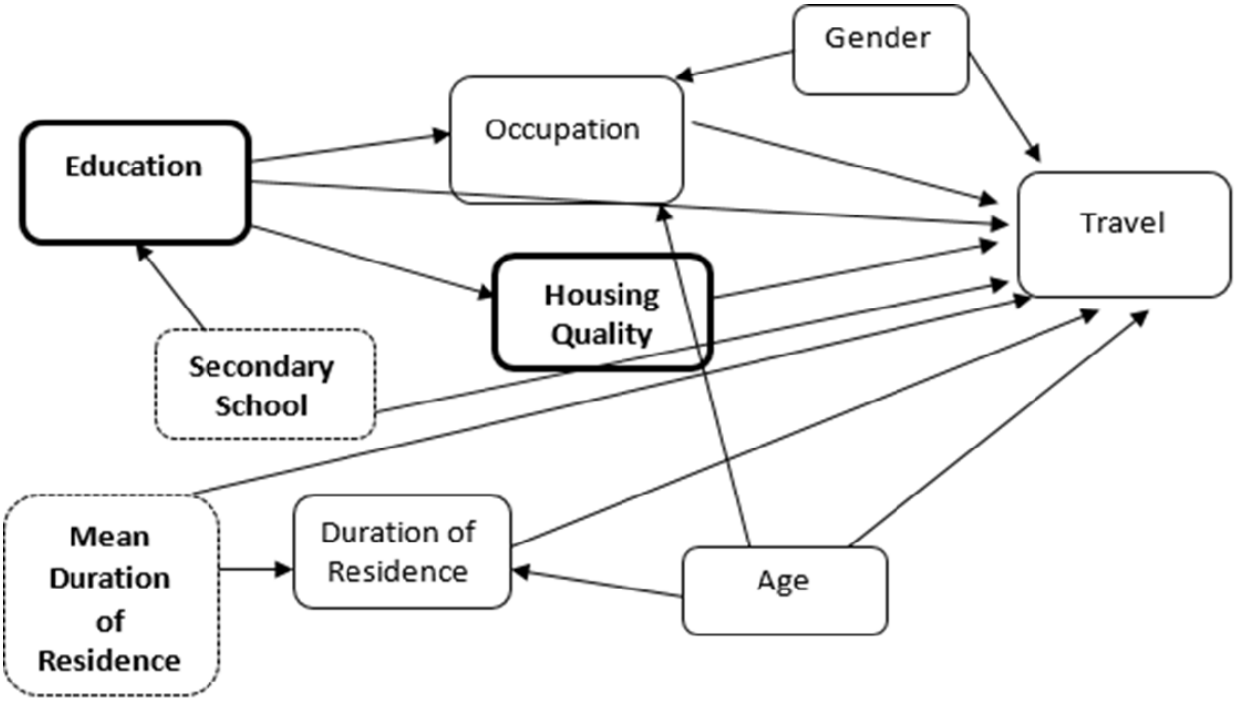
Supplemental Information

Causal Diagram

The causal diagram relating demographic variables to travel is shown in eFigure1 below. Baseline remoteness is considered a common cause of all demographic variables as well as a cause of travel patterns (not shown).



eFigure 1: Demographic Variables Conceptual Diagram. Dashed boxes, thick boxes, and think boxes are shown for community, household, and individual variables respectively.

Transmission Model Details

We developed a deterministic, compartmental susceptible, infected, recovered (SIR) transmission model with two patches at a time. One patch was parameterized to reflect a generic close, medium, or far village, and the other patch had the same size and population structure as Borbón. All disease processes were parameterized to represent rotavirus. Contact between the communities occurred

when members of the in-region village traveled to the out-of-region city. Each simulation only considered one village (close, medium, or far) in addition to the city; interactions between the modeled village and other villages were not considered both because these communities villages are relatively small and would not be expected to sustain an outbreak in isolation and because of the geography of our study region. Although villagers might leave our study region to travel to other remote communities and not urban centers, due to the geography and transportation network of our study region, these travelers would have to pass through the larger city of Borbon to access transportation to other more distant communities. We stratified the model by age, dichotomized at five years, to capture heterogeneity in risk of infection and travel. Children under five were more likely to become infected and less likely to travel. Both travel patterns and disease transmission parameters were allowed to vary by age. For this model, we allow travelers to infect and be infected during travel. To do this, we allowed members of the village compartments to move back and forth to a corresponding "traveling" compartment imagined to be physically located at the other patch; the transmission rates of villagers in the traveling compartment are the same as the people in that new patch until they return, following the method of Knipl [1]. We also assume that infected and susceptible individuals have the same travel probabilities.

This model includes three different transmission relevant processes: (1) infection, (2) recovery, and (3) travel. We used survey data to estimate parameters for infection and travel. The recovery rate and relative transmission rates for adults and children were taken from prior literature [31]. To estimate the transmission β terms for each community, we used age-specific prevalence of infection data from stool samples collected as part of the population-based case control study in 2007, before the vaccine was introduced. Because our study region is near its endemic equilibrium, we used steady state formulas to calculate the value of R_0 based on life expectancy and average age at first infection [37], and corrected for the fraction of cases that were symptomatic. We used prevalence ratios by remoteness [16] to fix the value of relative β terms by remoteness. To parameterize the travel portion of our model, we used the survey data for each survey year (2003, 2007, 2010, and 2013) to derive daily travel rates to the city for close, medium, and far villages. Additionally, we used census data collected as part of the EcoDess study to account for the relative population sizes of the different community population sizes and the fraction of the village in each age group for each of the four study years.

In this model, the $S_{i,j}$, $I_{i,j}$, and $R_{i,j}$ compartments represent susceptible, infectious, and recovered individuals in patch i, where $i = 1$ is the within-region village and $i = 2$ is the city, and of age group j (where A is $\leq$5 and C is < 5). Compartments $S^*_{1,j}$, $I^*_{1,j}$, and $R^*_{1,j}$ denote villagers of age j who have traveled to the city. The daily transmission rate from age group j to age group k in community i is given by $\beta_{jk,i}$. The travel rate from community i to ℓ is given by $\tau_{i\ell}$. The recovery rate is given by γ and is independent of age or community. The ordinary differential equations for this model are shown below.

$$
\begin{aligned}
\dot{S}_{1,C} =& - S_{1,C}[\beta_{CC,1} I_{1,C} + \beta_{AC,1} I_{1,A}] - \tau_{12,C} S_{1,C} + \tau_{21,C} S^*_{1,C} \\
\dot{I}_{1,C} =& S_{1,C}[\beta_{CC,1} I_{1,C} + \beta_{AC,1} I_{1,A}] - \tau_{12,C} I_{1,C} + \tau_{21,C} I^*_{1,C} - \gamma I_{1,C} \\
\dot{R}_{1,C} =& \gamma I_{1,C} - \tau_{12,C} R_{1,C} + \tau_{21,C} R^*_{1,C} \\
\dot{S}_{1,A} =& - S_{1,A}[\beta_{AA,1} I_{1,A} + \beta_{CA,1} I_{1,C}] - \tau_{12,A} S_{1,A} + \tau_{21,A} S^*_{1,A} \\
\dot{I}_{1,A} =& S_{1,A}[\beta_{AA,1} I_{1,A} + \beta_{CA,1} I_{1,C}] - \tau_{12,A} I_{1,A} + \tau_{21,A} I^*_{1,A} - \gamma I_{1,A} \\
\dot{R}_{1,A} =& \gamma I_{1,A} - \tau_{12,A} R_{1,A} + \tau_{21,A} R^*_{1,A} \\
\dot{S}^*_{1,C} =& - S^*_{1,C}[\beta_{CC,2} (I^*_{1,C} + I_{2,C}) + \beta_{AC,2} (I^*_{1,A} + I_{2,A})] - \tau_{21,C} S^*_{1,C} + \tau_{12,C} S^*_{1,C} \\
\dot{I}^*_{1,C} =& S^*_{1,C}[\beta_{CC,2} (I^*_{1,C} + I_{2,C}) + \beta_{AC,2} (I^*_{1,A} + I_{2,A})] - \tau_{21,C} I^*_{1,C} + \tau_{12,C} I_{1,C} - \gamma I_{1,C} \\
\dot{R}^*_{1,C} =& \gamma I *_{1,C} - \tau_{21,C} R *_{1,C} + \tau_{12,C} R_{1,C} \\
\dot{S}^*_{1,A} =& - S^*_{1,A}[\beta_{AA,2} (I^*_{1,A} + I_{2,A}) + \beta_{CA,2} (I^*_{1,C} + I_{2,C})] - \tau_{21,A} S^*_{1,A} + \tau_{12,A} S^*_{1,A} \\
\dot{I}^*_{1,A} =& S^*_{1,A}[\beta_{AA,2} (I^*_{1,A} + I_{2,A}) + \beta_{CA,2} (I^*_{1,C} + I_{2,C})] - \tau_{21,A} I^*_{1,A} + \tau_{12,A} I_{1,A} - \gamma I_{1,A} \\
\dot{R}^*_{1,A} =& \gamma I^*_{1,A} - \tau_{21,A} R^*_{1,A} + \tau_{12,A} R_{1,A} \\
\dot{S}_{2,C} =& - S_{2,C}[\beta_{CC,2} (I_{2,C} + I^*_{1,C}) + \beta_{AC,2} (I^*_{1,A} + I_{2,A})] \\
\dot{I}_{2,C} =& S_{2,C}[\beta_{CC,2} (I_{2,C} + I^*_{1,C}) + \beta_{AC,2} (I^*_{1,A} + I_{2,A})] - \gamma I_{2,C} \\
\dot{R}_{2,C} =& \gamma I_{2,C} \\
\dot{S}_{2,A} =& - S_{2,A}[\beta_{AA,2} (I_{2,A} + I^*_{1,A}) + \beta_{CA,2} (I^*_{1,C} + I_{2,C})] \\
\dot{I}_{2,A} =& S_{2,A}[\beta_{AA,2} (I_{2,A} + I^*_{1,A}) + \beta_{CA,2} (I^*_{1,C} + I_{2,C})] - \gamma I_{2,A} \\
\dot{R}_{2,A} =& \gamma I_{2,A}
\end{aligned}
$$

To fully understand the impact of travel and transmissibility on disease risk (measured by cumulative incidence for all models), we used a series of models. To assess the overall changes over time and their effects on risk, we used survey data from all time points in our study region for close, medium, and far villages to estimate the net effects of these on risk of rotavirus. The derivation of all parameters and their values are described in detail below. However, because many variables changed for each time point simultaneously, we also conducted what we call a pure effects analysis where we changed only the travel parameters and kept all other factors constant. For our first pure effects model, we systematically increased the average travel. For the second pure effects model, we fixed the travel patterns of children and increased the travel of adults by a proportionality constant to explore the role of heterogeneity. This model allowed us to assess the effect of heterogeneity in travel on disease risk. Using these models, we were able to assess to what extent the changes in our study region were driven by demographic change versus changes in travel.

Model Parameterization

We used a combination of literature and survey data to parameterize our model. Parameter values and their sources are shown in eTable 1 below. When the source is "Estimated," we describe the estimation in a later section.

eTable 1: Parameters used in model simulations. All parameters have units of days^{-1}.

Parameter	Value				Source
	Close	Medium	Far	City	
Child-Child Transmission Rate (β_{CC})	0.275	0.5	0.331	0.66	Estimated
Child-Adult Transmission Rate (β_{CA})	0.138	0.25	0.166	0.33	Estimated
Adult-Child Transmission Rate (β_{AC})	0.275	0.5	0.331	0.66	Estimated
Adult-Adult Transmission Rate (β_{AA})	0.138	0.25	0.166	0.33	Estimated
Village-city Travel Rate for Adults ($\tau_{12,A}$)	0.029	0.027	0.014	—	Estimated
City-Village Travel Rate for Adults ($\tau_{21,A}$)	1	1	1	—	Estimated
Village-city Travel Rate for Children ($\tau_{12,C}$)	0.017	0.03	0.0063	—	Estimated
City-Village Travel Rate for Children ($\tau_{21,C}$)	1	1	1	—	Estimated
Recovery Rate (γ)	0.2	0.2	0.2	0.2	[2]

In addition to differences in disease and travel parameters by community, we accounted for their different population sizes by fixing the relative proportions of susceptibles in the out of region community and the community of interest. To get the size of the 'close', 'medium' and 'far' villages, we used the mean population size for all time points for the 15 communities in our study, stratified by remoteness. We elected to use the mean for all time points and not the last time point because there were no significant time trends in population size. The mean population sizes were 575, 108, and 246 for close, medium, and far villages respectively. We assume that the out of region village has a population size of 5,000, comparable to Borbón.

Transmission Rate Calculations

Because our study region is close to its endemic equilibrium, we used the following steady state formula to calculate the value of R_0 [3]:

$$R_0 = L/A$$

In this equation, L is the average life expectancy and A is the average age of first infection. Other researchers have used an alternative formulation of R_0 ($R_0 = 1 + \frac{L}{A}$) but later analysis by Dietz has suggested that omitting the 1 ($R_0 = \frac{L}{A}$) is more accurate [3]. We also conducted sensitivity analysis for a wide range of β values and these results are presented in a later section. We assumed a life expectancy of 70 years and approximated the average age of first infection in our study region using the average age of those infected in 2007, corrected for the fraction symptomatic. In 2007, the average age of first infection was 15.7 for all cases and 11.4 for symptomatic cases. Since the only pathway of infection in this model is direct transmission, the average R_0 is simply $\frac{\beta}{\gamma}$. Fixing the value of γ to 0.2, we estimated that the average β value for the region overall was 0.315. To estimate the local transmission parameters for close, medium, and far villages, we used ratios of prevalence of infection by remoteness as a proxy for the β term of that village [4]. In this paper, the prevalence of symptomatic infection was found to be 0.6, 0.9, and 0.5 per 100 persons for remote, medium, and close villages respectively and 1.2 per 100 persons in Borbón. Taking Borbón as the reference, these values translate to prevalence ratios of 0.5, 0.75, and 0.42 for remote, medium and close villages versus Borbón respectively. To be consistent with prior literature, we assumed

that transmission to children had a β term that was twice as high as transmission to adults [2]. Thus the estimated β term for a given age group within a particular village was a weighted average of the transmission patterns of adults and children based on their population size.

Patch Specific R_0

For the full model, the absolute value of R_0 is not especially informative. What is more useful is the patch-specific R_0 values, which suggest the transmission potential within a community after the outbreak is seeded by travel to the city. The reduced model for a single village and the resulting R_0 calculations are shown below.

$$\begin{aligned}\dot{S}_A =& - S_A[\beta_A I_A + \beta_{CA} I_C] \\ \dot{I}_A =& S_A[\beta_A I_A + \beta_{CA} I_C] - \gamma I_A \\ \dot{R}_A =& \gamma I_A\end{aligned}$$

$$\begin{aligned}\dot{S}_C =& - S_C[\beta_C I_C + \beta_{AC} I_A] \\ \dot{I}_C =& S_C[\beta_C I_C + \beta_{AC} I_A] - \gamma I_C \\ \dot{R}_C =& \gamma I_C\end{aligned}$$

Using the next generation matrix approach, we calculate the matrices F, V, and V^{-1}. The largest eigenvalue of the matrix product FV^{-1} is R_0. The calculations that lead to this result are shown below.

$$f = \begin{bmatrix}\beta_A S_A I_A + \beta_{CA} S_A I_C \\ \beta_C S_C I_C + \beta_{AC} S_C I_A\end{bmatrix}$$

$$v = \begin{bmatrix}\gamma I_A \\ \gamma I_C\end{bmatrix}$$

Now we take the Jacobians of both matrices at the disease free equilibrium to produce F and V.

$$\mathcal{F} = \begin{bmatrix}\beta_A S_A & \beta_{CA} S_A \\ \beta_{AC} S_C & \beta_C S_C\end{bmatrix}$$

$$\mathcal{V} = \begin{bmatrix}\gamma & 0 \\ 0 & \gamma\end{bmatrix}$$

Then, we take the inverse of V

$$V^{-1} = \begin{bmatrix} 1/\gamma & 0 \\ 0 & 1/\gamma \end{bmatrix}$$

Then the matrix product FV^{-1} is

$$FV^{-1} = \begin{bmatrix} \frac{\beta_A S_A}{\gamma} & \frac{\beta_{CA} S_A}{\gamma} \\ \frac{\beta_{AC} S_C}{\gamma} & \frac{\beta_C S_C}{\gamma} \end{bmatrix}$$

The largest Eigenvalue of this matrix was calculated using Mathematica and is:

$$\frac{\sqrt{(\beta_A)^2 (S_A)^2 - 2\beta_A \beta_{AC} S_A S_C + (\beta_{AC})^2 (S_C)^2} + \beta_A S_A + \beta_{AC} S_C}{2\gamma}$$

Global Reproduction Number

In all cases, the system R_0 once travel is included are greater than 1 because of travel to the city.

Although the basic reproduction number for the full model will always be greater than 1 in our study villages, we show our basic setup to allow other researchers to apply our findings to other settings in which travel is needed to bring the global R_0 above 1.

To calculate R_0, we begin by calculating the matrices F and V representing the rates of new infections and net compartment transfer, respectively.

$$f = \begin{bmatrix} S_{1,C} \left(\beta_{CC,1} I_{1,C} + \beta_{AC,1} I_{1,A}\right) \\ S_{1,A} \left(\beta_{CA,1} I_{1,C} + \beta_{AA,1} I_{1,A}\right) \\ S^*_{1,C} \left(\beta_{CC,2}(I_{2,C} + I^*_{1,C}) + \beta_{AC,2}(I_{2,A} + I^*_{1,A})\right) \\ S^*_{1,A} \left(\beta_{CA,2}(I_{2,C} + I^*_{1,C}) + \beta_{AA,2}(I_{2,A} + I^*_{1,A})\right) \\ S_{2,C} \left(\beta_{CC,2}(I_{2,C} + I^*_{1,C}) + \beta_{AC,2}(I_{2,A} + I^*_{1,A})\right) \\ S_{2,A} \left(\beta_{AC,2}(I_{2,C} + I^*_{1,C}) + \beta_{AA,2}(I_{2,A} + I^*_{1,A})\right) \end{bmatrix}$$

$$v = \begin{bmatrix} \tau_{12,C} I_{1,C} - \tau_{21,C} I^*_{1,C} + \gamma I_{1,C} \\ \tau_{12,A} I_{1,A} - \tau_{21,A} I^*_{1,A} + \gamma I_{1,A} \\ \tau_{12,C} I^*_{1,C} - \tau_{21,C} I_{1,C} + \gamma I^*_{1,C} \\ \tau_{12,A} I^*_{1,A} - \tau_{21,A} I_{1,A} + \gamma I^*_{1,A} \\ \gamma I_{2,C} \\ \gamma I_{2,A} \end{bmatrix}$$

Now we take the Jacobians of both matrices at the disease free equilibrium

$$\mathcal{F} = \begin{bmatrix} \beta_{CC,1}\bar{S}_{1,C} & \beta_{AC,1}\bar{S}_{1,C} & 0 & 0 & 0 & 0 \\ \beta_{CA,1}\bar{S}_{1,A} & \beta_{AA,1}\bar{S}_{1,A} & 0 & 0 & 0 & 0 \\ 0 & 0 & \beta_{CC,2}\bar{S}^*_{1,C} & \beta_{AC,2}\bar{S}^*_{1,C} & \beta_{CC,2}\bar{S}^*_{1,C} & \beta_{AC,2}\bar{S}^*_{1,C} \\ 0 & 0 & \beta_{CA,2}\bar{S}^*_{1,A} & \beta_{AA,2}\bar{S}^*_{1,A} & \beta_{CA,2}\bar{S}^*_{1,A} & \beta_{AA,2}\bar{S}^*_{1,A} \\ 0 & 0 & \beta_{CC,2}\bar{S}_{2,C} & \beta_{AC,2}\bar{S}_{2,C} & \beta_{CC,2}\bar{S}_{2,C} & \beta_{AC,2}\bar{S}_{2,C} \\ 0 & 0 & \beta_{CA,2}\bar{S}_{2,A} & \beta_{AA,2}\bar{S}_{2,A} & \beta_{CA,2}\bar{S}_{2,A} & \beta_{AA,2}\bar{S}_{2,A} \end{bmatrix}$$

$$\mathcal{V} = \begin{bmatrix} \tau_{12,C}+\gamma & 0 & -\tau_{21,C} & 0 & 0 & 0 \\ 0 & \tau_{12,A}+\gamma & 0 & -\tau_{21,A} & 0 & 0 \\ -\tau_{12,C} & 0 & \tau_{21,C}+\gamma & 0 & 0 & 0 \\ 0 & -\tau_{12,A} & 0 & \tau_{21,A}+\gamma & 0 & 0 \\ 0 & 0 & 0 & 0 & \gamma & 0 \\ 0 & 0 & 0 & 0 & 0 & \gamma \end{bmatrix}$$

We then take the inverse of V.

$$V^{-1} = \begin{bmatrix} \frac{\gamma+\tau_{21,C}}{\gamma^2+2\tau_{12,C}\tau_{21,C}+\gamma(\tau_{12,C}+\tau_{21,C})} & 0 & -\frac{\tau_{21,C}}{\gamma^2+2\tau_{12,C}\tau_{21,C}+\gamma(\tau_{12,C}+\tau_{21,C})} & 0 & 0 & 0 \\ 0 & \frac{\gamma+\tau_{21,A}}{\gamma(\gamma+\tau_{12,A}+\tau_{21,A})} & 0 & \frac{\tau_{21,A}}{\gamma(\gamma+\tau_{12,A}+\tau_{21,A})} & 0 & 0 \\ \frac{\tau_{12,C}}{\gamma^2+2\tau_{12,C}\tau_{21,C}+\gamma(\tau_{12,C}+\tau_{21,C})} & 0 & \frac{\gamma+\tau_{12,C}}{\gamma^2+2\tau_{12,C}\tau_{21,C}+\gamma(\tau_{12,C}+\tau_{21,C})} & 0 & 0 & 0 \\ 0 & \frac{\tau_{12,A}}{\gamma(\gamma+\tau_{12,A}+\tau_{21,A})} & 0 & \frac{\gamma+\tau_{12,A}}{\gamma(\gamma+\tau_{12,A}+\tau_{21,A})} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\gamma} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{\gamma} \end{bmatrix}$$

To calculate R_0, the next steps are to calculate the matrix product FV^{-1} and its eigenvalues. The largest eigenvalue of the next generation matrix FV^{-1} is the global R_0. While this matrix product and its eigenvalues can be calculated, the next generation matrix is quite large and the resulting eigenvalues are on the order of 6000 terms long. For this reason, they are not presented here. Readers interested in calculating the global R_0 for similar systems might wish to do so numerically.

Travel Rate Calculations

The fraction of people in a given age group traveling to the out-of-region city in the last seven days is known from survey data and is denoted ω. Assume that the time between travel is exponentially distributed with mean $1/\tau$. Then, the relationship between ω and τ is

$$\omega = \int_0^7 \tau e^{-\tau t}\, dt = 1 - e^{-7\tau},$$

implying

$$\tau = -\frac{\log(1-\omega)}{7}.$$

For example, if 17.5% of people (the value for close villages) traveled in the last seven days, then τ =0.0275, with an average of travel every 36 days. Similarly, if only 8.5% of people (the value for far villages) traveled in the last seven days, then $\tau = 0.013$, with an average of travel every 79 days. We assume duration of travel is short, one day on average.

Simulation Approach

For each model, we simulated for 200 days with an infection seed of 1% of the population in the city and calculated the cumulative incidence for each age group as well as the fraction of infections acquired in the in-region village.

We investigated the impact of changing the average frequency of travel, the average rate of transmission, and heterogeneity in both factors by age (children vs. adults). To investigate the average effects of travel and transmission, we used the range of travel and infectivity parameters estimated from our data. For travel, we used the measured values of travel for each remoteness level. Because the village-specific R_0 values in our study region varied from 0.79 (close villages) to 1.43 (medium villages), we used the β parameters from these two village types to represent the range of transmission rates seen in our study region. To investigate heterogeneity in travel, we kept the travel rates of children constant at their measured value for close villages (τ=.017 visits/day) and increased travel rates for adults only by a proportionality constant. To investigate heterogeneity in transmission, we fixed the transmission rates of adults and increased transmission for children by a proportionality constant.

Net Effects: Parameter Values

The parameter values for each study year for the net effects models are shown in the tables below. Parameters for close, medium, and far villages are shown in eTable 2, eTable 3, and eTable 4 respectively. Parameters for the city are shown in eTable 5 below. The parameters that were varied for the pure effects models are also indicated with a † symbol in each table.

Sensitivity Analysis

To enhance the generalizability of our findings to other regions that might have different levels of disease transmissibility and population structures, we also conducted sensitivity analysis across a range of β values and population sizes. We also considered a model with three age groups to account for the greater complexity of our data (<5, 5–13, >13) but found that the results were qualitatively similar.

In all of our analyses, the city was the major driver of transmission. To determine if this effect was predominantly driven by the population size of the city relative to the village or the relative infectivity in the city, we systematically varied both characteristics.

eTable 2: Model parameters for close villages for each study year. Population sizes were based on census data. Travel was based on the proportion reporting travel in the sociometric survey, which was then used to derive the specific τ parameters, as described above. We used stool sample data to estimate transmission parameters but assumed the ratio children: adults was 2 and fixed gamma based on prior literature [2]

	2003	2007	2010	2013
Population size (Survey Data)				
Overall	616	683	741	686
Proportion $<$ 5 years	0.192	0.221	0.205	0.145
Proportion $>=$ 5 years	0.808	0.779	0.795	0.855
Ratio Population Village: City	0.11	0.12	0.13	0.12
Travel				
Survey Data				
Overall Travel	0.082	0.14	0.075	0.175
Travel $<$ 5	0.043	0.081	0.069	0.113
Travel $>=$ 5	0.093	0.156	0.077	0.185
Derived Model Parameters				
τ Average †	0.012	0.022	0.011	0.027
$\tau_{12,C}$	0.006	0.012	0.010	0.017
$\tau_{12,A}$	0.014	0.024	0.011	0.029
τ Ratio (Adult to Child)†	2.22	2.01	1.12	1.71
Transmission Parameters				
β_C	0.275	0.275	0.275	0.275
β_{AC}	0.275	0.275	0.275	0.275
β_A	0.138	0.138	0.138	0.138
β_{CA}	0.138	0.138	0.138	0.138
γ	0.2	0.2	0.2	0.2
Within Patch R_0	0.82	0.84	0.83	0.79

†Varied for pure effects models

eTable 3: Model parameters for medium villages for each study year. Population sizes were based on census data. Travel was based on the proportion reporting travel in the sociometric survey, which was then used to derive the specific τ parameters, as described above. We used stool sample data to estimate transmission parameters but assumed the ratio children: adults was 2 and fixed gamma based on prior literature [2]

	2003	2007	2010	2013
Population size (Survey Data)				
Overall	119	118	121	123
Proportion $<$ 5 years	0.214	0.221	0.196	0.135
Proportion $>=$ 5 years	0.786	0.779	0.804	0.865
Ratio Population Village: City	0.023	0.023	0.024	0.024
Travel				
Survey Data				
Overall Travel	0.071	0.08	0.092	0.155
Travel $<$ 5	0.015	0.013	0.044	0.189
Travel $>=$ 5	0.086	0.099	0.104	0.15
Derived Model Parameters				
τ Average†	0.011	0.012	0.014	0.024
$\tau_{12,C}$	0.002	0.002	0.006	0.030
$\tau_{12,A}$	0.013	0.015	0.016	0.023
τ Ratio (Adult to Child) †	5.95	7.97	2.44	0.776
Transmission Parameters				
β_C	0.5	0.5	0.5	0.5
β_{AC}	0.5	0.5	0.5	0.5
β_A	0.25	0.25	0.25	0.25
β_{CA}	0.25	0.25	0.25	0.25
γ	0.2	0.2	0.2	0.2
Within Patch R_0	1.52	1.53	1.5	1.42

† Varied for pure effects models

eTable 4: Model parameters for far villages for each study year. Population sizes were based on census data. Travel was based on the proportion reporting travel in the sociometric survey, which was then used to derive the specific τ parameters, as described above. We used stool sample data to estimate transmission parameters but assumed the ratio children: adults was 2 and fixed gamma based on prior literature [2]

	2003	2007	2010	2013
Population size (Survey Data)				
Overall	307	305	326	313
Proportion $<$ 5 years	0.196	0.197	0.187	0.141
Proportion $>=$ 5 years	0.804	0.803	0.813	0.859
Ratio Population Village: City	0.058	0.057	0.061	0.059
Travel				
Survey Data				
Overall Travel	0.04	0.063	0.054	0.086
Travel $<$ 5	0.021	0.023	0.034	0.043
Travel $>=$ 5	0.044	0.072	0.058	0.093
Derived Model Parameters				
τ Average†	0.006	0.009	0.008	0.013
$\tau_{12,C}$	0.003	0.003	0.005	0.006
$\tau_{12,A}$	0.006	0.011	0.009	0.014
τ Ratio (Adult to Child)	2.12	3.21	1.73	2.22
Transmission Parameters				
β_C	0.331	0.331	0.331	0.331
β_{AC}	0.331	0.331	0.331	0.331
β_A	0.166	0.166	0.166	0.166
β_{CA}	0.166	0.166	0.166	0.166
γ	0.2	0.2	0.2	0.2
Within Patch R_0	0.99	0.99	0.99	0.95

† Varied for pure effects models

eTable 5: Model parameters for city for each study year. Population sizes were assumed constant and based on Borbon and the ratio of children to adults was set to like close villages. We did not consider travel from the city except returning villagers. Since we assumed the travel duration was 1 day, all τ values for the city were fixed to 1. As for the villages, we used stool sample data to estimate transmission parameters but assumed the ratio children:adults was 2 and fixed gamma based on prior literature [2]

	2003	2007	2010	2013
Population size				
Overall	5000	5000	5000	5000
Proportion $<$ 5 years	0.192	0.221	0.205	0.145
Proportion $>=$ 5 years	0.808	0.779	0.795	0.855
Travel Parameters				
τ Average †	1	1	1	1
$\tau_{21,C}$	1	1	1	1
$\tau_{21,A}$	1	1	1	1
τ Ratio (Adult to Child)†	1	1	1	1
Transmission Parameters				
β_C	0.66	0.66	0.66	0.66
β_{AC}	0.66	0.66	0.66	0.66
β_A	0.33	0.33	0.33	0.33
β_{CA}	0.33	0.33	0.33	0.33
γ	0.2	0.2	0.2	0.2
Within Patch R_0	1.97	2.01	1.99	1.89

†Varied for pure effects models for the villages, but not the city

Population Size

For this sensitivity analysis, travel patterns for adults and children were fixed to the average for close villages (0.027) for both children and adults. The transmission terms for the village (patch 1) were allowed to differ by age and were set to the estimated values for medium villages. As stated in the text, as the population sizes of the city and village became more similar, the fraction of transmission attributable to local transmission increased proportionally with population size but the overall incidence decreased. See eTable 6 below.

Transmissibility

We also considered the impact of varying the transmissibility of rotavirus in each community. We considered three approaches: 1) increasing the transmissibility (β terms) in the village only, such that the local R_0 of the village and the city became more similar; 2) increasing transmissibility for both the village and the city proportionally, such that the ratio of the R_0 for the village and the city remained the same; and 3) allowing transmission parameters to be equal (such that the ratio of R_0 values was 1) in the city and the village but increasing the value of R_0. This increase in R_0 was done by keeping transmission to adults constant and increasing it for children only.

In scenario 1, we found that increasing the R_0 of the village had no effect on risk (see eTable 7), but impacted the distribution of cases for children and led to more local transmission. When both the transmissibility of the village and the city were increased (scenario 2, see eTable 8), the total

eTable 6: Cumulative incidence and proportion of locally acquired infection by relative population size.

Village population size (% of total population)	Low within-village transmission		High within-village transmission	
	Children	Adults	Children	Adults
Cumulative Incidence				
5%	0.056	0.029	0.057	0.030
10%	0.052	0.027	0.055	0.028
15%	0.048	0.025	0.053	0.027
20%	0.044	0.022	0.050	0.025
25%	0.038	0.019	0.046	0.023
Proportion Locally Acquired				
5%	0.03	0.03	0.06	0.06
10%	0.06	0.06	0.12	0.11
15%	0.10	0.09	0.17	0.17
20%	0.13	0.13	0.24	0.23
25%	0.16	0.16	0.30	0.30

number of cases increased. The fraction of cases that were locally acquired increased slightly, but the majority of cases were acquired in the city. When the values of R_0 for the city and village were kept equal, the city remained the primary source of cases although the total number of cases increased as the value of R_0 increased.

eTable 7: Sensitivity analysis for scenario 1. The R_0 of the city is constant and the R_0 of the village increases.

R_0 Ratio	City R_0	Village R_0	Fraction City Children	Fraction City Adults	Relative Infectivity (Children vs. Adults)	Cumulative Incidence (CI)
2.74	1.89	0.69	0.92	0.93	1	0.04
2.39	1.89	0.79	0.87	0.92	2	0.04
2.12	1.89	0.89	0.82	0.92	3	0.04
1.91	1.89	0.99	0.77	0.92	4	0.04
1.73	1.89	1.09	0.72	0.92	5	0.04
1.59	1.89	1.19	0.68	0.92	6	0.04
1.47	1.89	1.29	0.65	0.92	7	0.04
1.36	1.89	1.39	0.61	0.92	8	0.04
1.27	1.89	1.49	0.58	0.92	9	0.04
1.19	1.89	1.59	0.55	0.91	10	0.04

For all three scenarios, the majority of cases originated in the city, regardless of the local transmission parameters. Even when the village β term for children was over 3.5 times that of children in the city, the majority of village child cases still originated in the city. We therefore concluded that the major reason for the influence of the city was its population size.

References

[1] Knipl, D., 2016 A new approach for designing disease intervention strategies in metapopulation models. *J Biol Dyn* **10**, 71–94.

eTable 8: Sensitivity analysis for scenario 2. The R_0 and the village and the city are both increased but their ratio is kept constant

R_0 Ratio	City R_0	Village R_0	Fraction City Children	Fraction City Adults	Relative Infectivity (Children vs. Adults)	Cumulative Incidence (CI)
2.39	1.65	0.69	0.88	0.92	1	0.03
2.39	1.89	0.79	0.87	0.92	2	0.04
2.39	2.13	0.89	0.86	0.92	3	0.04
2.39	2.37	0.99	0.85	0.91	4	0.05
2.39	2.61	1.09	0.84	0.91	5	0.05
2.39	2.85	1.19	0.83	0.91	6	0.05
2.39	3.08	1.29	0.83	0.90	7	0.05
2.39	3.32	1.39	0.81	0.89	8	0.06
2.39	3.56	1.49	0.80	0.89	9	0.06
2.39	3.80	1.59	0.79	0.89	10	0.06

eTable 9: Sensitivity analysis for scenario 3. The transmission parameters for the village and the city are equal but the value of R_0 is increased.

R_0 Ratio	City/Village R_0	Fraction City Children	Fraction City Adults	Relative Infectivity (Children vs. Adults)	Cumulative Incidence (CI)
1.00	0.69	0.96	0.97	1	<0.01
1.00	0.79	0.95	0.97	2	<0.01
1.00	0.89	0.95	0.97	3	<0.01
1.00	0.99	0.94	0.97	4	<0.01
1.00	1.09	0.94	0.96	5	<0.01
1.00	1.19	0.94	0.96	6	<0.01
1.00	1.29	0.93	0.96	7	0.01
1.00	1.39	0.93	0.96	8	0.01
1.00	1.49	0.93	0.96	9	0.01
1.00	1.59	0.92	0.95	10	0.01

[2] de Blasio, B. F., Kasymbekova, K. & Flem, E., 2010 Dynamic model of rotavirus transmission and the impact of rotavirus vaccination in kyrgyzstan. *Vaccine* **28**, 7923–32.

[3] Dietz, K., 1993 The estimation of the basic reproductive number for infectious diseases. *Statistical Methods in Medical Research* **2**, 23–41.

[4] Eisenberg, J. N. S., Cevallos, W., Ponce, K., Levy, K., Bates, S. J., Scott, J. C., Hubbard, A., Vieira, N., Endara, P., Espinel, M. *et al.*, 2006 Environmental change and infectious disease: how new roads affect the transmission of diarrheal pathogens in rural ecuador. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 19460–5.