

Supplementary Information

Effect of Modeling Slum Populations on Influenza Spread in Delhi

Survey Data

A local vendor “Indiamart.com” conducted the slum survey in Delhi. Field workers on the ground interviewed a total of 992 households in slum areas. The non-slum survey was conducted by “MayMyIndia.com” on the ground and covered a total of 2,366 households in Delhi. The surveys collected responses on questions related to demographics and daily sets of activities. The respondents were asked to report the activities performed every hour for a period of 24 hours on a typical day. The activities were categorized into travel, home, work, school, college, shop and other. Category “Other” included all activities that were not included in the other six categories. Age distributions of respondents are given in Figure S1.

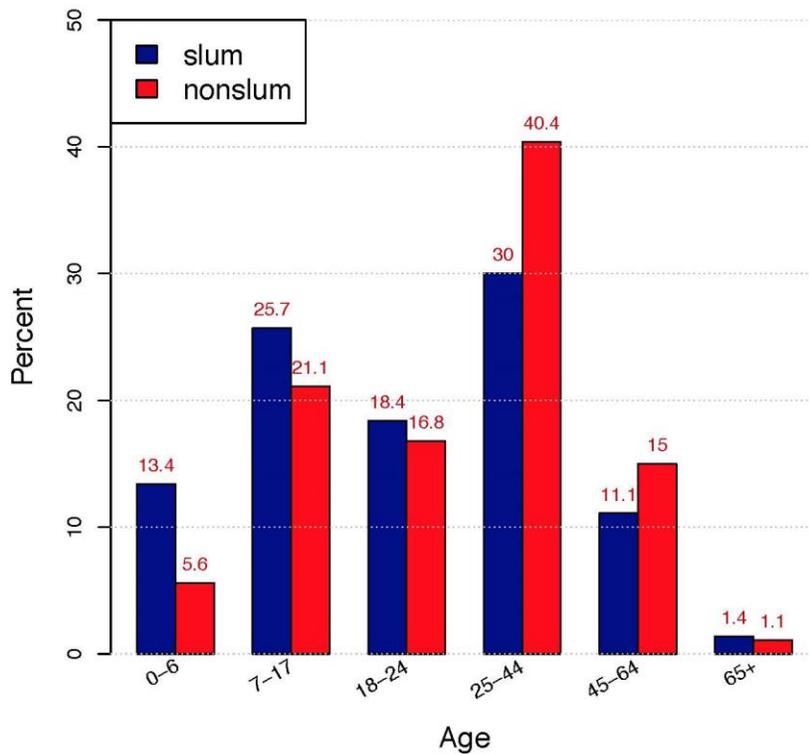


Figure S1: Fractions of total survey respondents in each age group. The fractions total 100% for each of slum and non-slum.

Population Generation Methodology

The synthetic population is created by integrating a variety of data sets from commercial and public sources into a common architecture for data exchange. The process preserves the confidentiality of the individuals in the original data sets, yet produces realistic attributes and demographics for each synthetic individual. The synthetic population is a set of synthetic people and households located geographically, each associated with demographic variables recorded in the census. Joint demographic distributions are reconstructed from the marginal distributions available in typical census data using an iterative proportional fitting (IPF) technique. Each household is located geographically using land use data and data pertaining to transportation networks.

In the next step, a set of activity templates for households are determined, based on survey responses to an activity or time-use survey. Each synthetic household is then matched with one of the survey households, using a decision tree based on demographics such as the number of people in the household, number of children of various ages, income, etc. The synthetic household is assigned the activity template of its matching survey household. For each household and each activity performed by this household, an assignment of a location is made based on observed land-use patterns, transportation network etc. A social network is formed based on the co-location of agents at various times throughout the day, as described in the main text. The network is dynamic and changes as people visit different locations and come in contact with individuals at these locations. Additional discussion may be found in [1],[2].

Network and Population Data

Data in this section are confined to Network 2, the social contact network that includes Delhi slum individuals, their attributes, and activities.

Figure S2 contains a histogram of slum zone population sizes. There are 298 slums, ranging in population sizes from about 10 to 46000. Bins in the plot are in 5000 person increments.

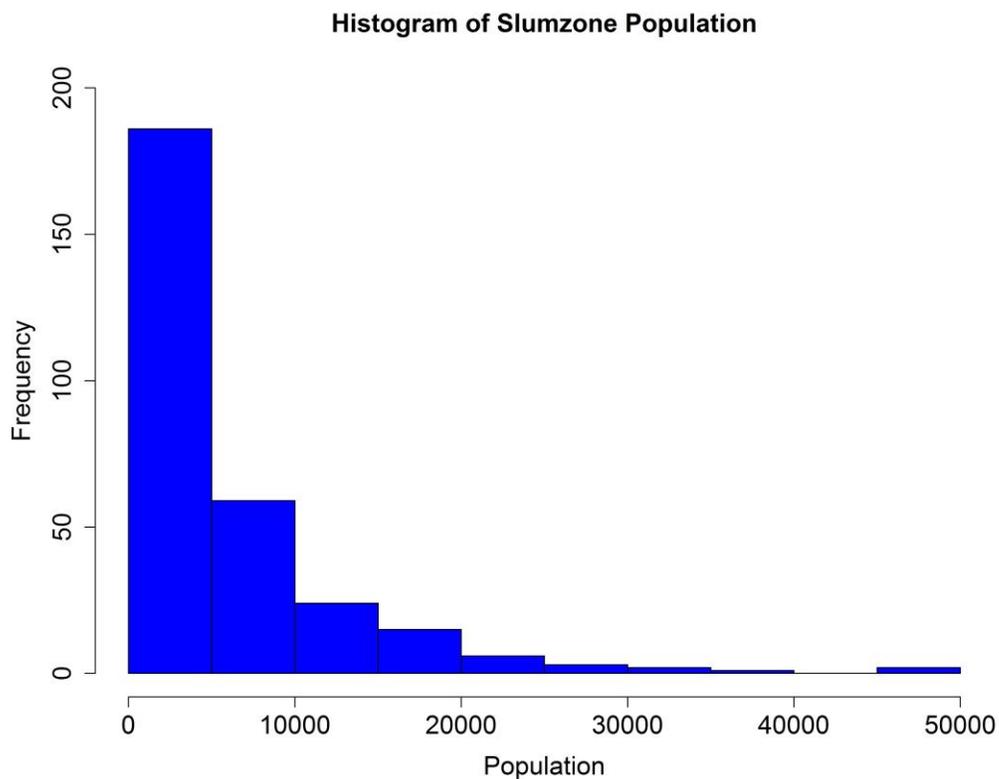


Figure S2: Histogram of populations in 298 slum zones in Network 2. Slum zone populations are binned in 5000-person size increments. Only two slum zones have population sizes larger than 40,000 people.

Household size distributions for slum and non-slum households are provided in Figure S3. The slum curve is based on aggregating the households across all 298 slum zones. Nonslum refers to the non-slum regions of Delhi.

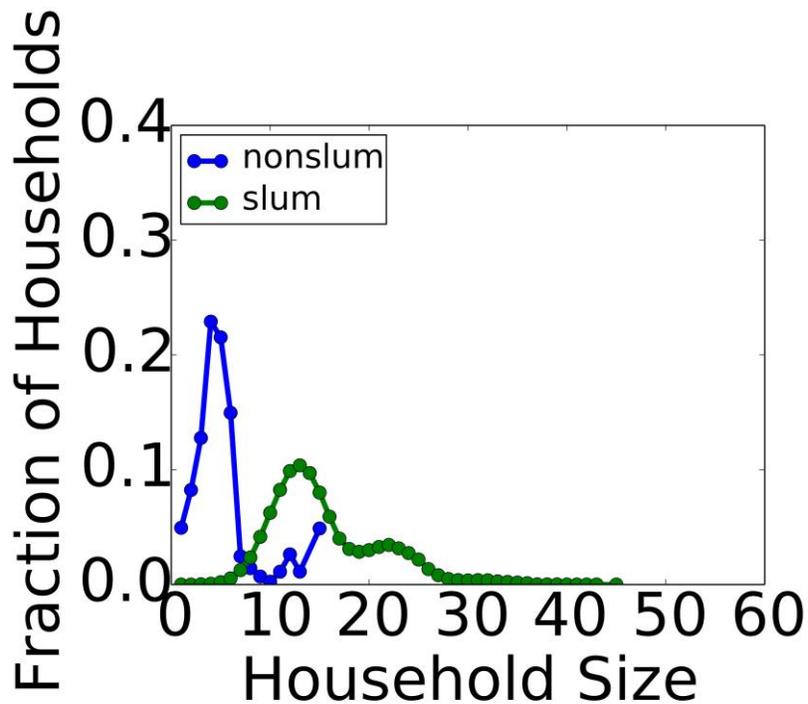


Figure S3: Household size distributions for slum and non-slum regions. Slum data corresponds to aggregated households across all slum zones.

Figure S4 provides the number of edges in Network 2 for the six activity types. The All category is a sum of the edges in the six categories. For each type, the four curves correspond to both vertices of an edge representing people residing in slums (slum), both vertices of an edge representing people residing outside of slums (non-slum), a slum and non-slum vertex forming an edge (mix), and total number of edges (Delhi).

Note that for most activity types (e.g., work, shopping) in Figure S4, the number of edges for slums is a small fraction of those for non-slum. This is because about 87% of vertices (individuals) in Network 2 are non-slum, so these edges should dominate. However, for home activities, the ratio of slum:non-slum is about 1:2. This ratio is greater because of the larger number of people in slum households, so the number of edges in slum households is much greater than that in non-slum households. Every household forms a clique on n vertices, where n is the household size. The other data that stand out are the “other” activities; there are many more edges of this type in

slums than non-slums, even with fewer people. Thus, we conclude that slum residents have more varied activities than non-slum residents.

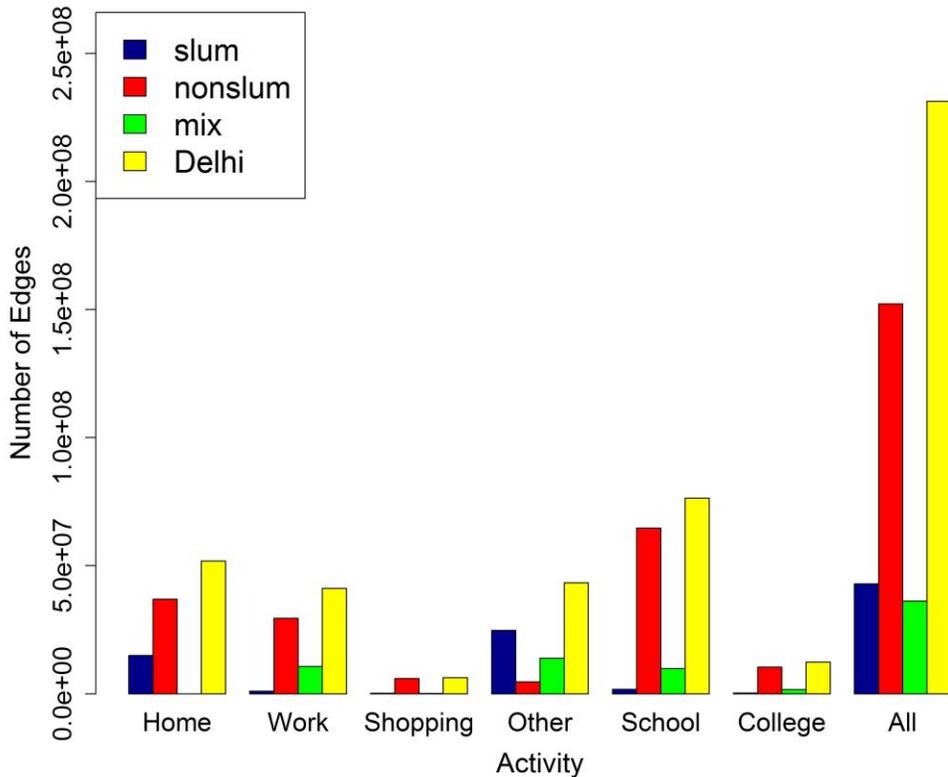


Figure S4: Number of edges for slum, non-slum, mix, and the entire Delhi network for the six activity categories. Slum edges are those where both incident vertices (individuals) are slum dwellers; non-slum edges are those where both incident vertices (individuals) are non-slum dwellers; mix denotes edges having one node as a slum resident and the other node as a non-slum dweller; and Delhi denotes total edges.

Figure S5 contains data for the six activity types, broken down into the same four types of edges as in Figure S4, but now the data are the average duration of each edge (i.e., interaction) in Network 2. The differences between slum and non-slum, across activity types, are not great, except for perhaps type “other”. Also, except for type “work”, the non-slum edges have slightly greater durations on average. The mixed edges are

generally of least duration. The ``all'' category is referred to in the main text: slum and non-slum durations are 2.5 and 3.2 times that of the mixed edge duration.

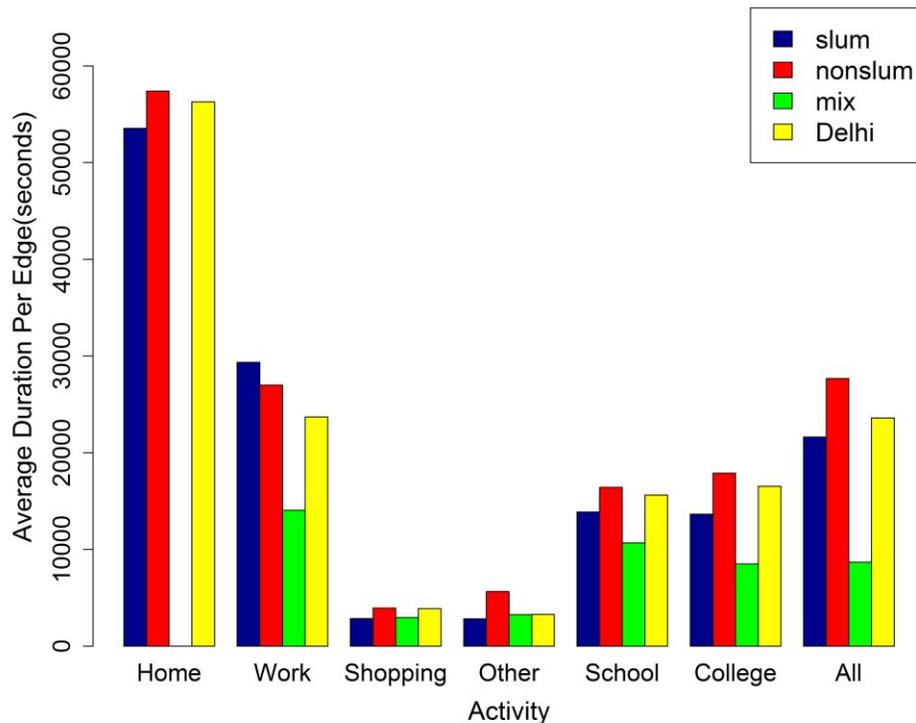


Figure S5: Average duration per edge (i.e., contact) for slum, non-slum, mix, and the entire Delhi network for the six activity categories.

Figure S6 provides average total interaction duration per person per day in Network 2. This differs from the data in Figure S5, as follows. Suppose there is a total of 500 work seconds of interaction for people in slums based on social contact Network 2. Suppose there are 20 work edges. Then the average duration per edge is 25 seconds ($=500/20$); this kind of data is plotted in Figure S5. Now, suppose 10 people generate the 20 work edges. Then, the average duration per person is 50 seconds ($=500/10$); this type of data is plotted in Figure S6. Note that in Figure S6, the average duration per person per day of interaction can be greater than the number of seconds in a day (86,400). This is because a person can interact with multiple people simultaneously, as when a family sits down to dinner. If a family has four members, then a one-hour dinner results in 3-

person-hours of interaction duration for each person at the table. In Figure S6, for home and other activity types, the average duration per person is greater for slum dwellers, and these provide the dominant contributions to the overall results (last category). Note that the averages are close to the non-slum data, since 87% of people are non-slum residents.

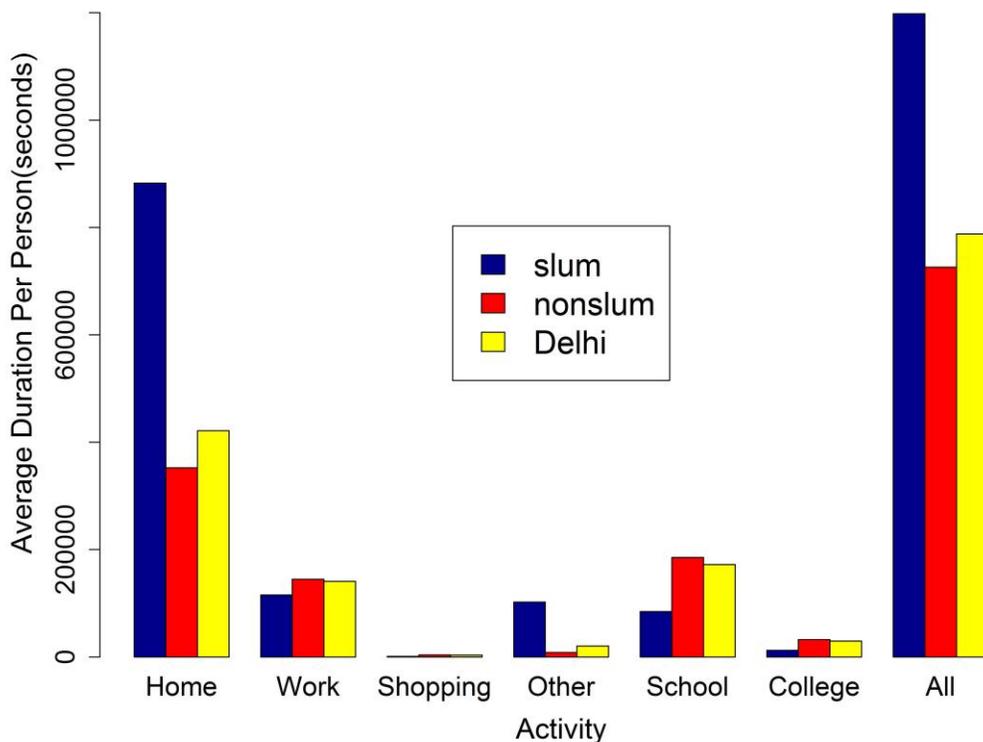


Figure S6: Average total interaction duration per person per day for slum, non-slum, and all Delhi individuals for the six activity categories.

Figure S7 contains, for Network 2, the average degrees, by age, of four groups of people: all combinations of male or female and those who do and do not live in slums. Average degree is synonymous with “average number of interactions.” The slum dwellers have much higher degrees across almost all ages than do those that do not live in slums. There are several trends. First, for slum people, the average degree of women is greater than that of men over adulthood: roughly the years 20 through 60. Outside this range, for the young and older adults, the comparison is mixed. For non-

slum people, again at younger and older ages, the average degrees of men and women are comparable. However, during the more active adult years (20 through 60), now men have greater degrees than woman, although not by as wide a margin as in the slum comparison.

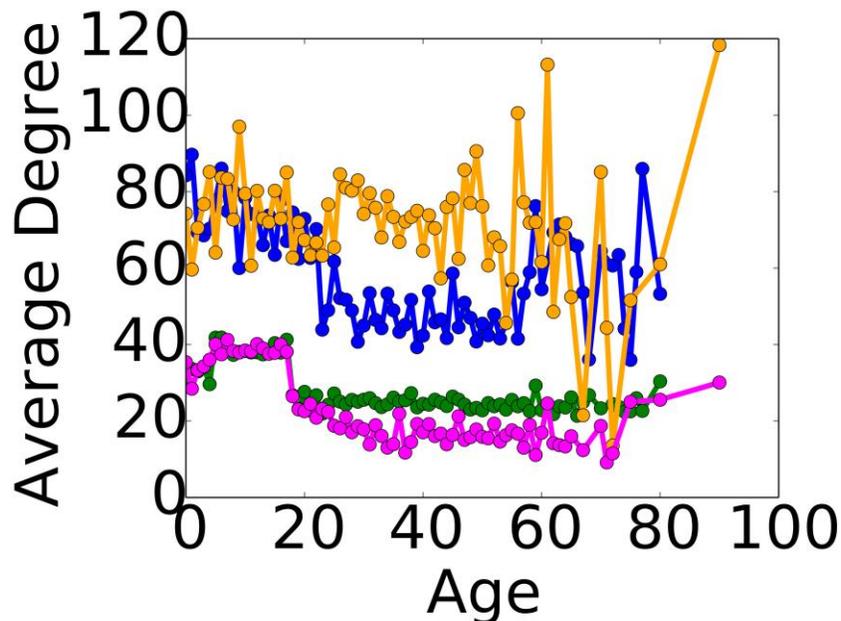


Figure S7: Average degree, or average number of interactions, for males and females that live in and outside of slums, as a function of age. The legend is: (male, slum) in blue; (male, non-slum) in green; (female, slum) in orange; and (female, non-slum) in magenta. Slum dwellers have much more interaction than do non-slum dwellers. For the most active adult years, roughly 20 through 60 years, women in slums have more interactions than men on average, while men outside of slums have slightly more interactions on average than women.

Figure S8 depicts the average number of contacts for male and females, 18 years of age and older, for home activities and for non-home activities in Network 2. Non-home activities in this figure are those in Figure S6 that are not “home” activities. Interactions of males and females are computed for all of their contacts, regardless of age. Hence, if a 5-year old female child interacts with her mother who is 33 years, then the mother’s contacts include that with the 5-year old and the mother is represented in the data. The

5-year old, however, is not represented in the populations that form the bars because the child is younger than 18. The data show that on average, a female adult is more connected to others than are males. In this figure, only people who have contacts at home or non-home are included in the averages. Thus, for example, a person living alone will have no home contacts and is not included in the “home activities” bars, but may have work activities and if so is included in the “non-home activities” bars.

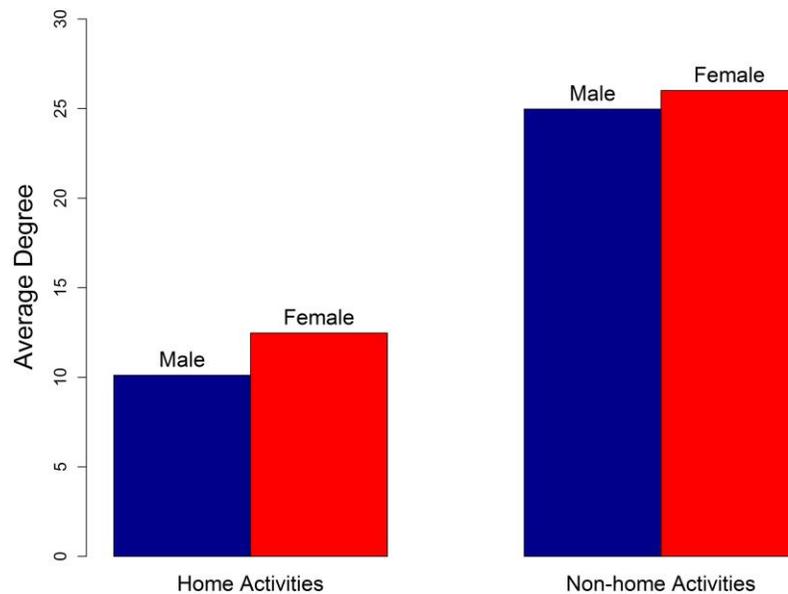


Figure S8: Average degrees (i.e., average number of interactions per person) for males and females 18 years and older, separated by human interactions that occur during a person’s home activities and non-home activities.

Figure S9 shows degree distributions for men (i.e., males ≥ 18 years) and women (i.e., females ≥ 18 years) that are connected to children (i.e., < 18 years) of either gender in Network 2. Hence, these data are related to those of Figure S8, except that now interactions of adults are confined to those with children. Women are better connected to children than are men.

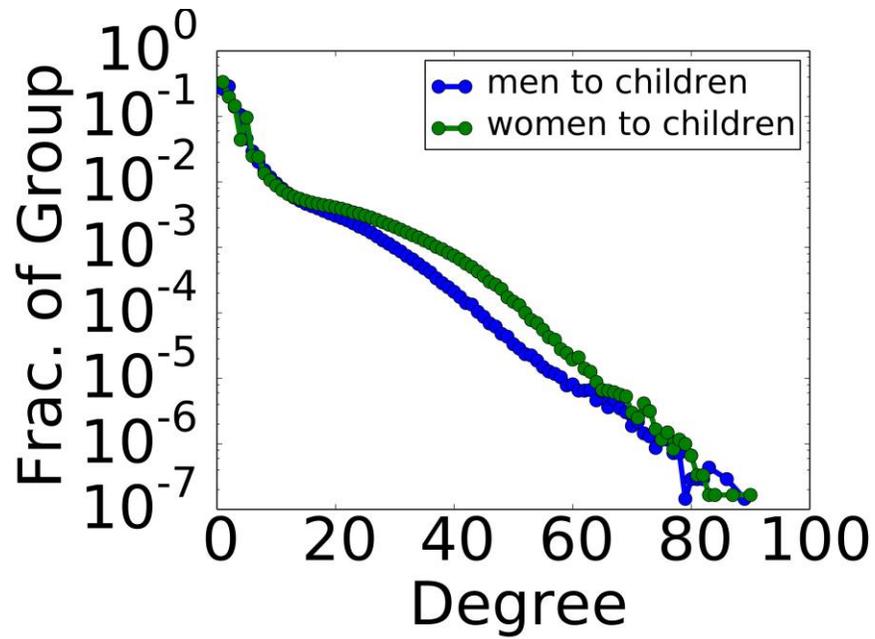


Figure S9: Degree distributions for men and women (i.e., those of age ≥ 18 years) that are connected to children (i.e. under age 18) for all activities.

Epidemic Simulation Parameters

□

Table S1 contains transmissibility parameters, R_0 , and attack rates for Network 1 (that does not include slum-specific data) and Network 2 (that contains slum-specific data).

TABLE S1: Simulation parameters and cumulative attack rates.

Area	Delhi		
Network	Network 1, Network 2		
Transmissibility	0.0000215(Mild)	0.000027(Strong)	0.00003(Catastrophic)
R_0 (Network 1)	1.015	1.266	1.40
R_0 (Network 2)	1.123	1.398	1.546
Attack Rate (Network1)	16.82%	41.70%	49.90%
Attack Rate (Network2)	29.77%	48.43%	55.50%
Simulation Day	400		
Initial Seeding	20 persons first day		

□

REFERENCES

- 1 Bisset K and Marathe M. [A cyber-environment to support pandemic planning and response](#). DOE SciDAC Magazine 2009; 13: 36-47.
- 2 Barrett C, Bisset K, Leidig J, Marathe A, Marathe M. [Economic and social impact of influenza mitigation strategies by demographic class](#). *Epidemics* 2011;3 (1): 19-31.