

**Authors**

Joshua Blumenstock  
University of California, Berkeley

Dean Karlan  
Northwestern University

Christopher Udry  
Northwestern University

**Article**

**Machine learning and phone data can improve targeting of humanitarian aid**

<https://doi.org/10.1038/s41586-022-04484-9> Emily Alkon<sup>1</sup>, Suzanne Bekker<sup>2</sup>, Dean Karlan<sup>3</sup>, Christopher<sup>4</sup> & Joshua E. Blumenstock<sup>1,5\*</sup>

Received: 15 July 2021

Accepted: 25 January 2022

Published online: 18 March 2022

Open access

Check for updates

The COVID-19 pandemic has devastated many low- and middle-income countries, causing widespread food insecurity and a sharp decline in living standards<sup>1</sup>. In response to this crisis, governments and humanitarian organizations worldwide have distributed social assistance to more than 1.5 billion people<sup>2</sup>. Targeting is a central challenge in administering these programmes: it remains a difficult task to rapidly identify those with the greatest need given available data<sup>3</sup>. Here we show that data from mobile phone networks can improve the targeting of humanitarian assistance. Our approach uses traditional survey data to train machine-learning algorithms to recognize patterns of poverty in mobile phone data; the trained algorithms can then predict the aid of the poorest mobile subscribers. We evaluate this approach by studying a flagship emergency cash transfer program in Togo, which used these algorithms to distribute millions of U.S. dollars worth of COVID-19 relief aid. Our analysis compares outcomes—the targeting exclusion errors, total social welfare and measures of fairness—under different targeting regimes. Relative to the geographic targeting options considered by the Government of Togo, the machine-learning approach reduces errors of exclusion by 4–28%. Relative to methods requiring a comprehensive social registry (a hypothetical exercise; no such registry exists in Togo), the machine-learning approach increases exclusion errors by 9–35%. These results highlight the potential for new data sources to complement traditional methods for targeting humanitarian assistance, particularly in crisis settings in which traditional data are missing or out of date.

The COVID-19 pandemic has led to a sharp decline in living standards across the world, as public health officials stop the spread of the disease we have disrupted normal economic activity. Economically vulnerable households in low- and middle-income countries have been severely impacted, with more than 100 million individuals estimated to have transitioned into extreme poverty since the onset of the pandemic<sup>1</sup>. To offset the most severe consequences of this sudden decline in income, governments and humanitarian organizations around the world have mobilized relief efforts. It has been estimated that more than 1,300 new social assistance programmes have been launched<sup>2</sup> since early 2020, providing more than 1.5 billion people in cash transfer payments to over 1.1 billion people (roughly one-fifth of the world's population). The overwhelming majority of COVID-19 response efforts—and the majority of cash transfer programmes globally—provide targeted social assistance<sup>3</sup>. In other words, specific criteria—typically a proxy for socioeconomic status—are used to determine potential eligibility. In more wealthy nations, governments commonly use tax or household income data to determine programme eligibility<sup>4</sup>. However, in low- and lower-middle-income countries (LICs), where economic activity is often informal and based on home-produced agriculture, governments typically do not collect income for the vast majority of the population<sup>5</sup>. Other potential sources of targeting data are often incomplete or out of date<sup>6</sup>. For example, only half of the poorest countries have completed a census

in the past 10 years<sup>7</sup>. In such contexts, data gaps provide governments with an opportunity to improve social assistance programmes<sup>8,9</sup>.

Here we develop, implement and evaluate an approach to targeting social assistance based on machine-learning algorithms and non-traditional big data from satellite and mobile phone networks. This approach leverages recent advances in machine learning that show that such data can help accurately estimate the wealth of small geographic regions<sup>10–12</sup> and individual mobile subscribers<sup>13–15</sup>. It also builds on a rich economic literature on the design of appropriate mechanisms for targeting social assistance<sup>16–18</sup>. See Supplementary Discussion section 1 for a summary of previous work.

**Humanitarian response to COVID-19 in Togo**

Our results are based on the design and evaluation of *Novitas*, a flagship emergency social assistance programme carried out in Togo. The Government of Togo launched *Novitas* in July 2020, shortly after the first cases of COVID-19 appeared in the country. An economic lockdown had reduced many Togolese to stop working and led to widespread food insecurity (Supplementary Fig. 1). *Novitas* used to provide substance cash relief to those most affected (see <https://novitas.gov.tg>). Eligible beneficiaries received one weekly payments of roughly US\$6. In an effort to minimize its personnel costs, *Novitas* used text and payments were implemented

\*School of Information, University of California, Berkeley, CA, USA. <sup>1</sup>Department of Economics, University of Berkeley, Berkeley, California. <sup>2</sup>Technology School of Management, Global Poverty Research Center, Northwestern University, Evanston, IL, USA. <sup>3</sup>Department of Economics, Global Poverty Research Center, Northwestern University, Evanston, IL, USA. <sup>4</sup>These authors equally contributed to this work. <sup>5</sup>World Bank, Washington, DC, USA. <sup>6</sup>World Bank, Washington, DC, USA. <sup>7</sup>World Bank, Washington, DC, USA. <sup>8</sup>World Bank, Washington, DC, USA. <sup>9</sup>World Bank, Washington, DC, USA. <sup>10</sup>World Bank, Washington, DC, USA. <sup>11</sup>World Bank, Washington, DC, USA. <sup>12</sup>World Bank, Washington, DC, USA. <sup>13</sup>World Bank, Washington, DC, USA. <sup>14</sup>World Bank, Washington, DC, USA. <sup>15</sup>World Bank, Washington, DC, USA. <sup>16</sup>World Bank, Washington, DC, USA. <sup>17</sup>World Bank, Washington, DC, USA. <sup>18</sup>World Bank, Washington, DC, USA.

# Machine Learning and Phone Data Can Improve Targeting of Humanitarian Aid

The COVID-19 pandemic has devastated many low- and middle-income countries, causing widespread food insecurity and a sharp decline in living standards. In response to this crisis, governments and humanitarian organizations worldwide have distributed social assistance to more than 1.5 billion people. Targeting is a central challenge in administering these programmes: it remains a difficult task to rapidly identify those with the greatest need given available data. Here we show that data from mobile phone networks can improve the targeting of humanitarian assistance. Our approach uses traditional survey data to train

machine-learning algorithms to recognize patterns of poverty in mobile phone data; the trained algorithms can then prioritize aid to the poorest mobile subscribers. We evaluate this approach by studying a flagship emergency cash transfer program in Togo, which used these algorithms to disburse millions of US dollars worth of COVID-19 relief aid. Our analysis compares outcomes—including exclusion errors, total social welfare, and measures of fairness—under different targeting regimes. Relative to the geographic targeting options considered by the Government of Togo, the machine-learning approach reduces errors of exclusion by 4–21%. Relative to methods requiring a comprehensive social registry (a hypothetical exercise; no such registry exists in Togo), the machine-learning approach increases exclusion errors by 9–35%. These results highlight the potential for new data sources to complement traditional methods for targeting humanitarian assistance, particularly in crisis settings in which traditional data are missing or out of date.

March 21, 2022