

Supplementary Information

This supplementary information provides implementation details, eligibility rules, sensitivity specifications and diagnostics for the analyses summarized in the main Methods.

S1. Study design, analytic frame and activity domains

The primary survey, age restriction and two placement domains are defined in the main Methods. For implementation, a respondent-day entered a domain analysis only when positive time was observed in that domain.

The strict eight-category sequence taxonomy assigned retained leisure activities to the leisure domain. Care/household obligation combines unpaid domestic activity and care. Personal maintenance, including sleep, eating and personal care, was excluded from the primary placement domains because these activities form a distinct set of strongly anchored routines. Recorded sleep was analyzed separately as an activity and as a source of waking-day boundaries.

S2. Education measurement and endpoint summary

The four ordered education categories and their descriptive role are defined in the main Methods. In the notation below, H denotes bachelor-plus and L less than high school, with the harmonized endpoint contrast

$$\Delta = \theta_H - \theta_L \quad (1)$$

Historical, spatial, decomposition and contextual analyses used this contrast. Education was treated as durable empirical stratification marker, not as a randomly assigned exposure or a complete measure of social position.

S3. Clock coordinate and primary placement estimand

The 04:00-to-following-04:00 midpoint coordinate and respondent-day placement estimator P_{id} are defined in the main Methods. Implementation used 96 fixed 15-minute bins and advanced after-midnight midpoints by 1,440 minutes before aggregation, thereby placing them at the end of the same analytic day.

For reproducible group aggregation, the duration-weighted group mean was

$$\mu_{gd} = \frac{\sum_i w_i \sum_b M_{idb} x_b}{\sum_i w_i \sum_b M_{idb}} \quad (2)$$

where w_i is the official respondent-day survey weight, M_{idb} is the number of domain minutes and x_b is the unwrapped bin midpoint. The endpoint placement gap was $\Delta_d^p = \mu_{Hd} - \mu_{Ld}$: positive values indicate later placement for the bachelor-plus group.

S4. Survey weights and replicate uncertainty

ATUS nonresponse and survey design require the supplied final and replicate weights. Full-period points use TU20FWGT in 2020 and TUFNWGTP in other years. The supplied 2020 ATUS replicate weights do not provide valid positive weights, so analyses requiring replicate recomputation exclude 2020 and retain all other supported years.

For an estimand $\hat{\theta}$ and its recomputation $\hat{\theta}_r$ under replicate weight $r=1,\dots,160$, the design variance used in these analyses is

$$Var(\hat{\theta}) = \frac{4}{160} \sum_{r=1}^{160} (\hat{\theta}_r - \hat{\theta})^2 \quad (3)$$

Displayed uncertainty ranges are $\hat{\theta} \pm 1.96SE(\hat{\theta})$. Window-share, decomposition and component displays use the replicate-supported non-2020 population for both their point estimates and replicate recomputations. Full-period primary placement and transport points retain 2020 through the official final weight. The manuscript labels the ranges as replicate-weight uncertainty ranges to distinguish the matched non-2020 uncertainty population from the full-period point-estimate population.

S5. Window-share summaries

Window shares link the continuous placement distribution to afternoon and evening contexts. For group g , domain d and window W ,

$$S_{gdW} = \frac{\sum_i w_i \sum_b M_{idb} 1(x_b \in W)}{\sum_i w_i \sum_b M_{idb}} \quad (4)$$

The primary windows are 12-18 and 18-24. Their endpoint differences are $\Gamma_{dW} = S_{HdW} - S_{LdW}$. These summaries describe the fraction of each group's weighted domain minutes occurring in each window; they are neither respondent prevalence nor substitutes for the continuous placement estimand.

S6. Optimal transport and duration standardization

For each endpoint and domain, survey-weighted domain minutes form a probability distribution over the 96 bin midpoints. The exact one-dimensional Wasserstein-1 and Wasserstein-2 distances, quantile displacement $D(u)$ and translation/residual-shape identity are defined in the main Methods. [Main Fig.2a](#) evaluated $D(u)$ at $u = 0.05, 0.10, \dots, 0.95$, preserved the 15-minute source resolution and overlaid the corresponding mean placement gap. The reported translation share was $(\mu_H - \mu_L)^2 / W_2^2$. Equal-day weighting first normalized each respondent-day's domain distribution to sum to one before survey-weight aggregation. Further sensitivities used a leisure/care common-domain sample and alternative 03:00 and 05:00 analytic origins.

Duration accounting assessed whether placement was redundant with respondent-day domain duration. Positive-domain respondent-days were assigned to pooled weighted duration strata. Let s_{gq} be endpoint g 's share of weighted domain minutes contributed by duration stratum q , and let m_{gq} be mean placement within that stratum. With $\bar{s}_q = (s_{Hq} + s_{Lq}) / 2$ and $\bar{m}_q = (m_{Hq} + m_{Lq}) / 2$, the symmetric identity is

$$P_H - P_L = \sum_q (s_{Hq} - s_{Lq}) \bar{m}_q + \sum_q (m_{Hq} - m_{Lq}) \bar{s}_q \quad (5)$$

The first term is the domain-duration-distribution component and the second the common-duration-standardized placement gap. Quintiles were primary; quartiles and deciles were sensitivity specifications. Cut points were frozen from the pooled point-weight distribution and reused in all 160 replicate recomputations.

S7. Recorded sleep and observed waking interval

Total recorded sleep among all aged-30-plus respondent-days was

$$S_i = \sum_{a \in \{010101, 010199\}} d_{ia} \quad (6)$$

where d_{ia} is the duration of an ATUS sleeping episode (010101) or sleep-not-elsewhere-classified episode (010199). Sleeplessness (010102) was reported separately and excluded from S_i . Endpoint means used respondent-day rather than activity-minute weights. Composition accounting first standardized the pooled endpoint distribution of age bands (30-44, 45-64, 65-plus) and weekday/weekend, then additionally included employed versus non-employed status as sensitivity.

The waking-interval analysis used diaries with sleep touching both 04:00 boundaries. Let A_i be the end of the initial boundary sleep episode and O_i the start of the terminal boundary sleep episode on the unwrapped 04:00-to-04:00 clock axis. The observed waking interval was $W_i = O_i - A_i$. The primary route retained intervals of 8-20 hours; a 6-22-hour route was a boundary sensitivity. Initial boundary sleep, terminal boundary sleep and sleep episodes inside W_i were separately reconstructed and reconciled with summary sleep minutes. Normalized waking-phase placement was calculated by subtracting wake time from the episode-based, duration-weighted mean activity time and dividing by the observed waking interval. The all-respondent sleep sample, boundary-eligible waking sample and domain-positive normalized-phase samples remain analytically distinct.

S8. Activity composition and component mapping

The symmetric decomposition of the broad-domain placement gap into component composition and within-component timing terms is defined in the main Methods. Point and replicate decompositions were calculated in the replicate-supported years excluding 2020 and had to reconcile with the corresponding broad-domain placement gap.

The four core leisure subtypes were screen-based, audio, gaming and computer leisure; relaxing and personal hobbies; social and communication activities; and sports, exercise and recreation. Religious, civic and volunteering activity was retained as a boundary subtype. Waiting, security and other leisure and unmapped leisure-like activities were excluded and counted. Care/household obligation was divided into home-anchored household tasks, direct care and service-facing household errands. Home-anchored tasks comprised unpaid-domestic activity. Service-facing errands comprised consumer, professional-service, household-service and government/civic-service activity. Direct care comprised household and non-household care. A small care/service communication boundary category was excluded and separately accounted for with a prespecified 0.10-minute reconciliation tolerance.

S9. Component equality tests and heatmap construction

Let θ contain the component gaps and θ_r the corresponding vector under replicate weight r . The design covariance is

$$\hat{V} = \frac{4}{160} \sum_{r=1}^{160} (\theta_r - \theta)(\theta_r - \theta)^T \quad (7)$$

For a contrast matrix C encoding equality of the included component gaps, the Wald statistic is

$$(C\theta)^T (C\hat{V}C^T)^+ (C\theta) \quad (8)$$

where the superscript $+$ denotes a generalized inverse. The primary leisure test used the four core subtypes with an all-component sensitivity retaining religious, civic and volunteering activity. The care test used the three retained subdomains. P values for the two primary domain tests are adjusted by the Holm method.

For group g , component k and bin b , the heatmap cell is

$$S_{gkb} = \frac{\sum_i w_i M_{ihkb}}{\sum_{b'} \sum_i w_i M_{ihkb'}} \quad (9)$$

where M_{igkb} is the number of component minutes contributed by respondent-day i to bin b , and w_i is the final ATUS weight. Each component-by-endpoint row therefore sums to one. The heatmap therefore shows where the group's weighted component minutes occur across the day, rather than prevalence, total duration or uncertainty. The component-equality tests were omnibus tests sensitive to differences in either direction; P values were calculated from the upper tail of the chi-squared reference distribution using the reported degrees of freedom.

S10. Daily-anchor construction and relative placement

Wake, last-work-end and child-bedtime contexts used separate eligible samples. The wake anchor was the end of the initial sleep episode touching the 04:00 boundary in diaries with sleep at both boundaries and an observed waking span of 8-20 hours. The last-work-end anchor was the

unwrapped end of the final paid-work or study bin among respondent-days with a paid-work or study episode and positive focal-domain time. The child-bedtime context required a household child aged 0-12, a valid first-child wake and last-child bedtime pair and positive focal-domain time. Child bedtime was mapped to the same 04:00-to-following-04:00 coordinate as the activity sequence.

A common-sample architecture aligned seven anchors: wake, first work start, principal work start, principal work end, last work end, child bedtime and final sleep onset. Eligible diaries came from employed respondents aged 30-64 years and recorded a weekday. Respondents had a child aged 0-12 years and valid child and sleep anchors. The observed waking interval lasted 8-20 hours, and the principal paid-work or study block lasted at least 60 minutes. First work start and last work end were the earliest start and latest end of any valid paid-work/study fixed-bin episode. The principal block was the contiguous paid-work/study block containing the most work minutes, with earlier start breaking ties.

For endpoint g , the observed work-envelope span was $E_g = L_g - F_g$, where F_g and L_g were the weighted mean first work start and last work end. The principal-block span was defined analogously. The work envelope may include intervening non-work episodes and is not continuous paid-work duration.

For each eligible anchor context, let P_{gda} denote the domain-minute-weighted mean placement for group g , domain d and anchor context a ; let A_{gda} denote the corresponding mean anchor time; and let $R_{gda} = P_{gda} - A_{gda}$ denote placement relative to that anchor. The bachelor-plus minus less-than-high-school contrast satisfies

$$\Delta P_{da} = \Delta A_{da} + \Delta R_{da} \quad (10)$$

All terms used the official respondent-day weight multiplied by domain minutes and were recomputed under all 160 replicate weights. The share of domain minutes after last work end or child bedtime was a secondary diagnostic. Anchor-eligible placement gaps were compared with broader reference samples to audit selection direction. Wake, work-end and child-bedtime estimates were not pooled because they arose from different eligible populations.

S11. Structural and work-time accounting

Endpoint estimates are standardized to common distributions over prespecified categories. Broad structure combines age, employment and household-child composition. Work-clock categories are constructed from paid-work/study fixed-bin episodes on the diary day. Occupational and job-structure routes are supported occupation, industry, class-of-worker, job-multiplicity and work-status/hour categories. A separate sensitivity restricts the sample to respondent-days without paid-work/study fixed-bin episodes.

For a given domain and standardization route, let D_{gc} denote the final-weighted domain minutes contributed by education group g in structure category c . Categories were retained when

both education groups had positive domain-time support and at least 30 eligible respondent-days. Among the supported categories C^* , the common target share was

$$t_c = \frac{D_{Lc} + D_{Hc}}{\sum_{k \in C^*} (D_{Lk} + D_{Hk})} \quad (11)$$

If μ_{gc} denotes mean placement within category c , the standardized endpoint gap was

$$\Delta^{std} = \sum_{c \in C^*} t_c (\mu_{Hc} - \mu_{Lc}) \quad (12)$$

We extended the same standardization to the full clock-time distributions. Let p_{gcb} denote education group g 's normalized distribution of domain minutes across the 96 clock bins within structure category c . The standardized distribution was

$$P_{gb}^{std} = \sum_{c \in C^*} t_c p_{gcb} \quad (13)$$

We then recalculated the mean placement gap, Wasserstein-1 and Wasserstein-2 distances, and residual-shape distance defined in Supplementary Methods S6. The supported categories and pooled target shares were fixed across all 160 replicate-weight recomputations. Point estimates retained the special 2020 final weight; replicate-weight ranges excluded 2020.

Broad structure, work-clock and selected job-structure estimates are parallel accounting routes, not sequential adjustments. Work-day, work-free-day, weekday and weekend comparisons are retained as context diagnostics and are not factorially crossed.

S12. Historical, regional and Canadian context

Period-level estimates were calculated for 2003-2007, 2008-2014, 2015-2019, 2021 and 2022-2024 using replicate-supported data. The fourth period-level estimate used 2021 alone. The annual companion analysis applies TU20FWGT in 2020 and TUFNWGTP otherwise. The 2020 estimate is retained as a final-weight point without replicate-weight uncertainty. For non-2020 years, endpoint-by-year contrasts use respondent final weight multiplied by domain minutes and are recomputed under all 160 replicate weights. A sensitivity standardizes supported annual cells to common age-band and weekday/weekend distributions.

State and District of Columbia displays use the replicate-supported non-2020 route. A jurisdiction is colored when each education endpoint contains at least 30 eligible respondent-days; unsupported jurisdictions are hatched. Map colors encode point estimates, while replicate diagnostics are tabulated separately.

Canadian Multinational Time Use Study waves are reported separately by country-wave rather than pooled across discontinuous years. Their harmonized high-minus-low education contrast is not design-equivalent to the ATUS bachelor-plus minus less-than-high-school contrast.

S13. OpenStreetMap service-clock benchmark

The service-clock benchmark uses OpenStreetMap opening hour records for public-facing facilities in all 50 states and the District of Columbia. Facilities are grouped into leisure-facing and care/household-facing families. Parseable weekly schedules are evaluated over a fixed reference weekday without holiday modeling. For each facility, the analysis calculates the fraction of 12-18 and 18-24 recorded as open; family and jurisdiction summaries average these facility-level shares.

Coverage accounting reports the number of relevant points of interest, the number with opening hours, the number with parseable schedules and the corresponding shares. OpenStreetMap is volunteered geographic information whose completeness varies across places and feature types (Barrington-Leigh,2017)(Herfort,2023). The benchmark represents contemporary facility context, is not linked to ATUS respondent-days and does not reconstruct historical opening hours for 2003-2024.

S14. Waiting-time and Well-Being Module context

Waiting summaries use the same 12-18 and 18-24 windows. The reported quantities are waiting minutes per respondent-day, waiting episodes per respondent-day and mean window-clipped minutes per waiting episode among observed waiting episodes. They describe realized waiting conditions within each window and education group rather than an individual-level pathway from placement to waiting.

The ATUS Well-Being Module assigns activity-level weights to selected diary activity reports (Kahneman, 2004)(Stone,2018). For indicator k , domain d and window W ,

$$M_{kdW} = \frac{\sum_{j \in J_{kdW}} v_j y_{jk}}{\sum_{j \in J_{kdW}} v_j} \quad (14)$$

where y_{jk} is the reported value and v_j the WBM activity weight. The displayed contrast is

$$C_{kd}^{WBM} = M_{kd,18-24} - M_{kd,12-18} \quad (15)$$

The indicators were happiness, meaningfulness, interaction during the activity (including by telephone), stress, tiredness, sadness and pain. This is an activity-context comparison, not an education-group experience comparison.

For the separate behavioral calibration, ATUS Who reports were joined to the sampled WBM leisure activities. The strict reported-alone category used Who codes 18 and 19. The comparison category required at least one recorded companion and no strict-alone code. WBM interaction was coded from the question asking whether the respondent was interacting with anyone during the activity, including over the phone. The primary pooled analysis used adults aged 30 years and older in 2010, 2012 and 2013 and the adjusted WBM activity weight. Interaction rates were recomputed with all 160 activity replicate weights. The common-context standardization retained year by weekday/weekend by activity-start time band by leisure-setting cells containing at least 20 unweighted observations in each companion status. It used the mean of

the two status-specific weighted cell distributions as the common target. The original activity weight, all-age sample and 2021 module were sensitivity or boundary analyses.

S15. Population-rhythm distance, prevalence alignment and reference-composition audit

For each domain d , education endpoint g and supported survey-year by weekday/weekend cell c , survey-weighted domain minutes formed a normalized 96-bin distribution P_{gdc} . Let s_{gdc} be endpoint g 's share of activity-minute mass in cell c , and let

$$\bar{s}_{dc} = \frac{S_{Hdc} + S_{Ldc}}{2} \quad (16)$$

The common-cell-standardized endpoint distribution was

$$P_{gd} = \sum_c \bar{s}_{dc} P_{gdc} \quad (17)$$

The observed population reference Q_d used all survey-covered respondent-days aged 15 years or older in the replicate-supported non-2020 years. It was aggregated with the same common cell weights:

$$Q_d = \sum_c \bar{s}_{dc} Q_{dc} \quad (18)$$

On the ordered 04:00-following-04:00 bin midpoints, exact discrete one-dimensional Wasserstein-1 distance was

$$W_1(P, Q) = \sum_{b=1}^{95} |F_P(x_b) - F_Q(x_b)| (x_{b+1} - x_b) \quad (19)$$

where F_P and F_Q are cumulative masses evaluated at bin midpoint x_b . The reference-distance contrast was

$$R_d = W_1(P_{Hd}, Q_d) - W_1(P_{Ld}, Q_d) \quad (20)$$

Negative R_d means that the bachelor-plus distribution is mathematically closer to the reference distribution. It does not imply better coordination, greater access or welfare.

The observed reference was partitioned exhaustively into respondents younger than 30 years, age-30-plus less-than-high-school respondents, age-30-plus middle-education respondents and age-30-plus bachelor-plus respondents. Within cell c , let Q_{kdc} be partition k 's normalized domain distribution and π_{kdc} its share of observed reference activity-minute mass. Partition k 's contribution to the standardized reference was

$$\Pi_{kd} = \sum_c \bar{s}_{dc} \pi_{kdc}, \quad \sum_k \Pi_{kd} = 1 \quad (21)$$

Two reference-balancing specifications were primary. The equal-adult reference assigned equal weight within each cell to the lowest, pooled middle and highest adult education-group distributions,

$$Q_{dc}^{eq} = \frac{Q_{Ldc} + Q_{Mdc} + Q_{Hdc}}{3} \quad (22)$$

The middle-education-only reference Q_{Mdc} excluded both focal endpoints. As additional sensitivities, the observed-adult reference excluded respondents younger than 30 years, and a hybrid reference retained the observed younger-than-30 activity-time share α_{dc} while equalizing the adult education groups:

$$Q_{dc}^{hyb} = \alpha_{dc} Q_{Udc} + (1 - \alpha_{dc}) Q_{dc}^{eq} \quad (23)$$

All endpoint distributions, reference distributions, common cell weights and partition contributions were recomputed under each of the 160 ATUS replicate weights. Reported ranges are the point estimate plus or minus 1.96 replicate standard errors and are used for direction accounting rather than formal confidence-interval claims.

The prevalence-retaining alignment score used the reference prevalence level rather than normalizing the reference clock to unit mass. Let r_{dcb} be the survey-weighted share of the reference population engaged in domain d at bin b in cell c , including respondent-days with zero domain time in the denominator, and let M_{gdc} be endpoint g 's weighted domain-minute mass. The endpoint alignment score was

$$A_{gd} = \sum_c \bar{s}_{dc} \frac{\sum_b M_{gdc} r_{dcb}}{\sum_b M_{gdc}}, \quad G_d = A_{Hd} - A_{Ld} \quad (24)$$

Thus, A_{gd} is the mean reference-population same-domain prevalence at the clock positions occupied by endpoint domain minutes. When G_d is expressed in percentage points, the illustrative difference per 1,000 reference persons is

$$T_d = 10G_d \quad (25)$$

This is a deterministic unit conversion. It is not an estimate of observed encounters, co-presence, people affected or cumulative population impact.

For the broad-domain decomposition, let p_{gk} be subtype k 's standardized share of domain minutes and let a_{gk} be the mean broad-domain reference prevalence at the clock positions occupied by that subtype. The symmetric identity was

$$G_d = \sum_k (p_{Hk} - p_{Lk}) \frac{a_{Hk} + a_{Lk}}{2} + \sum_k (a_{Hk} - a_{Lk}) \frac{p_{Hk} + p_{Lk}}{2} \quad (26)$$

The first term is the subtype-composition component and the second the within-subtype timing component. Separately, each subtype's endpoint minutes were evaluated against the reference prevalence clock for that same subtype. Those same-subtype gaps are not additive contributions to G_d . The full reference included survey-covered respondents aged 15 years or older; adult and middle-education-only references were retained as composition sensitivities.

Reported-alone clock calibration

ATUS Who reports identify whether the respondent described an activity as unaccompanied; they do not record unknown people present in a public setting. For each of five period bins crossed with weekday/weekend status, we estimated the reference population's leisure prevalence and the share of Who-eligible leisure minutes reported alone in each of the 96 clock bins. The primary reference included survey-covered people aged 15 years and older. A sensitivity reference excluded both focal education groups and retained respondents younger than 30 years plus the middle adult education groups.

The clock-cell association was summarized as the weighted change in reported-alone percentage points for a 10-percentage-point increase in population leisure prevalence. Relation weights combined the common endpoint period-by-day-type distribution with Who-eligible leisure support in each clock bin. We then projected each endpoint's normalized leisure timing distribution onto the same clock-specific reported-alone curve and compared the projected rates. Projection mass exceeded 99% in every retained comparison. This is a placement-only context projection; it is distinct from each group's observed companion composition. A non-household-adult companion outcome was retained as a boundary analysis because ATUS Who records companions but not ambient strangers or all social contact.

Supplementary Results

Supplementary Note 1. Annual recurrence and domain-qualified change

The annual analysis separates recurrence, whether the placement gap retains the same direction, from average change in its magnitude. Methods and weighting are specified in Supplementary Methods S12.

The endpoint placement gap was positive in all 22 survey years for both domains (Supplementary Table 1; Supplementary Fig.1). In each of the 21 replicate-supported years, the lower endpoint of the replicate-weight uncertainty range was also positive. The separate 2020 final-weight point estimates were 35.9 minutes for leisure and 52.2 minutes for care/household obligation.

The unstandardized annual interaction summarized average narrowing in both domains. After standardization for age and weekday or weekend composition, the uncertainty range for the estimated annual change included zero for leisure but remained negative for care or household

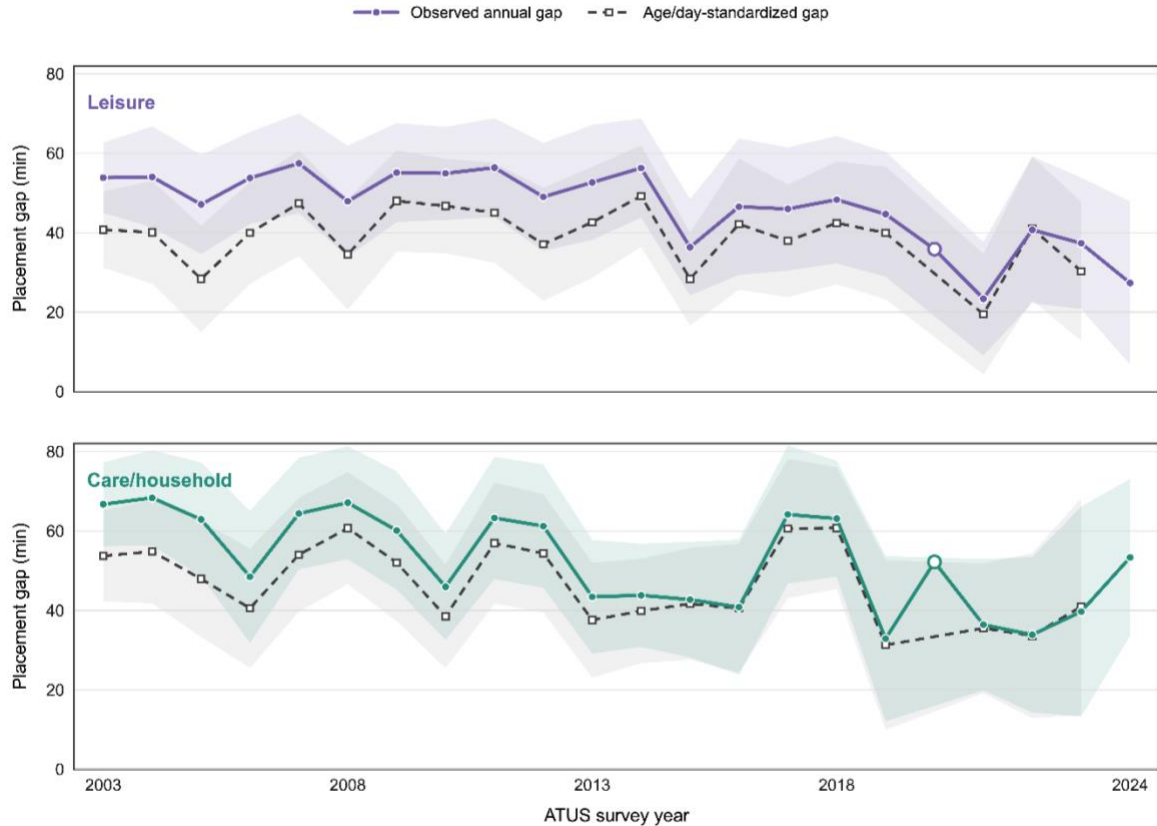
activity. The bounded interpretation is recurrence in both domains and domain-qualified average narrowing, not a common monotonic trend.

Supplementary Table 1. Annual education-group placement gaps. Values are bachelor-plus minus less-than-high-school placement gaps in minutes. Ranges are estimate plus or minus 1.96 replicate standard errors and are not available for 2020.

Year	Leisure gap	Leisure range	Care gap	Care range
2003	53.8	45.0 – 62.7	66.7	56.2 – 77.3
2004	54.0	41.3 – 66.7	68.4	56.4 – 80.4
2005	47.1	34.7 – 59.6	62.9	48.6 – 77.2
2006	53.8	42.2 – 65.4	48.5	31.8 – 65.1
2007	57.5	44.9 – 70.0	64.4	50.4 – 78.4
2008	47.9	33.9 – 61.9	67.1	52.9 – 81.3
2009	55.1	42.7 – 67.6	60.1	45.2 – 75.0
2010	54.9	43.2 – 66.7	45.9	32.6 – 59.3
2011	56.4	43.9 – 68.8	63.3	47.9 – 78.6
2012	49.0	35.5 – 62.6	61.2	45.7 – 76.8
2013	52.7	38.2 – 67.1	43.4	29.1 – 57.7
2014	56.3	43.8 – 68.7	43.8	30.8 – 56.8
2015	36.3	24.2 – 48.4	42.7	28.2 – 57.2
2016	46.5	29.4 – 63.7	40.8	23.8 – 57.8
2017	46.0	30.5 – 61.5	64.2	46.8 – 81.6
2018	48.3	32.3 – 64.3	63.1	48.5 – 77.7
2019	44.7	29.0 – 60.4	32.9	12.1 – 53.7
2020	35.9	not available	52.2	not available
2021	23.4	9.1 – 37.7	36.4	19.9 – 52.9
2022	40.7	22.3 – 59.2	33.9	14.3 – 53.4
2023	37.3	20.8 – 53.9	39.7	13.3 – 66.1
2024	27.4	6.8 – 47.9	53.3	33.6 – 73.1

Supplementary Table 2. Annual recurrence and average-change summaries.

Domain	Positive years	Primary change, min/year	Age/day standardized, min/year	Interpretation
Leisure	22/22	-0.99 (-1.49 to -0.50)	-0.35 (-0.88 to 0.17)	Narrowing unresolved after age/day standardization
Care/household obligation	22/22	-1.24 (-1.90 to -0.68)	-0.72 (-1.40 to -0.04)	Narrowing retained after age/day standardization



Supplementary Fig. 1. Annual education-group placement gaps and composition sensitivity. Solid colored lines show observed annual bachelor-plus minus less-than-high-school placement gaps. Shading shows replicate-weight uncertainty ranges. The open 2020 point uses the official final weight and has no replicate-weight range. Dashed square-marked lines show annual gaps standardized to common age-band and weekday/weekend distributions where the pre-specified support criterion was met.

Supplementary Note 2. Distributional transport details

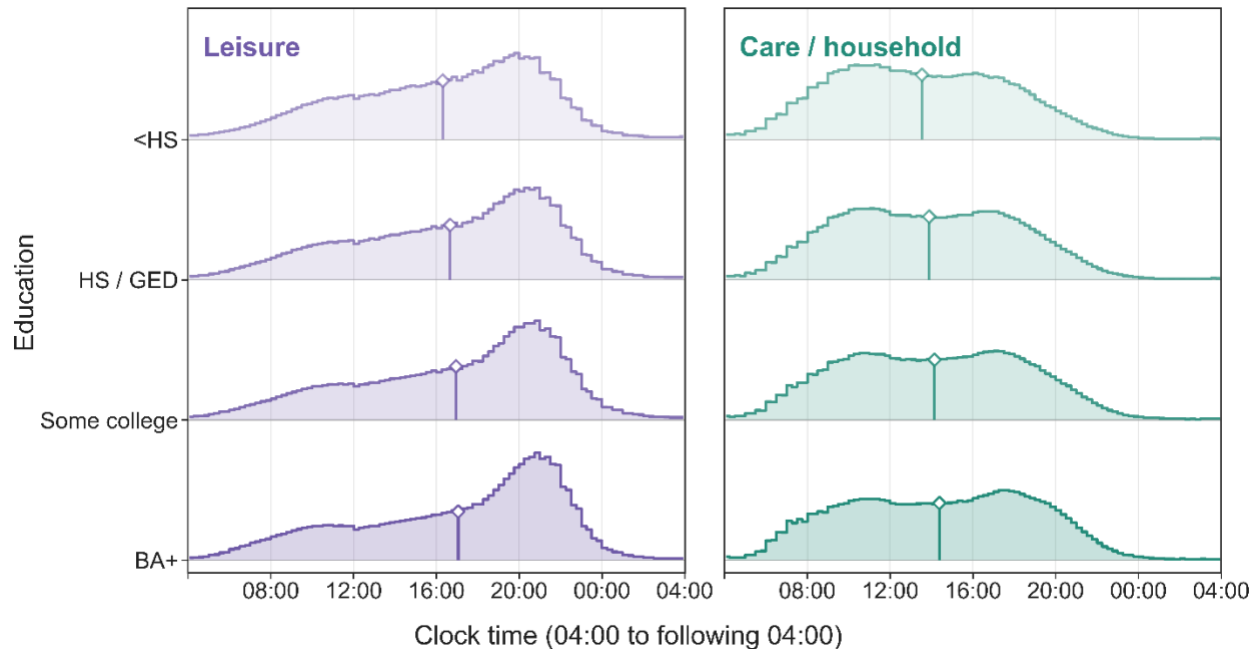
The main text reports the Wasserstein-2 distance and the residual-shape distance because these quantities directly distinguish broad mean translation from remaining distributional reorganization. For completeness, Supplementary Table 3 retains the parallel Wasserstein-1 distance, the mathematical translation share of squared Wasserstein-2 distance and the quantile-direction accounting. These are distributional descriptions, not explained or mediated proportions.

Supplementary Table 3. Full-period optimal-transport diagnostics.

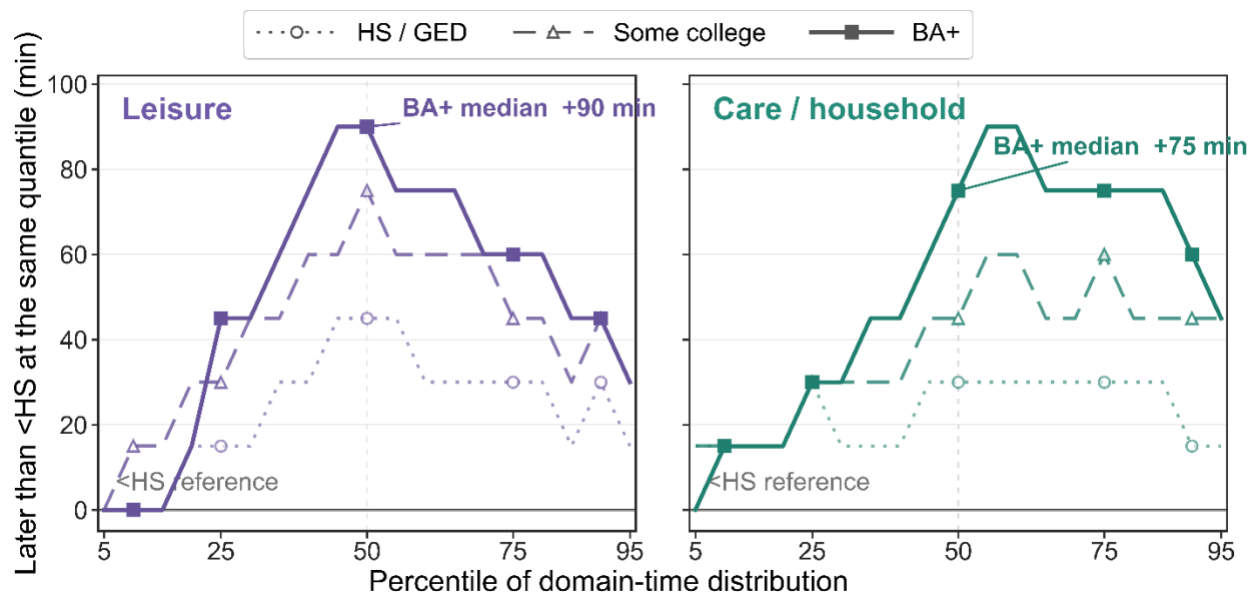
Domain	W1, min	W2, min	Residual shape, min	Mean-translation share of squared W2	Positive quantile displacements	Replicate-supported residual shape, min
Leisure	45.6	52.9	28.0	71.9%	16/19	28.0 (26.7 to 29.2)
Care/household obligation	51.1	57.4	26.1	79.2%	18/19	26.5 (24.7 to 28.3)

Full-period values use the official final-weight route. Replicate-supported residual-shape summaries exclude 2020 because the supplied 2020 replicate weights do not support the design-based recomputation. Positive quantile displacement means that the bachelor-plus quantile occurs

later than the corresponding less-than-high-school quantile on the unwrapped 04:00-to-following-04:00 clock.



Supplementary Fig. 2 Complete clock-time distributions across four education groups. Ridgelines show survey-weighted domain-minute mass in each observed 15-minute bin for leisure and care or household activity. All eight distributions use the same vertical scale. Vertical guides and diamond markers show duration-weighted mean placement. Curves are unsmoothed and the clock runs from 04:00 to the following 04:00. The display does not imply that the groups are ordered at every clock bin.



Supplementary Fig. 3 Quantile displacement across four education groups. Lines show each non-reference education group's quantile position minus the corresponding less-than-high-school quantile at cumulative shares from 5% to 95%. Positive values indicate later clock-time placement than the reference at that cumulative share. Line style and marker identify education group; color identifies domain. The bachelor-plus curves reconcile to main Fig. 2a. Values are point estimates.

Supplementary Note 3. Sensitivity analyses

The analyses named as sensitivities in the Methods retained the conclusions they were designed to check. Equal respondent-day weighting and 03:00 or 05:00 analytic origins preserved positive mean placement and residual distributional shape in both domains ([Supplementary Table 4](#)). The restricted common leisure/care workday sample also retained non-zero distributional distance, with W2 distances of 37.7 minutes for leisure and 29.8 minutes for care/household obligation. This diagnostic reports distance rather than a directional mean gap. Unlike the full-period estimates in Supplementary Table 3, the estimates in Supplementary Table 4 exclude 2020, which accounts for the small differences in baseline values.

Supplementary Table 4. Optimal-transport weighting and clock-origin sensitivities

Domain	Route	Mean gap, min	W2, min	Residual shape, min
Leisure	Domain-minute weighting	45.3	53.2	28.0
Leisure	Equal respondent-day weighting	53.4	64.5	36.1
Leisure	03:00 origin	46.7	54.0	27.0
Leisure	04:00 origin	45.3	53.2	28.0
Leisure	05:00 origin	42.4	52.0	30.2
Care/household	Domain-minute weighting	51.2	57.6	26.5
Care/household	Equal respondent-day weighting	71.9	82.5	40.4
Care/household	03:00 origin	51.5	57.9	26.4
Care/household	04:00 origin	51.2	57.6	26.5
Care/household	05:00 origin	47.2	55.8	29.7

Duration standardization was also insensitive to the number of strata. The leisure residual was smaller than the broad gap but remained positive under quartiles, quintiles and deciles; the care/household residual remained close to 50 minutes under all three specifications.

Supplementary Table 5. Duration-stratum sensitivity

Domain	Duration strata	Standardized gap, min	Replicate-weight range
Leisure	Quartiles	19.7	16.9 to 22.6
Leisure	Quintiles	18.7	15.9 to 21.5
Leisure	Deciles	18.1	15.3 to 21.0
Care/household	Quartiles	49.9	46.8 to 53.1
Care/household	Quintiles	49.9	46.8 to 53.0
Care/household	Deciles	50.1	47.0 to 53.2

Expanding the accepted observed waking interval from 8-20 to 6-22 hours left the normalized-phase and relative-offset conclusions nearly unchanged. The leisure component-equality test likewise remained decisive when religious, civic and volunteering activity was retained as a fifth subtype.

Supplementary Table 6. Waking-frame and component-test sensitivities

a. Waking-frame sensitivity

Domain	Frame	Normalized-phase gap	Wake-relative gap, min	Clock gap, min
Leisure	8-20 h primary	0.034	56.6	40.7
Leisure	6-22 h sensitivity	0.035	57.1	41.2
Care/household	8-20 h primary	0.048	60.2	50.9
Care/household	6-22 h sensitivity	0.048	60.9	51.4

b. Component equality tests

Test	df	Wald statistic	Unadjusted P
Four core leisure subtypes	3	216.2	1.3e-46
Five leisure subtypes	4	335.1	2.8e-71
Three care/household subdomains	2	129.0	9.9e-29

The two primary domain P values are Holm-adjusted in the main text. The all-component leisure row is a sensitivity result and is not included in that two-domain adjustment family.

Supplementary Note 4. Regional uncertainty and Canadian context

The state maps display point estimates only. Among supported jurisdictions, all point estimates were positive; replicate-weight uncertainty ranges were entirely positive in 39 of 47 leisure jurisdictions and 36 of 45 care/household jurisdictions ([Supplementary Table 7](#)). The maps therefore describe broad directional recurrence rather than state-level precision.

Supplementary Table 7. State and District of Columbia replicate diagnostics

Domain	Supported jurisdictions	Positive point estimates	Ranges entirely positive	Median gap, min	Point range, min
Leisure	47	47	39	51.1	2.0 to 82.8
Care/household	45	45	36	45.7	0.2 to 100.2

Canadian waves retained the same directional episode-midpoint placement contrast for both primary domains ([Supplementary Table 8](#)). These estimates use harmonized high-minus-low education categories and are not design-equivalent replications of the ATUS contrast.

Supplementary Table 8. Canadian country-wave placement context

Year	Domain	High-minus-low gap, min	Low-group days	High-group days
1992	Leisure	14.9	1634	2219
1992	Care/household	52.7	1752	3017
2005	Leisure	21.9	2477	6067
2005	Care/household	56.7	2664	8573
2010	Leisure	18.2	1731	5779
2010	Care/household	52.4	1854	8037
2015	Leisure	15.6	1547	6401
2015	Care/household	53.0	1728	9263

Supplementary Note 5. Population-rhythm reference composition and prevalence alignment

The observed population rhythm contains unequal contributions from the education endpoints, so proximity to that rhythm could partly reflect who defines the reference. [Supplementary Table 9](#) quantifies the standardized contribution of each age and education partition. The middle-education group contributed the largest activity-time share for both domains; bachelor-plus activity-time mass nevertheless exceeded the less-than-high-school contribution.

[Supplementary Table 10](#) reports the high-minus-low W1 distance contrast under all five reference templates. Negative values mean that the bachelor-plus distribution is closer. For leisure, the direction remained resolved under both equal adult education-group weighting and the middle-education-only reference that excluded both endpoints. For care/household obligation, both balanced contrasts were close to zero and their replicate-weight uncertainty ranges crossed zero. The population-rhythm result is therefore retained for leisure only; the care/household result is reference-dependent.

The complementary prevalence-retaining score preserves how many reference-population members occupy the domain at each clock-time. [Supplementary Table 11](#) reports both broad domains and the exact subtype-composition and within-subtype timing decomposition. For leisure, composition acted against the observed ordering, whereas within-subtype timing produced a larger positive component. [Supplementary Table 12](#) reports all seven same-subtype contrasts. The screen-based leisure contrast was the largest positive result, whereas care/household subtypes differed in direction.

Supplementary Table 9. Contributions to the observed standardized population activity rhythm

Domain	Under 30	Age 30+, <HS	Age 30+, middle	Age 30+, BA+	BA+ / <HS contribution ratio
Leisure	24.13 (23.95 to 24.30)	9.70 (9.53 to 9.87)	43.30 (43.05 to 43.55)	22.88 (22.67 to 23.08)	2.36
Care/household	17.15 (16.90 to 17.40)	8.80 (8.60 to 9.00)	45.32 (44.98 to 45.67)	28.72 (28.43 to 29.02)	3.26

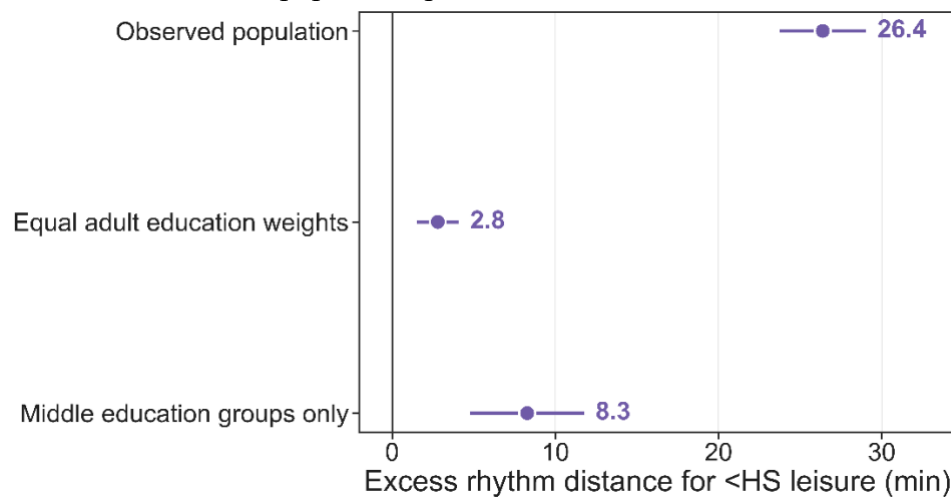
Values are percentage contributions with replicate-weight uncertainty ranges in parentheses. Contributions sum to 100% within domain after common survey-year by weekday/weekend standardization.

Supplementary Table 10. Endpoint distance contrast across reference templates

Domain	Observed 15+	age 30+	Observed 30+	age 15+	Equal education	adult	Middle only	Under-30 +equal adult	observed
Leisure	-26.4 (-28.9 to -23.8)	-	-12.3 (-14.6 to -10.0)	-	-2.8 (-4.0 to -1.6)	-	-8.3 (-11.7 to -4.9)	-20.0 (-21.9 to -18.0)	-
Care/household	-27.9 (-30.9 to -24.9)	-	-12.9 (-15.6 to -10.1)	-	-0.7 (-2.1 to 0.7)	-	-1.0 (-5.0 to 3.1)	-18.6 (-20.7 to -16.5)	-

Values are bachelor-plus minus less-than-high-school W1 distance contrasts in minutes, with replicate-weight uncertainty ranges in parentheses. Negative values indicate that the bachelor-

plus distribution is closer to the specified reference. Equal-adult and hybrid references are sensitivity estimands rather than population prevalence estimates.



Supplementary Fig. 4 summarizes the leisure contrast under the observed population, equal-adult-education and middle-education-only reference rhythms. It reports the same quantities as Supplementary Table 10 with the sign reoriented so that positive values mean that lower-education leisure is farther from the specified rhythm.

Supplementary Table 11. Prevalence-retaining population-rhythm alignment

Domain	<HS alignment,%	BA+ alignment, %	BA+ minus <HS, pp	Illustrative difference per 1,000	Composition component, pp	Within-subtype timing component, pp
Leisure	34.38	36.26	+1.88 (1.72 to 2.05)	+18.8	-0.65	+2.53
Care/household	18.68	18.31	-0.37 (-0.43 to -0.30)	-3.7	-0.04	-0.33

Alignment is the mean share of the survey-covered age-15-plus reference population engaged in the same broad domain at the clock times occupied by endpoint domain minutes. Values are standardized to a common survey-year by weekday/weekend distribution. Ranges are estimate plus or minus 1.96 replicate standard errors. The illustrative per-1,000 value equals ten times the percentage-point contrast; it is not an observed count of encounters or affected people. Composition and within-subtype timing components sum to the broad-domain gap apart from rounding.

Supplementary Table 12. Same-subtype population-rhythm alignment

Domain	Subtype		<HS alignmen t, %	BA+ alignmen t, %	BA+ minus <HS, pp	Illustrativ e difference per 1,000	Middle- only referenc e gap, pp
Leisure	Screen/audio/games/computer		22.41	26.06	+3.65 (3.47 to 3.82)	+36.5	+3.90
Leisure	Relaxing/personal hobbies		3.87	3.77	-0.10 (-0.13 to -0.07)	-1.0	-0.22
Leisure	Social/communication/events		6.35	6.51	+0.16 (0.10 to 0.23)	+1.6	0.00
Leisure	Sports/exercise/recreation		2.12	2.14	+0.02 (-0.02 to 0.06)	+0.2	+0.02
Care/househol d	Home-anchored tasks	household	12.04	11.72	-0.32 (-0.37 to -0.26)	-3.2	-0.53
Care/househol d	Direct care		3.53	3.59	+0.06 (0.02 to 0.10)	+0.6	+0.02
Care/househol d	Service-facing errands	household	3.97	3.97	0.00 (-0.06 to 0.05)	0.0	-0.03

Subtype labels are shortened for display; complete definitions are given in Supplementary Methods S8. In this table, Care denotes care/household obligation. Each row uses the population clock for the same subtype, including zero-subtype respondent-days in the reference prevalence denominator. The middle-only column excludes both focal education endpoints from the reference. These same-subtype gaps are separate descriptive contrasts and are not additive contributions to the broad-domain gap in [Supplementary Table 11](#). Social/communication/events changes direction under the middle-only reference; sports/exercise/recreation and service-facing household errands have full-reference replicate-weight ranges that cross zero.

Supplementary Note 6. Recorded sleep, waking span and daily-boundary analyses

Recorded sleep and the boundaries of the observed day did not support a uniformly delayed-day interpretation. Across all respondent-days among adults aged 30 years and older, recorded sleep averaged 548.6 minutes in the less-than-high-school group and 499.6 minutes in the bachelor-plus group. The four-category means declined in education order from 548.6 to 520.0, 507.6 and 499.6 minutes. In replicate-supported years excluding 2020, the endpoint contrast was -49.4 minutes (-52.3 to -46.5). Standardization to common age-band and weekday/weekend distributions left -47.0 minutes, and adding employment composition left -36.3 minutes. These diary episodes measure recorded sleep activity rather than sleep quality, need or deprivation.

In the narrower sample with sleep touching both 04:00 boundaries and an observed primary waking interval of 8--20 hours, the bachelor-plus group woke 20.0 minutes earlier and reached final sleep onset 26.3 minutes later than the less-than-high-school group. The observed waking interval was therefore 46.3 minutes longer (43.5 to 49.0). Leisure and care/household placement remained later after normalization to this interval, by 0.034 and 0.048 of the waking span.

Work-boundary estimates came from employed weekday respondents aged 30--64 with valid child, sleep and work anchors: 702 less-than-high-school and 7,699 bachelor-plus respondent-days. The difference in first work start was unresolved at -6.7 minutes (-21.6 to 8.3), as was principal work start at 8.0 minutes (-9.6 to 25.6). Principal work end occurred 31.8 minutes later (12.7 to 51.0) and last work end 61.4 minutes later (45.3 to 77.5). The first-to-last work envelope was 68.0 minutes longer (52.7 to 83.4), and the principal contiguous block was 23.8 minutes longer (14.0 to 33.6). The work envelope can include intervening non-work episodes and is not continuous paid-work duration.

Relative to each group's wake boundary, leisure and care/household activity occurred 56.6 and 60.2 minutes later for the bachelor-plus group. Relative to each group's child-bedtime boundary, the corresponding differences were 43.9 and 59.1 minutes. Around last work end, activity-relative terms of -35.4 and -19.4 minutes partly offset substantially later work endings, and the care/household uncertainty range crossed zero. These estimates arise from different eligible samples and are not pooled into a single pathway. Their role is to show that later leisure and care/household placement did not arise from an intact day shifted uniformly later.

Supplementary Note 7. Reported-alone and interaction calibration

The companion-context analysis was designed as a bounded calibration of the population-rhythm result. It does not turn temporal co-availability into an observed encounter measure. Clock cells with greater population leisure prevalence had a lower share of eligible leisure minutes reported alone. Under the observed age-15-plus reference, a 10-percentage-point increase in population leisure prevalence corresponded to a 4.94-percentage-point decrease in reported-alone leisure (-5.06 to -4.83). Applying the common clock-specific curve to the two education-group timing distributions projected 0.57 percentage points more reported-alone context for the less-than-high-school schedule (0.46 to 0.69). The corresponding endpoint-excluded estimate was 0.48 points (0.36 to 0.60).

The non-household-adult indicator did not support the same placement interpretation. Its clock-cell association with population leisure prevalence was positive, but projecting the education-group distributions produced more non-household-adult context for the lower-education schedule. ATUS Who reports known companions rather than ambient strangers, so this indicator cannot represent all opportunities for contact in public settings. We therefore retain it as a boundary result rather than use it to validate social mixing.

In the WBM pooled 2010, 2012 and 2013 sample, 27.31% of leisure activities reported alone included interaction, compared with 79.83% of leisure activities with a recorded companion. The 52.51-percentage-point contrast (50.00 to 55.02) was similar in each primary year. Standardizing both statuses to a common year, weekday/weekend, activity-start-time-band and leisure-setting distribution left a 48.37-point contrast (45.67 to 51.07). The standardization retained 33 of 66 context cells and 97.73% of valid activity weight. The WBM item includes telephone interaction and does not identify the interaction partner.

Supplementary Table 13 Clock-specific companion-context calibration

Reference	Outcome	Change per 10-point higher population prevalence, pp	<HS leisure projected rate, %	BA+ projected rate, %	<HS BA+, pp	minus
Observed 15+	age Reported alone	-4.94	44.11	43.54	0.57	(0.46 to 0.69)
Endpoints excluded	Reported alone	-4.73	44.13	43.65	0.48	(0.36 to 0.60)
Observed 15+	age Non-household adult	0.67	20.58	19.73	0.85	(0.79 to 0.91)
Endpoints excluded	Non-household adult	0.80	21.36	20.64	0.73	(0.66 to 0.80)

Projection contrasts are calculated by applying both education-group leisure timing distributions to the same clock-specific context curve. Ranges are estimate plus or minus 1.96 replicate-weight standard errors. The endpoint-excluded reference contains respondents younger than 30 years and the middle adult education groups. The non-household-adult rows are boundary results and do not identify strangers or cross-education contact.

Supplementary Table 14 WBM interaction rates by reported companion status

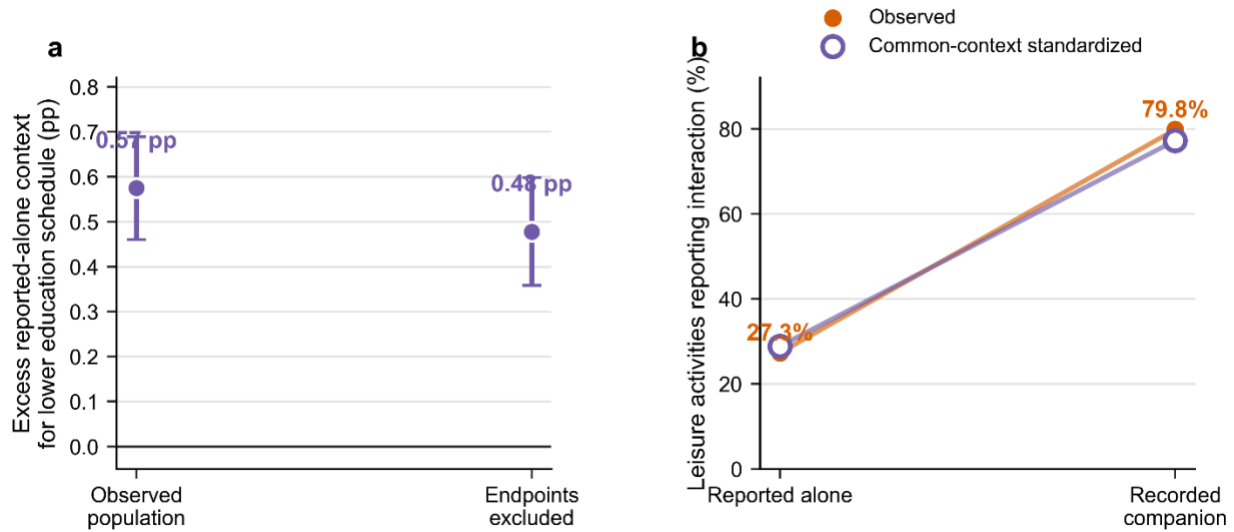
Analysis	Reported alone, % interactive	Recorded companion, % interactive	Companion minus alone, pp
2010/2012/2013, age 30+	27.31	79.83	52.51 (50.00 to 55.02)
2010, age 30+	25.84	78.66	52.82 (49.40 to 56.24)
2012, age 30+	28.59	78.29	49.69 (45.30 to 54.09)
2013, age 30+	26.78	81.56	54.77 (51.44 to 58.11)
2010/2012/2013, all ages	30.15	81.41	51.26 (49.12 to 53.40)
2021 boundary, age 30+	32.14	80.48	48.34 (44.33 to 52.36)

The primary pooled analysis uses the adjusted WBM activity weight. Ranges are estimate plus or minus 1.96 activity-replicate-weight standard errors. The 2021 estimate is a separate boundary analysis using the original activity weight.

Supplementary Table 15 Primary WBM support and context standardization

Companion status	Activity rows	Valid interaction rows	Valid interaction share, %
Reported alone	9719	9641	99.20
Recorded companion	10058	10032	99.74

The common-context standardized interaction rates were 28.87% for reported-alone leisure and 77.24% for leisure with a recorded companion. Their difference was 48.37 percentage points (45.67 to 51.07).



Supplementary Fig. 5 Companion-context and interaction calibration. a, Excess projected reported-alone leisure context for the less-than-high-school timing distribution relative to the bachelor-plus distribution under the observed age-15-plus and endpoint-excluded clock references. Ranges are estimate plus or minus 1.96 respondent replicate-weight standard errors. b, Weighted WBM interaction rates for leisure reported alone and leisure with a recorded companion in pooled 2010, 2012 and 2013 data among adults aged 30 years and older. Filled points show observed rates and open points show rates after common-context standardization. The WBM interaction item includes telephone interaction. The panels are separate aggregate calibrations and do not identify encounters, partner identity or cross-education contact.

Supplementary Note 8. Distributional separation after structural standardization

Main Fig. 3 reports how much of the mean placement gap remained after standardizing observed work, household and job structure. A smaller positive mean gap could still represent one common distribution shifted later. We therefore applied the same standardizations to the full 96-bin clock-time distributions.

Mean placement remained later in the bachelor-plus group under all six domain-by-standardization comparisons. For leisure, residual-shape distances ranged from 18.7 to 22.6 minutes and exceeded the 15-minute source-bin resolution under all three routes. For care or household activity, residual-shape distances ranged from 12.0 to 17.0 minutes. The work-clock result was below one source bin, whereas the broad-structure and selected-job results exceeded it. Observed structure therefore reduced both placement gaps. The remaining leisure difference consistently retained a change in distributional shape, whereas the care or household evidence differed across standardization routes.

Supplementary Table 16 Clock-time distribution differences after standardization to common observed demographic, household and work structures. Positive mean gaps indicate later placement in the bachelor-plus group relative to the less-than-high-school group. Wasserstein-2 measures the overall distance between the standardized clock-time distributions, and residual shape measures the distributional difference remaining after removing the mean gap. Point estimates retain the 2020 final weight; ranges are estimates plus or minus 1.96 replicate-weight standard errors from the non-2020 sample. Translation share is a mathematical component of squared Wasserstein-2 distance, not an explained or mediated proportion. The standardization routes are parallel descriptive comparisons.

Domain	Standardized characteristics	Categories	Mean gap, min	W2, min	Residual shape, min	Translation share
Leisure	Age, employment-day and household-child structure	32	23.0 (20.0–26.8)	32.2	22.6 (20.7–23.9)	50.9%
Leisure	Age, employment-day and last-work-end category	36	18.2 (15.6–22.1)	26.1	18.7 (17.2–19.9)	48.8%
Leisure	Occupation	7	21.3 (14.2–29.3)	29.5	20.3 (16.9–22.8)	52.5%
Care/household	Age, employment-day and household-child structure	32	34.1 (30.3–37.7)	38.1	17.0 (15.2–19.4)	80.0%
Care/household	Age, employment-day and last-work-end category	36	29.9 (25.9–34.1)	32.2	12.0 (10.5–14.1)	86.1%
Care/household	Work status and usual weekly hours	11	34.5 (31.0–38.6)	38.4	16.7 (14.9–19.3)	81.0%