The Effect of Population Density on Remote Humanitarian Mapping Activities: A Triple-Difference Analysis

ANONYMOUS AUTHOR(S)

The proliferation of OpenStreetMap (OSM) as a collaborative geographic dataset has been instrumental in addressing data gaps globally. However, disparities in map coverage persist, particularly in economically disadvantaged and disaster-prone regions. The emergence of the Humanitarian OpenStreetMap Team (HOT) in 2010 aimed to bridge these gaps by leveraging the collective efforts of volunteers through platforms like the HOT Tasking Manager. While previous research has highlighted the success of these initiatives in recruiting contributors and expanding map coverage, their implications for existing structural biases remain unclear, potentially hindering the regions benefiting from humanitarian activities. Thus, our study employs the difference-in-difference-in-difference(DDD) approach to empirically examine the pattern between contribution dynamics and population density in project regions involved in humanitarian mapping activities. By further investigating the participation of various levels of contributors in projects with different population densities, we aim to inform better design strategies to align contributor expectations and experiences, fostering more equitable and effective humanitarian mapping efforts.

CCS Concepts: • **Do Not Use This Code → Generate the Correct Terms for Your Paper**; *Generate the Correct Terms for Your Paper*; Generate the Correct Terms for Your Paper; Generate the Correct Terms for Your Paper.

Additional Key Words and Phrases: Do, Not, Us, This, Code, Put, the, Correct, Terms, for, Your, Paper

ACM Reference Format:
Anonymous Author(s). 2018. The Effect of Population Density on Remote Humanitarian Mapping Activities: A Triple-Difference Analysis. In . ACM, New York, NY, USA, 21 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

OpenStreetMap is one of the most successful examples of Volunteered Geographic Information (VGI) [15] platforms today, and has gained increasing popularity and importance due to its ability to meet the growing demand for accessible geographic data. For instance, OpenStreetMap data are widely used in consumer-facing applications and services, including Tesla, Amazon, and Craigslist, among others [8, 9, 43], as well as playing a crucial role in informing decision-making related to urban planning, public health, and climate change [37].

However, structural information disparities that follow dimensions such as population density have become significant obstacles in most peer production systems, including OpenStreetMap, potentially preventing them from reaching their full potential [34, 57]. OpenStreetMap content is voluntarily produced by contributors who are not evenly distributed globally [21, 28, 36, 58], such contribution dynamics follow "born, not made" patterns, where contributors predominantly focus on higher socioeconomic and more population dense areas across time in the system, facilitating to the creation of these information disparities. Making matters worse, these less well-covered regions often also face unavailable or out-of-date governmental data [1, 2], which increases their vulnerability to natural disasters and complicates post-disaster recovery efforts.

Manuscript submitted to ACM

To better support these vulnerable regions and in alignment with the Sustainable Development Goals (SDGs) of the United Nations, the Humanitarian OpenStreetMap Team (HOT) was established in 2010. Initially formed in response to the devastating earthquake in Haiti, HOT aims to bridge data gaps and focus mapping efforts to regions where they are most needed [25]. A primary mechanism for this purpose is a microtasking tool called the Tasking Manager (https://tasks.hotosm.org/), which helps to organize and manage collaborative mapping projects in over ten thousand regions, leveraging the efforts of contributors from around the globe. By decomposing mapping work into smaller, focused tasks (as shown in Figure 1), the HOT Tasking Manager allows individuals to contribute data about small regions, without lengthy time commitments [59]. This approach may help overcome the "born, not made" patterns, as the Tasking Manager facilitates anyone, from anywhere, to produce map content based on satellite imagery, meaning that contributors do not need localized knowledge or expertise to contribute [6, 17, 30, 52–54]. For example, during the 2015 Nepal earthquake, around 9,000 volunteers worldwide participated through Humanitarian OpenStreetMap, rapidly mapping critically affected areas of Nepal in three days [26]. The detailed map that resulted became instrumental in guiding the allocation of essential supplies and medication during subsequent relief operations [20].

However, the extent to which the community benefits from humanitarian mapping activities, particularly in mitigating structural biases in OpenStreetMap, such as those related to local population density [21, 28, 36, 58], remains unclear. Indeed, prior research has found that humanitarian project creation in various regions has successfully increased local participation and coverage [10, 24, 38, 41, 63], while also exacerbating the phenomenon of contribution concentration, where a small group of contributors conducts the majority of the work, following a power-law distribution [41, 63].

Our work here focuses on this central concept: if microtasking tools like the Tasking Manager are broadly effective at helping increase coverage of the map in places that have missing or incomplete data, are they *equally* effective in all places? We focus here on one well-known dimension of disparity in OpenStreetMap, population density, and investigate this driving question. Holistically, our work is guided by the following two research questions:

RQ1 How does the population density of project regions influence the contribution dynamics in humanitarian efforts?

RQ2 How do these patterns reflect contributor participation in project regions with different population densities?

To conduct this work, our study relies on a robust "natural experiment" enabled by the HOT Tasking Manager and employs a difference-in-difference-in-difference (DDD) model to explore how the population density of projects influences contribution dynamics in the humanitarian OpenStreetMap community. Additionally, through the lens of power-law dynamics, we investigate participation patterns in the context of humanitarian activities, making three primary contributions:

- Our results indicate that population density remains a significant factor in contribution disparity within the Humanitarian OpenStreetMap community, despite Tasking Manager project creation being helpful in increasing coverage. Project regions with higher population density tend to attract more contributors, but also exhibit a pronounced concentration of contributions.
- We show that the *contribution behavior mechanism* behind these trends varies according to the types of regions that projects are created in. These results suggest distinct participation patterns that shape how contributions are concentrated across volunteer contributors, extending and adding nuance to current CSCW understandings of contribution dynamics in peer production systems like OpenStreetMap.
- Holistically, the contribution and participation dynamics in humanitarian OpenStreetMap highlight that, despite the potential for the Tasking Manager to help overcome the "born, not made" bias, disparities persist along

population density in humanitarian efforts. Such patterns shed light on implications for practitioners and CSCW researchers, and suggest potential design directions for participation and the impact of peer production tools across varied regions.

2 BACKGROUND OF OPENSTREETMAP ECOSYSTEM

OpenStreetMap (OSM), established in 2004, is a collaborative mapping platform often referred to as the "Wikipedia of maps." It operates as a volunteer-driven Volunteered Geographic Information (VGI) system where people can remotely collaborate to create and maintain accessible geographical data. The platform has grown to be one of the most important geographic data sources globally [19], now encompassing over 10 million registered members, with approximately 2 million active contributors generating an average of 4 million daily map changes.

The OSM ecosystem supports a variety of diverse contribution mechanisms that accommodate varying levels of expertise and engagement. Contributors can map remotely by tracing satellite imagery, or collect on-the-ground data using using GPS-enabled devices, or even through importing authorized open-source geographical information [39]. These contribution methods serve distinct mapping needs [29], ranging from creating new geometries to validating existing data, thereby ensuring comprehensive geographic coverage and data quality.

Within this broader ecosystem, the Humanitarian OpenStreetMap Team (HOT) operates as a specialized initiative focused on humanitarian and disaster response scenarios. Founded after the 2010 Haiti earthquake [65], HOT has grown significantly over the decade. Through its Tasking Manager system [12], HOT has facilitated more than 10,000 humanitarian projects in partnership with various organizations such as Red Cross and humanitarian aid campaigns, enabling coordinated mapping efforts across the globe. Unlike the general OSM community where contributors freely choose mapping areas, HOT implements a structured approach to project discovery and participation through the Tasking Manager (Fig 1), which organizes and coordinates mapping efforts more systematically. According to Dittus et al. [11], potential contributors engage with HOT projects through three primary channels. First, high-profile humanitarian initiatives attract broad public participation through substantial online and offline media coverage. Second, strategic partnerships with large organizations enable direct recruitment of contributors, often bringing specialized expertise to specific humanitarian mapping needs. Third, contributors independently discover HOT projects through the platform's project listing, driven by personal interests or humanitarian concerns. Overall, the structured approach to project discovery and participation distinguishes HOT from the broader OSM community, enabling focused humanitarian mapping efforts while maintaining the collaborative spirit of the larger OSM ecosystem.

3 RELATED WORK

While prior research extensively explores contributor and community dynamics within broader Volunteered Geographic Information (VGI) contexts, less attention has been given to understanding these dynamics within smaller, specialized communities such as the humanitarian OpenStreetMap community. Our work addresses this gap by investigating variations in humanitarian mapping activities, with a particular focus on the impact of population density. This study builds upon and extends three primary areas of prior research: (1) Geographic Disparities in OpenStreetMap, (2) Contribution Disparities in OpenStreetMap, and (3) Background and Contributor Activities in Humanitarian OpenStreetMap.

3.1 Geographic Disparities in OpenStreetMap

Prior research has consistently demonstrated significant disparities in the quality and quantity of content within OpenStreetMap, often focusing on socioeconomic status and urban/rural dimensions of analysis [11, 24, 24, 57]. For instance, prior work shows that areas with lower socioeconomic status tend to have fewer mapping contributions and less engagement from contributors [11, 17]. Moreover, Herfort et al. [24] identified geographical disparities in both the quality and type of contributions, noting that regions with higher Human Development Indexes receive more attention in OpenStreetMap's mapping efforts. Conversely, regions with low and medium levels of human development, where the majority of the population resides, are often neglected, with only a minor portion of roads and buildings mapped [24]. Further study conducted by Thebault-Spieker et al. [57] shows that contribution trends in OpenStreetMap follow "born, not made" patterns, indicating that contributors predominantly focus on urban and higher socioeconomic areas from the onset of their participation.

Another study by Thebault-Spieker et al. [56] also found that the types of content being mapped are, in some cases, subject to highly localized contribution patterns, which illustrates some of the underlying causes of such disparities. One theory explaining the relationship between data and participation disparities in OpenStreetMap is self-focus bias [7], which posits that individuals tend to contribute information about places local to them [22, 56]. In other words, the self-focus bias concept would suggest that most contributors in OpenStreetMap live in urban and wealthier places, and thus tend to contribute in urban and wealthier places, thereby causing areas with socioeconomic disadvantages and rural regions tend to exhibit lower data coverage.

3.2 Contribution Dynamics in OpenStreetMap

OpenStreetMap, like many other peer production systems [22, 30, 31, 49, 55], exhibits the power-law dynamics in how contribution occurs [56]. Small numbers of contributors tend to produce the majority of the contributions, accounting for a large proportion of the overall contribution effort in the system [13]. In OpenStreetMap, Yang et al. [62] evaluated contributions across four countries, finding that despite their differing trajectories, their Gini coefficient — a metric capturing contribution inequality — can reach a high level (0.95 out of 1). Moreover, prior research has also explored the participation patterns in the OpenStreetMap community, revealing that less than 10% of contributors remain active six months after their initial contributions to the project [5, 35].Additionally, Sim and Biddle [51] and Arazy et al. [4] found that participation levels are typically associated with the social status and identities of contributors, as well as their assigned responsibilities and access privileges.

Other studies have explored the influence of contribution inequality on peer production systems. They found that power-law dynamics imply contribution inequality can influence data disparity [56], increase heterogeneity [18], and create barriers for new participants [18]. While these power-law contribution dynamics may be common to peer production settings, there is a risk that they also are a mechanism of disparity within these systems, though prior research has not yet fully characterized that mechanism.

3.3 Background and Contributor Activities in Humanitarian OpenStreetMap

Unlike the broad OpenStreetMap community, which engages in global map-making, the Humanitarian OpenStreetMap Team (HOT) [25] focuses specifically on regions where mapping efforts are critically needed. By creating and releasing microtasking projects in the Tasking Manager(https://tasks.hotosm.org/), HOT facilitates more targeted mapping efforts to support disaster relief and humanitarian goals. Beyond responding to immediate disaster events such as Typhoon

Yolanda and the Ebola virus epidemic [11], HOT also undertakes a variety of long-term, mission-focused mapping projects covering extensive areas [24]. For instance, HOT's Climate-Ready Cities program launches projects in East and Southern Africa to enhance local mapping capabilities for responding to and mitigating climate risks. Additionally, projects initially launched as localized emergency responses, such as those for the Ebola outbreak, have expanded into mission-focused initiatives that improve maps in affected regions to achieve long-term objectives [33].

With the growing popularity and importance of the Humanitarian OpenStreetMap Team (HOT), an increasing body of work focuses on the dynamics of contributions in the context of humanitarian mapping activities [10]. Prior research has observed that while microtasking creation has led to increased participation and better coverage in projects [24, 63], contribution patterns in the humanitarian OpenStreetMap community still exhibit a power-law distribution, even worse than before [41, 63]. Moreover, while humanitarian mapping initiatives improve geographic coverage, they may inadvertently lead to a reduction in ongoing maintenance and enhancement of the map [38].

However, engagement patterns for humanitarian purposes display unique dynamics compared to general contributions in OpenStreetMap. For instance, Gary Esworthy [14] found that following disasters such as earthquakes and hurricanes, contributor activity on platforms like OpenStreetMap and Wikipedia typically spikes shortly after the event and then declines over time. Prior studies have also focused on well-known campaigns in the Tasking Manager, analyzing contribution dynamics within these projects [10, 41, 45]. They found that while newcomer mappers generally contribute at lower rates than prolific contributors, their efforts are essential for comprehensive data collection, particularly in humanitarian mapping activities [10, 41, 45]. Dittus et al. [10] found that event-centric campaigns, such as those responding to Typhoon Haiyan/Yolanda, tend to attract more contributors and reactivate previous contributors. However, newcomer contributors may produce lower quality data. Overall, while humanitarian mapping initiatives improve geographic data coverage, there may be unintended risks around data quality [10] or on-going maintenance of the map data [38].

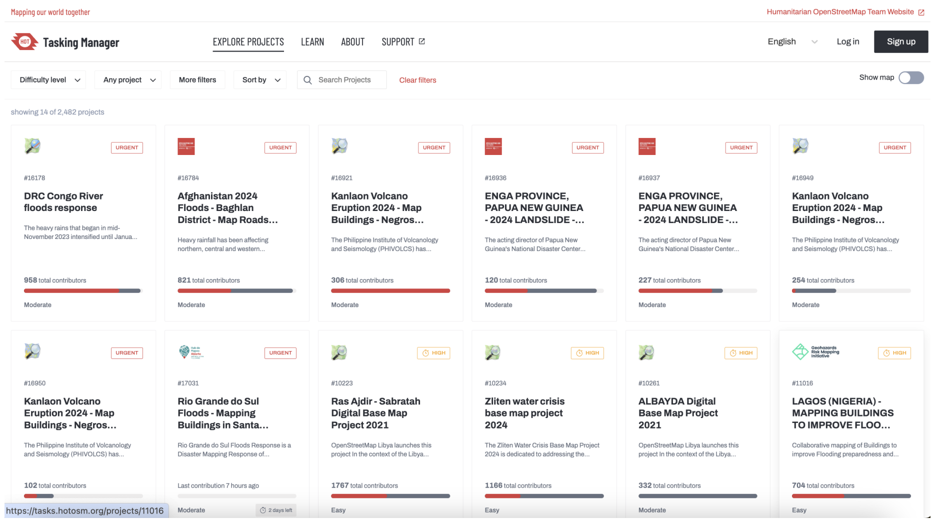
4 METHODOLOGY

Prior work conducted by Yin et al. [63] using quasi-experimental methods — a difference-in-differences (DID) approach, commonly used in the social sciences to control for potential confounding factors — causally studied how HOT project creation influenced contributor and contribution dynamics within Humanitarian OpenStreetMap. This work found that the creation of microtasking projects in the Tasking Manager indeed led to increased participation in humanitarian mapping efforts, with a higher average contribution rate, echoing prior observational work [10, 24, 38, 41, 63]. However, Yin et al. [63] also found that project creation exacerbates contribution inequality, as measured by the Gini coefficient, a phenomenon characterized by a power-law distribution.

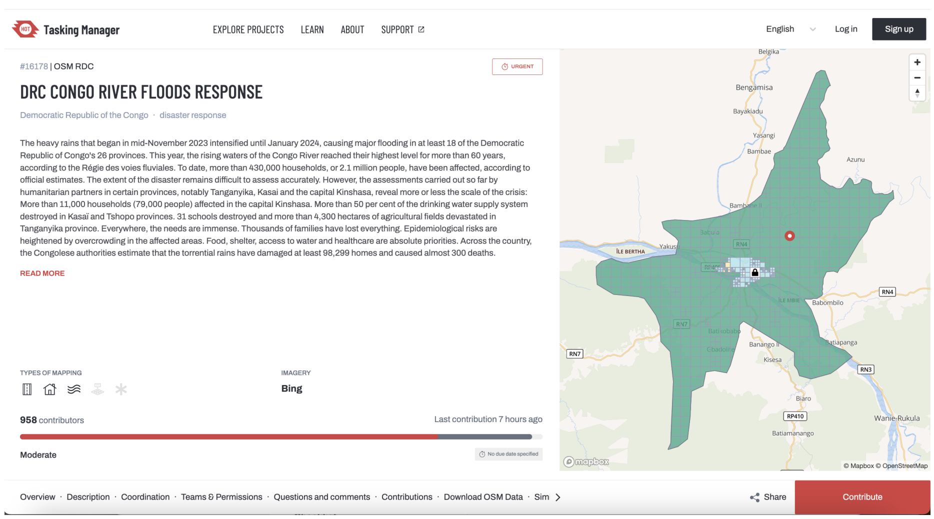
Guided by our research questions and with the goal of extending and adding nuance to Yin et al. [63]'s prior work, here we also adopt the difference-in-difference-in-difference (DDD) method. Whereas prior work relied on HOT project creation as a way to causally understand contributor dynamics, our work here adopts a similar study paradigm — the difference-in-difference-in-difference (DDD) approach — to control for the causal mechanisms, and to focus our analysis on the influence of population density in a microtasking setting.

4.1 Setting up the difference-in-difference-in-difference (DDD) Model

= To empirically understand how population density of project regions influences the contribution dynamics in humanitarian efforts, we adopted a difference-in-differences (DID) approach, commonly used in social sciences and economics [3, 27, 38, 47]. The DID approach leverages a "natural experiment" setting where an "experimental treatment"



(a) List of Active Projects in Tasking Manager



(b) Interface of a Specific Project

Fig. 1. Examples of Tasking Manager Interface

(in our case, HOT project creation) occurs in some regions but not others, enabling the construction of control groups. This method allows us to causally explore trends by comparing changes between treatment and control groups over time.

To further examine how population density mediates these effects, we extended our analysis to a difference-in-difference-in-difference (DDD) model. The DDD model builds upon the DID framework by incorporating additional variables to understand interaction effects with the causal trends. Specifically, by including population density as an interaction term, we can analyze not only the causal effects of project creation but also how these effects vary across regions with different population densities. Following the methodological framework established by Yin et al. [63], our DDD model is formally defined as:

$$
\begin{aligned}
Y_{sit} = {} & \beta_0 + \beta_1 Treat + \beta_2 Post + \beta_3 PopDensity \\
& + \beta_4 (Treat \times Post) + \beta_5 (Treat \times PopDensity) + \beta_6 (PopDensity \times Post) \\
& + \beta_7 (Treat \times Post \times PopDensity) + \epsilon_{sit}
\end{aligned}
\tag{1}
$$

Here, Y_{sit} represents the dependent variables that reflect the dynamics of community contribution, which have been widely investigated in prior work [10, 24, 38, 41, 63], including the number of contributors, individual productivity and the Gini coefficient (see 4.3). The variables $Treat$ and $Post$ are dummy variables. $Treat$ equals 1 if the regions where the projects have been created are in Tasking Manager; otherwise, it equals 0. $Post$ equals 1 if it is after the project creation date, which we obtained from the Tasking Manager API (see 4.2.1); otherwise, it equals 0. $PopDensity$ represents the population density within a region. We applied a $\log_2$ transformation to this variable to ensure compliance with the necessary modeling assumptions. The term $Treat \times Post \times PopDensity$ is our DDD estimator. The coefficient β_7 measures the effect of population density on the contribution dynamics among project regions. If β_7 is positive, it suggests that regions with higher population density have a positive estimate on the metrics. In contrast, negative impacts would yield negative estimates with statistical significance for this interaction coefficient.

4.2 Data Collection

4.2.1 Experimental Treatment Group.

To construct our "treatment" group consisting of HOT projects, we accessed the Tasking Manager API to retrieve all 11,894 projects published up to May 2024. During our data collection, we encountered 624 projects that were unavailable through the API, likely due to deletion or removal from public access. Consequently, our dataset comprises 11,270 projects that constitute our experimental treatment group. For each of these projects, we collected the project creation date and their geographic region where each project was initiated, which was essential for constructing the "control" group and obtaining population density measures.

4.2.2 Experimental Control Group.

To establish "experimental pairs" [64], with one member belonging to the "control" group and the other receiving the "treatment", we followed the method used in studies such as [16, 23, 63], which adhere to the principles of the "First Law of Geography"[16, 23] and "local production"[18, 57]. The "First Law of Geography" suggests that regions in close proximity are likely to exhibit geographical similarities [16, 23].

More specifically, in the first stage, we focused on geographic matching. For each project region, we defined its area as a circle with radius R from its center coordinates. We then identified candidate control regions as adjacent areas that form tangent circles with the same radius R, ensuring no overlap with HOT project areas. This geometric arrangement ensures that paired regions share fundamental characteristics such as climate and geological structure while maintaining independence from humanitarian projects.

In the second stage, we refined our selection using population density as an additional criterion. This refinement is grounded in previous research showing that OpenStreetMap contribution patterns are significantly influenced by local population density [18, 57, 63]. Among the geographically adjacent candidate regions identified in stage one, we selected the region with the most similar population density to its corresponding project region as control group.

In summary, for each "treatment" region, we identified a neighboring region with the same geographic outline as the project region, and ensured that they did not overlap. On average, our "control" region differed by differing by less than 12 people/km^2 in comparison to our "experimental" project region, indicating substantial similarity both geographically and in terms of population of the area.

4.2.3 Population Density of A Region.

Population density is a crucial data source in our study, used both to define "control" regions and as an independent variable in our analysis. Globally, however, not all countries provide official administrative data for population. Moreover, Humanitarian OpenStreetMap focuses on specific regions that do not necessarily align with administrative boundaries, so we need a more globally available dataset. Therefore, we used the Global Gridded Population of the World, Version 4 (GPWv4) dataset from NASA's Socioeconomic Data and Applications Center (SEDAC) and computed the population density per km^2 within each project region and "control" region. We then applied a log_2 transformation to the population density variable, in order to adjust for distributional skew in our DDD model. This transformation means that a one-unit increase in the log_2-transformed population density corresponds to a doubling of the actual population density.

4.3 Observation Period and Dependent Variables

With our "experimental pairs" established, we proceeded to collect and analyze OpenStreetMap data for each region to define our observation period and capture variables related to contribution dynamics. Data was extracted from the OpenStreetMap history planet dump as of May 2024. We initially calculated the number of contributions for both "treatment" and "control" regions over a 30-day period, spanning 15 days before and 15 days after project initiation. Upon further analysis, as depicted in Figure 2, we refined our observation window to a more focused two-week period, specifically 7 days before and after project creation. This adjustment follows best practices [38, 63], aiming to balance the need for a sufficiently broad window to detect causal effects while minimizing the influence of external variables that could lead to spurious correlations.

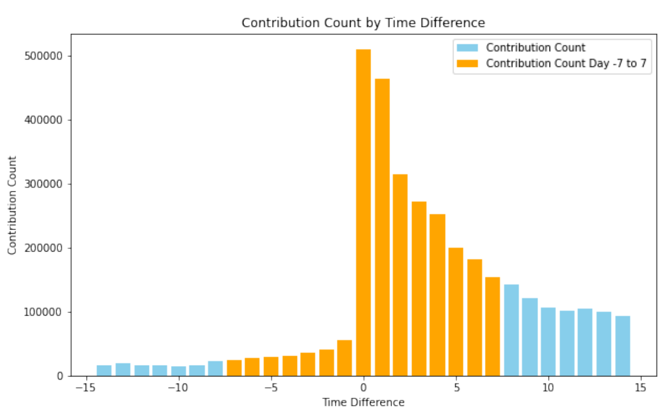


Fig. 2. Total Number of Contributions in Both Project Regions and Control Regions Over Time.

We then computed three dependent variables to represent different aspects of contribution dynamics, following metrics used in prior work [63]: the number of contributors, the average number of contributions per person, and the

Gini coefficient. The first two variables reflect our interest in well-understood patterns in OpenStreetMap — that the number of contributions and the number of contributors increase with sociodemographic variables like population density. The third variable, the Gini coefficient, is a measure of distributional skew that is widely used in Economics [13]. Yin et al. [63] also showed that project creation can influence how contribution is distributed within a HOT project, so we include this variable here as well. When the Gini coefficient value is near 0, it indicates an equal distribution of contributions among all mappers in the project in our time window, while a Gini coefficient near 1 indicates that the majority of contributions are concentrated among a few highly active contributors [13, 63]. Specifically, we measured the daily number of contributors, the daily average number of changesets per contributor, and the daily Gini coefficient in each region. We computed these variables for all regions in both the "treatment" and "control" groups, resulting in a dataset containing 14 daily measurements for each of our three dependent variables for each paired region.

4.4 The Parallel Trends Assumption in DDD

In order to reliably draw causal interpretations from the DDD model construction, there is a key assumption that needs to be met. Namely, the "control" and the "treatment" groups need to exhibit similar or "parallel" trends *prior* to the creation of the project, in our case. However, because the DDD approach also focuses on interaction effects with the causal trends, it requires that the relative outcome of the interaction variables also exhibit this parallel trend prior to the intervention. In our case, this means that our "treatment" regions and our "control" regions need to have similar trends, *in terms of population density*, before the "experimental intervention" [40]. Therefore, to implement this analysis, we first filtered our dataset to include only observations from before HOT projects were created. We then constructed a regression to evaluate the parallel trends assumption:

$$Y_{sit} = \beta_0 + \beta_1 \text{Day} + \beta_2 \text{PopDensity} + \beta_3 (\text{Day} \times \text{PopDensity}) + \epsilon_{sit} \tag{2}$$

In this equation, Y_{sit} represents the three outcome variables. The variable Day_t is a continuous measure ranging from -7 to 0, representing the days prior to project creation. PopDensity_s is transformed using log_2 , representing the average population density of that region. The coefficients β_0, β_1, β_2, and β_3 represent the intercept, the effect of day, the effect of population density, and the interaction effect between day and population density, respectively. Finally, ϵ_{sit} denotes the error term, capturing the variation in the outcome variable not explained by the model.

The most crucial coefficient for testing the parallel trend assumption is β_3. This coefficient captures how the population density trend, over time, differs between the two population density groups. Our models consistently show thatβ_3 is not statistically significant, suggesting that the trends in population density over time are statistically indistinguishable — or parallel — across the different population density groups in the pre-treatment period. In short, the parallel trend assumption is supported.

5 RESULTS

5.1 RQ1: How does the population density of project regions influence the contribution dynamics in humanitarian efforts

To address our first research question, we focus on our difference-in-difference-in-difference (DDD) analysis, and the results of our model are shown in Table 1. Examining this table in detail, immediately visible is the baseline causal effect of project creation in the HOT Tasking Manager. This is evident through the "treated × time" coefficients in our three models, serving as a replication of the results found in prior work [10, 63]. All other things held constant, this model predicts that when a Humanitarian OpenStreetMap project is created, we would expect a causal increase of 3.47 contributors, 247.56 contributions per person, and an increase in the Gini coefficient of 0.095. These baseline causal findings echo those of Yin et al. [63] in direction, significance, and size of the coefficients.

Our first research question, however, focuses more directly on how variations in population density do, or do not, relate to these same contribution dynamics. Focusing on the "treated × time × Population Density (log2)" term in our three models, the results we find are mixed. First, we find no significant relationship between this interaction term and individual contribution rates. That is, we find no evidence that volunteers' contribution rates vary with population density, above and beyond the increase caused by the creation of the project. However, we see different trends for our contributors and Gini coefficient models. In our number of contributors model, this interaction effect suggests for two project regions that are otherwise equivalent, but one has twice the population of the other, we would expect a slight increase in the average number of contributors, with a coefficient of 0.1 contributors (p<0.001). Similarly, in our Gini coefficient model, for two project regions that are otherwise equivalent, but one has twice the population, we would expect the Gini coefficient in that region to increase by 0.002 ($p < 0.0001$).

Taken holistically, our results in this analysis replicate prior findings that suggest that creating a Humanitarian OpenStreetMap project in a region via the HOT Tasking Manager not only helps achieve better data coverage but also exacerbates contribution concentration. Moreover, we find that these causal trends in the number of contributors and the Gini coefficient vary with the population density of project regions. A project in a region with higher population density tends to attract more contributors but also exacerbates the concentration of contributions within that region. To contextualize the superficially small coefficients described above, we turn to Table 2. An increase of 0.1 contributors per unit in our log_2 population density variable would mean that a 5-unit increase in log_2 transformed population density would result in an increase of 0.5 contributors. This size increase is very possible within our dataset, and would be equivalent to the comparison between "low" population dense regions like Oslo, Norway. vs "high" population dense regions like Augusta, US. Similarly, a 5-unit increase in log_2 transformed population density would result in an increase of the Gini coefficient by 0.02. Despite the benefits of eliciting more remote and distanced mapping efforts, our results suggest that higher population dense regions still receive different treatment above and beyond the causal benefits of the HOT Tasking Manager. This suggests that the disparities caused by of "born, not made" style patterns [57] persist, even in a microtasking setting.

Predictors	Model 1 Number of Contributors	Model 2 Productivity	Model 3 Gini Coefficient
(Intercept)	0.142	59.607***	0.022
treated × time	3.470***	247.564***	0.095***
treated × time ×Population Density ($\log_2$)	0.102**	-2.884	0.002***
treated	0.250	24.584**	0.006**
time	1.193	3.132	0.023***
Population Density ($\log_2$)	0.018	-1.227	0.000
time × Population Density ($\log_2$)	-0.038	-2.254	-0.001***
treated × Population Density ($\log_2$)	0.014	-0.509	0.001*

Table 1. Difference-in-difference-in-difference (DDD) Results *Notes:* *p<0.05; **p<0.01; ***p<0.001.

5.2 RQ2: How do these patterns reflect contributor participation in project regions with different population densities?

Our findings in RQ1 point to structured variations in the impacts of project creation, which follow the population density of the regions where projects are created, for both the number of contributors and the Gini coefficient. However, while the trend for the number of contributors is intuitive, the Gini coefficient trend is more complex to unpack. This is because the Gini coefficient, as a metric of concentration, can change in a number of ways. To further explore this trend and aid in interpretation, we categorized projects by population density and contributors by how prolific they are, following prior work [58].

To systematically analyze population density's influence on contribution patterns, we extend previous geographic analysis approaches in peer production systems [28, 56] by developing a more granular categorization that captures nuanced differences across the population density spectrum, rather than using traditional urban/rural divisions. Specifically, we classified project regions into five categories: Very Low, Low, Medium, High, and Very High density. We established category boundaries based on the mean (μ) and standard deviation (σ) of the $\log_2$ population densities: regions with density 0–6 people/km^2 were classified as Very Low density (below $\mu - 1.5\sigma$), 6–54 people/km^2 as Low density (between $\mu - 1.5\sigma$ and $\mu - 0.5\sigma$), 54–150 people/km^2 as Medium density (between $\mu - 0.5\sigma$ and $\mu + 0.5\sigma$), 150–407 people/km^2 as High density (between $\mu + 0.5\sigma$ and $\mu + 1.5\sigma$), and 407–55,413 people/km^2 as Very High density (above $\mu + 1.5\sigma$). This categorization method ensures statistical robustness while maintaining meaningful distinctions between categories, aligning with previous studies on population density effects in spatial crowdsourcing [28**?**]. The descriptive statistics for each category are shown in Table 2.

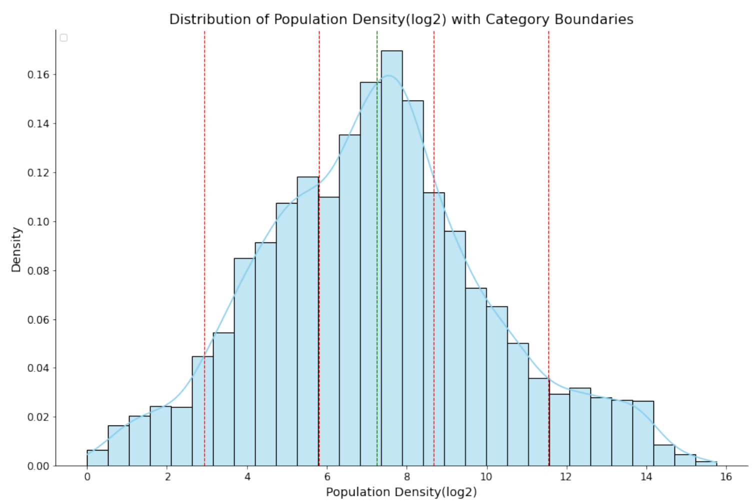


Fig. 3. Distribution of Population Density (log_2) with Category Boundaries

With regard to contributors, we also follow prior work [42, 57], and categorize contributors based on 1) engagement consistency and 2) contribution amount within a 15-day timeframe surrounding the project's creation. We designate contributors who have not made any contributions during a period extending 7 days prior to the project's creation as "new" contributors. This includes individuals who are either reactivating their participation or joining OpenStreetMap for the first time. To categorize contributors by how prolific they are, we apply the power law distribution definition and

Table 2. Population Density Ranges for Each Category with City Examples and Number of Projects

Category	Population Density ($\log_2$)	Original Population Density	City Example	Number of Projects
Very Low	Less than 2.58	0 – 6	Alberta, Canada	724
Low	2.58 – 5.75	6 – 54	Oslo, Norway	2871
Medium	5.75 – 7.23	54 – 150	Budapest, Hungary	2133
High	7.23 – 8.99	150 – 407	Augusta, US	2501
Very High	Greater than 8.99	410 – 55413	Manila, Philippines	3041

follow practices used in prior work to determine the cutoffs, as shown in Figure 4. In the humanitarian OpenStreetMap community, the top 5% of individuals contribute 54.03% of the total contributions, the next 15% contribute 22.19%, and the bottom 80% contribute 23.79%. This process results in 6 categories of users, three main groups based on the aforementioned percentages (5%, 15%, and 80% contributors) and three corresponding groups of new contributors who mirror these participation levels, namely 5%, 15%, 80%, New 5%, New 15%, and New 80%.

Based on these population density and user categories, we now turn to addressing our second research question, seeking to understand the underlying contributor behaviors that lead to the trends we observed in Section 5.1.

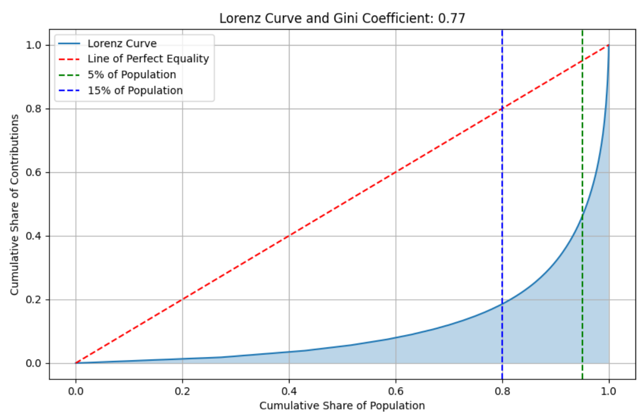
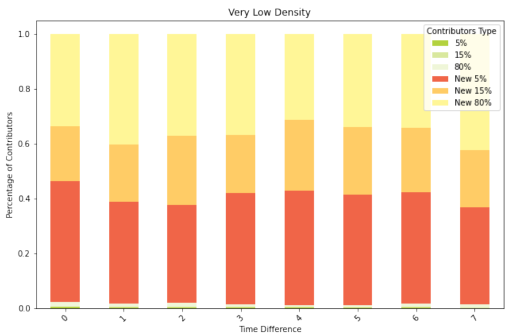


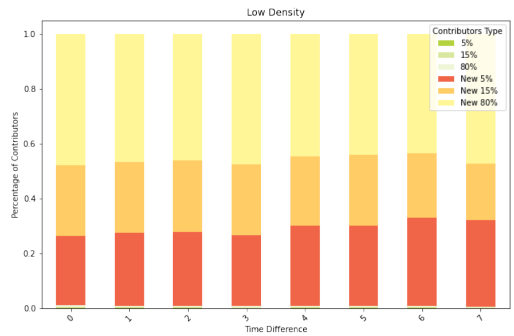
Fig. 4. Distribution of Contribution

5.2.1 Widening Contributor Inequality with Increasing Population Density. In order to more deeply understand the trends we found in Section 5.1, we first split our data according to the different population density groups shown in Figure 3, and plot the distributions of contributors within each of our six groups across the week following the creation of the HOT project.

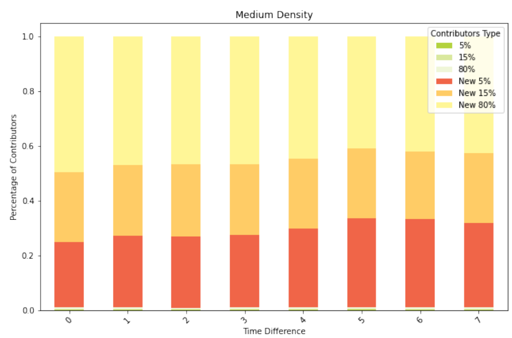
Examining Figure 5 in detail, we see that the majority of humanitarian efforts made after project creation predominantly come from "new" contributors. More specifically, these individuals, either inactive in the seven days preceding the project or completely new to Humanitarian OpenStreetMap's mapping activities, eventually become the majority of active contributors post-project creation and key players in humanitarian mapping efforts.



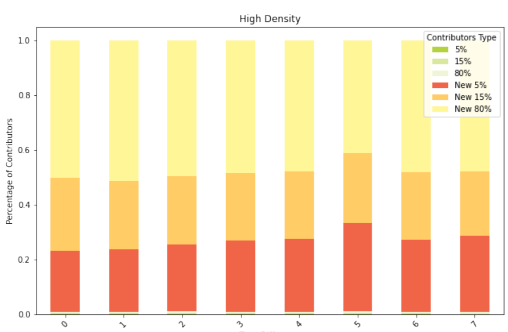
(a) Composition of Contributors in Projects with Very Low Density



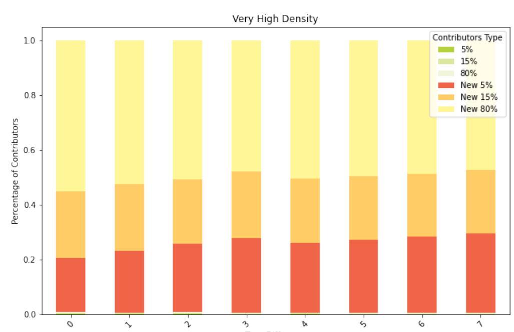
(b) Composition of Contributors in Projects with Low Density



(c) Composition of Contributors in Projects with Medium Density



(d) Composition of Contributors in Projects with High Density



(e) Composition of Contributors in Projects with Very High Density

Fig. 5. Composition of Contributors Across Five Population Density Categories

In project regions with lower population densities — categorized as Very Low and Low, the proportion of `New 5%` contributors can range from 24% to 43% of the total contributors. Additionally, the remaining composition includes 19% to 24% `New 15%` contributors and 31% to 48% `New 80%` contributors.

As population density increases, the inequality gap in contribution patterns widens. In project regions with medium population density, the proportion of `New 80%` contributors starts to expand, ranging from 41% to 50%. While the composition of `New 15%` remains consistent at 20%, the proportion of `New 5%` contributors shrinks to range from 23% to 32%.

In projects within higher densely populated areas — categorized as High and Very High — the proportion of `New 5%` contributors again remains moderately low by comparison to less densely populated areas, ranging from 20% to 32%. Furthermore, the `New 15%` group holds a similar share of the contributor composition, accounting for 18% to 25%. However, `New 80%` contributors comprise 43% to 57% of the contributor base in these high-density projects, which is much more than the proportion in low-density regions. In other words, in higher population dense regions, nearly half of the contributors are `New 80%` contributors, who contribute the least work and are less prolific in humanitarian OpenStreetMap activities. Conversely, in low population dense regions, we see larger proportions of `New 5%` contributors — who are the most prolific and contribute the largest amount of content — reaching 43% of contributors in some cases.

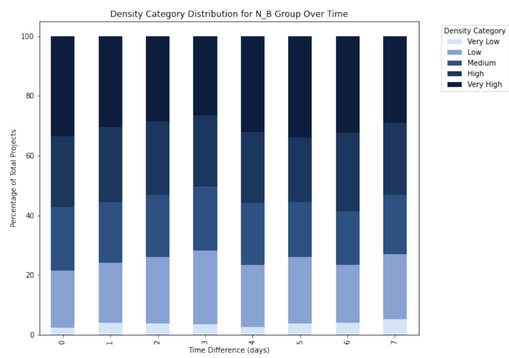
Stepping back, these results paint a surprising picture for the *underlying causes* of an increase in the Gini coefficient in lower population dense regions. The regions where the largest range of prolific contributors (`New 5%` contributors) contribute are the projects in lower population dense areas. In other words, prolific contributors focus at higher rates in less populated areas when new projects are created, and because these contributors are highly productive, this seems to result in more tightly concentrated work in the hands of relatively few prolific editors.

5.2.2 Extending the Long Tail with Increasing Population Density. Of course, participation across the population density spectrum is only one dimension of understanding the underlying dynamics at play in our findings from Section 5.1. While the charts above illustrate one mechanism of how a Gini coefficient might increase, they do not fully capture participation behavior. Therefore, we also wanted to understand how contribution behavior differs across different groups of contributors. Since the large majority of contributors within our 7-day window are "new" (or re-active) contributors, we focus our analysis here on only the `New 80%`, `New 15%`, and `New 5%` contributors. In Figure 6, we plot the extent to which these groups of contributors contribute to projects across the population density spectrum. Broadly, we find that across all three groups of contributors, a consistently minimal proportion participate in mapping work in very low-populated regions, with all percentages ranging from 2% to 4%.

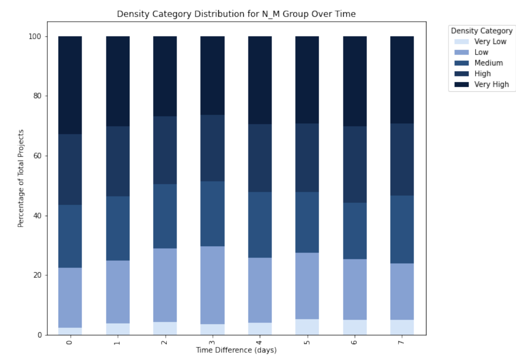
When looking at projects in low-populated regions, both `New 5%` and `New 15%` contributors have similar levels of involvement, ranging from 18% to 29%. However, the `New 5%` contributors exhibit a higher participation rate in low-populated project regions, ranging from 29% to 39%.

In medium and higher population density project regions, `New 80%` and `New 15%` contributors display comparable participation patterns: 17% to 20% are involved in medium-populated regions, 17% to 21% in higher-populated regions, and 27% to 32% in very high-populated regions. Combined, 71% to 79% of `New 80%` and `New 15%` contributors are engaged in projects within these higher density categories. In contrast, the `New 5%` contributors show lower engagement in medium to high population density projects: 15% to 20% in medium and higher-populated regions, and only 27% to 29% in very high-populated regions. Overall, 61% to 64% of `New 5%` contributors participate in projects in these medium to very high-density categories.

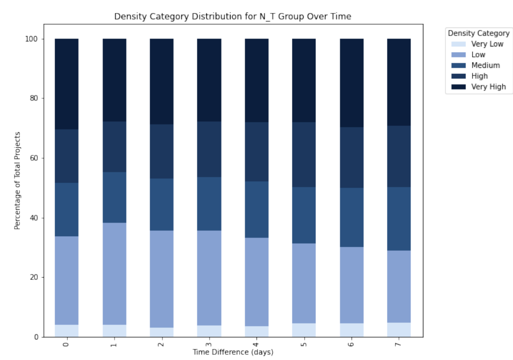
Taken holistically, these results tell a different story than before, namely, lower-productivity individuals (`New 15%` and `New 80%`) tend to join projects in high population density areas. While they do serve to help increase the number of contributors, their contributions do not rise proportionately. These contributors are typically less productive, extending the long tail of the Lorenz curve and further leading to a higher Gini coefficient. In other words, participation widens



(a) Distrbution of New 80% Contributors in Population Density Categories



(b) Distrbution of New 15% Contributors in Population Density Categories



(c) Distrbution of New 5% Contributors in Population Density Categories

Fig. 6. Contributor Behavior Dynamics in Population Density

with the creation of new projects, but the new contributors are largely less productive, resulting in an increased concentration of the overall work in the hands of relatively few more prolific editors.

6 DISCUSSION

Previous research [24, 63] has shown that using microtasking mechanisms through Tasking Manager can help overcome structural disparities, where regions that are less populated and have less content get better data coverage through remote humanitarian work. Our work adds nuance to our understanding of Humanitarian OpenStreetMap by highlighting that the "born, not made" bias, as well as unfortunate trends that advantage more densely populated places, persist within humanitarian mapping. Despite Humanitarian OpenStreetMap largely focusing on less populated areas worldwide, and despite the microtasking "anyone, anywhere can contribute" concept, our results suggest that the population density of the project regions still influences the number of contributors participating, and the *types of contributors* who contribute. We find evidence of a more nuanced set of processes that seem to facilitate population density biases in this setting.

For instance, we find that disparities in contribution along population density lines seems to be influenced by how prolific the contributors who focus on these projects are. Specifically, projects targeting higher populated regions tend to have a wider pool of contributors, but a larger proportion of these are among the bottom 80% of contributors, who only contribute 23% of the data overall. This results in the contributions within these regions being more highly concentrated in the hands of relatively few contributors. Despite having more contributors, most contributions are still made and concentrated within the top 5% of contributors. Conversely, although projects targeting less populated regions tend to have a narrower contributor pool overall, a higher proportion of these contributors are more prolific, belonging to the top 5% of contributors who contribute 54% of the data. This results in a more equitable distribution of contributions among a relatively small group of prolific contributors.

6.1 Self-focus Bias in Humanitarian OpenStreetMap Work

While Humanitarian OpenStreetMap has been broadly effective at helping to focus contributors' efforts in places where there is insufficient map data [24, 63], our findings complicate this success. We find that the effectiveness of participation varies with the population density of the regions where HOT projects are created. Project regions with higher population density tend to have more contributor participation, potentially resulting in better coverage than projects in lower population density regions. This pattern introduces another consideration: under the goal of humanitarian aid, such disparity may undermine the success of projects in low population density areas, especially when targeting large amounts of work for time-sensitive post-disaster efforts.

The trends we find here suggest that even though the HOT Tasking Manager, and HOT overall, are designed to help facilitate better remote mapping efforts in underrepresented areas, the efficacy of these tools remains somewhat constrained by population density. Unlike the "born, not made" patterns we see in broader peer production settings [57], our results here may reflect a self-focus bias [7], where regions with higher population density are more well-known and mainstream and thus attract more attention and participation from the public, compared to less known, marginalized regions.

By no means do these results suggest the HOT Tasking Manager is ineffective, merely that it is not as effective for some places as might be anticipated. There may be small design changes that could help better direct and focus HOT contributor effort and minimize the population density biases we find here. For example, Yin et al. [63] found that project attributes such as priority and difficulty can influence contributor participation. Project attributes such as "rural" or "extra eyes" may help draw the necessary attention to less population dense places and ensure that the Tasking Manager serves all regions well.

6.2 Follow-Up Actions Needed for Projects Targeting High Population Densities

Our findings reveal a paradoxical relationship in humanitarian mapping projects: while higher population density regions attract more contributions and contributors, they simultaneously exhibit higher concentration of contributions among fewer individuals [63]. This pattern demonstrates a clear trade-off between work quantity and the equity of perspectives represented in information production,highlighting concerns about the emergence of oligarchic structures in these digital communities [44, 46, 48, 50]. In our results, the concentration of contributions in the hands of relatively few contributors manifests in two ways. First, in high-density regions, despite having a larger pool of potential contributors, contribution patterns become more concentrated rather than distributed, suggesting that project creation actually facilitates this concentration rather than democratizing participation. Second, this concentration becomes self-reinforcing — as early and active contributors establish their presence, there is risk of establishing informal authority through their extensive contributions, creating increasing barriers for newcomers to achieve similar influence levels.

Moreover, a growing body of research has highlighted the risks of such contribution inequality in peer production systems. Haklay [18] warns about potential data quality risks when contributor pools become too homogeneous, while Thebault-Spieker et al. [57] demonstrates how these power-law dynamics can create geographic disparities in data coverage. Beyond data quality concerns, these concentration patterns can also influence how the community develops over time. Top contributors may inadvertently establish community norms that enforce their standards or viewpoints, potentially creating entry barriers for newcomers [18]. Further if these top-contributors stop participating, there is risk of creating serious challenges in data maintenance and production sustainability. Recent research by Li et al. [32] has explored the economic value of labor in platforms like OpenStreetMap, suggesting potential compensation mechanisms for contributors. However, implementing such mechanisms may exacerbate existing biases, transforming the Gini coefficient from a measure of contribution concentration to one of economic value concentration in peer production systems. These findings highlight the need for strategic interventions in projects focusing on higher population density regions, where severe contribution inequality exists.

To maintain community sustainability and humanitarian effort reliability, we suggest that sustaining engagement among the `New 80%` contributors — those contributors who do activate but do relatively little work — by focusing their efforts on maintenance and validation tasks. Of course, validation may necessitate additional expertise in OpenStreetMap, and different tools or interaction modalities may be more effective at building that expertise than the micro-tasking approach used in the Tasking Manager. While these contributors may not produce as much content as the `New 5%` contributors, more effective allocation of their efforts could enhance overall project sustainability and data quality, particularly crucial as mapping efforts directly influence disaster relief and local safety outcomes.

Moreover, we see opportunities for the CSCW community to better understand what this contribution inequality means for the community. Taking the context of humanitarian mapping efforts as an example, it is unclear if "equity in effort" is considered a goal of the community on par with, or perhaps even above, mapping coverage. While our results here illustrate a mechanism of disparity, whether that disparity is *harmful* is an open question.

6.3 Understanding the Participation Patterns Across Different Contributor Groups

Our results also show that the dynamics of how contributions are, or are not, concentrated in the hands of relatively few contributors seem to be driven by the variation in participation interest among different contributor types in Humanitarian OpenStreetMap projects. Specifically, more prolific contributors (top 5%) tend to participate in projects located in low-density regions, while less prolific contributors (bottom 80%) are more likely to contribute to projects in high-density regions. The factors, motivations, and goals that underpin participation across different contributor groups remain unclear. However, our preliminary investigation suggests that different types of contributors may be driven by different motivations, even within the broader context of humanitarian mapping work.

Prior work suggests that the goals and organization of projects may facilitate different contribution patterns as well. For instance, projects targeting long-term goals might have sustainable and high retention rates, whereas projects focusing on urgent goals might quickly reach a peak in contributions but have a lower likelihood of continued participation [5, 35]. Additionally, prior studies indicate that factors such as a sense of responsibility and the challenge involved also play significant roles in contributor motivation and participation [4, 51].

Overall, there may be opportunities to design the composition of contribution pools more contextually. Fully understanding these participation patterns can offer opportunities to enhance contributor experience and retention through the development of role-specific mapping tools and workflows. This approach is already exemplified in several editing support tools, such as Maproulette [61], which assigns specific editing tasks to OpenStreetMap users, fostering collaborative challenge-based participation. Such tools not only increase user engagement and improve mapping quality but also enhance contributor experience.

6.4 Future Work: Toward Contextualizing Contribution Inequity in Humanitarian Mapping

While evidence is clear that the HOT Tasking Manager mitigates crucial spatial data gaps in OpenStreetMap [24, 63], our results suggest it also exacerbates contribution inequality in the humanitarian community. Furthermore, the population density of project regions adds another dimension to these dynamics — more populated dense project regions tend to have a higher power-law distribution and more concentrated contributions.

However, it is not clear how the OpenStreetMap community broadly, or the Humanitarian OpenStretMap community more specifically, understands, interprets, and values the concentration of contributions in the hands of relatively few contributors. Of course, this is a common pattern of contribution within peer production settings, but prior work has also found that such power-law dynamics can risk data quality, participation barriers, and community sustainability Haklay [18], Thebault-Spieker et al. [57]. Conversely, Warncke-Wang et al. [60] suggest that frequent and active contributors, who have gained proficiency through experience, are more likely to contribute larger amounts of higher-quality data.

Looking forward, our study highlights the need for further research to consider community goals and values in how the research community evaluates power-law dynamics in peer production. Specifically, our work suggests the importance of contextualizing contribution concentration within specific contexts and communities, such as Humanitarian OpenStreetMap. Aligning our scholarly evaluation of communities like Humanitarian OpenStreetMap with the community's own goals enables more focused research impact and contributions. For instance, our work here may have implications for how the Humanitarian OpenStreetMap community continues to develop, including issues of growth, diversity of perspectives, representation, and data quality and coverage. Furthermore, this research highlights the importance of examining how contribution concentration aligns with varying project objectives, such as those distinguishing long-term projects from short-term projects or mission-based from event-based initiatives. While the CSCW community has largely been the intellectual home for scholarship on peer production and patterns of collaborative data production more broadly, the relationships between our work and community values and goals remain somewhat unclear.

7 CONCLUSION

In conclusion, our findings add nuance to the evaluation of humanitarian mapping activities, showing that even though most work is conducted remotely, the number of contributors and power-law dynamics are still associated with the population density of project regions, similar to the broader OpenStreetMap community. Furthermore, by investigating the variation of power-law dynamics across differently populated projects, we uncover the participation patterns of different contributor groups. By unpacking these dynamics, our work underscores the importance of considering geographic context and community dynamics in the design and implementation of humanitarian mapping initiatives. Additionally, we pave the way for more informed decision-making and more effective humanitarian interventions, aiding the community and practitioners in continuous and sustainable humanitarian OpenStreetMap efforts.

REFERENCES

[1] Irasema Alcántara-Ayala. 2002. Geomorphology, natural hazards, vulnerability and prevention of natural disasters in developing countries. *Geomorphology* 47 (2002), 107–124.

[2] Nezih Altay and Melissa T. Labonte. 2014. Challenges in humanitarian information management and exchange: evidence from Haiti. *Disasters* 38 Suppl 1 (2014), S50–72.

[3] Joshua Angrist and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion.*

[4] Ofer Arazy, Hila Lifshitz-Assaf, and Adam Balila. 2018. Neither a Bazaar nor a cathedral: The interplay between structure and agency in Wikipedia's role system. *Journal of the Association for Information Science and Technology* 70 (2018). https://doi.org/10.1002/asi.24076

[5] Daniel Bégin, Rodolphe Devillers, and Stéphane Roche. 2018. The life cycle of contributors in collaborative online communities -the case of OpenStreetMap. *International Journal of Geographical Information Science* 32 (2018), 1611 – 1630. https://doi.org/10.1080/13658816.2018.1458312

[6] Camille Cobb, Ted McCarthy, Annuska Perkins, Ankitha Bharadwaj, Jared Comis, Brian Do, and Kate Starbird. 2014. Designing for the deluge: understanding & supporting the distributed, collaborative work of crisis volunteers. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. ACM, Baltimore Maryland USA, 888–899. https://doi.org/10.1145/2531602.2531712

[7] Maitraye Das, Brent J. Hecht, and Darren Gergle. 2019. The Gendered Geography of Contributions to OpenStreetMap: Complexities in Self-Focus Bias. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (2019).

[8] Claribelle Deveza. 2023. *Tesla Owners Improve Smart Summon Routes by Updating Open Street Maps.* https://www.tesmanian.com/blogs/tesmanian-blog/tesla-owners-smart-summon-routes-open-street-maps-full-self-driving

[9] Corey Dickinson. 2023. Inside the 'Wikipedia of Maps,' Tensions Grow Over Corporate Influence. https://www.bloomberg.com/news/articles/2021-02-19/openstreetmap-charts-a-controversial-new-direction. [Accessed 16-July-2023].

[10] Martin Dittus, Giovanni Quattrone, and Licia Capra. 2016. Analysing Volunteer Engagement in Humanitarian Mapping: Building Contributor Communities at Large Scale. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (2016).

[11] Martin Dittus, Giovanni Quattrone, and Licia Capra. 2017. Mass Participation During Emergency Response: Event-centric Crowdsourcing in Humanitarian Mapping. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (2017).

[12] Featured HOT projects. 2024. Humanitarian OpenStreetMap Team Projects. https://www.hotosm.org/projects/, Last accessed on 2024-10-27.

[13] Frank A. Farris. 2010. The Gini Index and Measures of Inequality. *The American Mathematical Monthly* 117, 10 (2010), 851. https://doi.org/10.4169/000298910x523344

[14] Gary Esworthy. 2016. COMPARING OPENSTREETMAP AND WIKIPEDIA EDIT PATTERNS DURING MAJOR EVENTS. http://mars.gmu.edu/bitstream/handle/1920/10251/Esworthy_thesis_2016.pdf;sequence=1, Last accessed on 2022-07-14.

[15] Michael Goodchild. 2007. Citizens as Sensors: The World of Volunteered Geography. *GeoJournal* 69 (08 2007), 211–221. https://doi.org/10.1007/s10708-007-9111-y

[16] Michael F. Goodchild. 2004. The Validity and Usefulness of Laws in Geographic Information Science and Geography. *Annals of the Association of American Geographers* 94 (2004), 300 – 303.

[17] Muki Haklay. 2013. Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation. In *Crowdsourcing Geographic Knowledge*, Daniel Sui, Sarah Elwood, and Michael Goodchild (Eds.). Springer Netherlands, Dordrecht, 105–122. https://doi.org/10.1007/978-94-007-4587-2_7

[18] Muki Haklay. 2016. Why is participation inequality important.

[19] Muki Haklay and Patrick Weber. 2008. OpenStreetMap: User-Generated Street Maps. *Haklay, M. and Weber, P. (2008) OpenStreetMap: user-generated street maps. IEEE Pervasive Computing, 7 (4). pp. 12-18. ISSN 15361268* 7 (10 2008). https://doi.org/10.1109/MPRV.2008.80

[20] Mordechai (Muki) Haklay, Vyron Antoniou, Sofia Basiouka, Robert Soden, Vivien Deparday, Ryan Sheely, and Peter Mooney. 2018. Identifying Success Factors in Crowdsourced Geographic Information Use in Government.

[21] Mordechai (Muki) Haklay, Sofia Basiouka, Vyron Antoniou, and Aamer Ather. 2010. How Many Volunteers Does it Take to Map an Area Well? The Validity of Linus' Law to Volunteered Geographic Information. *The Cartographic Journal* 47 (2010), 315 – 322.

[22] Brent Hecht and Darren Gergle. 2009. Measuring self-focus bias in community-maintained knowledge repositories. In *Proceedings of the fourth international conference on Communities and technologies - C&T '09*. ACM Press, University Park, PA, USA, 11. https://doi.org/10.1145/1556460.1556463

[23] Andreas Henrich and Volker Lüdecke. 2009. Measuring Similarity of Geographic Regions for Geographic Information Retrieval. In *Advances in Information Retrieval*, Mohand Boughanem, Catherine Berrut, Josiane Mothe, and Chantal Soule-Dupuy (Eds.). Vol. 5478. Springer Berlin Heidelberg, Berlin, Heidelberg, 781–785. https://doi.org/10.1007/978-3-642-00958-7_85 Series Title: Lecture Notes in Computer Science.

[24] Benjamin Herfort, Sven Lautenbach, João Porto de Albuquerque, Jennings Anderson, and Alexander Zipf. 2021. The evolution of humanitarian mapping within the OpenStreetMap community. *Scientific Reports* 11, 1 (Dec. 2021), 3037. https://doi.org/10.1038/s41598-021-82404-z

[25] hotosm. 2010. hotosm introduction. https://www.hotosm.org/, Last accessed on 2024-05-14.

[26] Humanitarian OpenStreetMap Team. 2015. DISASTER ACTIVATION: NEPAL EARTHQUAKE 2015. https://www.hotosm.org/projects/nepal_2015_earthquake_response, Last accessed on 2024-07-01.

[27] Guido Imbens. 2023. Causal Inference in the Social Sciences. *Annual Review of Statistics and Its Application* 11 (11 2023). https://doi.org/10.1146/annurev-statistics-033121-114601

[28] Isaac L. Johnson, Yilun Lin, Toby Jia-Jun Li, Andrew Hall, Aaron L Halfaker, Johannes Schöning, and Brent J. Hecht. 2016. Not at Home on the Range: Peer Production and the Urban/Rural Divide. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (2016).

[29] Jamal Jokar Arsanjani, Peter Mooney, Marco Helbich, and Alexander Zipf. 2015. An Exploration of Future Patterns of the Contributions to OpenStreetMap and Development of a Contribution Index. *Transactions in GIS* (01 2015).

[30] Brian Keegan, Darren Gergle, and Noshir Contractor. 2013. Hot Off the Wiki: Structures and Dynamics of Wikipedia's Coverage of Breaking News Events. *American Behavioral Scientist* 57, 5 (May 2013), 595–622. https://doi.org/10.1177/0002764212469367

[31] Aniket Kittur, Ed Chi, Bryan A. Pendleton, and Bongwon Suh. 2006. Power of the Few vs . Wisdom of the Crowd : Wikipedia and the Rise of the Bourgeoisie.

[32] Hanlin Li, Brent Hecht, and Stevie Chancellor. 2022. *Measuring the Monetary Value of Online Volunteer Work.* Technical Report arXiv:2205.14528. arXiv. http://arxiv.org/abs/2205.14528 arXiv:2205.14528 [cs] type: article.

[33] Liat Clark. 2014. The race to contain West Africa's Ebola outbreak. https://www.wired.co.uk/article/ebola-open-street-map, Last accessed on 2022-07-14.

[34] Allen Yilun Lin, Kate Kuehl, Johannes Schöning, and Brent J. Hecht. 2017. Understanding "Death by GPS": A Systematic Study of Catastrophic Incidents Associated with Personal Navigation Technologies. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (2017).

[35] Zarif Mahmud, Aarjav Chauhan, Dipto Sarkar, and Robert Soden. 2022. Revisiting Engagement in Humanitarian Mapping: An Updated Analysis of Contributor Retention in OpenStreetMap. *CHI Conference on Human Factors in Computing Systems Extended Abstracts* (2022).

[36] Afra Jahanbakhsh Mashhadi, Giovanni Quattrone, Licia Capra, and Peter Mooney. 2012. On the accuracy of urban crowd-sourcing for maintaining large-scale geospatial databases. In *WikiSym '12.*

[37] Nikola Milojevic-Dupont, Nicolai Hans, Lynn Kaack, Marius Zumwald, François Andrieux, Daniel Soares, Steffen Lohrey, Peter-Paul Pichler, and Felix Creutzig. 2020. Learning from urban form to predict building heights. *PLOS ONE* 15 (12 2020), e0242010. https://doi.org/10.1371/journal.pone.0242010

[38] Abhishek Nagaraj. 2017. Information Seeding and Knowledge Production in Online Communities: Evidence from OpenStreetMap. *IRPN: Open Collaborative Innovation (Topic)* (2017).

[39] Pascal Neis and Alexander Zipf. 2012. Analyzing the Contributor Activity of a Volunteered Geographic Information Project — The Case of OpenStreetMap. *ISPRS International Journal of Geo-Information* 1 (12 2012), 146–165. https://doi.org/10.3390/ijgi1020146

[40] Andreas Olden and Jarle Men. 2022. The triple difference estimator. *The Econometrics Journal* 25 (03 2022). https://doi.org/10.1093/ectj/utac010

[41] Leysia Palen, Robert Soden, T. Jennings Anderson, and Mario Barrenechea. 2015. Success & Scale in a Data-Producing Organization: The Socio-Technical Evolution of OpenStreetMap in Response to Humanitarian Events. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems.* ACM, Seoul Republic of Korea, 4113–4122. https://doi.org/10.1145/2702123.2702294

[42] Katherine Panciera, Aaron Halfaker, and Loren Terveen. 2009. Wikipedians are born, not made: a study of power editors on Wikipedia. In *Proceedings of the 2009 ACM International Conference on Supporting Group Work* (Sanibel Island, Florida, USA) *(GROUP '09).* Association for Computing Machinery, New York, NY, USA, 51–60. https://doi.org/10.1145/1531674.1531682

[43] Paul Sawers. 2018. OpenStreetMap now lets you edit maps with Microsoft's Bing Streetside imagery. https://venturebeat.com/mobile/openstreetmap-now-lets-you-edit-maps-with-microsofts-bing-streetside-imagery/,.

[44] Shruti Phadke, Mattia Samory, and Tanushree Mitra. 2022. Pathways through Conspiracy: The Evolution of Conspiracy Radicalization through Engagement in Online Conspiracy Discussions. https://doi.org/10.48550/arXiv.2204.10729

[45] Thiago Henrique Poiani, Roberto Dos Santos Rocha, Livia Castro Degrossi, and Joao Porto De Albuquerque. 2016. Potential of Collaborative Mapping for Disaster Relief: A Case Study of OpenStreetMap in the Nepal Earthquake 2015. In *2016 49th Hawaii International Conference on System Sciences (HICSS).* IEEE, Koloa, HI, 188–197. https://doi.org/10.1109/HICSS.2016.31

[46] Emiel Rijshouwer, Justus Uitermark, and Willem Koster. 2021. Wikipedia: a self-organizing bureaucracy. *Information, Communication Society* 26 (10 2021), 1–18. https://doi.org/10.1080/1369118X.2021.1994633

[47] Jonathan Roth, Pedro Sant'Anna, Alyssa Bilinski, and John Poe. 2023. What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics* 235 (04 2023). https://doi.org/10.1016/j.jeconom.2023.03.008

[48] Nathan Schneider. 2021. Admins, Mods, and Benevolent Dictators for Life: The Implicit Feudalism of Online Communities. https://doi.org/10.33767/osf.io/sf432

[49] Abel Serrano, Javier Arroyo, and Samer Hassan. 2018. Participation Inequality in Wikis: A Temporal Analysis Using WikiChron. In *Proceedings of the 14th International Symposium on Open Collaboration* (Paris, France) *(OpenSym '18).* Association for Computing Machinery, New York, NY, USA, Article 12, 7 pages. https://doi.org/10.1145/3233391.3233536

[50] Aaron Shaw and Benjamin Hill. 2014. Laboratories of Oligarchy? How the Iron Law Extends to Peer Production. *Journal of Communication* 64 (03 2014). https://doi.org/10.1111/jcom.12082

[51] Francis M. Sim and Bruce J. Biddle. 1982. Role Theory: Expectations, Identities, and Behaviors. http://www.loc.gov/catdir/toc/els031/79006930.html

[52] Robert Soden and Leysia Palen. 2014. From Crowdsourced Mapping to Community Mapping: The Post-earthquake Work of OpenStreetMap Haiti. In *COOP 2014 - Proceedings of the 11th International Conference on the Design of Cooperative Systems, 27-30 May 2014, Nice (France),* Chiara Rossitto, Luigina Ciolfi, David Martin, and Bernard Conein (Eds.). Springer International Publishing, Cham, 311–326. https://doi.org/10.1007/978-3-319-06498-7_19

[53] Kate Starbird and Leysia Palen. 2011. "Voluntweeters": self-organizing by digital volunteers in times of crisis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, Vancouver BC Canada, 1071–1080. https://doi.org/10.1145/1978942.1979102

[54] Kate Starbird and Leysia Palen. 2013. Working and sustaining the virtual "Disaster Desk". In *Proceedings of the 2013 conference on Computer supported cooperative work - CSCW '13.* ACM Press, San Antonio, Texas, USA, 491. https://doi.org/10.1145/2441776.2441832

[55] Bongwon Suh, Gregorio Convertino, Ed H. Chi, and Peter Pirolli. 2009. The singularity is not near: slowing growth of Wikipedia. In *Proceedings of the 5th International Symposium on Wikis and Open Collaboration - WikiSym '09.* ACM Press, Orlando, Florida, 1. https://doi.org/10.1145/1641309.1641322

[56] Jacob Thebault-Spieker, Aaron Halfaker, Loren G. Terveen, and Brent Hecht. 2018. Distance and Attraction: Gravity Models for Geographic Content Production. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems.* ACM, Montreal QC Canada, 1–13. https://doi.org/10.1145/3173574.3173722

[57] Jacob Thebault-Spieker, Brent Hecht, and Loren Terveen. 2018. Geographic Biases are 'Born, not Made': Exploring Contributors' Spatiotemporal Behavior in OpenStreetMap. In *Proceedings of the 2018 ACM Conference on Supporting Groupwork*. ACM, Sanibel Island Florida USA, 71–82. https://doi.org/10.1145/3148330.3148350

[58] Jacob Thebault-Spieker, Brent J. Hecht, and Loren G. Terveen. 2018. Geographic Biases are 'Born, not Made': Exploring Contributors' Spatiotemporal Behavior in OpenStreetMap. *Proceedings of the 2018 ACM International Conference on Supporting Group Work* (2018).

[59] Rajan Vaish, Keith Wyngarden, Jingshu Chen, Brandon Cheung, and Michael S. Bernstein. 2014. Twitch crowdsourcing: crowd contributions in short bursts of time. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2014).

[60] Morten Warncke-Wang, Vladislav R. Ayukaev, Brent Hecht, and Loren G. Terveen. 2015. The Success and Failure of Quality Improvement Projects in Peer Production Communities. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM, Vancouver BC Canada, 743–756. https://doi.org/10.1145/2675133.2675241

[61] OpenStreetMap Wiki. 2024. MapRoulette — OpenStreetMap Wiki,. https://wiki.openstreetmap.org/w/index.php?title=MapRoulette&oldid=2664306 [Online; accessed 3-July-2024].

[62] Anran Yang, Hongchao Fan, Ning Jing, Yeran Sun, and Alexander Zipf. 2016. Temporal Analysis on Contribution Inequality in OpenStreetMap: A Comparative Study for Four Countries. *ISPRS Int. J. Geo-Inf.* 5 (01 2016).

[63] Yaxuan Yin, Longjie Guo, and Jacob Thebault-Spieker. 2024. Productivity or Equity? Tradeoffs in Volunteer Microtasking in Humanitarian OpenStreetMap. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW1, Article 113 (apr 2024), 34 pages. https://doi.org/10.1145/3637390

[64] Haiyi Zhu, Robert Kraut, and Aniket Kittur. 2012. Effectiveness of shared leadership in online communities. In *Proceedings of the ACM 2012 conference on computer supported cooperative work*. 407–416.

[65] Matthew Zook, Mark Graham, Taylor Shelton, and Sean Gorman. 2010. Volunteered Geographic Information and Crowdsourcing Disaster Relief: A Case Study of the Haitian Earthquake. *World Medical Health Policy* 2 (01 2010). https://doi.org/10.2202/1948-4682.1069

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009