Predictability of population displacement after the 2010 Haiti earthquake

Xin Lu[a,b,1,2], Linus Bengtsson[a,1,2], and Petter Holme[a,b,c,d]

[a]Department of Public Health Sciences, Karolinska Institutet, 17177 Stockholm, Sweden; [b]Department of Sociology, Stockholm University, 10691 Stockholm, Sweden; [c]Department of Physics, Umeå University, 90187 Umeå, Sweden; and [d]Department of Energy Science, Sungkyunkwan University, Suwon 440-746, Korea

Most severe disasters cause large population movements. These movements make it difficult for relief organizations to efficiently reach people in need. Understanding and predicting the locations of affected people during disasters is key to effective humanitarian relief operations and to long-term societal reconstruction. We collaborated with the largest mobile phone operator in Haiti (Digicel) and analyzed the movements of 1.9 million mobile phone users during the period from 42 d before, to 341 d after the devastating Haiti earthquake of January 12, 2010. Nineteen days after the earthquake, population movements had caused the population of the capital Port-au-Prince to decrease by an estimated 23%. Both the travel distances and size of people's movement trajectories grew after the earthquake. These findings, in combination with the disorder that was present after the disaster, suggest that people's movements would have become less predictable. Instead, the predictability of people's trajectories remained high and even increased slightly during the three-month period after the earthquake. Moreover, the destinations of people who left the capital during the first three weeks after the earthquake was highly correlated with their mobility patterns during normal times, and specifically with the locations in which people had significant social bonds. For the people who left Port-au-Prince, the duration of their stay outside the city, as well as the time for their return, all followed a skewed, fat-tailed distribution. The findings suggest that population movements during disasters may be significantly more predictable than previously thought.

trajectory | human mobility | disaster informatics | disaster relief

In 2010, natural disasters displaced 42 million people, directly affected an estimated 217 million people, and resulted in USD 120 billion worth of damage (1, 2). The humanitarian response to natural disasters relies critically on data on the geographic distribution of affected people (3). During the early response phase, data on population distributions are fundamental to the delivery of water, food, and shelter, and to the creation of sampling frames for needs assessment surveys (4). During later stage reconstruction efforts, population distribution data is required for the allocation of schooling resources, delivery of seeds, construction of houses, and the like (5, 6).

Despite a number of studies on human mobility patterns during small-scale, short-term emergencies such as crowd panics (7, 8) and fires (9, 10), research on the dynamics of population mobility during large-scale disasters such as earthquakes, tsunamis, floods, and hurricanes has been limited (11). Existing research on population movements after large-scale disasters has been hampered by difficulties in collecting representative longitudinal data in places where infrastructure and social order have collapsed (12, 13), and where study populations are moving across vast geographical areas (14). Existing research has found that people displaced by natural disasters typically stay within their country of residence, that sudden-onset disasters often lead to more short-term displacement than do slow-onset disasters (15), and that postdisaster reconstruction programs in the long

run can cause populations to move into disaster-affected areas rather than moving away from them (11).

The increased use of mobile phones, even in low- and middle-income countries (16), offers a new way to circumvent methodological problems of earlier research. Data from mobile phones have the advantage of high resolution in time and space, being instantaneously available with no interview bias, and they provide longitudinal data for very large numbers of persons (12, 17–23). Even more importantly, cellphone data allows for statistics based on trajectories of individuals. This means one can, as we will in this paper, study how the disaster affects people's daily behavior and routines.

Pioneering work using mobile phone data to describe human mobility patterns has been carried out during stable social conditions (17–19). One major conclusion from these studies is that, despite a broad distribution of average travel distances among people, the movements of individuals are surprisingly predictable (17). In this paper, we study mobile phone data from Haiti collected before and after the tragic Haiti earthquake on Tuesday, January 12, 2010, which left an estimated 1.8 million people homeless and killed between 65,000 and 300,000 persons (24, 25).

We collaborated with the largest mobile phone operator in Haiti, Digicel, to analyze the positions of 2.9 million anonymous subscribers during the period from 42 d before the earthquake to 341 d after (December 1, 2009, to December 19, 2010). Specifically, we obtained the locations of all anonymous Digicel mobile phone users at the time of their first call each day. To exclude relief workers entering Haiti after the earthquake and people who died or whose SIM cards stopped functioning, we excluded people who did not call at least once before the earthquake and at least once during the last month of the study period. After this filtering, we obtained 1.9 million individuals across Haiti [10 million inhabitants (26)] out of which 0.8 million were located within Port-au-Prince [2.6 million inhabitants (26)] on the day of the earthquake. We assume in this paper that the mobile phone movements were representative of the general population movements. Although this issue requires additional research, we showed in a separate paper (13), using the same data source, that mobile phone movements after the Haiti earthquake corresponded well with comparable movement data from a large retrospective household survey of 12,250 persons, performed by UNFPA eight months after the earthquake. The spatial resolution of people's locations is that of the coverage areas of the mobile phone towers in the network (ranging from less than 100 m in urban areas to a few tens of kilometers in the hinterland).

In order to understand fundamental changes in mobility patterns after a large-scale disaster, we analyze how the earthquake changed the aggregate mobility of individuals in the severely hit capital Port-au-Prince (PaP), to what extent the chaotic conditions after the earthquake influenced the disorder and predictability of the population's movements, and the dynamics of the population flows out of and back into PaP. We address both the larger-scale prediction of population displacements and the predictability of the trajectories of individuals. Surprisingly, we find that despite large changes in the population distribution across the country, the mobility of the PaP population contained several highly regular features, and most individuals' movements remained highly predictable.

Results

Daily Travel Distances and Population Flows. To get an overview of the aggregate travel patterns before and after the earthquake, we show (Fig. 1B) the observed distribution of travel distances over the sampling period. One day after the earthquake (January 13, 2010), 6.5% of the observed individuals had traveled more than 20 km as compared to the preceding day, while the corresponding figure before the earthquake (December 1–2, 2009) was 3.5%. The increase in average daily travel distances lasted for two to three weeks after the earthquake. It is worth noting that other periods also saw sudden increases in average daily travel distances. These periods coincided with Christmas and New Year from around December 20 to January 3—just before the earthquake—as well as the Easter holidays (early April).

The earthquake did not directly affect large parts of Haiti. In the rest of our analyses, we therefore focus on the population of the heavily affected capital region (PaP). As we show in Fig. 1C, the population movements after the earthquake on January 12, 2010, led to a rapid decrease in the PaP population. Nineteen days after the earthquake (January 31), the net population decrease was an estimated 23% compared to the stable level before Christmas (December 1–20, 2009), assuming the phone movements to be representative of the population movements. The net flow into PaP again became positive 20 d after the earthquake (February 1), and the PaP population increased approximately linearly over the following three months (February 1 to April 30). After this time, the population increase gradually leveled off and stabilized at the end of the year, with two short deviations around All Saints Day (November 1) and the election day (November 28).

There was a similar but smaller population decrease in PaP during the preceding Christmas and New Year (Fig. 1C). As we saw in Fig. 1B, this was also a period of generally increased travel in Haiti. A similar but smaller decrease was also seen during Easter. The population decrease in PaP during holidays is likely explained by many people leaving the capital to spend time with family and friends outside PaP. It is interesting to observe that the PaP population at the time of the earthquake had not yet fully recovered after the Christmas and New Year holidays. Assuming that the people who left PaP over the holidays were all going to return in the absence of the disaster, approximately 70,000 persons (2.5% of the PaP population) managed for this reason to avoid being in PaP on the day of the earthquake.

There is a strong weekly regularity in the number of mobile phone users in PaP. Increased numbers of people are present in PaP during working days, with corresponding smaller numbers present during weekends (Fig. 1C). This pattern was restored as early as three weeks after the earthquake.

To get a detailed view of the daily travel distances, d, we plot for a few different dates the cumulative probability distributions of d for two groups of people: persons present and not present in PaP on the day of the earthquake. The distributions are basically the same for both groups before the earthquake as well as eight months after the earthquake, when social life had stabilized considerably. However, right after the disaster there is a striking deviation in the distribution of travel distances (Fig. 1D), which is not present for people located outside PaP on the day of the earthquake (Fig. 1E). We fitted the curves in panels D and E
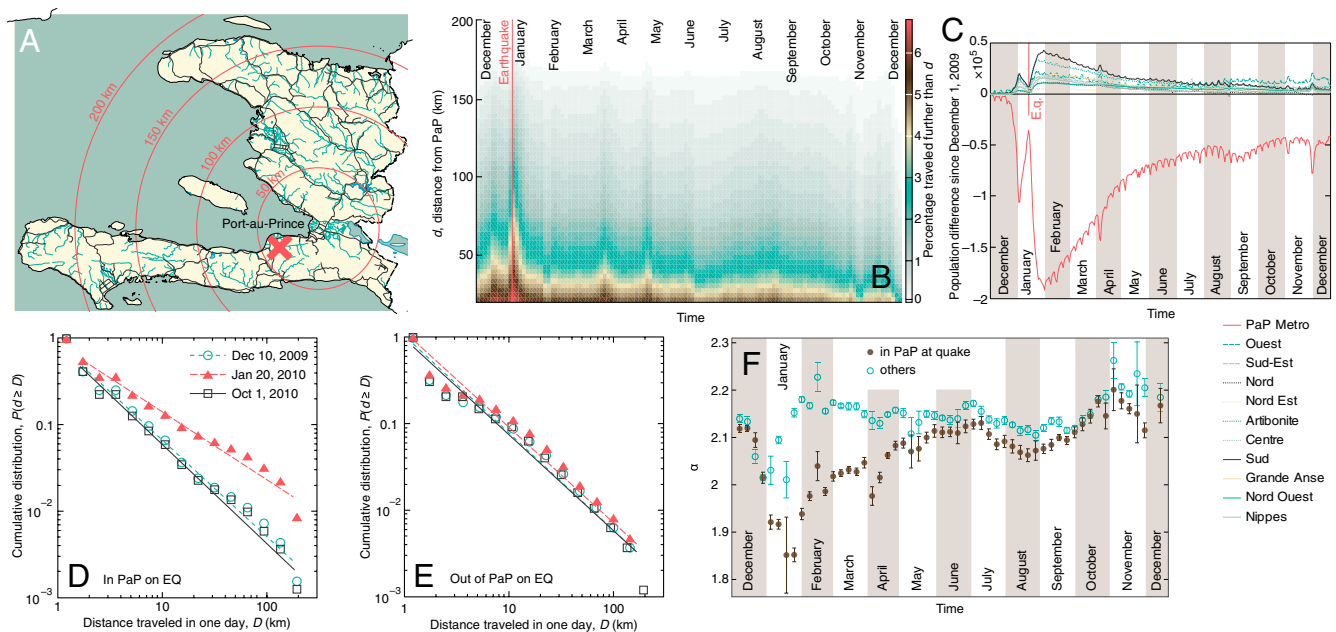
Fig. 1. Overview of population movements. (A) Shows the geography of Haiti, with distances from PaP marked. The epicenter of the earthquake is marked by a cross. (B) Gives the proportion of individuals who traveled more than d km between day $t − 1$ and t. Distances are calculated by comparing the person's current location with his or her latest observed location. In (C), we graph the change in the number of individuals in the various provinces in Haiti. (D) Gives a cumulative probability distribution of the daily travel distances d for people in PaP at the time of the earthquake. (E) Shows the cumulative probability distribution of d for people outside PaP at the time of the earthquake. Finally, (F) gives the exponent α of the power-law dependence of d—the probability of d is proportional to $d^{-\alpha}$. These are obtained by a maximum-likelihood method (33), and differ from the slopes of the lines in (D) and (E) by unity since these are the cumulative distributions.

(or rather the corresponding probability mass functions) to a power law $p(d) \sim d^{-\alpha}$. The coefficient α indicates the slope of the distribution. The smaller α is, the more fat-tailed is the distribution. In other words, a small α means larger differences between the average and the largest travel distances. In Fig. 1F, we show the α-values of the daily travel distances. We see that the α-values dip around the holidays for both the people present and not present in PaP on the day of the earthquake. However, while α goes back to normal in mid-January for the population outside PaP, it stays low until the early summer for the Port-au-Princians.

Movement patterns were heterogeneous both under normal conditions (echoing results of refs. 17 and 19), and after this large-scale disaster. In the latter case, however, the heterogeneity is even more pronounced than during normal times. Taken together, the analyses in Fig. 1, consistently shows that the movement patterns changed primarily for the directly affected people and returned back to normal after 4 ∼ 5 mon.

Regularity and Predictability of People's Trajectories. The above analysis reveals that the earthquake triggered large population movements, caused increases in daily travel distances of people in the affected area, and produced an increase in the heterogeneity of travel distances. Since a large-scale disaster throws much of society into disorder, one may assume that the underlying cause for these changes is an increase in the disorder in people's movements—that a large-scale disaster causes people to move irregularly while fleeing unrest and searching for material support. In this section, we will show that this hypothesis is incorrect.

To analyze the changes in people's trajectories caused by the earthquake, we divided the sampling period into three equally long periods: *spring*—from January 12 (the day of the earthquake) to May 5; *summer*—from May 6 to August 27; and *fall*—from August 28 to December 19. To get good statistics on the trajectories, we restrict ourselves to those who called 70% or more of the days during all three periods. To characterize the travel patterns, we follow ref. 17 and focus on two classes of measures—radius of gyration (a measure of the size of trajectories) and entropy measures for analyses of disorder and predictability.

The radius of gyration captures the size of the trajectory as if it was a physical object. (See the *Methods* section for a definition.) It is, in theory, rather different from the daily average travel distance. Someone who moves in a comparatively confined space will have a small radius of gyration even though he or she covers a large distance. R is also different from the physical extent of the trajectory in that it weighs the contribution from a sector by the time a person has been there. For example, someone that spends most of the time at one location (A) but makes one trip to another location (B) on the other side of the country will have a smaller radius of gyration than someone who constantly travels between A and B. The probability distribution of the radius of gyration for PaP residents is presented in Fig. 2A. We see that, just like the daily travel distance, the average radius of gyration was higher immediately following the earthquake (*spring*) than during the *summer* and *fall* periods. The distributions of the radius of gyration during these later periods (*summer* and *fall*) are very similar to each other, indicating the return to normality after earthquake. For example, during the *spring* period, right after the earthquake, 43% of the studied population had a radius of gyration of more than 10 km, while the corresponding figures for the *summer* and *fall* periods were 32% and 29%, respectively. To summarize, in terms of the shape of the distributions of the trajectories (like the daily travel distances), the disaster enlarged the travel patterns, an effect that lasted into the summer. Even if this is not a true for each individual, other quantities, as we will soon see, support this general picture.

The trajectories of people in stable societal conditions are often quite regular—people visit the same places (home, work, their favorite grocery store, etc.) in the same order (17). Since history repeats itself in this respect, the movements of people are predictable. To study this, we use entropy (disorder) and an information theoretic definition of predictability (described in the *Methods* section). Note that, the sense in which we use the word "predictability" here relates to the regularity of people's movements after the earthquake. Further below we describe how data from before the earthquake could be used to predict people's destinations as they left PaP after the earthquake.

Comparing the entropies S of the three periods (Fig. 2B), we discover that their probability distributions are conspicuously unchanged over the three periods, and the predictability Π (Fig. 2C) is also very similar during the three periods. The only difference is that people are actually slightly more predictable right after the disaster than during the rest of the year. As already mentioned,

Fig. 2. Trajectory analysis of mobile phone users who were present in PaP on the day of the earthquake. (*A*) Shows the cumulative distribution of the radius of gyration: r_g. (*B*) Displays the distribution of entropy S. (*C*) Gives the distribution of the maximal predictability Π. (*D*) Shows the correlation between radius of gyration r_g and Π. In (*E*), we graph the fraction of time a person spent in the top n visited communes $\tilde{\Pi}$. (*F*) gives the averaged R/R^{rand} versus the radius of gyration r_g, showing a relative stable dependence.

people did not become more predictable because they became immobile—they moved even more after the disaster. The probability distribution of entropy $P(S)$ peaks at around 1.5 (Fig. 2*B*). One interpretation is that a typical mobile phone user present in PaP on the day of the earthquake had an uncertainty of $2^{1.5} \sim 2.8$ locations for his or her next destination. The predictability $P(\Pi)$ peaks at around 0.85 (Fig. 2*C*), meaning that we have an upperlimit of 85% to predict the typical person's next destination during each of the three periods. These findings show that, even in this extreme disaster, human movements over the three 3-mon periods remained highly predictable. People moved farther, but not less regularly, during the tumultuous time after the disaster.

Data from a high-income country during stable social conditions have shown that Π is almost constant for people with r_g ranging from 10 up to 1,000 km (17). This is, however, not the case in our data. In Fig. 2*D*, we see that predictability increases with increasing r_g during all three periods (i.e., people traveling farther are more predictable). Furthermore, predictability remains slightly higher after the earthquake for most of the r_g range.

Predictability, based on the regularity in people's movements, gives us a theoretical upper limit of how well we can forecast a person's trajectory, but it does not tell us how to forecast it. The simplest prediction technique is to count the visiting frequency of a person's historical trajectory, taking the most frequently visited location as a predictor of the person's next destination. Because towers are not uniformly distributed across the country, we use a more relevant division of the country, the Haitian "commune" (in total, 140 communes). On average, a mobile phone user spent 75% of the time in his or her most frequently visited commune (see Fig. 2*E*). In agreement with the results above, this pattern is even stronger after the earthquake than during later periods. The top three most visited communes constituted, on average, 95% of the visited locations during *spring* and 90% during *summer* and *fall*. The frequency curves are almost identical for *summer* and *fall*, providing additional evidence that the mobility patterns returned to normal by this time.

Information about an individual's top visited locations provides the simplest way to make predictions about a person's future location. As with Fig. 2*D*, we checked whether the accuracy of such a predictive procedure is dependent on the people' travel distances. However, the more locations someone visits, the lower is the expected frequency of the most visited locations. We compensate for this effect by measuring the ratio between the probability of finding a person in his or her most visited location and the probability of finding an individual at a randomly chosen, previously visited, location—the *relative regularity* R/R^{rand}. The results are presented in Fig. 2*F*, where we can see that the difference between the time periods is negligible. On the other hand, the travel distances have, as expected, little effect—the relative regularity is around 6 for people in PaP with r_g ranging from 1 to 50 km and 4 to 6 for people with $r_g > 50$ km. This means that mobile phone users in PaP were, on average, at least four times more likely to spend their time in the most frequent location during the three periods than in a random location he or she visited during that period.

In sum, we have found that despite the social disorder, the increases in radius of gyration and the increases in average daily travel distance that we observed after the earthquake, the movements of the population remained highly regular and predictable. We have also made the same analysis by taking all the days after the earthquake as a single period and verified that the results are similar (see the *SI Appendix*).

Evacuation and Return Behavior. We now turn to research questions that explain and contextualize the high predictability shown in the previous section. The issues analyzed here are also of direct relevance to relief agencies responding to disasters. We investigate how soon after the earthquake people moved out, how far from

PaP they moved, what proportion of people returned, and how long time people stayed outside the capital after leaving. We also look into a strong predictor of the specific geographical area to which people decided to move after the earthquake, namely their location during the preceding Christmas and New Year.

To investigate how soon people started to move out of PaP, we select mobile phone users in PaP on the day of the earthquake who subsequently left PaP at some point between the earthquake and the end of June 2010 (170 d after the earthquake). We include all mobile phone users irrespective of their calling frequency. We plot the proportion, $P(t)$, of people who left PaP for the first time t days after the earthquake and compare this distribution to distributions later during the year when considerable stabilization had taken place. These five reference periods start on the same weekdays on June 1, 8, 15, 22, and 29 and all end 170 d later (Fig. 3*A*).

Interestingly, we see that the largest proportion of people left not immediately, but three days after the disaster. Although this finding is highly noteworthy, the delay may be partly due to reduced network capacity during the first few days after the earthquake. For $t > 3$, the distribution of the fraction of evacuated individuals is close to a power-law distribution $P(t) \sim t^{-\alpha}$, and reveals that the earthquake caused PaP residents to leave the city much earlier than on normal days. Another interesting finding is, again, the existence of weekly cycles in the reference data. These cycles were absent after the earthquake and then reappeared more than a month afterwards, indicating a return to normality.

So, how far did people move? Using the same inclusion criteria and reference periods, we plot the proportional distribution of the maximum distances the mobile phone users traveled after the earthquake, measured from the center of PaP (Fig. 3*B*). A majority (about 70%) of the individuals traveled quite short distances, maximally within 50 km of PaP center (note, however, the small size of Haiti, Fig. 1*A*). The distribution of maximum distances traveled by affected individuals is almost identical with those traveled during normal times, suggesting that the extremes
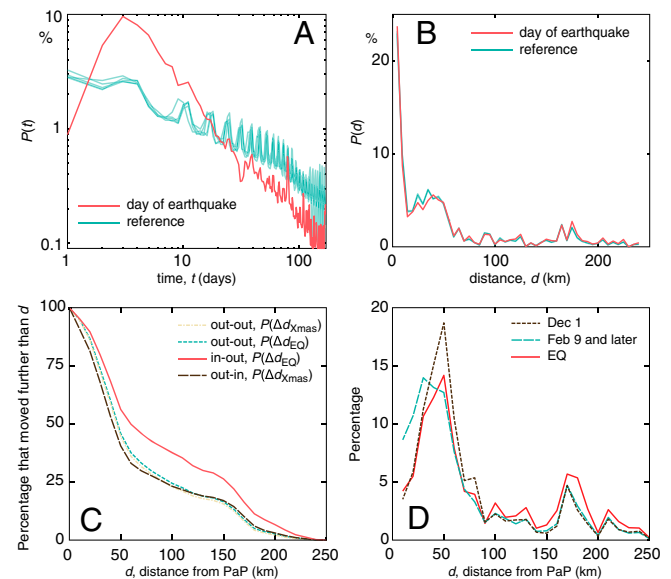

Fig. 3. Analysis of population movements out of PaP. (*A*) Shows the distribution of PaP residents moving out of PaP for the first time by *t* days after the day of the earthquake. In (*B*), we plot the maximum distance to the center of PaP traveled by PaP residents. Reference curves represent sample periods from June 1, 8, 15, 22, and 29 to 170 d after these dates. (*C*) Gives the cumulative distribution of people's relative distances on January 3 and 31 to their locations on the day of the earthquake for four different categories of people. (*D*) Gives the distribution of distance to the center of PaP for individuals present in PaP on the sampled day and outside PaP 19 d later. Results for the period after February 9 are averaged for clarity.

of the trajectories are the same between the PaP population during this first period after the earthquake and during normal times. This is thus another aspect (apart from predictability) that was conserved in the disaster.

The maximal travel distances do not, however, tell us where people travelled. Studies from the 1985 Mexico City earthquake, for example, have shown that people left the city to stay with relatives or friends elsewhere (27, 28). Recent analyses have also shown a high correlation between the locations of people's social contacts and their mobility patterns (29). In our data, we can test this hypothesis by comparing where people were at the time of the earthquake and during the preceding Christmas and New Year holidays (assuming most people spent this time with their family and friends). In Fig. 3C, we select the people who were in PaP on the day of the earthquake and group them into four categories, depending on where they were on January 3 and 31: *in-in* for those in PaP on both these dates; *out-out* for those not in PaP on January 3 and 31; *in-out* for those in PaP on January 3 and not in PaP on January 31; and *out-in* for those not in PaP on January 3 and in PaP on January 31. For these groups, we show the cumulative probability distributions of distances between people's locations on January 3 and their locations in PaP on the day of the earthquake (Δd_{Xmas}) as well as the distribution of distances between their PaP locations on the day of the earthquake and on January 31 (Δd_{Eq}). Results from this analysis are shown in Fig. 3C. First, we notice that the distances for the *out-out* group are almost identical on January 3 and 31, implying that the people who spent the holidays outside PaP are quite likely to have moved to these locations after a disaster. Second, people who were in PaP during Christmas and New Year, but moved out after earthquake (*in-out*), tend to have had longer travel distances than the others, possibly because they had safe social connections farther away and were thus less likely to travel to them during the holidays. Finally, the Δd_{Xmas} distance distribution of the *out-in* group is very similar to that of the *out-out* group, indicating that the reasons these people did not move out of PaP after the earthquake was not determined by their having a different geographic distribution of social contacts than the *out-out* group.

If one can generalize these findings, they point to a way of using the population distributions during the holidays as a way to estimate the distribution of displacement during a disaster. We have confirmed our results on the level of the smallest administrative area in Haiti, the communal section (a total of 570 communal sections). For people in PaP on the day of the earthquake who were outside PaP on both days (*out-out*), 68.5% were in the same communal section on January 31 as they were on January 3. The percentage rises to 85.5% when we use the larger administrative area of department (province).

We now relate the travel distributions in Fig. 3C to normal mobility patterns. We analyze two groups of people—individuals in PaP on December 1, 2009 and individuals in PaP on the day of the earthquake (January 12, 2010), and we use as reference the groups of individuals present in PaP on Tuesday after February 9. For these groups, we plot (Fig. 3D) the cumulative distribution of distances from PaP center to the individuals' locations 19 d afterwards (on December 19, January 31, etc.). Again, the postearthquake distributions are very similar to those of the reference periods. Differences consist of a slightly higher proportion travelling longer distances, possibly for the same reasons as discussed for the Christmas and New Year analysis above. There is also a small variation in the short to intermediate distances (30 to 70 km). The difference when comparing the postearthquake distributions in Fig. 3B (red curve) and Fig. 3D (red curve) is the large proportion of very short trips in the former. This is due to that Fig. 3B also captures a large number of persons who stray outside the city boundaries for short periods of times, while Fig. 3D captures predominantly those who left for a longer per-

iod. A further analysis of the duration people stay outside PaP can be found in the *SI Appendix*.

In order to investigate the potential bias from differences in calling frequencies, we have repeated all analyses for subgroups of users with different calling frequencies. These analyses reveal that the results overall are robust. For analyses where time or duration is an outcome (Fig. 3A), the conclusions in the text remain when analyzing groups with different calling frequencies, although the exact size of the differences between the earthquake and reference periods should be interpreted with caution (see *SI Appendix*).

Discussion

We have shown that, despite sharp changes in people's mobility patterns after the tragic earthquake of January 12, 2010, in Haiti, the predictability of people's movements over the three-month period after the earthquake was very high and remained unchanged in comparison to later parts of the year after considerable stabilization had taken place. We show that the destinations of people who left the capital early (during the first three weeks after the earthquake) were highly correlated to their mobility patterns during normal times, and specifically to the locations at which people had significant social bonds, as measured by where they spent Christmas and New Year holidays.

The above findings imply that population movements following large-scale disasters may be significantly more predictable than previously thought. Given the fundamental importance of knowing the locations of affected populations during disaster relief operations, these findings suggest that disaster planning and response may be significantly improved. On a more abstract level, the results force a change in our conceptualization of disasters as fundamentally chaotic events. People's movements are highly influenced by their historic behavior and their social bonds, and this fact remained even after one of the most severe disasters in history.

In the quest for a globally applicable model of human mobility, the present study confirms several findings from high-income countries, including the power-law distribution of travel distances and the high predictability of travel trajectories during stable social conditions (17). A further investigation in the *SI Appendix* shows that for people in PaP, the waiting times until leaving PaP, as well as the time until they returned, was power-law distributed, both during normal days and after the earthquake, albeit with different exponents. The results also highlight the very high mobility of people during normal conditions and the importance of taking these movements into account when investigating postdisaster movements.

Several limitations and avenues for future work exist (for additional discussions, see ref. 13). First and foremost, the exact correlation between disaster-related movements of people with and without mobile phones needs to be better understood. We have, however, shown very promising results in this area in earlier research (13). Additionally, mobility patterns in different types of disasters and different social contexts may vary considerably. The present dataset covered 42 d before the disaster. Additional data, including a longer period before the earthquake, would have allowed better differentiations of people's permanent place of living, social contacts, and predisaster mobility patterns.

Another limitation is that the data set included only one location update per day. This means that movements taking place during the course of 24 h, especially people's movements inside their cities and home communes, will not be included in the data. Granted this limitation, since our interest in this paper is primarily on movements over distances that are sufficiently large to prevent access to relief supplies, a daily resolution should provide the most important features of mobility relevant to relief coordinators. Yet another issue is how the societal changes during the disaster affect the sampling. The mobile phone network suffered

reduced capacity immediately after the earthquake, but was functioning again within a few days. This may have contributed to bias in the first week's data, but does not alter our conclusions. Lack of access to electrical charging could perhaps have reduced the number of calls. However, power was also frequently interrupted before the earthquake, and existing electric generators seem to have supplied considerable charging capacity. The effects of fatalities and the loss of phones in the disaster were circumvented by only studying users present at both the beginning and end of the dataset. Lack of possibilities to put credit on the phones shortly after the earthquake could have been another bias. However, the mobile phone operator, Digicel, supported their customers by adding five USD in calling credit to all accounts after the disaster.

In summary, the results show that population movements following the Haiti disaster had a high level of predictability and seemed highly influenced by people's social support structures. These findings form an important first step in forecasting the effects of large-scale disasters. With future research in other disaster contexts, such forecasts are likely to become an important part of national disaster preparedness planning and in predicting population movements during ongoing disaster relief operations.

Methods

Radius of Gyration. Let $T_i = \{t_1, t_2, \cdots, t_{L_i}\}$ be the sequence of mobile phone towers that person i visited during a period. Let $\mathbf{r}(t)$ be the location of t. Then the radius of gyration of i's trajectory in the specific period is

$$r_g(i) = \sqrt{\frac{1}{L_i} \sum_{k=1}^{L_i} (\mathbf{r}(t_k) - \bar{\mathbf{r}})^2}, \qquad [1]$$

where $\bar{\mathbf{r}} = \frac{1}{L_i} \sum_{k=1}^{L_i} \mathbf{r}(t_k)$ is the center of mass of the trajectory.

Predictability. To evaluate the predictability, we (following ref. 17) use a measure of entropy, or disorder, that accounts for both the relative frequency of the visited locations and the order of the visits:

$$S_i = - \sum_{T_i' \subset T_i} P(T_i') \log_2[P(T_i')], \qquad [2]$$

where $P(T_i')$ is the probability of finding a subsequence T_i' in T_i. Based on this measure of entropy, one can estimate the upper bound of the success rate in predicting the future location of the mobile phone user immediately after T_i. We get the maximum predictability, Π_i, by solving a limiting case of Fano's inequality (a relation derived from calculation of the decrease in information in a noisy information channel):

$$S_i = H(\Pi_i) + (1 - \Pi_i) \log_2(N - 1), \qquad [3]$$

where

$$H(\Pi_i) = -\Pi_i \log_2(\Pi_i) - (1 - \Pi_i) \log_2(1 - \Pi_i), \qquad [4]$$

and N is the number of distinct locations visited by person i (30–32).

1. Guha-Sapir D, Vos F, Below R, Ponserre S (2011) *Annual Disaster Statistical Review 2010: The Numbers and Trends* (CRED, Brussels).
2. IDMC and Norwegian Refugee Council (2011) *Displacement Due to Natural Hazard-Induced Disasters: Global Estimates for 2009 and 2010* (IDMC and Norwegian Refugee Council, Geneva).
3. Brown V, et al. (2001) Rapid assessment of population size by area sampling in disaster situations. *Disasters* 25:164–171.
4. Depoortere E, Brown V (2006) *Rapid Health Assessment of Refugee or Displaced Populations* (Médecins Sans Frontières, Paris), 3rd Ed.
5. Levine JN, Esnard AM, Sapat A (2007) Population displacement and housing dilemmas due to catastrophic disasters. *J Plann Lit* 22:3–15.
6. DMC and Norwegian Refugee Council (2011) *Internal Displacement: Global Overview of Trends and Developments in 2010* (DMC and Norwegian Refugee Council, Geneva).
7. Johansson A, Helbing D (2008) From crowd dynamics to crowd safety: A video-based analysis. *Adv Complex Syst* 11:497–527.
8. Moussaid M, Garnier S, Theraulaz G, Helbing D (2009) Collective information processing and pattern formation in swarms, flocks, and crowds. *Top Cogn Sci* 1:469–497.
9. Bickman L, McDaniel MA (1976) Model of human-behavior in fire emergencies. *Bull Psychon Soc* 8:254–254.
10. Hahm J, Lee JH (2009) Human errors in evacuation behavior during a traumatic emergency using a virtual fire. *Cyberpsychol Behav* 12:98–98.
11. Naik A (2009) Migration and natural disasters. *Migration, Environment and Climate Change: Assessing the Evidence*, eds F Laczko and C Aghazarm (IOM, Geneva).
12. Bagrow JP, Wang DS, Barabási AL (2011) Collective response of human populations to large-scale emergencies. *Plos ONE* 6:e17680.
13. Bengtsson L, Lu X, Thorson A, Garfield R, von Schreeb J (2011) Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: A post-earthquake geospatial study in Haiti. *PLoS Med* 8:e1001083.
14. Hori M, Schafer MJ, Bowman DJ (2009) Displacement dynamics in southern Louisiana after hurricanes Katrina and Rita. *Popul Res Policy Rev* 28:45–65.
15. Raleigh C, Jordan L, Salehyan I (2008) *Assessing the Impact of Climate Change on Migration and Conflict* (World Bank, Washington, DC).
16. International Telecommunication Union (2011) World Telecommunication/ICT Indicators Database Online. (International Telecommunication Union, Geneva).
17. Song CM, Qu ZH, Blumm N, Barabási AL (2010) Limits of predictability in human mobility. *Science* 327:1018–1021.
18. González MC, Hidalgo CA, Barabási AL (2008) Understanding individual human mobility patterns. *Nature* 453:779–782.
19. Song C, Koren T, Wang P, Barabási AL (2010) Modelling the scaling properties of human mobility. *Nat Phys* 6:818–823.
20. Eagle N, Pentland A, Lazer D (2009) Inferring friendship network structure by using mobile phone data. *Proc Natl Acad Sci USA* 106:15274–15278.
21. Eagle N, Pentland AS (2009) Eigenbehaviors: Identifying structure in routine. *Behav Ecol Sociobiol* 63:1057–1066.
22. Pentland A, Choudhury T, Eagle N, Singh P (2005) Human dynamics: Computation for organizations. *Pattern Recognit Lett* 26:503–511.
23. Wesolowski A, Eagle N (2010) Parameterizing the dynamics of slums. *AAAI Spring Symposium 2010 on Artificial Intelligence for Development* (AAAI, Stanford).
24. Archibold RC (January 13, 2011) Haiti: Quake's toll rises to 316,000. *NY Times*, http://www.nytimes.com/2011/01/14/world/americas/14briefs-Haiti.html.
25. Schwartz T (2011) Building assessments and rubble removal in quake-affected neighborhoods in Haiti. *BARR Survey Final Report* (LTL Strategies, Washington, DC) .
26. Institut Haïtien de Statistique et d'Informatique (2009) Population totale, population de 18 ans et plus menages et densités estimés en 2009. (Institut Haïtien de Statistique et d'Informatique, Port-au-Prince).
27. Quarantelli LE (1992) Human behavior in the Mexico City earthquake: Basic themes from survey findings. (Disaster Research Center, University of Delaware, Newark).
28. Dynes RR, Quarantelli EL, Wenger D (1990) Individual and organizational response to the 1985 earthquake in Mexico City, Mexico. (Disaster Research Center, University of Delaware, Newark).
29. Cho E, Myers SA, Leskovec J (2011) Friendship and mobility: User movement in location-based social networks. *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM, San Diego, CA), pp 1082–1090.
30. Kontoyiannis I, Algoet PH, Suhov YM, Wyner AJ (1998) Nonparametric entropy estimation for stationary processes and random fields, with applications to English text. *IEEE Trans Inf Theory* 44:1319–1327.
31. Navet N, Chen S-H (2008) On predictability and profitability: Would GP induced trading rules be sensitive to the observed entropy of time series? *Natural Computing in Computational Finance, Studies in Computational Intelligence*, eds A Brabazon and M O'Neill (Springer, Berlin/Heidelberg), Vol 100, pp 197–210.
32. Fano R (1961) *Transmission of Information; A Statistical Theory of Communications* (MIT Press, Cambridge, MA).
33. Clauset A, Shalizi CR, Newman MEJ (2009) Power-Law distributions in empirical data. *SIAM Rev* 51:661–703.

Supporting Information

S1. Entropy analysis

In this section, we present further details relating to the entropy analyses of the movements of mobile phone user present in PaP on earthquake day. For convenience, we present our results in the same order as in Ref. (1).

A. Data Processing. For each person the history trajectories are a series of geo-tagged location IDs, which in our case are tower IDs. Let $X_i = \{x_1, x_2, ..., x_T\}$ be sequence of daily locations observed for person i during the sampled T days. x_j equals the tower ID if person i is known on day j, otherwise we mark x_j "unknown". Each tower ID x_j is associated with the projected coordinate $r(x_j)$. Since we analyze the travel patterns of PaP individuals after earthquake, the total length of string X_i, T, is then 342 days (from January 12, 2010 to December 9, 2010). To compensate for the relative short overall time, we segment the 342 days into three equal periods and restrict our analysis to individuals who had less than $q = 30\%$ unknown days in all periods to achieve good statistics (compared to $q = 80\%$ used in Ref. (1)). The final sample includes 303,623 qualified individuals.

B. Results. *Radius of Gyration.* The radius of gyration captures the size of the trajectory, that is, the "average" traveling distance, as if the individual were a physical object. It is defined as:

$$r_g = \sqrt{\frac{1}{L}\sum_{k=1}^{L}\left(\mathbf{r}\left(x_k\right) - \bar{\mathbf{r}}\right)^2} \qquad [\mathbf{1}]$$

where L is the number of observed locations for individual i, and $\bar{\mathbf{r}} = \frac{1}{L}\sum_{k=1}^{L}\mathbf{r}\left(x_k\right)$ is the center of mass of the trajectory X_i. We can see from Fig. S1 that r_g follows a fat-tailed distribution for PaP individuals, consistent with the result of (1, 2). The clear cut off for the distribution well around $r_g = 100$ km, is due to the limits of Haiti's borders.

Entropy. Entropy is a common measurement for disorderedness, the larger the entropy, the greater the disorder, and consequently it implies lower predictability. Here, we inherit the series of entropy measurements studied in Ref. (1): (i) the *random entropy*, $S^{\text{rand}} = \log_2 L_i$, capturing the predictability by assuming each person's whereabouts is uniformly distributed among the L_i distinct locations; (ii) the *temporal-uncorrelated entropy*, $S^{\text{unc}} = -\sum_{k=1}^{L_k} p_k \log_2 p_k$, where p_k is the frequency at which the person visited tower k, characterizing the heterogeneity of visiting patterns between different locations; (iii) the *true entropy* (3), $S = -\sum_{X_i' \subset X_i} P(X_i')\log_2[P(X_i')]$, where $P(X_i')$ is the probability of finding a sub-sequence X_i' in X_i, considering both spatial and temporal patterns.

For each person i, we calculate S^{rand}, S^{unc} and S according to the above definitions and the obtained distribution $P(S^{\text{rand}})$, $P(S^{\text{unc}})$, and $P(S)$ are shown in Fig. S2. Similar to Ref. (1), we observe a prominent shift from $P(S^{\text{rand}})$ to $P(S)$. The distribution of random entropy $P(S^{\text{rand}})$ peaks around $S^{\text{rand}} \approx 5$, indicating that if we assume that individuals randomly choose locations the next day, a typical individual could be found on average in any of $2^{S^{\text{rand}}} \approx 32$ locations. On the other hand, if we utilize information contained in the frequency and sequence order of the trajectory of individuals, the uncertainty in a typical individual's whereabouts is only $2^{S^{\text{unc}}} = 2^{2.35} \approx 5$, and $2^S = 21.4 \approx 2.6$, respectively.

Maximum Predictability. Fano's inequality (4, 5) gives an upper limit for the predictability of an individual (Π), with entropy E moving between N locations:

$$\Pi \leq \Pi^{\text{Fano}}(E, N) \qquad [\mathbf{2}]$$

where Π^{Fano} is given by

$$E = H(\Pi^{\text{Fano}}) + (1 - \Pi^{\text{Fano}})\log_2(N - 1) \qquad [\mathbf{3}]$$

and

$$H(\Pi^{\text{Fano}}) = -\Pi^{\text{Fano}}\log_2(\Pi^{\text{Fano}}) - (1 - \Pi^{\text{Fano}})\log_2(1 - \Pi^{\text{Fano}})$$
$$[\mathbf{4}]$$

Fano's inequality reveals that, providing information given in E, the accuracy of the best possible predictive algorithm cannot exceed Π^{Fano}.

Let $\Pi^{\text{rand}} = \Pi^{\text{Fano}}(S^{\text{rand}}, N)$, $\Pi^{\text{unc}} = \Pi^{\text{Fano}}(S^{\text{unc}}, N)$ and $\Pi^{\text{max}} = \Pi^{\text{Fano}}(S, N)$, since $S^{\text{rand}} \geq S^{\text{unc}} \geq S$, Π^{max} then provides the ultimate best possible predictive power since it utilize maximum information from S. We consequently refer to Π^{max} as the "maximum predictability" in the context. The distribution of the above predictability quantities, $P(\Pi^{\text{rand}})$, $P(\Pi^{\text{unc}})$ and $P(\Pi^{\text{max}})$ are shown in Fig. S3.

The distribution of $P(\Pi^{\text{max}})$ peaks around $\Pi^{\text{max}} \approx 0.85$, meaning that we typically have an upper limit of 85% accuracy to predict the typical person's next destination following the earthquake. These findings show that, even in a extreme disaster, human movements over the three three-month periods remained highly predictable. People moved farther but not less regularly in the tumultuous time right after the disaster. We also observe a wide distribution of $P(\Pi^{\text{unc}})$, which peaks around $\Pi^{\text{unc}} \approx 0.67$, and an extremely left shifted $P(\Pi^{\text{rand}})$ with the peak $\Pi^{\text{rand}} = 0$. The above analysis reveals that neither Π^{rand} nor Π^{unc} are effective predicative tools, and a significant share of predictability is encoded in the temporal order of the visitation pattern, as concluded in Ref. (1).

r_g vs. the maximum predictability. From above analysis, we have found that despite the fat-tailed distribution of the radius of gyration among PaP inhabitants, the maximum predictability, Π^{max}, narrowly peaks around $\Pi^{\text{max}} \approx 0.85$, implying the potential independence of predictability on the travel distances. We investigate this correlation by plotting the log-binned average of Π^{max} over r_g, as shown in Fig. S4. Strikingly, contrary to what we observed in Ref. (1), which found that Π^{max} is largely independent of r_g for $r_g \geq 10$ km, for the PaP inhabitants after earthquake, there is a steady increase in maximum predictability when r_g, varying from 10 to 110 km, indicating that individuals with r_g covering a hundred kilometers are even much more predictable than those whose lives are limited to several kilometers.

Frequency of top visited locations. The simplest predictive algorithm is to use the most frequently visited locations to predict the person's next destination. Thus, the frequency of the top n visited locations, $\tilde{\Pi}(n)$, provides an upper bound for Π^{max}. In addition to tower, we also use a more practical location indicator, the lowest Haitian administrative region level "commune sections" to calculate $\tilde{\Pi}(n)$. We can see from Fig. S5 that a typical PaP individual spent 77% and 90% of his or her time on the most frequently visited towers and sections when $n = 2$, respectively. Compared to $\tilde{\Pi}(2) \approx 60\%$ in Ref. (1), these values are relatively higher. The cause of such

a difference may be varied: first, the density of towers in PaP may be relatively sparser than in developed countries; second, there are fewer transportation facilities in Haiti, which hinders the possibility of visiting diverse towers; and third, the trajectory data was collected on a daily basis and are much shorter than those used in Ref. (1). These reasons also yield a much quicker convergence of $\tilde{\Pi}(n)$ to $\tilde{\Pi} = 1$.

r_g *vs. relative regularity.* Relative regularity is defined by the ratio between the probability of finding an individual at the most visited location, $R = \tilde{\Pi}(1)$, and the probability of finding an individual at a randomly picked tower, $R^{\text{rand}} = 1/L_i$: R/R^{rand}. The relative regularity compensates for the limit of the above analysis in that the more locations an individual visits, the lower is the expected frequency of the most visited locations. Similar to Fig. S4, we plot the mean relative regularity, $\langle R/R^{\text{rand}} \rangle$, against r_g, as shown in Fig. S6. Strikingly, again, we find a steady increase in $\langle R/R^{\text{rand}} \rangle$ for $r_g \in [1\text{km}, 80\text{km}]$. There is a drop off in $r_g \approx 100$ km; however, the value of $\langle R/R^{\text{rand}} \rangle$ remains higher in a large range of r_g ($r_g \in [1\text{km}, 20\text{km}]$).

S2. Analysis of return behavior

Now we turn to the question of the pattern of return to PaP. We select all individuals who were present in PaP on the day of the earthquake, who left PaP some time after the earthquake and who then returned to PaP before July 1 (170 days after earthquake). We plot the probability distribution of the durations of people's first stay outside PaP and compare this with sample periods of the same length starting on the same weekdays on June 1, 8, 15, 22 and 29, respectively (Fig. S7A). The differences between the reference curves (green) are extremely small – all lines basically fall on top of each other. Interestingly, the distributions follow a power-law functional form quite closely. This means that most people who left the city during normal times stayed outside for relatively short periods of time, while the differences between people were very large. This general pattern remained after the earthquake, but the difference between people became even larger with a higher proportion spending a longer period away.

As people moved out at different points in time, it is important for government decision makers and relief agencies to know the proportion of people returning on each day after a disaster. The overall trend in these analyzes (Fig. S7B) is that the fraction of returning people decreases exponentially with time, although the exact shape of the curve is sensitive to people's calling frequency. The proportion of persons who returned to PaP is higher one week to a month after the earthquake than during the reference periods. However, this larger fraction can be explained by the fact that more people moved out during the beginning of the earthquake period, as we discussed in the context of Fig. 3A. The Pearson correlation between the dates the peoples leave and return is quite high at 0.71.

S3. Sensitivity analysis on data sparseness

A. Entropy and predictability. The validity of the algorithm used for calculating S has been evaluated comprehensively over different lengths L (from $L_{\min} = 48$ data points to $L_{\max} = 2352$ data points), as well as for different proportions of unknown locations in the trajectories (1).

It has been shown that for a typical user with a trajectory as short as 48 data points and with half of the locations in the trajectory unknown, the estimation error of the algorithm-based S is only about 25% (1). Comparatively, in our study, the length of the trajectories are $L = 114$ data points (cor-

responding to a 4.75 days hourly trajectory in Ref. (1)), and the proportion of unknown locations, q, in our analyzes on entropy and predictability is at most 30%. The estimation error based on these parameters is thus expected to be considerably smaller than 10%. See Fig. S6D in Ref. (1).

To confirm the inferences above and to investigate how sensitive our results are to inclusion of additional and increasingly sparse trajectories, we have produced separate analyzes for groups with different proportions of missing data. We construct three groups of users with $q \leq 30\%$, $q \leq 40\%$ and $q \leq 50\%$. The distributions of S^{rand}, S^{unc}, and S during each period, are shown in Fig. S8. We can see that the distributions of S^{rand} and S^{unc} are virtually unchanged when the proportions of unknown locations increase and this holds true for all three periods. The distributions of S become slightly left-shifted with increasing values of q. This tendency is somewhat more pronounced during the spring period although the overall changes are small.

We have also made the same separate analyzes on distributions of the maximum predictabilities, Π^{max}, with increasing data sparseness (Fig. S9). As expected from the entropy distributions, Π^{max} changes very little when the proportions of unknown locations increase. The left-shifted distribution of S above results in a slight increase in Π^{max}, which remain at around 0.8 to 0.9.

In summary, these sensitivity analyzes do not provide evidence for serious bias due to the relatively low temporal resolution of this data set.

B. Travel patterns. Calling frequency directly affects the number of observations in our data and users who rarely call are consequently under-observed. Consequently, if there were a difference in the travel patterns of mobile phone users with different calling frequencies, the results of our study would be biased towards the behavior of people who call often. We therefore investigated whether we observe different travel patterns in groups with high calling frequency compared to groups with low calling frequency. To do this, we divide the mobile phone users into four groups according to the number of active phone calls. Specifically, if f is the proportion of days a user made a call during the study period, then each user i is categorized into a quartile group where f fits in *group 1* for $0 < f \leq 25\%$, *group 2* for $25\% < f \leq 50\%$, *group 3* for $50\% < f \leq 75\%$, and *group 4* for $75\% < f \leq 100\%$.

We then produce, for the four groups, the same analyzes on daily traveling distances, evacuation and return behaviors as shown in the main paper. Like in the main paper, we differentiate between people present in PaP on the earthquake day (PaP group) and people outside PaP on this day (non-PaP group).

Distributions of daily travel distances. We start by looking at the changes in daily travel distance distributions of the four groups (of the preceding section) before and after the earthquake by comparing the fitted power law exponent α, for the PaP group and the non-PaP group, see Fig. S10. Here we show the daily value of α, as opposed to the weekly average in Fig. 1F. The patterns for the four groups with different calling frequencies are in concurrence with the results shown in Fig. 1F. As in Fig. 1F, there are clear differences when comparing the α of the PaP group and the non-PaP group. This difference starts around the Christmas-New Year time and then becomes suddenly more pronounced at the time of the earthquake, after which the differences between the groups largely disappeared by the end of the spring. We also note a short sharp change during Easter.

Remaining analyzes where time and durations are outcomes. Groups with low calling frequency are under-observed

in the data and this fact may be particularly problematic when studying events taking place during short time intervals. To investigate the effects of low calling frequency on the analyzes that are time-dependent (Fig. 3*A* and Fig. S7), we plot separate curves for the four groups with different calling frequencies (see Fig. S11). We can see that during the early part of the analysis periods, larger proportions of users from the frequently calling group (*group 4*) are observed, compared to users from *group 1*, *group 2*, *group 3*. This difference is very clear for the reference periods. For example, under normal conditions, only about 1% of the most frequent callers in the PaP groups are observed to leave PaP on the second day after the start of observations during the reference periods, while the proportion for *group 4* is above 5%. These analyzes show that the exact proportions of people leaving at a certain day should be interpreted with caution, especially also since the network experienced substantial problems during the first days. However, it is difficult to imagine a plausible scenario where changes in calling behaviors would invalidate the main conclusions of these analyzes i.e. the earlier evacuation times and the increased duration of time spent outside PaP after the earthquake, as well as the functional power-law shapes observed.

Remaining analyzes where travel distances are outcomes. We then proceed to check whether the under-observation of infrequent callers will affect our conclusions regarding the remaining traveling distance distributions. We thus replotted Fig. 3*B*, *C* and *D* and divided each group into the same four sub-groups as above, based on their calling frequency, i.e. *group 1* to *group 4*. The analyzes show that for all travel distance distributions there are only very small differences between groups of people with different calling frequency. This is the case for all the reference periods and the earthquake periods. These results strongly suggest that differences and changes in calling frequencies did not bias the results regarding travel distance distributions.

1. Song CM, Qu ZH, Blumm N, Barabási AL (2010) Limits of predictability in human mobility. Science 327:1018-1021.

2. Gonzalez MC, Hidalgo CA, Barabási AL (2008) Understanding individual human mobility patterns. Nature 453:779-782.

3. Kontoyiannis I, Algoet PH, Suhov YM, Wyner AJ (1998) Nonparametric entropy estimation for stationary processes and random fields, with applications to English text. IEEE Transactions on Information Theory 44:1319-1327.

4. Fano R (1961) Transmission of Information; A Statistical Theory of Communications (MIT Press, Cambridge, MA).

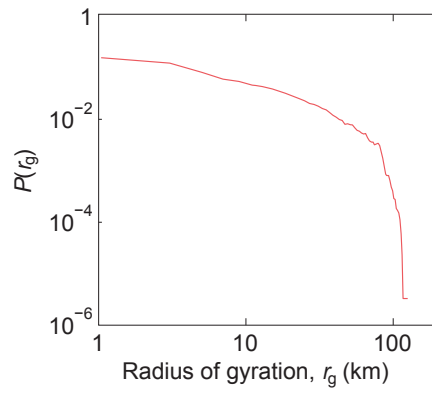5. Navet N, Chen SH (2008) Natural computing in computational finance (Springer, Berlin).

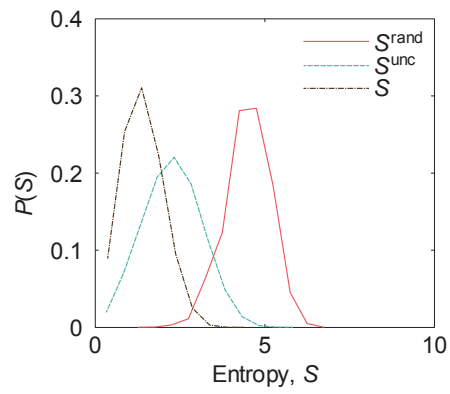**Fig.S1.** The cumulative distribution of radius of gyration r_g.
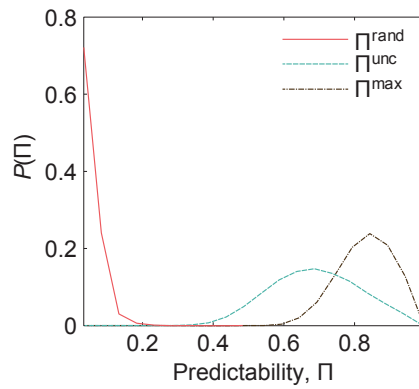

Fig.S2. Distribution of entropies S^{rand}, S^{unc} and S.

Fig.S3. Distribution of Π^{rand}, Π^{unc}, and Π^{max}.
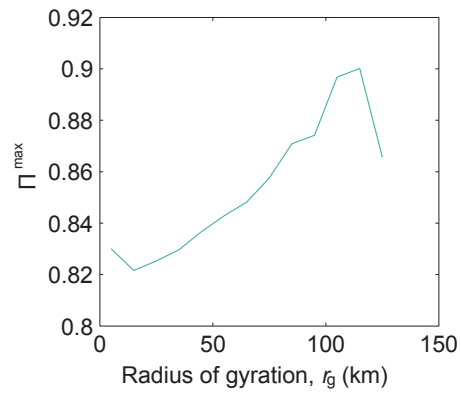
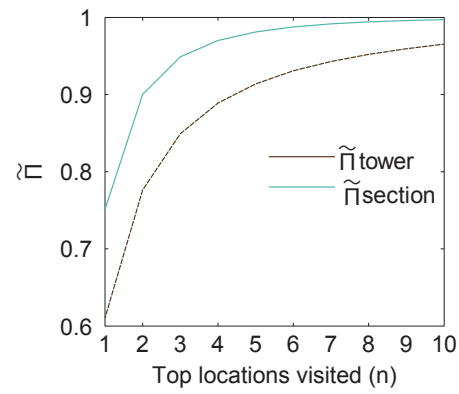**Fig.S4.** The dependence of predictability Π^{max} on r_g.

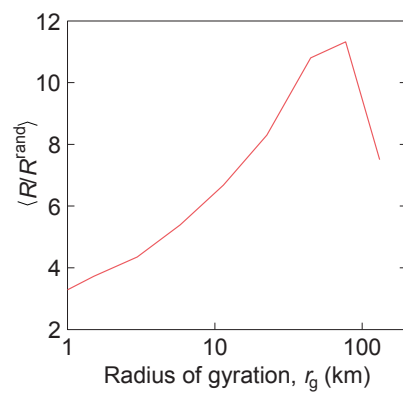**Fig.S5.** The fraction of time an individual spent in the top n visited towers or communal sections.

Fig.S6. The averaged relative regularity, $\langle R/R^{\mathrm{rand}}\rangle$, versus the radius of gyration r_g.
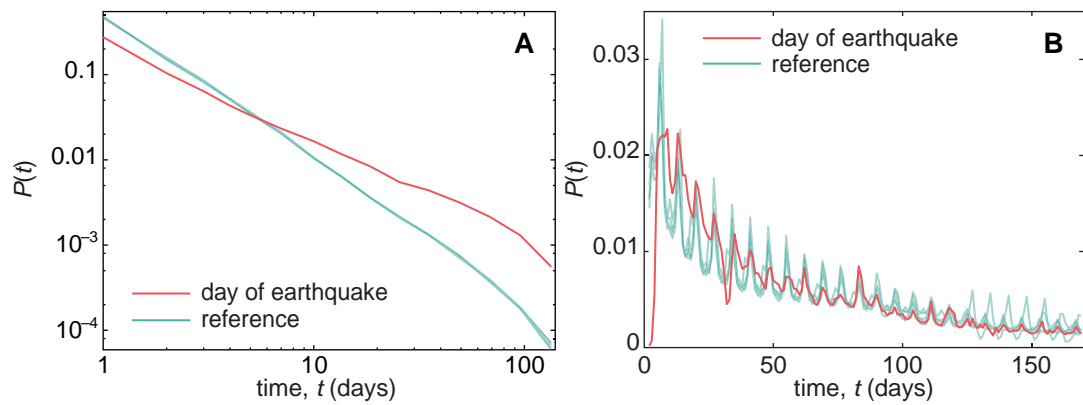
Fig.S7. Analysis of return behavior. Panel A shows the distribution of duration of PaP persons' first stay outside PaP during the post-earthquake and autumn periods. B shows the number of days after the first day outside PaP when a person first returned to PaP. The reference curves represent sample periods from June 1, 8, 15, 22, 29, 2010 and 170 days after.
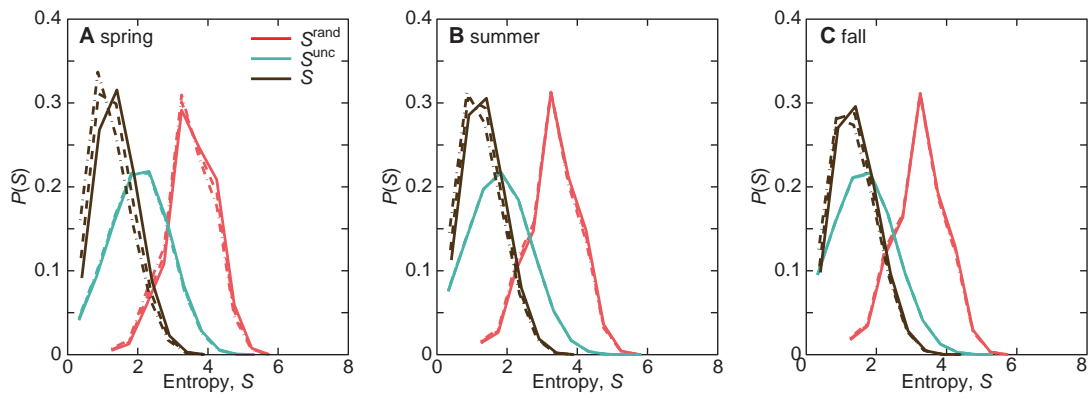
Fig.S8. A Distribution of entropies S^{rand}, S^{unc}, S, during the spring period, for users with q less than 30% (solid line), less than 40% (dashed line) and less than 50% (dash-dotted line), B and C shows the corresponding distributions for the summer and autumn periods.
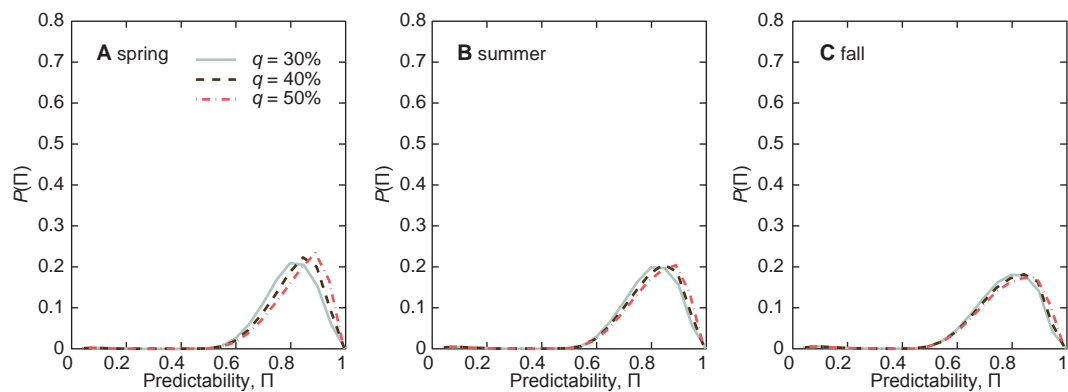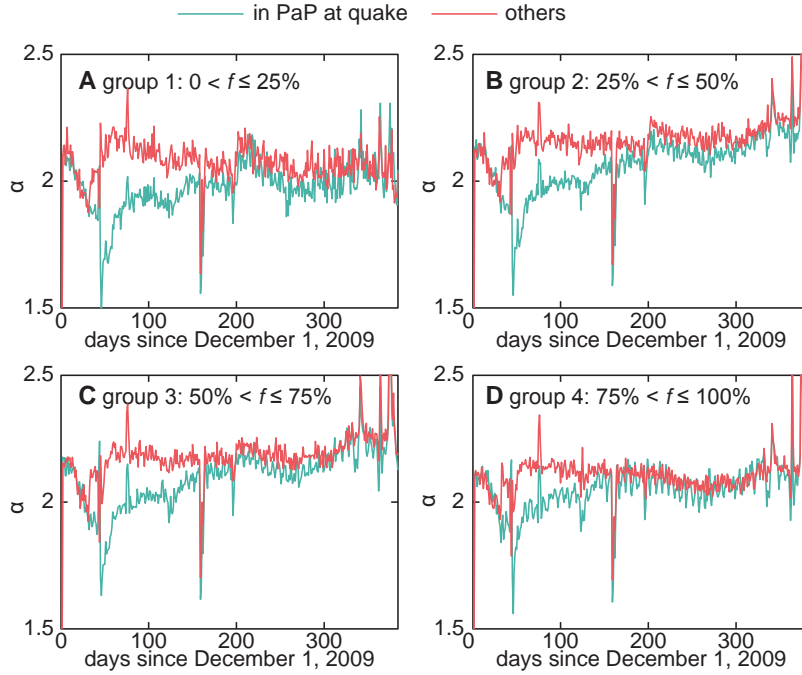

Fig.S9. A Distributions of Π^{max} during the spring period, for users with q less than 30% (solid line), less than 40% (dashed line) and less than 50% (dash-dotted line), B and C shows the corresponding distributions for the summer and autumn periods.

Fig. S10. The change of the exponent α of the power-law distribution of daily travel distances, $P(d) \sim d^{-\alpha}$, for people with different calling frequency f: A $0 < f \le 25\%$; B $25\% < f \le 50\%$; C $50\% < f \le 75\%$; D $75\% < f \le 100\%$.
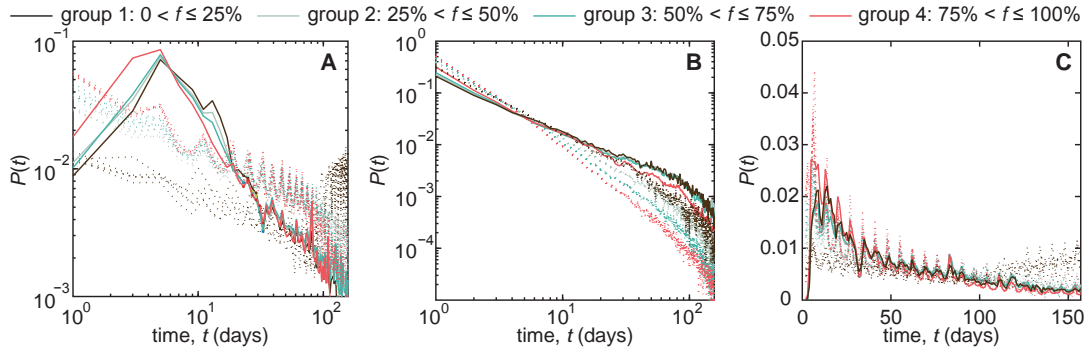

Fig. S11. Analyzes focusing on time as an outcome for people with different calling frequency: $0 < f \le 25\%$ (*group 1*), $25\% < f \le 50\%$ (*group 2*), $50\% < f \le 75\%$ (*group 3*), and $75\% < f \le 100\%$ (*group 4*). A shows the distribution of PaP residents moving out of PaP for the first time by t days after the day of the earthquake. B illustrates the distribution of duration of PaP persons' first stay outside PaP during the post-earthquake and autumn periods. C shows the number of days after the first day outside PaP when a person first returned to PaP. The reference curves (shown as dotted lines of the same color for a group) represent sample periods from June 1, 8, 15, 22, 29, 2010 and 170 days after.
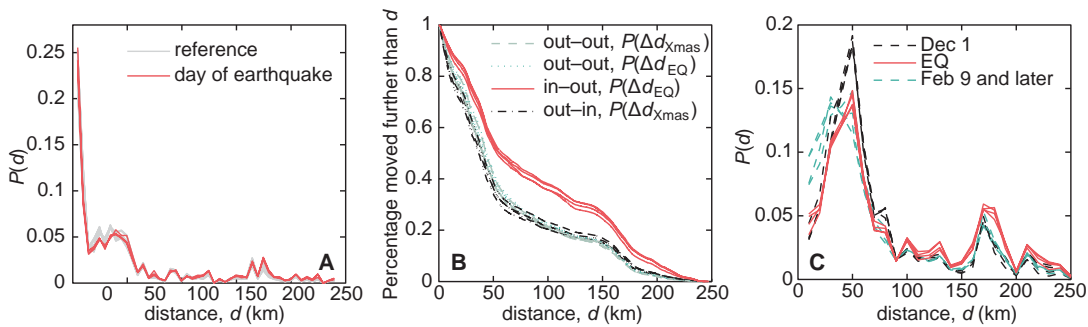

Fig. S12. Analyzes focusing on distances as an outcome for people with different calling frequency: $0 < f \le 25\%$ (*group 1*), $25\% < f \le 50\%$ (*group 2*), $50\% < f \le 75\%$ (*group 3*), and $75\% < f \le 100\%$ (*group 4*). Since the distributions are extremely similar between the four groups, the same color and shape is used for all groups to increase visual clarity. A maximum distance to the center of PaP traveled by PaP residents. Reference curves represent sample periods from June 1, 8, 15, 22, and 29 to 170 days after these dates. B cumulative distribution of people's relative distances on January 3 and 31 to their locations on the day of the earthquake. C distribution of distance to the center of PaP for individuals presented in PaP on the sampled day and outside PaP 19 days later. Results for the period after February 9 are averaged for clarity.