

# Multimedia Appendix 1. Supplementary data.

## Local indicators

We found substantial correlations between most of the economic and touristic indicators used in the construction of the travelers' index (Figure S1). Among the economic variables, GDP per capita displayed a much lower correlation with the remaining statistics, that shared very high pair-to-pair correlations. The variables quantifying the importance of tourism in the regions were also highly correlated, with those corresponding to hotels showing a higher variability, both among each other and with respect to those corresponding to apartments. Variables across the two different categories showed much smaller correlations in general, with the exception of the high correlations found between the economic variables (other than the GDP per capita) and some of the indicators measuring hotel activity.

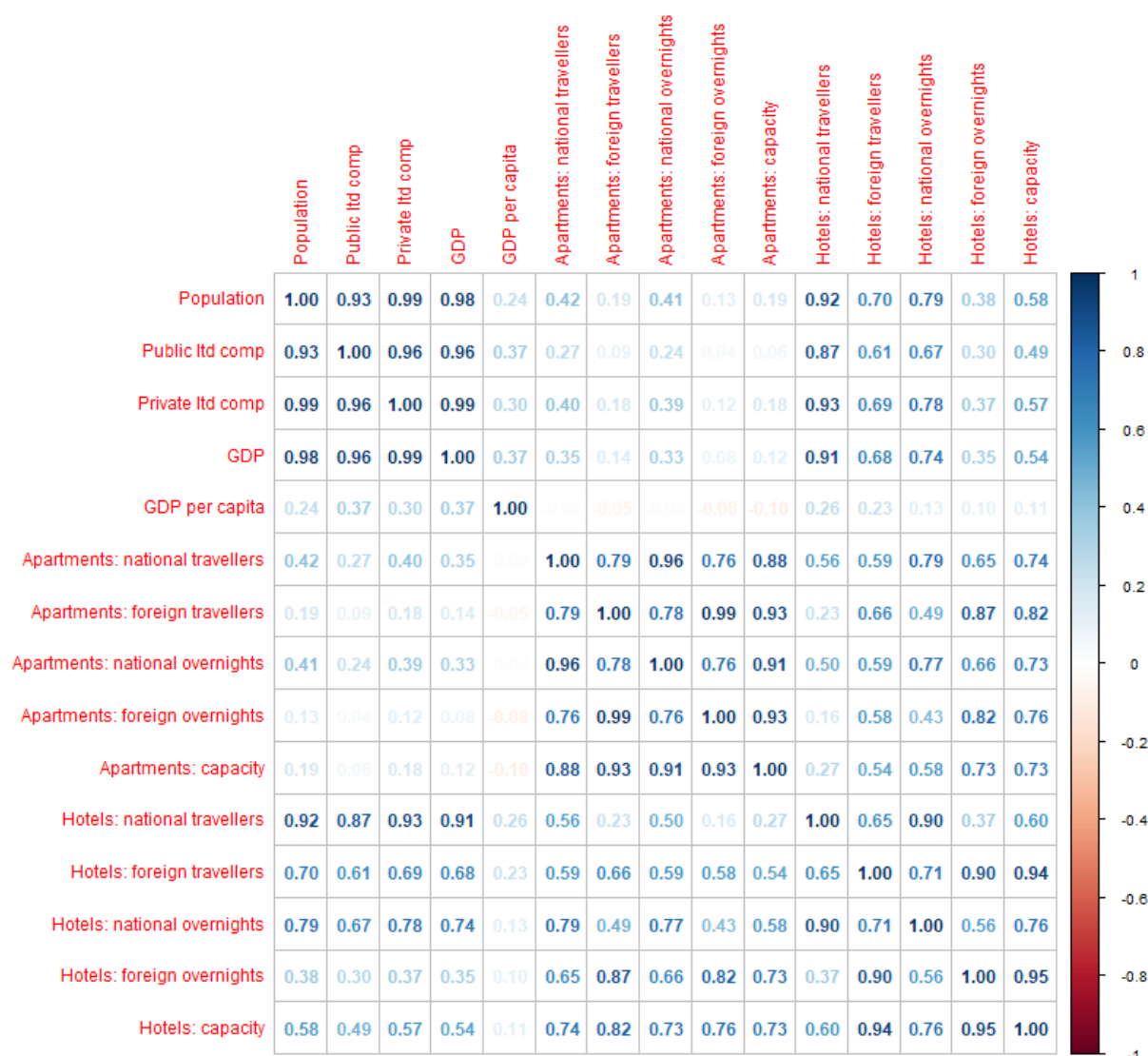


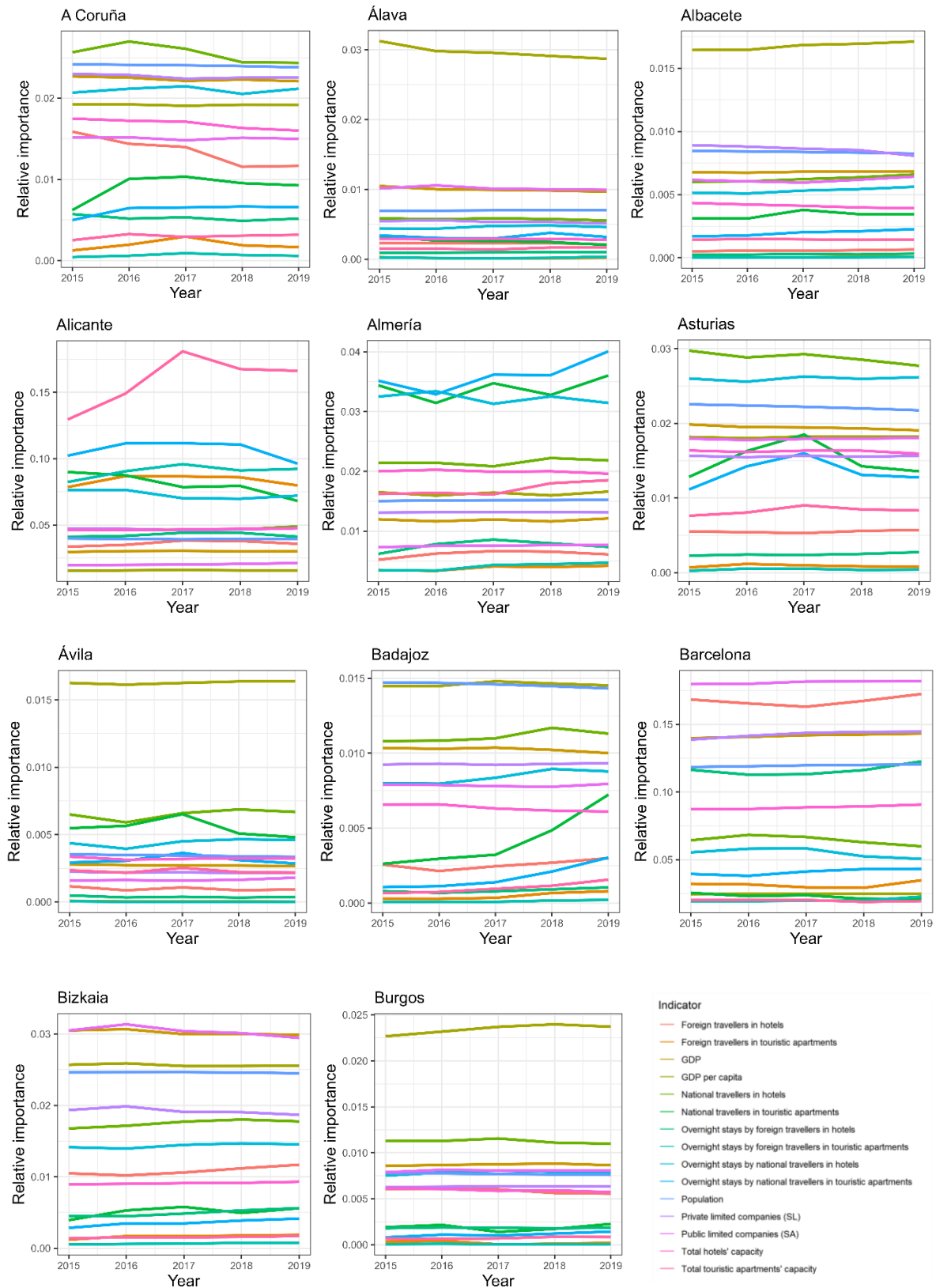
Figure S1. Pair-to-pair correlations between local touristic and economic indicators obtained from INE [40].

The numbers of foreign residents by birthplace showed a very high correlation with those by nationality, as expected (Figure S2). A significantly smaller, nevertheless large correlation was found for nationalized individuals born abroad.

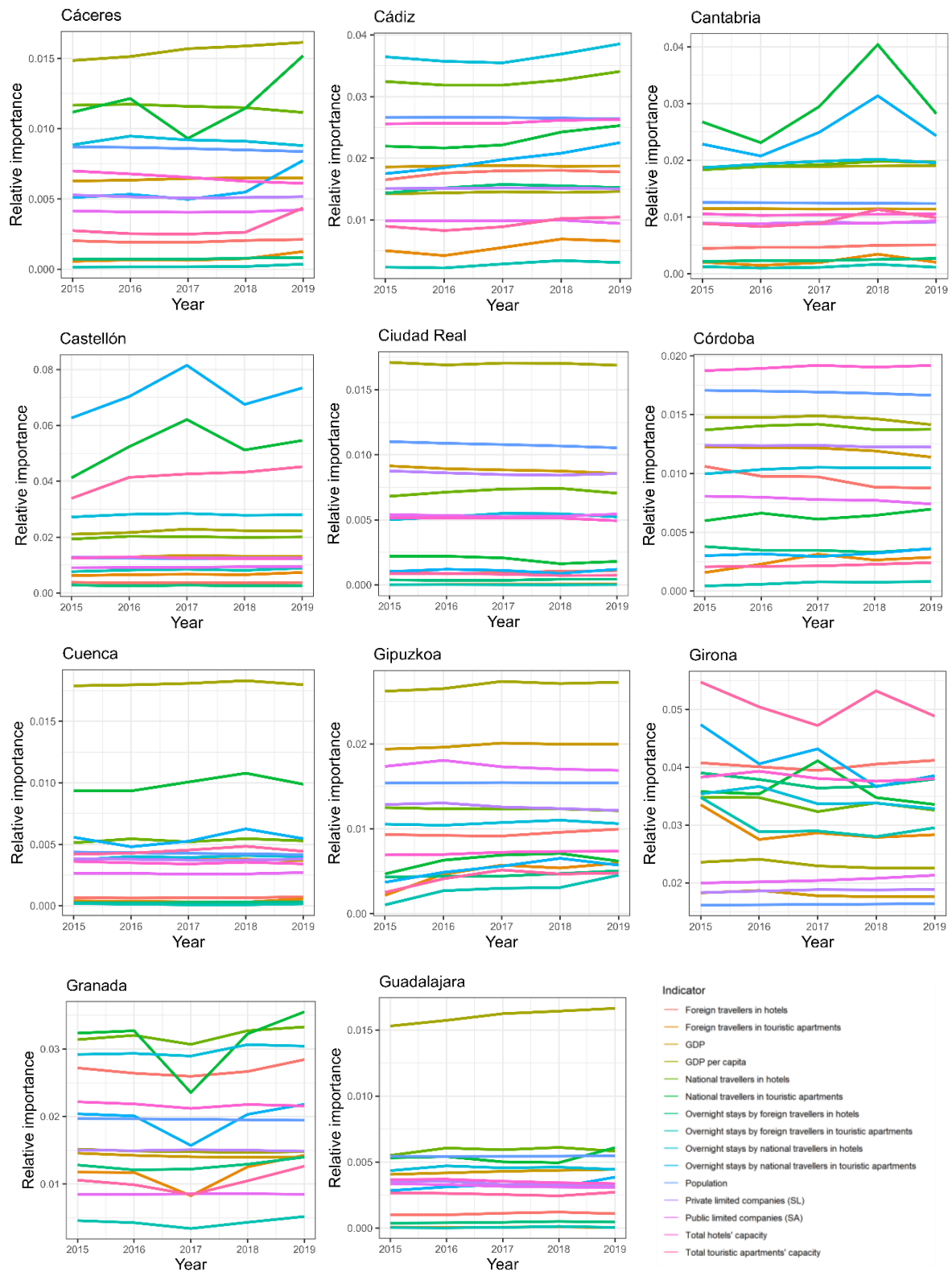


**Figure S2.** Pair-to-pair correlations between indicators concerning foreign and born-abroad population obtained from INE [40].

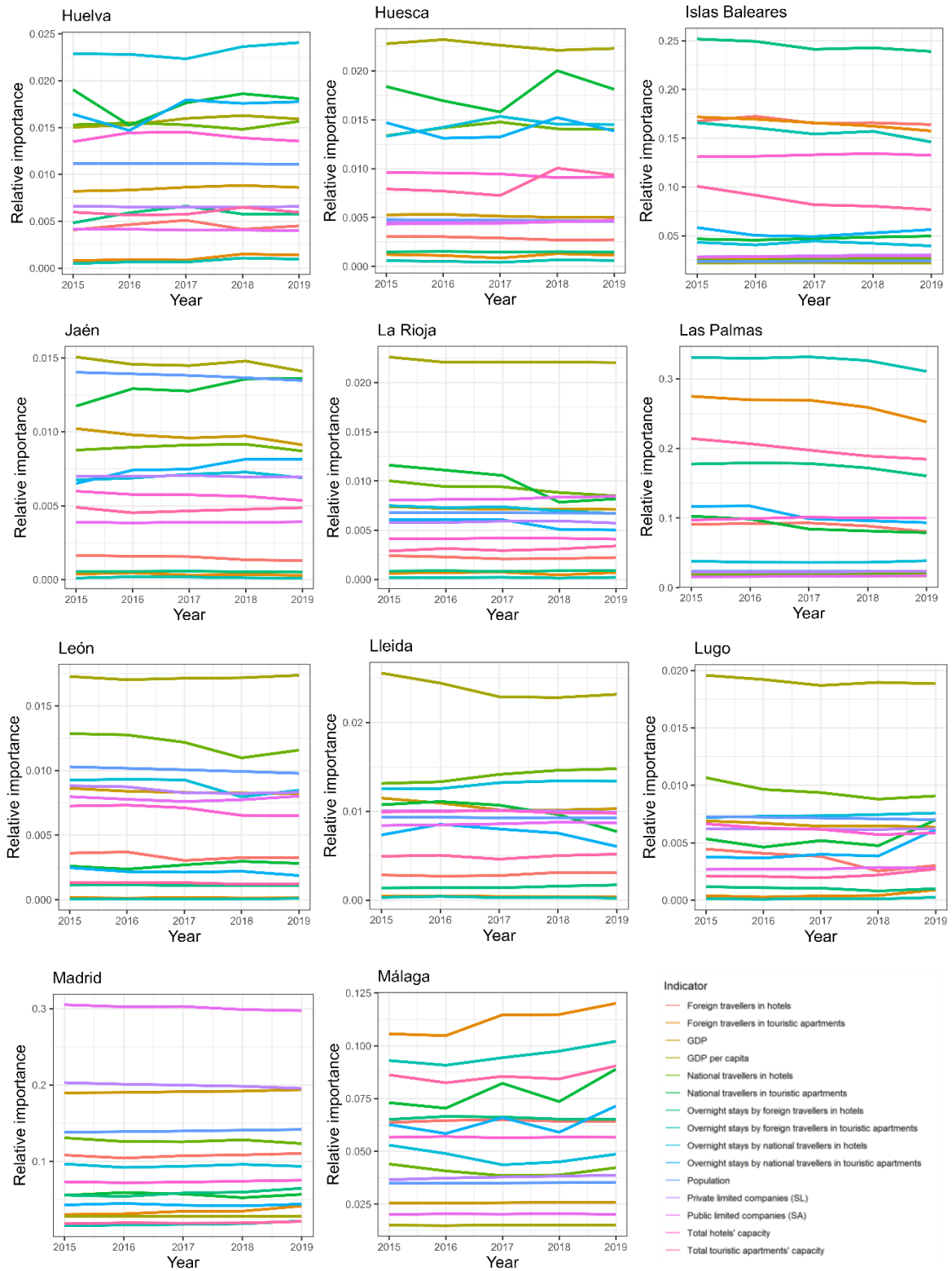
As stated in the main text, we found relatively small variations on the relative importances  $t_i, e_i, r_{ik}$  over the years for most provinces and choice of statistics (Figure S3). This suggests that the role of the provinces within the national context may be more stable than the global evolution displayed by the country as a whole, and that these indicators may provide reliable projections in the short-term.



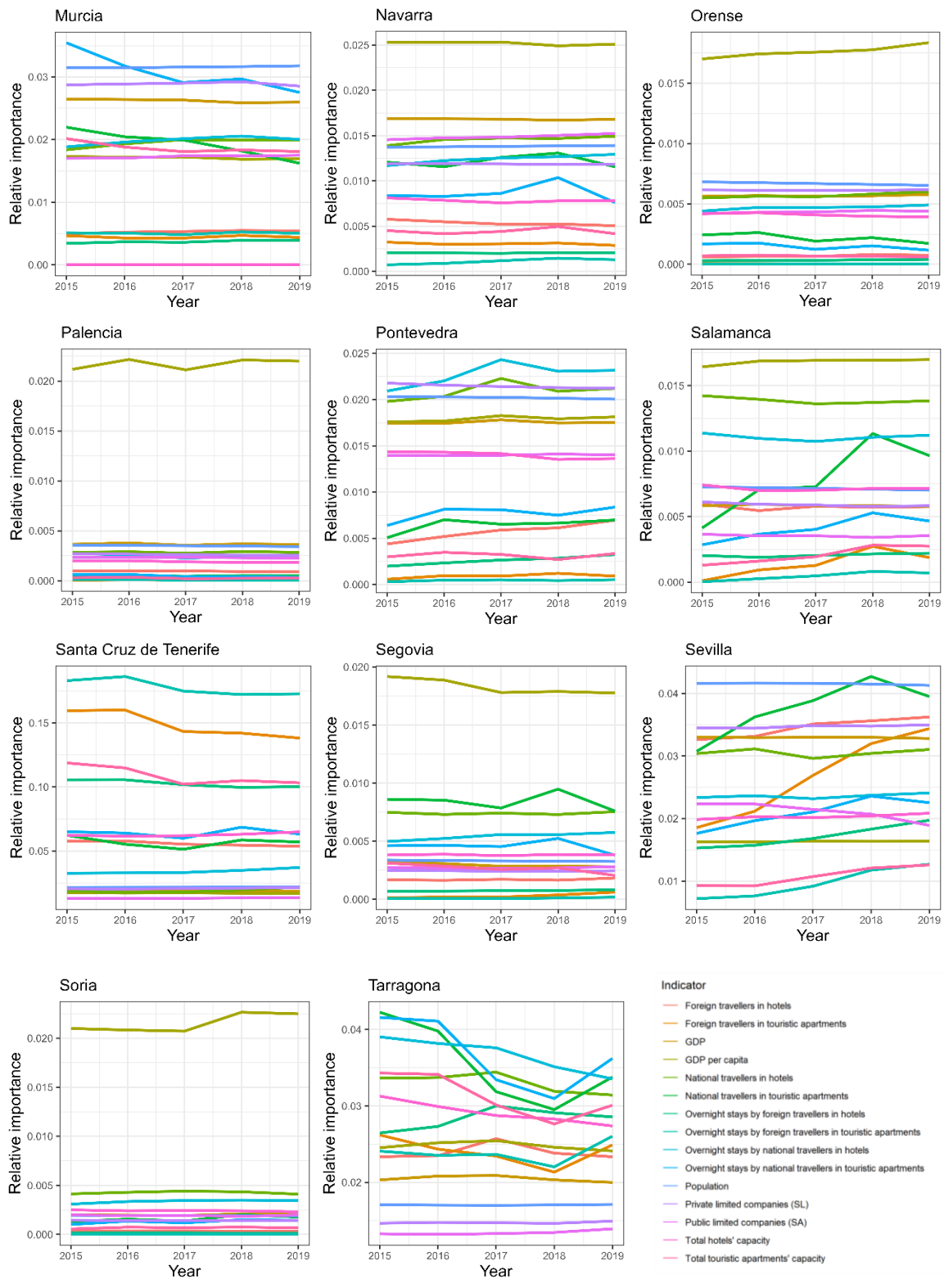
**Figure S3.1: Relative importances of provinces in Spain in terms of economic and touristic indicators.** Evolution of the indicators  $e_i$  and  $t_i$  during the years 2015-2019 for several provinces in Spain (see equation (2) in main text) for several choices of statistics.



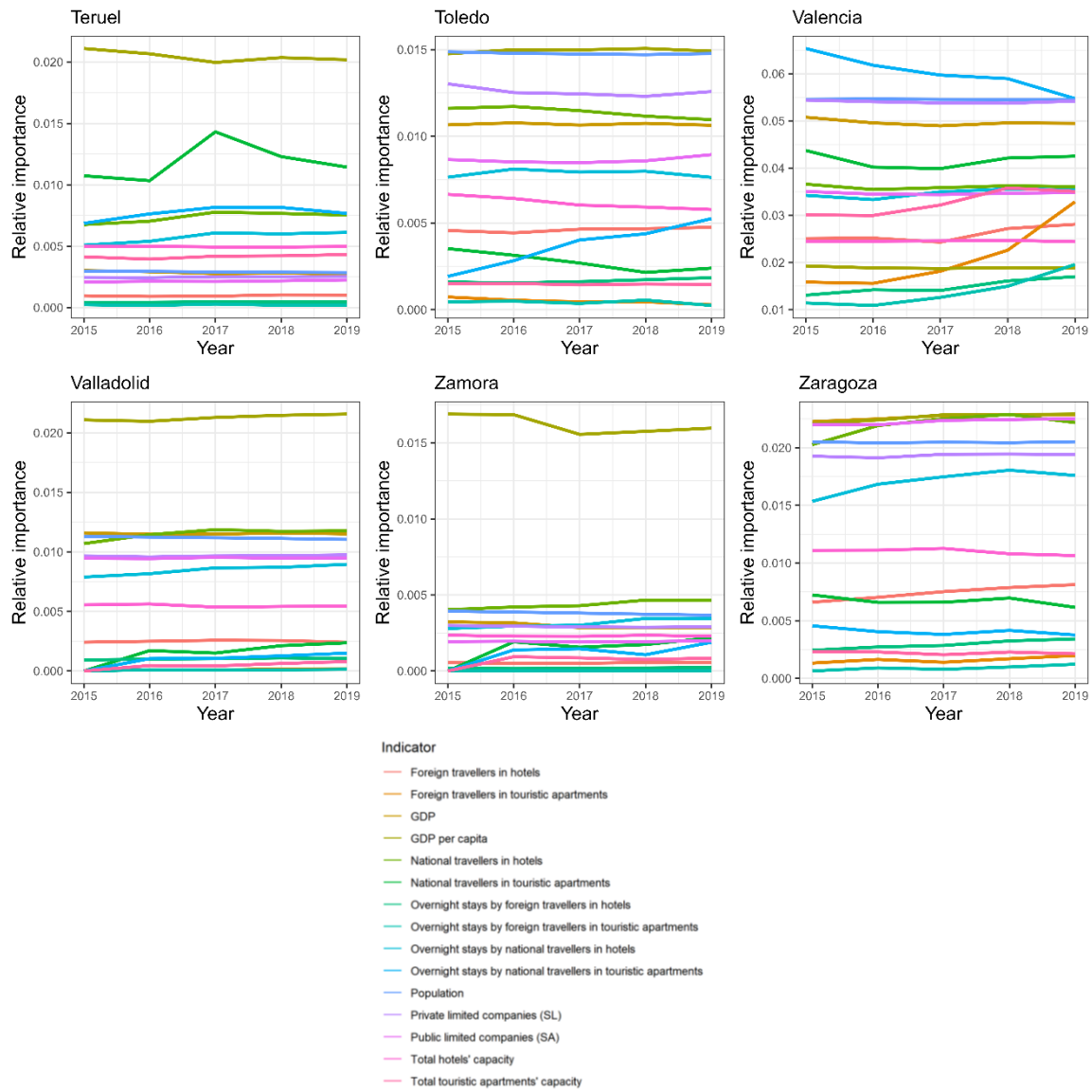
**Figure S3.2: Relative importances of provinces in Spain in terms of economic and touristic indicators.** Evolution of the indicators  $e_i$  and  $t_i$  during the years 2015-2019 for several provinces in Spain (see equation (2) in main text) for several choices of statistics.



**Figure S3.3: Relative importances of provinces in Spain in terms of economic and touristic indicators.** Evolution of the indicators  $e_i$  and  $t_i$  during the years 2015-2019 for several provinces in Spain (see equation (2) in main text) for several choices of statistics.



**Figure S3.4: Relative importances of provinces in Spain in terms of economic and touristic indicators.** Evolution of the indicators  $e_i$  and  $t_i$  during the years 2015-2019 for several provinces in Spain (see equation (2) in main text) for several choices of statistics.



**Figure S3.5: Relative importances of provinces in Spain in terms of economic and touristic indicators.** Evolution of the indicators  $e_i$  and  $t_i$  during the years 2015-2019 for several provinces in Spain (see equation (2) in main text) for several choices of statistics.

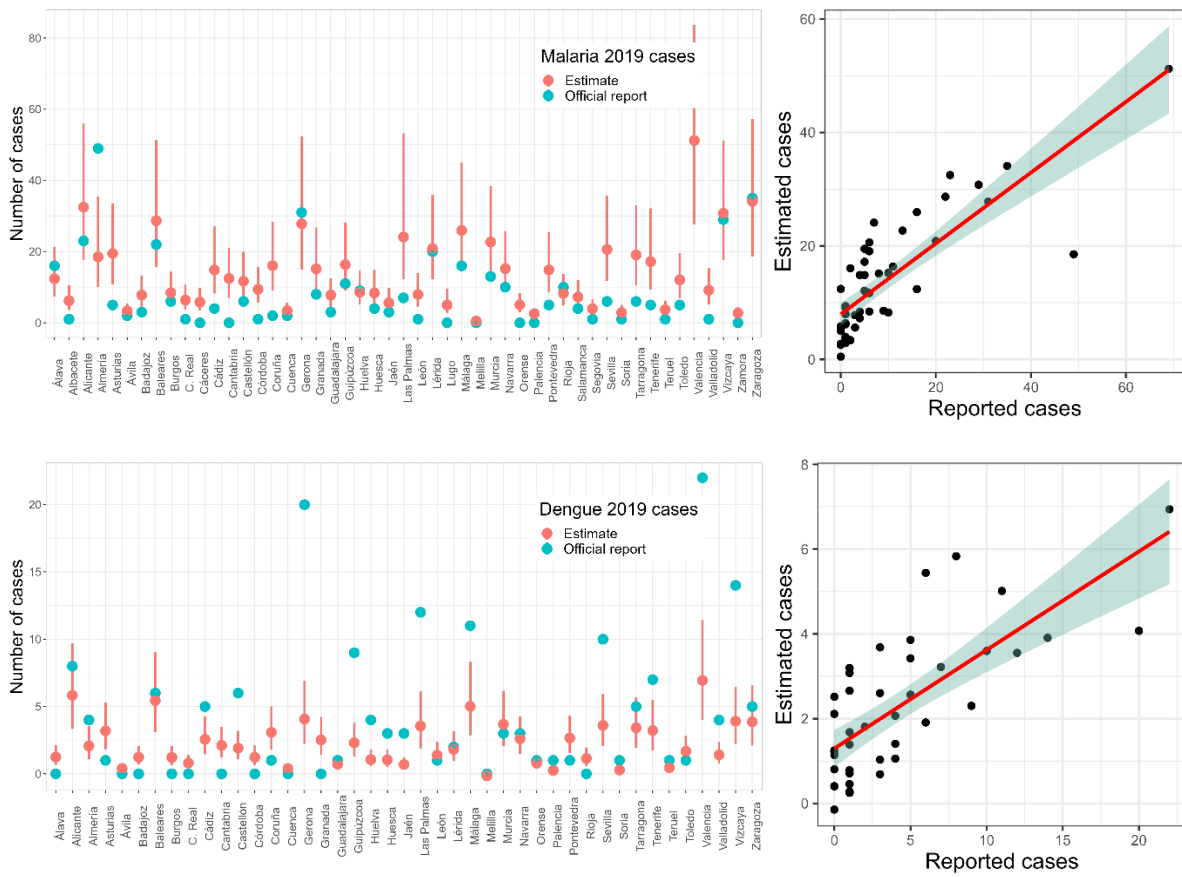
## Excluding outliers from the analysis

Table S1 shows the relevant features of the best models after excluding Madrid and Barcelona from the analysis. This was done in order to test the robustness of our approach, checking to what degree the accuracy of our models was due only to the extreme values of these provinces, both in terms of reported cases and travelers' index (see Figure S3).

Figure S4 shows the resulting models' fit and estimates for 2019 data, compared with the officially reported number of cases at each of the remaining provinces.

Disease (model)	Economic indicator (weight)	Touristic indicator (weight)	VFR indicator (weight)	Correlation with 2019 data	Overestimation
Malaria (simple)	Public limited companies (SA)	National travelers at hotels	Foreign residents by nationality	0.82	99.3%
Malaria (weighted)	No contribution (0)	No contribution (0)	Foreign residents by birthplace (1)	0.89	99.5%
Dengue (simple)	Public limited companies (SA)	National travelers at hotels	Foreign residents by birthplace	0.74	95.7%
Dengue (weighted)	Public limited companies (SA)	Overnight stays by national travelers at touristic apartments (0.1)	No contribution (0)	0.78	95.7%

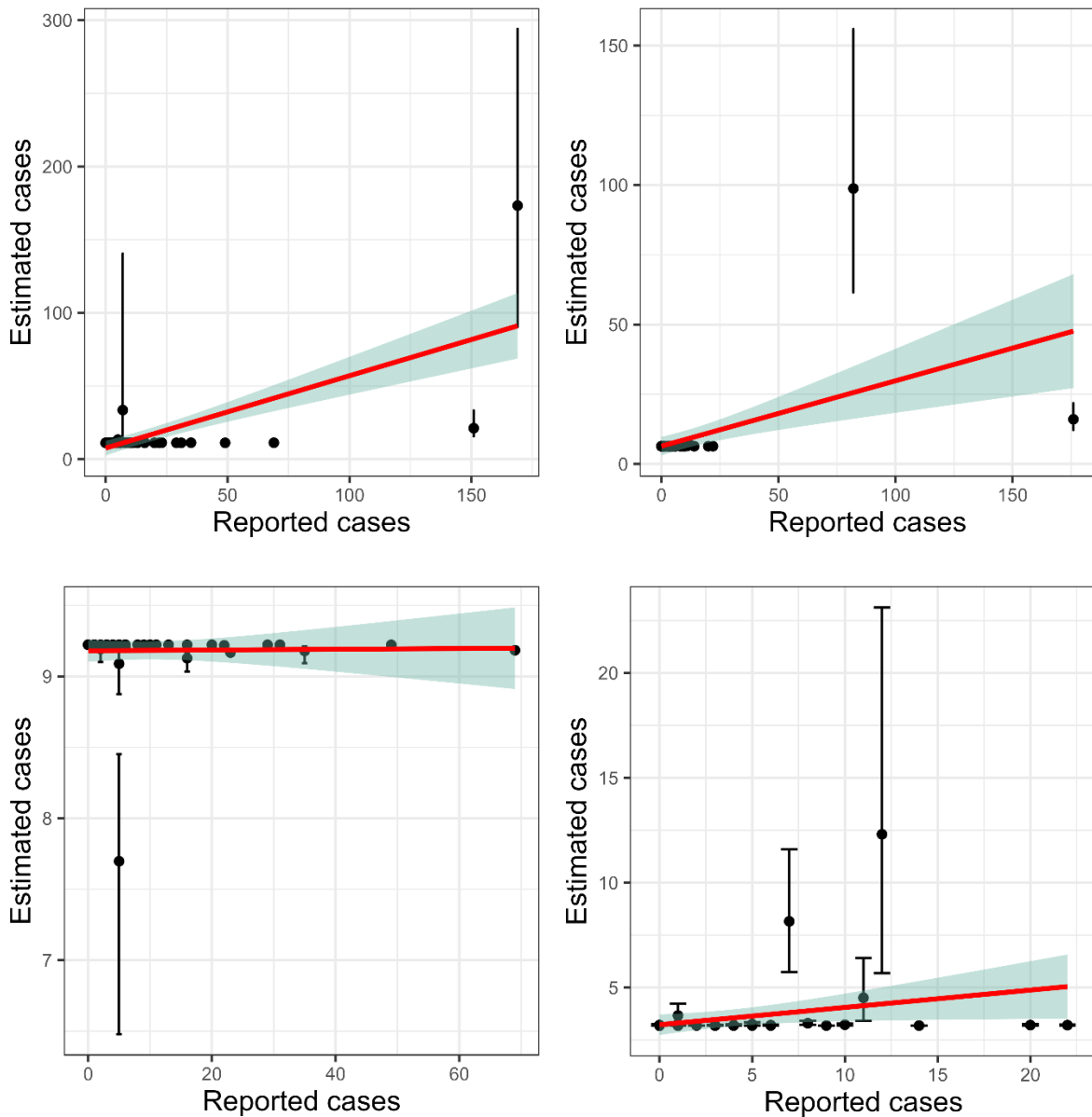
**Table S1: Summary of the models that most accurately approximated the 2015-2018 reported cases after excluding Madrid and Barcelona from the analysis.** Each row shows the statistics that provided the best estimate of imported cases of each disease), the correlation with the actually reported data in 2019, and the proportion of overestimation as obtained from the linear models. For the models including weighted averages, the weight  $\alpha_i$  of each indicator is included in parenthesis. If the weight of a given indicator is zero, this means that no contribution to the estimate is provided by the corresponding indicator.



**Figure S4. Summary of the best simple models for imported cases of malaria (top row) and dengue (bottom row), after excluding Madrid and Barcelona from the analysis.** The left column shows the models' predictions (in red), together with the actually reported cases (in blue) for 2019 for each province in Spain. The right column shows the fit between the models' estimates and the official records.

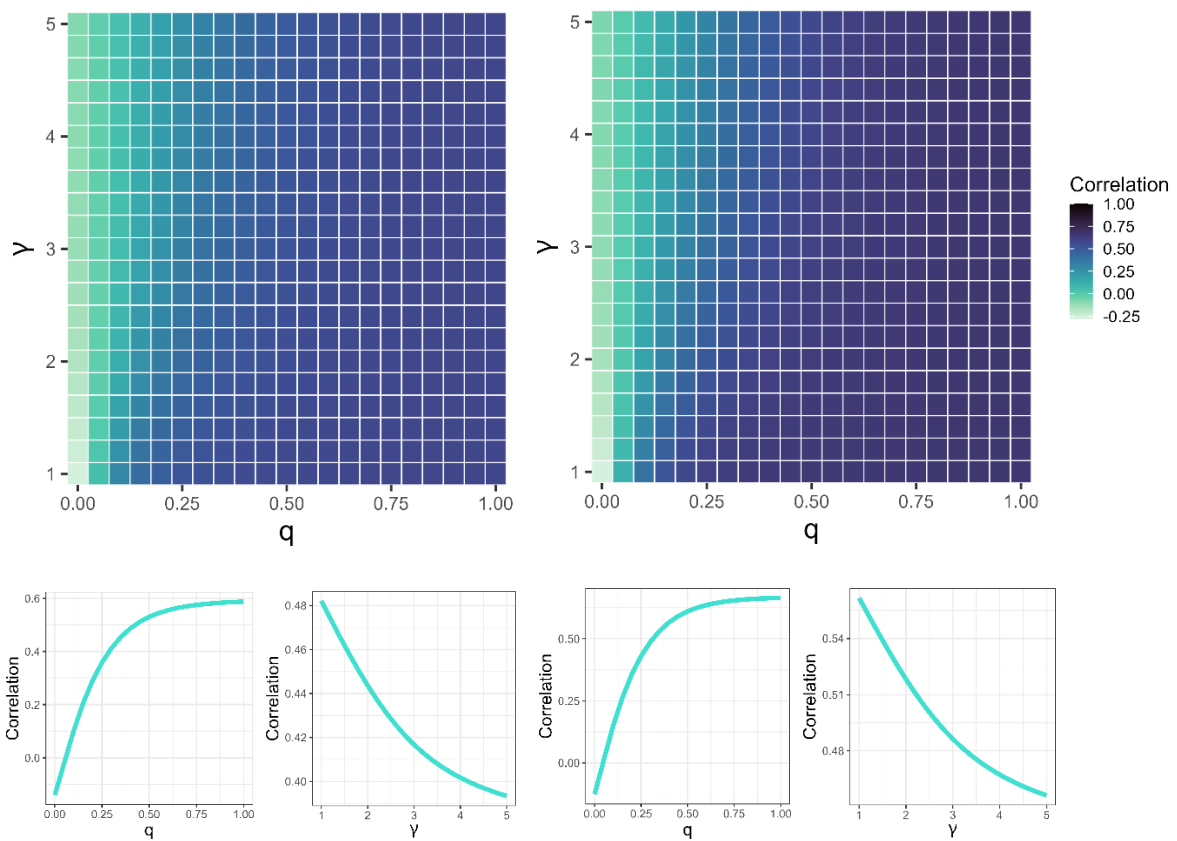
### Human mobility models

Despite the relatively good score in terms of correlation with 2019 data found for the most accurate human mobility models (Table 3 in the main text), visual inspection of the models' fit showed that these results could be due to the dominance of some provinces from the dataset (Figure S5). After excluding these provinces from the analysis, the models' accuracy decreased significantly (Table S2). Average correlations with 2019 data in terms of models' parameters are provided in Figure S6 for general diagnostics.



**Figure S5. Models' fit for the human mobility models for malaria (left column) and dengue (right column).**

Top row shows the fit for the complete dataset of 52 provinces in Spain; bottom row shows the fit after removing the outliers from the analysis. Madrid, Barcelona and Las Palmas were removed from the malaria dataset, and Madrid and Barcelona were removed from the dengue dataset.



**Figure S6. Accuracy of the human mobility model in terms of the models' parameters for malaria (left column) and dengue (right column).** In each column, the top row shows the average correlation of the mobility models' estimates of imported cases with the 2019 official data in terms of the proportion of travelers that do not move after their arrival ( $q$ ) and the exponent of the power law distribution ( $\gamma$ ). Bottom row shows the average correlation of the models in terms of  $q$  and  $\gamma$  separately (left plot for  $q$ , right plot for  $\gamma$ ).

Model	Proportion of cases that do not move after arrival ( $q$ )	Exponent of power law distribution ( $\gamma$ )	Correlation with 2019 data	Overestimation/underestimation
Malaria	1	Any	0.003	97.7%
Dengue	1	Any	0.12	17%

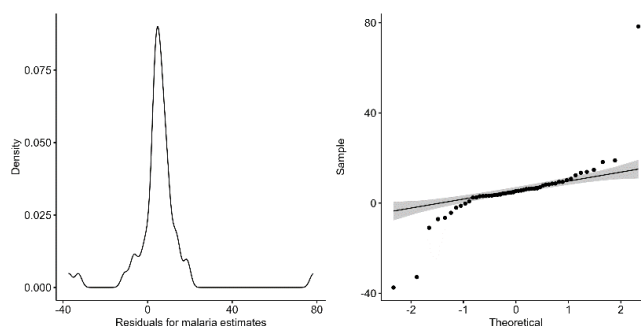
**Table S2: Summary of the human mobility models that most accurately approximated the 2019 reported cases after excluding outliers from the analysis.** Each row shows the parameters of the model that provided the best estimate of imported cases of each disease, the correlation with the actually reported data in 2019, and the overestimation/underestimation of cases as obtained from the linear models. Madrid, Barcelona and Las Palmas were removed from the malaria dataset, and Madrid and Barcelona were removed from the dengue dataset.

## Residual analysis

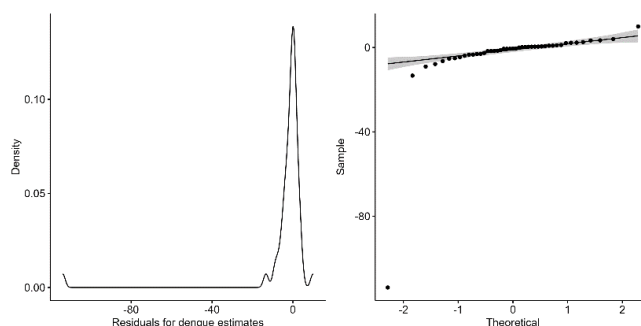
We analyze further the models' performance by inspecting their residuals, computed as the difference between the models' estimates and the reported cases. We performed a Shapiro-Wilk test, to check for normality of the residuals, and a Durbin-Watson test to check for autocorrelation. Table S3 shows the results of this analysis, and Figures S7-S9 show the distribution and QQ-plots of the models' residuals.

Model	Shapiro-Wilk test		Durbin-Watson test	
	Statistic	p-value	Statistic	p-value
Malaria	0.67	<0.001	2.03	0.543
Dengue	0.31	0.02	2.25	0.799
Dengue (excluding Barcelona)	0.94	<0.001	1.87	0.33

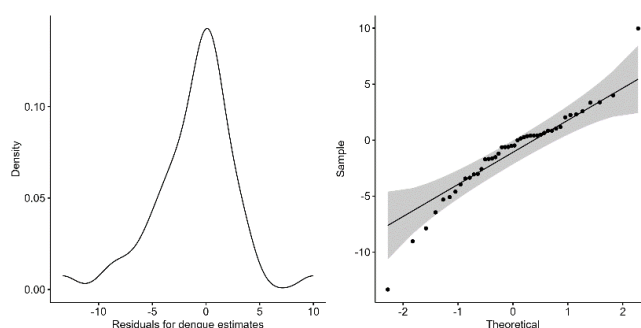
**Table S3: Results of the residual analysis.** Scores and p-values reported by the residuals of the weighted models in the Shapiro-Wilk and Durbin-Watson tests for normality and autocorrelation.



**Figure S7. Residuals' density (left) and QQ-plot (right) for the malaria estimates.** Plots obtained from the difference between the weighted model's estimate and the reported cases in 2019.



**Figure S8. Residuals' density (left) and QQ-plot (right) for the dengue estimates.** Plots obtained from the difference between the weighted model's estimate and the reported cases in 2019.



**Figure S9. Residuals' density (left) and QQ-plot (right) for the dengue estimates, upon removing Barcelona from the analysis.** Plots obtained from the difference between the weighted model's estimate and the reported cases in 2019.