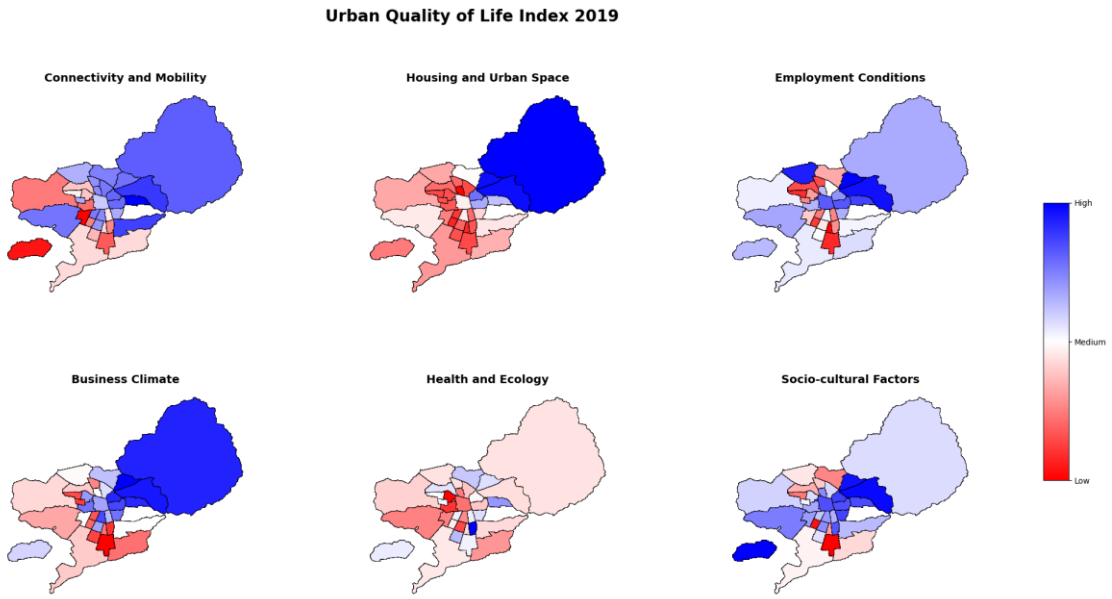
Supplementary Information of Beyond Mobility: Socioeconomic Context Shapes the Dynamics and Hidden States of COVID-19 Transmission.

Mauricio Herrera-Marín, mherrera@udd.cl[1,*,+], Constanza Neira-Urrutia, c.neira@udd.cl[2,+], and Fernando Lagos-Alvarado, f.lagosa@udd.cl[1,+]

[1]Faculty of Engineering. Universidad del Desarrollo (UDD). Santiago de Chile, Avda. Plaza 700, Las Condes, Chile.
[2]Faculty of Health Sciences. Universidad del Desarrollo (UDD). Ainavillo 456, Concepci ´on, Chile.
[*]Corresponding author
[+]These authors contributed equally to this work.

Supplementary Figures

Supplementary Figure N°1. Urban Quality of Life Index (ICVU) 2019 across comunes of the Metropolitan Region, disaggregate by dimension.



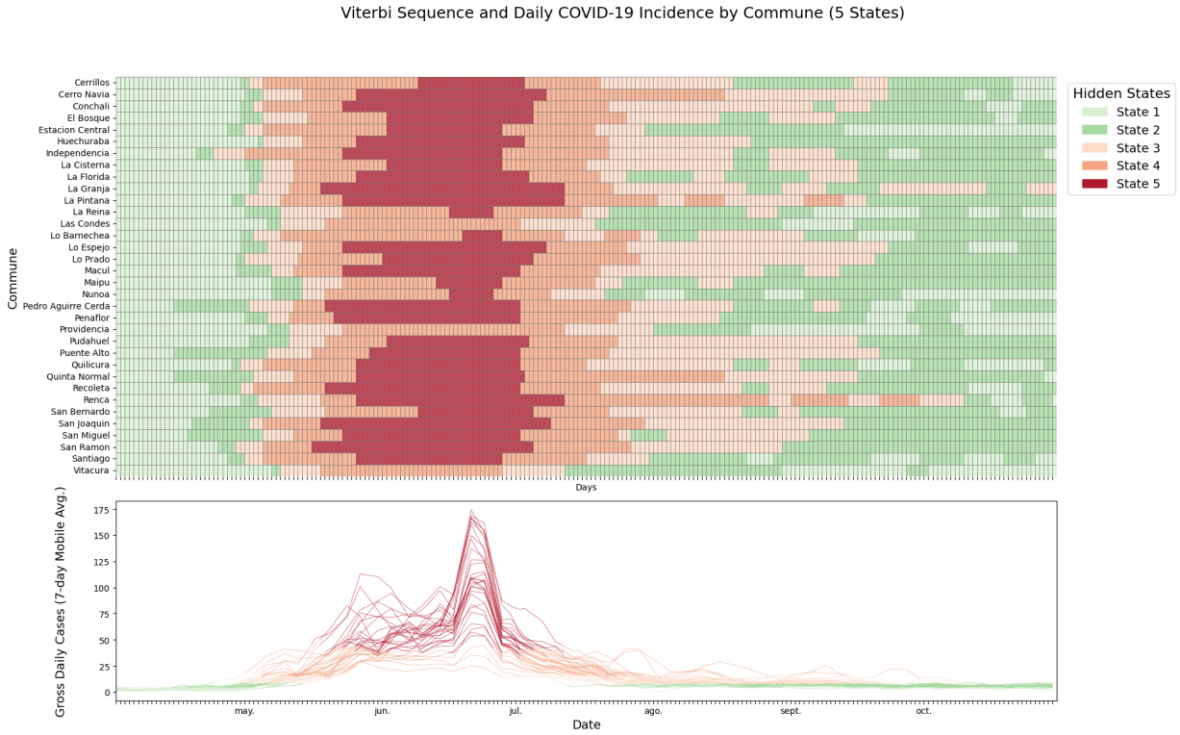
The Supplementary Figure N°1 illustrates the spatial distribution of the six components that comprise the 2019 Urban Quality of Life Index (ICVU) across the communes of the Santiago Metropolitan Area. The ICVU is a composite index based on 36 statistical variables selected through expert criteria and is structured around six dimensions: Employment Conditions, which refer to residents' income levels, job stability, contract types, cost of living, and indebtedness; Business Climate, which captures a commune's ability to attract economic activity, new ventures, real estate development, and both public and private services; Socio-cultural Factors, encompassing educational outcomes, civic participation, and indicators of social cohesion; Connectivity and Mobility, measuring access and proximity to public transportation, internet connectivity, and exposure to traffic accidents; Health and Environment, evaluating access to healthcare networks, prevalence of health deficiencies, and environmental exposure; and Housing and Urban Space, assessing housing conditions, public space maintenance, and neighborhood safety.

The color gradient ranges from red (low values) to blue (high values), indicating relative performance in each dimension. A clear east-west gradient is observed: eastern comunes consistently score higher in Connectivity and Mobility, Housing and Urban Space, Employment Conditions, and Business Climate, while central and southern comunes exhibit lower values across most domains. These spatial inequalities provide crucial context for understanding the differenciated epidemiological trajectories analyzed in the main text.

Supplementary Figure N°2. Viterbi-decoded sequence by commune using a four-state Non-Homogeneous Hidden Markov Model (nHMM).

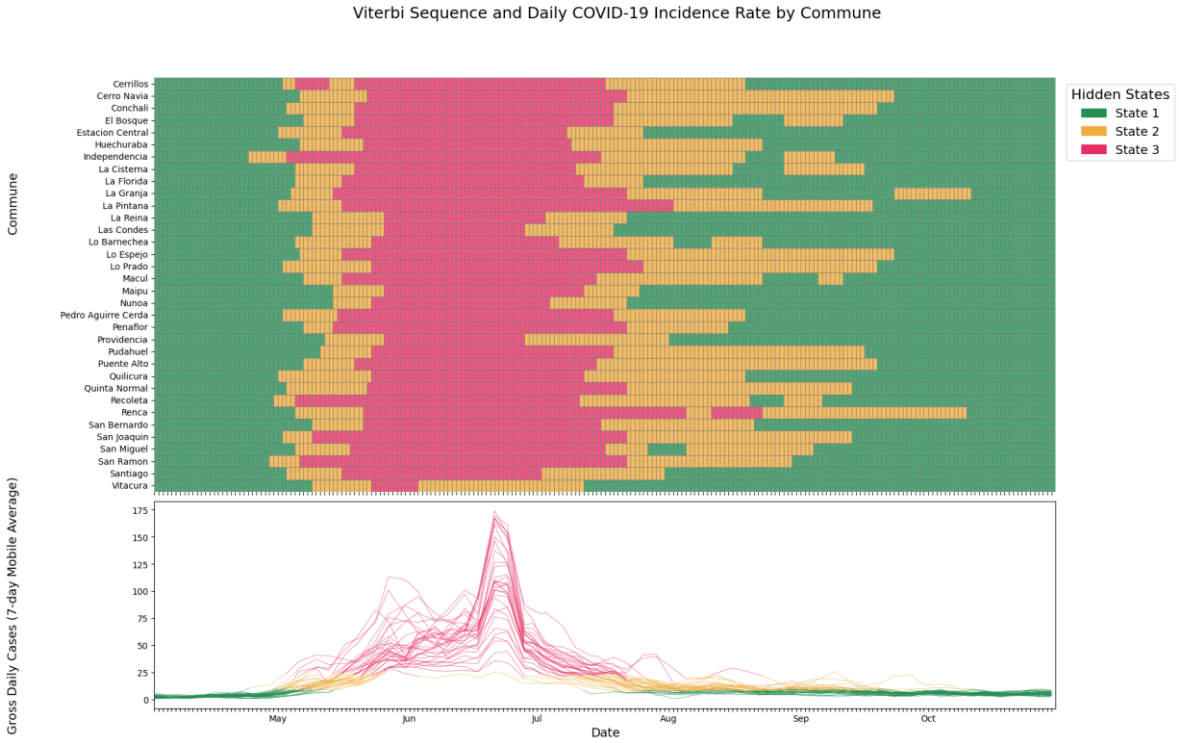
The supplementary figure N°2 displays the Viterbi sequence for each commune based on a Non Homogeneous Hidden Markov Model (nHMM) with four latent states. The model converged with a log-likelihood of -20448.6 (df=443), and yielded AIC=41783.21 and BIC=44836.52. Compared to the three-state model, the four-state model improves the differentiation of transitional epidemic phases, especially during the mid-period of the timeline. Visually, it introduces a higher leve lof detail while preserving the interpretability and temporal coherence across comunes. This model offers a favorable trade-off between statistical performance and visual clarity, avoiding the excessive fragmentation observed in more complex specifications.

Supplementary Figure N°3. Viterbi-decoded sequence by commune using a five-state Non-Homogeneous Hidden Markov Model (nHMM).



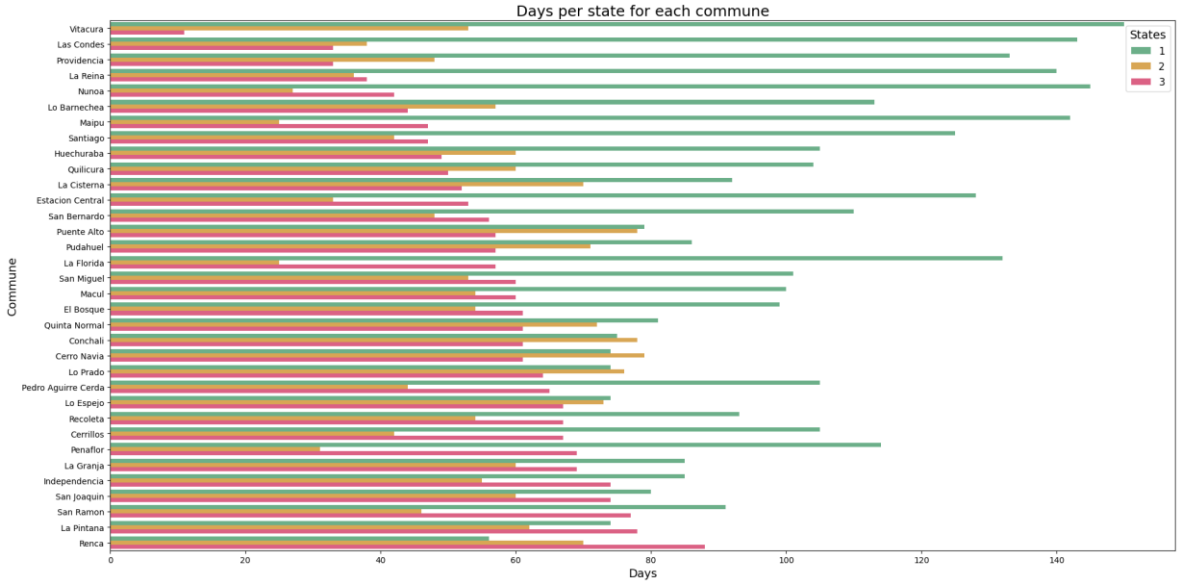
The supplementary figure N°3 shows the Viterbi-decoded hidden states under a five-state Non Homogeneous Hidden Markov Model. The model converged with a log-likelihood of -18959.39 (df=734), achieving the lowest AIC (39386.77) and BIC (44445.75) among all candidate models. Despite this statistical improvement, visual inspection reveals signs of overfitting. Several comunes exhibit erratic transitions and abrupt state changes lacking epidemiological plausibility. These patterns suggest the model may be capturing noise rather tan meaningful shifts in epidemic dynamics. Consequently, although the five-state model scores better in AIC/BIC, its interpretability is compromised, limiting its utility for comparative and policy-oriented analyses.

Supplementary Figure N°4. Viterbi-decoded sequence by commune using a three-state Non-Homogeneous Hidden Markov Model (nHMM).



Supplementary Figure N°4. Spatiotemporal evolution of COVID-19 across 34 central communes in Santiago during 2020. Daily incidence rates (per 100,000 population), with each curve representing one commune, colored by the inferred latent epidemic state. Viterbi-decoded sequences of epidemic states obtained from a nonhomogeneous hidden Markov model (nHMM), showing transitions among three levels: State 1 (green): Moderate transmission, State 2 (yellow): Severe transmission, State 3 (red): Critical transmission. Transitions were modeled using inter-communal mobility as a covariate. The figure reveals asynchronous epidemic trajectories across communes.

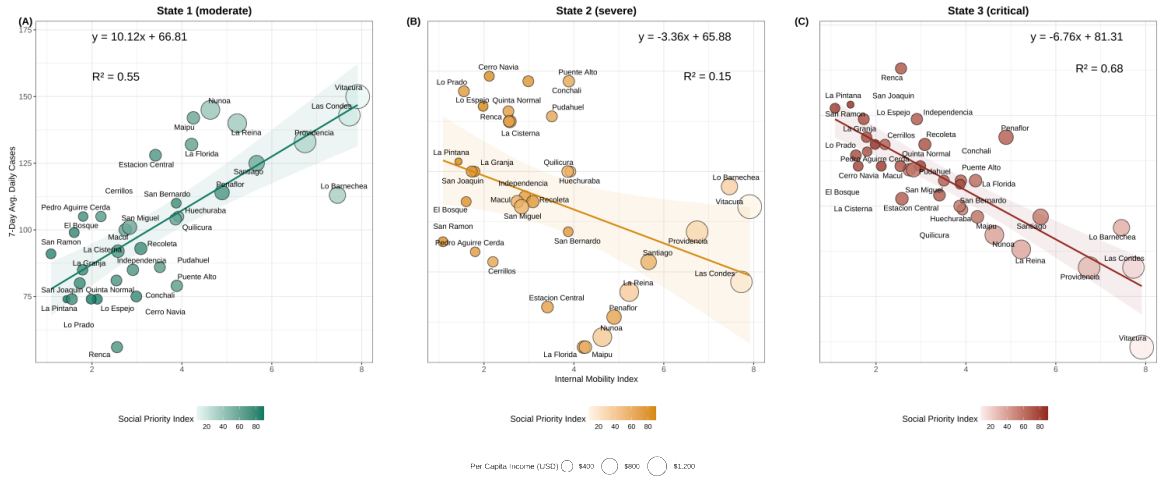
Supplementary Figure N°5: Total number of days each commune spent in the three epidemic states identified by the non-homogeneous Hidden Markov Model (nHMM).



The supplementary figure N°5 shows the cumulative duration (in days) that each commune in the Santiago Metropolitan Area remained in each of the three latent epidemic states inferred by the nHMM. Communes are ordered by the total number of days spent in the most critical state (State 3, red). State 1 (green) corresponds to periods of moderate transmisión, State 2 (yellow) to severe transmission, and State 3 (red) to phases of critical transmission intensity.

A clear territorial pattern emerges: comunes with lower socioeconomic status – such as La Pintana, Renca, San Ramon, and San Joaquin – accumulated a significantly higher number of days in the most severe epidemic pase, indicating prolonged exposure to elevated transmission risk. Conversely, high-income comunes such as Vitacura, Las Condes, and Providencia spent the majority of the period in the moderate state (State 1), with relatively brief durations in the most critical phase. These disparities highlight the unequal burden of the pandemic and underscore the role of structural socioeconomic conditions in shaping the epidemic trajectories at the local level.

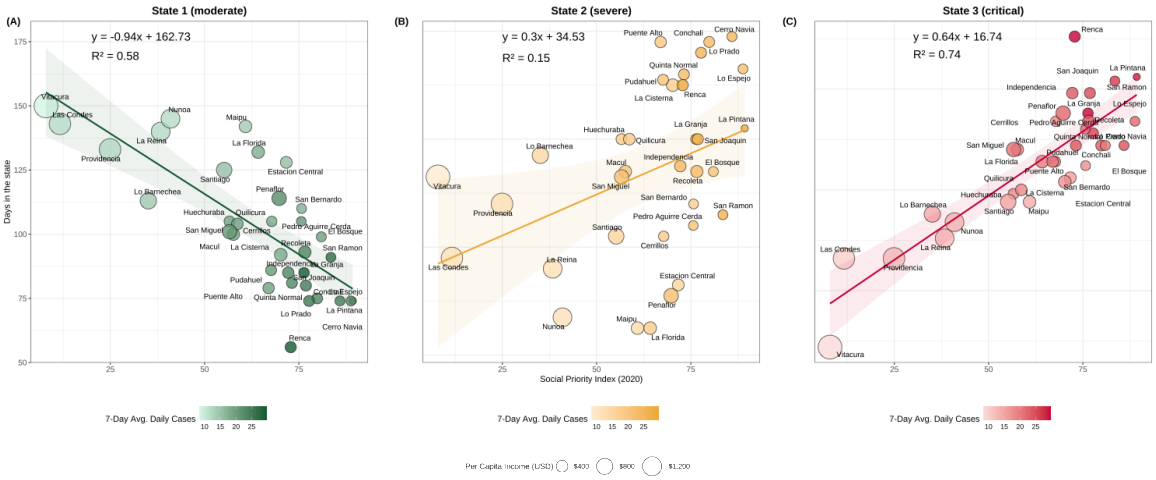
Supplementary Figure N°6. Association between Internal Mobility Index and average daily cases across comunes in each latent epidemic state.



The supplementary figure N°6 displays the linear relationship between the Internal Mobility Index and the average number of COVID-19 daily cases (per 100,000 population) observed during the periods classified under each of the three latent states inferred by the non-homogeneous Hidden Markov Model (nHMM): State 1 (moderate) in panel (A), State 2 (severe) in panel (B), and State 3 (critical) in panel (C). Each dot represents a commune, with its size and color intensity corresponding to the Social Priority Index, a proxy for socioeconomic vulnerability (higher values indicate greater social deprivation).

Panel (A) shows a strong positive association (R^2 = 0.55) between mobility and case rates in State 1, indicating that mobility was a key driver of transmission in periods of moderate intensity. Panel (B) displays a weak and statistically les relevant association in State 2 (R^2 = 0.15). Panel (C), by contrast, reveals a strong negative correlation (R^2 = 0.68), where comunes with higher internal mobility tend to report fewer cases during critical phases. This counterintuitive pattern likely reflects the influence of preexisting structural conditions, as communes with lower mobility – but higher social vulnerability – experience higher case burdens during the most intense epidemic phases. Together, these results suggest that the relationship between mobility and contagion is non-linear and state-dependent, being mediated by socioeconomic inequalities and structural exposure risks.

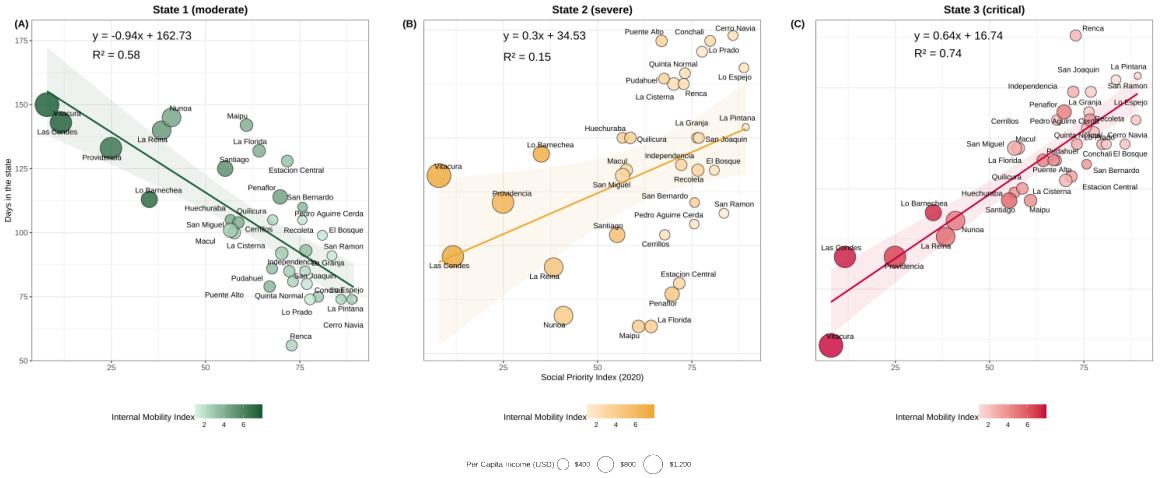
Supplementary Figure N°7. Association between Social Priority Index and duration in each epidemic state, stratified by average daily COVID-19 cases.



Supplementary Figure N°8. Association between Social Priority Index and duration in each epidemic state, stratified by External Mobility Index.

Supplementary Figure N°9. Association between Social Priority Index and duration in each epidemic state, stratified by Housing Precariousness Index.

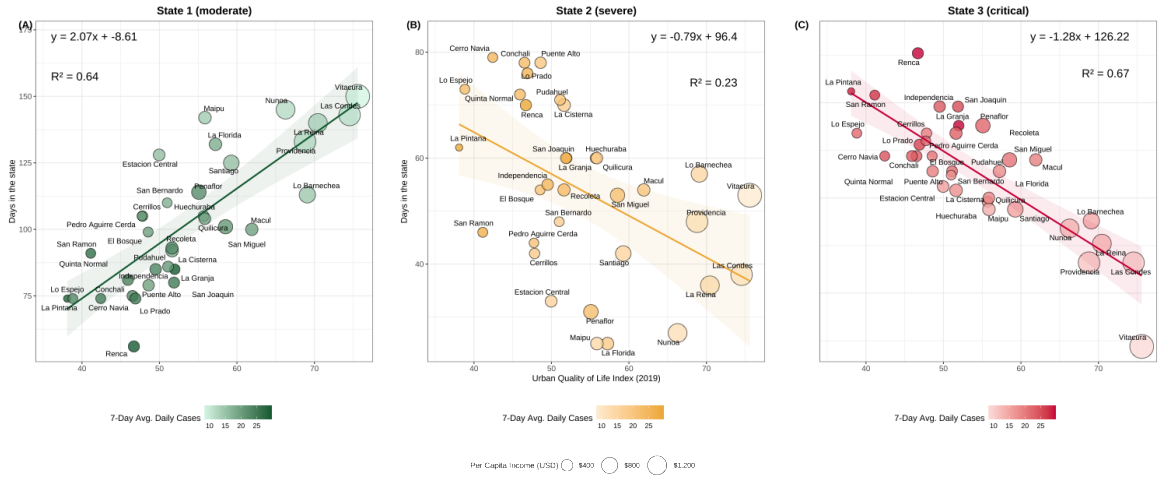
Supplementary Figure N°10. Association between Social Priority Index and duration in each epidemic state, stratified by Internal Mobility Index.



The supplementary figures N°7, 8, 9, and 10 explore the relationship between the Social Priority Index – a composite indicator of socioeconomic vulnerability – and the number of days each commune spent in the three latent epidemic states identified by the non-homogeneous Hidden Markov Model: State 1 (moderate), State 2 (severe), and State 3 (critical). Each figure incorporates a different contextual variable: the 7-day average of daily COVID-19 cases (supplementary figure N°7), the External Mobility Index (supplementary figure N°8), the Housing Precariousness Index (supplementary figure N°9), and the Internal Mobility Index (supplementary figure N°10). In all panels, the x-axis corresponds to the Social Priority Index and the y-axis to the cumulative number of days in the respective epidemic state. Bubble size reflects per capita income (USD), highlighting the economic disparities between comunes.

Despite the varying contextual variables, a consistent pattern is observed across all four figures. A strong positive correlation emerges in State 3 (R^2 = 0.74), where comunes with higher Social Priority Index values – representing greater socioeconomic deprivation – tend to spend significantly more time in the most critical transmission phase. Conversely, a strong negative association is found in State 1 (R^2 = 0.58), where comunes with lower Social Priority Index scores remain longer in moderate transmission conditions. State 2 presents only a weak association (R^2 = 0.15), consistent with its role as a transitional phase between moderate and critical states. These findings reveal that crucial vulnerability is a robust and state-dependent determinant of epidemic burden regardless of the specific structural or mobility-related variable considered.

Supplementary Figure N°11. Association between 2019 Urban Quality of Life Index (ICVU) and cumulative days spent in each latent epidemic state.



The supplementary figure N°11 assesses the association between the 2019 Urban Quality of Life Index (ICVU) and the cumulative number of days each commune spent in the three latent epidemic states. The x-axis corresponds to the ICVU score for each commune, and the y-axis to the total number of days classified under each state. Buble size represents per capita income in USD, while color intensity reflects the 7-days average of daily COVID-19 cases, serving as a proxy for local epidemic intensity.

A strong positive relationship is observed in State 1 (R^2 = 0.64) indicating that comunes with higher quality of life indicators experienced more days in moderate transmission pahses. In contrast, State 3 shows a strong negative association (R^2 = 0.67), suggesting that structurally disadvantaged communes endured longer periods under critical transmission conditions. The association in State 2 is relatively weak (R^2 = 0.23), consistend with its function as a transitional category. These results underscore the influence of structural urban inequalities on the temporal distribution of epidemic risk, demonstrating that higher quality of life is associated with reduced duration in critical states and prolonged exposure to moderate conditions.

Supplementary Tables

Supplementary Table N°1. Variance decomposition and model fit for each transition and model (A, B, C and D).

Transition	Model	Inter-Variance	Intra-Variance	Inter-Proportion (%)	Marginal R^2	Conditional R^2
$1 \rightarrow 2$	A	0.1153	0.0101	91.95	0.2516	0.9423
	B	0.2190	0.0101	95.60	0.2871	0.9345
	C	0.1420	0.0101	93.37	**0.3796**	0.9011
	D	0.1512	0.0101	93.74	0.3361	0.9046
$2 \rightarrow 2$	A	0.1421	0.0149	90.50	0.1223	0.9339
	B	0.2593	0.0149	94.43	0.1715	0.9268
	C	0.2089	0.0149	93.34	0.2331	0.8500
	D	0.2269	0.0149	93.84	0.2348	0.8717
$2 \rightarrow 3$	A	0.2807	0.0327	89.57	0.2593	0.9315
	B	**1.0706**	0.0326	**97.04**	**0.4543**	**0.9363**
	C	0.3340	0.0326	91.10	0.4484	0.8962
	D	0.7275	0.0326	95.71	0.4746	0.9233
$3 \rightarrow 3$	A	0.1086	0.0197	84.62	0.4579	0.9404
	B	**0.9885**	0.0197	**98.05**	0.5385	0.9242
	C	0.1882	0.0197	90.52	**0.5894**	0.9282
	D	0.4173	0.0197	95.50	0.4903	0.9445
$3 \rightarrow 2$	A	0.0070	0.0012	85.22	0.2374	0.9080
	B	0.0337	0.0012	**96.50**	0.1799	0.9056
	C	0.0197	0.0012	94.19	**0.3356**	0.8586
	D	0.0211	0.0012	94.55	0.2862	0.8723
$2 \rightarrow 1$	A	0.0034	0.0002	**95.93**	0.3270	0.9876
	B	0.0766	0.0002	**99.81**	0.3269	**0.9955**
	C	0.0132	0.0002	98.89	**0.3637**	0.9805
	D	0.0153	0.0002	99.06	0.2141	0.9756

Supplementary table N°1 presents the results of the linear mixed models (LMMs) applied to the transition probabilites between epidemic latent states, estimated using an nHMM model. Each transition was modeled independently under four specifications: model A included exclusively mobility indicators (internal and external mobility indexes of each comunne); model B incorporated sociodemographic covariates derived from the 2017 Census; model C considered the six dimensions of the 2019 Urban Quality of Life Index (ICVU); and model D combined variables from both structural domains. The selection of covariates for models B, C and D was performed by Elastic Net regularization (supplementary table N°2) which allowed us to identify an optimal subset of predictors by simultaneously penalizing the magnitude and redundancy of the coefficients.

This procedure effectively reduced dimensionality without sacrificing explanatory capacity, with validation RMSE errors close to 22.7 for the structural models.

In the transition from moderate to severe states (1→2), all models captured a high proportion of intercommunity variance, above 91%, which confirms the existence of relevant structural differences between comunes. Model C, base don ICVU dimensions, achieved the highest marginal R^2 (0.37958), while model D achieved a comparable performance (0.33605), integrating information from both census conditions and the urban environment.

As for the severe state persistence transitions (2→2), they also presented high levels of intercommunal variance. Although the marginal fit of the models was more modest compared to other transitions, the models that incorporated urban structures (C and D) show a better explanatory capacity than those that included only mobility or census data, suggesting that permanence in adverse epidemic situations depends on structural determinants that are difficult to modify in the short term.

In the case of the transition from severe to critical (2→3), the greatest structural dispersion between communes was observed. Model D, with a hybrid architecture, achieved both a high proportion of intercommune variance (0.727473) and the highest marginal R^2 (0.47456), surpassing the partial versions. The covariates selected by Elastic Net in this model reflect both conditions of social vulnerability and aspects of the urban environment related to employment, health and connectivity, therefore, it is clear that the progression towards critical epidemic scenarios is a complex interaction of the structural and social domain.

In the critical state persistence transition (3→3), model C again demonstrated a remarkable explanatory capacity, reaching the highest marginal R^2 (0.58939), reinforcing the hypothesis that factors related to the urban environment play a central role in the spread of severe epidemic conditions, especially in territories with structural deficits in housing, services and health.

Finally, in the regressive transitions (3→2, 2→1), the proportions of intercommunal variance were particularly high, exceeding 99% in some cases. In 2→1, where a substantial improvement in the epidemiological situation is observed, model B achieved the highest conditional R^2 (0.99552), reflecting the direct influence of sociodemographic characteristics such as per capita income, education, housing precariousness index, among others. Although model C obtained the best marginal performance (0.36370), model D maintained a balance between fit and explanatory capacity, confirming its overall robustness.

Supplementary Table N°2. Elastic Net Coefficients and RMSE for Models B, C, and D including Socio-demographic and Urban Quality variables.

Model	RMSE	Variable	Elastic Net Coef
B	22,92972	Internal Mobility Index	-6,44389
		External Mobility Index	-5,03525
		Per Capita Income (USD)	2,61930
		Percentage Overcrowding	0,09401
		Average Persons Households	0,18306
		Housing Precariousness Index	1,93892
		Percentage Higher Education Completed	-1,21381
		Percentage Households with Immigrants	0,42076
C	22,77792	Internal Mobility Index	-4,21061
		External Mobility Index	-6,35017
		Employment Conditions	-4,22224
		Business Climate	2,28145
		Socio-cultural Factors	1,64871
		Connectivity and Mobility	-0,58358
		Health and Ecology	2,12862
		Housing and Urban Space	-1,65299
D	22,69835	Internal Mobility Index	-4,77879
		External Mobility Index	-6,55418
		Per Capita Income (USD)	1,63088
		Percentage Overcrowding	0,00000
		Average Persons Households	0,00000
		Housing Precariousness Index	0,00000
		Percentage Higher Education Completed	-0,87650
		Percentage Households with Immigrants	1,65439
		Employment Conditions	-4,53999
		Business Climate	1,87196
		Socio-cultural Factors	1,55538
		Connectivity and Mobility	-0,92635
		Health and Ecology	2,29027
		Housing and Urban Space	-1,66662