Unraveling the impact of the Hamas-Israel war on short and mid-term population internal displacements in the Gaza strip

Abstract

In the wake of a violent incident, such as the 2023 conflict between Israel and Hamas, accurate estimation of population displacement becomes imperative for orchestrating an effective crisis response. However, this task is beset with challenges stemming from outdated data and the practical difficulties associated with data collection in conflict zones. Our study harnesses Meta's marketing API to gauge population dynamics, with a specific focus on audience sizes on Facebook and Instagram, disaggregated by age, gender, and subnational regions. In addition to providing near-real-time assistance in addressing humanitarian crises, this research extends prior research developed for the Ukraine war by scrutinizing the impact of conflicts on internal migration and utilizes linguistic attributes to further differentiate displacement among various linguistic groups. Furthermore, we suggest a statistical methodology to combine indicators of internal displacement derived from two sources, Facebook and Instagram, thereby creating a more representative picture of the underlying population.

1 Introduction

In the wake of the Hamas intervention in Israel on October 7th 2023, estimating population forced displacement at the subnational level has been critical to plan crisis responses from setting up temporary shelter accommodation to delivering adequate humanitarian aid. However it is an exceptionally challenging task because of the rapid obsolescence of preexisting population data and the impracticality of conducting new data collection in conflict-ridden regions (Checchi et al. 2013; Ratnayake, Abdelmagid, and Dooley 2022; Leasure et al. 2023). The acute phase of conflict exacerbates this challenge, as displacement trends change rapidly from day to day.

To address this challenge, various methods have been employed to assess population dynamics. Techniques such as satellite and aerial imagery offered the ability to map buildings and total settled areas for assessing the scale of destruction ("Guerre Israël-Hamas : la carte des destructions, d'une ampleur inédite, en onze jours de guerre à Gaza" 2023) or tracking crowds gatherings at borders (Leatherby et al. 2023). This can lay the groundwork for population estimation when field data from hard-to-access areas can still be collected to determine current population densities, such as people per building (Checchi et al. 2013). Mobile phone location data also known as call detail records has been proven to be a solution to monitor population flows without needing to go to the field and thus have been used by the Israel Defence Forces to monitor the displacement of Gazans following the military order for civilians to evacuate North Gaza (Kingsley and Bergman 2023). This however requires strong collaboration with phone companies which is not often available to the wider public.

Digital traces left behind when people use internet-based services presented a novel opportunity to fill the data gaps, offering anonymized and aggregated information on user locations in near real-time. This emerging field of research harnessing various forms of digital trace data, has already focused on migration dynamics, by leveraging sources including email IP addresses (Zagheni and Weber 2012) or geolocated tweets (Zagheni et al. 2014). These digital data sources held the advantage of providing real-time, dynamic measurements without the need for extensive primary data collection on a population that by definition is harder to reach due to its higher mobility rate.

The study under consideration focused on a specific source of digital trace data: aggregate data from

Meta's marketing tools, accessible through its marketing application programming interface (API). This data source provided estimates of current audience sizes on Facebook, segmented by age, sex, and subnational geographic areas, catering to targeted advertising on the social media platform. Meta's marketing API had been previously harnessed to quantify stocks of international migrants (Rampazzo, Rango, and Weber 2023) and monitor refugee flows in crisis situations, such as the exodus of people from Venezuela in 2018–2019 (Palotti et al. 2020). However nowcasting methods developed by Leasure et al. were used to monitor the internal displacement of Ukrainians following the 2022 war sparked by the Russian invasion to help deliver targeted humanitarian aid. They built a real-time monitoring system using Facebook daily active users count calibrated with pre-conflict population data to estimate daily population sizes beyond Facebook users for age-sex demographic groups within subnational administrative units (Leasure et al. 2023).

Our research builds upon this prior work focusing on utilizing Meta marketing API for monitoring conflict-driven displacement processes. In our approach, we extend this research in three key ways. First, our study does not restrict the examination of internal population movement to a single country. Instead, we broaden the scope by investigating how the resurgence of violence in the Middle East impacts the internal migration patterns of Lebanese, Israeli, Palestinian, and Egyptian populations, with a specific focus on age and gender demographics. Second, we introduce a novel component to the model that combines the analysis of audience changes from both the Facebook and Instagram platforms. This integrated approach aims to provide a more representative population sample, particularly for younger generations who have shifted away from Facebook, towards Instagram. Lastly, our research introduces an innovative method for dissecting conflict-induced internal migration patterns. We achieve this by leveraging the language attribute available through the Meta Marketing API. This allows us to disentangle the movement of the Hebrew-speaking population from that of the Arabic-speaking population, particularly in regions such as Israeli and West Bank territories.

2 Data

We systematically gathered data on the number of active users on Facebook and Instagram within specific geographic subregions and at the city level in Israel, Palestine, Lebanon, Jordan, and Egypt. Our data collection efforts commenced on October 9th, 2023, utilizing the Meta marketing API (Meta 2023) and were facilitated through the utilization of the pySocialWatcher software (Araujo et al. 2017). The data we amassed pertained to individuals aged 13 and above, and it was further disaggregated into distinct age brackets for males and females. These age categories encompassed: 13+, 18+, 18–60, 15–49, 15–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–59, 60–64, and 65+. Our dataset encompassed all users of Facebook and Instagram whose "recent" location was determined to be within a specific city or region based on their device data, in accordance with Meta's methodology. These data collection procedures were systematically scheduled to operate continuously on a server with centralized database storage, maintained in a standardized format. Notably, our observations indicated an initial decline in overall user counts, particularly in the Gaza strip, which reflects challenging material conditions that hinder access to the internet. Nevertheless, it's important to note that a complete blackout of usage was not observed.

The data we sourced from the Meta marketing API is publicly accessible to individuals with Meta accounts and serves the primary purpose of gauging audience sizes for targeted advertising campaigns on the social media platform. These counts pertain to active Facebook accounts, as opposed to individual users, and are susceptible to potential biases arising from factors such as multiple accounts and fraudulent profiles. Additionally, the data contains occasional anomalous entries, including sporadic zeros. The specific data cleaning and model-based estimation techniques employed by Meta to generate these counts remain undisclosed, introducing an element of uncertainty regarding data quality, the presence of random noise, and the potential for unaccounted biases.

3 Method

Governorate-level population estimates were derived from number of active users utilizing an adapted methodology outlined by Leasure et al. (2023) .

Commencing October 13, 2023, we gathered daily data on active Facebook users from five distinct cities: Beit Lahia (North Gaza Governorate), Gaza City (Gaza City Governorate), Bureij (Deir Al-Balah Governorate), Khan Younis (Khan Younis Governorate), and Rafah (Rafah Governorate). These data served as proxies for population dynamics within the respective governorates. Ratios were computed between the count of active users in each city on October 13, 2023, and the established baseline population figures for the corresponding governorate, referencing data from 2024 (Palestinian Authority Ministry of Planning, n.d.; Palestinian Central Bureau of Statistics, n.d.; Palestinian Central Bureau of Statistics, n.d.).

Incorporating non-Facebook users necessitated estimating ratios between the population of each governorate and the count of active Facebook users from the corresponding cities. These ratios were calculated for each age-sex group within each governorate using baseline Facebook user counts from October 13, 2023. Population estimates were derived by aggregating the user-to-population ratios with smoothed estimates of active user counts over a three-day moving window. To accommodate temporal variations in social media usage rates, a daily national scaling factor was applied to ensure accurate summation of population estimates across governorates, presupposing no border crossings into or out of the Gaza Strip.

We employed this methodology to generate adult population estimates from three distinct daily metrics of active Facebook users: daily active users, upper and lower bounds of monthly active users. The preferred model, based on the lower bound of monthly active users, was selected due to its resilience against network outages and daily fluctuations, although it may exhibit slower responsiveness to population shifts compared to daily active user counts.

Child population estimates (under 18 years) were indirectly inferred, given the scarcity of Facebook users in this demographic, by assuming proportional changes to adult female populations. Utilizing baseline ratios of boys and girls per adult woman within each governorate, estimates of child populations were derived from daily estimates of adult female populations.

Although we explored adjustments to daily population estimates to accommodate fluctuations in network connectivity, as gauged by NetBlocks.org NetBlocks (n.d.), incorporating such adjustments proved impractical due to issues arising during periodic network outages when Facebook user counts exhibited limited decline, potentially attributed to data smoothing by Meta. Consequently, adjustments for network connectivity were omitted from the final analysis.

4 Results

Figure 1 illustrates the evolution of population totals at governorate level in the Gaza strip as derived from Facebook user count as the war unfolds (timeline events derived from "Timeline: Major Events During 100 Days of War Between Israel and Hamas" (2024) and A et al. (2024)). The legend is organised from the Northest to Southest governorate. If before the war, the Gaza governorate concentrated the maximum number of people (767 828 - ie 33% of the Gaza strip population), its population massively decreases following the Israeli Defense Forces (IDF) evacuation order of the 13th October reaching an all time low on the 4th December of 58 881 people - a 90% decrease of its population. When the IDF ground invasion starts in the North of the Gaza strip on the 27th October, we see a gradual increase of the population of the Southern governorates, reaching its peak on the 4th of December for Khan Younis just before the IDF invades it and on the 7th February for Rafah, before the Israeli governement announced its plan to invade Rafah. We see clearly the impact of the truce on the 21st November which allowed Palestinian to return to their home and gather their belonging from the most impacted governorates as shown by the short bump in population number for the Gaza and the North governorate.

Figure 2 decomposes the forced displacement in the Gaza strip by gender and represents it as the percentage of population change compared to day 1 in order to account for different total population sizes of women and men. We see that in the first two weeks of the conflict displacement patterns are similar for women and men. However afterwards men are more mobile than women as seen by the almost always lower population change percentage of women population (in average women governorate level population has changed by 30% compared to 46% for the men). The exception is for the two governorates that have experienced the

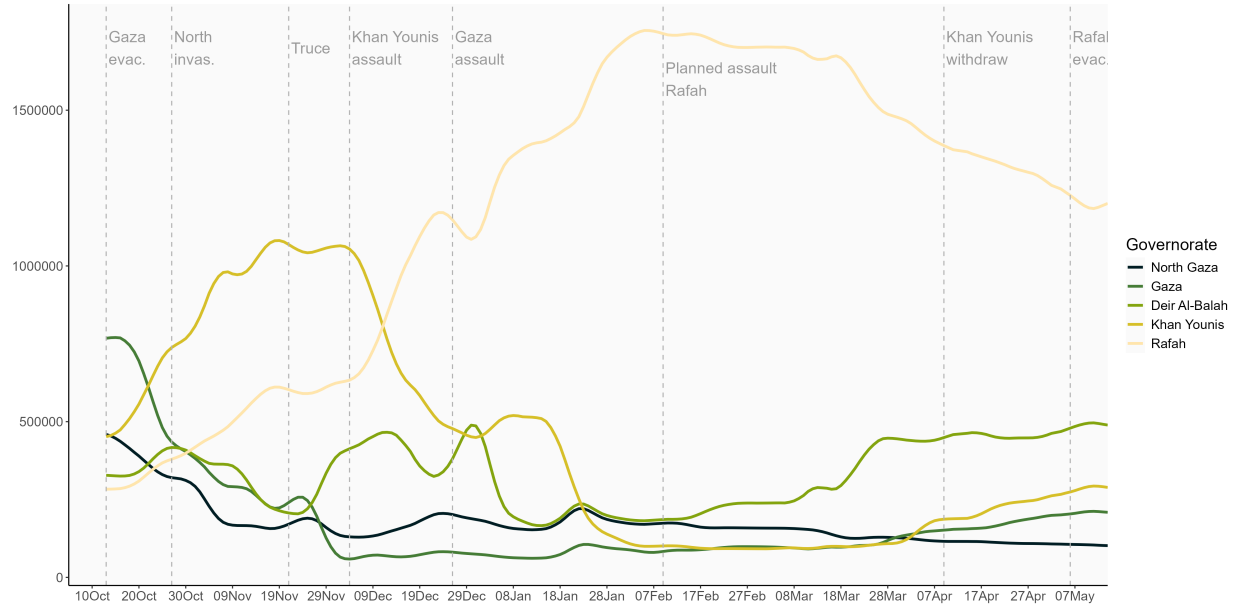

Figure 1: Governorate total population in the Gaza strip since 13th October 2023

most prolonged decline of population (Gaza and North Gaza) where the men population decline at a slower pace than women population.

5 Modelling limitations

Children displacement

Due to Facebook's policy prohibiting user accounts for individuals under 13 years old, we indirectly inferred the population sizes of children under 18 by extrapolating from daily population estimates for adult women (ages 18 and above), assuming consistent proportionalities with child populations. Therefore we are making the underlying assumption that children mobility is the same as their mothers.

Issue with input data

1. As mentionned in previous section, our methodology does not explicitly accommodate fluctuations in network connectivity. While resilient to short-term network disruptions and proportional declines in connectivity across all governorates, it may lead to underestimation of populations in governorates experiencing higher decreases in network connectivity compared to others. We explore the use of a connectivity indicator but its volatility had no counterpart in the observed user count. Given the size of the territory and the use of monthly active user count we are confident that this assumption should not have a large impact on the population number.
2. Pre-conflict baseline population data by age and sex within the five governorates of Gaza were not directly available. We derived this information by combining governorate-level (admin-2) total population sizes by sex with age-sex proportions for Gaza (admin-1). For age-groupings deviating from 5-year intervals (e.g., 18+), we applied single-year age proportions for Palestine (admin-0).
3. Periods of zero or low daily counts of active users, falling below Facebook's privacy threshold of 1,000 monthly active users, present challenges in discerning trends in usage. To mitigate these effects, we applied data smoothing over a 5-day moving window and replaced values of 1,000 active users with 500.

Constant national totals

Figure 2: Governorate adult population by gender in the Gaza strip since 13th October 2023

Our analysis relies on constant population totals to enable temporal variations in social media usage rate. It assumed therefore no crossings of the Gaza border, ensuring that admin-2 population estimates consistently sum up to the total baseline population for the Gaza Strip projected to 2024 by COD-PS. This assumption could be revisited in the future pending availability of border crossing data if those happens. And it does not account for casualties that, despite being massive, have been reported at less than 1.5% of the total population (34 844 Palestinians) according to United Nations as of the 8th of May 2024 United Nations Office for the Coordinational of Humanitarian Affairs (2024).

6 Next steps

The focus in recent months has been on providing weekly estimates to humanitarian response teams in the Gaza Strip.

In the time left for research development, our team has been developing a prototye Bayesian model to integrate observations of user counts retrieved both from Facebook and Instagram platform in order to strengthen the signal of population displacement. In the last month, we have also started a data pipeline to retrieve municipality level user counts, laying the ground to estimates population displacement at a finer spatial scale. We also gathered time series data on conflict locations (Raleigh et al. 2010), building damages (Asi et al. 2024), airstrikes (Airwars 2023) and military operations (Institute for the Study of War and AEI's Critical Threats Project 2024). And as previously mentionned we have collected a wealth of social media data yet to be explored such as population movement in the surrounding areas by age, sex and language spoken (Hebrew vs Arab vs English).

We are interested by different research directions:

1. Nowcasting with higher reliability and at finer spatial scale, daily population displacement in the Gaza strip by leveraging social media data from different platforms to strengthen population-related data signal as well as combining it with additional covariates to inform about the direction of population movement
2. Investigating the demographics of war-induced displacement and its relationship with external factors over time. For example do population subgroups react differently by age and sex to airstrikes? Is there a tipping point on building damages that stop being a push factor? Has the Hamas-Israel war pushed for relocation further away from the strip on Israeli side?

Bibliography

Airwars. 2023. "Methodology Note: Civilian Harm from Explosive Weapons Use in Gaza." https://airwars.org/research/methodology-note-civilian-harm-from-explosive-weapons-use-in-gaza/.

Araujo, Matheus, Yelena Mejova, Ingmar Weber, and Fabricio Benevenuto. 2017. "Using Facebook Ads Audiences for Global Lifestyle Disease Surveillance: Promises and Limitations." In, 253–57.

Asi, Yara, David Mills, P. Gregg Greenough, Dennis Kunichoff, Saira Khan, Jamon Van Den Hoek, Corey Scher, et al. 2024. "'Nowhere and No One Is Safe': Spatial Analysis of Damage to Critical Civilian Infrastructure in the Gaza Strip During the First Phase of the Israeli Military Campaign, 7 October to 22 November 2023." *Conflict and Health* 18 (1): 24. https://doi.org/10.1186/s13031-024-00580-x.

A, y, a, B, a, t, r, a, w, and y. 2024. "A Timeline of Events Leading up to Israel's Rafah Offensive." *NPR*, May. https://www.npr.org/2024/05/08/1249657561/rafah-timeline-gaza-israel-hamas-war.

Checchi, Francesco, Barclay T. Stewart, Jennifer J. Palmer, and Chris Grundy. 2013. "Validity and feasibility of a satellite imagery-based method for rapid estimation of displaced populations." *International Journal of Health Geographics* 12 (January): 4. https://doi.org/10.1186/1476-072X-12-4.

"Guerre Israël-Hamas : la carte des destructions, d'une ampleur inédite, en onze jours de guerre à Gaza." 2023. *Le Monde.fr*, October. https://www.lemonde.fr/international/article/2023/10/19/guerre-israel-hamas-la-carte-des-destructions-d-une-ampleur-inedite-en-onze-jours-de-guerre-a-gaza_6195433_3210.html.

Institute for the Study of War, and AEI's Critical Threats Project. 2024. "Interactive Map: Israel's Operation in Gaza." https://storymaps.arcgis.com/stories/2e746151991643e39e64780f0674f7dd.

Kingsley, Patrick, and Ronen Bergman. 2023. "Tracking Cellphone Data by Neighborhood, Israel Gauges Gaza Evacuation." *The New York Times*, October. https://www.nytimes.com/2023/10/16/world/middleeast/gaza-invasion-israel-cellphone-data.html.

Leasure, Douglas R., Ridhi Kashyap, Francesco Rampazzo, Claire A. Dooley, Benjamin Elbers, Maksym Bondarenko, Mark Verhagen, et al. 2023. "Nowcasting Daily Population Displacement in Ukraine Through Social Media Advertising Data." *Population and Development Review* 49 (2): 231–54. https://doi.org/10.1111/padr.12558.

Leatherby, Lauren, Karen Yourish, Elena Shao, Eli Murray, Scott Reinhard, and Josh Holder. 2023. "Maps: Tracking the Attacks in Israel and Gaza." *The New York Times*, October. https://www.nytimes.com/interactive/2023/10/07/world/middleeast/israel-gaza-maps.html.

Meta. 2023. *Meta Marketing API*.

NetBlocks. n.d. "Network Connectivity by Region - Palestine. 2023-10-01 to 2023-11-15."

Palestinian Authority Ministry of Planning. n.d. "State of Palestine - Subnational Administrative Boundaries."

Palestinian Central Bureau of Statistics. n.d. "State of Palestine - Subnational Population Statistics."

Palotti, Joao, Natalia Adler, Alfredo Morales-Guzman, Jeffrey Villaveces, Vedran Sekara, Manuel Garcia Herranz, Musa Al-Asad, and Ingmar Weber. 2020. "Monitoring of the Venezuelan Exodus Through Facebook's Advertising Platform." Edited by Federico Botta. *PLOS ONE* 15 (2): e0229175. https://doi.org/10.1371/journal.pone.0229175.

Raleigh, Clionadh, rew Linke, Håvard Hegre, and Joakim Karlsen. 2010. "Introducing ACLED: An Armed Conflict Location and Event Dataset." *Journal of Peace Research* 47 (5): 651–60. https://doi.org/10.1177/0022343310378914.

Rampazzo, Francesco, Marzia Rango, and Ingmar Weber. 2023. "New Migration Data: Challenges and Opportunities." In, edited by Eleonora Bertoni, Matteo Fontana, Lorenzo Gabrielli, Serena Signorelli,

and Michele Vespe, 345–59. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-16624-2_18.

Ratnayake, Ruwan, Nada Abdelmagid, and Claire Dooley. 2022. "What We Do Know (and Could Know) about Estimating Population Sizes of Internally Displaced People." *Journal of Migration and Health* 6 (January): 100120. https://doi.org/10.1016/j.jmh.2022.100120.

"Timeline: Major Events During 100 Days of War Between Israel and Hamas." 2024. *The Straits Times*, January. https://www.straitstimes.com/world/middle-east/major-events-during-100-days-of-war-between-israel-and-hamas.

United Nations Office for the Coordinational of Humanitarian Affairs. 2024. "Hostilities in the Gaza Strip and Israel | Flash Update #163." http://www.ochaopt.org/content/hostilities-gaza-strip-and-israel-flash-update-163.

Zagheni, Emilio, Venkata Rama Kiran Garimella, Ingmar Weber, and Bogdan State. 2014. "Inferring International and Internal Migration Patterns from Twitter Data." In, 439–44.

Zagheni, Emilio, and Ingmar Weber. 2012. "You Are Where You e-Mail: Using e-Mail Data to Estimate International Migration Rates." In, 348351. WebSci '12. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/2380718.2380764.