Symptoms and Stereotypes: Perceptions and Responses to Covid–19 in Malawi and Zambia

Karen E. Ferree[*], Boniface Dulani[†], Adam S. Harris[‡], Kristen Kao[§], Ellen Lust[¶], Cecilia Ahsan Jansson[‖] and Erica Ann Metheney[**][††]

August 22, 2021

[*]Associate Professor, University of California, San Diego.
[†]Associate Professor, University of Malawi.
[‡]Lecturer (Assistant Professor), University College London.
[§]Senior Research Fellow, University of Gothenburg.
[¶]Professor, University of Gothenburg.
[‖]PhD Student, University of Gothenburg.
[**]Statistician, University of Gothenburg.
[††]This research was supported by Swedish Research Council Recruitment Grant (Swedish Research Council – E0003801), PI: Pam Fredman; Social Institutions and Governance: Lessons from Sub-Saharan Africa (Swedish Research Council – 2016-01687); The Governance Challenge of Urbanization (FORMAS – 2016-00228), PI: Ellen Lust. We gratefully acknowledge helpful advice from Kim Yi Dionne, Adia Benton, Lauren Honig, Claire Adida, and attendees of the Spring 2021 Governance and Local Development Seminar Series.

Abstract

A growing literature documents Covid–19's health and economic effects. Can Covid–19 also exacerbate identity divisions? Psychologists argue that contagious disease increases threat perceptions and provokes policing of group boundaries and discrimination against perceived outsiders. We focus on the emergence of disease-based stereotypes as a mechanism underlying this work. We explore how insider/outsider status and symptoms of illness shape perceptions of infection, reported willingness to help, and desire to restrict free movement of an ailing neighbor using a phone-based survey experiment administered three times in two neighboring African countries during different stages of the pandemic: Malawi, from May 5 to June 2, 2020 (n=4,641); Zambia, from July 2 to August 13, 2020 (n=2,198); and Malawi again, from March 9 to May 1, 2021 (n=4,356). We find that symptoms more strongly shape perceptions and projected behavior than insider/outsider status in both countries and across time, suggesting that objective risk matters more than identities in shaping responses to the illness.

1 Introduction

A large and growing literature documents Covid–19's catastrophic health and economic effects. In this paper, we focus on its political effects, in particular the potential for Covid–19 to exacerbate identity group divisions. Research on epidemics by psychologists (Schaller, 2011) and political scientists (Aarøe, Petersen, & Arceneaux, 2017; Adida, Dionne, & Platas, 2018; Albertson & Gadarian, 2015; Arriola & Grossman, 2020; Casey, 2015; Dionne & Turkmen, 2020; Kam, 2019; Lieberman, 2007; Onoma, 2020) suggests that Covid-19 could intensify group conflict and lead to greater discrimination against already marginalized groups. We consider a mechanism underlying this work: the emergence of disease-based stereotypes around insider-outsider identities. Stereotypes precede and enable discriminatory behaviors and the stigmatizing of groups as bearers of disease. We pay particular attention to how insider/outsider identities interact with symptoms to shape perceptions about who has disease and behavioral responses to perceived disease threat. Earlier research suggests that disease threat may not reflect epidemiological data and actual risk of infection, with individuals dismissing or ignoring symptoms of insiders while associating disease with outsiders even when symptoms are absent (Dionne & Turkmen, 2020; Lieberman, 2009; Onoma, 2020). The relationship between identity and symptoms seems particularly important in the case of Covid-19, a disease often transmitted through asymptomatic individuals. A core contribution of our work is therefore to explore identity, symptoms, and the interaction between them in perceptions about and reaction to disease.

We examine the emergence of stereotypes through a phone-based survey experiment administered three times in two neighboring African countries during the pandemic in 2020 and 2021: Malawi, from May-June, 2020 (n=4,641); Zambia, from July-August, 2020 (n=2,198); and Malawi again, from March-May, 2021 (n=4,356). Our surveys occurred at different stages of the pandemic, ranging from very early, when cases were low, to advanced, when the disease had spread extensively, allowing us to examine how time shaped responses. In an experimental vignette about an ailing neighbor seeking help getting to the hospital, we manipulate three aspects of identity: insider/outsider status, gender and age. We consider outsider identities based on newer (versus more established) residency in the neighborhood; and those associated with foreign or racial minorities. We also manipulate the nature of the neighbor's ailment, with symptoms ranging from highly suggestive of infectious disease and Covid-19 (cough and high fever) to symptoms not suggestive of infectious disease and Covid-19 (a badly injured and infected leg). We evaluate the main effects of identity and symptoms as well as their interaction and explore three outcomes: 1) whether the respondent believed the neighbor had Covid–19; 2) whether the respondent would be willing to accompany the neighbor to the hospital; and 3) whether the respondent believed the neighbor should be restricted from moving freely around the community. The first captures associations of identity characteristics with disease and the early emergence of stereotypes; the last two consider how identity characteristics shape hypothetical behaviors.

As anticipated by disease threat theories, we find that identities shape perceptions about and behavioral responses to Covid-19, particularly in Malawi. These effects were surprisingly small in size, however, and remained so even as the pandemic grew in magnitude. Symptoms, on the other hand, strongly and consistently shaped reactions to disease in the hypothetical vignettes in both countries and across time periods. We find only limited evidence that stereotypes and symptoms interact: in the first Malawian survey, identity may have a slightly smaller effect on perceptions of disease when symptoms point unambiguously to Covid-19, but this interaction effect does not persist in the second Malawi survey, nor does it show up in Zambia or for any of our behavioral indicators. We therefore conclude that identity has a small but persistent main effect on disease related perceptions and behaviors, one that is dwarfed by the effect of symptoms.

Our results thus provide a counterpoint to previous work arguing that actual disease risk has less influence over perceptions and behavior than identity.

2 Theorizing Disease Stereotypes: Identity and Symptoms

Humans have long associated "others" with disease, blaming contagions on community outsiders, justifying exclusionary policies and even violence on the basis of real or imagined disease threats (Schaller, 2011). Communities attributed outbreaks of plague in Medieval Europe to ethnic or religious outsiders (Nohl, 2006). Ideas about disease, contagion, and purity accompanied the Holocaust (Suedfeld & Schaller, 2002) and informed longstanding practices of excluding immigrant groups in the United States (Adida et al., 2018). During the 2014 West African Ebola outbreak, Senegalese targeted the migrant Peul population because of their roots in neighboring Guinea (Onoma, 2020), while people in Western countries associated the disease with Africa itself (Benton & Dionne, 2015). Likewise, Asian-Americans suffered acts of discrimination and racism because of their purported association with Covid-19's origins, a connection encouraged by politicians like President Donald Trump (Dionne & Turkmen, 2020; Reny & Barreto, 2020).

Psychological theories explain the association of outsiders with contagion as a function of the "behavioral immune system," a set of "psychological processes that infer infection risk from perceptual cues, and that respond to these perceptual cues through the activation of aversive emotions, cognitions and behavioural impulses," (Schaller, 2011: 3418). Such cues include obesity, advanced age, physical disabilities, and, most relevant here, being perceived as foreign or an outsider. Schaller (2011) offers an evolutionary mechanism for the association of foreignness with disease: when humans lived in small, isolated communities, "exotic" visitors may have been a source of new and virulent pathogens. Such visitors may also have violated local cultural norms around hygiene and food preparation believed to reduce disease transmission, enhancing perceptions of threat. Schaller (2011) argues that this association between outsiders and disease became hardwired into human psychology as an unconscious instinct that continues to shape behaviors today, even as the original context has disappeared.

Empirical tests of behavioral immune system theories explore the effects of disease threats on outcomes like prejudice and xenophobia (Faulkner, Schaller, Park, & Duncan, 2004; Navarrete & Fessler, 2006; Navarrete, Fessler, & Eng, 2007; Schaller, 2011). Political scientists have built upon these foundations to explore the effect of disease threat on political outcomes like support for immigration restrictions, limits on civil liberties, and foreign policy spending, finding mixed support (Aarøe et al., 2017; Adida et al., 2018; Albertson & Gadarian, 2015; Casey, 2015; Kam, 2019).

A growing strand of this work points to the potential *conditionality* of the link between disease threat and attitudes towards outsiders. Onoma (2020), in his work on the Peul in Senegal, emphasizes the role of proximity in conditioning beliefs about who has the disease: while ethnic Peul in general were believed to be disease carriers, specific Peul who were well known neighbors and community members escaped this association. Dionne and Turkmen (2020) point instead to prior patterns of discrimination, suggesting that othering based on disease threat should focus on groups who are already marginalized, particularly racial and religious minorities.

We consider another potential conditioning effect: disease symptoms. Health threats do not always reflect objective risks (Faulkner et al., 2004). The behavioral immune system may "not only lead to the social stigmatization of people who truly are infectious but also equally pernicious prejudices directed against people who are not" (Schaller, 2011: 3420). Dutta and Rao (2015) find that disease threat can provoke xenophobic responses against outsider groups, *even when these groups pose no increased risk of contagion.* In his study of Senegal, Onoma (2020: 369) notes "the rather marginal influence (if any) that the epidemiology of EVD [Ebola Virus Disease] had on which people were targeted by xenophobic acts," a conclusion echoed by Dionne and Turkmen (2020) about other pandemics. These effects may reflect the ambiguity of symptoms for some diseases and the use of identity characteristics as an informational shortcut to resolve it, or a form of motivated reasoning wherein individuals discount disease risk for members of their own group and overstate disease risk for outsiders (Cohen, 1999; Lieberman, 2009).

Previous work on symptoms and identity has largely been historical or ethnographic. In the research that follows, we employ a survey experimental approach in Malawi and Zambia during different stages of the Covid-19 pandemic. Survey experiments allow us to independently manipulate symptoms and identity to gain greater insight into the independent and interactive effects of each, holding constant additional considerations like age and gender.

We explore four pre-registered hypotheses about identities, symptoms, and their interaction on perceptions about and responses to an ailing neighbor.[1]

Hypothesis 1. *Identity shapes perceptions and responses: Malawians and Zambians will be (a) more likely to believe a neighbor has Covid-19 when they are members of outsider rather than insider groups; (b) less likely to help a neighbor get to the hospital if that neighbor is described as an outsider rather than an insider; (c) less likely to support free movement for an ailing neighbor described as an outsider rather than an insider.*

We also consider a second identity hypothesis consonant with Dionne and Turkmen (2020)'s argument that identity effects are strongest for marginalized or racialized groups:

[1] For the Zambian pre-analysis plan, see: `https://osf.io/azpvk/` and for the Malawian pre-analysis plan, see `https://osf.io/wxtyd/`.

Hypothesis 2. *Identity effects will be stronger for previously marginalized and/or racially different groups versus less marginalized and/or racially similar groups. This will be true across outcomes (a) perceptions; (b) willingness to help an ailing neighbor; (c) support free movement of an ailing neighbor.*

Hypothesis 3. *Symptoms shape perceptions and responses: Malawians and Zambians will be (a) more likely to believe a neighbor has Covid-19 when they are described as having symptoms of fever and/or cough; (b) less likely to help a neighbor get to the hospital if they are described as having a fever and/or cough; (c) less likely to support free movement for a neighbor described as having a fever and/or cough.*

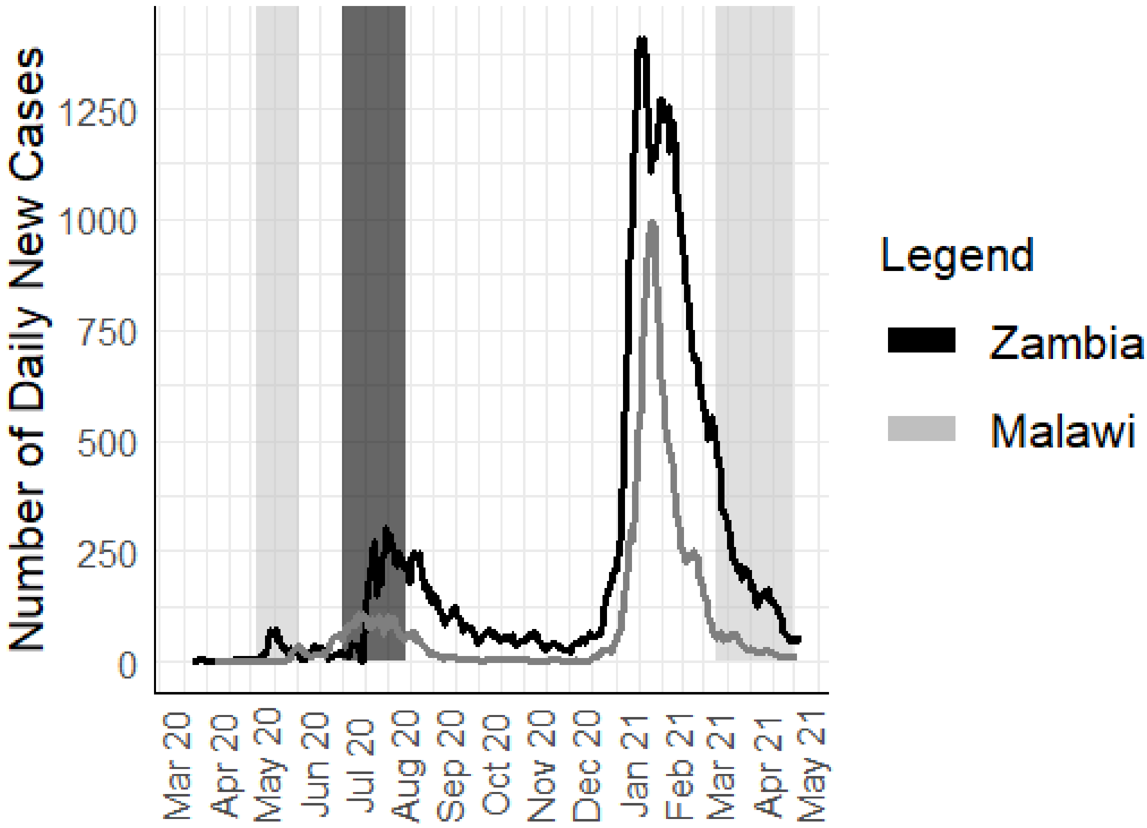
Hypothesis 4. *Identity and symptoms interactively shape perceptions and responses. When symptoms are clear, insider/outsider distinctions will cease to have an effect on (a) perceptions of disease; (b) willingness to help an ailing neighbor; (c) support for free movement of an ailing neighbor.*

In the next section we provide background on our two cases to situate our experiment before turning to the experimental design.

3 Covid–19 in Malawi and Zambia

Zambia reported its first cases of Covid–19 in March 2020, followed by Malawi a month later. Zambia experienced an early spike in May, reaching over 1000 cases by the end of that month, and a second, larger wave in late July and August. Malawi's first wave hit in July and early August. Both countries then experienced substantial waves in January and February of 2021 (see Figure 1).

Figure 1: Reported New Cases of Covid–19 in Malawi and Zambia Over Time



Note: The figure shows the 7-day rolling average of new cases. The shaded regions indicate when our surveys were in the field. Malawi and Zambia each have between 17-18 million inhabitants. Original data obtained from *Our World in Data*, Hasell (2020).

Even before the first confirmed cases, the Zambian government responded to the growing global pandemic by implementing nationwide restrictions in an attempt to curb the spread. They activated the National Public Health Emergency Operations Centre, implemented active surveillance of all international passengers, procured supplies of personal protective equipment, launched a series of community engagement campaigns to increase awareness about the disease, and closed all educational institutions. Following the confirmation of the first cases, they added new measures to regulate public interactions, implementing a partial lockdown of the country on March 13, 2020. As more information became available, the government restricted non-essential travel, instructed the use of facemasks in public, and issued a general stay-at-home advisory with police officers deployed to ensure compliance.

The Malawian government responded to the pandemic in late March by declaring a State of National Disaster, which banned public gatherings, closed schools, and introduced a hand-washing and social distancing campaign. The government further established a Ministerial Committee to spearhead efforts to contain the pandemic at the national level. Following confirmation of the first cases, the government added new measures to regulate public interactions. These included the closure of all national borders and airports, reduction in the number of passengers on public transportation, and further limiting the size of public and social gatherings.

As the number of cases continued to increase in Malawi, the government announced plans to impose a 21–day nationwide lockdown from April 18, 2020. However, this announcement was met with sporadic demonstrations and protests across the country, mostly by small businesses operating in the country's many markets. The protestors complained that a lockdown without government support for small businesses and poor households would create economic hardships for millions of Malawians and could result in mass starvation. A coalition of human rights civil society organizations, operating under the banner of the Human Rights Defenders Coalition (HRDC), obtained a court injunction against the proposed lockdown by arguing that the proposal was unconstitutional. The lockdown was accordingly suspended.

The first wave of COVID-19 in Malawi peaked by mid August 2020, with the total number of cases reaching 2,273. A new government that came to power in July 2020 introduced a raft of fresh COVID-19 preventative measures that were similar to those that came into force at the beginning of the pandemic. In the ensuing months, the number of active cases gradually declined, reaching a lowest point of 31 active cases by mid December 2020. Towards the second half of December 2020, however, Malawi began to experience a second wave of COVID-19 infections and fatalities with the more contagious Beta variant as the main strain. The number of active cases reached 14,631 by early February 2021. In March 2021, Malawi received a first batch of Covishield vaccines under the COVAX scheme and immediately launched a vaccination campaign that initially targeted health personnel and other essential service providers.[2]

Our surveys, which were fielded from May 5 to June 2 2020 in Malawi, July 2 to August 13 2020 in Zambia, and March 9 to May 1 2021 in Malawi, thus coincided with very different periods in the outbreak of the pandemic in each country. The First GLD–IPOR Malawi Covid–19 survey – indicated by the first vertical gray bar in Figure 1 – was implemented during the early period of the pandemic. At this time, the region witnessed cases of Covid–19, but the incidence of the disease itself was low. The First GLD–SAIPAR Zambia Covid–19 survey – the blue vertical bar in Figure 1 – took place at a later stage of the pandemic. Cases were spiking, and information about the disease was more widespread. The Second GLD–IPOR Malawi Covid–19 survey – indicated by the second gray vertical bar – occurred just after the large January-February wave in both countries. Covid-19 was truly "novel" during our first surveys in Malawi in March

[2]For more on the early Malawian response, see Dionne, Dulani, and Fischer (2021) and Tengatenga, Tengatenga Duley, and Tengatenga (2021).

2020. A year later this was no longer the case. Staggering surveys across twelve months and different pandemic stages allows us to explore the role of time and learning, two relatively under-studied factors in current research on Covid-19.

4 Empirical Strategy

We employ single profile conjoint experiments (Hainmueller, Hopkins, & Yamamoto, 2013) embedded in phone surveys on knowledge and fears about and behavioral responses to Covid–19.

4.1 Experimental Design

The experiments presented each respondent with a hypothetical profile about a neighbor, which randomly assigned the neighbor's age, gender, length of residency in the community, racial and/or national identity, and symptoms of illness. All attributes were concrete, easy to understand, and conveyed in a single sentence. Table 1 presents the attributes and their levels. After listening to the scenario read over the phone, we asked each respondent how likely he or she would be to help the neighbor to the hospital and whether the neighbor's movement should be restricted. To end the experiment, we asked respondents whether they thought the neighbor had Covid–19. Answers to these questions, which report on projected behavioral responses to and perceptions about a hypothetical scenario, constitute our dependent variables of interest. Below Table 1 is the full text of the experimental script; the post-treatment questions can be found in Table 2.

Table 1: Conjoint Experiment Attributes and Levels

Attributes	Levels (Malawi)	Levels (Zambia)
Age	25	25
	60	60
Gender	Man	Man
	Woman	Woman
Time in Community	Many years	Many Years
	A few months	A few months
Identity	Malawian	Zambian
	Mmwenye	Tanzanian
	Zambian	Malawian
Symptoms of Illness	Badly injured and infected leg	Badly injured and infected leg
	Bad cough and high fever	Bad cough and high fever
	High fever	High fever

The Malawian experimental script read as follows, with the randomized components in bold and in brackets:

> Now I would like you to imagine the following situation. Your neighbor, a [**25 year old/60 year old**] [**Malawian/Mmwenye/Zambian**] [**man/woman**] who has lived in your community for [**many years/a few months**], has [**a badly injured and infected leg; a bad cough and high fever; a high fever**].

A nearly identical script was read in Zambia, the only difference being the three randomized identity groups. This script was followed by three questions presented in the order given in Table 2 [3].

[3] RTA = Refuse to Answer

Table 2: Post-Treatment Questions

Question	Responses	
	Malawi	Zambia
If this person needed you to accompany him/her to the hospital, would you help him/her?	Yes No Don't Know/RTA	Yes No Don't Know/RTA
Do you think this person should be allowed to move freely about the community or should be made to stay inside at home?	Move freely Stay at home Don't know/RTA	Move freely Stay at home Don't Know/RTA
Do you think this person has Covid–19/ Novel Corona virus?	Yes No Not sure what Covid–19/ corona virus is RTA	Yes No Don't Know RTA

4.1.1 Choosing the Levels for the Identity Attribute

To choose appropriate levels of the Identity attribute in our experiment, we needed to know how Malawians and Zambians think about insiders and outsiders. Previous research emphasized ethnic (Dionne, 2015; Ejdemyr, Kramon, & Robinson, 2018; Posner, 2004), village (Dionne, 2015), and regional (Ferree & Horowitz, 2010; Posner, 2005) divisions as critical drivers of Malawian and Zambian politics. Cross-border migration and flows of refugees may also create divisions within African communities (Adida, 2014).

For insight into how these divisions map to perceptions of insiders and outsiders in both countries, we draw on a 2019 study, the Local Governance Performance Index (LGPI) survey,[4] administered to over 20,000 Malawians and Zambians in the regions along the Zambian/Malawi border and the Lilongwe and Lusaka metropolitan areas (Lust et al., 2020). The LGPI asked respondents: "Sometimes certain people are considered to be 'outsiders' from a community. When you hear the term 'outsider,' who do you think of?" Read out loud response options included "one from a different eth-

[4]For more information, please visit www.gld.gu.se.

nic group," "one from a different country," "one from a different neighborhood in this region," and "one from a different region." Multiple responses were possible, as was an "other" option. See Table 3.

Table 3: 'Outsiders' in Malawi and Zambia (from LGPI 2019)

	Malawi	**Zambia**
Who do you think of as an outsider?	**Proportion**	**Proportion**
One from another ethnic group	30%	14%
One from another country	57%	72%
One from another neighborhood in this region	35%	20%
One from another region	38%	15%
Other	3%	3%
Do not understand the question	5%	5%
Observations	10,302	9,864

"One from a different country" was the most common answer in both countries, with more than half of respondents in both samples providing this response, and somewhat higher prevalence in Zambia. Given the salience of the "different country" category and early narratives characterizing Covid-19 as "imported,"[5] we sought groups with clear connections to other countries that would nonetheless be familiar to survey respondents and plausible as residents within their communities.

For Malawians, we chose Zambians and Malawians of South Asian descent, or *Amwenye*. Populations frequently cross borders in Southern Africa, and the border between Zambia and Malawi is particularly porous (Posner, 2004). Zambians therefore seemed a logical candidate for local "outsiders" in the Malawian context. They have an immediate connection to a different country, yet would not be unusual residents in Malawian neighborhoods. Amwenye represent a racialized "other" with a distinct foreign connection. They are primarily descendants of Indian migrants to Malawi who arrived in the

[5]Capital FM Malawi, a radio station, characterized most Covid-19 cases as "imported," (Tengatenga et al., 2021). The index case in Malawi was a woman who had travelled to India to see family and an early center of disease outbreak was at the border in Mwanza, the exit and entry point for the migration route to South Africa (Dionne et al., 2021).

late 19th and early 20th centuries as small traders. There is a long history of othering Amwenye (Power, 1993) that persists today. Resentment focuses on their elite economic status and stereotypes about business practices.[6] The first cases of Covid–19 in Malawi coincidentally involved an Amwenye couple and their domestic worker. Amwneye thus represent a racialized other, if not a truly marginalized one,[7] and as such offer a way to test (at least partially) Hypothesis 2 that identity effects are strongest for marginalized and/or racially different groups.

For Zambia, we operationalized outsiders as Malawians and Tanzanians. Both groups share a porous border with Zambia. It would not be unusual in Lusaka or a border town to have migrants from either country living in communities. Additionally, Tanzania was believed to have a very high rate of Covid-19 infection at the time of the Zambian survey (Mwai & Giles, 2020). Malawi, on the other hand, likely had equal or lower levels of Covid-19 infection than Zambia, suggesting little reason to associate Malawians with the virus. We therefore have four different outsider groups, where we might expect stronger effects in Malawi for Amwenye vs. Zambians, and stronger effects in Zambia for Tanzanians versus Malawians.

We also manipulated the neighbor's time in the community (a few months, many years). This allows us to explore another dimension of "outsiderness." It also anchors the individual as a neighbor and resident in the community instead of a visitor and ensures that their objective risk of introducing contagion to the community is the same as any other resident (Dutta & Rao, 2015), and neutralizes "proximity" (Onoma, 2020) as a potential confound.

To evaluate prejudicial attitudes toward these identities, in each of the three Covid surveys we asked about willingness to have as neighbors individuals from the relevant

[6] For example, see *Malawi Nyasa Times* editorial by Mike Fiko (2020) from October 24, 2020 "Don't Blame it on 'Amwenye': It's Malawi Public Servants and Politicians, stupid!"

[7] Cohen (1999) and Dionne et al. (2021) see marginalized groups as those systematically excluded from access to resources and power. Amwenye, as an economic elite, do not fit this characterization

outgroups. This is a common technique for measuring prejudice in surveys (Schuman & Bobo, 1988; Strabac & Listhaug, 2008).[8] As a comparison, we also asked about willingness to have "people with Covid-19" as neighbors.

The results in Tables 4 and 5 suggest that the outsider groups we chose for the experiment tap into prejudicial attitudes for many respondents, particularly in Malawi, where only 24 percent of respondents in the first survey expressed willingness to live next to people of Indian origin, 27 percent expressed willingness to live next to people of Zambian origin, and 21 percent next to people with Covid-19. These numbers increased somewhat in the second round. Zambians, in contrast, were relatively more willing to live next to Malawians (60 percent) and Tanzanians (56 percent) versus those with Covid-19 (23 percent). These differences across countries could reflect variations in sampling or they could indicate less xenophobia toward the particular groups chosen for the experiment in Zambia versus Malawi or lower latent prejudice in Zambia.

Table 4: Attitudes in Malawi (from Covid phone surveys)

	Malawi Round 1	**Malawi Round 2**
Would you be willing to have them as neighbors?	**Proportion Yes**	**Proportion Yes**
Amwenye	24%	40%
People of Zambian origin	27%	43%
People with Covid-19	21%	35%
Observations	4,641	4,372

4.2 Sampling Method

The first Malawian survey, administered from May 5, 2020 through June 2, 2020, reached 4641 respondents in all regions of the country. The second Malawian survey, administered from March 9, 2021 to May 1, 2021, surveyed 4356 respondents, of which 1974

[8]These questions came at the end of each survey. They are therefore measured post-treatment; a large amount of unrelated content nonetheless separated the experiment and these questions, hopefully mitigating any treatment effects.

Table 5: **Attitudes in Zambia (from Covid phone survey)**

	Zambia
Would you be willing to have them as neighbor?	**Proportion Yes**
People of Malawian origin	60%
People of Tanzanian origin	56%
People with Covid-19	23%
Observations	2,198

also took the first wave survey. The sampling frame for the Malawi surveys consisted of phone numbers obtained in the LGPI 2019 and LGPI 2016 Malawi surveys.[9] The Zambian survey, in the field from July 2 through August 13 in 2020, drew from the Lusaka region and the areas bordering Malawi and Tanzania and reached 2198 respondents. The sampling frame for the Zambia Covid–19 telephone survey consisted of telephone numbers obtained from the LGPI 2019 Zambia survey. The Zambian sampling frame did not contain individuals in the Copper-Belt, North-Western, Western, or Southern regions of the country. (See Appendix B for more details on sampling.)

5 Data

5.1 Contextual Data

Nearly all (99 percent) of the respondents in both countries had heard of Covid–19. Anxiety about falling ill was substantially higher in the Zambian sample (82 percent of the Zambian sample were somewhat or very worried, compared with 54 percent in Malawi Wave 1 and 55 percent in Wave 2). Covid–19 was thus very much on the minds of our respondents, even when the disease itself had not yet penetrated deeply into the

[9]The original sampling frame for the Malawian surveys did not contain ten of the country's 28 districts. Due to movement of respondents and the allowance of replacement respondents, we were able to survey some individuals in districts not in the original sampling frame such that we collected data from 27 out of 28 district (8 of these had very small sample sizes).

population, see Figure 1.

5.2 Demographic Data

Summaries of respondent demographics can be found in Table 6. Comparing the Malawian waves, we see a slight change in the gender distribution and similar age and education distributions. The Zambian sample is more educated and a bit younger than the Malawian samples.

Table 6: Survey Sample Demographics

Demographic	Malawi Wave 1	Malawi Wave 2	Zambia
Male	56	49	51
Female	44	51	49
18-34 Years	48	43	52
35-54 Years	40	40	33
55 Years +	12	17	16
Little Formal Education	7	6	4
Primary Education	56	53	27
Secondary Education	35	36	53
University Education	1	4	17
Observations	4641	4356	2198

5.3 Experimental Data

Turning to our outcome variables, when averaged across experimental conditions, we find that 78 percent in Malawi Wave 1, 76 percent in Malawi Wave 2 and 71 percent in Zambia were willing to help those in need get to the hospital. We also find relatively low levels of support for free movement of ailing individuals (15 percent Malawi Wave 1; 18 percent Malawi Wave 2; 10 percent Zambia). A third of the Malawian Wave 1 sample thought the neighbor had Covid–19 and another 20 percent were unsure. These percentages held steady in Wave 2, in which a third thought that the neighbor had Covid-19 and 17 percent were unsure. A third of the Zambian sample suspected Covid-

19 and another 30 percent were not sure. The results we present for the outcome (Has Covid–19) excludes respondents who stated that they were unsure.[10]

Reflecting the discussion in Section 4.1.1, we define a new variable `Outsider`, based on the experimental attribute `Identity`. `Outsider` is equal to 0 if `Identity` is equal to the national identity of the sample (Malawian or Zambian depending) and 1 otherwise. We estimate group specific effects to evaluate Hypothesis 2 but otherwise discuss results using the `Outsider` variable.

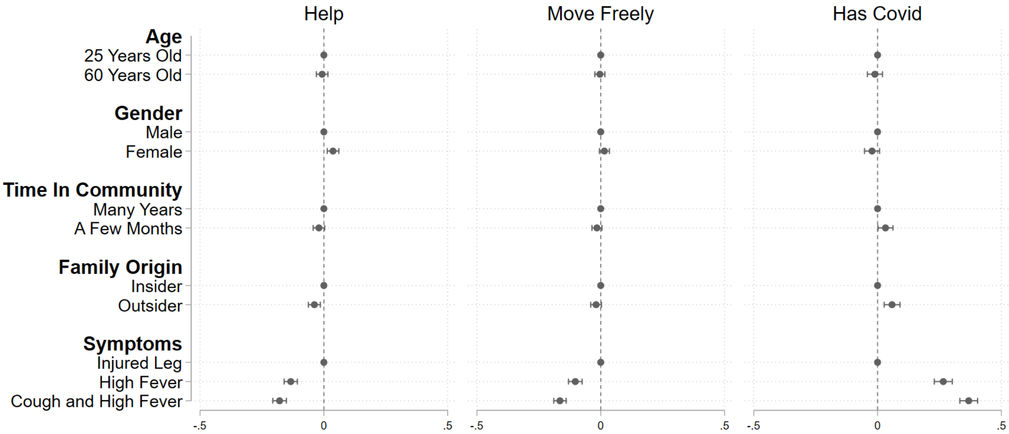
6 Results

6.1 Main Effects

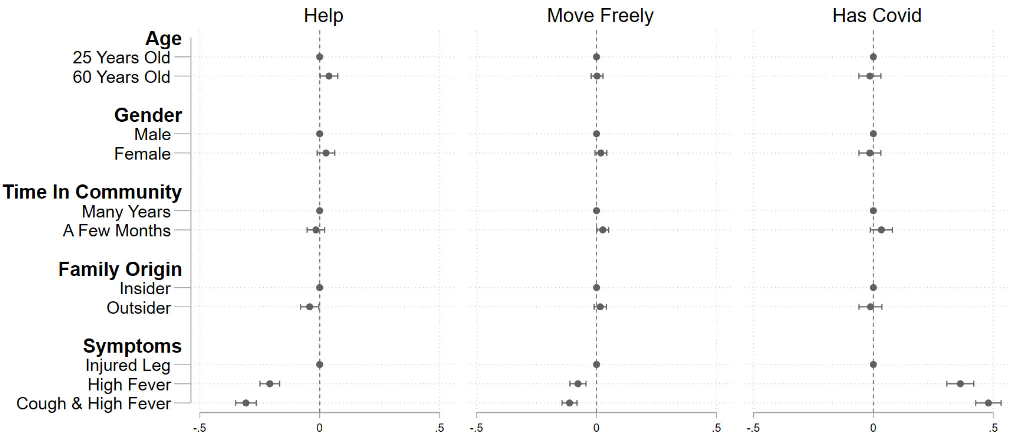
Following Hainmueller et al. (2013), we use ordinary least squares (OLS) regression to estimate average marginal component effects (AMCE). Subfigures 2a, 2b, and 2c display the results for our Malawi Wave 1, Zambia, and Malawi Wave 2 samples, respectively. They indicate point estimates with a dot and 95% confidence intervals with horizontal lines for the AMCE of each of the randomized profile attributes on the respondent's willingness to help the neighbor to the hospital, belief that the neighbor should be allowed to move about freely in the community, and belief that the neighbor has Covid–19. Baseline categories for each attribute are indicated with a dot at the vertical dashed line at 0. (See Table A1 in Appendix A for full results).

[10]A multinomial logistic regression analysis shows similar results to dropping these observations (See Appendix C for details).

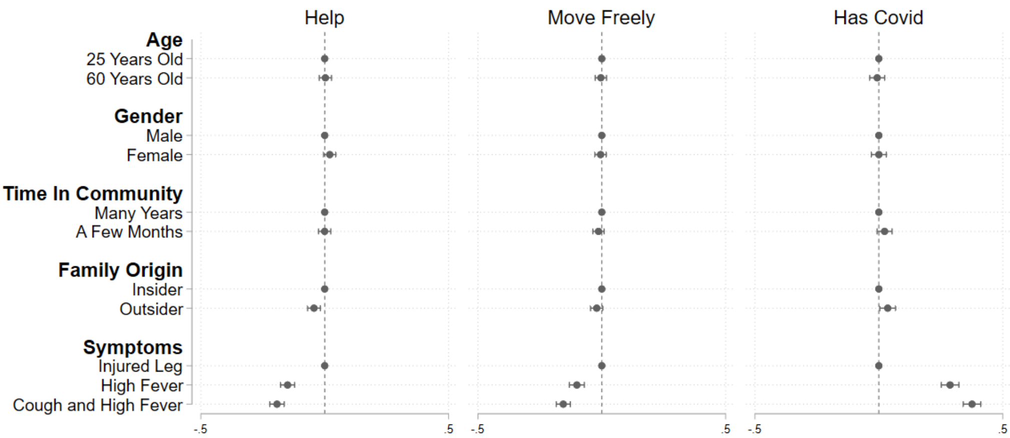
Figure 2: AMCEs of Profile Attributes on Outcomes in Malawi and Zambia



(a) Malawi First Wave



(b) Zambia



(c) Malawi Second Wave

We find mixed support for Hypothesis 1 that identity shapes perceptions and responses. Consistent with the hypothesis, Malawians are 6 (Wave 1) and 4 (Wave 2) percentage points (SE=0.02) more likely to perceive that an Amwenye or Zambian neighbor has Covid–19 versus a Malawian neighbor. We find weak evidence that Malawians also associate community newcomers with disease more than long-term residents, but this effect is small (3 percentage points, SE=0.015) and disappears in the second wave. Zambians, on the other hand, do not associate disease with any of our measures of outsider status: they are not more likely to perceive that Malawians or Tanzanians have Covid-19 than Zambians, nor are they more likely to perceive the disease in community newcomers vs. long-term residents. We also find that Malawians and Zambians are marginally more likely (4 percentage points with SE=0.01 and 0.02, respectively) to say they would help insiders versus outsiders get to the hospital. In our second wave in Malawi, this effect increased to 6 percentage points (SE=0.01). However, neither Malawians nor Zambians differentiate between insiders and outsiders when it comes to moving around the community, although Zambians do (unexpectedly) report slightly *more* willingness to allow newcomers to move freely than longer term residents. Altogether, we find limited evidence of identity effects, and these are strongest in Malawi.

Hypothesis 2 suggests that effects would be strongest for previously marginalized and/or racially similar groups. Amwenye in Malawi are the only racially distinct group of the four we analyze. We therefore consider whether effects are stronger in Malawi for Amwenye versus Zambians. For Wave 1, with Malawians as the base category, the estimated effect of the Amwenye treatment is 7 percentage points (SE=0.02), whereas that for Zambians was 5 percentage points (SE=.02), but this difference is not statistically significant (see Appendix A.2). For Wave 2, we find that Zambians, not Amwenye, are seen as significantly more likely to have contracted the disease (by 5 percentage points, SE = 0.02) (see Table A4 in Appendix A). We do not find significant differences in willingness to help or allow free movement across subgroups in any sample. Evidence for H2 is therefore weak. The only significant difference we find is for Zambians in Malawi in Wave 2. As the Wave 2 Malawi survey fielded just after a regional wave that hit Zambia harder (at least in terms of sheer numbers) than Malawi, these findings suggest that beliefs about who has disease reflect local and temporal patterns rather than static and unchanging stereotypes.[11]

While our findings on identity are surprisingly small (especially in light of prior literature), we do find strong support for Hypothesis 3 on symptoms. Malawians and Zambians clearly associate Covid–19 with high fever and cough. Malawians are 37 (Wave 1) and 38 (Wave 2) percentage points (SE = 0.02) more likely to judge neighbors with a high fever and bad cough than those with an injured leg to have the disease. They are 27 and 29 percentage points (SE=0.02) more likely to do so for those with just a high fever. Zambians display a similar, even starker pattern. They are 48 percentage points (SE=0.03) more likely to consider those with a high fever and bad cough to have Covid–19 than those with an injured leg. Similarly, they are 36 percentage points (SE=0.03) more likely to consider those with just a high fever to have Covid–19 than those with an injured leg.

Malawians and Zambians also report less willingness to help neighbors described with clear symptoms. In Malawi, respondents were 18 percentage points (SE=0.01) less likely to indicate that they would help someone with the strongest Covid–19 symptoms (a high fever and a bad cough) to the hospital, as compared to someone with an injured leg. This finding persists in our second wave, with the negative effect of strong symptoms at 19 percentage points (SE=0.02). Zambians respond even more forcefully to symptoms: they are 31 percentage points (SE=0.02) less likely to say they would help a neighbor with strong Covid–19 symptoms. These results are particularly noteworthy given norms in both countries favoring helping others in the community. The effect of symptoms carry

[11] It remains possible that findings would be different for truly marginalized groups.

through to restrictions on movement. Malawians are 15 percentage points (SE=0.01) (Wave 1 and Wave 2) less likely to endorse free movement for those with high fever and a cough, while Zambians are 11 percentage points (SE=0.02) less likely to endorse free movement.

Our results thus show that outward symptoms of disease drive perceptions and responses more than identities. Malawians and Zambians respond strongly to disease symptoms, associating high fever (alone or accompanied by a cough) with Covid–19 and indicating that they would sharply curtail interactions with neighbors who display these symptoms. On the other hand, with symptoms experimentally specified, we only find partial support for the role of insider/outsider stereotypes, with results strongest in Malawi.

6.2 Symptoms and Outsider Identity Interactions

Next we consider Hypothesis 4 and whether ambiguity about Covid–19 symptoms enhances the effects of insider/outsider status in Malawi.[12] If it does, this suggests that insider/outsider stereotypes are a heuristic used to overcome uncertainty about contagion risk, not a hardwired psychological response rooted in prejudice.

To evaluate, we estimate Average Component Interaction Effects (ACIEs) using the procedure described by Hainmueller et al. (2013). If individuals use insider/outsider status as a heuristic for disease, as expected by behavioral immune system theories (Schaller, 2011), then we would expect more ambiguous symptoms to induce stronger insider/outsider effects across our three dependent variables. Across our three sets of symptoms, "fever and cough" is the least ambiguous as it clearly indicates Covid-19 (indeed effects are strongest for it). In contrast, "fever" alone is more ambiguous as fevers can be caused by many diseases. "Injured and infected leg" could be seen as

[12]We do not do a similar analysis for Zambia given null insider/outsider results.

unambiguously indicating a lack of Covid-19 (and indeed, this was our assumption in our Pre-Analysis Plan), but given asymptomatic transmission, it might plausibly be seen as ambiguous as well.

Table 7: Malawi: Average Component Interaction Effects

	Help	Move Freely	Has Covid
Base: Insider			
Outsider	-0.0382*	-0.0420*	0.0861**
	(0.0162)	(0.0191)	(0.0279)
Base: Injured Leg			
Cough and High Fever	-0.183***	-0.194***	0.420***
	(0.0231)	(0.0219)	(0.0328)
High Fever	-0.130***	-0.120***	0.270***
	(0.0221)	(0.0220)	(0.0329)
Interactions of Origin and Symptoms			
Outsider × Cough and High Fever	0.00560	0.0433	-0.0776^{+}
	(0.0290)	(0.0268)	(0.0400)
Outsider × High Fever	-0.00636	0.0259	-0.00650
	(0.0280)	(0.0269)	(0.0403)
Constant	0.910***	0.266***	0.158***
	(0.0157)	(0.0179)	(0.0260)
Observations	4627	4548	3652

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Not Shown: Age, Gender, Time in Community

Table 7 shows the estimation of the interaction effects. We find weak evidence of an interaction between symptoms and insider/outsider status on the 'Has Covid–19' outcome with an estimated interaction effect of Outsider and Cough and High Fever of -0.0776 (SE = 0.04) which almost entirely negates the Outsider effect of 0.0861 (SE = 0.03). We find no evidence of an interaction between Outsider and High Fever. This suggests that Malawians may only rely on insider/outsider identities to determine infection status when clear Covid–19 symptoms are not available.

Confronted with clear disease symptoms, they judge insiders and outsiders the same way. As Kam (2019) argues for Americans, there may be a point at which disease risk becomes high enough that insider/outsider distinctions fall to the wayside, however we emphasize that these results are estimated with some uncertainty and should be viewed as suggestive. We did not find evidence of any interaction effects (symptoms X identity) for our other outcomes and the interaction effect on disease perceptions disappears in the second wave of the Malawian survey (see Table A7).

7 Scope Conditions and Other Limitations

Malawians and Zambians in our samples responded more strongly and consistently to overt indications of disease than the information conveyed about identity. We draw cautious optimism from these findings about the ability of the pandemic to exacerbate identity divisions, at least through the mechanism of disease threat. We nonetheless acknowledge some important aspects to our study design that might have lead to smaller identity effects than would have been observed had we studied different countries, groups, or samples. We also explore whether social desirability bias might have dampened identity effects.

Our findings pertain to two countries, Malawi and Zambia, without histories of extensive identity-based conflict. Moreover, while Amwenye are a racialized minority that has been subjected to some degree of discrimination in the past, they are not a truly marginalized group. It may be the case that identity results would be stronger for more marginalized groups, or groups that are both racially distinct and economically disadvantaged, in countries with greater history of identity violence. They might also be stronger had we not limited our experiment to "neighbors," as less proximate individuals from outsider groups may provoke stronger boundary policing (Onoma, 2020). While we cannot rule out the possibility of stronger identity effects in other contexts, we nonetheless can say that, even for racial minorities like the Amwenye, disease threat effects do not appear to be universal or automatic.

We also consider the sample itself. Given our use of cell-phones, which are not universal in Malawi or Zambia, we may have slightly over-sampled more educated and resourced respondents. Indeed, our Malawian sample is somewhat more educated on average than the more representative sample of the 2016 Malawian LGPI survey (14% of our sample had completed secondary education, versus only 5% of the 2016 LGPI sample). We do not have a comparable benchmark for Zambia, but we expect the same dynamic there to also lead to over-representation of educated respondents. If more educated respondents in turn respond more to symptoms and less to identity, then identity effects could be stronger in a more fully representative sample. To evaluate this possibility, we interacted respondent education level with the insider/outsider treatment, using a combined sample of Malawi and Zambia (See Appendix A). The results suggest that respondents with at least a secondary education are more willing to let outsiders move about freely than respondents with a primary education or less when all other attributes are at their baseline levels. However, with regards to disease perceptions and willingness to help outsiders, we do not find any significant differences across education groups. We therefore do not believe that the mild skewing of our sample toward more educated Malawians and Zambians has significantly attenuated identity effects.

Finally, it is possible that social desirability bias has masked the true effects of identity in our results. We cannot rule out attenuation from this source but are nonetheless skeptical that it plays a large role. Our respondents revealed fairly high levels of prejudice in direct questions about willingness to live next to people from the outsider groups in the study (see Table 5). It therefore seems unlikely that they would conceal prejudice in a more subtle experiment that manipulated several dimensions at once. Moreover, given the strong culture of mutual reciprocity in Malawian and Zambian communities, we would expect people to be the least willing to reveal bias in helping neighbors get medical help, yet we find the most consistent identity effects for this outcome.

8 Reflections and Conclusions

Will the Covid–19 pandemic, which has sickened and killed millions of people and deeply damaged economies around the world, also have long-term social and political effects? In this paper, we explored one possible fallout of the disease: the exacerbation of group conflict. Prior work in a variety of fields—most notably psychology—suggests that threats posed by disease provoke anti-outsider sentiments and lead to the hardening of boundaries between groups.

We have shown small but persistent identity effects on perceptions of disease in Malawi (but not Zambia) and willingness to help neighbors in both countries. We do not find identity effects for the free movement outcome. These effects do not seem particularly acute for the racially distinct group we analyze, and they do not shift significantly in response to the waxing and waning of the pandemic. In contrast, symptoms, which have been ignored or downplayed in prominent previous research, exert large, persistent, and robust effects on perceptions of disease, willingness to help, and support for free movement.

Our results thus offer something of an assurance about the potential identity effects of the pandemic, albeit one conditioned by the limitation of our method, which studies hypothetical responses to a vignette, not actual behavior in the real-world, and does so in two countries without histories of significant identity–based violence between the groups analyzed. At a minimum, we can say that disease threat effects are not universal. In the two countries we study, for the groups we study, we do not find evidence of major exacerbation of identity boundaries.

References

Aarøe, L., Petersen, M. B., & Arceneaux, K. (2017). The Behavioral Immune System Shapes Political Intuitions: Why and How Individual Differences in Disgust Sensitivity Underlie Opposition to Immigration. *American Political Science Review*, *111*(2), 277–294. Retrieved 2020-06-23, from `https://www.cambridge.org/core/product/identifier/S0003055416000770/type/journal_article` doi: 10.1017/S0003055416000770

Adida, C. L. (2014). *Immigrant exclusion and insecurity in Africa.* New York, NY: Cambridge University Press.

Adida, C. L., Dionne, K. Y., & Platas, M. R. (2018). Ebola, elections, and immigration: How politicizing an epidemic can shape public attitudes. *Politics, Groups, and Identities*, 1–27. Retrieved 2020-05-10, from `https://www.tandfonline.com/doi/full/10.1080/21565503.2018.1484376` doi: 10.1080/21565503.2018.1484376

Albertson, B., & Gadarian, S. (2015). Ebola, Anxiety, and Public Support for Protective Policies. *Political Science and Politics*, *48*, 8–9.

Arriola, L. R., & Grossman, A. N. (2020). Ethnic Marginalization and (Non)Compliance in Public Health Emergencies.
(Publisher: Harvard Dataverse type: dataset)

Benton, A., & Dionne, K. Y. (2015). International Political Economy and the 2014 West African Ebola Outbreak. *African Studies Review*, *58*(1), 223–236.

Casey, L. (2015). Emtions and the Politics of Ebola. *Political Science and Politics*, *48*, 7–8.

Cohen, C. J. (1999). *The Boundaries of Blackness: AIDS and the Breakdown of Black Politics* (1edition ed.). Chicago: University of Chicago Press.

Dionne, K. Y. (2015). Social Networks, Ethnic Diversity, and Cooperative Behavior in

Rural Malawi. *Journal of Theoretical Politics*, *27*(4), 522–543. Retrieved 2020-07-08, from http://journals.sagepub.com/doi/10.1177/0951629814556173 doi: 10.1177/0951629814556173

Dionne, K. Y., Dulani, B., & Fischer, S. E. (2021). Pandemic Amid Political Crisis: Malawi's Experience with and Response to COVID-19. In S. L. Greer, E. J. King, E. M. d. Fonseca, & A. Peralta-Santos (Eds.), *Coronavirus Politics* (pp. 541–559). University of Michigan Press. Retrieved 2021-07-08, from http://www.jstor.org/stable/10.3998/mpub.11927713.32

Dionne, K. Y., & Turkmen, F. F. (2020). The Politics of Pandemic Othering: Putting COVID-19 in Global and Historical Context. *International Organization*, *74*(S1), E213–E230. doi: 10.1017/S0020818320000405

Dutta, S., & Rao, H. (2015). Infectious Diseases, Contamination Rumors and Ethnic Violence: Regimental mutinies in the Bengal Native Army in 1857 India. *Organizational Behavior and Human Decision Processes*, *129*, 36–47. Retrieved 2020-05-10, from https://linkinghub.elsevier.com/retrieve/pii/S0749597814000806 doi: 10.1016/j.obhdp.2014.10.004

Ejdemyr, S., Kramon, E., & Robinson, A. L. (2018). Segregation, Ethnic Favoritism, and the Strategic Targeting of Local Public Goods. *Comparative Political Studies*, *51*(9), 1111–1143. Retrieved 2020-07-08, from http://journals.sagepub.com/doi/10.1177/0010414017730079 doi: 10.1177/0010414017730079

Faulkner, J., Schaller, M., Park, J. H., & Duncan, L. A. (2004). Evolved Disease-Avoidance Mechanisms and Contemporary Xenophobic Attitudes. *Group Processes & Intergroup Relations*, *7*(4), 333–353. Retrieved 2020-05-10, from http://journals.sagepub.com/doi/10.1177/1368430204046142 doi: 10.1177/1368430204046142

Ferree, K., & Horowitz, J. (2010). Ties That Bind? The Rise and Decline of Ethno-

Regional Partisanship in Malawi, 1994–2009. *Democratization*, *17*(3), 534–563. Retrieved 2020-01-10, from http://www.tandfonline.com/doi/full/10.1080/13510341003700394 doi: 10.1080/13510341003700394

Fiko, M. (2020). 'Don't Blame it on 'Amwenye': It's Malawi Public Servants and Politicians, Stupid! *Malawi Nyasa Times*. Retrieved 2021-07-08, from https://www.nyasatimes.com/dont-blame-it-on-amwenye-its-malawi-public-servants-and-politicians-stupid/

Hainmueller, J., Hopkins, D. J., & Yamamoto, T. (2013). Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices Via Stated Preference Experiments. *Political Analysis*, *22*(1), 1–30.

Hasell, M. E. B. D. e. a., J. (2020). *A cross-country database of covid-19 testing.* Sci Data 7, 345. (https://doi.org/10.1038/s41597-020-00688-8)

Kam, C. D. (2019). Infectious Disease, Disgust, and Imagining the Other. *The Journal of Politics*, *81*(4), 1371–1387. Retrieved 2020-06-23, from https://www.journals.uchicago.edu/doi/10.1086/704438 doi: 10.1086/704438

Lieberman, E. S. (2007). Ethnic Politics, Risk, and Policy-Making: A Cross-National Statistical Analysis of Government Responses to HIV/AIDS. *Comparative Political Studies*, *40*(12), 1407–1432. Retrieved 2020-05-10, from http://journals.sagepub.com/doi/10.1177/0010414007306862 doi: 10.1177/0010414007306862

Lieberman, E. S. (2009). *Boundaries of Contagion: How Ethnic Politics Have Shaped Government Responses to AIDS.* Princeton, N.J: Princeton University Press. (OCLC: ocn276340693)

Lust, E., Harris, A., Kao, K., Landry, P., Dulani, B., Zgambo, A., ... Nickel, S. (2016). Malawi truncated dataset. *The Program on Governance and Local Development*. Retrieved from http://gld.gu.se/en/publications-and-data/data/

lgpi-malawi-2016-truncated/

Lust, E., Kao, K., Landry, P. F., Harris, A., Dulani, B., (co PIs), ... Lueders, H. (2020). The Local Governance and Performance Index (LGPI) 2019: Kenya, Malawi, Zambia. *The Program on Governance and Local Development*. Retrieved from http://gld.gu.se

Mwai, P., & Giles, C. (2020). "Coronavirus in Tanzania: What do we know?". *BBC News*. Retrieved 2020-07-08, from https://www.bbc.com/news/world-africa-52723594

Navarrete, C. D., & Fessler, D. M. (2006). Disease Avoidance and Ethnocentrism: The Effects of Disease Vulnerability and Disgust Sensitivity on Intergroup Attitudes. *Evolution and Human Behavior*, *27*(4), 270–282. Retrieved 2020-05-10, from https://linkinghub.elsevier.com/retrieve/pii/S109051380600002X doi: 10.1016/j.evolhumbehav.2005.12.001

Navarrete, C. D., Fessler, D. M., & Eng, S. J. (2007). Elevated Ethnocentrism in the First Trimester of Pregnancy. *Evolution and Human Behavior*, *28*(1), 60–65. Retrieved 2020-06-23, from https://linkinghub.elsevier.com/retrieve/pii/S1090513806000651 doi: 10.1016/j.evolhumbehav.2006.06.002

Nohl, J. (2006). *The Black Death: A Chronicle of the Plague*. Westholme ; Windsor [distributor. (OCLC: 1063340490)

Onoma, A. K. (2020). Xenophobia's Contours During an Ebola Epidemic: Proximity and the Targeting of Peul Migrants in Senegal. *African Studies Review*, *63*(2), 353–374. doi: 10.1017/asr.2019.38

Posner, D. N. (2004). The Political Salience of Cultural Difference: Why Chewas and Tumbukas Are Allies in Zambia and Adversaries in Malawi. *American Political Science Review*, *98*(4), 529–545. Retrieved 2020-07-08, from https://www.cambridge.org/core/product/identifier/

S0003055404041334/type/journal_article doi: 10.1017/S0003055404041334

Posner, D. N. (2005). *Institutions and Ethnic Politics in Africa*. Cambridge University Press.

Power, J. (1993). Race, Class, Ethnicity, and Anglo-Indian Trade Rivalry in Colonial Malawi, 1910-1945. *The International Journal of African Historical Studies*, *26*(3), 575–607.

Reny, T. T., & Barreto, M. A. (2020). Xenophobia in the Time of Pandemic: Othering, Anti-Asian Attitudes, and COVID-19. *Politics, Groups, and Identities*, *0*(0), 1-24. doi: 10.1080/21565503.2020.1769693

Schaller, M. (2011). The Behavioural Immune System and the Psychology of Human Sociality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1583), 3418–3426. Retrieved 2020-06-23, from https://royalsocietypublishing.org/doi/10.1098/rstb.2011.0029 doi: 10.1098/rstb.2011.0029

Schuman, H., & Bobo, L. (1988). Survey-Based Experiments on White Racial Attitudes Toward Residential Integration. *American Journal of Sociology*, *94*(2), 273–299.

Strabac, Z., & Listhaug, O. (2008). Anti-Muslim Prejudice in Europe: A Multilevel Analysis of Survey Data from 30 Countries. *Social Science Research*, *37*(1), 268–286.

Suedfeld, P., & Schaller, M. (2002). Authoritarianism and the Holocaust: Some Cognitive and Affective Implications. In *Understanding Genocide: The Social Psychology of the Holocaust* (pp. 68–90). Oxford, United Kingdom: Oxford University Press.

Tengatenga, J., Tengatenga Duley, S. M., & Tengatenga, C. J. (2021). Zimitsani Moto: Understanding the Malawi COVID-19 Response. *Laws*, *10*(2), 20.

Online Appendix

A Additional Analyses

A.1 Regression Tables

This section reports the full regression tables associated with the figures in the paper.

Table A1: Malawi Wave 1: Average Marginal Component Effects for Figure 1. Binary Outsider Variable.

	Help	Move Freely	Has Covid
60 Years Old	-0.00703	-0.00346	-0.0108
(Base: 25 Years Old)	(0.0119)	(0.0103)	(0.0155)
Female	0.0369**	0.0148	-0.0223
(Base: Male)	(0.0119)	(0.0103)	(0.0155)
A Few Months	-0.0200	-0.0152	0.0320*
(Base: Many Years)	(0.0119)	(0.0103)	(0.0155)
Outsider	-0.0384**	-0.0187	0.0584***
(Base: Insider)	(0.0123)	(0.0111)	(0.0162)
High Fever	-0.134***	-0.103***	0.265***
(Base: Injured Leg)	(0.0136)	(0.0140)	(0.0186)
Cough and High Fever	-0.179***	-0.165***	0.368***
(Base: Injured Leg)	(0.0140)	(0.0128)	(0.0182)
Constant	0.910***	0.250***	0.176***
	(0.0148)	(0.0159)	(0.0197)
Observations	4627	4548	3652

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A2: Malawi Wave 1: Average Marginal Component Effects. Three Level Outsider Variable.

	Help	Move Freely	Has Covid
60 Years Old	-0.00700	-0.00350	-0.0108
(Base: 25 Years Old)	(0.0119)	(0.0103)	(0.0155)
Female	0.0369**	0.0149	-0.0216
(Base: Male)	(0.0119)	(0.0103)	(0.0155)
A Few Months	-0.0200^{+}	-0.0153	0.0319*
(Base: Many Years)	(0.0119)	(0.0103)	(0.0155)
Mmwenye	-0.0407**	-0.0140	0.0715***
(Base: Malawian)	(0.0145)	(0.0128)	(0.0190)
Zambian	-0.0362*	-0.0233	0.0459*
(Base: Malawian)	(0.0143)	(0.0126)	(0.0187)
Cough and High Fever	-0.179***	-0.165***	0.368***
(Base: Injured Leg)	(0.0140)	(0.0128)	(0.0182)
High Fever	-0.134***	-0.103***	0.265***
(Base: Injured Leg)	(0.0136)	(0.0140)	(0.0186)
Constant	0.890***	0.235***	0.208***
	(0.0154)	(0.0162)	(0.0206)
Observations	4627	4548	3652

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A3: Malawi Round 2: Average Marginal Component Effects. Binary Outsider Variable.

	Help	Move Freely	Has Covid
60 Years Old	0.00255	-0.00379	-0.00664
(Base: 25 Years Old)	(0.0126)	(0.0116)	(0.0154)
Female	0.0200	-0.00479	0.000248
(Base: Male)	(0.0126)	(0.0116)	(0.0154)
A Few Months	-0.000473	-0.0133	0.0232
(Base: Many Years)	(0.0126)	(0.0116)	(0.0154)
Outsider	-0.0433**	-0.0205	0.0358*
(Base: Insider)	(0.0132)	(0.0126)	(0.0163)
High Fever	-0.150***	-0.101***	0.288***
(Base: Injured Leg)	(0.0144)	(0.0154)	(0.0181)
Cough and High Fever	-0.193***	-0.155***	0.376***
(Base: Injured Leg)	(0.0149)	(0.0144)	(0.0180)
Constant	0.898***	0.292***	0.146***
	(0.0168)	(0.0180)	(0.0205)
Observations	4345	4281	3628

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A4: Malawi Round 2: Average Marginal Component Effects. Three Level Outsider Variable.

	Help	Move Freely	Has Covid
60 Years Old	0.00259	-0.00376	-0.00668
(Base: 25 Years Old)	(0.0126)	(0.0116)	(0.0154)
Female	0.0199	-0.00485	0.000660
(Base: Male)	(0.0126)	(0.0116)	(0.0154)
A Few Months	-0.000529	-0.0133	0.0236
(Base: Many Years)	(0.0126)	(0.0116)	(0.0154)
Mmwenye	-0.0369*	-0.0173	0.0176
(Base: Malawian)	(0.0154)	(0.0146)	(0.0189)
Zambian	-0.0494**	-0.0235	0.0530**
(Base: Malawian)	(0.0153)	(0.0143)	(0.0187)
High Fever	-0.150***	-0.101***	0.288***
(Base: Injured Leg)	(0.0144)	(0.0154)	(0.0181)
Cough and High Fever	-0.193***	-0.155***	0.376***
(Base: Injured Leg)	(0.0149)	(0.0144)	(0.0180)
Constant	0.899***	0.292***	0.146***
	(0.0168)	(0.0180)	(0.0205)
Observations	4345	4281	3628

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A5: Zambia: Average Marginal Component Effects. Binary Outsider Variable.

	Help	MoveFreely	HasCovid
60 Years Old	0.0387*	0.00236	-0.0145
(Base: 25 Years Old)	(0.0187)	(0.0126)	(0.0233)
Female	0.0264	0.0180	-0.0145
(Base: Male)	(0.0187)	(0.0125)	(0.0233)
A Few Months	-0.0157	0.0261*	0.0332
(Base: Many Years)	(0.0186)	(0.0126)	(0.0233)
Outsider	-0.0416*	0.0156	-0.0121
(Base: Insider)	(0.0194)	(0.0130)	(0.0244)
High Fever	-0.208***	-0.0773***	0.363***
(Base: Injured Leg)	(0.0209)	(0.0172)	(0.0288)
Cough and High Fever	-0.308***	-0.113***	0.480***
(Base: Injured Leg)	(0.0219)	(0.0160)	(0.0270)
Constant	0.886***	0.127***	0.214***
	(0.0244)	(0.0191)	(0.0317)
Observations	2181	2169	1540

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A6: Zambia: Average Marginal Component Effects. Three Level Outsider Variable.

	Help	MoveFreely	HasCovid
60 Years Old	0.0387*	0.00237	-0.0149
(Base: 25 Years Old)	(0.0187)	(0.0126)	(0.0234)
Female	0.0267	0.0180	-0.0149
(Base: Male)	(0.0187)	(0.0125)	(0.0233)
A Few Months	-0.0158	0.0261*	0.0336
(Base: Many Years)	(0.0186)	(0.0126)	(0.0233)
Malawian	-0.0223	0.0127	-0.0277
(Base: Zambian)	(0.0222)	(0.0151)	(0.0277)
Tanzanian	-0.0623**	0.0187	0.00428
(Base: Zambian)	(0.0231)	(0.0154)	(0.0290)
High Fever	-0.208***	-0.0773***	0.364***
(Base: Injured Leg)	(0.0209)	(0.0172)	(0.0287)
Cough and High Fever	-0.307***	-0.113***	0.479***
(Base: Injured Leg)	(0.0219)	(0.0160)	(0.0270)
Constant	0.886***	0.127***	0.214***
	(0.0245)	(0.0191)	(0.0317)
Observations	2181	2169	1540

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A7: Malawi Wave 2: Average Component Interaction Effects

Base: Insider			
Outsider	-0.0464**	-0.0172	0.0426
	(0.0173)	(0.0256)	(0.0227)
Base: Injured Leg			
High Fever	-0.153***	-0.0752**	0.281***
	(0.0236)	(0.0282)	(0.0309)
Cough and High Fever	-0.196***	-0.176***	0.397***
	(0.0250)	(0.0258)	(0.0315)
Interactions of Origin and Symptoms			
Outsider × High Fever	0.00432	-0.0382	0.00969
	(0.0297)	(0.0336)	(0.0381)
Outsider × Cough and High Fever	0.00475	0.0293	-0.0303
	(0.0311)	(0.0311)	(0.0384)
Constant	0.901***	0.290***	0.141***
	(0.0178)	(0.0239)	(0.0228)
Observations	4345	4281	3628

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A8: Zambia: Average Component Interaction Effects

Base: Insider			
Outsider	-0.0337	0.0141	0.0282
	(0.0244)	(0.0293)	(0.0380)
Base: Injured Leg			
High Fever	-0.227***	-0.0793**	0.415***
	(0.0345)	(0.0288)	(0.0484)
Cough and High Fever	-0.272***	-0.114***	0.507***
	(0.0359)	(0.0268)	(0.0452)
Interactions of Origin and Symptoms			
Outsider × High Fever	0.0298	0.00303	-0.0794
	(0.0433)	(0.0359)	(0.0600)
Outsider × Cough and High Fever	-0.0546	0.00131	-0.0413
	(0.0452)	(0.0334)	(0.0563)
Constant	0.881***	0.128***	0.187***
	(0.0256)	(0.0259)	(0.0370)
Observations	2181	2169	1540

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.2 Outsider Identity in Malawi

In this section we test if the estimated coefficients on the outsider identities in the regressions presented in Table A2 and Table A4 are statistically equal. The test was conducted in STATA using the following command after each regression `test i2.Identity == i3.Identity` where 2 = Mmwenye and 3 = Zambian.

Table A9: Test

Outcome	Malawi Wave 1 prob > F	Malawi Wave 2 prob > F
Help	0.7597	0.4261
Move Freely	0.4509	0.7597
Has Covid–19	0.1817	0.4509

A.3 Interactions with Respondent Education[13]

Table A10: Outsider Education Interactions Malawi and Zambia

	Help	Move Freely	Has Covid
Outsider	-0.0349*	-0.0254*	0.0439*
	(0.0141)	(0.0123)	(0.0191)
Higher Education	-0.0401*	-0.0112	-0.00497
	(0.0170)	(0.0142)	(0.0221)
Outsider × Higher Education	-0.00845	0.0371*	-0.0177
	(0.0210)	(0.0171)	(0.0271)
Observations	6796	6706	5183

Standard errors in parentheses
$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[13] We also analyzed the interaction between outsider and education using a four factor education variable and did not find the results to significantly differ.

A.4 Multinomial Logistic Regression

Table A11: Main Effects Only Malawi (three level outsider)

	Dependent variable:	
	Not Sure	Yes
	(1)	(2)
60 years old	0.033 (−0.126, 0.191)	−0.030 (−0.168, 0.108)
Female	0.180^{*} (0.022, 0.339)	−0.115 (−0.252, 0.023)
Few Months	0.185^{*} (0.026, 0.343)	0.145^{*} (0.007, 0.282)
Mmwenye	0.193^{+} (−0.001, 0.387)	0.336^{***} (0.166, 0.506)
Zambian	0.073 (−0.121, 0.266)	0.222^{**} (0.053, 0.391)
High Fever	0.836^{***} (0.643, 1.029)	1.225^{***} (1.049, 1.402)
Cough and High Fever	0.946^{***} (0.747, 1.145)	1.642^{***} (1.465, 1.818)
Constant	-1.696^{***} (−1.925, −1.467)	-1.492^{***} (−1.698, −1.286)
Akaike Inf. Crit.	9,106.948	9,106.948

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A12: Main Effects Only Zambia (three level outsider)

	Dependent variable:	
	Don't Know	Yes
	(1)	(2)
60 Years Old	0.090 (−0.122, 0.302)	−0.063 (−0.277, 0.151)
Female	0.085 (−0.127, 0.297)	−0.032 (−0.245, 0.181)
A Few Months	0.220* (0.008, 0.432)	0.164 (−0.049, 0.378)
Malawian	0.034 (−0.222, 0.289)	−0.075 (−0.334, 0.184)
Tanzanian	0.041 (−0.221, 0.304)	0.044 (−0.218, 0.306)
Cough and High Fever	0.737*** (0.470, 1.003)	2.128*** (1.845, 2.411)
High Fever	0.778*** (0.531, 1.026)	1.625*** (1.345, 1.904)
Constant	−0.849*** (−1.148, −0.550)	−1.361*** (−1.686, −1.035)
Akaike Inf. Crit.	4,547.676	4,547.676

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A13: Main Effects Only Malawi (binary outsider)

	Dependent variable:	
	Not Sure	Yes
	(1)	(2)
60 years old	0.033 (−0.125, 0.192)	−0.030 (−0.168, 0.108)
Female	0.178^{*} (0.020, 0.337)	-0.116^{+} (−0.254, 0.021)
Few Months	0.185^{*} (0.027, 0.344)	0.145^{*} (0.007, 0.283)
Outsider	0.132 (−0.036, 0.300)	0.278^{***} (0.131, 0.426)
High Fever	0.836^{***} (0.644, 1.029)	1.226^{***} (1.049, 1.403)
Cough and High Fever	0.947^{***} (0.749, 1.146)	1.643^{***} (1.467, 1.820)
Constant	-1.696^{***} (−1.925, −1.467)	-1.492^{***} (−1.698, −1.287)
Akaike Inf. Crit.	9,105.297	9,105.297

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A14: Main Effects Only Zambia (binary outsider)

	Dependent variable:	
	Don't Know	Yes
	(1)	(2)
60 Years Old	0.090 (−0.122, 0.302)	−0.064 (−0.277, 0.150)
Female	0.085 (−0.127, 0.297)	−0.031 (−0.244, 0.183)
A Few Months	0.220^{*} (0.008, 0.432)	0.164 (−0.049, 0.378)
Outsider	0.037 (−0.187, 0.261)	−0.016 (−0.241, 0.208)
Cough and High Fever	0.737^{***} (0.470, 1.003)	2.131^{***} (1.848, 2.413)
High Fever	0.778^{***} (0.531, 1.026)	1.625^{***} (1.345, 1.904)
Constant	-0.849^{***} (−1.148, −0.550)	-1.362^{***} (−1.688, −1.037)
Akaike Inf. Crit.	4,544.649	4,544.649

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A15: Treatment Interactions Malawi (three level outsider)

	Dependent variable:	
	Not Sure	Yes
	(1)	(2)
60 years old	0.031 (−0.128, 0.189)	−0.033 (−0.171, 0.105)
Female	0.174* (0.015, 0.333)	−0.115 (−0.253, 0.023)
Few Months	0.183* (0.024, 0.341)	0.143* (0.005, 0.281)
Mmwenye	0.354* (0.007, 0.702)	0.588*** (0.242, 0.934)
Zambian	0.194 (−0.154, 0.542)	0.499** (0.157, 0.842)
High Fever	0.884*** (0.542, 1.226)	1.386*** (1.055, 1.717)
Cough and High Fever	1.213*** (0.865, 1.562)	1.994*** (1.664, 2.325)
Mmwenye X High Fever	−0.182 (−0.661, 0.297)	−0.181 (−0.628, 0.266)
Zambian X High Fever	0.038 (−0.434, 0.511)	−0.260 (−0.704, 0.184)
Mmwenye X Cough and High Fever	−0.342 (−0.827, 0.144)	−0.509* (−0.956, −0.063)
Zambian X Cough and High Fever	−0.478+ (−0.971, 0.014)	−0.508* (−0.950, −0.066)
Constant	−1.787*** (−2.075, −1.498)	−1.679*** (−1.965, −1.393)
Akaike Inf. Crit.	9,110.914	9,110.914

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A16: Treatment Interactions Zambia (three level outsider)

	Dependent variable:	
	Don't Know	Yes
	(1)	(2)
60 Years Old	0.102 (−0.110, 0.315)	−0.065 (−0.279, 0.150)
Female	0.083 (−0.129, 0.295)	−0.032 (−0.245, 0.182)
A Few Months	0.223* (0.011, 0.435)	0.170 (−0.044, 0.383)
Malawian	0.333 (−0.071, 0.737)	−0.170 (−0.730, 0.389)
Tanzanian	0.252 (−0.176, 0.679)	0.485+ (−0.033, 1.003)
Cough and High Fever	0.979*** (0.514, 1.445)	2.288*** (1.792, 2.783)
High Fever	1.116*** (0.678, 1.554)	1.884*** (1.391, 2.377)
Malawian X Cough and High Fever	−0.221 (−0.873, 0.432)	0.203 (−0.518, 0.924)
Tanzanian X Cough and High Fever	−0.480 (−1.142, 0.182)	−0.621+ (−1.299, 0.057)
Malawian X High Fever	−0.688* (−1.289, −0.087)	−0.079 (−0.782, 0.624)
Tanzanian X High Fever	−0.289 (−0.913, 0.335)	−0.647+ (−1.331, 0.037)
Constant	−1.034*** (−1.396, −0.673)	−1.495*** (−1.926, −1.063)
Akaike Inf. Crit.	4,549.642	4,549.642

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A17: Treatment Interactions Malawi (binary outsider)

	Dependent variable:	
	Not Sure	Yes
	(1)	(2)
60 years old	0.031 (-0.127, 0.190)	-0.032 (-0.170, 0.106)
Female	0.176^{*} (0.017, 0.335)	-0.118^{+} (-0.256, 0.020)
A Few Months	0.183^{*} (0.025, 0.342)	0.143^{*} (0.005, 0.281)
Outsider	0.272^{+} (-0.032, 0.577)	0.542^{***} (0.236, 0.848)
High Fever	0.884^{***} (0.542, 1.226)	1.386^{***} (1.055, 1.717)
Cough and High Fever	1.213^{***} (0.865, 1.562)	1.995^{***} (1.664, 2.325)
Outsider X High Fever	-0.068 (-0.482, 0.346)	-0.219 (-0.611, 0.174)
Outsider X Cough and High Fever	-0.402^{+} (-0.827, 0.022)	-0.508^{*} (-0.899, -0.116)
Constant	-1.788^{***} (-2.077, -1.500)	-1.678^{***} (-1.964, -1.392)
Akaike Inf. Crit.	9,105.007	9,105.007

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A18: Treatment Interactions Zambia (binary outsider)

	Dependent variable:	
	Don't Know	Yes
	(1)	(2)
60 Years Old	0.098 (−0.115, 0.310)	−0.059 (−0.273, 0.155)
Female	0.080 (−0.132, 0.292)	−0.034 (−0.248, 0.179)
A Few Months	0.221^{*} (0.009, 0.433)	0.165 (−0.048, 0.379)
Outsider	0.297 (−0.066, 0.659)	0.178 (−0.288, 0.643)
Cough and High Fever	0.979^{***} (0.513, 1.444)	2.288^{***} (1.793, 2.783)
High Fever	1.115^{***} (0.677, 1.554)	1.884^{***} (1.391, 2.378)
Outsider X Cough and High Fever	−0.357 (−0.926, 0.211)	−0.235 (−0.839, 0.369)
Outsider X High Fever	-0.499^{+} (−1.031, 0.032)	−0.388 (−0.987, 0.211)
Constant	-1.029^{***} (−1.391, −0.667)	-1.494^{***} (−1.925, −1.063)
Akaike Inf. Crit.	4,548.615	4,548.615

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

B Sampling

These experiments were embedded in phone surveys conducted in May 2020 (Wave 1) (n = 4,641) and in March to April 2021 (Wave 2) (n = 4,356) in Malawi and June to July 2020 in Zambia (n = 2,198). Respondents were drawn from three pools: 1) We re-contacted participants from the LGPI, a survey conducted in Malawi and Zambia in 2019 (Lust et al., 2020) who had expressed willingness to participate in future surveys and had provided phone numbers to do so. We obtained 5,100 phone numbers from Malawian respondents and approximately 4,500 phone numbers from Zambian respondents through this process; 2) In some instances, we could not contact the original respondent through the phone number(s) provided, but we found a new participant willing to take the survey and administered it to him/her; 3) As the LGPI 2019 Malawi 2019 survey did not sample from the south of Malawi, we drew in additional participants by re-visiting villages from the LGPI 2016 Malawi survey (Lust et al., 2016) and collected phone numbers for the phone survey.

For the 2020 respondents who participated in the 2016 or 2019 LGPI surveys, we have a rich set of previously collected data on both the individual respondent and their community. Both the 2016 and 2019 LGPI surveys were implemented to allow local-level indicators; the 2016 survey sample was drawn in villages, and the 2019 survey was drawn in 1 km2 areas. Both surveys were also coupled with factual and local elite surveys, which provided additional information on the nature of the community and its leadership.

B.1 Locating 2019 participants

The sample included phone numbers that had been collected from participants in the 2019 LGPI survey (see below for sampling strategy). At the end of that survey, in preparation for a panel study, we had asked individuals if they would be willing to participate in a follow-up survey.

We created a dataset that included the individual's name, telephone number(s), how long the individual had lived in the area, gender, age, and education. These questions were used to verify whether the individual answering the phone was the same respondent from 2019. Where the respondent existed but was not available, enumerators set a call-back time. Where the respondent was not available but the individual answering the phone was over 18 years of age, the individual was asked if s/he wanted to participate in the study. Where the individual was under 18 years of age and the initial respondent was not available, the enumerator asked if an adult was available. That adult was then given the chance to participate in the survey. Replacement individuals were asked at the end of the survey if they are willing to participate in future studies.

B.2 Revisiting 2016 villages

The phone numbers collected in the 2019 LGPI survey only included respondents in an area within a 50 km radius of each of the capitol cities (Lilongwe, Malawi and Lusaka, Zambia) and 100 km distance from the Malawi-Zambian border. In Malawi, in order to incorporate southern districts, we sent teams to the southern region and to two southern-central region districts where we had conducted a 2016 LGPI survey but that had not been included in the 2019 LGPI survey. At a point when the incidence of Covid in Malawi was still low, researchers were given and instructed to wear masks, use hand sanitizer, and maintain social distancing measures and were sent to the same villages that were included in the 2016 survey.

For each village, they were given lists of the first names of the adults who were in the household in 2016 and their ages (drawn from the Kish grid), and the name of the original respondent chosen. They met with the village head, who then helped them to contact and hire a person from the village. This person went to village houses to ask previous respondents if they would be willing to be contacted. The telephone numbers were collected from those who were willing. Where an individual was not willing or available, another adult in the household was asked to participate and, if s/he agreed, demographic information and the phone number was collected. If no one existed in the original household (e.g,, the family had moved or passed away) or if no one agreed to be contacted, the village contact was asked to find another household in the village willing to be contacted. Telephone numbers and demographics were entered into a database for use in the survey.

B.3 LGPI 2019 Sampling Plan: Malawi and Zambia

The sampling of the LGPI was performed independently in 4 regions in Malawi and Zambia. We name the 4 regions:

- Lilongwe
- Malawi Border
- Lusaka
- Zambia Border

As seen in Figure B.1, there are two different types of regions in the sample. There are two capital regions (Lilongwe and Lusaka) and two border regions (Malawi Border and Zambia Border). The LGPI 2019 survey utilized two sampling plans, one for the capital regions and one for the border regions. The two sampling plans are very similar with two major differences:

1. Definition of Bins
2. Selection Method for Choosing Square Kilometers

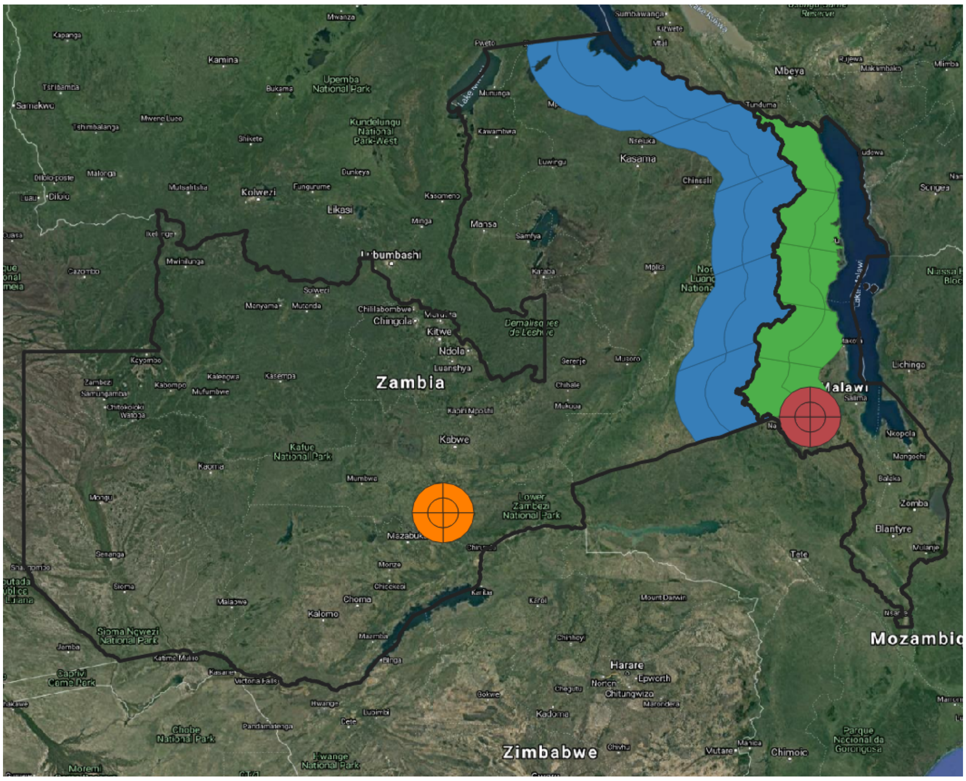


Figure B.1: Malawi and Zambia Sampling Areas

B.3.1 Overview of Sampling Plans

The sampling plans for the capital and border regions can be described as a multi-stage probability proportional to size sampling. Each sampling plan follows the same general steps:

1. Define the bins. (these are the strata)
2. Choose square kilometers within the bins using PPS. (stage 1)
3. In each selected sqkm, create a grid of hectares.
4. Remove all empty/underpopulated hectares from consideration.
5. Within each sqkm, randomly assign an order to the viable hectares using PPS (stage 2)

To see the code used to implement the sampling plan visit: `https://github.com/senickel/sampling_documentation`
and
`https://senickel.github.io/sampling_documentation/index.html#prerequisites`
The subsequent sections provide a detailed description of the sampling plan.

B.3.2 Defining the Bins

The LGPI 2019 Survey does not cover the entirety of Malawi or Zambia. Therefore we must first define where in the country we will be conducting our survey. We call these large defining areas Bins. In the capital areas there are 8 bins and in the border areas there are 10 bins.

Creating Urban Bins. When surveying in a capital region, we want to ensure sampling of areas close to and further away from the city center. Therefore, we structure the capital region bins as a "bull's eye" over the city center. The bull's eyes are centered at the following locations:

- Lilongwe: 33.783333 Longitude, -13.983333 Latitude
- Lusaka: 28.283333 Longitude, -15.416667 Latitude

Each bull's eye consists of two concentric circles: the first is 25km from the center and the second is 50 kilometers from the center. Then a vertical line and a horizontal line are drawn through the city center.

Creating Rural Bins. To ensure that we sample areas both close to and far away from the border, we begin by defining two regions. The first is 50km from the border and the second is 50-100km from the border (as geography allows).

Then to ensure that we have equal sampling in the north-south direction we divide each of the first two regions into 5 smaller regions as seen in Figure B.1.

B.3.3 Selection of Square Kilometers

The square kilometer sampling unit is used as a proxy of a community. Since boundaries of villages and communities are unclear and in some cases dynamic, we define a square kilometer area to be a community. Sampling these units will allow us to gather data from individuals living in close proximity to each other.

Capital Regions. We perform a stratified sample (stratum = bin) using PPS to select square kilometers (SQKM). First we create a grid of 1 kilometer squares over the bins and compute the expected population density for that areas using WorldPop data. We then sample a predetermined number of squares using probability proportional to size (where size = population density) sampling. The number of square kilometers sampled was equal to the desired number of communities plus a small number of supplemental units. The supplemental units serve as backups in the case that a chosen square kilometer is not sufficiently populated to obtain the necessary number of observations.

Border Regions. The border regions are notably larger than the capital regions, so traveling through the region will be intensive and expensive for the enumerators. To help mitigate this issue, we we sample 5 kilometer squares instead of sampling individual square kilometers. That is, we sample 25 square kilometers at a time. To do this, we create a grid of 5km squares in each bin, and then using PPS select an appropriate number of 5km squares.

Sampled Square Kilometer Evaluation. Once the square kilometers have been sampled, they must be reviewed to ensure they are sufficiently populated. While PPS should minimize the probably of selecting underpopulated areas, there is still a small probability they could end up in the sample. A google map image of each selected square kilometer was reviewed to ensure the area was populated. If the area was not sufficiently populated it was removed from the sample.

B.3.4 Selection of Hectares

In every verified square kilometer from the previous step, we create a grid of hectares. Each hectare is checked visually using satellite imagery to ensure it is populated. We then use PPS to randomly assign an order to the inhabited hectares within each square kilometer. This second stage of sampling helps ensure that the entire square kilometer is being sampled and thus providing a more complete picture of the community.

B.3.5 Selection of Houses Respondents

This stage of the sampling plan was implemented on the ground by teams of enumerators. Teams were sent to a specified hectare by their team leader. They were instructed to enter the hectare using tablets to track their locations and confirm they were in the correct area. They would then go to the center of the hectare and then move outward in a random walk. Once a household had been selected (and agreed to participate) a Kish grid was used to randomly choose a respondent from all reported adults (at least 18 years old) in the household.

B.3.6 Reaching Target Number of Observations

Starting from the center of a hectare, the team of enumerators would go in separate directions to additional houses using a random walk. The team continued surveying in the hectare until at least 8 surveys had been completed. Once the hectare was completed, if the target number of observations for the current square kilometer had not yet been met, the team would move onto the next listed hectare for the current square kilometers as outlined in Subsection B.3.4. The teams would continue to complete hectares until the target number of observations for a square kilometer had been met.

The target number of observations for each region were:

	Total # Obs	**# of SQKM**	**# Obs Per SQKM**	**# Obs Per Hect***
Lusaka	4500	150	30	8
Zambia Border	6000	200	30	8
Lilongwe	4500	150	30	8
Malawi Border	6000	200	30	8

*–One will note that the number of observations per hectare does not equally divide the number of observations per square kilometer. This is due to a last minute change to increase the number of observations per hectare from 5/6 to 8. This was done to increase the chances that enumerators would be able to reach the target number of observations per square kilometer.

Team Size. It is important to note that the size of the teams sent to hectares was large enough to minimize the chances that a small number of enumerators complete all of the observations in a square kilometer. By doing so, we help address the issue of enumerator effects confounding with community (sqkm) effects.

B.3.7 Special Considerations

We now describe additional sampling rules that were implemented to help mitigate issues in the field.

One–Hectare–West

Justification: We added this rule as a way to combat the issue of low population density areas.

Rule: In the event that the team of enumerators could not obtain 8 observations in a given hectare, they were allowed to move one hectare west to make up the missing observations.

Implementation Period: This rule was in place in all regions throughout the entirety of the respective fielding period.

Adjacent–Square–Kilometer

Justification: During fielding in Kenya and Zambia it become apparent there were some issues with the population density of selected and rejected areas. Namely that many selected areas were not sufficiently populated and numerous rejected areas were in fact populated. This resulted in a large number of incomplete (number of obs ¡ 30/25) square kilometers.

Rule: If after an entire square kilometer has been exhausted, the square kilometer is still not complete, then enumerators may go to any *adjacent* sqkm not in the sampling plan to obtain the remaining observations. Enumerators are instructed to obtain the observations as close to the border as possible.

Implementation Period: This rule was first implemented in Kenya and Zambia on July 27, 2019. It was implemented during the entire Malawi fielding.

Open–Square–Kilometer

Justification: During fielding in Kenya and Zambia we learned that many of the selected hectares were underpopulated while some of the rejected hectares were populated.

Rule: If after all hectares in a given sqkm have been exhausted the target number of observations for that sqkm have not been met, enumerators may go anywhere within the sqkm

Implementation Period: This rule was never used in Kenya. It was first implemented in Zambia on September 5, 2019. It was implemented during the entire Malawi fielding.

B.4 LGPI Malawi 2016 Survey

The survey was conducted in Malawi during March and April 2016. We implemented the survey using tablet computers. See here for more information on this survey.

This survey sought to measure and better understand governance and service delivery at the local level. This is a highly clustered survey, which facilitates measurement and inference at the local (in this case, village) level. The survey covers: political participation, social norms and institutions, education, health, security, welfare, corruption, land, and dispute resolution.

The sample was stratified on: region (North, Central, South); the presence of matrilineal and patrilineal ethnic groups; and the urban/rural divide. Because patrilineal groups are rare in Malawi, and we wanted to maximize variation in matrilineal and patrilineal heritage, we oversampled Primary Sampling Units (PSUs) from the patrilineal stratum. We sampled 22 PSUs, namely 'Traditional Authorities' (TAs). These 22 sampled TAs are located in 15 of Malawi's 28 districts. Districts are the largest sub-national administrative units in Malawi. Within each TA (i.e., PSU), we randomly selected four enumeration areas (EAs) as Secondary Sampling Units (SSUs). EAs are comparable to census tracts. Both PSUs and SSUs were selected without replacement, according to the principle of Probability of Selection Proportional to Measure of Size (PPMS). Within each EA, we sampled four villages, based on known geographical points provided on the maps of the EAs produced for Malawi's latest population census. Once in the village, enumerators followed a random walk pattern to select households. After they entered the household, the interviewer collected the necessary data about composition of the household. Both the contact and main questionnaires were programmed on digital tablets, including the selection of the final respondent in the household through a digital version of the Kish grid. The target was to interview 22 respondents in each village. This process produced a sample of 8,100 respondents. See Table 1 for a list of the districts and TAs included in the sample and Table 2 for a list of the villages.

While the sampling procedures were planned as presented, of course in practice this was not always the case. In total the research team had to draw 11 replacement EAs. One replacement EA was drawn because enumerators were chased out of a village and forced to withdraw from the EA. In the remaining 10 cases, EAs were not accessible (e.g. in one instance our team was unable to reach the designated EA because a bridge had washed away during heavy rains). In these instances, backup enumeration areas were randomly selected within the same EAs (excluding already selected and inaccessible zone) and were used as replacements. In total, only 11 of the 99 sampled EAs are replacement EAs. In addition, given that multiple enumerators conducted surveys in the same village, the target number of 22 respondents per village (neighborhood in urban areas) was not always precisely reached. In some instances, more were surveyed and in others slightly fewer than 22 households were surveyed. In addition, the boundaries between villages and neighborhoods were not always clear, which also caused our teams to deviate from the target of 22 per village/neighborhood.

B.5 Sampling Results

Figure B.2 shows the results of the Covid 2020 survey sampling procedure in Zambia. The dark blue dots represent LGPI 2019 respondents who also took the Covid 2020 survey. The light blue dots represent LGPI 2019 respondents we did not reach. The yellow dots represent Covid 2020 survey respondents that were replacements for the LGPI 2019 respondents. Since we do not know the exact location of replacement respondents, there is one yellow dot per district and its size is proportional to the number of replacement respondents in the district.

Figure B.3 shows the same map for Malawi.

Figure B.2: Sampling Results in Zambia.

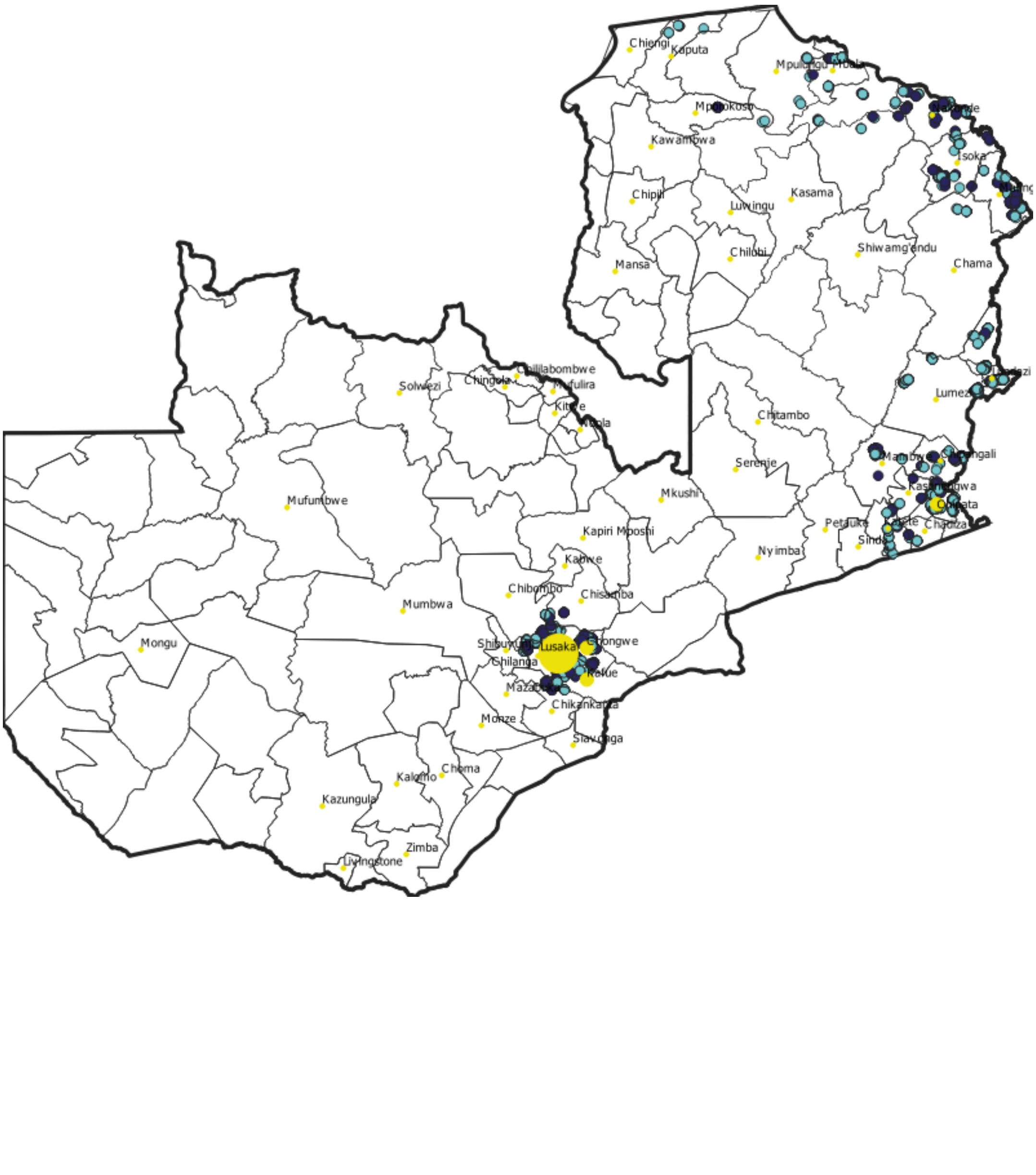


Figure B.3: Sampling Results in Malawi.

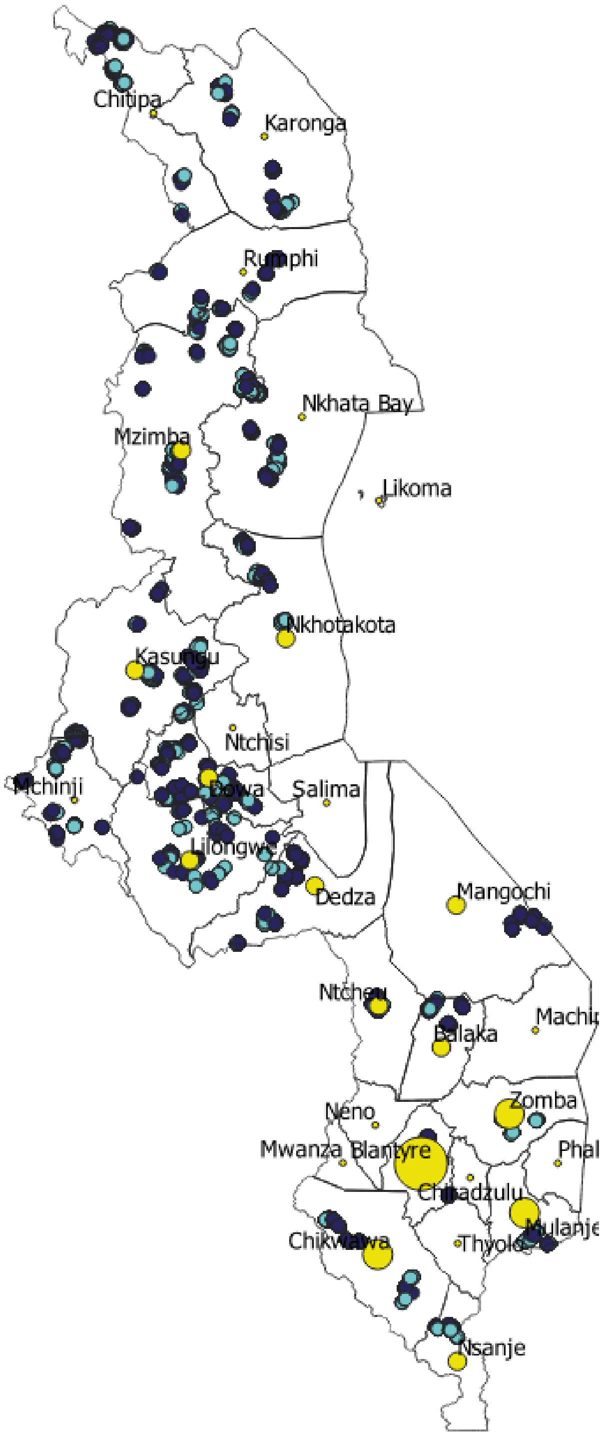
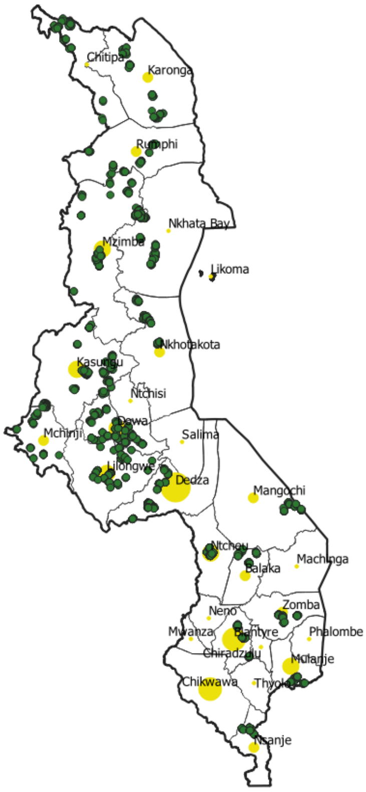


Figure B.4: Sampling Results in Malawi 2021.



B.6 Ethics and Consent

This research seeks to maximize benefit for Malawian society while minimizing risk to participants in the study. The project leaders are trained in courses on the ethical treatment of human research participants. All research activities complied with the Swedish Data Services regulations and guidelines for research ethics of the national data service regulatory body, and we ensured the survey complied with their high ethical standards. The numbers we called were of respondents from the 2019 LGPI Survey, conducted with Malawian Institutional Review Board approval. At the end of that survey, we asked respondents if they would be willing to be contacted again. We only interviewed respondents who are over 18 years of age, and for whom we have informed consent. If the respondent we reached was not the same as in the original survey, we asked the respondent's age and declined to interview if they were underage. All data collected will be kept anonymous and stored in encrypted files. We will not distribute anything with names or GPS coordinates, and all data will be retained on encrypted University of Gothenburg servers. We understand that there is always risk when handling confidential data, and we did all in our power to mitigate that risk by ensuring encrypted data storage and enforcing communication regulations. Additionally, all enumerators signed non-disclosure agreements and were subject to GDPR guidelines.

We asked for new consent from all respondents. Part of this consent section clearly explained that participation is voluntary and that the respondents have the right to seek clarification of their rights or to withdraw participation consent at any time. Each question also allowed for a do not know/refuse to answer response to mitigate discomfort for the participants. The consent scripts were as follows, changing slightly depending on the respondent type reached:

1. CONSENT (New Respondent, First Person) We are currently conducting a survey on Covid–19 and would like to talk to you today. Your answers will be confidential. They will be put together with about 3500 other people we are talking to, to get an overall picture. It will be impossible to pick you out from what you say, so please feel free to tell us what you think. This interview will take about 15 minutes. There is no penalty for refusing to participate. We would like your opinion with the knowledge that there are no right or wrong answers to these questions and that you may stop the survey at any time. Are you willing to participate in this survey, either now or at another time?

Yes, now
Yes, at another time
No

2. Hello, my name is {enumerator's name}. I am calling from the Institute of Public Opinion and Research , a research organization based in Zomba. Last year, we conducted a study looking at development and service provision and we also requested our participants to call them and discuss a few other issues this year. Do you remember to have talked to any of our survey team member between September and November last year?

Yes
No

3. CONSENT (Prior Respondent) We are currently conducting a survey on Covid–19 and would like to talk to you today. Your answers will be confidential. They will be put together with about 3500 other people we are talking to, to get an overall picture. It will be impossible to pick you out from what you say, so please feel free to tell us what you think. This interview will take about 15 minutes. There is no penalty for refusing to participate. We would like your opinion with the knowledge that there are no right or wrong answers to these questions and that you may stop the survey at any time. Are you willing to participate in this survey, either now or at another time?

Yes, now

Yes, at another time

No

4. CONSENT (New Respondent, Second Person) Hello, my name is {enumerator's name}. I am calling from the Institute of Public Opinion and Research , a research organization based in Zomba. We are currently conducting a survey on Covid–19 and would like to talk to you today. Your answers will be confidential. They will be put together with about 3500 other people we are talking to, to get an overall picture. It will be impossible to pick you out from what you say, so please feel free to tell us what you think. This interview will take about 15 minutes. There is no penalty for refusing to participate. We would like your opinion with the knowledge that there are no right or wrong answers to these questions and that you may stop the survey at any time. Are you willing to participate in this survey, either now or at another time?

Yes, now

Yes, at another time

No