UC Berkeley
CEGA Working Papers

Title
Privacy Guarantees for Personal Mobility Data in Humanitarian Response

Authors
Kohli, Nitin
Aiken, Emily
Blumenstock, Joshua

Privacy Guarantees for Personal Mobility Data in Humanitarian Response

Nitin Kohli, Emily Aiken, and Joshua Blumenstock

CEGA
Center for Effective Global Action

Working Paper Series

Center for Effective Global Action
University of California

eScholarship
University of California

Recommended Citation:

Privacy Guarantees for Personal Mobility Data in Humanitarian Response

Nitin Kohli[1,+], Emily Aiken[2,+], and Joshua Blumenstock[1,2,*]

[1]Center for Effective Global Action, UC Berkeley, 94704 USA
[2]School of Information, UC Berkeley, 94704 USA
[+]These authors contributed equally to this work.
[*]jblumenstock@berkeley.edu

ABSTRACT

Personal mobility data from mobile phones and other sensors are increasingly used to inform policymaking during pandemics, natural disasters, and other humanitarian crises. However, even aggregated mobility traces can reveal private information about individual movements to potentially malicious actors. This paper develops and tests an approach for releasing private mobility data, which provides formal guarantees over the privacy of the underlying subjects. Specifically, we (1) introduce an algorithm for constructing differentially private mobility matrices, and derive privacy and accuracy bounds on this algorithm; (2) use real-world data from mobile phone operators in Afghanistan and Rwanda to show how this algorithm can enable the use of private mobility data in two high-stakes policy decisions: pandemic response and the distribution of humanitarian aid; and (3) discuss practical decisions that need to be made when implementing this approach, such as how to optimally balance privacy and accuracy. Taken together, these results can help enable the responsible use of private mobility data in humanitarian response.

Introduction

Personal mobility data has the potential to provide critical information to guide the response to humanitarian crises, and to advance several of the Sustainable Development Goals (SDGs). For instance, recent work has shown that mobility data can be used to model and prevent the spread of epidemics[1–3], monitor and assist displaced populations after natural disasters[4,5], target emergency cash transfers[6,7], and flag riots and violence against civilians[8,9]. The COVID-19 pandemic in particular demonstrated that aggregated personal mobility data can provide critical insight into population movement[10–14], and resulted in a push for collaborations among governments, technology companies, and researchers to leverage these data to inform evidence-based pandemic response measures[15].

At the same time, personal mobility data is, by definition, personal, so its analysis – even for humanitarian purposes – may create privacy risk. Individual mobility data can expose personally identifying and sensitive information, including individuals' home and work locations, travel patterns, and interpersonal interactions. Moreover, since certain locations are endowed with social meaning, mobility data can be used to draw inferences about an individual's life, including their political preferences[16] and sexual orientation[17,18]. Even coarsening the spatial and temporal characteristics of individual locations traces provides little privacy protection, due to the idiosyncratic nature of human movements throughout daily life[19]. There has thus been widespread criticism and controversy about the use of such data by governments[20], private companies[16], civil society[21], and researchers[22].

To mitigate such personal inferences about individuals from digitally-derived mobility data – and the potential harms that could follow from these revelations – active measures must be taken to ensure that individual privacy is protected when such data are analyzed or released. The traditional approach to protecting privacy is to release only aggregated statistics about group movement rather than individual mobility trajectories[23]. However, such aggregation is typically insufficient to de-identify individuals in mobility traces[19,24], so other papers have turned to other techniques including distorting spatiotemporal information[25,26], obfuscating individual trajectories[27], and data slicing and mixing[28].

This paper develops and tests an approach to producing *provably private* mobility statistics from personal data based on differential privacy[29]. A number of previous studies have shown how differentially private approaches can add carefully calibrated noise to aggregated mobility statistics to ensure that the amount of individual information leaked by such statistics is bounded[30–39]. In comparison to past work, the focus of our study is on understanding the implications and tradeoffs that arise when using private mobility statistics in humanitarian settings. We are specifically interested in how differential privacy impacts *intervention accuracy*, i.e., the accuracy (and effectiveness) of policy decisions that are based on private mobility data rather than standard (non-private) data.

This paper makes three main contributions. First, we develop an algorithm for computing differentially private mobility information, and derive the guarantees that this algorithm provides for both data privacy and intervention accuracy. Second, we provide experimental results on the tradeoff between privacy and the accuracy of downstream policy decision in two real-world humanitarian contexts: response to (1) pandemics and (2) natural disasters and violent events. Third, we discuss the nuanced implementation choices that policymakers and algorithm developers are likely to grapple with in deploying such systems, and their effects on the privacy-accuracy tradeoff and ability to effectively deliver humanitarian aid.

Results

Building and Testing a Differentially Private Mobility Matrix

Our first set of results formalize the notion of a *private mobility matrix*, propose a novel algorithm for generating such matrices, and derive guarantees on the privacy and accuracy provided by this algorithm. In general, a *mobility matrix* is a data structure that quantifies the extent to which a population moves between regions in a fixed period of time – see Figure 1A for an example of a an O-D matrix derived from mobile phone records. The most common form of mobility matrix, an origin-destination matrix (hereafter abbreviated as an O-D matrix), indicates for each pair of regions (i, j) the number of trips taken by individuals from i to j. Mobility matrices are widely used by scholars and policymakers, and are playing an increasingly critical role in many humanitarian settings[1, 2, 12, 13]. However, as discussed above, the raw data in mobility matrices are sensitive, and can compromise the privacy of the individuals whose data are used to construct those matrices[33, 40].

Algorithm 1, described in detail in the "Methods" section, provides a computationally efficient method to construct *differentially private*[29] origin-destination matrices from individual mobility data. At its core, this algorithm adds very carefully calibrated noise to the underlying data, in a specific way that both protects the privacy of individuals and ensures the fidelity of the resulting O-D matrix. In the *Supplementary Information*, we prove that our algorithm satisfies ε-differential privacy (Theorem 1, *Supplementary Information*). This proof mathematically guarantees that no one with access to the private O-D matrix could infer "too much" about any individual trip (or any individual) whose data was contained in the matrix. The algorithm quantifies "too much" using a privacy loss parameter $\varepsilon \geq 0$ (Definition 1, *Supplementary Information*). This privacy loss parameter ε is thus critical to understanding the function of the algorithm, as it governs the likelihood that information about an individual trip (or an individual person) is leaked by a private mobility matrix. Smaller values of ε provide more robust privacy guarantees while larger values increase the risk that individual data could be compromised.

However, a central concern — and a focal point of our work — is that increases in privacy come at a cost. In particular, as ε decreases, the resulting private mobility matrix looks less and less like the true (non-private) mobility matrix. In the *Supplementary Information*, we derive formal guarantees of the worst-case accuracy loss of the private O-D matrix algorithm as a function of ε. In particular, Theorems 2, 3, and 4 show that the accuracy loss of our algorithm exponentially decreases as ε increases; Theorem 5 further shows that, when using multiple private mobility matrices to estimate changes in population flows, the detection error decreases according to $\Theta(\varepsilon \exp(-\varepsilon))$ as ε increases.

To provide intuition for the tradeoff between privacy and accuracy, Figure 2 illustrates how the privacy loss parameter ε relates to differences between the original O-D matrix and the private O-D matrix. To construct the figure, we use a large dataset of mobile phone records provided by one of the largest mobile network providers in Afghanistan, covering roughly 4 million individuals and containing over 3 billion phone transactions over 305 days (for details about the data used in this paper, see Table 1). We use our algorithm to construct a private O-D matrix from these data, and compare this private matrix to a non-private O-D matrix constructed from the raw data. In the figure, we observe that the errors introduced by privatization are quite small. For $\varepsilon = 0.1$ the median absolute error (i.e., the median difference between any given private matrix count and the corresponding non-private matrix count) is 6 trips at the admin-2, and the median relative error (the absolute error divided by the non-private matrix count) is 2.54% at the admin-2 level. For $\varepsilon = 1$ — a weaker level of privacy protection — the median absolute error is 0 trips and the median relative error is under 0.01% at the admin-2.

How does privatizing mobility data affect humanitarian interventions based on those data?

The preceding results suggest that our algorithm introduces only modest errors when creating a private version of the O-D matrix. However, it is not clear if and how those modest decreases will impact downstream policy decisions based on private O-D matrices. Our next set of results therefore illustrate how policies in two key humanitarian settings are affected by the privatization of mobility data. Specifically, we evaluate private O-D matrices in the context of a hypothetical pandemic scenario in Afghanistan, and in natural disaster scenarios in Afghanistan and Rwanda. These case studies leverage location data recorded in call detail records (CDR) collected by mobile phone operators (see Table 1) — the same sort of data that has raised privacy concerns in a wide range of contexts discussed in the Introduction[19–22]. The raw CDR indicate, for each phone call made on the operator's network, the date, time, and duration of the call, identifiers for each subscriber involved in each call, and the location of the cell tower through which each call was placed. This last piece of information is what makes it possible to infer the approximate location of each originating subscriber at the time of the phone call. From these raw location trace data, we

calculate daily O-D matrices recording daily movements between each pair of administrative subdivisions in each country (see Figure 1C for an example), and then show how those matrices – and the private version of them – influence downstream humanitarian interventions.

Non-pharmaceutical interventions in pandemic response

Our first case study compares the effectiveness of private and non-private mobility data in guiding public health interventions intended to stop the spread of contagious disease. This is an increasingly common application of mobility data, and an approach that was used to inform non-pharmeceutical interventions to the COVID-19 pandemic[10–13]. Our analysis simulates a pandemic scenario, where the spread of the disease is modeled using a classic SIR (Susceptible, Infected, Recovered) model[41], adapted to account for the fact that many modern diseases are influenced by human mobility[42]. We consider policy decisions that may be enacted when disease prevalence estimates from the mobility-informed SIR model exceed a certain threshold (such as travel restrictions, medication distribution, and so on), and test the extent to which the timing of such decisions differs when private and non-private O-D matrices are used for mobility modeling.

To conduct these simulations, we use the same large dataset of mobile phone records from Afghanistan in 2020 used to calculate errors in matrix counts from privatization (Figure 2). Daily O-D matrices at the province (admin-2) and district (admin-3) level derived from the CDR for 305 days are used as inputs to the mobility-informed SIR model, assuming that mobility data is available on a daily basis (but epidemiological data is available only at the start of the pandemic). We simulate three possible pandemics — pandemics initiating in Hirat, in Kabul, and in a randomly selected origin region — with 1% of individuals in the original region infected on the first day. In these simulations, we measure intervention accuracy by studying the daily (binary) decisions policymakers would make about whether or not to institute local anti-contagion policies, comparing the decisions that would be made with private data to those that would be made with non-private data. We measure the accuracy, precision, and recall of these binary decisions, with the assumption that policies are put into place at a threshold of 20% estimated disease prevalence.

As shown in Table 2, we find that intervention accuracy remains fairly high when using differentially private O-D matrices at the admin-2 (province) level (e.g. accuracy of policy decisions = 96-99% for $\varepsilon = 0.5$). However, when simulations are conducted with a finer spatial granularity — modeling disease spread at the district, or admin-3, level — intervention accuracy drops, and there is higher accuracy variance across pandemic simulations (e.g. accuracy of policy decisions = 72-91% for $\varepsilon = 0.5$). This difference highlights the importance of implementation choices like spatial granularity in deployments of the private O-D matrix algorithm: in general, the accuracy loss incurred by privatization will be larger at high spatial resolutions, when individual entries in the O-D matrix counts are smaller.

The value of the privacy loss parameter ε also has important implications for intervention accuracy. Higher values of ε — which correspond to less noise being added to mobility matrices in the private O-D matrix algorithm — lead to higher intervention accuracy (e.g. accuracy of policy decisions = 97-99% at the admin-2 level and 77-91% at the admin-3 level for $\varepsilon = 1.0$). Lower values of ε — meaning more noise added to mobility matrices — lead to lower intervention accuracy (e.g. accuracy of policy decisions = 90-98% at the admin-2 level and 64-83% at the admin-3 level for $\varepsilon = 0.1$). The choice of ε ultimately comes down to a policy decision on the relative importances of privacy and accuracy; we discuss methods for tuning the privacy-accuracy tradeoff at the end of the results section.

Geographic targeting of humanitarian aid after natural disasters and violent events

Our second case study illustrates how the privatization of mobility data can impact the allocation of humanitarian aid (such as medical resources, temporary shelter, and food) following natural disasters and other shocks. This application is motivated by the fact that mobility data derived from mobile phones and social media have been used to inform post-disaster humanitarian interventions, including following the 2010 earthquake in Haiti[4], the 2019 wildfires in California[43], and most recently during the 2023 earthquake in Turkey[44]. In our hypothetical aid-targeting scenario we simulate a short-term shock — such as an earthquake, landslide, flood, or violent event — and measure the degree of out-migration (as the total out-trips observed in an O-D matrix) in the following week. We assume a government or other policy actor seeks to provide humanitarian aid to the migrants in the areas with the most out-migration.

For our aid targeting simulations, we use two CDR datasets related to real-world natural disasters and violent events in simulating the implications of our private O-D matrix algorithm for targeting humanitarian aid. In Afghanistan, we study the battle of Kunduz in 2015, which resulted in the displacement of more than 100,000 people according to official statistics[45]. We focus on displacement in the week following the largest single day of the battle, using a CDR dataset of around 64 million transactions associated with 2.8 million mobile subscribers (Table 1). In Rwanda, we study mobility in the week following the Lake Kivu Earthquake of 2008, using a CDR dataset that includes around 13 million transactions associated with half a million mobile subscribers (Table 1). In these settings we measure intervention accuracy using two metrics: first, we compare estimated total out-migration from the affected area using private data to non-private data, and second we calculate the accuracy of the top 3 regions of out-migration identified in private data in comparison to those identified in non-private data.

In Table 3 we show that in both the Afghanistan and Rwanda contexts, private O-D matrices have relatively low error in determining total out-migration following natural disasters and violent events in comparison to non-private data (e.g. percent error = 2.54-8.27% for $\varepsilon = 0.5$). All private counts are biased downwards relative to non-private counts due to suppression of small values in private O-D matrices (Figure S2, *Supplementary Information*). As in the simulations of non-pharmaceutical interventions in pandemic response, we observe that intervention accuracy using private mobility matrices is higher at the admin-2 level (percent error = 2.54-7.38% for $\varepsilon = 0.5$) than at the admin-3 level (percent error = 4.98-8.27% for $\varepsilon = 0.5$). Unintuitively, however, we actually observe *higher* intervention accuracy for small values of ε (which correspond to large amounts of noise added to mobility matrices) than for large values of ε (which correspond to adding large amounts of noise). This arises from the use of a suppression threshold. By Theorem 2 (*Supplementary Information*), a non-private count below the threshold is more likely to become non-suppressed in Step 3 of the algorithm when a smaller ε is used compared to a larger one, thereby making the total out-migration flow more accurate for $\varepsilon = 0.1$ compared to $\varepsilon \in \{0.5, 1\}$.

Private O-D matrices also have relatively low error in identifying the areas of largest out-migration for targeting humanitarian aid. We focus in our simulations on selecting the top 3 geographies with the most out-migration for targeting post-disaster resources. We find that selection with private data is 90-100% accurate for $\varepsilon = 0.5$ (Table 3). Unlike in measuring total out-migration, we observe little sensitivity to the spatial granularity of simulation or the value of ε in identifying the top regions of out-migration. This result may be at least partially explained by the fact that identifying the top regions of out-migration from a relatively small number of regions with large out-migration flows is a less nuanced task than identifying the exact degree of outflow. In general we also do not observe much sensitivity to the number of regions targeted for humanitarian aid (for example, if 5 or 10 regions are identified instead of 3, as shown in Figure S3 (*Supplementary Information*).

How should policymakers navigate the tradeoff between privacy and accuracy?

Our main results show that policymakers can switch from non-private to private mobility matrices — which guarantee a certain degree of protection of individual mobility traces — with relative little accuracy loss. For example, in simulating the accuracy of nonpharmaceutical interventions in mobility-informed pandemic response, we found an accuracy of 72-99% for policies enacted using private mobility matrices. Accuracy is similar for policy decisions related to mobility-informed targeting of humanitarian resources following natural disasters, at 77-99%. However, our results also underscore that there is a key tradeoff between privacy and accuracy (shown in Figure 2 and made explicit in Theorems 1-5, *Supplementary Information*). In the case of the private O-D matrix algorithm, this tradeoff is quantified by the privacy loss parameter ε.

What is the benefit of increasing privacy by reducing ε? So far, we have shown that increasing ε generally comes with a decrease in accuracy for downstream policy interventions. All else being equal, we would like to maintain the highest standard of accuracy in policy interventions informed by mobility data — but doing so comes at a loss to the privacy of individual mobility traces. In this section, we use properties of differential privacy (Lemmas 2 and 3, *Supplementary Information*) to derive worst-case privacy guarantees for the policy intervention simulations from the previous section. We then compare these worst-case privacy guarantees to the expected individual-level privacy loss, and discuss practical methods for policymakers to determine the appropriate tradeoff between privacy and accuracy.

Worst Case vs. Expected Privacy Loss

The privacy guarantees for our differentially private O-D matrix algorithm (Lemmas 2 and 3, *Supplementary Information*) suggest that worst-case privacy loss for individuals is quite high — but our empirical data indicate that expected privacy loss is substantially lower. For example, in the pandemic simulation in Afghanistan, 305 private mobility matrices are released. The theoretical worst-case privacy loss for an individual over the entire simulation is $\varepsilon(n_1 + ... + n_{305})$, where n_j is the number of inter-region trips an individual took on day j that appear in the CDR. Likewise, in our case studies of the Lake Kivu Earthquake in Rwanda and the Battle of Kunduz in Afghanistan, where 7 days of O-D matrices are released, the theoretical worst-case privacy loss over the entire simulation is $\varepsilon(n_1 + ... + n_7)$, where n_j is the number of inter-region trips an individual took on day j that appear in the CDR.

In both case studies, we find that the actual privacy loss is much lower than the worst-case characterization. In the Afghanistan dataset used in the pandemic case study, the average individual makes 14 province-to-province trips and 52 district-to-district trips over the course of the simulation; these trips are aggregated in mobility matrices (Table 1), so the average individual-level privacy loss is 14ε at the admin-2 level and 52ε at the admin-3 level. For a value of $\varepsilon = 0.5$, this corresponds to a total privacy loss of 7 for provinces and 26 for districts. In our case study of the Lake Kivu earthquake in Rwanda, only seven days of mobility matrices are released, and the average individual makes 1.53 trips at an admin-2 level and 2.88 trips at an admin-3 level (Table 1). For a value of $\varepsilon = 0.5$, this corresponds to a total privacy loss of 0.77 for provinces and 1.44 for districts. Privacy loss in our case study of the Battle of Kunduz in Afghanistan is expected to be even smaller: over the seven days of data released, the average individual makes 0.26 trips at an admin-2 level and 1.73 trips at an admin-3 level (Table 1). For a value of $\varepsilon = 0.5$, this corresponds to a total privacy loss of 0.13 for provinces and 0.87 for districts.

For comparison, Google's COVID-19 Community Mobility Reports incurs a total privacy loss of 2.64 per day[32]. If Google's COVID-19 Community Mobility Reports were released for seven days, as in our disaster response scenarios, the total privacy loss would accumulate to 18.48. If Google's COVID-19 Community Mobility Reports were released for 305 days, as in our pandemic response scenario, the total privacy loss would accumulate to 805.2. Thus the expected privacy loss in our setting is well within the range of privacy loss in previously-deployed settings using differential privacy on personal mobility data.

Methods to tune the privacy-accuracy tradeoff

How should policymakers choose a value of ε to navigate the privacy-accuracy tradeoff? Across applications of differential privacy, there is no consensus on the "right" way to set ε[46–49]. In our setting, where policymakers may have a maximum tolerance for accuracy loss, one might imagine using estimates of intervention accuracy loss from the datasets being released to set ε. For instance, consider an aid organization seeking to target relief to the top-3 Admin-3 locations with the highest out-migration during the Battle of Kunduz. Suppose the aid organization wants to provide the strongest privacy protections possible, provided the algorithm has at least 90% intervention accuracy. The aid organization generates Table 3 and chooses the smallest value of ε that yields 90% accuracy, namely $\varepsilon = 0.5$. Then, the aid organization uses our algorithm with $\varepsilon = 0.5$ to generate private O-D matrices to determine the 3 locations to receive relief. (Note that this approach, while intuitively appealing, has a subtle caveat: technically, the value of ε selected can itself leak privacy when chosen based on the data[49]. One approach to address this would be to compute private mobility matrices on pre-program data, set ε based on simulations on those data, before applying the chosen ε to the deployment data).

Alternatively, policymakers could set ε using heuristics based on the our algorithm. For example, one simple heuristic approach is to determine ε based on a policymaker's tolerance for error of an O-D matrix count using the typical noise introduced by Step 2 of Algorithm 1. The standard deviation of the Laplace distribution with $\lambda = \varepsilon^{-1}T$ is given by $T\sqrt{2}/\varepsilon$, so we can bound on the set of feasible ε for a given O-D matrix count accuracy α, as $\alpha \geq \sqrt{2}T/\varepsilon \iff \varepsilon \geq \sqrt{2}T/\alpha$. For example, if policymakers require that the typical error in each entry of the O-D matrix is at most 10, then then the can set $\varepsilon \geq \sqrt{2}/10 \approx 0.14$. However, if policymakers instead can tolerate a typical error of at most 50, then any $\varepsilon \geq \sqrt{2}/50 \approx 0.028$ will suffice. While simple, this heuristic has two drawbacks: First, it ignores the impact of the post-processing steps of our algorithm; And second, this heuristic only utilizes the first two moments of the Laplace distribution to find a rough balance between the privacy level ε and the typical error.

A final option is to use the privacy-accuracy tradeoff theorems provided in the *Supplementary Information* to configure the algorithm to behave optimally. Theorems 6 and 7 (*Supplementary Information*) provide explicit formulas for policymakers to set ε based on context-specific needs. For example, suppose policymakers require 95% confidence that the error in an entry of the private O-D matrix be at most 10. Using Theorem 6, the policymakers can utilize any $\varepsilon \geq -(10.5)^{-1}\ln(0.05) \approx 0.285$. Compared to the value of 0.14 from the simple heuristic above, Theorem 6 enables policymakers to achieve a similar goal of bounding errors in each cell by 10 while utilizing a larger ε, thereby enabling the construction of more accurate private matrices. Theorem 7 can be similarly utilized to readily translate a policymaker's tolerance for error in terms of the parameter ε.

Discussion

In this paper, we presented a differentially private algorithm for releasing O-D matrices to learn about the mobility of a population without learning "too much" about any individual (or any individual trip). We analytically derived closed-form expressions for the privacy-accuracy tradeoff and showed that our algorithm can be configured to produce accurate private mobility matrices for a wide-range of parameter settings.

We then tested the performance of our differentially private O-D matrix algorithm using mobility data from two humanitarian settings. In these settings, we compare policy decisions made using the original mobility data to decisions made from a privatized version of our data. This allows us to empirically characterize the privacy-accuracy tradeoff, and calibrate privacy in a way that allows the policymaker to still make effective decisions. Our closing discussion provides more general and practical guidance for calibrating privacy (i.e., setting ε) based on the policymaker's tolerance for uncertainty.

As digital devices proliferate throughout the developing world, the data generated by those devices are increasingly being used to advance the Sustainable Development Goals. In these and other "social good" applications, it is imperative to consider and respect the privacy of the individuals behind the data. Our algorithm, analysis, and discussion highlights one way to provide strong privacy guarantees while still allowing for policymakers to make informed decisions. We hope future work can improve and expand upon these methods, to provide a more robust set of options for using private data in effective humanitarian response.

Methods

Differentially private O-D matrices

The main text provides an intuitive description of the differentially private O-D matrix algorithm. Here we formalize the algorithm, and describe a number of accuracy and privacy guarantees. We defer the formal presentation of these statements, as well as their proofs, to the *Supplementary Information*.

Our Private O-D Matrix algorithm (Algorithm 1) requires four inputs: the dataset d which is used to construct the private O-D matrix, a value of ε to control the privacy-accuracy tradeoff in the private matrix, a value T which has two interpretations based on whether trip-level differential privacy or individual-level differential privacy is required, and a threshold parameter $\tau \in \mathbb{Z}_{\geq 0}$ that ensures the noisy matrix entries are non-negative and enables suppression of small counts when desired. We describe these last two parameters in more detail below.

When individual-level differential privacy is required, T represents the maximum number of trips that any one individual can contribute to the mobility matrix. In order to ensure that the dataset d does not contain more than T trips for any individual, a preprocessing transformation can be applied to censor additional trips. One such algorithm that does so runs as follows. For each individual, the algorithm determines the number of trips θ they took. If $\theta > T$, the algorithm selects T trips at uniformly random to keep and drops the remaining $\theta - T$ trips from the dataset. If $\theta \leq T$, the algorithm leaves all θ trips in d. Alternatively, when trip-level differential privacy is required, setting $T = 1$ achieves this requirement, as the presence or absence of any trip changes any count in the O-D matrix by 1.

For any choice of $\tau \in \mathbb{Z}_{\geq 0}$, the entries of the private matrix are guaranteed to be non-negative. The magnitude of τ determines the level of suppression present in the private output. When $\tau = 0$, there is no suppression of the values in the matrix. As we increase τ, we increase the set of small counts to be suppressed. Prior to provable privacy techniques, it was common practice in statistical disclosure control to suppress information derived from a small number of individuals[50]. While suppression alone offers no provable privacy guarantee, it can inform end-users that the statistic may be unreliable due to the limitations in the data collection methods, such as having low phone coverage in a particular region. Additionally, for certain types of humanitarian responses – such as those involving disease spread – regions with small mobility counts are likely to provide fewer vectors for disease spread (and hence could be less relevant for mitigation strategies) compared to those regions with higher mobility counts. Consequently, suppression may be a useful signalling apparatus to policymakers.

Algorithm 1: Private O-D Matrix

Input:

- Dataset d

- Privacy parameter $\varepsilon \geq 0$

- Trip threshold $T \in \mathbb{Z}_{\geq 1}$ that corresponds to the maximum number of trips an individual can contribute when individual-level differential privacy is required. When trip-level differential privacy is required, set $T = 1$.

- Suppression threshold $\tau \in \mathbb{Z}_{\geq 0}$

Output: ε-differentially private O-D matrix $\hat{M}(d)$

1 (*Non-private computation*) For a geographic region with k zones, compute the $k \times k$ matrix $M(d)$, where $M(d)_{a,b}$ equals the number of trips from zone a to zone b when $a \neq b$, and 0 when $a = b$.

2 (*Incorporate privacy*) Add independent Laplace noise with scale $\lambda = \varepsilon^{-1} T$ to each non-diagonal entry of $M(d)$. Call this resulting matrix $M'(d)$.

3 (*Post-processing*) Next,
 (I) Round each of the resulting entries in $M'(d)$ to the nearest integer. In the event where there is no unique nearest integer, we adopt the convention of rounding up to the nearest integer.
 (II) Then, for all resulting values that are below a threshold of τ, map these values to 0.
 Call this resulting matrix $\hat{M}(d)$

4 **Return** $\hat{M}(d)$

This algorithm is accompanied with a several provable guarantees. For starters, our algorithm satisfies ε-differential privacy (Theorem 1, *Supplementary Information*). In addition to the privacy guarantee, our algorithm has strong accuracy guarantees, which we call the *privacy-accuracy tradeoff theorems*. With high probability: (1) the algorithm preserves cells that would have been suppressed absent differential privacy; (2) the algorithm does not suppress cells that would not have been suppressed absent differential privacy; (3) for non-suppressed cells, the algorithm's error decays exponentially as ε increases; and (4) for

non-suppressed cells, when examining the difference in cell counts across different time periods, the algorithm's error decreases according to $\Theta(\varepsilon \exp(-\varepsilon))$ as ε increases. These statements are made rigorous in Theorems 2, 3, 4, and 5 respectively in the *Supplementary Information*. Theorems 6 and 7 utilize Theorems 4 and 5 to provide policymakers with exact formulas to calculate ε based on their context-specific needs.

Empirical Simulations

Data

Our simulations of using the private O-D matrix algorithm for real-world policy decisions rely on location traces derived from mobile phone metadata (call detail records, or CDR). We use three call detail record datasets: one from Afghanistan in 2015, one from Afghanistan in 2020, and one from Rwanda in 2008. Table 1 summarizes these data. Our dataset from Afghanistan in 2020 is 305 days long and contains transactions from around 7 million subscribers; the remaining two datasets are only 7 days long (as the focus on specific natural disasters or violent events) and contain data for half a million (Rwanda) and around three million (Afghanistan) subscribers, respectively. In each dataset the CDR include the date, time, and duration of each call placed on the mobile phone networks, along with pseudonymized IDs for the originating subscriber and the recipient. They also record the cell tower through which each call was placed, providing a measure of the approximate location of each originating subscriber. We derive daily O-D matrices at the admin-2 and admin-3 level for each dataset by geolocating cell towers to administrative subdivisions, and then counting the number of trips observed in the data on a daily basis (where a trip occurs if a subscriber places a call or text in one administrative subdivision and a subsequent call or text in a different subdivision). Across all simulations, we use a suppression threshold of 15.

Non-pharmaceutical interventions in pandemic response

For this case study, we rely on the mobility-based SIR (Susceptible, Infected, and Recovered) model introduced by Goel and Sharma[42], which adapts the classic SIR model of Kermck and McKendrick[41] by including features to model human mobility. In the mobility-based SIR model, at any time t, individuals in region $i \in [k]$ are categorized into one of three groups: they are either susceptible to the disease, infected with the disease, or recovered from the disease. The number of individuals in these groups is given by $S_i(t), I_i(t)$, and $R_i(t)$ respectively. We denote the population of region i at time t as $N_i(t) = S_i(t) + I_i(t) + R_i(t)$. As time goes on, the number of individuals in these compartments can change. The epidemic dynamics are a function of both intra- and inter-region disease spread, described by the following differential equations involving a $k \times k$ O-D matrix M, that are governed by three parameters: β, α, and μ.

$$
\frac{dS_i(t)}{dt} = -\frac{\beta S_i(t) I_i(t)}{N_i(t)} - \frac{\alpha \beta S_i(t) \sum_{j \in [k]} \frac{M_{i,j} I_j(t)}{N_j(t)}}{N_i(t) + \sum_{j \in [k]} M_{i,j}}
$$

$$
\frac{dR_i(t)}{dt} = \mu \frac{I_i(t)}{N_i(t)}
$$

$$
\frac{dI_i(t)}{dt} = \frac{dS_i(t)}{dt} - \frac{dR_i(t)}{dt}
$$

The parameter β describes the frequency of mixing of intra-region populations in combination with the virality of the disease: how likely is an infected person to meet a non-infected person from the same region and transfer the pathogen? The α term in the differential equations describes the frequency of mixing with inter-region visitors in relation to the frequency of mixing with intra-region residents: how likely is an infected person to meet a visitor and transfer the pathogen? The last parameter μ describes the rate of recovery from disease: what proportion of the infected population recovers each day?

In principle, β, α, and μ could all vary from region to region on the basis of a number of factors (for example, community cohesion, access to healthcare in the area, and mobility restrictions), but for simplicity we keep them constant across regions. To configure this model, we set $\alpha = 1$, meaning that there is no difference between the degree of mixing with residents and the degree of mixing with visitors. We additionally set β and μ based on estimated dynamics from the COVID-19 pandemic. SIR estimates of β range between 0.06 and 0.39 in SIR models fit in 2020 in a set of nations; estimates of μ range between 0.04 and 0.19 in the same set of studies[51–53]. For this case study, we set $\beta = 0.10$ and $\mu = 0.04$.

The mobility-based SIR model relies additionally on mobility matrices M and a time-invariant regional population values $N_1, ..., N_k$. To compare the intervention differences between the private and non-private regimes, we constructed non-private mobility matrices using the CDR from Afghanistan along with private versions with $\varepsilon \in \{0.1, 0.5, 1\}$ using our private O-D matrix algorithm at the trip level. We estimate the regional population values $N_1, ..., N_k$ by inferring a home location for each subscriber in the CDR dataset using the most common cell tower through which they placed calls between the hours of 8pm and 6am. These regional counts $N_1, ..., N_k$ could leak information about individuals, so differentially private estimates of $N_1, ..., N_k$ could in principle be constructed using the Laplace mechanism. However, since we are interested in focusing on the intervention

differences that arise from the use of private mobility matrices, we omit doing so to keep an apples-to-apples comparison of the effect induced solely from the private mobility matrix.

As described in the results section, we use the mobility-based SIR model with daily origin-destination matrices to calculate epidemic curves (Figure S1, *Supplementary Information*), and compare the accuracy of policy decisions taken when SIR models are deployed with private and non-private OD matrices.

Geographic targeting of humanitarian aid after natural disasters and violent events

Our second simulated policy context is the targeting of humanitarian aid after natural disasters, violent events, and other shocks. As described in the results section, we calculate daily O-D matrices for time periods associated with two such shocks in our datasets: the Battle of Kunduz in Afghanistan in 2015, and the Lake Kivu Earthquake in Rwanda in 2008. We calculate daily O-D matrices for seven days following each event and calculate the two statistics using both private and non-private O-D matrices: (1) the total amount of out migration from the affected area, and (2) the top-*k* districts with the most out migration. We assess the accuracy of both statistics using private data in comparison to non-private data. We also assess the sensitivity of these results to the choice of *k* (Figure S3, *Supplementary Information*).

References

1. Wesolowski, A. *et al.* Impact of human mobility on the emergence of dengue epidemics in Pakistan. *Proc. Natl. Acad. Sci.* **112**, 11887–11892 (2015).

2. Wesolowski, A. *et al.* Quantifying the impact of human mobility on malaria. *Science* **338**, 267–270 (2012).

3. Bengtsson, L. *et al.* Using mobile phone data to predict the spatial spread of cholera. *Sci. reports* **5**, 1–5 (2015).

4. Lu, X., Bengtsson, L. & Holme, P. Predictability of population displacement after the 2010 Haiti earthquake. *Proc. Natl. Acad. Sci.* **109**, 11576–11581 (2012).

5. Pastor-Escuredo, D. *et al.* Flooding through the lens of mobile phone activity. *Proc. 2014 IEEE Glob. Humanit. Technol. Conf. (GHTC)* (2014).

6. Aiken, E., Bellue, S., Karlan, D., Udry, C. R. & Blumenstock, J. Machine Learning and Mobile Phone Data Can Improve the Targeting of Humanitarian Assistance. Working Paper 29070, National Bureau of Economic Research (2021). DOI: 10.3386/w29070. Series: Working Paper Series.

7. Blumenstock, J. Machine learning can help get COVID-19 aid to those who need it most. *Nature* **581**, DOI: 10.1038/d41586-020-01393-7 (2020). Publisher: Nature Publishing Group.

8. Dobra, N., A. Williams & Eagle, N. Spatiotemporal detection of unusual human population behavior using mobile phone data. *PLoS ONE* **10** (2015).

9. Gundogdu, D., Incel, O., Salah, A. & Lepri, B. Countrywide arrhythmia: emergency event detection using mobile phone data. *EPJ Data Sci.* **5** (2016).

10. Kraemer, M. U. *et al.* The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* **368**, 493–497 (2020).

11. Warren, M. S. & Skillman, S. W. Mobility changes in response to COVID-19. *arXiv preprint arXiv:2003.14228* (2020).

12. Chinazzi, M. *et al.* The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* **368**, 395–400 (2020).

13. Ilin, C. *et al.* Public mobility data enables COVID-19 forecasting and management at local and global scales. *Sci. Reports* **11**, 1–11, DOI: 10.1038/s41598-021-92892-8 (2021). Cc_license_type: cc_by Number: 1 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Diseases;Health care Subject_term_id: diseases;health-care.

14. Lai, S. *et al.* Effect of non-pharmaceutical interventions to contain COVID-19 in China. *nature* **585**, 410–413 (2020).

15. Oliver, N. *et al.* Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. *Sci. Adv.* (2020).

16. Thompson, S. A. & Warzel, C. Twelve million phones, one dataset, zero privacy. *The New York Times* **19**, 2019 (2019).

17. Ovide, S. The Nightmare of Our Snooping Phones. *The New York Times* (2021).

18. Boorstein, M., Iati, M. & Shin, A. Top U.S. Catholic Church official resigns after cellphone data used to track him on Grindr and to gay bars. *The Wash. Post* (2021).

19. de Montjoye, Y.-A., Hidalgo, C., Verleysen, M. & Blondel, V. Unique in the crowd: The privacy bounds of human mobility. *Sci. Reports* **3** (2013).

20. Privacy & Board, C. L. O. Report on the government's use of the call detail records program under the usa freedom act (2020).

21. Hosein, G. & Nyst, C. Aiding surveillance: an exploration of how development and humanitarian aid initiatives are enabling surveillance in developing countries. *Available at SSRN 2326229* (2013).

22. Taylor, L. No place to hide? the ethics and analytics of tracking mobility using mobile phone data. *Environ. Plan. D: Soc. Space* **34** (2016).

23. de Montjoye, Y.-A. *et al.* On the privacy-conscientious use of mobile phone data. *Sci. data* **5**, 1–6 (2018).

24. Xu, F., Tu, Z., Zhang, P., Fu, X. & Jin, D. Trajectory recovery from ash: User privacy is not preserved in aggregated mobility data. In *Proceedings of the 26th World Wide Web Conference*, 1241–1250 (2017).

25. Abul, O., Bonchi, F. & Nanni, M. Never walk alone: Uncertainty for anonymity in moving objects databases. In *2008 IEEE 24th international conference on data engineering*, 376–385 (Ieee, 2008).

26. Primault, V., Mokhtar, S. B., Lauradoux, C. & Brunie, L. Time distortion anonymization for the publication of mobility data with high utility. In *2015 IEEE Trustcom/BigDataSE/ISPA*, vol. 1, 539–546 (IEEE, 2015).

27. Meyerowitz, J. & Roy Choudhury, R. Hiding stars with fireworks: location privacy through camouflage. In *Proceedings of the 15th annual international conference on Mobile computing and networking*, 345–356 (2009).

28. Shi, J., Zhang, R., Liu, Y. & Zhang, Y. Prisense: privacy-preserving data aggregation in people-centric urban sensing systems. In *2010 Proceedings IEEE INFOCOM*, 1–9 (IEEE, 2010).

29. Dwork, C., McSherry, F., Nissim, K. & Smith, A. Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference*, 265–284 (Springer, 2006).

30. Mir, D. J., Isaacman, S., Cáceres, R., Martonosi, M. & Wright, R. N. DP-Where: Differentially private modeling of human mobility. In *2013 IEEE international conference on big data*, 580–588 (IEEE, 2013).

31. Qardaji, W., Yang, W. & Li, N. Differentially private grids for geospatial data. In *2013 IEEE 29th international conference on data engineering (ICDE)*, 757–768 (IEEE, 2013).

32. Aktay, A. *et al.* Google covid-19 community mobility reports: Anonymization process description (version 1.0). *arXiv preprint arXiv:2004.04145* (2020).

33. Machanavajjhala, A., Kifer, D., Abowd, J., Gehrke, J. & Vilhuber, L. Privacy: Theory meets practice on the map. In *2008 IEEE 24th international conference on data engineering*, 277–286 (IEEE, 2008).

34. Pratesi, F. *et al.* Privacy-aware distributed mobility data analytics (2014).

35. Jiang, K., Shao, D., Bressan, S., Kister, T. & Tan, K.-L. Publishing trajectories with differential privacy guarantees. In *Proceedings of the 25th International conference on scientific and statistical database management*, 1–12 (2013).

36. Ghane, S., Kulik, L. & Ramamohanarao, K. Tgm: A generative mechanism for publishing trajectories with differential privacy. *IEEE Internet Things J.* **7**, 2611–2621 (2019).

37. Shao, D., Jiang, K., Kister, T., Bressan, S. & Tan, K.-L. Publishing trajectory with differential privacy: A priori vs. a posteriori sampling mechanisms. In *Database and Expert Systems Applications: 24th International Conference, DEXA 2013, Prague, Czech Republic, August 26-29, 2013. Proceedings, Part I 24*, 357–365 (Springer, 2013).

38. Li, M., Zhu, L., Zhang, Z. & Xu, R. Achieving differential privacy of trajectory data publishing in participatory sensing. *Inf. Sci.* **400**, 1–13 (2017).

39. Savi, M. K. *et al.* A standardised differential privacy framework for epidemiological modelling with mobile phone data. *medRxiv* 2023–03 (2023).

40. Shaham, S., Ghinita, G. & Shahabi, C. Differentially-Private Publication of Origin-Destination Matrices with Intermediate Stops. *arXiv preprint arXiv:2202.12342* (2022).

41. Kermack, W. O. & McKendrick, A. G. A contribution to the mathematical theory of epidemics. *Proc. royal society london. Ser. A, Containing papers a mathematical physical character* **115**, 700–721 (1927).

42. Goel, R. & Sharma, R. Mobility based sir model for pandemics-with case study of covid-19. In *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 110–117 (IEEE, 2020).

43. Schroader, A. The Kincade Fire: Population Movement and Displacement (2019).

44. CrisisReady. 7.8 and 7.5 Magnitude Earthquakes strike Turkey, Cause Mass Destruction and Growing Death Toll (2023).

45. Afghan forces struggle as Taliban seeks northern stronghold. *Mil. Times* (2015).

46. Dwork, C., Kohli, N. & Mulligan, D. Differential privacy in practice: Expose your epsilons! *J. Priv. Confidentiality* **9** (2019).

47. Naldi, M. & D'Acquisto, G. Differential privacy: An estimation theory-based method for choosing epsilon. *arXiv preprint arXiv:1510.00917* (2015).

48. Hsu, J. *et al.* Differential privacy: An economic method for choosing epsilon. In *2014 IEEE 27th Computer Security Foundations Symposium*, 398–410 (IEEE, 2014).

49. Kohli, N. & Laskowski, P. Epsilon voting: Mechanism design for parameter selection in differential privacy. In *2018 IEEE Symposium on Privacy-Aware Computing (PAC)*, 19–30 (IEEE, 2018).

50. Matthews, G. J., Harel, O. & Aseltine Jr, R. H. A review of statistical disclosure control techniques employed by web-based data query systems. *J. public health management practice: JPHMP* **23**, e1 (2017).

51. Bertozzi, A. L., Franco, E., Mohler, G., Short, M. B. & Sledge, D. The challenges of modeling and forecasting the spread of covid-19. *Proc. Natl. Acad. Sci.* **117**, 16732–16738 (2020).

52. Bagal, D. K., Rath, A., Barua, A. & Patnaik, D. Estimating the parameters of susceptible-infected-recovered model of covid-19 cases in india during lockdown periods. *Chaos, Solitons & Fractals* **140**, 110154 (2020).

53. Lounis, M. & Bagal, D. K. Estimation of SIR model's parameters of COVID-19 in Algeria. *Bull. Natl. Res. Centre* **44**, 1–6 (2020).

54. Dwork, C. & Roth, A. The algorithmic foundations of differential privacy. *Foundations Trends Theor. Comput. Sci.* **9**, 211–407 (2014).

55. Chatzigeorgiou, I. Bounds on the lambert function and their application to the outage analysis of user cooperation. *IEEE Commun. Lett.* **17**, 1505–1508 (2013).

Acknowledgements

Author contributions statement

NK and JB conceived of the research idea. NK developed the private mobility matrix algorithm and proved each of the theorems in the paper. EA implemented the empirical measurements of accuracy and simulations of intervention accuracy. NK, EA, and JB wrote the manuscript.

Additional information

The author(s) declare no competing interests.

Figures and Tables

A. Non-private mobility matrix
(January 15, 2015)

	Dih Sabz	Kabul	Kalakan	Mir Bacha Kot	Paghman	Qarabagh	Shakardara
Dih Sabz		1,639	28	16	3	36	56
Kabul	1,835		386	1,011	7,872	518	12,039
Kalakan	17	215		468	0	271	169
Mir Bacha Kot	15	781	489		15	188	719
Paghman	9	7,717	1	6		12	130
Qarabagh	41	400	293	175	6		196
Shakardara	52	10,948	165	870	135	208	

B. Privatization (Algorithm 1)

C. Private mobility matrix
(January 15, 2015), ε = 1

	Dih Sabz	Kabul	Kalakan	Mir Bacha Kot	Paghman	Qarabagh	Shakardara
Dih Sabz		1,639	37	41	26	----	55
Kabul	1,841		381	927	7,868	506	12,029
Kalakan	21	227		464	---	274	179
Mir Bacha Kot	16	776	501		31	203	721
Paghman	---	7,710	15	---		---	124
Qarabagh	36	394	295	177	27		204
Shakardara	55	10,942	155	874	137	208	

D. Privatized map of movement
(January 15, 2015), ε = 1

Trips in CDR
- 0-10
- 10-20
- 20-30
- 30-40
- 40-50
- 50-100
- 100-500
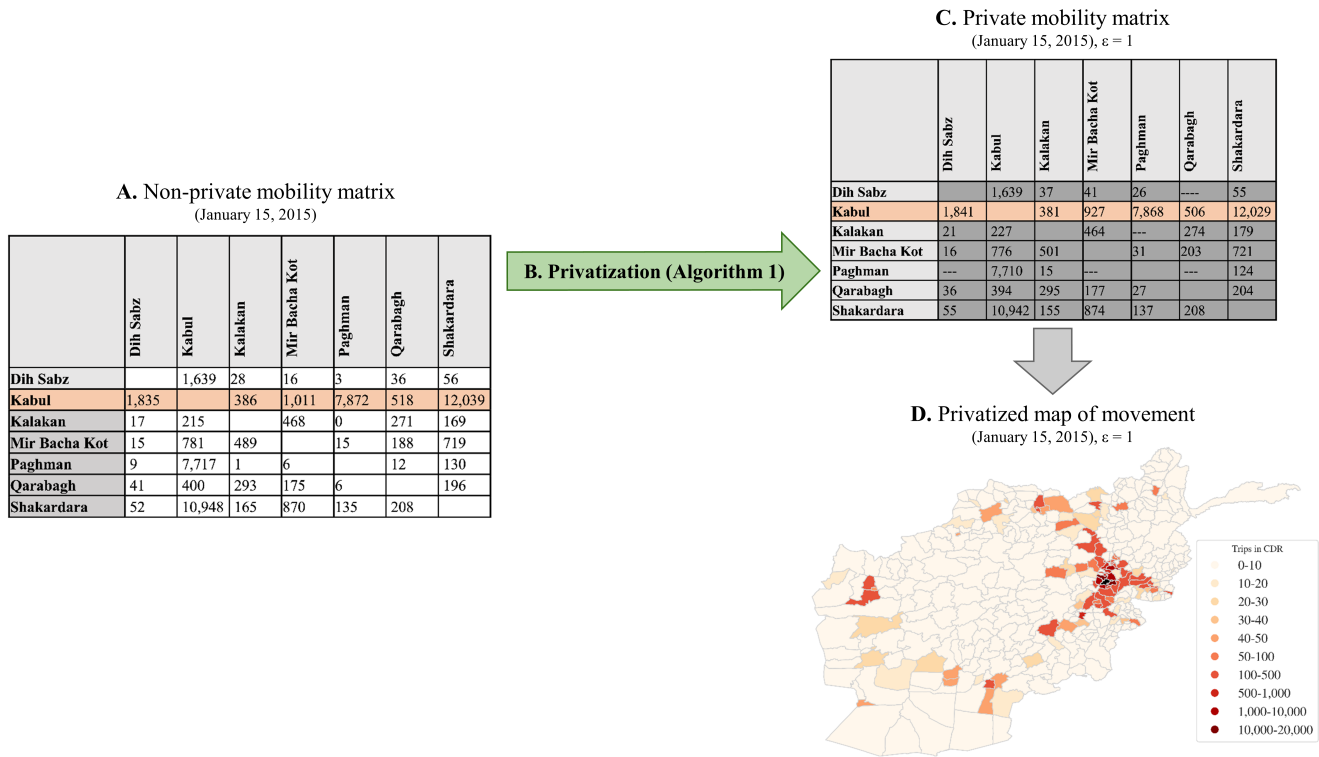- 500-1,000
- 1,000-10,000
- 10,000-20,000

Figure 1. Panel A: A portion of an origin-destination matrix, showing movement between 10 districts of Afghanistan on January 1, 2015, as calculated from data provided by a mobile phone operator. Panel B: The origin-destination matrix is passed into our private O-D matrix algorithm (as described in Algorithm 1) for trip-level protection (T =1) with privacy parameter $\varepsilon = 1$ and suppression threshold $\tau = 15$. Panel C: The same portion of the origin-destination matrix, this time after being privatized by our private O-D matrix algorithm (with $\varepsilon = 1$). Panel D: Movement from Kabul to other districts of Afghanistan on January 1, 2015, based on the highlighted row of the private origin-destination matrix in Panel B. Kabul, the origin district, is shown in black.
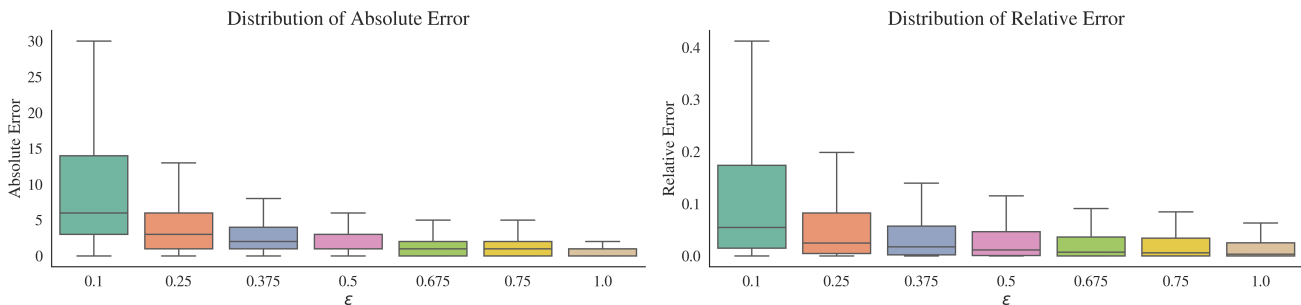

Figure 2. Left: Absolute errors in matrix entries for differentially private O-D matrices, relative to non-private matrices (derived over all days in our Afghanistan 2020 dataset and all origin-destination pairs at the admin-2 or province level). Absolute error is calculated as the difference between an O-D matrix count in the private matrix and the corresponding O-D matrix count in the non-private matrix. The distributions shown are taken over all 305 days in our 2020 CDR dataset (see Table 1) and all origin-destination province pairs. Right: Relative errors in matrix entries for differentially private O-D matrices, relative to non-private matrices. Relative error is calculated as twice the absolute error in a private O-D matrix count, divided by sum of the private and non-private O-D matrix counts for the same cell. Again, the distributions shown are taken over all 305 days in our 2020 CDR dataset (see Table 1) and all origin-destination province pairs.

Table 1. Summary of the call detail record datasets used in our empirical simulations.

	Rwanda 2008	Afghanistan 2015	Afghanistan 2020
Dates	Feb. 3 - Feb. 9, 2008	Sep. 28 - Oct. 4, 2015	Jan. 1 - Oct. 31, 2020
Number of days	7	7	305
Number of transactions	13 million	64 million	3.2 billion
Number of subscribers	541 thousand	2.79 million	7.12 million
Admin-2 level	30 districts	34 provinces	34 provinces
Admin-3 level	416 sectors	421 districts	421 districts
Mean trips per subscriber (admin-2 level)	1.53	0.26	14
Mean trips per subscriber (admin-3 level)	2.88	1.73	52.00
95th percentile of trips per subscriber (admin-2 level)	10	2	52
95th percentile of trips per subscriber (admin-3 level)	20	9	251

Table 2. Average errors in pandemic response (measured in accuracy, precision, recall) by degree of privacy introduced in the mobility matrix, at the admin-2 (left) and admin-3 level (right). Accuracy, precision, and recall are calculated for each province/district across all 305 timesteps in the simulation, and averages are taken across all provinces/districts. Standard deviations across provinces/districts are reported in parentheses.

	Admin-2 Level (Provinces)			Admin-3 Level (Districts)		
	Accuracy	Precision	Recall	Accuracy	Precision	Recall
Panel A: Pandemic initiating in Kabul						
Non-Private	100%	100%	100%	100%	100%	100%
$\varepsilon=0.1$	98% (2%)	94% (10%)	95% (10%)	83% (12%)	61% (27%)	64% (27%)
$\varepsilon=0.5$	99% (0%)	98% (3%)	98% (3%)	91% (8%)	79% (21%)	80% (20%)
$\varepsilon=1.0$	99% (0%)	98% (2%)	98% (2%)	91% (9%)	78% (22%)	80% (22%)
Panel B: Pandemic initiating in Hirat						
Non-Private	100%	100%	100%	100%	100%	100%
$\varepsilon=0.1$	90% (8%)	64% (33%)	63% (33%)	64% (12%)	16% (33%)	18% (29%)
$\varepsilon=0.5$	96% (5%)	85% (21%)	85% (21%)	72% (12%)	33% (29%)	35% (30%)
$\varepsilon=1.0$	97% (5%)	89% (20%)	89% (20%)	77% (11%)	44% (29%)	45% (29%)
Panel C: Pandemic initiating randomly						
Non-Private	100%	100%	100%	100%	100%	100%
$\varepsilon=0.1$	93% (8%)	74% (31%)	74% (32%)	71% (15%)	33% (36%)	34% (37%)
$\varepsilon=0.5$	97% (5%)	89% (19%)	89% (19%)	79% (13%)	52% (31%)	53% (32%)
$\varepsilon=1.0$	98% (3%)	93% (12%)	93% (12%)	82% (12%)	58% (30%)	59% (30%)

Table 3. Average errors in total migration counts and top-*k* regions of out-migration, by degree of privacy introduced in the mobility matrix, at the admin-2 level (left) and admin-3 level (right).

	Admin-2 level				Admin-3 level			
	Non-Private	$\varepsilon = 0.1$	$\varepsilon = 0.5$	$\varepsilon = 1$	*Non-Private*	$\varepsilon = 0.1$	$\varepsilon = 0.5$	$\varepsilon = 1$
Panel A: Battle of Kunduz in Afghanistan								
Total out-migration	49,994	49,001	48,725	48,722	87,007	83,327	82,671	82,609
Percent error in total out-migration	0.00%	1.98%	2.54%	2.61%	0.00%	4.23%	4.98%	5.05%
Accuracy of top-*k* regions ($k = 3$)	0.00%	90.47%	90.47%	90.47%	0.00%	85.71%	90.47%	90.47%
Panel B: Lake Kivu Earthquake in Rwanda								
Total out-migration	51,102	48,085	47,331	47,320	32,627	30,595	29,930	29,801
Percent error in total out-migration	0.00%	5.90%	7.38%	8.66%	0.00%	6.23%	8.27%	8.66%
Accuracy of top-*k* regions ($k = 3$)	0.00%	100.00%	100.00%	100.00%	0.00%	90.47%	95.24%	90.47%

Supplementary Information: Mathematical Details and Supplementary Figures

Mathematical Details

In this section, we provide the mathematical foundation of Private O-D Matrix Algorithm, as well as formalize and prove all the theorems referenced in the main manuscript. In order to do so, we begin by first presenting the necessary mathematical preliminaries on differential privacy. We then leverage these preliminaries to formalize the privacy guarantees of our algorithm. Lastly, we provide formal proofs of of our privacy-accuracy tradeoff theorems.

Mathematical Preliminaries

Differential privacy provides a mathematical guarantee that a privacy-preserving statistic $A(d)$ that approximates a non-private statistic $M(d)$ derived from dataset d will not reveal "too much" information about any underlying observation in dataset d[46].

Definition 1. (ε-Differential Privacy[29]) A randomized algorithm A mapping datasets in $\mathbb{D}$ to an outcome space $\mathbb{O}$ satisfies ε-differential privacy if, for all datasets $d, d' \in \mathbb{D}$ that differ in one element, and for all events $O \subseteq \mathbb{O}$, $\mathbb{P}(A(d) \in O) \leq \exp(\varepsilon)\mathbb{P}(A(d') \in O)$.

Thus the privacy-preserving statistic $A(d)$ is the output of a randomized algorithm A whose noisy-behavior is governed by the parameter ε.

It is worth noting that differential privacy is not a particular technique. Rather, it is a mathematical standard that an algorithm either satisfies or not[46]. As such, differential privacy provides us with a promise about the worst-case privacy loss incurred from running a computational process. As a general matter however, it does not tell us how to implement the promise. There are many possible ways to construct an algorithm to achieve the privacy guarantee. A foundational technique for privatizing a statistic $S : \mathbb{D} \to \mathbb{R}^m$ is to use noisy values sampled from a Laplace distribution[1] with a carefully crafted scale parameter λ (denoted as $Lap(\lambda)$) to ensure that contribution of *any* possible data point is masked.

Lemma 1 (Laplace Mechanism[29]). *For a statistic $S : \mathbb{D} \to \mathbb{R}^m$, let Δ_S be the supremum of $||S(d) - S(d')||_1$, where $d, d' \in \mathbb{D}$ are datasets that differ in one element. For any $\Delta_S \in (0, \infty)$, the algorithm $A(d) = S(d) + (\zeta_1, ..., \zeta_m)^T$, where each ζ_i is drawn independently from $Lap(\varepsilon^{-1}\Delta_S)$, satisfies ε-differential privacy.*

It is also worth highlighting that the definition of differential privacy quantifies over datasets that differ in one *element*. There are two ways to consider this elemental difference in our setting, and each has a different consequence for the level of privacy afforded. When the elemental difference in mobility datasets is a trip, differential privacy affords *trip-level privacy protection* by protecting each trip with parameter ε. Hence, if an individual took t trips, they experience $t\varepsilon$ privacy loss. This can result in *heterogeneous privacy loss* for data subjects, with individuals who travel more incurring larger privacy losses. Alternatively, when the elemental difference in datasets is an individual, differential privacy protects all the trips an individual took with parameter ε, yielding *homogeneous privacy loss* for all data subjects. This *individual-level privacy protection* is a stronger privacy guarantee, but requires more noise in order to hide multiple trips.

Differentially private algorithms are accompanied with a host of provable guarantees that enable rigorous reasoning about privacy losses that occur from the computation of statistics. Three important properties for this study are the *composition result*, the *post-processing result*, and the *group privacy*.

Lemma 2 (Composition[54]). *Given any collection of ε_i-differentially private algorithms $A_i : \mathbb{D} \to \mathbb{O}_i$, the algorithm $A = (A_1, ..., A_r) : \mathbb{D} \to \mathbb{O}_1 \times ... \times \mathbb{O}_r$ is $(\varepsilon_1 + ... + \varepsilon_r)$-differentially private.*

Lemma 3 (Group-Privacy[54]). *Suppose $d, d' \in \mathbb{D}$ are datasets that differ in g elements. Then any ε-differentially private algorithm $A : \mathbb{D} \to \mathbb{O}$ further guarantees that $\mathbb{P}(A(d) \in O) \leq \exp(g\varepsilon)\mathbb{P}(A(d') \in O)$ for all events $O \subseteq \mathbb{O}$.*

Lemma 4 (Post-Processing[54]). *Given an ε-differentially private algorithm $A : \mathbb{D} \to \mathbb{O}$ and any (potentially randomized) function $f : \mathbb{O} \to \mathbb{O}'$, $f \circ A$ is still ε-differentially private.*

Practically speaking, if humanitarian response requires multiple computations to be run on the data, then the composition result allows us to keep track of the total privacy loss accumulated by adding up the ε's used. Additionally, if the output of a differentially private analysis subsequently needs to be analyzed or manipulated in any way without the original data, then the post-processing result ensures no additionally privacy loss has occurred. And lastly, the group privacy result informs us that: (1) when trip-level privacy is used, the total privacy incurred scales linearly with the number of trips g that an individual took; and (2) when individual-level privacy is used, the total privacy incurred by a group of g individuals scales linearly as well.

---

[1]The Laplace distribution with scale parameter $\lambda \in (0, \infty)$ corresponds to a real-valued random variable with probability density function $f(z) = \frac{1}{2\lambda}\exp(-|z|/\lambda)$[29]. This distribution has mean 0 and a standard deviation of $\sqrt{2}\lambda$.

Privacy Guarantee of the O-D Matrix Algorithm

Using the mathematical tools from Section 4, we deduce that our algorithm satisfies ε-differential privacy.

Theorem 1. *For any value $\tau \in \mathbb{Z}_{\geq 0}$ that is selected independent of data, the Private O-D Matrix algorithm satisfies (1) individual-level ε-differential privacy for any $T \in \mathbb{Z}_{\geq 1}$ that is also determined independently of the underlying data, and (2) trip-level ε-differential privacy when $T = 1$.*

Proof. It is sufficient to privatize the non-diagonal entries of $M(d)$, as these are the only entries that are responsive to d. When individual-level privacy is required, the inclusion or deletion of any record in d can change the sum of the non-diagonal entries of $M(d)$ by $\pm T$. And when trip-level privacy is required, any record in d can change the sum of the non-diagonal entries of $M(d)$ by ± 1 (which equals $\pm T$ as per the algorithm). Hence, $\Delta_M = T$. Adding $k^2 - k$ independently sampled $Lap(\varepsilon^{-1}T)$ values to the non-diagonal entries of $M(d)$ is equivalent to reshaping the non-diagonal entries of $M(d)$ as a vector $S(d) \in \mathbb{R}^{k^2-k}$ and then utilizing the Laplace mechanism, yielding ε-differential privacy. After this, there is no additional privacy loss by the post-processing result, as τ was set agnostic of d. $\qquad\square$

Privacy-Accuracy Tradeoff Theorems

Our first two privacy-accuracy tradeoff theorems demonstrate that our algorithm preserves suppression and non-suppression of the non-private entries with high probability.

To do so, we consider a dataset d that produces a non-private matrix $M(d)$ and a private matrix $\hat{M}(d)$ using some threshold τ and some privacy parameter ε. Unless $\tau = 0$ we cannot directly compare these two outputs, as the private matrix will not contain any counts between 1 and $\tau - 1$ due to the suppression. For an apples-to-apples comparison, let $\bar{M}(d)$ be the result of mapping all entries in $M(d)$ that are below τ to 0.

Theorem 2. *For any $T \in \mathbb{Z}_{\geq 1}$, $\tau \in \mathbb{Z}_{\geq 0}$, and for any suppressed non-diagonal entry (a,b) in $\bar{M}(d)$, $\hat{M}(d)_{a,b}$ is also suppressed with probability $1 - 0.5\exp(-\varepsilon T^{-1}(\tau - 0.5 - M(d)_{a,b}))$.*

Proof. $\bar{M}(d)_{a,b}$ is suppressed $\iff M(d)_{a,b} < \tau$. For $\eta \sim Lap(\varepsilon^{-1}T)$, we have $\mathbb{P}(\hat{M}(d)_{a,b} = 0) = \mathbb{P}(\text{round}(M(d)_{a,b} + \eta) < \tau) = \mathbb{P}(\eta < \tau - 0.5 - M(d)_{a,b}) = 1 - 0.5\exp(-\varepsilon T^{-1}(\tau - 0.5 - M(d)_{a,b}))$. $\qquad\square$

Theorem 3. *For any $T \in \mathbb{Z}_{\geq 1}$, $\tau \in \mathbb{Z}_{\geq 0}$, and for any non-suppressed non-diagonal entry (a,b) of $\bar{M}(d)$, $\hat{M}(d)_{a,b}$ is also not suppressed with probability $1 - 0.5\exp(\varepsilon T^{-1}(\tau + 0.5 - M(d)_{a,b}))$.*

Proof. $\bar{M}(d)_{a,b}$ is not suppressed $\iff M(d)_{a,b} \geq \tau$. For $\eta \sim Lap(\varepsilon^{-1}T)$, $\mathbb{P}(\hat{M}(d)_{a,b} \neq 0) = \mathbb{P}(\text{round}(M(d)_{a,b} + \eta) \geq \tau) = \mathbb{P}(\eta \geq \tau + 0.5 - M(d)_{a,b}) = 1 - 0.5\exp(\varepsilon T^{-1}(\tau + 0.5 - M(d)_{a,b}))$. $\qquad\square$

Our third privacy-accuracy tradeoff theorem quantifies the probability that the error between a non-private matrix entry and private matrix entry exceeds α, in cases where the (a,b) entry is non-suppressed in both the non-private and private matrices.

Theorem 4. *Suppose (a,b) is a non-suppressed non-diagonal entry of both $M(d)$ and $\hat{M}(d)$. Then for any $\alpha \in \mathbb{Z}_{\geq 0}$, the chance that $|\hat{M}(d)_{a,b} - M(d)_{a,b}| > \alpha$ is $\exp(-\varepsilon T^{-1}(\alpha + 0.5))$.*

Proof. For $\eta \sim Lap(\varepsilon^{-1}T)$ and $\alpha \in \mathbb{Z}_{\geq 0}$, $\mathbb{P}(|\hat{M}(d)_{a,b} - M(d)_{a,b}| > \alpha) = \mathbb{P}(|\text{round}(\eta)| > \alpha) = 2\mathbb{P}(\eta \geq \alpha + 0.5) = \exp(-\varepsilon T^{-1}(\alpha + 0.5))$. $\qquad\square$

The previous three results can be viewed as demonstrating "static" accuracy guarantees. However, mobility matrices can be used to inform policy decisions based on differences in population flow at two points in time. Our privatization method can still allow for strong accuracy guarantees, even when looking at such differences. For any non-diagonal entry (a,b), let $M_{t_1}(d)_{a,b}$ and $M_{t_2}(d)_{a,b}$ be represent the non-private counts of trips from a to b at time periods t_1 and t_2 respectively. The following privacy-accuracy tradeoff theorem allows us to quantify the chance that the observed difference in the entry (a,b) in private matrices and the actual difference the (a,b) entry of non-private matrices is larger than α.

Theorem 5. *Suppose the counts in $M_{t_1}(d)_{a,b}, M_{t_2}(d)_{a,b}, \hat{M}_{t_1}(d)_{a,b}$, and $\hat{M}_{t_2}(d)_{a,b}$ are all non-suppressed. Then for any $\alpha \in \mathbb{Z}_{\geq 0}$ and $T \in \mathbb{Z}_{\geq 1}$, the chance that $|(\hat{M}_{t_2}(d)_{a,b} - \hat{M}_{t_1}(d)_{a,b}) - (M_{t_2}(d)_{a,b} - M_{t_1}(d)_{a,b})| > \alpha$ is given by*

$$\exp(-\varepsilon T^{-1}(\alpha + 1))\frac{\varepsilon T^{-1}(\alpha + 1) + 2}{2}$$

To streamline the proof of the Theorem 5, we first provide a helpful technical lemma.

Lemma 5. *For any $\alpha \in \mathbb{R}$ and independent $\eta_1, \eta_2 \sim Lap(\varepsilon^{-1}T)$,*

$$\mathbb{P}(\eta_2 - \eta_1 \geq \alpha) = \exp(-\text{sgn}(\alpha)\varepsilon\alpha T^{-1})\frac{\varepsilon\alpha T^{-1} + 2(\text{sgn}(\alpha) + \mathbb{I}(\alpha = 0))}{4}$$

Proof. Each $\eta_i \sim Lap(\varepsilon^{-1}T)$, so they can be written as the difference of two independent Exponential random variables with rate parameter εT^{-1}. So $\eta_2 - \eta_1$ can be written as the difference of independent Gamma random variables X' and Y', each with shape parameter 2 and rate parameter εT^{-1}. To simplify calculations, we rewrite $X' = (\varepsilon^{-1}T)X$ and $Y' = (\varepsilon^{-1}T)Y$, where X and Y independent Gammas, each with shape 2 and rate 1. Hence, $\eta_2 - \eta_1 = (\varepsilon^{-1}T)(X - Y)$. The pdf for a Gamma distribution with shape 2 and rate 1 is given by $f(t) = t\exp(-t)$. So, when $z < 0$,

$$f_{X-Y}(z) = \int_0^\infty f_X(x)f_Y(x-z)dx = \frac{1}{4}\exp(z)(1-z)$$

Since $X - Y$ is symmetric about 0, when $z \geq 0$ we deduce from the above integration that $f_{X-Y}(z) = \frac{1}{4}\exp(-z)(1+z)$. Hence, $f_{X-Y}(z) = \frac{1}{4}\exp(-|z|)(1+|z|)$ for all $z \in \mathbb{R}$, so then

$$\mathbb{P}(\theta_2 - \theta_1 \geq \alpha) = \mathbb{P}(X - Y \geq \varepsilon\alpha T^{-1}) = \int_{\varepsilon\alpha T^{-1}}^\infty \frac{1}{4}\exp(-|z|)(1+|z|)dz$$

When $\alpha \geq 0$,

$$\mathbb{P}(\theta_2 - \theta_1 \geq \alpha) = \int_{\varepsilon\alpha T^{-1}}^\infty \frac{1}{4}\exp(-z)(1+z)dz = \exp(-\varepsilon\alpha T^{-1})\frac{\varepsilon\alpha T^{-1} + 2}{4}$$

And when $\alpha < 0$,

$$\mathbb{P}(\theta_2 - \theta_1 \geq \alpha) = \int_{\varepsilon\alpha T^{-1}}^0 \frac{1}{4}\exp(z)(1-z)dz + \frac{1}{2} = \exp(\varepsilon\alpha T^{-1})\frac{\varepsilon\alpha T^{-1} - 2}{4}$$

Consolidating these cases delivers the advertised claim. □

Proof of Theorem 5

Proof. Let $D_{1,2} := (\hat{M}_{t_2}(d)_{a,b} - \hat{M}_{t_1}(d)_{a,b}) - (M_{t_2}(d)_{a,b} - M_{t_1}(d)_{a,b})$. By construction,

$$\hat{M}_{t_j}(d)_{a,b} = \text{round}(M_{t_j}(d)_{a,b} + \eta_j)$$

for independently drawn $\eta_j \sim Lap(\varepsilon^{-1}T)$ for $j \in \{1,2\}$. Then $D_{1,2} = \text{round}(\eta_2) - \text{round}(\eta_1)$. As η_1 and η_2 are identically distributed, so too are $\text{round}(\eta_1)$ and $\text{round}(\eta_2)$. So then $(\text{round}(\eta_2) - \text{round}(\eta_1))$ and $(\text{round}(\eta_1) - \text{round}(\eta_2))$ are also identically distributed. Thus,

$$\mathbb{P}(|D_{1,2}| > \alpha) = \mathbb{P}(|\text{round}(\eta_2) - \text{round}(\eta_1)| > \alpha) = 2\mathbb{P}(\text{round}(\eta_2) - \text{round}(\eta_1) > \alpha) = 2\mathbb{P}(\eta_2 - \eta_1 \geq \alpha + 1)$$

Replacing α with $(\alpha + 1)$ from Lemma 5 completes the proof. □

These privacy-accuracy tradeoff theorems described above can be leveraged to *proactively* enable policymakers to set ε based on context-specific goals. In particular, we derive *exact formulas* to set ε in Theorems 6 and 7. To elucidate discussion, suppose policymakers can tolerate some maximal error α with confidence $100(1-\beta)\%$ and still provide effective interventions. Then we can set $\beta \geq \exp(-\varepsilon T^{-1}(\alpha + 0.5))$. Solving for ε yields $\varepsilon \geq -T(\alpha + 0.5)^{-1}\ln(\beta)$. This yields the following result.

Theorem 6. *Any $\beta \in [0, 1)$ upper-bounds the result from Theorem 4 $\iff \varepsilon \geq -T(\alpha + 0.5)^{-1}\ln(\beta)$.*

Alternatively, if policymakers are instead interested in ensuring that trends across different O-D matrices are α-inaccurate with $100(1-\beta)\%$ confidence, then we can use Theorem 5 to derive feasible values of ε to meet this goal using the Lambert-W function[2].

Theorem 7. *For $\beta \in [0, 0.5\exp(-1)]$, β upper-bounds the result from Theorem 5 $\iff \varepsilon \geq T(\alpha + 1)^{-1}(-2 - W_{-1}(-2\beta\exp(-2)))$ where W_{-1} is the lower branch of the Lambert-W function.*

---

[2] The Lambert-W function is a countable family of functions $\{W_k : \mathbb{C} \to \mathbb{C} \mid k \in \mathbb{Z}\}$ that satisfies the product-exponential equation $x = W_k(x)\exp(W_k(x))$. That is, for $x, y \in \mathbb{C}$, $x = y\exp(y) \iff y = W_k(x)$ for some $k \in \mathbb{Z}$. In the language of complex analysis, each W_k is a branch of the Lambert-W function. When x is restricted to the real line instead of the complex plane, the domain of the Lambert-W function is $x \geq -\exp(-1)$. When $x < 0$ as well, it is sufficient to consider the principal branch W_0 and the lower branch W_{-1} to solve for $y \in \mathbb{R}$. The principal branch W_0 corresponds to all x such that $W_0(x) \geq -1$, while the lower branch W_{-1} corresponds to all x such that $W_{-1}(x) \leq -1$. See Chatzigeorgiou[55] for a further mathematical details.

Proof. Let $A = -\varepsilon T^{-1}(\alpha+1) - 2$. Then,

$$\beta \geq \exp(-\varepsilon T^{-1}(\alpha+1)) \frac{\varepsilon T^{-1}(\alpha+1)+2}{2} \iff -2\beta \exp(-2) \leq A\exp(A)$$

Since $\beta \leq 0.5\exp(-1)$, we have $-2\beta \exp(-2)) \geq -\exp(-1)$, so $-2\beta \exp(-2))$ is within the domain of the Lambert-W function. Since $-2\beta \exp(-2)) < 0$ there are two potential solutions for A over the reals: one involving the principal branch W_0 and a second involving the lower branch W_{-1}.

For the first case, we will show that there is no feasible positive ε that works. Since W_0 is an increasing function, $-2\beta \exp(-2) \leq A\exp(A) \iff W_0(-2\beta \exp(-2)) \leq A \iff \varepsilon \leq T(\alpha+1)^{-1}(-2 - W_0(-2\beta \exp(-2)))$. But this bound is always negative, as $-2\beta \exp(-2) < 0$ implies $W_0(-2\beta \exp(-2)) < 0$. So there is no feasible positive ε from the use of the principal branch.

For the second case, we will derive the inequality stated in the theorem. As W_{-1} is a decreasing function, $-2\beta \exp(-2) \leq A\exp(A) \iff W_{-1}(-2\beta \exp(-2)) \geq A \iff \varepsilon \geq T(\alpha+1)^{-1}(-2 - W_{-1}(-2\beta \exp(-2)))$. Unlike the prior case, this bound is indeed a positive quantity. By definition of the Lambert-W function, $W_{-1}(-2\exp(-2)) = -2$. Since $\beta \leq 0.5\exp(-1) < 1$, we have $-2\exp(-2) < -2\beta \exp(-2)$. Since W_{-1} is decreasing, we deduce that $-2 = W_{-1}(-2\exp(-2)) \geq W_{-1}(-2\beta \exp(-2))$ which implies that the bound is positive. □
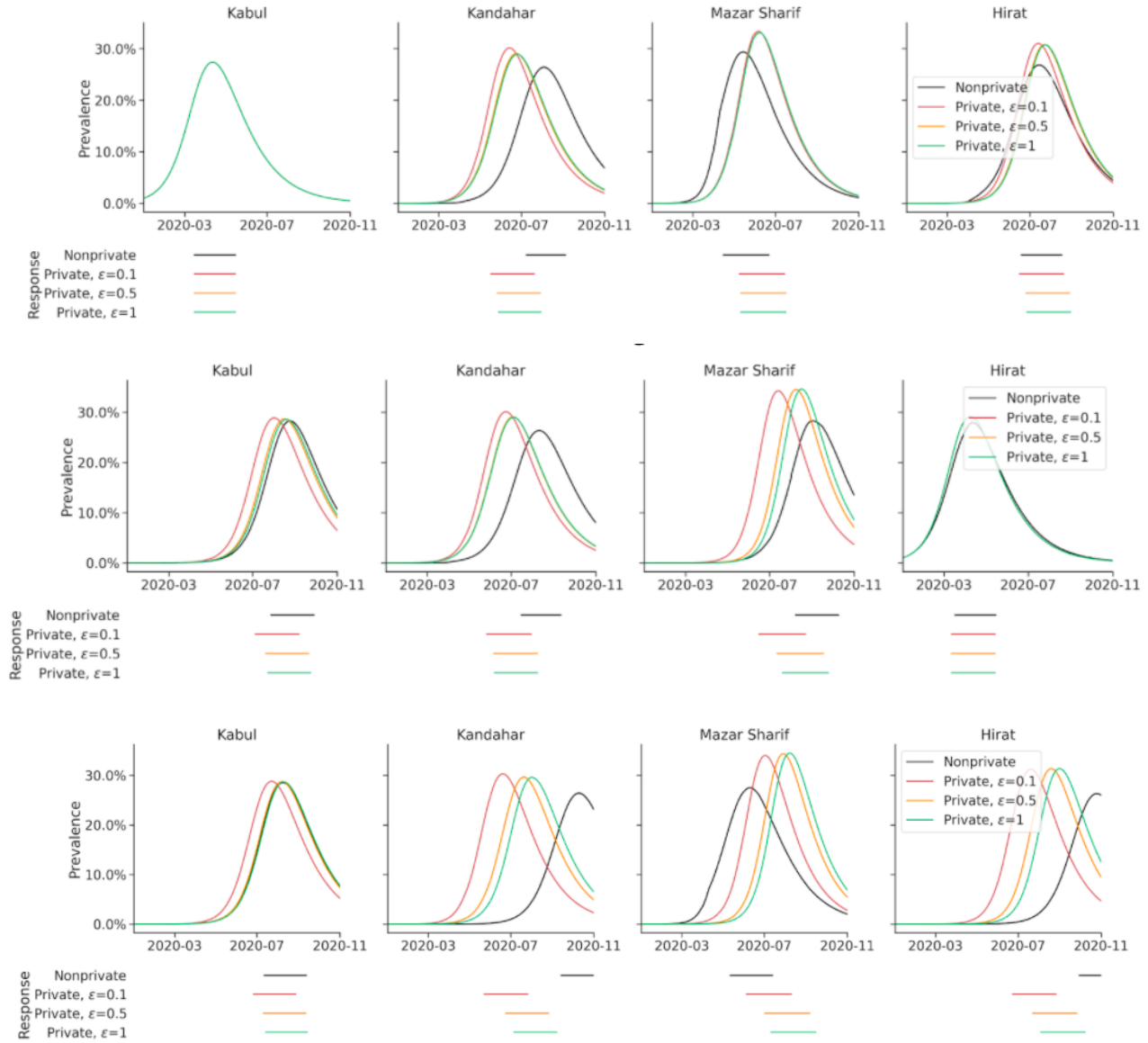
Figure S1. Epidemic curves based on mobility-based SIR models and O-D matrices derived from call detail records. Top: Pandemic initiating in Kabul. Middle: Pandemic initiating in Hirat. Bottom: Pandemic initiating randomly.
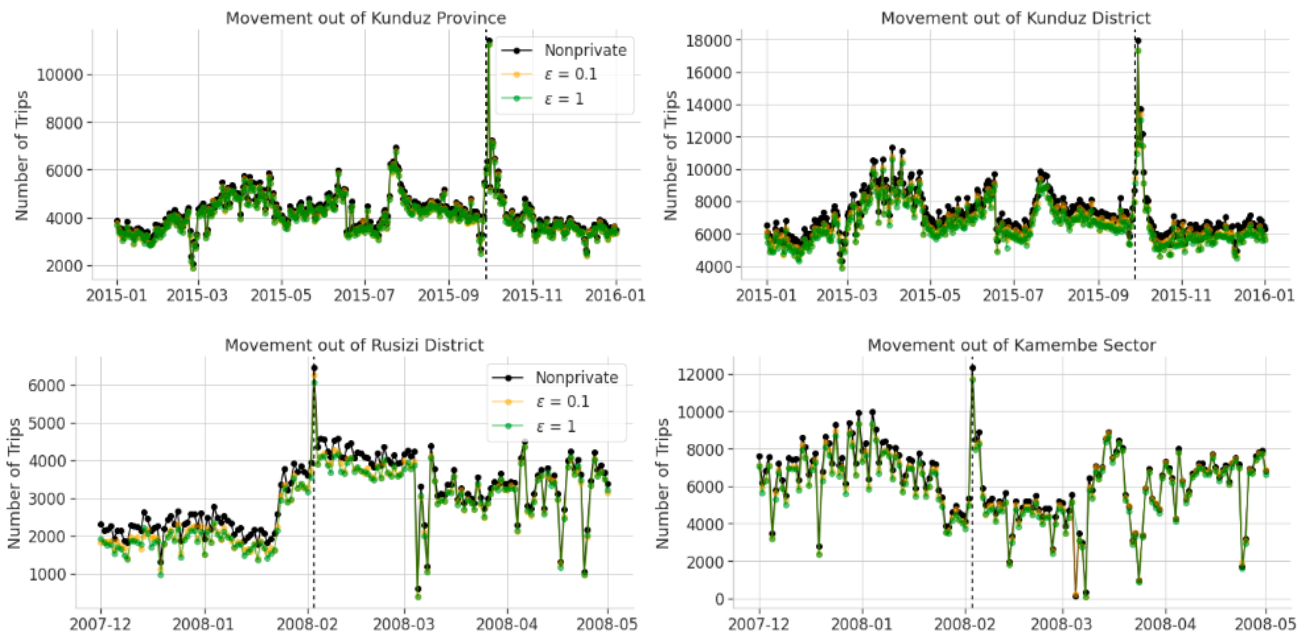
Figure S2. Total out-migration from areas used in simulations of humanitarian response to natural disasters and violent events, calculated from call detail records. Top: Out-migration during the Battle of Kunduz in Afghanistan. Bottom: Out-migration following the Lake Kivu Earthquake in Rwanda.
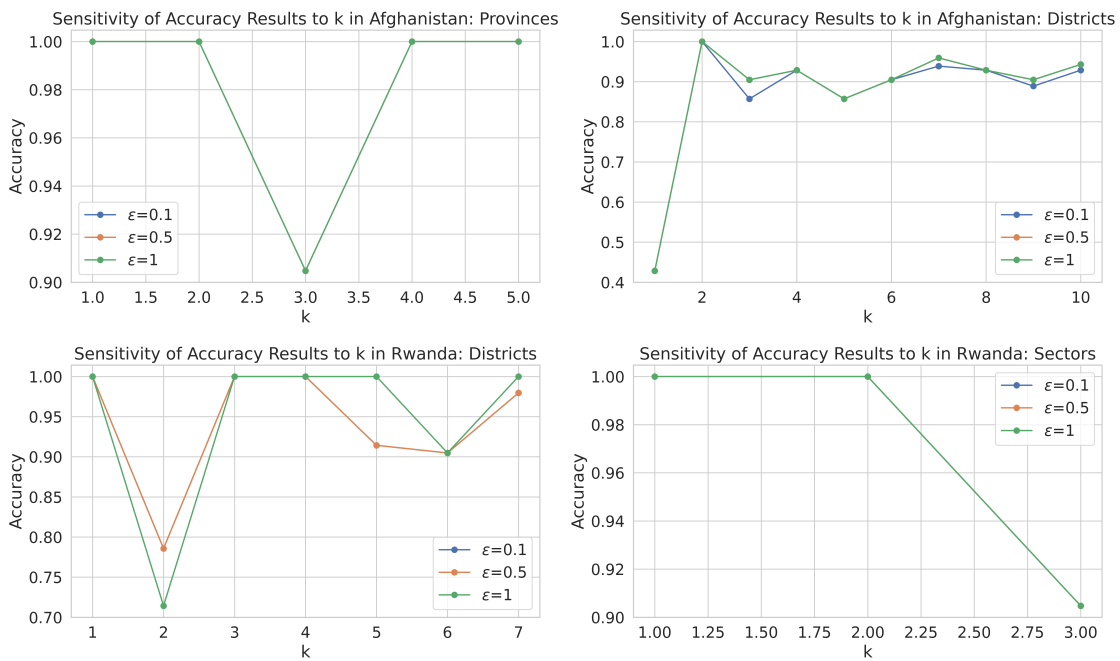

Figure S3. Sensitivity of results on accuracy for identifying top-*k* regions of out migration in simulations of natural disasters and violent events to value of *k*. Left: admin-2 level. Right: admin-3 level. Lines are not present where there are fewer than *k* regions of out-migration after supression of small counts.