\ effect = \hat{g}(R_s(D_e)) - g(actual_s) \quad (41)$$

$$Within-cell\ sequence\ effect = g(actual_e) - \hat{g}(R_s(D_e)) \quad (42)$$

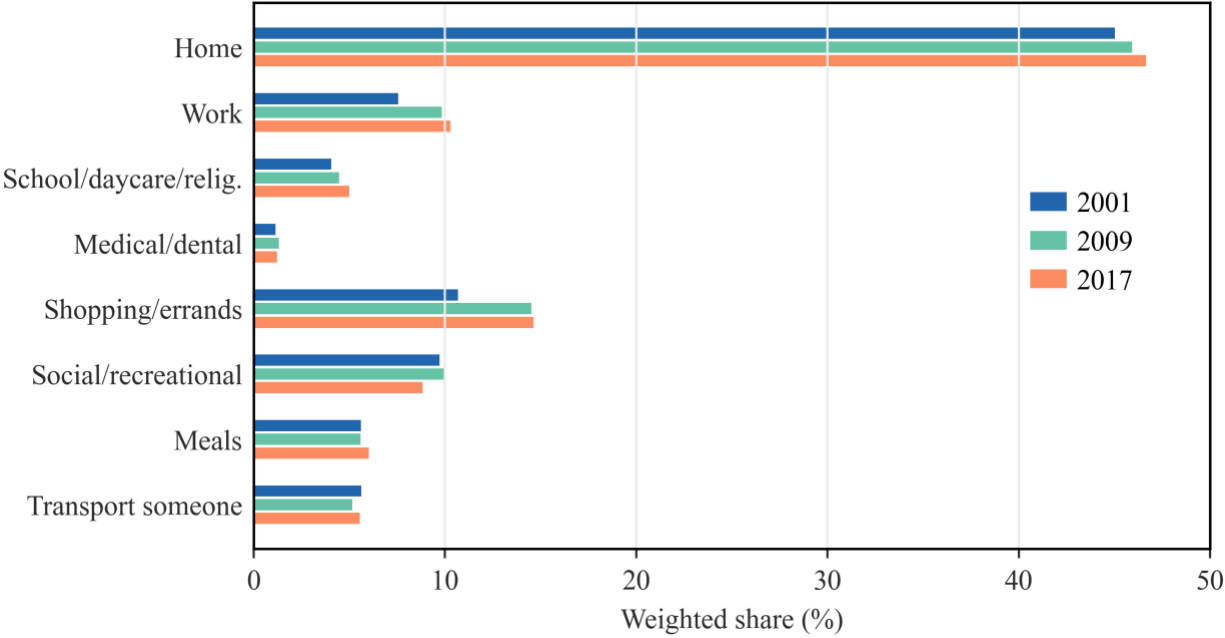
Hence,

$$\Delta g = Composition\ effect + Within-cell\ sequence\ effect \quad (43)$$

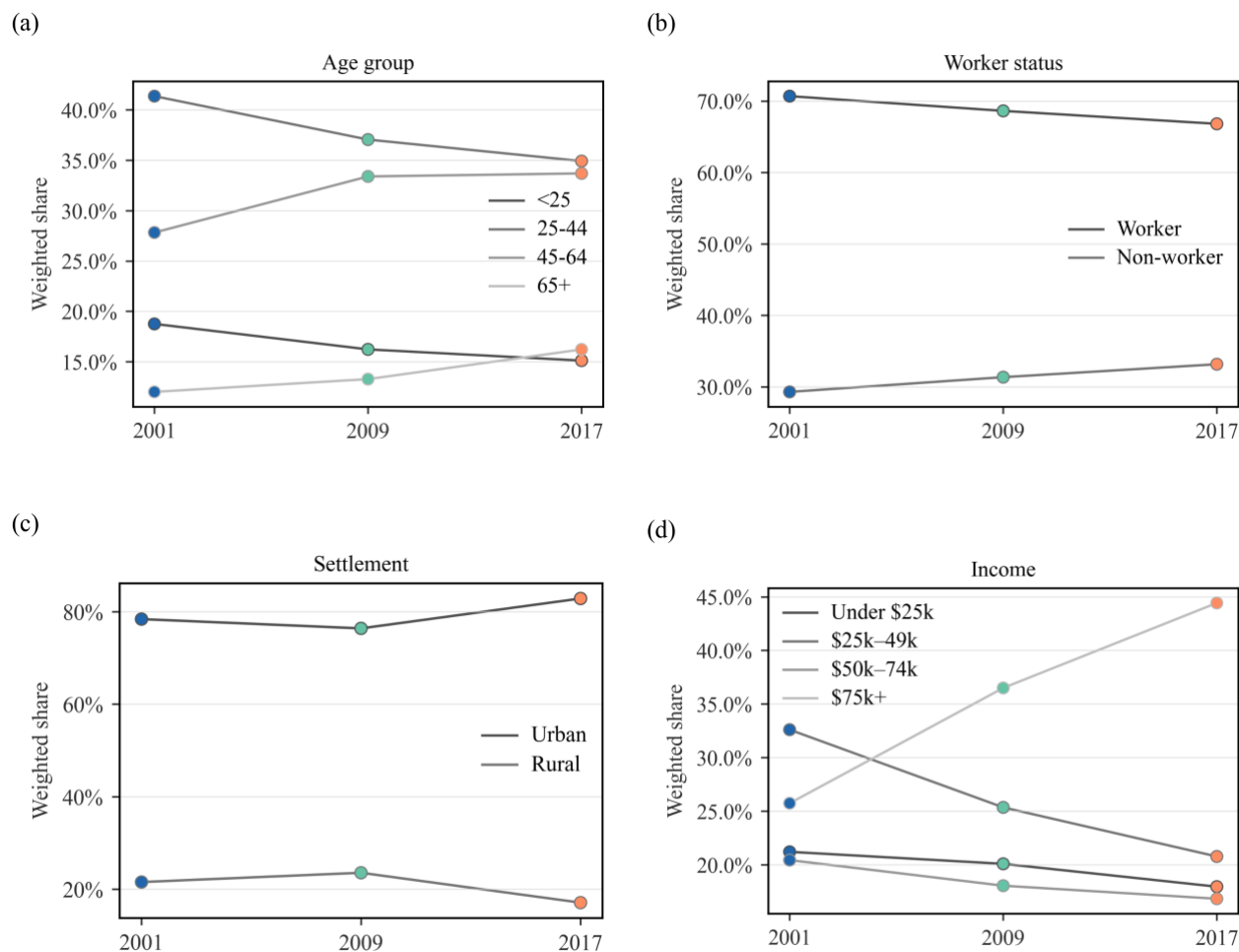
S6 Stratified analysis and robustness checks

Observed G was repeated across 15 prespecified series: weekdays (TRAVDAY 2–6) and weekends (TRAVDAY 1 or 7); survey-defined urban and rural strata; workers and non-workers; age 16–34, 35–54 and 55 years or older (respondents younger than 16 were excluded from the age-series comparison); three income terciles; and household vehicle counts of 0, 1 and 2 or more. Income terciles were calculated separately within each survey year from the person-day-weighted cumulative distribution of the raw standardized household-income code, with cutpoints at one-third and two-thirds of the cumulative weight. Counterfactual decompositions by stratum reused the fitted Markov framework rather than estimating separate transition systems. Robustness checks varied the demographic partition, weighting and decomposition reference. The income-collapsed weighted specification used 16 Age $\times$ Worker $\times$ Settlement cells within the baseline complete sample; the expanded weighted specification added sex to the baseline partition, yielding 128 cells; reverse Oaxaca assessed path dependence; and the unweighted specification assessed survey weighting. These checks evaluated whether the dominance of the residual within-cell sequence component depended on cell design, weighting or decomposition direction.

Jurisdiction-level sequential entropy gaps were computed by linking person-day sequences to household state FIPS codes, subject to at least 300 person-days per jurisdiction-wave. Comparisons involving 2001 covered 35 available states; the 2009–2017 comparison covered 50 states and the District of Columbia. Estimates were used descriptively to summarize observed directions. Global Moran's I and local indicators of spatial association used row-standardized Queen-contiguity weights for the 48 contiguous states plus the District of Columbia and 999 permutations. Cross-jurisdiction Pearson correlations examined persistence in change in G and its relation to the reported aggregate covariates.

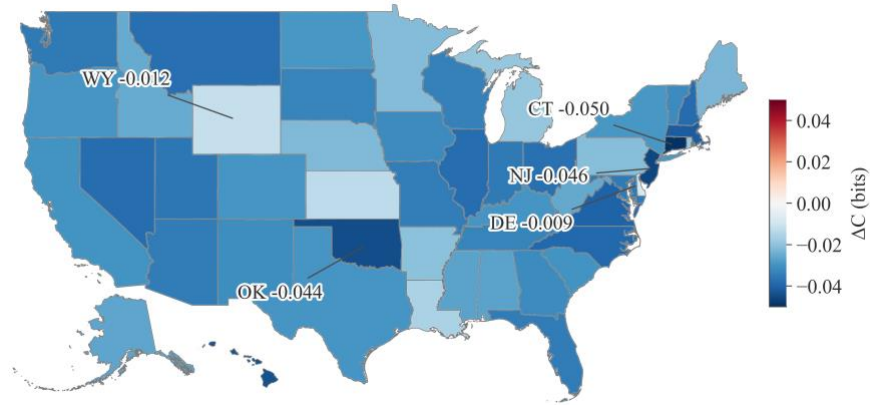


221
222 Supplementary Fig. 1. Weighted marginal activity-purpose shares by survey year. The residual other/unaligned category is omitted because it partly
223 reflects cross-wave harmonization and coding differences rather than a substantive activity-purpose category.



224

225 Supplementary Fig. 2. Demographic composition of the complete four-dimension model-fitting sample. Panels show weighted shares by age group,
 226 worker status, survey-defined settlement stratum and income group in 2001, 2009 and 2017. Income groups are survey-reported nominal-dollar bands;
 227 the analysis pipeline did not apply an inflation adjustment.

Home-to-Home ΔC , jurisdiction-level (2009–2017)

Supplementary Fig. 3. Jurisdiction-level change in the all-record strict-MI contribution of the Home-to-Home transition from 2009 to 2017. Home-to-Home was the largest negative transition-specific contributor in 35 of 51 jurisdictions. This auxiliary diagnostic is measurement-sensitive because 2017 Home-to-Home pairs overlap substantially with loop-flagged records; it is not a loop-excluded estimate or a causal mechanism test.

Supplementary Table 1. Sample coverage after harmonization and quality filtering.

Year	Traveling retained	person-days	Trips retained
2001	140,747		641,302
2009	262,557		1,164,432
2017	219,192		923,556

Supplementary Table 2. Full set of primary purpose-sequence and secondary mode-based information-theoretic measures by survey year, with household-cluster bootstrap 95% confidence intervals where reported and pairwise permutation p values. Chain-entropy intervals were not reported.

Metric	Year	Point estimate	95% CI	p vs 2001	p vs 2009	p vs 2017
H ₀ purpose	2001	2.543	[2.537, 2.548]	--	<0.001	<0.001
H ₀ purpose	2009	2.476	[2.469, 2.483]	<0.001	--	<0.001
H ₀ purpose	2017	2.438	[2.432, 2.445]	<0.001	<0.001	--
H ₁ purpose	2001	2.427	[2.419, 2.435]	--	<0.001	0.002
H ₁ purpose	2009	2.359	[2.349, 2.369]	<0.001	--	<0.001
H ₁ purpose	2017	2.4	[2.391, 2.409]	0.002	<0.001	--
H ₂ purpose	2001	2.19	[2.176, 2.197]	--	<0.001	<0.001
H ₂ purpose	2009	2.099	[2.083, 2.106]	<0.001	--	<0.001
H ₂ purpose	2017	2.145	[2.129, 2.152]	<0.001	<0.001	--
Sequential entropy gap G	2001	0.115	[0.112, 0.120]	--	0.704	<0.001
Sequential entropy gap G	2009	0.117	[0.112, 0.122]	0.704	--	<0.001
Sequential entropy gap G	2017	0.038	[0.034, 0.043]	<0.001	<0.001	--
H ₀ mode	2001	0.879	[0.861, 0.899]	--	<0.001	<0.001
H ₀ mode	2009	1.019	[0.996, 1.043]	<0.001	--	<0.001
H ₀ mode	2017	1.074	[1.053, 1.096]	<0.001	<0.001	--
H ₁ mode	2001	0.479	[0.467, 0.490]	--	<0.001	<0.001
H ₁ mode	2009	0.517	[0.503, 0.529]	<0.001	--	<0.001
H ₁ mode	2017	0.558	[0.544, 0.571]	<0.001	<0.001	--
MI purpose-mode	2001	0.056	[0.054, 0.059]	--	<0.001	0.799
MI purpose-mode	2009	0.065	[0.062, 0.069]	<0.001	--	<0.001
MI purpose-mode	2017	0.057	[0.054, 0.060]	0.799	<0.001	--
NMI purpose-mode	2001	0.064	[0.062, 0.067]	--	0.898	<0.001
NMI purpose-mode	2009	0.063	[0.061, 0.068]	0.898	--	<0.001
NMI purpose-mode	2017	0.053	[0.051, 0.056]	<0.001	<0.001	--
Chain entropy	2001	10.549	not reported	--	<0.001	<0.001
Chain entropy	2009	10.076	not reported	<0.001	--	<0.001
Chain entropy	2017	9.649	not reported	<0.001	<0.001	--

Secondary mode-based analysis

As a supplementary specificity analysis, we examined whether the weakening of purpose-sequence dependence coincided with a general simplification of travel-mode organization. Marginal mode entropy increased from 0.879 in 2001 to 1.019 in 2009 and 1.074 in 2017, while first-order mode conditional entropy increased from 0.479 to 0.517 and 0.558. Absolute purpose–mode mutual information returned to approximately its 2001 level by 2017 (0.056, 0.065 and 0.057), whereas normalized mutual information declined from 0.064 and 0.063 to 0.053. Thus, the weakening of purpose-sequence dependence did not reflect a uniform simplification across all dimensions

of travel organization: modal diversity increased, while purpose and mode became less tightly coupled after accounting for marginal mode uncertainty.

We first assessed robustness of the 2001–2017 fitted decomposition across alternative specifications (Supplementary Table 3). The mean-chain-length result was stable in every specification: the residual within-cell sequence component remained dominant and continued to move opposite to the cell-share component in the baseline weighted 64-cell, income-collapsed weighted 16-cell, sex-expanded weighted 128-cell, reverse-Oaxaca and unweighted variants. H_0 purpose was less stable in magnitude, although the forward weighted, 16-cell, 128-cell and unweighted specifications retained the same sign pattern. Only the reverse-Oaxaca entropy decomposition showed stronger path dependence. Residual within-cell dominance for chain length was therefore robust, whereas entropy-based partitions were more sensitive to decomposition direction.

Supplementary Table 3. Robustness summary for the 2001–2017 fitted sequence decomposition.

Check	Mean chain length: residual within-cell sequence dominates?	H_0 purpose: residual within-cell sequence dominates?
Baseline weighted 64-cell	Yes	Yes
Weighted 16-cell (Age × Worker × Settlement; income collapsed within baseline complete sample)	Yes	Yes
Weighted 128-cell (Age × Worker × Settlement × Income × Sex)	Yes	Yes
Reverse Oaxaca	Yes	Path-dependent
Unweighted 64-cell	Yes	Yes

We next evaluated whether the sequential entropy gap depended on the purpose taxonomy (Supplementary Table 4). Under a five-category scheme, G was 0.162 in 2001, 0.159 in 2009 and 0.078 in 2017, retaining both the small 2001–2009 difference and the 2009–2017 decline. Excluding the residual other/unaligned category changed the early-wave comparison, yielding 0.009, 0.071 and 0.040 bits, but retained the 2009–2017 decline. Under residual-category exclusion, 40 of 51 jurisdictions still had lower 2017 point estimates, and baseline and alternative jurisdiction-level changes remained strongly correlated ($r = 0.853$).

Supplementary Table 4. Sequential entropy gap G across alternative purpose specifications.

Specification	Year	G (95% CI)
Baseline 9-category	2001	0.115 [0.112, 0.120]
Baseline 9-category	2009	0.117 [0.112, 0.122]
Baseline 9-category	2017	0.038 [0.034, 0.043]
Drop residual 8-category	2001	0.009 [0.005, 0.015]
Drop residual 8-category	2009	0.071 [0.066, 0.076]
Drop residual 8-category	2017	0.040 [0.036, 0.045]
Coarse 5-category	2001	0.162 [0.158, 0.165]
Coarse 5-category	2009	0.159 [0.155, 0.164]
Coarse 5-category	2017	0.078 [0.074, 0.083]

A convergent check based on official NHTS trip-chain and tour files also indicated lower tour complexity between 2009 and 2017 (Supplementary Table 5). After deduplication of official tour identifiers, mean STOPS fell from 0.375 to 0.335, the share of tours with $\text{STOPS} \geq 2$ fell from 6.45% to 5.84%, and the share with $\text{STOPS} \geq 1$ fell from 27.8% to 24.6%. In 2017, deduplicated trip-chain records matched official tour-level counts across all STOPS categories. These official tour measures are not the same construct as the person-day chain motifs but provide convergent evidence of later-period simplification.

Supplementary Table 5. Official NHTS trip-chain/tour validation, 2009–2017.

Year	Mean STOPS	Tours with $\text{STOPS} \geq 1$ (%)	Tours with $\text{STOPS} \geq 2$ (%)	Unweighted tours
2009	0.375	27.8	6.45	831,317
2017	0.335	24.6	5.84	678,458

The lower 2017 sequential entropy gap was not confined to aggregate mixing across demographic groups (Supplementary Table 6). All 64 demographic cells were present in every core wave. Between 2009 and 2017, 57 of 64 cells had lower point estimates of G , and 26 of 64 declined by more than half relative to 2009. The weighted mean of cell-specific G fell from 0.149 in 2009 to 0.075 in 2017. These cell-averaged values are distinct from the pooled complete-sample values in Supplementary Table 10.

267 Supplementary Table 6. Weighted mean of cell-specific sequential entropy gap G and direction counts across the complete 64-cell demographic partition.

Metric	Entry	Estimate	Additional detail
Weighted mean within-cell G	2001	0.155	Included person-days=108,805
Weighted mean within-cell G	2009	0.149	Included person-days=214,130
Weighted mean within-cell G	2017	0.075	Included person-days=189,928
Cells with G decline	2001-2009	38/64	2/64 declined by more than 50%
Cells with G decline	2001-2017	59/64	31/64 declined by more than 50%
Cells with G decline	2009-2017	57/64	26/64 declined by more than 50%

268 As a pair-consistent validation, strict adjacent-pair mutual information showed the same cross-wave pattern (Supplementary
269 Table 7). The estimate was 0.355 bits in 2001 and 0.356 bits in 2009 ($p = 0.748$), then 0.289 bits in 2017, an 18.7% reduction relative to
270 2009 ($p < 0.001$). The direct pairwise component accounted for approximately 85% of the absolute 2009–2017 decline in the primary
271 sequential entropy gap G.

272 Supplementary Table 7. Strict adjacent-pair mutual information by survey year.

Year	Strict adjacent-pair MI (bits)	95% CI	p vs 2001	p vs 2009	p vs 2017
2001	0.355	[0.351, 0.359]	--	0.748	<0.001
2009	0.356	[0.351, 0.362]	0.748	--	<0.001
2017	0.289	[0.285, 0.295]	<0.001	<0.001	--

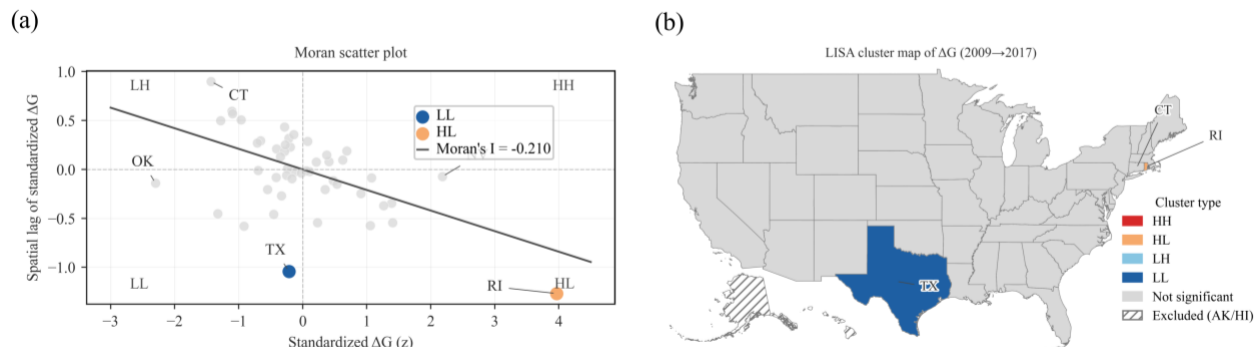
273 We tested whether the 2009–2017 contrast depended on retained 2017 trip records flagged LOOP_TRIP = 1. These records
274 comprised 20,616 of 923,556 trips (2.23%; weighted share 1.88%). After excluding them and rebuilding person-day sequences with the
275 same boundary rule, G remained lower in 2017 (0.069 bits) than in 2009 (0.117 bits), and strict adjacent-pair MI likewise remained lower
276 (0.319 versus 0.356 bits). Under the 300-person-day threshold, lower 2017 point estimates remained in 47 and 45 of 51 jurisdictions,
277 respectively (Supplementary Table 8).

278 Supplementary Table 8. Sensitivity to 2017 trip records flagged as loop trips.

Measure	2009, all retained	2017, all retained	2017, flagged loops excluded	Jurisdictions lower after exclusion
Sequential entropy gap G (bits)	0.117	0.038	0.069	47/51
Strict adjacent-pair MI (bits)	0.356	0.289	0.319	45/51
Traveling person-days	262,557	219,192	216,697	--
Observed trip records	1,164,432	923,556	902,940	--

279 **Spatial and geographic diagnostics for the sequential entropy gap**

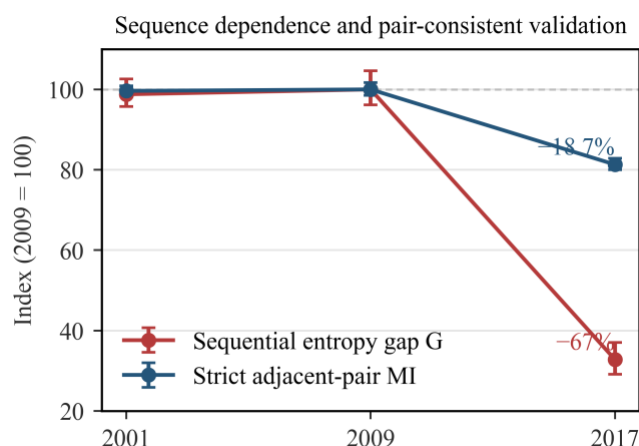
280 The 2009–2017 jurisdiction pattern was spatially dispersed rather than concentrated in broad contiguous regions. Global Moran's
281 I for change in G was -0.210 ($p = 0.030$, 999 permutations) under row-standardized Queen-contiguity weights for the 48 contiguous
282 states plus the District of Columbia (Supplementary Fig. 4a). Local indicators of spatial association identified only two significant
283 jurisdictions at the 0.05 level: Texas as Low–Low and Rhode Island as a High–Low outlier (Supplementary Fig. 4b). Additional
284 diagnostics indicated that the change in G was not closely related to the examined aggregate jurisdiction characteristics. Baseline G in
285 2001 did not predict subsequent change ($r = 0.052$, $p = 0.766$), and 2009–2017 change in G was essentially uncorrelated with change in
286 mean chain length ($r = 0.020$, $p = 0.890$). It was also weakly related to changes in urban-household share ($r = 0.16$, $p = 0.25$), changes in
287 internet access ($r = -0.035$, $p = 0.807$), and baseline 2013 internet access ($r = -0.015$, $p = 0.917$).



Supplementary Fig. 4. Spatial diagnostics for 2009–2017 change in the sequential entropy gap G. a Moran scatter plot under row-standardized Queen contiguity (Moran's $I = -0.210$, permutation $p = 0.030$). b Local indicators of spatial association from 999 permutations. Significant clusters and outliers are colored; non-significant jurisdictions are gray.

Additional sequence diagnostic

Supplementary Fig. 5 compares the primary sequential entropy gap G with strict adjacent-pair mutual information after indexing each measure to its 2009 value. Both changed little through 2009 and were lower in 2017, confirming the same timing and direction under different positional supports.



Supplementary Fig. 5. Indexed comparison of the sequential entropy gap G and strict adjacent-pair mutual information. Each measure is indexed to its 2009 estimate (=100); points show survey-weighted estimates and whiskers show 1,000-replicate household-cluster bootstrap 95% confidence intervals. Strict MI is interpreted as a pair-consistent validation of timing and direction rather than an interchangeable estimate of G.

Destination-specific portability contrasts

The primary counterfactual used eight directly aligned destination purposes, excluding observed adjacent pairs when either endpoint belonged to the residual other/unaligned category and refitting the source and target matrices on the retained pairs (Supplementary Table 9, Panel A). Transitions ending at work exceeded the 2009-rule projection by 4.14 per 100 weighted person-days (95% CI, 2.76–5.51), and transitions ending in the harmonized school/daycare/religious category exceeded it by 1.86 (95% CI, 1.10–2.67). Together, the observed rate exceeded the projection by 5.99 transitions per 100 weighted person-days (95% CI, 4.50–7.60), or 8.9% of the projection, corresponding to approximately 12.4 million transitions on an average day under survey weights (95% CI, 9.3–15.7 million). The full nine-category specification yielded a 12.5% gap, about 17.0 million transitions per day, but the residual category contributed -7.36 transitions per 100 weighted person-days and partly reflects cross-wave harmonization and coding differences (Panel B). We therefore use Panel A as the primary estimand and report Panel B as a sensitivity analysis. Signed gaps sum to zero within each panel because retained transition volume is fixed.

311 Supplementary Table 9. Destination-specific 2017 observed-versus-2009-rule counterfactual. Panel A reports the primary directly aligned eight-purpose
312 comparison; Panel B reports the full nine-category sensitivity analysis.

Destination purpose	Observed 2017 (per 100 weighted person-days)	2009-rule projection (per 100 weighted person-days)	Observed - projected (95% CI)	Relative gap
Panel A. Primary directly aligned eight-purpose comparison				
Home	136.96	138.27	-1.30 (-3.12, 0.33)	-0.9%
Work	57.96	53.82	4.14 (2.76, 5.51)	7.7%
School/daycare/religious	15.04	13.18	1.86 (1.10, 2.67)	14.1%
Medical/dental	6.32	7.44	-1.11 (-1.63, -0.63)	-15.0%
Shopping/errands	76.07	77.07	-1.00 (-2.53, 0.57)	-1.3%
Social/recreational	41.27	45.58	-4.32 (-5.49, -3.14)	-9.5%
Meals	30.42	29.24	1.18 (0.27, 2.18)	4.0%
Transport someone	27.86	27.30	0.56 (-0.44, 1.61)	2.1%
Work + school/daycare/religious	73.00	67.01	5.99 (4.50, 7.60)	8.9%
All retained destinations	391.90	391.90	0.00 (0.00, 0.00)	0.0%
Panel B. Full nine-category sensitivity analysis				
Home	139.94	141.35	-1.41 (-3.33, 0.21)	-1.0%
Work	58.54	52.77	5.77 (4.40, 7.11)	10.9%
School/daycare/religious	15.30	12.88	2.42 (1.65, 3.22)	18.8%
Medical/dental	6.40	7.23	-0.83 (-1.33, -0.36)	-11.5%
Shopping/errands	77.04	76.35	0.69 (-0.86, 2.24)	0.9%
Social/recreational	41.97	44.77	-2.80 (-3.97, -1.62)	-6.2%
Meals	31.02	29.01	2.01 (1.09, 2.98)	6.9%
Transport someone	28.28	26.75	1.52 (0.51, 2.56)	5.7%
Other	7.78	15.14	-7.36 (-7.99, -6.72)	-48.6%
Work + school/daycare/religious	73.84	65.65	8.19 (6.68, 9.78)	12.5%
All destinations	406.25	406.25	0.00 (0.00, 0.00)	0.0%

313 Values are transitions per 100 weighted person-days. The counterfactual uses the complete four-dimension sample. For Panel A,
314 any originally adjacent pair with a residual other/unaligned endpoint was excluded before the 2009 and 2017 transition matrices and
315 origin exposures were refitted. Panel B retains all nine harmonized categories. Both projections hold the observed 2017 demographic-
316 cell-by-origin retained transition exposure and weighted person-day denominator fixed and apply the corresponding 2009 demographic-
317 conditioned transition probabilities. Parentheses report 95% percentile intervals from 1,000 household-cluster bootstrap resamples. Signed
318 destination gaps sum to zero within each panel because the origin exposures, and therefore total retained transition volume, are fixed.

319 **Sample-definition sensitivity**

320 The information measures were first estimated for all traveling person-days and then re-estimated in the complete four-
321 dimension sample used for the fitted Markov and decomposition analyses. The 2009–2017 decline was nearly unchanged: G fell by 67.2%
322 in the full sample and 67.3% in the complete sample, while strict adjacent-pair MI fell by 18.7% and 19.5%, respectively.

323 Supplementary Table 10. Pooled information-measure sensitivity to the complete four-dimension model-fitting sample definition.

Sample definition	Year	Person-days	Unweighted share of full (%)	Weighted share of full (%)	G (bits)	Direct MI (bits)
All traveling person-days	2001	140,747	100.0	100.0	0.115247	0.354717
All traveling person-days	2009	262,557	100.0	100.0	0.116667	0.356040
All traveling person-days	2017	219,192	100.0	100.0	0.038251	0.289475
Complete four-dimension cells	2001	108,805	77.3	74.2	0.115494	0.338155
Complete four-dimension cells	2009	214,130	81.6	79.9	0.108374	0.333226
Complete four-dimension cells	2017	189,928	86.6	82.8	0.035487	0.268138

324 Shares use all traveling person-days in the same survey year as the denominator. G denotes the sequential entropy gap; direct
325 MI denotes strict adjacent-pair mutual information. The G values are pooled within each sample definition rather than weighted means
326 of cell-specific G. Values are weighted except person-day counts and unweighted shares.

327 **Same-year generator reconstruction**

328 We evaluated the fitted sequence generator by drawing sequences from each survey year's own fitted model and comparing
329 reconstructed summaries with the observed complete-sample values. Mean chain length was recovered closely, whereas G and the exact
330 simple-commute motif retained residual differences. These checks support interpreting the decomposition as reduced-form model-based
331 attribution rather than exact reproduction of every higher-order sequence feature.

332 Supplementary Table 11. Same-year reconstruction diagnostics for the fitted sequence generator.

Year	Metric	Observed	Reconstruction mean	Residual	Simulation SD
2001	Mean chain length	4.789941	4.792925	+0.002984	0.015019
2001	Sequential entropy gap G	0.115494	0.153325	+0.037831	0.001856
2001	Simple-commute motif share	0.090947	0.023821	-0.067127	0.000688
2009	Mean chain length	4.517206	4.518572	+0.001366	0.014844
2009	Sequential entropy gap G	0.108374	0.144807	+0.036433	0.002197
2009	Simple-commute motif share	0.104454	0.029404	-0.075050	0.001137
2017	Mean chain length	4.195259	4.194982	-0.000277	0.012180
2017	Sequential entropy gap G	0.035487	0.064296	+0.028808	0.001952
2017	Simple-commute motif share	0.120605	0.033779	-0.086825	0.001149

333 Residual equals reconstruction mean minus observed. Each reconstruction mean and simulation SD summarize 50 Monte Carlo draws
334 in the complete four-dimension sample.

335 **Settlement-stratified destination-purpose counterfactual**

336 Stratifying the directly aligned counterfactual by the harmonized NHTS urban–rural settlement classification located most of
337 the national destination-purpose mismatch in the survey-defined urban stratum. Its combined observed-minus-projected rate was 6.93
338 transitions per 100 weighted person-days (95% CI, 5.23–8.75), or 10.4% of the projection, corresponding to 11.9 million transitions on
339 an average day (95% CI, 9.0–15.0 million). At the point estimate, the survey-defined urban stratum contributed 95.9% of the national
340 survey-weighted average-day transition-count gap for the selected work plus school/daycare/religious subtotal. The rural combined
341 estimate was 1.45 transitions per 100 weighted person-days (2.1%; 95% CI, –1.33 to 4.77), and its interval included zero. The paired
342 urban-minus-rural contrast was 5.49 transitions per 100 weighted person-days (95% CI, 2.04–8.86) and was concentrated in work
343 transitions (5.86; 95% CI, 2.70–8.76), whereas the school/daycare/religious contrast spanned zero (–0.38; 95% CI, –2.24 to 1.64).

344 For Supplementary Table 12, the first three rows report the combined work plus school/daycare/religious subtotal for the national,
345 urban and rural strata in the complete four-dimension sample. Subsequent rows report destination-specific contrasts within the urban and
346 rural strata. Urban and rural follow the harmonized NHTS settlement classification and do not identify individual cities or metropolitan
347 areas. Observed and projected values are transitions per 100 weighted person-days. Projections hold the observed 2017 demographic-
348 cell-by-origin retained transition exposure and weighted person-day denominator fixed while applying 2009 transition probabilities.
349 Parentheses give 95% percentile intervals from 1,000 household-cluster bootstrap resamples. In the paired-contrast block, the third and
350 fourth columns give the urban and rural observed-minus-projected gaps, and the fifth column gives their paired urban-minus-rural
351 difference. Paired contrasts were computed within each bootstrap replicate before percentile intervals were taken. Signed destination gaps
352 sum to zero within each settlement stratum because the origin exposures, and therefore total retained transition volume, are fixed.

353

355 Supplementary Table 12. Settlement-stratified destination-purpose counterfactual.

Settlement	Destination purpose	Observed per 100 weighted person-days	Projected per 100 weighted person-days	Gap (95% CI)	Relative gap, % (95% CI)	Average-day gap, million (95% CI)
National	Work + school/daycare/religious	73.00	67.01	+5.99 (4.50 to 7.60)	8.9 (6.7 to 11.5)	12.41 (9.31 to 15.73)
Urban	Work + school/daycare/religious	73.69	66.75	+6.93 (5.23 to 8.75)	10.4 (7.7 to 13.3)	11.90 (8.97 to 15.02)
Rural	Work + school/daycare/religious	69.69	68.24	+1.45 (-1.33 to 4.77)	2.1 (-1.9 to 7.2)	0.51 (-0.47 to 1.69)
Urban	Home	138.95	141.09	-2.14 (-4.06 to -0.25)	-1.5 (-2.9 to -0.2)	-3.66 (-6.96 to -0.44)
Urban	Work	58.66	53.52	+5.14 (3.54 to 6.68)	9.6 (6.5 to 12.7)	8.82 (6.07 to 11.45)
Urban	School/daycare/religious	15.03	13.23	+1.79 (0.91 to 2.76)	13.6 (6.7 to 21.5)	3.08 (1.56 to 4.73)
Urban	Medical/dental	6.22	7.43	-1.21 (-1.76 to -0.67)	-16.3 (-22.6 to -9.5)	-2.08 (-3.01 to -1.14)
Urban	Shopping/errands	74.95	76.43	-1.48 (-3.14 to 0.25)	-1.9 (-4.1 to 0.3)	-2.54 (-5.39 to 0.44)
Urban	Social/recreational	42.16	45.88	-3.72 (-5.04 to -2.44)	-8.1 (-10.8 to -5.3)	-6.39 (-8.64 to -4.19)
Urban	Meals	30.50	29.38	+1.11 (0.08 to 2.14)	3.8 (0.3 to 7.5)	1.91 (0.14 to 3.67)
Urban	Transport someone	28.63	28.13	+0.50 (-0.70 to 1.69)	1.8 (-2.4 to 6.2)	0.86 (-1.21 to 2.90)
Rural	Home	127.35	124.63	+2.72 (-1.30 to 6.07)	2.2 (-1.0 to 5.0)	0.97 (-0.46 to 2.15)
Rural	Work	54.58	55.30	-0.72 (-3.05 to 2.09)	-1.3 (-5.4 to 3.9)	-0.26 (-1.08 to 0.74)
Rural	School/daycare/religious	15.11	12.94	+2.17 (0.36 to 3.77)	16.8 (2.8 to 30.6)	0.77 (0.13 to 1.34)
Rural	Medical/dental	6.81	7.46	-0.65 (-1.63 to 0.42)	-8.8 (-20.6 to 5.8)	-0.23 (-0.58 to 0.15)
Rural	Shopping/errands	81.48	80.16	+1.32 (-2.38 to 4.84)	1.6 (-2.9 to 6.1)	0.47 (-0.85 to 1.72)
Rural	Social/recreational	36.95	44.13	-7.18 (-9.68 to -4.51)	-16.3 (-21.5 to -10.4)	-2.55 (-3.44 to -1.60)
Rural	Meals	30.03	28.54	+1.50 (-0.68 to 3.59)	5.2 (-2.4 to 13.0)	0.53 (-0.24 to 1.27)
Rural	Transport someone	24.13	23.28	+0.85 (-1.08 to 2.89)	3.6 (-4.5 to 12.9)	0.30 (-0.38 to 1.03)
Paired urban-rural contrasts		Urban gap	Rural gap	Urban - rural gap (95% CI)	-	-
Urban - rural	Work + school/daycare/religious	6.93	1.45	+5.49 (2.04 to 8.86)	-	-
Urban - rural	Work	5.14	-0.72	+5.86 (2.70 to 8.76)	-	-
Urban - rural	School/daycare/religious	1.79	2.17	-0.38 (-2.24 to 1.64)	-	-
Urban - rural	Home	-2.14	2.72	-4.86 (-8.91 to -0.41)	-	-
Urban - rural	Medical/dental	-1.21	-0.65	-0.56 (-1.66 to 0.52)	-	-
Urban - rural	Shopping/errands	-1.48	1.32	-2.80 (-6.83 to 1.09)	-	-
Urban - rural	Social/recreational	-3.72	-7.18	+3.45 (0.50 to 6.32)	-	-
Urban - rural	Meals	1.11	1.50	-0.38 (-2.79 to 2.03)	-	-
Urban - rural	Transport someone	0.50	0.85	-0.35 (-2.79 to 1.98)	-	-