Supplementary Materials:

Home-based HIV Counseling and Testing Dataset

Data from six districts in western Kenya, Bungoma East, Burnt Forest, Kapsaret, Bunyala, Chulaimbo, and Teso North, are part of the catchment area of the Academic Model Providing Access to Healthcare. These districts vary by population density, economic activities, access to health services (including a district hospital and between 7 to 11 peripheral health facilities within each district), and disease burden. Kapsaret and Burnt Forest are more urban. Bunyala borders Lake Victoria and in this district, along with Chulaimbo, represent districts where disease burden is high. Teso North and Bungoma East fall somewhere between these extremes.

Geographic data: Travel times to the nearest large settlement and health facility

We calculated the distance to the nearest health facility using Euclidean distance in ArcGIS. We also calculated the expected travel time maps for health facilities, large settlements (population of 10,000 or more) and schools, generated through standard cost-distance-based spatial analysis methods (see Methods).

In the majority of households with children, these children were immunized (number of households missing immunization: 686, 2% of eligible households). Almost 40% of households with pregnant women were missing antental care care (39%, N=1,374).

Table S1: Basic summary statistics about the number of households missing immunizations and antenatal care for the full data set and each catchment.

	Full Data Set	Burnt Forest	Chulaimbo	Kapsaret	Port Victoria	Teso
Number of Sublocations	94	18	16	6	18	35
Number of Households	78635	20019	15516	15276	17308	10516
Number of Households w. Pregnant Women	3533	574	731	692	1027	509
Number of Households w. Children	40088	4947	10035	8181	9989	6936
Number of HH Missing Immunization	686	124	144	179	149	90
Number of Househoulds Missing Antenatal Care	1374	286	233	254	437	164
Population Density	497.5	182.1	497.9	535.2	441.7	688.3

per Sublocation	(100.2, 1506.5)	(57, 324)	(283.2, 838.2)	(72.3, 1710.4)	(141.7, 809.1)	(239.3, 2797.5)
Percentage of Households Missing Immunization	1 (0, 3)	2 (0, 7)	2 (0.7, 2)	2 (2, 4)	1 (0, 2)	1 (0, 3)
Percentage of Households Missing Antental Care	31 (0, 55)	38 (0, 65)	31 (16, 40)	37 (30, 46)	39 (28, 51)	22 (0, 51)

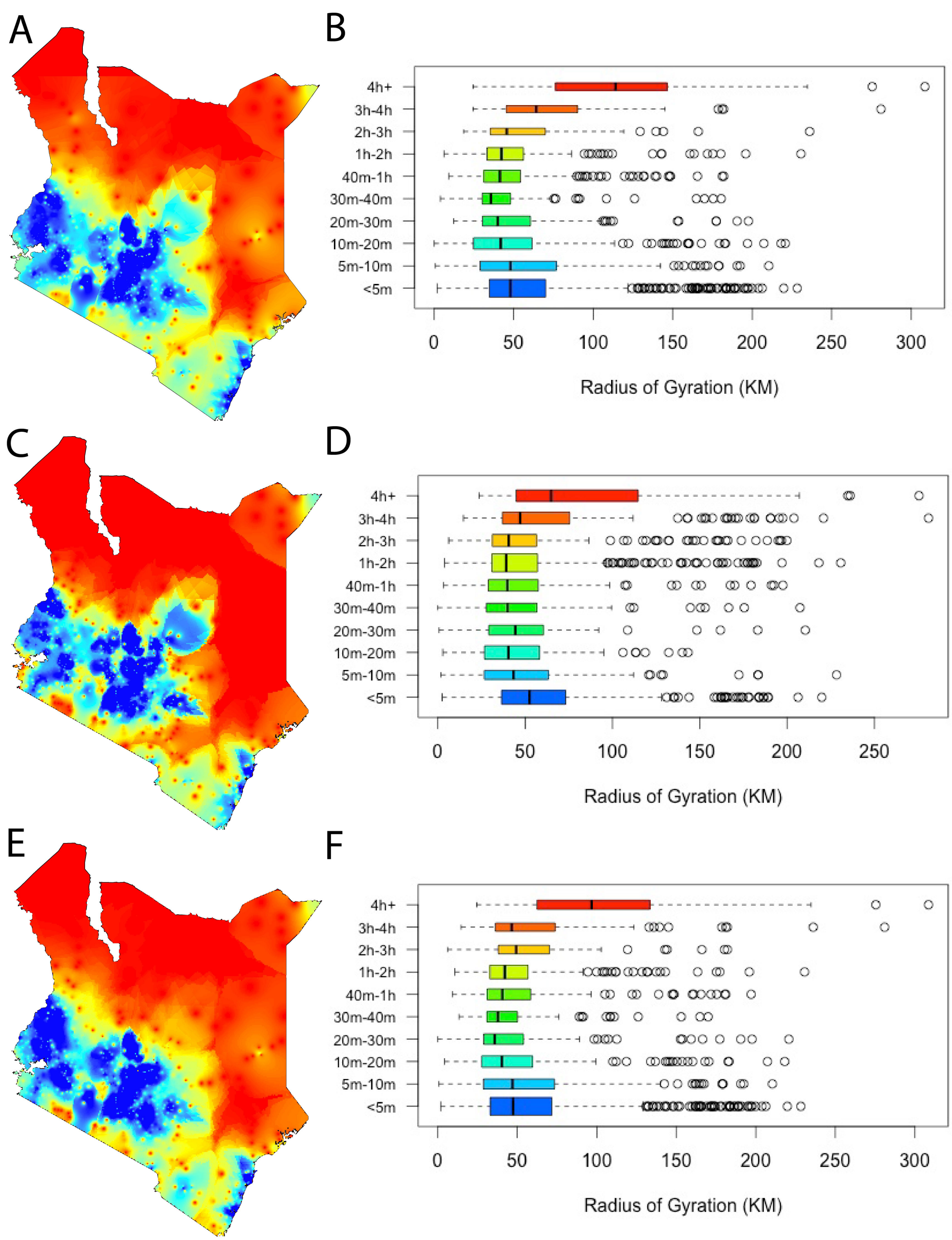

Figure S1: The relationship between travel times and radius of gyration. The travel time to the nearest health facilities A), large settlements with a population of at least 10,000 individuals C), and school E) per mobile phone tower are shown. Mobile phone towers near health facilities, large settlements, and schools are shown in blue and more remote locations shown in red. All three variables are strongly correlated with one another (correlation coefficient: health facility and large settlement 0.92, health facility and school 0.96, large settlement and school 0.95, $p<0.001$ for all comparisons). The relationship between travel time and radius of gyration are shown for health facilities B), large settlements D), and schools F). In general, individuals from the most and least remote mobile phone towers traveled the most (mean radius of gyration for towers at least 4 hours from the nearest health facility = 116, large settlement = 83, and school = 103) (mean radius of gyration for towers within 5 minutes from the nearest health facility = 63, large settlement = 68, and school = 64).

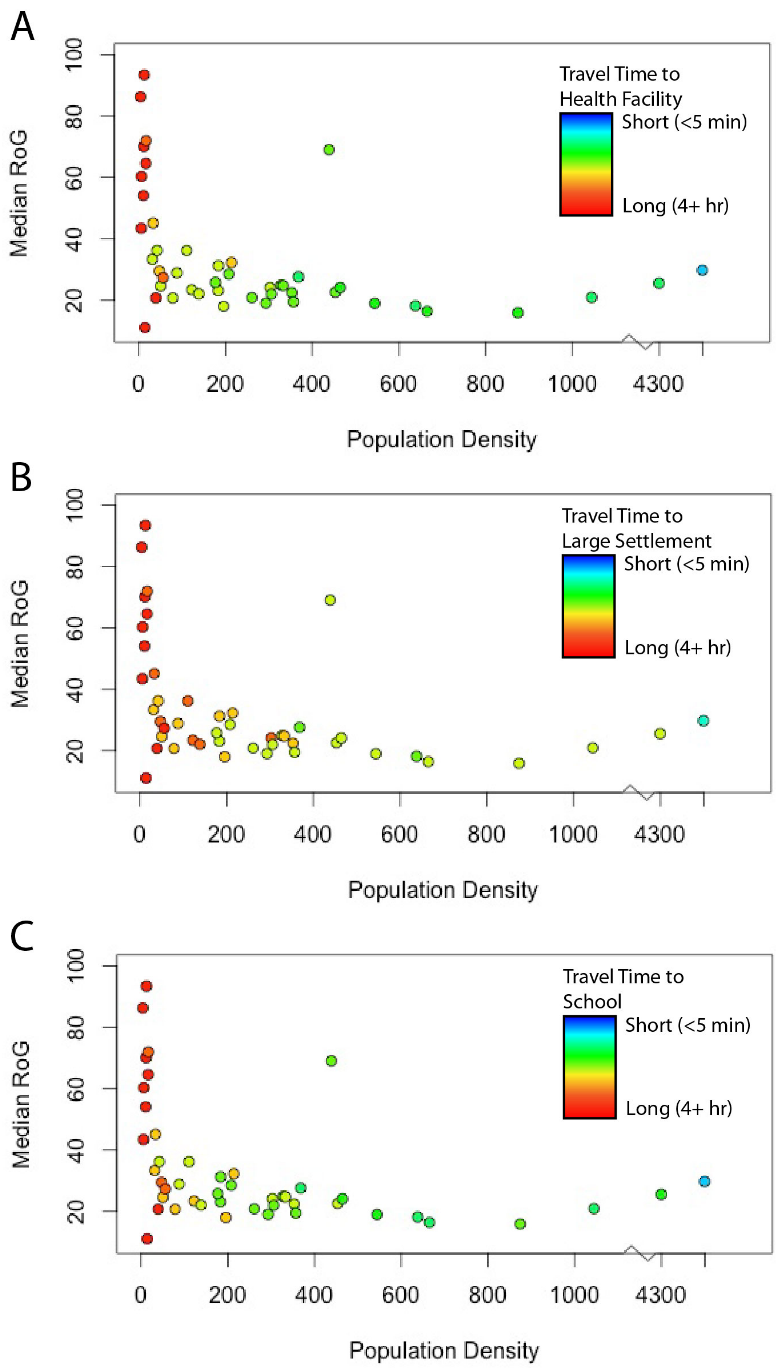

Figure S2: The relationship between population density and radius of gyration

(KM) for counties. For each county, the relationship between population density and radius of gyration is shown. Counties are colored according to the average travel time to the nearest A) health facility, B) large settlement, and C) school. Red counties have long travel times to these locations whereas blue counties have shorter travel times.

The impact of tower density on radius of gyration

Areas where the density of mobile phone towers is high can inherently measure smaller movement patterns than areas where the towers are sparser. To determine the effect of heterogeneous tower density across Kenya, we simulated 1,000 data sets each comprised of 100 synthetic mobility patterns.

We simulated the effect of the number of mobile phone towers within a set area on radius of gyration estimates using the following algorithm:

1. Generate synthetic movement data within a 100 km^2 area.
2. Randomly choose *n* tower locations such that the tower in in the 100 km^2 area such that the coverage area of each tower does not overlap any other tower's coverage area. The assumed the coverage area of each tower was approximately 3 km.
3. For each location in the movement data, determine the location of the closest tower and if this tower was within range (within 3 km) of the movement location. If the location was not within 3 km of any tower, then the movement was not recorded in the call records.
4. Using the synthetic call records, calculate a center of mass and radius of gyration. This is the 'true' radius of gyration value.
5. For each number of possible towers within the area, determine the radius of gyration associated with the number of towers that is the 'measured' value.

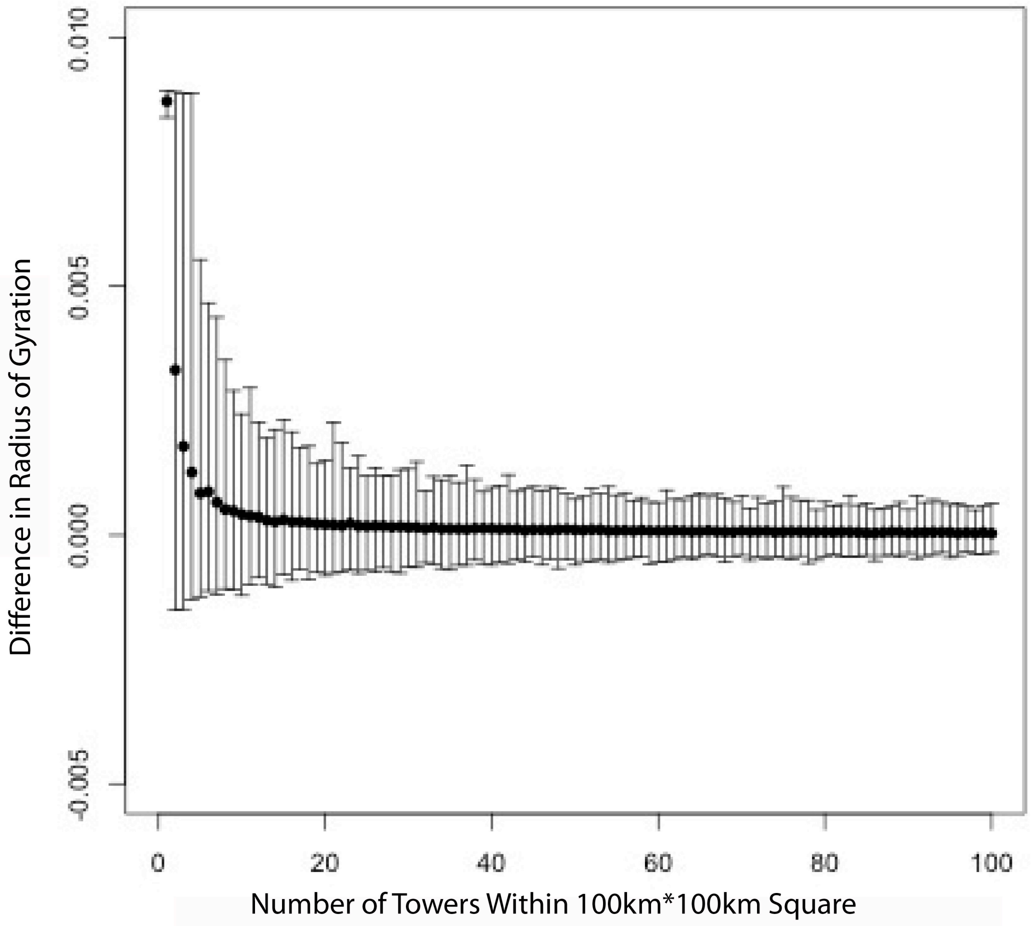

Figure S3: Results from the simulation increasing the number of mobile phone towers within a square geographic area. We varied the number of mobile phone towers within a 100 km^2 area. We generated movement patterns within this area by randomly choosing coordinates within the space. We then randomly chose n (from 1 – 100) tower locations within the space such that the coverage area of each tower did not over lap any other tower's coverage area. For each movement in the synthetic movement data, if this movement was within a tower's coverage area, then this was added to the synthetic call data. Otherwise, the movement was not recorded in the synthetic call data. Using the synthetic call data, the radius of gyration value was calculated and compared to the true radius of gyration value from the synthetic movement data. The accuracy to quantify the movement patterns of an individual subscriber greatly increases as the number of mobile phone towers increases.

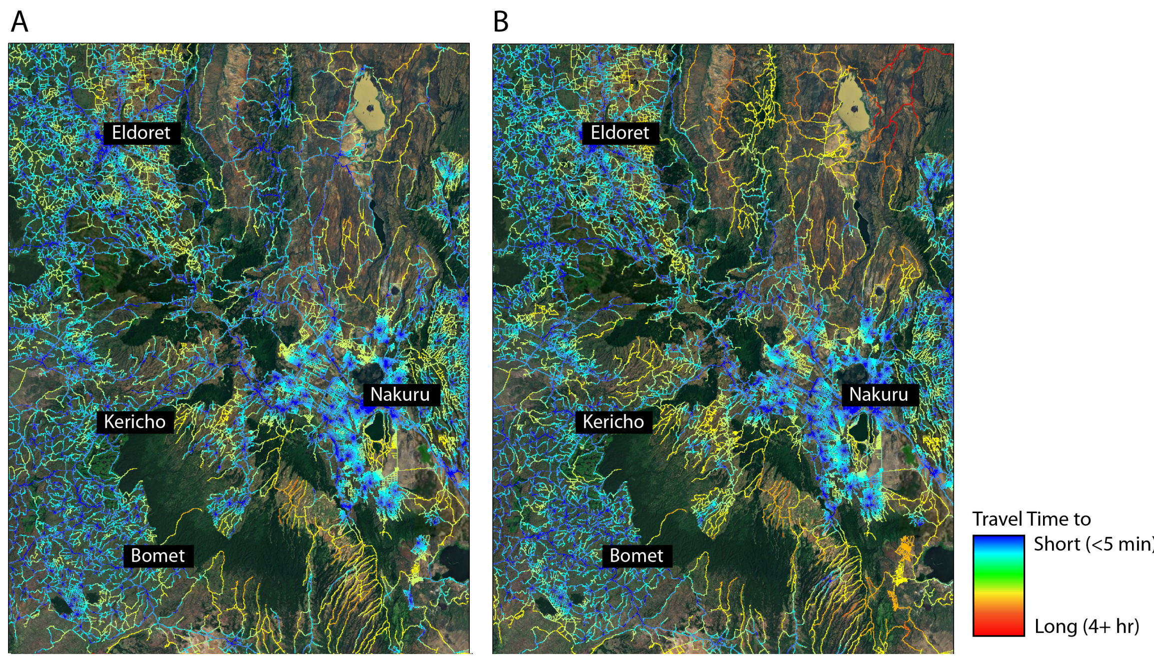

Figure S4: The road networks of Western Kenya. The road networks in a region of Western Kenya are shown colored by the travel time to the nearest A) health facility and B) school. The area shown in each map is outlined in Figure 1A.

Regression Models

We constructed four linear regression models to predict the percentage of eligible households within a sublocation missing either immunizations or antenatal care using either the travel time to the nearest health facility or mobility values for the sublocation as predictors.

Table S2: The results from each regression model predicting either the percentage of eligible households missing immunizations or antenatal care.

Immunization Model	**Adjusted R2**	**F-statistic**	**DF**	**p-value**	**Reduction in Deviance (%)**
Predictor: Travel Time	0.095	9.8	82	0.002	0.41
Predictor: Mobility	-0.0086	0.28	82	0.6	3.3
Antenatal Care Model					
Predictor: Travel Time	-0.0079	0.34	82	0.6	0.34
Predictor: Mobility	0.022	2.8	82	0.10	11