More Data, Better Policies? Dealing with Migration Challenges of 2050

AMMAR A. MALIK

APRIL 2018 GLOBAL SHIFTS COLLOQUIUM

More Data, Better Policies? Dealing with Migration Challenges of 2050

Ammar A. Malik, Senior Research Associate in the Center for International Development and Governance at the Urban Institute

The Big Picture: Climate Change, Conflict, and Digital Device Proliferation

The global migration challenges of 2050 will emerge through shocks and stressors of climate change, growing inter- and intra-regional income disparities, and cross-border or civil conflict. On the other hand, unprecedented volumes of digital data are being generated every day from millions of devices. Critically, harnessing this data for improving policies and practices will require the following. First, developing a deeper understanding of the local political economies around data. Second, overcoming barriers to institutionalized cross-jurisdictional data sharing. Third, resolving serious challenges of data security and privacy. And fourth, forging effective partnerships with the private sector.

The World Bank (2018)[1] recently projected that climate change could triggered internal movement of 143 million people just in Sub-Saharan, South Asia and Latin America. When combined with rising inter- and intra-country income disparities, high fertility in poorer countries and potential armed conflicts, this prospect demands a paradigm shift in the global humanitarian response system based on deeper analytic insights. As more people around the world, particularly in developing countries, subscribe to mobile phones and gain access to the internet, amounts of digital data will skyrocket through trillions of Call Detail Records (CDRs), financial transactions and social media footprints. Latest global statistics and projections for 2017 and 2025 demonstrate the scale of this data avalanche: unique mobile subscribers from 5 to 5.9 billion, share of subscribers owning smartphones from 57 to 77%, mobile internet users from 3.3 to 5 billion, and internet of things devices 7.5 to 25.1 billion (GSMA 2018)[2].

With each device leaving digital footprints, the immense potential for data analytics has rightly created excitement among humanitarians and development practitioners alike. But **technical conversations seldom feature basic political economy questions** regarding this data: Why and in what form would entities share it? Who owns it, subscribers or providers? How could it be utilized, both commercially and for development outcomes? Regardless, researchers studying international development and humanitarian responses ought to agree that new data sources offer clear benefits over traditional alternatives.

Limitatons of Traditional Datasets: Censuses, Government Surveys, and Registrations

In national statistical agency activities, both censuses and sampled surveys, **foreign and other hidden populations are often excluded** for logistical, legal or political reasons. When fearing persecution, refugees or asylum seekers would make every effort to avoid census or survey teams, which in turn have little incentive to pursue them. For example, both in the 1998 and 2017 censuses, Karachi's population counts were widely contested as they purportedly did not include Bihari, Afghan and/or some Pakistani Pashtuns due to their transient stay in the city. But if host governments have incentives to systematically record this information, e.g. part of conditional aid package, they are likely to respond in earnest. However, since many refugee origin and hosting countries remain in turmoil, administrative incapacities and security challenges prevent them from

---

[1] Source: https://openknowledge.worldbank.org/handle/10986/29461

[2] https://www.gsmaintelligence.com/research/2018/02/the-mobile-economy-2018/660/

More Data, Better Policies? Dealing with Migration Challenges of 2050

Ammar A. Malik, Senior Research Associate in the Center for International Development and Governance at the Urban Institute

increase in national population, much higher than anticipated. Nationally representative surveys, which are held more regularly than censuses, although based on disaggregated sampling frames, simply do not provide enough statistical power to estimate impacts at the local (district or county) level. Even if district level indicators could be distilled, governments in national governments are concerned that releasing subnational GDP will reignite erstwhile debates on dealing with lagging and leading regions.

Similarly, refugee **registration data systems are often unlinked to national and international identity or other databases** owned by governments. To the best of my knowledge, there are no internationally agreed protocols for data ownership and incentives for various parties to collect, report on and share data are disparate. For example, in 2005, more than two decades after Soviet invasion of Afghanistan, which triggered the then largest force displacement crisis in a generation, UNHCR and government of Pakistan decided to undertake a 'census' of Afghan refugees living all over the country. In the context of a tripartite agreement between UNHCR and the governments of Pakistan and Afghanistan, resulting information was intended to guide returns which Pakistan was keen on organizing soon. In recent fieldwork, we found many undocumented Afghans who wished to stay in-country and hence avoided census officials, or were integrated in Pakistan's national identity system through marriage, or other means. While censuses have been conducted in a few other places, e.g. Rwanda (2005) and Dominican Republic (2012), and provide valuable information on all aspects of refugees' lives, most typical registration datasets are insufficient to understand their social and economic lives, needs, or aspirations.

Unsurprisingly in this data environment, **researchers continue relying on household surveys**, focus groups or semi-structured interviews to get deeper insights into refugees' social networks, livelihoods, self-reliance potential, aspirations, and so on. During the last two years, the Urban Institute has surveyed thousands of refugees in Kenya, Pakistan, Turkey and the United States through phone, mail, in-person using paper forms and in-person using tablets. Such surveys are costly (e.g., $30 per 40-minute survey in Gaziantep), labor intensive and time consuming – requiring careful questionnaire design, field staff training, pilot testing and obtaining official permissions. Given refugees' disparate location patterns and absence of refugee focused baselines, government sampling frames are unusable, making it impossible to draw representative samples and significantly increasing survey costs as field staff must make many unsuccessful contacts before finding refugees. In places where refugees face risk of deportation or persecution, surveys can be outright dangerous, both for field staff (being spies) and respondents (suspecting leakage of information). Researchers must spend considerable resources in ensuring that risks associated with these surveys do not outweigh potential benefits to refugees by working with Institutional Review Boards and implementing strict protocols for data security.

Promising Applications of New Data Sources: Understanding Social Networks, Supporting Financial Inclusion, and Measuring Economic Impacts on Host Communities

But as shown in recent studies, new digital datasets (CDRs, airtime transfers) offer novel opportunities for overcoming these barriers

More Data, Better Policies? Dealing with Migration Challenges of 2050

Ammar A. Malik, Senior Research Associate in the Center for International Development and Governance at the Urban Institute

and offering robust evidence to improve salient migration policies. They are comprehensive (millions of subscribers, billions of records), high frequency (multiple records per day) and mostly generated without physical contact with users – making them safe, reliable, more rapidly usable and low cost. Three promising recent applications are cases in point.

First, enabling greater refugee self-reliance to reduce burdens on formal assistance is arguably the most sustainable long-term solution to the migration crisis. But understanding the forms and evolution of refugee social networks through surveys is expensive and methodologically challenging. Using **50 billion CDRs covering all text messages, phone calls and peer-to-peer airtime transfers made in Rwanda** between 2005 and 2009, Blumenstock et al. (2016) found robust evidence that following natural disasters, wealthy individuals are more likely to receive airtime transfers, and there are greater transfers between "pairs of individuals with strong history of reciprocal exchange." By mapping refugee social network structures, their evolution following disasters and specific ways they contribute to self-reliance, by 2050 humanitarian agencies and governments could much improve targeting of assistance programs.

Second, vast literature has shown that the poor's financial situation is so precarious that even a small exogenous shock, e.g. as short-term disability or poor seasonal crop yield, can have serious long-term repercussions. Recent work by Jack and Suri (2014) on Kenya's poor including refugees shows that in the wake of such shocks, **non-users of mobile money technology suffer a 7 percent drop in consumption, whereas users are unaffected** and have greater likelihood of receiving remittances. With over 300 active cash transfer programs benefiting thousands each year, linking refugee IDs with mobile money accounts could open fresh possibilities for consumption smoothing microinsurance

products. Through ID matching, companies could access information on refugees' financial histories, income levels, socioeconomic backgrounds etc. from other datasets to vastly improve actuarial models (Kumar and Muhota 2012). Just as affordable international mobile roaming has become a reality in recent years, by 2050 it is plausible that we will witness universally accepted financial IDs for all humanity, linked global credit histories, and no-cost international money transfers regardless of physical distance. This could open unparalleled economic opportunities for the world's poor, or the 1.7 billion mobile subscribers without a bank account (GSMA 2018).

Third, host-refugee community relations are hurt by arguments that **migration 'kills local jobs' and imposes economic burdens** on host communities' already weak public service delivery institutions. Refugee hosting governments such as Jordan and Pakistan wonder what sectors and how could low-cost refugee workforce help improve local firm profitability – but appropriately detailed labor market and firm productivity data is unavailable in most places. After Turkey's statistical authority included 1.6 million refugees in at least two rounds of labor force surveys however, Del Caprio et al. (2015) found clear evidence of refugee-induced job displacements in certain sectors and labor categories. Relatively low skilled Turkish workers, who were previously employed in low wage informal economy jobs, moved to better paying formal jobs as they were replaced by Syrian refugees. But no such impacts were identifiable for Turks with highly specialized skillsets, or women in general. Even if most refugee hosting governments do not collect or share labor force microdata on refugees, this research has shown that greater adaption of digital payment platforms including mobile money could help evaluate refugees' net economic contributions, including through detailed financial diaries of incomes, expenditures, and savings.

More Data, Better Policies? Dealing with Migration Challenges of 2050

Ammar A. Malik, Senior Research Associate in the Center for International Development and Governance at the Urban Institute

Challenges to Scaling Up New Data Analytics: Local Politics, Global Standardization, Data Security, and Cross-Sector Partnerships

While these examples demonstrate the exciting prospects for new waves of research using open data and machine learning, researchers must consider a series of factors that will likely limit their potential. First, even though the raw supply of data is evidently exploding, **local political realities** determine how created, shared, analyzed and disseminated. Kingsley (2017) argues that even industrialized country governments are notoriously inadept at data sharing, even on basic public services such as water and sanitation, but not without reasons. In refugee hosting countries for instance, sharing performance data on public services could: expose poor performance of bureaucrats, create backlash from host communities pressurizing elected representatives to justify resource allocations away from citizens, and erode refugees' trust in humanitarian agencies in case private data is leaked. In some cases, leakage of information regarding refugee locations and their 'illegal' or 'informal' economic activities could seriously jeopardize their security, even resulting in loss of life. To analyze city data environments, Edwards et al. (2016) suggest considering "permissions, incentives and institutionalization" to understand whether and why various stakeholders would cooperate for open data based transparent governance. In migration policy, sorting data ownership issues, aligning stakeholder incentives and institutionalizing processes is particularly problematic due to the plethora of political complications introduced above.

Second, while the March 2018 draft of the Global Compact for Migration calls for "standardization" and "harmonization" of migration focused data systems, it is unclear why member countries might agree to do so? While the collective benefits to the humanitarian systems are obvious, it appears to be a classic case of **tragedy of the commons**, that too at a time when UNHCR's funding gap crossed $ 3.8 billion. Which entity within the international system has the financial resources, intellectual capacity and universally recognized integrity to design and implement global refugee data collection and dissemination protocols? The Urban Institute's National Neighborhoods Indicator Project, which pools resources through a network of local data gatekeepers, offers a potentially replicable model. But at the international level, how can countries be persuaded to opt-in to a universally implemented biometrics-based digital ID programs covering every individual on earth? With the click of a button, such a system could allow humanitarian agencies, employers, social service providers and people themselves to access and verify key credentials such as social security identification, passport/immigration records, educational diplomas, work experience certificates and so on. This could greatly improve the prospects of the forcibly displaced to make fresh starts in new places, even if original documentation is lost to conflict.

Third, when the world's most sophisticated and resourceful technology companies are unable to protect customers' information from social media profiles or online shopping portals, **it is unclear how humanitarian agencies could protect personal information**. But as the recent Cambridge Analytica scandal shows, even without leaks to hackers there are serious concerns regarding invasion of privacy as digital devices touch every

More Data, Better Policies? Dealing with Migration Challenges of 2050

Ammar A. Malik, Senior Research Associate in the Center for International Development and Governance at the Urban Institute

aspect of our daily lives. In the case of vulnerable populations like refugees, this raises even more serious ethical questions: Do refugees own digital data created from participation in cash transfer programs? Are they in a position to give consent to the use of this data? For what purposes and for how long should access be authorized? Under what laws and regulations would telecommunication companies, governments, researchers and nonprofits access it? As most refugees reside in developing countries, where data privacy laws even for citizens are underdeveloped, perhaps the global compact for migration framework could support technical assistance to support governments in this regard?

Fourth, since most digital data is being generated on devices and software created by private companies, they are fast emerging as a key player in the humanitarian system. Many are employing innovative ways to harness creative energies for utilizing data, such as TurkCell's innovation competition, Data for Refugees, soliciting proposals for creating public value from cellphone data, with the reward being data access and financial resources for researcher's time. While the private sector has worked with humanitarian agencies for decades, most of this engagement has been either through philanthropic ventures where humanitarians receive in-kind or cash grants, or through procurement channels where private companies become contracted service providers. Hence there is a need to fundamentally reimagine **ways in which the private and humanitarian sectors could work together through mutually beneficial arrangements**. Initial findings from ongoing Urban Institute research suggests that synergies could be created particularly in the technology sector, but successful partnerships require significant upfront investments in building organizational trust through mutual due diligence, identifying

specific forms of collaboration through carefully negotiated contractual clauses.

References

Blumenstock, J. E., Eagle, N., & Fafchamps, M. (2016). Airtime transfers and mobile communications: Evidence in the aftermath of natural disasters. *Journal of Development Economics,* 120(C), 157–181.

Carpio, D., Vanessa, X., & Wagner, M. C. (2015). *The impact of Syrians refugees on the Turkish labor market* (No. WPS7402) (pp. 1–47). The World Bank. Retrieved from http://documents.worldbank.org/curated/en/505471468194980180/The-impact-of-Syrians-refugees-on-the-Turkish-labor-market

Edwards, B., Greene, S., & Kingsley, G. T. (2016). A Political Economy Framework for the Urban Data Revolution (Research Report). Washington D.C.: Urban Institute. Retrieved from https://www.urban.org/research/publication/political-economy-framework-urban-data-revolution

GSMA. (2018). Mobile Economy 2018 [Data]. Retrieved April 9, 2018, from https://www.gsma.com/mobileeconomy/

Jack, W., & Suri, T. (2014). Risk Sharing and Transactions Costs: Evidence from Kenya's Mobile Money Revolution. *American Economic Review,* 104(1), 183–223. https://doi.org/10.1257/aer.104.1.183

Kabir Kumar, & Kim Muhota. (2012). *Can Digital Footprints Lead to Greater Financial Inclusion?* (Brief). Washington D.C.: CGAP. Retrieved from http://www.cgap.org/sites/default/files/CGAP-Brief-Can-Digital-Footprints-Lead-to-Greater-Financial-Inclusion-Jul-2012.pdf

More Data, Better Policies? Dealing with Migration Challenges of 2050

Ammar A. Malik, Senior Research Associate in the Center for International Development and Governance at the Urban Institute

Kingsley, G. T. (2017). *A Broader View of the Data Revolution and Development Agenda* (Research Report). Washington D.C.: Urban Institute. Retrieved from https://www.urban. org/research/publication/broader-view-data-revolution-and-development-agenda