# Malaria and Chinese Economic Activities in Africa

**Matteo Cervellati** (University of Bologna, CEPR and IZA)
**Elena Esposito** (HEC and University of Lausanne)
**Uwe Sunde** (LMU Munich, CEPR and IZA)
**Song Yuan** (University of Warwick)

Discussion Paper No. 293

November 11, 2021

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Georg Weizsäcker, Ph.D., Humboldt University Berlin, 10117 Berlin, Germany
info@rationality-and-competition.de

# Malaria and Chinese Economic Activities in Africa*

Matteo Cervellati† — University of Bologna, CEPR, IZA

Elena Esposito‡ — HEC, University of Lausanne

Uwe Sunde§ — LMU Munich, CEPR, IZA

Song Yuan¶ — University of Warwick

**Abstract**

We present novel evidence for the influence of malaria exposure on the geographic location of Chinese economic activities in Africa. The hypothesis is based on the observation that many Chinese aid projects and infrastructure contractors rely on Chinese personnel. High malaria exposure might constitute an important impediment to their employment and productivity. Combining data on Chinese aid and construction projects with geo-localized information about the presence of individuals from internet posts reveals a lower density of Chinese activities and of Chinese workers in areas with a high malaria exposure. This effect is mitigated partly through heterogeneity across sectors and immunity of the local population, through the selection of Chinese workers from regions in China with historically high malaria risk, and through the availability of malaria treatment.

JEL-classification: F2, F6, J2, J6

Keywords: Infrastructure Projects, Malaria, Disease Prevalence, Immunity, Weibo

*The authors are grateful to Axel Dreher, Kevin Donovan, Roland Hodler, Jonas Hjort, Omer Moav, Sharun Mukand and especially James Fenske, as well as seminar participants at the German Institute of Global and Area Studies, Hamburg, University of Warwick, University of Lausanne, Jinan University, Sichuan University, and the audiences at the CSAE Annual Conference (Oxford), the Royal Economic Society Annual Conference, The European Winter Meeting of the Econometric Society (Nottingham) and the Chinese Young Economist Society (Guangzhou) for helpful comments. The authors would like to thank Jianghao Wang for sharing the Weibo Data and Andrea Tesei for sharing the 2G signal coverage data. Uwe Sunde gratefully acknowledges financial support by Deutsche Forschungsgemeinschaft through CRC TRR 190 (project number 280092119).

†Department of Economics, University of Bologna. Email: m.cervellati@unibo.it

‡Department of Business and Economics, University of Lausanne. Email: elena.esposito.1@unil.ch

§Department of Economics, LMU Munich. Email: Uwe.Sunde@econ.lmu.de

¶Department of Economics, University of Warwick. Email: s.yuan.2@warwick.ac.uk

# 1 Introduction

Despite decades of international development aid, Africa remains the economically least developed continent. The need for improvement of infrastructure has been identified as a key necessity for poverty reduction and economic development in Africa (World Bank, 2017). Over the past decade, African economies indeed saw a large scale-up of foreign investments, with infrastructure commitments exceeding US$100 billion per year for the first time in 2018 (African Development Bank, 2018). Empirical evidence suggests that these infrastructure projects had a considerable impact on economic development over the past two decades, particularly for the poorest countries (see, e.g., Kodongo and Ojah, 2016; African Development Bank, 2020).

China plays an increasingly important role in this context. Different from Western donors and investors, the focus of China's foreign engagement lies on the financing and realization of large infrastructures, e.g., as part of the Belt and Road Initiative. China has become the biggest contributor and contractor in Africa, with Chinese infrastructure financing in Africa exceeding US$25 billion in 2018, which accounts for about a quarter of total infrastructure funding (ICA, 2018). This has spurred the discussion about the factors behind this "New Scramble for Africa".[1]

A specific feature of Chinese activities in Africa is the twin role of China as donor and of Chinese firms as private contractors of infrastructure and aid projects irrespective of the source of funding. For instance, in 2013, total infrastructure financing in Africa from all sources summed to US$ 99.6 billion, of which US$ 13.4 were from China. The value of infrastructure projects contracted by Chinese companies in Africa summed to $37.2 billion. A key element of the success of Chinese firms in this context is their reliance on Chinese personnel. In addition to materials and equipment, Chinese construction contractors typically source a large fraction of their work force, particularly skilled and managerial workers, directly from China (Chen, Goldstein and Orr, 2009). In spite of moderate profit margins in infrastructure building, this makes Chinese firms highly competitive (Corkin and Burke, 2006) and the market share of Chinese companies among infrastructure construction projects in Africa rose from below 10% in 2002 to nearly 62% in 2019 (ENR, 2019), while Africa is China's main overseas construction market (Huang and Chen, 2016). Importantly, Chinese firms realize projects financed by all donors and governments. For instance, they covered more than 20% of the value of all World Bank transportation infrastructure projects in 2018. This rising involvement in infrastructure projects went along with a massive inflow of Chinese workers in Africa. Despite caveats related to data quality and reliability, the number of Chinese migrant workers in infrastructure projects, mining and other Chinese activities in Africa is estimated to be over a million (Dollar, 2016).

[1] See, e.g., The Economist (2019), https://www.economist.com/leaders/2019/03/07/the-new-scramble-for-africa.

Given the impact on local development perspectives, a better understanding of the drivers of the spatial location of Chinese foreign activities in terms of projects and personnel is of primary importance. Existing work surveyed below has focused on the role of economic stability, political institutions, and natural resources for Chinese aid and infrastructure financing, mostly at the country level and with mixed findings. The implications of the presence of Chinese personnel for the spatial location and intensity of Chinese activities in Africa and the role of factors that primarily affect the presence of Chinese personnel, such as the local disease exposure, have not been investigated so far.

In this paper, we investigate whether and to what extent the exposure to malaria affects the location of Chinese foreign activities in Africa, in terms of aid projects, infrastructure projects, and the presence of Chinese migrant workers. The investigation is based on the hypothesis that Chinese economic activities heavily rely on personnel from China, while malaria infections involve serious health risks for Chinese workers. Malaria is exclusively transmitted by anopheles mosquitoes and the exposure depends on local environmental and climatic factors that favor the transmission of the pathogen. Malaria has plagued African regions since thousands of years, which has led to the emergence of protective genetic and acquired immunities among the local population. In the absence of an effective vaccine, malaria exposure poses a serious immanent threat for Chinese workers in Africa who have not developed such protective immunities and who typically work on their projects outside of buildings. Differently from other diseases, like HIV, which can be prevented with simple behavioral strategies, malaria imposes considerable immanent costs for the deployment of Chinese workers, which manifest in lost working days and costs for treatments and medication.

The analysis leverages existing information about the location of Chinese aid projects, and novel data on infrastructure projects contracted by Chinese firms that include information about the number and value of projects, the exact geo-referenced location of infrastructure projects, the sector and the contracting companies. We also use a new geo-referenced data set on the presence of Chinese migrant workers across Africa, for which official statistics do not exist. To this end, we construct geo-localized data on the presence of Chinese individuals in Africa at the individual level by exploiting information from micro-blog posts on Weibo, the Chinese equivalent to Twitter, in the year 2014. The blog posts provide geo-referenced information about the location of Chinese individuals in Africa. This is combined with anomymized information about the Chinese users, particularly their region of origin in China, and an application of text analysis to the blog-posts is used to classify individuals by their working status and sector using a supervised cloud of words approach. The data on projects and Chinese workers is complemented by a large set of geo-referenced information on the characteristics of the destination locations in Africa, of the regions of origin in China, and by individual characteristics of Weibo users.

The empirical analysis makes use of information about malaria exposure in terms of malaria transmission risk, which is based on the epidemiology of malaria and provides a source of plausibly exogenous variation at the geographic grid cell-level. Malaria is an infectious disease with severe symptoms. The disease is caused by various strains of *Plasmodium* parasites, which are transmitted exclusively through female anopheles mosquitos. The reproduction cycle of *Plasmodium* parasites involves various stages of development both in the vector and in the human host, so that the exposure to, and incidence of, malaria is closely related to environmental factors, in particular climatic conditions. Malaria can be fatal and malaria infections caused by *Plasmodium falciparum* are the leading cause of disease-related death in Africa. WHO reports suggest that globally malaria caused more than half a million deaths per year in 2010, and still caused more than 400,000 deaths in 2018; the economic losses related to malaria are substantial and amount to more than US$12 bn per year for Africa over the past decades (WHO, 2017, 2019). The risk of infection is reduced by immunity, which can either be due to genetic mutations that prevent the acute infection and completion of the reproduction cycle, or acquired through earlier infections. Persistent exposure to malaria promotes (genetic or acquired) immunity among the local population. Conversely, high malaria exposure entails high infection risk and frequent infections among individuals that do not exhibit resistance due to immunity. Likewise, epidemics are likely to occur in places where the population has low or no resistance to malaria and the sudden occurrence of particularly suitable weather conditions creates a favorable environment for the disease.

The empirical strategy applies these insights by making use of various indices of malaria exposure that are based on exogenous variation in biological and climatic conditions. The empirical analysis is executed at the level of $1^\circ \times 1^\circ$ grid cells for all Africa. This setting complements earlier work on Chinese infrastructure and aid projects, which used existing official data at the country level or, more recently, at the level of sub-national administrative regions (see the literature review below). The fine spatial resolution of the analysis that exploits geo-referenced information about the location of Chinese foreign activities enables an identification strategy that accounts for the typical confounders for an effect of malaria exposure on Chinese activities. In particular, the fine resolution of the data allows us to not only control for country fixed effects, but to account for systematic variation in overall economic activity and institutional quality by including fixed effects for sub-national administrative regions, for cultural groups to account for historical and cultural factors, and by controlling for systematic variation in economic activity and geographic features across individual cells. Moreover, the use of geo-coded web post data from the Chinese equivalent of Twitter, Weibo, can be used to investigate the role of malaria exposure for the local presence of Chinese workers across different regions in Africa, and the link between the distribution of Chinese infrastructure construction projects and the distribution of Chinese workers. This allows us to make progress on the determinants of the presence of Chinese personnel and its relation to Chinese aid and infrastructure projects, which have been largely neglected in existing research.

The analysis also exploits exogenous variation in the prevalence of genetic immunity to malaria that exclusively emerged in response to persistent exposure to the pathogen in the past. Information on the prevalence of genetic mutations that are malaria protective and whose spread is related to past local exposure to the disease is used to further isolate the specific role of malaria and the existence of possible attenuation mechanisms. This analysis makes use of the prevalence of immunities among the African population, and also among the Chinese population in areas that were affected by *Plasmodium falciparum* in the past. This analysis is based on the hypothesis that the adverse effects of malaria might be attenuated by the prevalence of resistance to high malaria exposure of the local African workforce. This resistance affects the profitability of projects and is related to the skill intensity of the respective Chinese projects, which might differ across sectors. Alternatively, the effect is expected to be attenuated by relying more on Chinese personnel from regions in China that exhibit greater resistance in terms of prevalence of acquired or genetic immunity. By exploiting information about the home region of Chinese workers, analysis at the level of individual workers allows us to condition on individual characteristics and include fixed effects for region of origin and destination region in Africa. The identification then exploits variation within locations and across individuals, and accounts for confounding factors at the cell level. Additional analysis accounts for other diseases for which the channel under investigation should not hold, like HIV and trypanosomiasis, as placebos.

The empirical analysis produces novel evidence regarding the effect of malaria exposure on the spatial distribution of Chinese activities in Africa, on the mechanisms that are related to the deployment of Chinese migrant workers, and on the existence of attenuation mechanisms. The results show, first, that higher malaria exposure significantly reduces the local prevalence of Chinese aid projects and of infrastructure projects contracted by Chinese firms. The effect is heterogeneous across sectors. Second, the results reveal a strong association between the presence of Chinese individuals in a cell and the number of aid and infrastructure projects in that cell. Third, the results provide evidence for a negative effect of the exposure to malaria on the presence of Chinese workers. These findings are robust to alternative specifications and a rich set of control variables that rules out potential confounders. In addition, the finding of a negative effect of malaria exposure to the presence of Chinese workers also holds when conditioning on Chinese aid or infrastructure projects, which provides further support for the influence of malaria exposure to Chinese activities and suggests mechanisms of adaptation and attenuation. Additional analysis suggests that the effect is weaker in more skill-intensive sectors. Moreover, Chinese workers from regions in China with historically greater malaria exposure and with more prevalent genetic immunity against malaria are relatively more likely to be deployed in cells in Africa with a higher malaria risk. Finally, the findings suggest that the negative effect of malaria exposure on the presence of Chinese workers is mediated by the local availability of anti-malaria health policies.

**Contribution to the Literature.** The analysis complements a large and rapidly growing body of research that documents the relevance of aid financing, particularly related to infrastructure projects, for African development, and the increasingly important role of China in this context. This literature has found a positive, although often small, effect of development aid on economic growth (see, e.g., Busse, Erdogan and Mühlen, 2016; Galiani et al., 2017; Dreher and Langlotz, 2020), and the impact of aid on political stability and conflict is subject to an ongoing debate (see, e.g. Crost, Felter and Johnston, 2014; Nunn and Qian, 2014; Bluhm et al., 2021). Besides the central role for development of Chinese outward direct investment (ODI) in Africa, it has changed traditional patterns of economic influence (Brautigam, 2009; Donou-Adonsou and Lim, 2018). Recent work has investigated the determinants and consequences of Chinese aid finance and infrastructure investments. Empirical evidence at the country level suggests that infrastructure projects favored growth over the last two decades, particularly for the poorest countries (see, e.g., Kodongo and Ojah, 2016; African Development Bank, 2020). The disbursement of Chinese aid and development finance implies better development outcomes and boosts short-term growth at the level of countries (Dreher et al., 2021*b*). More recent work has also investigated this issue using data at subnational administrative regions (Bluhm et al., 2020; Dreher et al., 2021*a*). We add to this literature by contributing an analysis on Chinese activities in different dimensions of infrastructure projects and on the spatial distribution of Chinese workers at the cell level.

By focusing on the spatial determinants of Chinese aid and infrastructure projects, our study contributes to a growing literature on the role of natural resources, and of institutional and politico-economic mechanisms as drivers of the intensity and location of Chinese activities in Africa. Access to mineral resources is often seen as the main motivation for Chinese aid and infrastructure projects, or for reserves-backed lending at low cost (see, e.g., Habiyaremye, 2013, for a discussion of the "Angola Mode"). While some evidence suggests that the location of Chinese aid is not driven more by political factors or natural resources than that by other donors (Dreher and Fuchs, 2015), other work shows that Chinese official development assistance (ODA) is driven more by foreign policy considerations, whereas contracting and other forms of official financing are driven by resources and economic interests (Chen, Dollar and Tang, 2018; Dreher et al., 2018). Cheung et al. (2012) document that besides motives related to market access and investment risk, Chinese activities are concentrated in areas where there are opportunities for resource exploitation. Existing evidence also points at an important role of corruption and favoritism for the location of Chinese activities in Africa (Dreher et al., 2019; Anaxagorou, Efthyvoulou and Sarantides, 2020). By documenting a role of malaria exposure, the analysis below contributes a distinct channel through which local geo-climatological conditions affect the location of Chinese activities in terms of projects and of the presence of Chinese personnel. This channel is related to employment patterns and contributes to a better understanding of the determinants of the presence of foreign economic activities in Africa and the interactions of these activities with the local workforce. At the same time, as consequence of the identification strategy, the mechanism is orthogonal to other mechanisms that have been documented at the level of countries or sub-national regions.

With the focus on the role of malaria exposure for the location of Chinese activities in Africa, our paper relates to the literature on the implications of geography for economic development (see, e.g., Henderson, Shalizi and Venables, 2001; Desmet and Henderson, 2015; Desmet, Nagy and Rossi-Hansberg, 2018; Cervellati et al., 2018). The empirical approach follows a recent strand of papers that apply a similar identification strategy to obtain credible identification while relying on time-invariant, cross-sectional variation (including work by Gennaioli et al., 2013; Alsan, 2015; Alesina, Michalopoulos and Papaioannou, 2016; Galor and Özak, 2016; Michalopoulos, Naghavi and Prarolo, 2018; Enke, 2019; Arbatlı et al., 2020; Mayshar et al., 2021, among others). In comparison to work along this line of research, we adopt a very restrictive empirical specification that controls for sub-national administrative regions and cultural group fixed effects and thereby accounts for unobserved variation in institutional quality and economic development across regions and cultural groups. Moreover, by relying on variation in the exposure to malaria, our approach relates to work on the consequences of the disease environment for development. This work has shown that disease exposure affects settlement strategies, the development of colonial and post-colonial institutions, and ultimately long-run development (Acemoglu, Johnson and Robinson, 2001). Other work has pointed at an important role of disease exposure that mainly affected life stock, providing evidence for a channel related to pre-colonial political centralization and the origins of institutions (Alsan, 2015). By focusing on Chinese aid and infrastructure activities and applying an identification strategy that accounts for these channels, our results provide evidence for a distinct and novel channel through which malaria affects development today and which can be addressed by policy interventions.

This paper also relates to recent work on the potentially detrimental implications of exposure to epidemic disease for development and that has delivered mixed findings. Micro-evidence suggest that early exposure to diseases shapes economic outcomes of adults (Bleakley, 2010; Cutler et al., 2010; Hong, 2011; Lucas, 2010). Other work has considered the impact of diseases across countries and found negative correlational effects on health and human capital and, indirectly, on growth and foreign investments (Azemar and Desbordes, 2009; Asiedu, Jin and Kanyama, 2015; Ghosh and Renna, 2015). Moreover, the direct causal impact of health at the macroeconomic level has been a matter of debate (Acemoglu and Johnson, 2007; Weil, 2007; Cervellati and Sunde, 2011). Our work also contributes to a fast-growing literature on the impacts of malaria and its eradication for development. The role played by malaria is intensely debated. Early evidence of a detrimental effect on growth across countries (see, e.g. Gallup and Sachs, 2001; Sachs and Malaney, 2002) has been subsequently put into question. Investigations of the role of differential exposure to malaria within African locations that use disaggregated data at the level of grid cells find no effect of malaria on historical development patterns and pre-colonial institutions (Depetris-Chauvin and Weil, 2018; Alsan, 2015), and a moderate, possibly non-linear, effect on modern development (Cervellati, Esposito and Sunde, 2017). Related work considered the impact of malaria epidemics on infant mortality (Kudamatsu, Persson and Strömberg, 2012) and outbreaks of violence (Cervellati, Esposito and Sunde, 2021).

Our evidence also contributes to the literature on the determinants of Chinese work migration to Africa (Cook et al., 2016) and, more generally, to the literature on selective migration patterns, which has mostly focused on skills rather than health and immunity (see, e.g. Grogger and Hanson, 2011; Moraga, 2013; Bartolucci, Villosio and Wagner, 2018). In contrast, by focusing on Africa and documenting the role of malaria for the selection and location of Chinese work migrants to Africa, we provide evidence for a pull factor that, to our knowledge, has received little attention so far.

The remainder of the paper is structured as follows. Section 2 provides a discussion of the conceptual background and describes the data and their sources. Section 3 presents the main results for the location of infrastructure projects and Chinese workers in Africa, respectively. Section 4 explores the mechanisms underlying the main results, and Section 5 concludes.

# 2 Background and Data

## 2.1 China's Economic Activities in Africa

**China's Engagement in Africa.** China has been playing an increasingly important role in African development. Chinese economic activity in Africa includes FDI, trade, construction projects, financing support, and migrant workers. The FDI inflow from China rose from US$0.08 billion to US$2.98 billion between 1997 and 2015, which accounts for more than 5 percent of FDI to Africa in 2015 (UNCTAD, 2016). China's FDI stock increased from US$ 0.25 billion to US$39.9 billion during the same period.[2] Compared to FDI, infrastructure projects and development financing from China is even larger in scope and size. Construction projects (including infrastructure construction) undertaken by China reached US$51.5 billion in 2015. Chinese infrastructure financing commitments in African countries rose from less than US$1 billion in 2001 to US$20.9 billion in 2015, which accounted for 25 percent of the total infrastructure funding in Africa (ICA, 2015). The official number of Chinese workers in Africa by the end of 2015 was 263,700 according to National Bureau of Statistics in China, while estimates of the total stock of Chinese workers, including those who stay after the end of the official engagement exceed one million (Dollar, 2016).

Infrastructure investment from China is usually in the form of preferential and commercial loans from the China EXIM Bank and China Development Bank, which are intertwined with trade deals and resource extraction.[3] Loans are used for constructing infrastructure. In contrast to traditional loans from Western countries, Chinese infrastructure financing is disbursed without any policy conditions attached, which is critical from the perspective of many African governments. However, there is usually a requirement that the construction work be contracted to a Chinese company. A significant proportion of equipment and materials must be imported from China, and most importantly Chinese construction companies mainly employ Chinese nationals in managerial positions and employ a large number of Chinese workers, with typically half of the workforce and more than 90% of the skilled workforce positions being Chinese nationals (Chen, Goldstein and Orr, 2009). In light of the relatively low cost of Chinese workers and the lack of a reliable supply of locally-sourced skilled workers this leads to a high labor productivity, which is also ensured by temporary deployments with high work intensity (typically shift work with low rates of absenteeism), a high skill level thanks to intensive training in

[2] See http://data.stats.gov.cn/index.htm.

[3] This business model is also known as "Angola Mode" or "Resources for Infrastructures" in the literature (Jansson, 2013; Nissanke and Söderberg, 2011). In this mode, an African government concludes official loan agreements with Chinese government with natural resources as the credit guarantee. In return for the building of infrastructure, the African governments use the revenues from exporting natural resources to pay back the loans or grant Chinese companies the rights to exploit natural resources. Trade deals and FDI from China are also usually added to the agreements.

China prior to deployment, an attitude for hard work, and less language-related communication frictions, which results in Chinese firms being very competitive (see, e.g., Zhao and Shen, 2008; Chen, Goldstein and Orr, 2009; Huang and Chen, 2016). From the perspective of the Chinese government, this approach can produce win-win outcomes for both parties. On the one hand, it allows African countries to relieve critical financing bottlenecks for infrastructure projects that are crucial for local development; on the other hand, it allows China to secure a stable supply of resources and enlarge its influence in Africa via various forms of economic engagement, including the development of a construction sector operating abroad.

**Employment Patterns.** Information about the location of Chinese infrastructure projects provides a unique opportunity to investigate the determinants of the destination of Chinese investments in Africa. Typically, Chinese companies bring laborers from China to Africa for their infrastructure and investment projects in Africa, at least at the beginning. Chinese workers are familiar with the processes and organization of their company and possess the skills and experience to install structures and to train local workers. There are complaints that even low-skilled jobs involved in Chinese projects are filled by Chinese nationals, but infrastructure projects employ large shares of local workers. According to a report published by the Angolan Ministry of Finance in 2007, out of a total of 3136 employees in Chinese companies, 1872 (almost 60%) were Angolans (Tang, 2016). A survey of 49 firms in South Africa shows that the share of local workers is as high as 80 percent (Huang and Ren, 2013). In Kenya, it is estimated that more than 85 percent of employees in Chinese firms in the services and manufacturing sectors in 2013 were locals (Sanghi and Johnson, 2016).

The employment of locals in Chinese firms varies across sectors. Chinese companies in the labour-intensive sectors such as mining, construction and manufacturing rely more on local workers. In sectors with advanced technology and management requirements such as hydropower and telecommunication, Chinese companies employ larger shares of Chinese workers, who make up more than 60 percent of the workforce at the African sites (Tang, 2016, 2010). Across all sectors, however, there is a mix of Chinese and local personnel, and a greater immunity of local personnel reduces the risk of disruptions in the project execution and thus likely attenuates the effect of malaria exposure on Chinese workers. The employment composition also changes with the development phase of firms. A survey in the DRC by Tang (2016) found that in the first five years, Chinese companies use Chinese workers for almost one-third of positions. After five years, this proportion declines to 15 percent or less. This trend is especially common in sectors such as manufacturing.

There is considerable variation in the share of Chinese workers in construction as well. Chen, Goldstein and Orr (2009) find that Chinese construction firms employ equal numbers of Chinese and locals in unskilled occupations, but less than 10 percent of skilled workforce positions are occupied by locals. Data from Deborah Brautigam's blog "China in Africa: The Real Story" shows that the proportion of Chinese workers is very high in the construction of large public buildings, which requires professional management and techniques.[4] For example, Chinese workers represented more than 70 percent of employees in a stadium project in Angola. In contrast, in road, harbour and mine projects Chinese workers represent less than 10 percent.

**Malaria and the Presence of Chinese Workers.** The risk of malaria infections has gained increasing attention among Chinese workers and firms in recent years. Newspaper reports and blog posts have raised awareness about malaria risk in Africa.[5] There have been blog posts warning Chinese expatriates of the prevalence of malaria and giving advice how to prevent infection.[6] In a recent case, a Chinese court sentenced an employer to pay a compensation of 100,000 yuan (which corresponded to approximately US$ 16,000 in 2014) to an employee who was diagnosed with malaria *falciparum* upon returning to China from a work-related deployment in Africa. The reasons given for the judgement state that "arid countries such as Africa often have many outbreaks of malignant infectious diseases that are not common in China. Because our citizens generally lack natural immunity to such diseases, they are easily infected and even killed. [...] It is necessary to pay full attention to the prevention and control of severe infectious diseases in the tropics, timely injection of vaccines, special attention to maintaining hygiene and good living habits, and symptoms such as fever and chills must be treated promptly. At the same time, overseas labor companies and organizers should also have corresponding awareness of prevention and control, timely organize labor services to carry out infectious disease prevention and control work. Migrant workers cannot be allowed to have inadequate disease prevention and control while living in the tropics according to Chinese lifestyle. It is best to organize the injection of vaccines in advance so as to avoid the occurrence of huge medical expenses compensation after the workers are infected".[7] This suggests that malaria exposure might influence Chinese economic activities in Africa, particularly when activities require Chinese

[4]See https://csisprosper.com/2015/06/17/chinese-investment-africa-jobs/ and http://www.chinaafricarealstory.com/.

[5]Examples can be found at http://china-africa-jinghao.blogspot.com/2012/02/impressive-photos-about-chinese-in.html, or http://china-africa-jinghao.blogspot.com/2011/08/chinese-in-takoradi-1-building-boat.html.

[6]For instance, a "letter to travellers in high-risk areas such as Africa and South-East Asia" at http://www.jshzcdc.com/html/news/yuqing/2015-02-11/3242.html warns that "In recent years, with the increasing number of people leaving the country, especially labor exporters, many people who have gone to areas abroad with high malaria risk have been infected with malaria, which has not only affected work and life, but also some malaria patients have not been diagnosed and treated in time. The phenomenon of death has caused significant losses to the personnel of the country and their families. For your health, please understand the following knowledge: First, the epidemic range and harm of malaria – malaria is an infectious disease and parasitic disease caused by malaria parasites and transmitted by mosquito bites. (...)" (own translation).

[7]See http://rmfyb.chinacourt.org/paper/html/2017-06/09/content_126372.htm?div=-1, own translation.

workers to be on site.

**Data on Infrastructure Projects.** Our analysis draws on two data sources for Chinese infrastructure activities in Africa. The first data set is the "AidData's Geocoded Global Chinese Official Finance" (Dreher et al., 2019, Version 1.1.1). While our analysis is mostly cross-sectional in nature, we combine information on aid projects in the period 2010 to 2014 to have a sufficiently large number of aid projects for estimation at the cell level while maintaining comparability with the second data set on infrastructure projects, which is only available for 2013. From 2010 to 2014, the data contains information for 1296 Chinese aid projects in Africa in total. Of these, 1048 are projects that correspond to the criteria for official development assistance (ODA-like projects), 109 are other official flows that do not meet the official development assistance criteria (OOF-like projects), and 139 are Chinese official finance projects (Chinese OF).[8] Among all projects, the data comprises information on the exact locations (precision code 1) for 555 projects, 99 projects are up to 25 km away from an exact location (precision code 2), 141 are in a second-order administrative division (ADM2) (precision code 3), 143 are in first-order administrative division (ADM1) corresponding to a province (precision code 4), 14 are more than 25 km away from a specific location (precision code 5), 108 are independent of political entity (precision code 6), and 236 are projects that stipulate that the central government will be the only direct beneficiary (precision code 8). Since our analysis of projects is at the cell level, we can only include projects that can be unambiguously located within a cell, which leaves us with 952 aid projects in the sample.

We categorize the projects into infrastructure and non-infrastructure aid projects using a categorization compatible with the second data set on infrastructure construction project data described below. In particular, we classify 576 projects that are devoted to infrastructure construction such as roads, dams, or public buildings, as infrastructure aid projects. Alternatively, we classify infrastructure aid projects and aid projects as projects involving Chinese workers based on information about "free-standing technical assistance". ODA-like projects are aid projects aimed at development as defined by AidData. Finally, donations are projects in the form of monetary and material donations, scholarships, and training in China. These projects do not necessarily involve Chinese workers in Africa, but around half of the donations projects have no clear geo-localization below the country level and are therefore excluded from the sample. Hence, this category is likely to be measured with considerable error. This classification of aid projects is not fully exclusive.

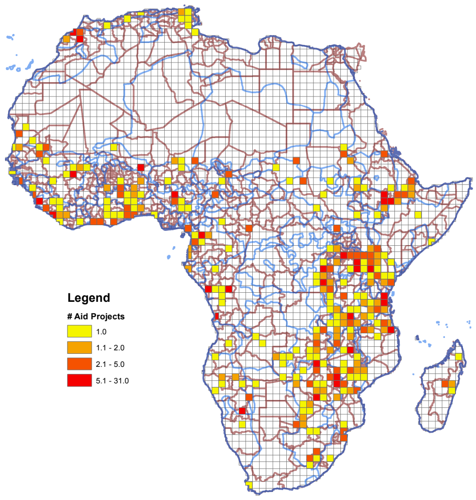
The second data set on infrastructure construction projects comes from the *China Overseas*

[8]Chinese official finance projects (Chinese OF) include concessional and non-concessional funding from Chinese government institutions (including central, state or local government institutions) with development, commercial or representational intent.

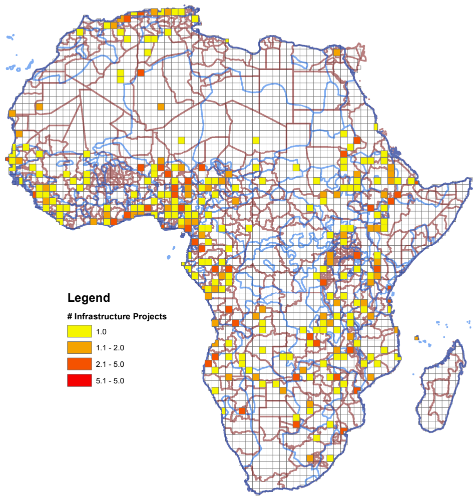
*Engineering Projects and Labour Service Yearbook 2014*. The Yearbook records all engineering construction projects with value of at least US$ 5 million undertaken by Chinese companies in Africa in 2013. It includes the name, value, host country and company name for each project. Table D1 in the Supplementary Appendix provides detailed information about the data source and the different variables constructed from this source. We obtained the geographic coordinates for each project by searching for the project information on search engines such as Google and Baidu to determine exact locations. With the location name, we identified the coordinates using Google Maps. The project name is self-explanatory regarding the sector. As the finance from China predominantly goes to infrastructure construction, we focus our analysis on the projects belonging to the four infrastructure sectors – transport, water supply, energy and telecommunication – as defined by the Infrastructure Consortium for Africa, and on an additional sector of social infrastructure – public buildings. Projects are classified as hydropower, public buildings, transportation or utility. Hydropower includes the construction of hydroelectric power plants, and dam and river projects for water supply, which are usually constructed together with hydropower plants. Public buildings include stadiums, churches, but also schools, hospitals, or other buildings for public use. Transportation encompasses roads, railways, harbors, airports and bridges. Utilities are power and gas supply, water supply and sewage. Telecommunication only represents a small number among all projects but is closely related to citizens' daily lives.

To compute our primary variables of interest, we geo-locate all the projects and then aggregate the contracted value and number of projects in $1° \times 1°$ grid cells. The geographic distribution of aid projects and infrastructure projects, as well as the number of the respective projects per cell are displayed in Figures 1a and 1b. The summary statistics are contained in Table C1 in the Supplementary Appendix.

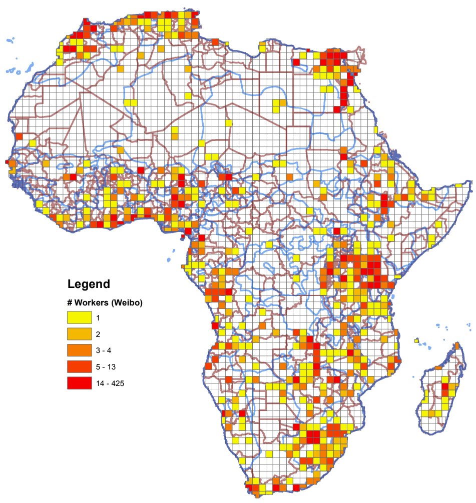
There are significant differences between the projects in the aid data set and the infrastructure projects in the yearbook data. The infrastructure aid projects are projects financed by the Chinese government through loans or grants. In contrast, the infrastructure projects listed in the yearbook consist of all infrastructure construction projects with contract value above $5 million that are contracted by Chinese companies in Africa in 2013. The financing of these projects stems from various sources such as local African governments, international development banks, and international aid donors (including China) or private parties. Hence, infrastructure aid projects are directly associated with Chinese financing, whereas the infrastructure projects are primarily associated with Chinese contracting. To illustrate the differences, the total value of 756 infrastructure projects contracted by Chinese companies in Africa in 2013 summed to $37.2 billion. In comparison, the total number of infrastructure aid projects from




(a) Aid Projects




(b) Infrastructure Projects




(c) Weibo Users (Workers)

Figure 1: Location of Chinese Activity in Africa

*Note*. Panel (a): the shades represent the number of aid projects in a cell. Panel (b): the shades represent the number of infrastructure projects in a cell. Panel (c): the shades represent the number of individuals that are identified as workers on the basis of their Weibo posts in a cell. All Panels: Blue polygons represent ethnic cultural groups, brown polygons administrative regions. See Section 2.1 for details on data sources and construction.

2010 to 2014 is 586, and the total value is $42.1 billion. In 2013, the value of 91 infrastructure aid projects amounted to $10.5 billion.

**Data on Individual Workers: Weibo Posts.** To find out where Chinese workers are located in Africa, we take location information from posts on the Chinese equivalent of Twitter, Weibo. The Weibo data comes from Weibo.com, SINA Corporation, a Chinese microblogging website. Weibo is one of the largest social media and network platforms in China. By July 2018, Weibo had over 431 million monthly active users.[9]

Our dataset includes all the geo-tagged Weibo posts of individuals located in Africa between February and December 2014, which corresponds to about 94,000 messages written by around 9,000 individuals. Each post has detailed information on the user id, the date of the post, geographic coordinates (latitude and longitude) of the post, post text, the location in China where the user was registered, gender, user level, number of followers, number of friends, number of replies, number of comments, and number of likes. Using Weibo data provides us with a unique source of information that allows tracing of the disaggregated geographic distribution of Chinese workers in Africa.

For the empirical analysis, the variable of main interest is the number of Weibo users who are identified as workers in different sectors in a particular grid cell. We apply supervised text analysis to search the content of each message in order to elicit whether the individual is located in Africa for work or for other reasons (e.g., tourism), and to identify the sector of employment of the individual. To encode the occupations of users, we construct a list of more than 2,900 keywords which are closely related to industry sectors classified by International Standard Industrial Classification of All Economic Activities (ISIC) Revision 4.[10] A user is classified as a worker if there is at least one word from this list in any of their posts or in their user description. By the same logic, a user is allocated to a specific sector if a post or user description has at least one word from the keyword list for this sector.[11] Appendix Table D1 contains a list of variables constructed from this data set. We geo-localized all posts and computed the number of users from different sectors instead of posts in a given grid cell. The geographic distribution and intensity of Weibo posts in Africa is shown in Figure 1c. Figure 2 shows a map of the home regions of the Chinese individuals who published Weibo posts while located in Africa and who are identified as workers (Weibo users who can be unambiguously

[9]Other studies that used these data include an investigation of happiness in China by Zheng et al. (2019) and a study of state influence by Qin and Wu (2017). To our knowledge, ours is the first study to use Weibo data in Africa and track the footprint of Chinese overseas activities.

[10]The detailed list is available upon request.

[11]For simplicity and transparency, we adopt a binary coding depending on whether a key word appears in an individual's posts or not. This implies that the same individual can be assigned to more than one sector. In principle, the assignment of individuals to a worker status and a particular sector is probabilistic. We refrain from a more involved assignment procedure in light of the complications that would arise for estimation.

associated with work-related information).[12] Summary statistics are reported in Tables C1 and C2 in the Supplementary Appendix, respectively.

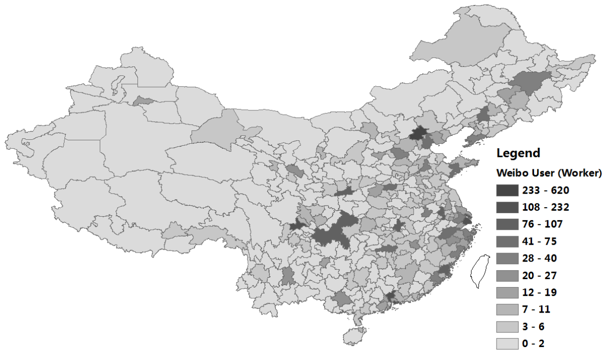



Figure 2: Chinese Prefectures of Origin of Weibo Users (Workers) in Africa

*Note.* The figure shows the prefectures of origin of Weibo users in Africa that can be associated with posts with an unambiguous relation to work (workers) and the frequency for each Chinese prefecture.

We next investigate the role of malaria exposure for the distribution of Chinese economic activities in Africa.

## 2.2 Malaria in Africa

**Epidemiology of Malaria.** Malaria is an infectious disease that is caused by *Plasmodium* parasites and that is widespread in tropical regions around the world. Of the five existing strains of *Plasmodium* parasites, *Plasmodium falciparum* is the most prevalent parasite in Africa and is responsible for most malaria deaths worldwide.[13] *Plasmodium* parasites can only be transmitted by bites of infected female anopheles mosquitos, which transmit the parasite into the human bloodstream. The reproduction cycle of the parasite takes place in both human hosts and the vector. Malaria infections are associated with severe symptoms within seven to eighteen days after infection.[14] Malaria infections can lead to death of the infected individual if not diagnosed and treated promptly. Other complications include severe anemia and cerebral malaria. People who survive an acute malaria infection acquire partial immunity and resistance to reinfection that is related with considerably milder symptoms (see, e.g., Lell et al., 2018). Without continued exposure to malaria, this acquired partial immunity vanishes over time.

Malaria imposes substantial costs on both individuals and governments. Although malaria deaths declined by 60 percent from 2000 to 2015 due to widespread prevention and treatment

[12]The corresponding map for Weibo users where this association of posts with work is not obvious (unidentified Weibo users) can be found in Figure C1 in the Supplementary Appendix.

[13]According to WHO (2017), more than 3.2 billion people live in areas at risk of malaria transmission in 106 countries. There were 216 million malaria cases which resulted in an estimated 445,000 deaths worldwide in 2016. More than 90 percent of clinical cases and deaths worldwide occur in Africa.

[14]Once an individual is infected, the parasite enters the bloodstream and travels to the liver, where it grows and multiplies before re-entering the bloodstream and multiplying in the red blood cells. At regular intervals, the infected blood cells burst, causing symptoms including fever, nausea, shaking chills, vomiting, and headaches.

(WHO, 2015), malaria imposes substantial economic costs estimated to amount to US$12 billion per year to Africa.

**Latent Exposure to Malaria.** Exposure to malaria is related to the prevalence of the parasite and of the anopheles vector, which also requires suitable geo-climatic conditions for reproduction.[15] Anopheles mosquitoes lay their eggs in water, which hatch into pupae, eventually emerging as adult mosquitoes. The length of this process is sensitive to temperature, with the suitable temperature ranging from 14 to 28°C and a peak for the speed of the emergence of adult anopheles mosquitos around 24°C (Lyons, Coetzee and Chown, 2013). Female anopheles mosquitos require blood meals from vertebrate to nurture their eggs. During these blood meals they transmit the parasite. This transmission is also temperature sensitive.

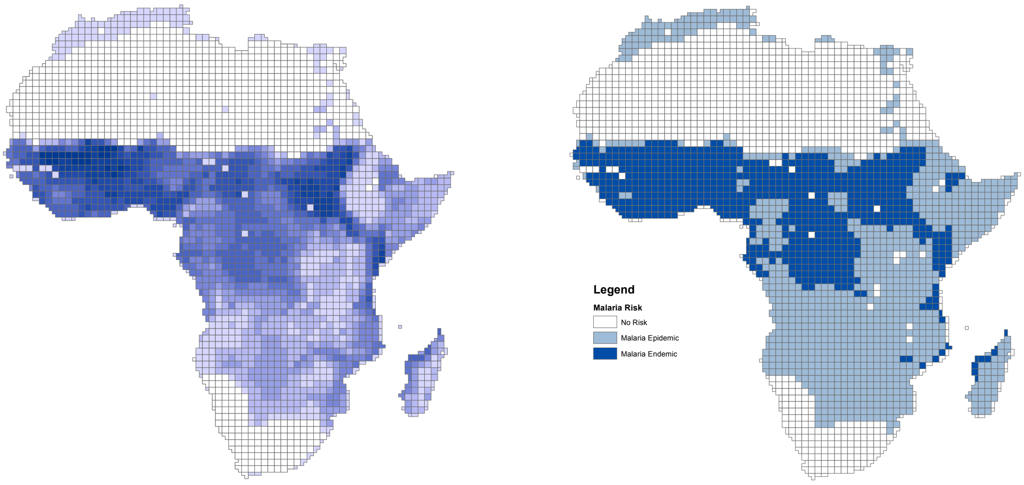
Given that the transmission speed and intensity on biological characteristics of the parasites and mosquitoes and on climatic conditions, the literature has developed measures of malaria exposure that are based on the local variation in conditions related to mosquito prevalence, types and climate. A frequently used index, constructed by Kiszewski et al. (2004), uses biological characteristics and climatic conditions to construct a measure of the force and stability of malaria transmission. This time-invariant index measures the stability and force of malaria transmission globally, as a function of the characteristics of the respective mosquito vectors for malaria and climate conditions.[16] The resulting time-invariant stability index implies that the larger the index, the stronger the stability and force of the malaria transmission. According to the epidemiology of malaria, malaria resistance (genetic and acquired) is higher in regions historically most exposed to the disease. Hence malaria ecology is an indirect measure of the level of immunity to malaria. The lower the malaria ecology index, the lower the exposure to frequent and repeated malaria infections, and, correspondingly, the lower the resistance of the adult population. Figure 3a depicts this index on a grid of $1^\circ \times 1^\circ$ cells for the entire African continent. At the cell level, the index ranges from 0 to 36.

To investigate the impact of the malaria ecology index on Chinese economic activities, we consider a baseline specification in which the index enters linearly. In additional analyses, we use variation of the index in various alternative ways, including a binary classification of malaria

[15] Adult anopheles mosquitoes live for about 1-2 weeks during which the transmission is possible. In many places, transmission is seasonal, with the peak during and just after the rainy season. The extrinsic incubation period, reflecting the reproduction cycle, lasts from 10 to 21 days varying with the temperature.

[16] In particular, the malaria ecology index developed by Kiszewski et al. (2004) is an ecology-based spatial index of stability and force of malaria transmission. The index is constructed based on the characteristics of the regional dominant vector mosquito and the sum of the transmission intensity index for each month. It consists of three factors: (1) duration of the extrinsic incubation period for each month which is determined by the average monthly temperature; (2) daily survival rate of mosquitos; (3) the biting rate of humans. The index is computed as $\sum_{m=1}^{12} a_{i,m}^2 p_{i,m}^E / -\ln(p_{i,m})$, where $m$ is month (1 - 12), $i$ represents the identity of the locally dominant vector mosquito, $p$ is the daily survival rate (0 - 1), $a$ is the preference for humans in terms of biting rate (0 - 1), $E$ is the length of the extrinsic incubation period in days. The time span of temperature is measured from 1901 to 1990.

risk where we code cells as endemic or epidemic areas. This coding captures the emergence of acquired immunities due to frequent exposure to malaria infections. In endemic areas, the force and stability of malaria transmission implies a fairly constant exposure to malaria, so that a relatively large portion of the local population exhibits acquired resistance or genetic immunity. Where the exposure to malaria is more infrequent, acquired resistance or immunity is much less prevalent, so when weather conditions that favor the transmission of the disease occur, epidemic outbreaks of malaria follow. As a way to operationalize the latent risk of such epidemic outbreaks, we construct a binary index for risk of malaria epidemics that takes value one for cells with an average malaria ecology index greater than 0 and smaller than 15; cells with a malaria ecology index of 15 and higher are coded as endemic areas. This classification follows previous work (Cervellati et al., 2018; Cervellati, Esposito and Sunde, 2021). The respective map is shown in Figure 3b.




(a) Malaria Ecology (b) Malaria Risk

Figure 3: Malaria Exposure in Africa

*Note.* Panel (a) maps the spatial distribution of malaria ecology index in Africa. Darker shades indicate higher and more stable malaria exposure. Panel (b) displays the degree of malaria risk in terms of endemic and epidemic areas. Darker cells correspond to cells where malaria is endemic in terms of a malaria ecology index equal to or greater than 15. Lighter cells correspond to cells where malaria is epidemic in terms of a malaria ecology index greater than zero and smaller than 15. White cells correspond to cells with no malaria risk in terms of a malaria ecology index equal to zero.

**Risk of Epidemic Malaria Outbreaks.** As proxy measure for the risk of epidemic malaria outbreak, we combine the binary index of epidemic malaria exposure with information about the occurrence of weather conditions that are necessary conditions for elevated malaria transmission (Tanser, Sharp and Le Sueur, 2003).[17] As a proxy of epidemic outbreak risk, we construct a

[17] These are: (a) the average monthly rainfall during the past 3 months (the current month and the two previous months) is at least 60mm; (b) rainfall in at least one of these three months is at least 80 mm; (c) none of the past 12 months (the current month and the previous 11 months) had an average temperature below 5°C; and (d) the average temperature in the past 3 months exceeds the sum of 19.5°C and the standard deviation of monthly temperature in the past 12 months.

variable Malaria Suitable Months (MSM) that corresponds to the average number of months per year in the period 1998 to 2012 during which a given cell experienced these conditions. The resulting measure ranges from 0 to 12.[18] Following the same logic of the binary definition of endemic and epidemic areas, we construct a measure of outbreak risk by the interaction between the average number of months with suitable climatic conditions for malaria outbreaks and the binary indicator reflecting malaria epidemic cells that takes value 1 if the malaria ecology index in a cell is smaller than 15 and greater than 0. This measure captures that cells with a lower force and stability of malaria transmission, as reflected by the binary measure of epidemic areas, are populated by a population that shows little or no resistance to malaria outbreaks. In such an environment, malaria transmission spreads quickly once the climatic conditions for malaria are satisfied.[19]

**Genetic Immunity to Malaria.** While acquired resistance or immunity as consequence of repeated malaria infections wears out over time, immunity can also be the result of genetic mutations. In regions with high malaria exposure, malaria, and especially *Plasmodium falciparum*, has exerted measurable selective pressure on the human genome that resulted in genetically determined immunity (Kwiatkowski, 2005).

We use information for three different genetic mutations to measure the level of immunity in the populations of local Africans and of Chinese workers: the sickle cell trait, the absence of Duffy antigens, and glucose-6-phosphate dehydrogenase (G6PD) deficiency.

The sickle cell trait (HbAS) is a benign mutation that, for heterozygotes, has a protective effect against all malaria phenotypes without being associated with protection against other diseases or with negative health consequences in normal life (Taylor, Parobek and Fairhurst, 2012; Uyoga et al., 2019; Purohit et al., 2018).[20] Data on the sickle cell trait is from Piel et al. (2013) and measures the frequency of the sickle haemoglobin (HbAS) allele in 5×5 km areas. The sickle cell trait is the most effective protective genetic factor against *Plasmodium falciparum* malaria after a person has been infected (Ferreira et al., 2011). A low level of haem in the blood of people carrying the sickle cell gene stimulates the production of an enzyme that breaks it down and releases small quantities of carbon monoxide that kill the *Plasmodium* parasite. The sickle cell trait has an evolutionary advantage where malaria infection is frequent.

The Duffy antigen, or Duffy receptor for chemokines, is a protein in the membrane of red blood cells (erythrocytes) that plays an important role as receptor for the invasion by *Plasmodium* parasites. Individuals without the Duffy antigen exhibit resistance to the outbreak

[18]The spatial distribution of Malaria Suitable Months in Africa is illustrated in Figure C2a in the Appendix.

[19]The spatial distribution of this measure is depicted in Figure C2b in the Appendix.

[20]In contrast, homozygosity implies sickle cell anemia, which is a condition associated with high mortality, particularly at young ages.

of malaria (although not the infection). Data on the prevalence of the Duffy antigen negativity in the population is obtained from Howes et al. (2011).

Finally, G6PD deficiency is a genetic disorder that provides significant protection against severe, life-threatening malaria such as cerebral malaria for male but not female children (Guindo et al., 2007). Data on the global predicted allele frequency for G6PD deficiency globally is obtained from Howes et al. (2012). We compute the average of each index at the cell level.

These genetic disorders only provide partial immunity or resistance, and only for particular groups of individuals. As an average measure of the prevalence of genetic resistance at the $1° \times 1°$ degree cell level, we combine the measures at the cell level using a principal component analysis. Figure 4 shows the corresponding map for Africa.[21]

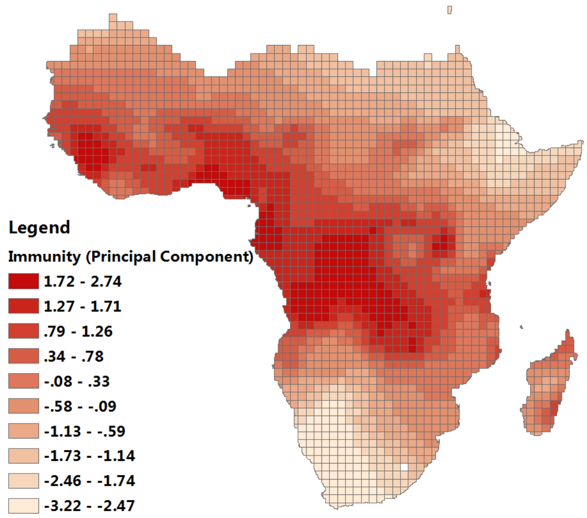



Figure 4: Immunity Level - Principal Component

*Note.* The figure displays the first principal component of sickle cell trait and Duffy negativity in Africa.

**Other Data.** The empirical analysis also uses information about other cell characteristics from various sources. The variables serve as control variables, for robustness checks, or for explorations of the mechanisms behind the main findings. They include, in particular, information about the prevalence of other diseases such as HIV or the Tsetse fly that transmits trypanosomiasis, geographic and climatological variables, institutional and historical characteristics, and information about anti-malaria policies. The different variables and the respective data sources are described in Appendix Table C1 and Appendix Table C3 contains summary statistics.

[21]Since for Africa, the distribution of G6PD displays a very similar geographic pattern to that of the malaria ecology index and for the sickle cell trait, we compute the principal component of sickle cell trait and Duffy antigen negativity only. Figure C3a displays the immunity level in terms of sickle cell in Africa and Figure C3b displays the corresponding map for the prevalence of the absence of the Duffy antigen.

# 3 Malaria and Chinese Activities in Africa

## 3.1 Empirical Strategy

We test the hypothesis that exposure to malaria has an influence on Chinese economic activities in Africa. With $1^\circ \times 1^\circ$ grid cells as unit of observation, the empirical framework is given by

$$Activity_{cjr} = \alpha Malaria\ Exposure_{cjr} + X'_{cjr}\gamma + \delta_{jr} + \epsilon_{cjr} \ , \tag{1}$$

with $Activity_{cjr}$ representing a continuous variable of the Chinese activity in cell $c$ in country $j$ and region $r$. As activities, we first focus on aid and infrastructure projects, and then on the location of Chinese workers in Africa. The coefficient of main interest is $\alpha$, which captures the effect of malaria exposure on Chinese economic activities in cell $c$. $Malaria\ Exposure_{cjr}$ is a continuous measure of the force and stability of malaria transmission index developed by Kiszewski et al. (2004) in the baseline specification.

The vector $X_{cjr}$ contains geographic and economic control variables. The geographic controls are average temperature and total precipitation in 2013, absolute latitude, land ruggedness, natural logarithm of the distance to the country capital, to the coast and to China, average land suitability for agriculture, the presence of petrol fields, and the presence of mines. Economic controls include population in terms of $\ln(population\ count{+}1)$ which corresponds to population density at the cell level and is a proxy of market size, a proxy of the local development level in terms of satellite-based data on night light intensity, $\ln(night\ light\ intensity + 0.01)$, as well as 2G internet coverage and fibre cable coverage. In addition, we include various fixed effects $\delta_{jr}$ to account for the effect of systematic variation in institutions, policies, and other related variables. In some specifications, we include country fixed effects, $\delta_j$, but as a baseline, we include fixed effects at lower levels of spatial aggregation, $\delta_{jr}$, namely at the level of administrative areas, and at the level of cultural groups. To account for administrative areas, we constructed region polygons using the GADM dataset, focusing on administrative region of level 1 (corresponding to province), which allows us to associate cells with the location in one of 728 regions. For constructing cultural group fixed effects, we rely on the maps constructed by Murdock (Murdock, 1959, 1967, Nunn, 2008). As baseline, we associated each cell with cultural groups by locating the centroid of the cell within the cultural group polygon. This way, we associate each cell with one of 105 cultural groups. We estimate the model using least squares. Robust standard errors are computed allowing for spatial dependencies and arbitrary heteroskedasticity of the error term of all cells in the baseline model.

In terms of identification, the usual assumptions of exogeneity of $Malaria\ Exposure_{cjr}$ in the sense of orthogonality to the error term conditional on the control variables contained in

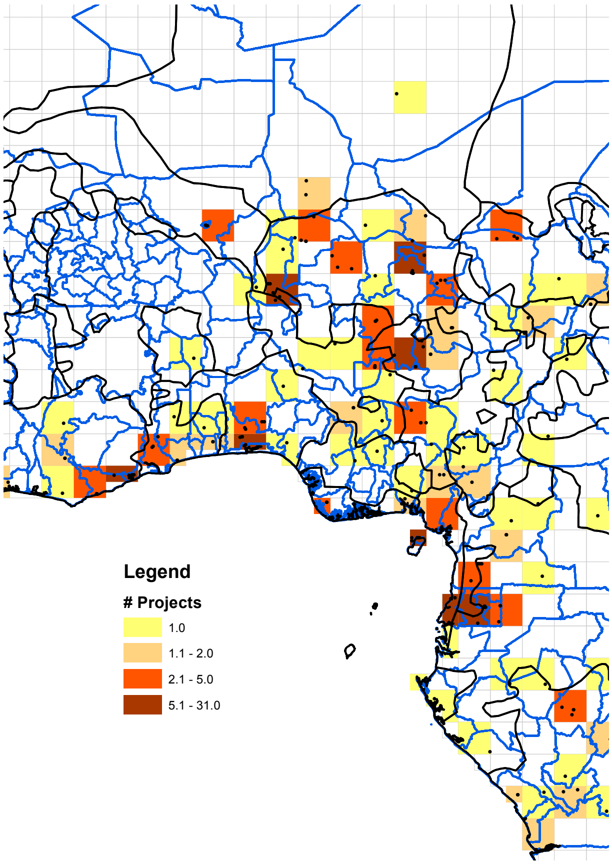



Figure 5: Illustration of Fixed Effects for Region and Ethnic Groups

*Note.* The figure shows a zoomed-in map of West Africa to illustrate the construction of fixed effects for administrative areas and cultural groups at the cell level. Blue polygons correspond to administrative areas at the subnational level (level 1), black polygons correspond to cultural group homelands. Dots within cells correspond to the geographic location of infrastructure projects.

the empirical framework (1) must be fulfilled. As our measure of malaria exposure is based on variation in malaria ecology and in the latent malaria outbreak risk (which is related to variation in weather conditions), reverse causality with respect to Chinese economic activities in Africa is ruled out by construction. The greater concern for identification is the presence of unobserved third factors that correlate with malaria exposure and foreign activities. These could include general economic activity, or institutional or historical factors. To account for such heterogeneity, the empirical framework includes an extensive set of geographic and economic controls, controls for sub-national regions and cultural groups. In particular, confounders would have to vary systematically above and beyond the administrative region and cultural group fixed effects, which account for systematic heterogeneity in institutions and cultural factors at a high resolution. Figure 5 provides an illustration of the identifying variation conditional on administrative region and cultural group fixed effects.[22] This constitutes a substantially more demanding specification than accounting for country fixed effects. In robustness checks, we

[22]Alternatively, we also constructed ethnic group fixed effects by associating each cell with the ethnic group occupying the cell centroid, with similar results.

present results for all cells, and when restricting the sample to grid cells with strictly positive values of the malaria ecology index.[23] Moreover, some of the later analysis conditions on the overall level of Chinese foreign activities and focuses on the composition of the Chinese workforce, which provides additional indirect evidence for the validity of the hypothesis, since any confounding factor, beyond affecting Chinese activities in terms of projects or presence of individuals, would have to account for the selection of Chinese workers from particular regions in China.

### 3.2 Malaria and Chinese Projects in Africa

Table 1 presents the estimation results for the impact of malaria exposure, measured by the malaria ecology index of the force and stability of malaria transmission, on Chinese aid projects (Panel A) and on Chinese infrastructure projects (Panel B). The dependent variable in Columns (1) and (2) is the value and the number of all projects of the respective type in a given cell. The dependent variable in the remaining columns is the number of projects in different categories. For the aid data in Panel A, we consider infrastructure projects, projects involving the presence of Chinese personnel, ODA-like projects, and donations. For the contracted infrastructure projects in Panel B, we present corresponding results for infrastructure projects across different sectors: hydropower, public buildings, transports, and utilities. All results are obtained with specifications that include the entire set of control variables and fixed effects for administrative regions and cultural groups. The estimates report Conley (1999) standard errors accounting for spatial autocorrelation for a cutoff of 500km in square brackets.[24]

Focusing on the results for the aid data in Panel A, greater malaria exposure has a negative impact on the aggregate value and number of aid projects. The effect is significant for the value of projects, and highly significant when considering the number of projects. The results in Columns (3)-(6) indicate that this negative effect pertains to different categories of aid projects, but reveals considerable heterogeneity in terms of effect size. Malaria exposure has a significantly negative effect on the number of infrastructure projects, projects involving Chinese personnel, and ODA-like projects; these effects are quantitatively comparable and largest for ODA-like projects, whereas the effect on donations is significantly smaller. The results are qualitatively similar but statistically weaker when considering infrastructure projects contracted by Chinese firms, as shown in Panel B. While the negative effect remains negative in all specifications, the

[23]Excluding the cells with malaria ecology index = 0 from the sample produces similar results. Cells with a malaria ecology index = 0 are usually deserts and have a very low population density, which implies that they do not add relevant variation regarding Chinese activities and the presence of Chinese workers. Additional robustness checks also consider subsamples that exclude very thinly populated cells (in terms of the lowest one or two deciles of the population density).

[24]In view of the recent debate on the appropriate choice of cutoffs for spatial clustering (Kelly, 2019; Colella et al., 2019), we experimented with different cutoffs ranging from 50km to 1,000km. The cutoff of 500km overall delivered the most conservative estimates of standard errors.

effect is statistically significant only for the number of all projects and for projects concerned with public buildings. The smaller and less significant results in Panel B may be due to the fact that infrastructure projects are financed from various sources, implying that the location of the projects is affected also by other parties such as African governments, foreign countries, and regional development banks. In comparison, aid projects in Panel A are financed by China, which implies a potentially stronger strategic component to allocate projects considering the exposure to malaria for Chinese personnel.

Instead of considering estimates for the full sample of cells, we replicated the analysis for the sample of cells with a strictly positive malaria exposure in terms of the malaria ecology index. Again, malaria exposure exhibits a negative effect on the number of projects, but estimates are less precise and significant effects are only found for ODA-like aid projects and public buildings.[25]

### 3.3 Projects and Weibo Users

As a next step, we establish that Chinese activities in Africa are associated with the presence of Chinese workers. In the following analysis, we use the number of Weibo users in a cell as proxy for the presence of Chinese individuals in a particular region. Since Chinese individuals working in Africa are relatively young and affine to social media like Weibo, information on the location of posts offers a unique opportunity to overcome the problem of reliable and sufficiently disaggregated data about the geographic location of Chinese workers.[26] Moreover, the analysis exploits information on the number of users in each location rather than information on the number of posts, because the frequency might be directly affected by individual exposure to malaria.

To establish that Chinese projects in Africa, particularly projects related to infrastructure, use Chinese workers, we regress the number of projects in a geographic cell on the number of Weibo users located in the cell. This cross-validates the usefulness of Weibo data for locating Chinese individuals and confirms the validity of the information that is used to classify Chinese individuals as workers on the basis of the Weibo data.

The results are in Table 2 for aid projects (Panel A) and infrastructure projects (Panel B).

[25] Table A1 in the Appendix reports the corresponding results. Similar results are obtained when focusing on cells with high population density, see Appendix Table A2. Unreported results for a less restrictive specification that accounts for country fixed effects rather than fixed effects for sub-national administrative regions and cultural groups also deliver qualitatively similar estimates.

[26] According to the official report of Sina Weibo (Weibo, 2015), more than 50% of all Weibo users in 2015 were between 24 and 45 years of age, more than 90% were between 17 and 45 years of age, and 50% of Weibo users were female. The gender distribution of Weibo users in our data set was 49.6% female and 50.4% male users. In our data, we have no information on age, but according to hiring advertisements of Chinese companies doing business in Africa, workers to be deployed in Africa are typically required to be younger than 45 years of age (see, e.g., `https://africa.rencaijob.com/`).

Table 1: Malaria Exposure and Chinese Projects in Africa

| Panel A: Aid Projects<br>Dependent Variables | Aid Projects | | Number of aid projects | | | |
|---|---|---|---|---|---|---|
| | Value | Number | Infrastructure | Involving Chinese | ODA-like | Donation |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Malaria ecology | -1581.345** | -0.254*** | -0.140*** | -0.173*** | -0.182*** | -0.047*** |
| | (655.012) | (0.064) | (0.045) | (0.052) | (0.039) | (0.017) |
| | [588.724] | [0.055] | [0.035] | [0.045] | [0.034] | [0.014] |
| Observations | 2524 | 2524 | 2524 | 2524 | 2524 | 2524 |
| $R^2$ | 0.574 | 0.665 | 0.590 | 0.594 | 0.677 | 0.759 |
| Panel B: Construction Projects<br>Dependent Variables | Construction Projects | | Number of Projects in | | | |
| | Value | Number | Hydropower | Public Building | Transport | Utility |
| Malaria ecology | -90.858 | -0.141** | -0.017 | -0.073** | -0.024 | -0.026 |
| | (488.866) | (0.062) | (0.011) | (0.035) | (0.026) | (0.016) |
| | [321.502] | [0.046] | [0.013] | [0.028] | [0.021] | [0.013] |
| Observations | 2524 | 2524 | 2524 | 2524 | 2524 | 2524 |
| $R^2$ | 0.684 | 0.662 | 0.409 | 0.656 | 0.596 | 0.515 |
| Both Panels: | | | | | | |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Economic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Admin. Area FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Ethnic Group FEs | Yes | Yes | Yes | Yes | Yes | Yes |

*Notes.* OLS estimates. The unit of observation is a 1 x 1 degree cell. All continuous control variables are standardized. The dependent variable is the total number of Chinese aid projects in Africa from 2010 to 2014, data from AidData (Panel A) or number or value of the infrastructure projects in various sectors contracted by Chinese companies in Africa in 2013 from Yearbook (Panel B). "Malaria ecology" is a time-invariant index measuring the stability and force of malaria transmission in a cell. "Admin. Area FEs" represent (dummy) controls for level-1 administrative areas at the sub-national level from the Database of Global Administrative Areas (GADM). "Ethnic Group FEs" represent (dummy) controls for cultural groups as defined by Murdock (1967) (see also Nunn, 2008). "Geographic Controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic Controls" include population, night lights, 2G internet coverage and Fibre cable area. Robust standard errors are reported in parentheses. Conley (1999) standard errors are reported in square brackets (cutoff = 500 km). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

They confirm that a higher number of projects in a cell is associated with more Weibo users in that cell. This effect is larger for Weibo users who can be unambiguously associated with work-related information (workers) than for Weibo users where this is not possible (unidentified). In a horse race, shown in Column (3), the effect of the presence of workers using Weibo is significantly positive and about three times larger than the effect of the presence of Chinese individuals that cannot be identified as workers using Weibo posts. In terms of effect size, an increase of 50% (half a log point) in Weibo users in a cell is associated with about 0.25 more infrastructure projects (which roughly corresponds to the unconditional mean of 0.26). While these findings reflect a statistical association rather than a causal link, they establish the close relationship between Chinese infrastructure projects and the presence of Chinese workers. If malaria affects the presence of Chinese workers, it is therefore conceivable that it has an indirect effect on the location of infrastructure projects.

Table 2: Chinese Projects and Weibo Users

| Panel A: Aid Projects | Number of Projects | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| ln Weibo user | 0.532*** | | | 0.475*** | | |
| | (0.090) | | | (0.085) | | |
| ln Weibo user (worker) | | 0.776*** | 0.594*** | | 0.696*** | 0.512*** |
| | | (0.134) | (0.127) | | (0.118) | (0.104) |
| ln Weibo user (unidentified) | | | 0.187** | | | 0.196** |
| | | | (0.079) | | | (0.082) |
| Observations | 2542 | 2542 | 2542 | 2524 | 2524 | 2524 |
| $R^2$ | 0.307 | 0.319 | 0.322 | 0.695 | 0.700 | 0.701 |

| Panel B: Construction Projects | Number of Projects | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| ln Weibo user | 0.405*** | | | 0.284*** | | |
| | (0.087) | | | (0.070) | | |
| ln Weibo user (worker) | | 0.600*** | 0.517*** | | 0.396*** | 0.298*** |
| | | (0.137) | (0.129) | | (0.107) | (0.103) |
| ln Weibo user (unidentified) | | | 0.085 | | | 0.105 |
| | | | (0.064) | | | (0.066) |
| Observations | 2542 | 2542 | 2542 | 2524 | 2524 | 2524 |
| $R^2$ | 0.251 | 0.263 | 0.264 | 0.678 | 0.678 | 0.679 |
| Both Panels: | | | | | | |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Economic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Country FEs | Yes | Yes | Yes | No | No | No |
| Admin. Area FEs | No | No | No | Yes | Yes | Yes |
| Ethnic Group FEs | No | No | No | Yes | Yes | Yes |

*Notes.* OLS estimates. The unit of observation is a 1 x 1 degree cell. All continuous control variables are standardized. The dependent variable is the total number of Chinese aid projects in Africa from 2010 to 2014, data from AidData (Panel A) or number or value of the infrastructure projects in various sectors contracted by Chinese companies in Africa in 2013 from Yearbook (Panel B). The independent variable "ln Weibo users (workers)" is the logarithm of the Weibo users identified as worker plus one in Africa in 2014 and "ln Weibo users (unidentified)" is the logarithm of the Weibo users unidentified plus one in Africa in 2014. "Geographic Controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic Controls" include population, night lights, 2G internet coverage and Fibre cable area. "Admin. Area FEs" represent (dummy) controls for level-1 administrative areas at the sub-national level from the Database of Global Administrative Areas (GADM). "Ethnic Group FEs" represent (dummy) controls for culture groups as defined by Murdock (1967) (see also Nunn, 2008). Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

### 3.4 Malaria and the Presence of Chinese Workers in Africa

We now investigate the hypothesis that malaria exposure affects projects through its influence on the physical presence of Chinese workers in Africa. We test our conjecture by adapting the empirical framework (1) to the use of Weibo data as a proxy of Chinese economic activities. In particular, as a measure of $Activity_{cj}$ we use the log of Weibo users (workers) in a cell in 2014. The remaining specification is the same, except that control variables are also measured in 2014 for consistency.

Table 3 presents the results for malaria exposure in terms of the malaria ecology index for different specifications of the empirical model for the full sample (Panel A) and for the sample of cells with a strictly positive malaria ecology index (Panel B). The estimates provide evidence for a significant negative effect of malaria exposure on the presence of Chinese workers in a cell, consistent with the conjecture that a harsher disease environment discourages Chinese activities that require Chinese workers.[27]

The result that malaria exposure affects the location of Chinese workers raises the question whether this might indeed be the channel behind the earlier results on the effect of malaria exposure on the location of Chinese aid and infrastructure projects. In fact, additional results from an extended specification suggest that the presence of Chinese personnel largely accounts for the effect of malaria exposure on Chinese projects.[28]

Table 3: Malaria and Weibo Users Identified as Workers - Linear Specification

| Dependent Variables | ln Weibo users (workers) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Panel A: Full Sample | | | | |
| Malaria ecology | -0.034** | -0.142*** | -0.172*** | -0.205*** |
| | (0.014) | (0.033) | (0.033) | (0.030) |
| | [0.034] | [0.041] | [0.035] | [0.030] |
| Observations | 2542 | 2524 | 2524 | 2524 |
| $R^2$ | 0.002 | 0.596 | 0.625 | 0.704 |
| Panel B: Cells with Malaria Ecology > 0 | | | | |
| Malaria ecology | -0.177*** | -0.243*** | -0.241*** | -0.192*** |
| | (0.023) | (0.057) | (0.068) | (0.057) |
| | [0.036] | [0.067] | [0.063] | [0.049] |
| Observations | 1621 | 1609 | 1609 | 1609 |
| $R^2$ | 0.041 | 0.648 | 0.678 | 0.777 |
| Controls: | | | | |
| Geographic Controls | No | No | Yes | Yes |
| Economic Controls | No | No | No | Yes |
| Admin. Area FEs | No | Yes | Yes | Yes |
| Ethnic Group FEs | No | Yes | Yes | Yes |

*Notes.* OLS estimates. The unit of observation is a 1 x 1 degree cell. All continuous control variables are standardized. The dependent variable "ln Weibo users (workers)" is the logarithm of the Weibo users identified as workers plus one in Africa in 2014. "Malaria ecology" is a time-invariant index measuring the stability and force of malaria transmission in a cell."Geographic Controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic Controls" include population, night lights, 2G internet coverage and Fibre cable area. "Admin. Area FEs" represent (dummy) controls for level-1 administrative areas at the sub-national level from the Database of Global Administrative Areas (GADM). "Ethnic Group FEs" represent (dummy) controls for culture groups as defined by Murdock (1967) (see also Nunn, 2008). Robust standard errors are reported in parentheses. Conley (1999) standard errors are reported in square brackets (cutoff = 500 km). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

### 3.5 Robustness and Additional Results

**Alternative Specifications.** To investigate the robustness of the finding of a negative effect of malaria exposure on the presence of Chinese workers, we conducted several alternative analyses. In particular, the finding of a negative effect of malaria exposure on the presence of

[27] Results for similar further restrictions on the sample as in the analysis of projects are reported in Table A3.

[28] See Appendix Table A4.

Chinese activities in a given cell in Africa are also robust to alternative estimation methods.[29] The results are also robust to the use of Weibo posts by all individuals and not just those by individuals who are identified as workers.[30] Earlier work has emphasized the role of disease for the formation of precolonial institutions, population density, or economic activity (see, e.g. Alsan, 2015). When investigating the respective analysis for precolonial institutions, we find no evidence for an effect of malaria exposure on centralization or historical development as measured by population density, intensive agriculture, or plow use.[31] We also investigated the role of malaria ecology for economic development, proxied by night light density. The results indeed suggest a negative association. At the same time, however, a sensitivity analysis reveals that the results are very similar regardless of whether night lights are included as control for the level of economic development in a cell.[32] We also extended specifications with controls for the prevalence of other diseases such as HIV or vectors of other diseases such as the Tsetse fly that transmits trypanosomiasis. The result of a negative effect of malaria ecology is confirmed, while neither HIV nor Tsetse suitability have a significant effect on Chinse activities.[33] We also replicated the results for extended specifications that account for a variety of other control variables.[34] Throughout, we find a consistently negative effect of malaria exposure on the presence of Chinese personnel.

To account for possibly non-linear effects of malaria exposure, we replicated the estimates for log number of Weibo users (workers) using a semi-parametric specification for malaria exposure with value 15 as the reference category, and controlling for country fixed effects and geographic and economic characteristics. The results reveal a visible decline in the presence of Chinese workers the higher the level of malaria exposure in terms of higher levels of the malaria ecology index and provide no evidence for strong non-linearities in the effect.[35] These results suggest that using a linear specification or a specification with a binary indicator that classifies cells into endemic and epidemic, with endemic cells exhibiting a malaria ecology index larger than 15 and epidemic cells exhibiting a positive value below 15, should deliver qualitatively similar results. To further explore the robustness of the results, which are based on cross-sectional variation in malaria exposure as measured by the malaria ecology index, we also conducted

[29] Table A5 in the Appendix replicates the analysis using a binary variable for the extensive margin of Chinese activities in terms of any project or Weibo post in a given cell or not. Unreported results for Tobit regressions that account for censoring of projects and Weibo posts at 0 are qualitatively similar.

[30] See Table A6 for the corresponding results for all Weibo users.

[31] See Appendix Table A7 for details.

[32] See Tables A8 and A9 in the Appendix.

[33] The results are shown in Appendix Table A10.

[34] See Table A11 in the Appendix.

[35] See Appendix Figure A1. Geographic characteristics include average temperature, total precipitation, absolute latitude, land ruggedness, natural logarithm of the distance to the country capital, to the coast and to China, average land suitability for agriculture, the presence of petrol fields, and the presence of mines. Economic characteristics are log(population count + 1) and log(night light intensity + 0.01).

additional exercises that make use of alternative measures in malaria exposure. The results for other cross-cell measures of malaria risk deliver similar findings.[36]

**Exposure to Epidemic Malaria Outbreaks.** For additional insights regarding the heterogeneity of the effect across cells and their sensitivity to outside conditions, we also explored the robustness of the results by considering variation in the exposure to epidemic malaria outbreaks. To do this, we adapt the approach taken by Cervellati, Esposito and Sunde (2021) and consider the sensitivity to the occurrence of suitable weather conditions in the context of a local population that shows little or no resistance to malaria outbreaks. This analysis is based on the conjecture that the effect might differ across cells with high malaria exposure and cells with less frequent exposure and a population with less acquired resistance that is more susceptible to epidemic outbreaks. To investigate this conjecture, we estimate an extended specification that accounts for endemic malaria exposure, in terms of a malaria ecology index of 15 and above, the long-run frequency of suitable weather conditions for epidemic outbreaks (malaria suitable months), and an interaction of epidemic exposure (an index below 15) with the frequency of suitable weather conditions. The findings confirm a negative effect of high malaria exposure (endemic malaria) on the presence of Chinese workers. In addition, the results show an additional negative effect in epidemic areas that experience suitable conditions for an immanent epidemic outbreak of malaria.[37]

The conjecture that malaria exposure affects Chinese foreign activities is based on the hypothesis that the presence of Chinese firms and, in particular, personnel, is prevented by the overall malaria exposure in certain regions. Correspondingly, the main analysis relies on cross-cell variation, and the corresponding results refer to structural heterogeneity that affects foreign activities in the long run. Short run fluctuations, such as epidemic outbreaks of malaria, are not expected to affect the overall strategy of Chinese activities. Instead, the results suggest that firms avoid certain areas. However, there might be even short-run variation in the presence of Chinese activities is related to temporal variation in malaria exposure. Part of the effect might come from concentrating activities in malaria-prone areas to the less risky periods of the year. The sub-period effect can be explained by firms avoiding certain seasons that entail elevated health risks for their workers. However, we do not expect an analysis using variation within cells over short periods of time to show a similarly pronounced pattern as a cross-cell analysis, because temporal adaptation is not always possible in view of large projects that often

[36]See Appendix Table A12.

[37]See Appendix Table A13. In particular, in comparison to cells with no malaria exposure, cells with endemic malaria (i.e., large positive values of the malaria ecology index), have fewer Chinese workers as measured by Weibo posts, consistent with the findings from the linear specification. This effect is amplified by malaria risk, as measured by the interaction between malaria suitable conditions and epidemic exposure (i.e., a positive but comparably low value of the malaria ecology index).

take years to realize and where operations and equipment cannot simply be suspended or left temporarily.

To explore whether the negative effect of malaria exposure is also detectable by exploiting within-cell variation within the observation period we estimated a difference-in-differences analysis that extends the analysis to variation in the exposure to epidemic malaria outbreaks. Since the Weibo data are only available for the period February–December 2014, this analysis makes use of the interaction between the extent of malaria exposure and the occurrence of suitable conditions for the outbreak of malaria epidemics, following a similar approach as Cervellati, Esposito and Sunde (2021). In particular, we split the period in 2014, for which information on Weibo users is available, into two sub-periods and conduct the equivalent analysis in a panel with two periods. Again, the results reveal a negative effect on the presence of Chinese personnel if the risk of epidemic malaria outbreaks is elevated by the coincidence of a non-resilient population (in malaria epidemic cells) and suitable weather conditions for malaria outbreaks (malaria suitable months).[38]

**Presence of Chinese Workers Conditional on Projects.** To further investigate the concern of unobserved third factors that might correlate with both malaria exposure and foreign activities, we also explored whether the influence of malaria exposure affects the presence of Chinese personnel exclusively through the allocation of projects, or whether there is an effect on the presence even when holding the number of projects fixed. This analysis allows us to disentangle the relationship between malaria exposure and the presence of Chinese individuals as reflected by Weibo posts while holding fixed the overall intensity of Chinese activities in terms of projects in a given cell in Africa. While using projects as explanatory variable might raise concerns about potential endogeneity through a problem of "bad controls", it is nevertheless insightful by documenting the robustness or sensitivity of the previous results.

Table 4 shows the results for an extended specification that accounts for aid projects and infrastructure projects (both in terms of value and number). The findings reveal that the negative relationship between malaria exposure and the presence of Chinese workers as proxied by Weibo posts is robust with respect to controlling for value or number of aid projects. In fact, the conditional correlation is quantitatively very similar to the baseline results. At the same time, a larger value or number of projects is associated with a significantly larger number of Chinese workers.[39] Taken together, this suggests that the realization of projects is linked to the presence of Chinese personnel, but that malaria exposure continues to exert a negative

[38] See Appendix Table A14. Notice that this specification exploits variation between two half-year periods in 2014 and is very taxing in terms of the variation that can be used for identification. This prevents the use of higher frequency data and restricts the analysis to the consideration of country and period effects.

[39] These results complement the corresponding finding that Chinese personnel largely accounts for the effect of malaria exposure on Chinese projects, see Appendix Table A4.

effect on the presence of Chinese workers even when controlling for projects. This finding is reassuring since projects should account for unobserved factors that might influence Chinese activities in general, such as variation in institutional quality or economic activity. This finding also rationalizes that adverse conditions in terms of malaria exposure might have a less direct effect on Chinese projects than on the presence of Chinese personnel if it is possible to substitute local workers for Chinese workers. This might vary, e.g., across projects that differ in terms of their skill-intensiveness, which might be linked to a greater reliance on Chinese personnel. In sum, the finding of a negative effect of malaria exposure on the presence of workers above and beyond projects raises the question about the substitution of Chinese personnel with other workers, and about other potential mechanisms that lead to a varying degree of the intensity with which Chinese workers are used in the realization of projects. We turn to this issue and other attenuation mechanisms in the next section.

Table 4: Malaria and Weibo Users Identified as Workers – Conditional on Chinese Projects

| Dependent Variables | ln Weibo users (workers) | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Malaria ecology | -0.204*** | -0.187*** | -0.181*** | -0.166*** | -0.183*** | -0.178*** |
| | (0.029) | (0.028) | (0.026) | (0.028) | (0.027) | (0.028) |
| Project value | 0.134*** | | | | | |
| | (0.049) | | | | | |
| Project number | | 0.164*** | | | | |
| | | (0.044) | | | | |
| Value of aid projects | | | 0.147*** | | | |
| | | | (0.024) | | | |
| Number of aid projects | | | | 0.242*** | | |
| | | | | (0.048) | | |
| Value of infrastructure aid projects | | | | | 0.141*** | |
| | | | | | (0.024) | |
| Number of infrastructure aid projects | | | | | | 0.166*** |
| | | | | | | (0.029) |
| Observations | 2524 | 2524 | 2524 | 2524 | 2524 | 2524 |
| $R^2$ | 0.712 | 0.718 | 0.720 | 0.735 | 0.719 | 0.723 |
| Controls: | | | | | | |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Economic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Admin. Area FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Ethnic Group FEs | Yes | Yes | Yes | Yes | Yes | Yes |

*Notes.* OLS estimates. The unit of observation is a 1 x 1 degree cell. All continuous control variables are standardized. The dependent variable "ln Weibo users (workers)" is the logarithm of the Weibo users identified as worker plus one in Africa in 2014. "Malaria ecology" is a time-invariant index measuring the stability and force of malaria transmission in a cell. "Project value" and "Project number" are the value and number of the infrastructure projects in various sectors contracted by Chinese companies in Africa in 2013 respectively. "Aid projects" are the total value and number of Chinese aid projects in Africa from 2010 to 2014. "Infrastructure aid projects" are the Chinese aid projects in Africa from 2010 to 2014 for the purpose of infrastructure construction. "Geographic Controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic Controls" include population, night lights, 2G internet coverage and Fibre cable area. "Admin. Area FEs" represent (dummy) controls for level-1 administrative areas at the sub-national level from the Database of Global Administrative Areas (GADM). "Ethnic Group FEs" represent (dummy) controls for culture groups as defined by Murdock (1967) (see also Nunn, 2008). Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

# 4 Attenuation Mechanisms

## 4.1 Chinese Workers and Immunity of the Local Population

The findings in the previous section revealed effects of malaria exposure on Chinese infrastructure projects, and in particular substantial and significant effects on the presence of Chinese workers. In this section we explore various mechanisms and moderating factors that attenuate the effects of malaria exposure. These mechanisms are based on the premise that malaria infections lower the labor productivity and imply higher labor costs due to higher turnover.

To explore this, we first investigate heterogeneity in the effect across sectors. The motivation for this is that sectors differ in terms of the intensity with which they use specific skills or raw labor, and hence in terms of the potential to substitute Chinese workforce with local workforce. If the local population has adapted to a more intense exposure to malaria, this implies that projects can be scheduled and executed more easily due to the lower risk of epidemic malaria outbreaks, which might reduce the sensitivity of Chinese activities, and particularly of the presence of Chinese workers, to malaria exposure. This attenuation is expected to be less pronounced in the case of Chinese individuals entering a particular region for other reasons than work. Moreover, in a particularly malarious region, a project might be implemented in spite of the malaria risk by employing fewer Chinese workers if the elasticity of substitution is high and malaria immunity is high among the local population.

Table 5 presents estimates of the effect of malaria exposure on the presence of Chinese workers in Africa in relation to the immunity of the local African population to malaria infections. The results in Column (1) confirm the earlier results regarding a negative effect of malaria exposure on the presence of Chinese workers. Overall, this effect is unaffected by the prevalence of genetic immunity of the local (African) population. The remaining columns present results for heterogeneity across sectors by classifying the Weibo users that are identified as workers into five sectors: construction, manufacturing, business, hydropower and telecommunication. The results reveal the largest negative effect of malaria exposure for construction, manufacturing and business. For these sectors, the immunity of the local population does not exhibit a significant effect. In contrast, for hydropower and telecommunication, the effects of malaria ecology are smaller, while a higher prevalence of immunity among the local population exerts a negative effect that is marginally significant. These results suggest that the immunity of local workers does not attenuate the presence of Chinese workers in construction projects, manufacturing or business projects, but there is a negative effect in hydropower and telecom. Unreported results for interaction effects do not provide evidence for a substitution effect. At the same time, the heterogeneity in the effects parallels those on heterogeneity in the effect of malaria on Chinese activity in different sectors, particularly in hydropower and public buildings (see Table 1).

A possible explanation for this pattern is heterogeneity in the skill intensity. Since data on skill intensity across sectors for Africa are not available, we rely on information provided by the OECD to classify sectors by skill intensity in non-OECD economies. Based on this information, skill intensity is relatively low in manufacturing, construction and business, but high in electricity (including hydropower) and communication.[40] Interestingly, qualitatively similar results for malaria ecology, but substantially larger coefficient estimates for the prevalence of immunity in the African population are obtained for less restrictive specifications that account for country fixed effects or for country and cultural group fixed effects, but not for administrative region fixed effects.[41] This could be an indication for a greater substitution of Chinese personnel with local workforce, which is mainly related to skill composition and which varies across administrative regions. Overall, the results are therefore consistent with the hypothesis that the effect of malaria on the presence of Chinese personnel is smaller for projects in sectors that are skill-intensive and rely relatively more on Chinese technicians, experts and managers than projects in other sectors (Chen, Goldstein and Orr, 2009; Tang, 2016). In line with this conjecture, the presence of Chinese workers is affected less by malaria ecology in cells with projects in hydropower and telecom. In addition, a greater immunity of the local population exerts a negative, marginally significant effect on the presence of Chinese workers in these sectors, which suggests a greater elasticity of substitution under difficult working conditions for skilled workers. Overall, these results are also consistent with findings in the literature (Tang, 2010) that the share of local employment in Chinese companies is usually the highest in manufacturing sector.[42]

These findings are confirmed when considering alternative measures of immunity of the local African population. In particular, unreported findings suggest that the prevalence of immunity against malaria of the local population has a negative effect on the presence of Chinese workers, to a varying degree, in manufacturing, construction and business. The results are also confirmed for the non-linear specifications within the subsample of cells that are at risk of malaria transmission. Again, a greater prevalence of genetic immunity among the local population further reduces the presence of Chinese workers. Additional results suggest that this substitution effect is not found for sample of Weibo users that reveal no obvious connection to work (unidentified), while the main effect of malaria exposure is negative and at least as

[40]This refers to the average skill intensity across different sectors in Argentina, Bulgaria, Cyprus, Peru, Romania, and South Africa, see https://stats.oecd.org/Index.aspx?DataSetCode=SKILLS_2018_INDUSTRY. Skill intensity comprises Basic Skills (content), Basic Skills (Process), Complex Problem Solving Skills, Technical Skills, System Skills, Resource Management Skills, and the average of all skills. To proxy for the skill-intensity in the business sector, we refer to data from the OECD sector Wholesale and Retail Trade, while for Hydropower, we examine the skills in OECD sector Electricity, Gas, Steam, Water Supply and Water Management.

[41]See Tables A15 and A16 in the Supplementary Appendix.

[42]We abstract from mining since mines are used as collateral for loans from China.

large as for workers. Together, this suggests that any presence of Chinese individuals, even not work-related, is negatively affected by malaria exposure, while the substitution effect is mainly confined to work-related presence. The opposite effects of the prevalence of genetic immunities, which emerged in response to the past exposure to the pathogen, in regions of origin and destination of Chinese Weibo users on their location in Africa provides additional evidence for the specific role of malaria exposure in contrast to other location-specific correlates or confounders. Additional results based on information about leisure or leisure-related activities in Weibo posts reveal similar patterns. In particular, the elasticity of the presence of Chinese individuals in a cell with respect to the exposure to malaria is considerably smaller but still negative for individuals that are identified with work but whose posts contain no leisure-related activities when compared to posts by workers that contain references to leisure.[43]

Table 5: Chinese Workers and the Immunity of the Local Population

| Dependent Variable | ln Weibo users (workers) in | | | | | |
|---|---|---|---|---|---|---|
| | All Sectors | Construction | Manufacturing | Business | Hydropower | Telecom |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Malaria ecology | -0.178*** | -0.102*** | -0.117*** | -0.107*** | -0.043*** | -0.037*** |
| | (0.030) | (0.021) | (0.024) | (0.022) | (0.014) | (0.011) |
| African immunity | 0.049 | -0.026 | 0.005 | 0.006 | -0.047* | -0.048* |
| | (0.062) | (0.044) | (0.047) | (0.046) | (0.027) | (0.025) |
| Observations | 2022 | 2022 | 2022 | 2022 | 2022 | 2022 |
| $R^2$ | 0.721 | 0.662 | 0.679 | 0.650 | 0.587 | 0.587 |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Economic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Admin. Area FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Ethnic Group FEs | Yes | Yes | Yes | Yes | Yes | Yes |

*Notes.* OLS estimates. The unit of observation is a 1 x 1 degree cell. All continuous predictor variables are standardized. The dependent variable "ln Weibo users (workers)" is the logarithm of the Weibo users identified as worker plus one in Africa in 2014. "Malaria ecology" is a time-invariant index measuring the stability and force of malaria transmission in a cell. "African immunity" is the principal component of the shares of the local African population that exhibit resistance in terms of the sickle cell traits or the negative Duffy antigen. "Geographic controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic controls" include population, night lights, 2G internet coverage and Fibre cable area. "Admin. Area FEs" represent (dummy) controls for level-1 administrative areas at the sub-national level from the Database of Global Administrative Areas (GADM). "Ethnic Group FEs" represent (dummy) controls for culture groups as defined by Murdock (1967) (see also Nunn, 2008). Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

## 4.2 Region of Origin and Immunity of Chinese Workers

A different attenuation mechanism relates to the immunity of Chinese workers. The *Plasmodium falciparum* strain, the most harmful malaria strain and the most widespread in Sub-Saharan Africa, was endemic throughout all the south of China (Tang, Qian and Xu, 1991), other strains, like malaria vivax, were also seasonally prevalent in North China. Transmission was facilitated by suitable climatic conditions and by the presence of Anopheles with the required

[43] See Table A17 in the Appendix.

characteristics. As consequence of numerous eradication campaigns implemented by the Chinese government starting in the 1950s and lasting until the 1990s, *falciparum* is endemic only in the very Southern part of the country (Gething et al., 2011). This implies that Chinese workers from regions in China that had a high malaria exposure historically, such as Hainan and Yunnan, exhibit a relatively higher level of resistance to malaria infections than workers from other regions in China due to similar adaptation mechanisms as mentioned before for the African population. Hence, if malaria exposure in Africa affects Chinese activities, one would expect sorting of workers, with Chinese workers from regions with high exposure to be more frequently employed in areas in Africa that exhibit high malaria exposure.[44]

**Malaria Exposure in China and in Africa.** To investigate this issue, we make use of the information about the home region of Chinese workers from the Weibo data and consider the very same index for exposure to malaria, which focuses on *Plasmodium falciparum*, for China that is used for Africa. This implies that immunity that is acquired through high exposure early in life refers to the same strain of malaria. Moreover, we consider the role of immunity among Chinese workers, as proxied by the prevalence of the corresponding genetic disorders in Chinese prefectures. Figure 6 depicts the distribution of the summary measure of immunity based on a principal component analysis of malaria exposure in terms of the malaria ecology index and the G6PD deficiency across Chinese prefectures.[45] Considerable exposure (and resistance) to malaria is mainly restricted to the South.

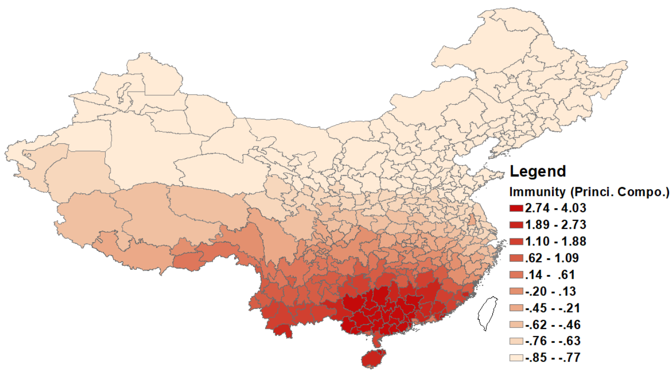



Figure 6: Immunity Level in China – Principal Component

*Note.* The figure displays the principal component of the level of immunity to malaria measured by the malaria ecology index and by G6PD deficiency in Chinese prefectures.

[44] While malaria has been widespread in China historically, with up to 30 million diagnosed per year in the 1940s, recent elimination campaigns have reduced malaria incidence significantly and eliminated it entirely with exception of few provinces such as Hainan and Yunnan, see Lai et al. (2016).

[45] Figure C4a plots the exposure to malaria as reflected in the Malaria Ecology Index for Chinese prefectures and Figure C4c depicts the spatial distribution of the frequency of G6PD deficiency in China. For China, we lack information for the prevalence of the absence of the Duffy antigen. Moreover, the prevalence of the sickle cell trait is almost zero in China as shown in Figure C4b. In the principal component analysis for China, we therefore use malaria ecology and G6PD prevalence.

To test the hypothesis that a greater level of immunity among Chinese workers is associated with a smaller sensitivity of the presence of Weibo users to malaria exposure, we estimate the empirical model

$$Location_{ipcjm} = \beta Immunity_{ip} + X'_{ip}\gamma + \delta_j + \theta_r + \eta_m + \lambda_{jm} + \rho_{rm} + \epsilon_{ipcjm} , \qquad (2)$$

with $Location_{ipcjm}$ denoting the malaria exposure in cell $c$ in an African country $j$ in which a Weibo user (worker) $i$ from a Chinese prefecture $p$ (in province $r$) is located during the month $m$ in which the Weibo posts by user $i$ is largest among the months he stayed in cell $c$. As a baseline, we use the malaria ecology index, in robustness we also use the binary variable that takes value 1 if cell $c$ is malaria-endemic (malaria ecology index equal to or greater than 15). $Immunity_{ip}$ is a measure of average immunity to malaria in user $i$'s prefecture of origin, $p$, in China. We use two of the indices discussed in Section 2.2 – malaria ecology and G6PD deficiency – to measure the level of immunity to malaria in a Chinese prefecture.[46] In addition, we construct a composite indicator of immunity using principal component analysis for malaria ecology and G6PD deficiency. As the principal component reflects the overall degree of immunity, we expect it to have the strongest effects on our outcomes of interest. $X_{ip}$ is a vector of controls for individual characteristics and home prefecture characteristics. Individual characteristics include sex, the number of followers on Weibo, the number of friends on Weibo, user level which reflects a user's active days on Weibo, and the number of Weibo posts a user posted in Africa in 2014. Controls for prefecture characteristics include population, share of tertiary sector, GDP per capita, and GDP growth. The specification also controls for country fixed effects $\delta_j$, Chinese province fixed effects $\theta_r$, as well as month fixed effects $\eta_m$. We also include more flexible country $\times$ month fixed effects, $\lambda_{jm}$, and province $\times$ month fixed effects, $\rho_{rm}$, depending on the model specifications. The standard error $\epsilon_{ipcjm}$ is clustered at the Chinese prefecture level which is the level at which the immunity level of users exhibits variation. The OLS regression is weighted by the inverse of the number of Weibo users (worker) in a cell to treat every individual equally.[47]

Table 6 shows the estimation results for the exposure to malaria in a Weibo user's location in Africa, as measured by the malaria ecology index for the cell (Panel A), or by whether malaria is endemic in the cell (Panel B), for different specifications of the empirical model in the different columns. All columns control for individual characteristics from information in the Weibo account. The results in Columns (1) and (2) reveal that the level of acquired immunity to malaria as proxied by the level of the malaria ecology index in the Weibo user's

[46] We do not use the prevalence of the sickle cell trait since the prevalence is very low across Chinese prefectures and there is essentially no variation.

[47] The purpose of weighting is to rule out the possibility that the results are driven by the fact that population size and immunity level of malaria in southern China is large and that, therefore, there are more Chinese workers with higher levels of immunity in malaria-endemic regions in Africa.

home prefecture has a significantly positive effect on the exposure to malaria in the users' location in Africa. The specification in Column (2) includes country × month and province × month fixed effects to account for all observed and unobserved Africa country specific and Chinese province-specific determinants of a users' location of deployment in Africa in a given month. In terms of magnitude, the coefficients of malaria exposure in Columns (1) and (2) of Panel B imply that a one standard deviation increase in the exposure to malaria in the user's home prefecture increases the probability that a Weibo user (worker) works in malaria-endemic regions by 4%. Notice that this analysis compares across cells with different exposure in Africa and shows that Chinese workers who experienced higher exposure in their home prefectures in China are located in cells in Africa with greater malaria exposure on average. To the extent that these results refer to the same measure of exposure to (*falciparum*) malaria, this is indicative of sorting based on malaria resistance. These results are confirmed for alternative measures. The coefficient of "Genetic immunity (G6PD, China)" in Columns (3) and (4) is also positive and even larger in size. Columns (5) and (6) display similar results for the alternative measure of malaria immunity in a user's home prefecture in China, using a principal component measure.[48]

**Malaria Exposure in Africa and Selection of Chinese Workers.** Another way to test the hypothesis that a greater level of immunity among Chinese workers is associated with a smaller sensitivity of the presence of Weibo users to malaria exposure is to regress the exposure in an individual Chinese worker's home prefecture on the malaria exposure in the cell in which the worker is located in Africa, again using individual-level observations rather than cell-level observations. Notice that the previous exercise showed that greater average exposure in their home prefecture (and hence resistance) of the Weibo users in a cell in Africa is associated with higher exposure in that cell, comparing across cells. Instead, the alternative exercise considered now investigates the relation between malaria exposure in Africa with resistance at the level of individual workers. This provides an arguably more direct test of the hypothesis, as it delivers direct evidence for the composition of Chinese personnel in a cell in Africa, while holding fixed their presence in the cell. In terms of identification, this restricts the focus exclusively on the composition of the Chinese workforce while conditioning on the overall presence, thereby accounting for confounds that might influence the overall presence of Chinese activities in a cell. Also here, the hypothesis is that in regions with higher malaria exposure in Africa one would expect more Chinese individuals with a background of malaria exposure in their prefecture of origin.

The results shown in Table 7 support this hypothesis and document that higher malaria

[48]The results are qualitatively robust to estimating an un-weighted OLS regression or when using as weight the reciprocal of the number of Weibo users (workers) in a cell in Africa (see Tables A18, A19, and A20 in the Appendix).

Table 6: Immunities of Weibo Users (workers) and Malaria Risk of Destination in Africa - User-level Analysis

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Panel A. Full sample | | | | | | |
| Dependent Variables | Malaria ecology index of Weibo user (worker)'s location in Africa | | | | | |
| Malaria exposure | 0.059*** | 0.058*** | | | | |
| (m. ecol., China) | (0.020) | (0.017) | | | | |
| Genetic immunity | | | 0.155* | 0.208*** | | |
| (G6PD, China) | | | (0.086) | (0.076) | | |
| Immunity index | | | | | 0.101*** | 0.103*** |
| (princ. comp., China) | | | | | (0.033) | (0.027) |
| Observations | 3825 | 3825 | 3825 | 3825 | 3825 | 3825 |
| $R^2$ | 0.792 | 0.915 | 0.791 | 0.915 | 0.792 | 0.915 |
| Panel B. Cells with malaria ecology > 0 | | | | | | |
| Dependent Variables | Weibo user (worker) location is malaria endemic | | | | | |
| Malaria exposure | 0.041*** | 0.038*** | | | | |
| (m. ecol., China) | (0.010) | (0.011) | | | | |
| Genetic immunity | | | 0.144*** | 0.097* | | |
| (G6PD, China) | | | (0.046) | (0.053) | | |
| Immunity index | | | | | 0.073*** | 0.065*** |
| (princ. comp., China) | | | | | (0.015) | (0.018) |
| Observations | 3058 | 3058 | 3058 | 3058 | 3058 | 3058 |
| $R^2$ | 0.709 | 0.880 | 0.708 | 0.879 | 0.709 | 0.880 |
| All Panels: | | | | | | |
| User characteristics | Yes | Yes | Yes | Yes | Yes | Yes |
| Home prefecture characteristics | Yes | Yes | Yes | Yes | Yes | Yes |
| African country FEs | Yes | No | Yes | No | Yes | No |
| Chinese province FEs | Yes | No | Yes | No | Yes | No |
| Month FEs | Yes | No | Yes | No | Yes | No |
| African country×Month FEs | No | Yes | No | Yes | No | Yes |
| Chinese province×Month FEs | No | Yes | No | Yes | No | Yes |

*Notes.* Weighted OLS estimates. The unit of observation is a Weibo user (worker). The weight is the reciprocal of the number of Weibo users (workers) in a cell in Africa. All continuous control variables are standardized. "Malaria exposure (m. ecol., China)" is the average malaria ecology index in a user's home prefecture in China. "Genetic immunity (G6PD, China)" is the average level of the median predicted allele frequency for G6PD deficiency in a user's home prefecture in China. "Immunity index (princ. comp., China)" is the principal component of Malaria exposure (m. ecol., China) and Genetic immunity (G6PD, China). "Genetic immunity (sickle cell, China)" is the average level of sickle cell disease in a user's home Chinese prefecture in % of the population. User characteristic controls include Male, # of Followers. # of Friends, User level, # of Posts. Home prefecture characteristic controls are Population, Share of Tertiary sector, GDP per capita and GDP growth. "African country FEs" controls for the country specific effect a user is located in Africa. "Month FEs" corresponds to the fixed effect of the month a user posted most Weibo posts in a cell. "Chinese province FEs" is the fixed effect of a user's home province in China. Standard errors clustered at the user's home Chinese prefecture level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

exposure in Africa is indeed associated with a higher exposure in the Weibo users' home prefecture in China. This suggests that the deployment of workers who are more resistant to malaria implies higher productivity and lower risk of sickness-related employment interruptions.[49] In addition, there is no evidence that the number of projects, Chinese personnel, or other factors related to foreign activities, like night light density, are relevant for the immunity of Chinese workers. This suggests that malaria exposure affects the distribution of foreigners in Africa,

[49] This finding is robust to the use of a continuous measure of malaria exposure (Table A21), to controlling for HIV and Tsetse fly prevalence (Table A22), to using a binary measure of malaria exposure below/above the median (Table A23), to considering different measures of immunity (Table A24), and to the consideration of all Weibo users and the use of binary measures of endemic and epidemic malaria exposure (Table A17).

Table 7: Malaria in Africa and Immunity of Chinese

| Dependent Variables | Chinese immunity (malaria ecology - binary variable) | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Malaria ecology | 0.059*** | 0.054*** | 0.062*** | 0.062*** | 0.060*** | 0.059*** |
| | (0.019) | (0.019) | (0.018) | (0.018) | (0.019) | (0.019) |
| African immunity | | | -0.044*** | -0.043*** | -0.040*** | -0.040*** |
| (principal component) | | | (0.015) | (0.015) | (0.015) | (0.015) |
| Number of projects | | | | 0.001 | 0.000 | -0.001 |
| | | | | (0.002) | (0.002) | (0.003) |
| ln Weibo user (worker) | | | | | 0.006 | 0.005 |
| | | | | | (0.006) | (0.006) |
| Number of aid projects | | | | | | 0.005 |
| | | | | | | (0.012) |
| Economic controls | No | Yes | Yes | Yes | Yes | Yes |
| User characteristics | Yes | Yes | Yes | Yes | Yes | Yes |
| Home prefecture characteristics | Yes | Yes | Yes | Yes | Yes | Yes |
| Chinese province×Month FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| African country×Month FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3889 | 3889 | 2579 | 2579 | 2579 | 2364 |
| $R^2$ | 0.867 | 0.868 | 0.899 | 0.899 | 0.899 | 0.917 |

*Notes.* Weighted OLS estimates. The unit of observation is a Weibo user (worker). The weight is the reciprocal of the number of Weibo users (workers) in a cell in Africa. All continuous predictor variables are standardized. The dependent variable "Chinese immunity (malaria ecology - binary variable)" is a binary variable equalling to 1 if the average malaria ecology index in a user's home prefecture in China is greater than zero, 0 otherwise. "Malaria ecology" is a time-invariant index measuring the stability and force of malaria transmission in Africa. "Economic controls" include population, night lights, 2G internet coverage and Fibre cable area. The unit of observation is a 1 x 1 degree cell. User characteristic controls include Male, # of Followers. # of Friends, User level, # of Posts. Home prefecture characteristic controls are Population, Share of Tertiary sector, GDP per capita and GDP growth. "African country FEs" controls for the country specific effect a user is located in Africa. "Month FEs" corresponds to the fixed effect of the month a user posted most Weibo posts in a cell. "Chinese province FEs" is the fixed effect of a user's home province in China. Standard errors clustered at the cell level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

alleviating concerns about endogeneity in the other results.

Taken together, the results of Tables 5 and 7 also support the hypothesis that the employment of local workers interacts with that of Chinese workers, particularly in areas of high malaria risk and greater resistance of the local population. We find a consistently negative effect of greater immunity among the African local population on the prevalence of malaria in the home regions of Chinese workers. This could be an indication for less selective deployment in these areas or of an attenuation of the effect that works through complementarities.[50] The results are consistent with our hypothesis that the risk of contracting malaria affects Chinese activities in Africa by influencing the location of Chinese workers, either through the location decision of these workers or through the recruitment and deployment policies of Chinese companies.

## 4.3 The Role of Anti-malaria Policies

A last question refers to the role of anti-malaria policies as a moderating factor. Improved protection and better treatment facilities are expected to reduce the effects of malaria exposure

[50] Unreported results reveal a positive but insignificant coefficient for the interaction term of malaria ecology with African immunity.

on the presence of Chinese workers. Just as a population with greater prevalence of immunity is less likely to suffer from epidemic outbreaks that might harm temporary migrants from China, prompt access to treatment and therapies reduces the risk of epidemic outbreaks and implies lower health risks for temporary migrants and, in particular, Chinese personnel. In fact, recent initiatives to reduce malaria were successful in reducing the number of malaria-related fatalities in Africa. These initiatives mainly involved the dissemination of artemisinin-based combination therapy coverage (ACT) and insecticide treated bednets. While these are mostly very local policies targeted at a resource-constrained population and potentially less relevant for Chinese workers, they still serve as proxy for the possibility of rapid treatment in case of infection and the availability of preventive measures, respectively. To explore whether the prevalence and availability of anti-malaria treatment affects the presence of Chinese workers, we replicate the previous analysis while controlling for a binary measure that is constructed as taking value 1 if the average coverage of each of these measures in a given cell is above the median across cells.[51]

Table 8 presents results that correspond to an extended version of the specifications in Column (4) in Table 3 that accounts for a coverage with policies above the cell median in 2014, and, in addition, for the value of Chinese health aid in a given cell. The estimates reveal a smaller coefficient for the main effect of malaria exposure once anti-malaria policies are included, regardless of the specification with country fixed effects or administrative area fixed effects. Moreover, a greater coverage with anti-malaria policies is correlated with a positive effect on the presence of Chinese workers in a cell. Additionally controlling for the value of Chinese health aid leaves these findings unaffected, but also exerts a positive effect on the presence of Chinese personnel.[52] The estimates are similar or even stronger when considering all Weibo users rather than focusing on workers.[53] To rule out that this effect is driven by Chinese health aid in cells with high malaria exposure, we conducted additional analysis. The results reveal a positive relation between the presence of Chinese workers and Chinese health aid, but they also document a negative effect of malaria exposure on the location of Chinese health aid projects.[54] This suggests that the effect of malaria exposure also extends to the domain of health projects.

[51]The data have been collected by Bhatt et al. (2015) and were retrieved from http://www.map.ox.ac.uk. ACT coverage measures the annual proportion of febrile children treated with ACT. The measure is predicted through a LGM model in order to deal with the sparse temporal and spatial resolution of the original data. The prediction of febrile children treated is a function of ACT pills per capita distributed in a given country and year, a nonparametric time trend, country specific random effects, and Africa macro-region specific random effects.

[52]We account for the value of health aid rather than the number of projects to better reflect the access to health measures. It is rather unlikely that the policy measures suffer from reverse causality since policy coverage, e.g., for ACT, is mainly organized by local health providers. In addition, Chinese health aid is mainly about facilities and not medication.

[53]See Appendix Table A25.

[54]See Tables A26 and A27 in the Appendix.

Table 8: Effects of Anti-malaria Policies

| Dependent Variables | ln Weibo users (workers) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Malaria ecology | -0.147*** | -0.140*** | -0.176*** | -0.168*** |
| | (0.029) | (0.028) | (0.037) | (0.037) |
| Policy (above median) | 0.102* | 0.092* | 0.126* | 0.112* |
| | (0.052) | (0.052) | (0.066) | (0.063) |
| Value of health aid | | 0.053*** | | 0.043*** |
| | | (0.012) | | (0.014) |
| Geographic Controls | Yes | Yes | Yes | Yes |
| Economic Controls | Yes | Yes | Yes | Yes |
| Ethnic Group FEs | Yes | Yes | Yes | Yes |
| Country FEs | Yes | Yes | No | No |
| Admin. Area FEs | No | No | Yes | Yes |
| $N$ | 1656 | 1656 | 1656 | 1656 |
| $R^2$ | 0.536 | 0.542 | 0.745 | 0.749 |

*Notes.* OLS estimates. The unit of observation is a 1 x 1 degree cell. All continuous predictor variables are standardized. "Policy (above median)" is a binary variable equal to 1 if the mean of anti-malaria policies - Coverage of ACT therapy and bednets in 2014 - is above the median. "Value of health aid" is the value of Chinese health aid projects in Africa from 2010 to 2014, data from AidData. "Geographic controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic controls" include population, night lights, 2G internet coverage and Fibre cable area. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

# 5 Conclusion

This paper presents evidence for the influence of the disease environment on foreign economic activities by documenting that local malaria exposure has a significantly negative impact on the local presence of Chinese aid and infrastructure projects, and particularly on the presence of Chinese workers in Africa. The analysis is based on a combination of information about malaria exposure with geographically disaggregated data for Chinese activities in Africa and with novel geo-localized data on the use of the Weibo social network among Chinese workers.

The results provide evidence for a negative effect of malaria exposure on Chinese foreign activities. This effect works at least partly through its effect on the deployment of Chinese personnel. The influence of malaria exposure varies across sectors, presumably reflecting skill requirements or other features that determine the degree to which Chinese workers can be substituted by locals. The findings also suggest that Chinese personnel is located selectively in regions of lower malaria exposure. The finding that the malaria exposure in their location in Africa predicts the malaria exposure in the prefecture of origin of Chinese workers provides evidence for sorting of Chinese personnel by malaria resistance. Likewise, the availability of medical malaria treatment mitigates the effect of high malaria exposure.

The results have relevant implications for policy. In particular, the findings provide support for a channel of malaria exposure that differs from earlier arguments of an indirect and historically determined effect of disease exposure on development that works through culture or institutions. In fact, by conditioning on both cultural groups and sub-national administrative units, our evidence suggests that the malaria channel has an effect that operates above and beyond the long-run influence of the disease environment. From a policy perspective, this channel, which works through foreign activities, provides more scope for policy interventions than channels that work through culture or historical institutions. Such interventions include policies aimed at fostering the training and employment of the local African population or health policies aimed at protecting or treating migrant workers. Second, the results are limited to malaria. Similar results are not found generically for any diseases affecting migrants. Notably, we do not find any significant effect of HIV on the distribution of Chinese activities nor on the region of origin of Chinese workers. This calls for policy interventions that focus on malaria and its specific epidemiological features, rather than infectious diseases in general.

The finding that malaria determines the location of Chinese workers in Africa and, through this channel, the uneven prospects of the massive Chinese engagement in Africa unfolds a novel channel through which malaria impacts development in Africa. While this evidence establishes a novel link between malaria and comparative development at the disaggregate level, more work is needed for a better understanding of the relationship between infrastructure investment and aid. The present analysis relied on cross-sectional variation as consequence of the lack of Weibo data as proxy for the presence of Chinese persons for more than one year. Promising avenues for future work concern the role of interactions between Chinese infrastructure project contracting and health interventions and the interplay of long-run exposure as studied here with short-run variation in conditions that favor epidemic outbreaks of malaria. More work is also needed to quantify the relevance of the indirect effect of malaria exposure on development through its negative influence on infrastructure projects. The moderate heterogeneity of the effects for aid projects and infrastructure contracting suggests that the effects on Chinese-financed aid is more pronounced, which might reflect a greater strategic component in the realization of these projects than in projects financed by other donors like the World Bank. However, more work is needed to disentangle this from heterogeneity that is the result of differences in data quality.

## References


**Acemoglu, Daron, and Simon Johnson.** 2007. "Disease and development: the effect of life expectancy on economic growth." *Journal of Political Economy*, 115(6): 925–985.

**Acemoglu, Daron, Simon Johnson, and James A Robinson.** 2001. "The colonial origins of comparative development: An empirical investigation." *American Economic Review*, 91(5): 1369–1401.

**African Development Bank.** 2018. "African Economic Outlook 2018."

**African Development Bank.** 2020. "African Economic Outlook 2020."

**Alesina, Alberto, Stelios Michalopoulos, and Elias Papaioannou.** 2016. "Ethnic inequality." *Journal of Political Economy*, 124(2): 428–488.

**Alsan, Marcela.** 2015. "The Effect of the Tsetse Fly on African Development." *American Economic Review*, 105(1): 382–410.

**Anaxagorou, Christiana, Georgios Efthyvoulou, and Vassilis Sarantides.** 2020. "Electoral motives and the subnational allocation of foreign aid in sub-Saharan Africa." *European Economic Review*, 127: 103430.

**Arbatlı, Cemal Eren, Quamrul H Ashraf, Oded Galor, and Marc Klemp.** 2020. "Diversity and conflict." *Econometrica*, 88(2): 727–797.

**Asiedu, Elizabeth, Yi Jin, and Isaac K Kanyama.** 2015. "The impact of HIV/AIDS on foreign direct investment: Evidence from Sub-Saharan Africa." *Journal of African Trade*, 2(1-2): 1–17.

**Azemar, Celine, and Rodolphe Desbordes.** 2009. "Public governance, health and foreign direct investment in Sub-Saharan Africa." *Journal of African Economies*, 18(4): 667–709.

**Bartolucci, Cristian, Claudia Villosio, and Mathis Wagner.** 2018. "Who migrates and why? Evidence from Italian administrative data." *Journal of Labor Economics*, 36(2): 239–281.

**Bhatt, S., S. Weiss, E. Cameron, D. Bisanzio, B. Mappin, U. Dalrymple, K. E. Battle, C. L. Moyes, A. Henry, P. A. Eckhoff, E. A. Wenger, O. Briet, M. A. Penny, T. A. Smith, A. Bennett, J. Yukich, T. P. Eisele, J. T. Griffin, C. A. Fergus, M. Lynch, F. Lindgren, J. M. Cohen, C. L. J. Murray, D. L. Smith, S. I. Hay, R. E. Cibulskis, and P. W. Gething.** 2015. "The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015." *Nature*, 526(7572): 207–211.

**Bleakley, Hoyt.** 2010. "Malaria eradication in the Americas: A retrospective analysis of childhood exposure." *American Economic Journal: Applied Economics*, 2(2): 1–45.

**Bluhm, Richard, Andreas Fuchs, Dreher Axel Strange, Austin, Bradley Parks, and Michael J. Tierney.** 2020. "Connective Financing: Chinese Infrastructure Projects and the Diffusion of Economic Activity in Developing Countries." *CESifo Working Paper*, 8344.

**Bluhm, Richard, Martin Gassebner, Sarah Langlotz, and Paul Schaudt.** 2021. "Fueling conflict? (De)Escalation and bilateral aid." *Journal of Applied Econometrics*, 36(2): 2447–261.

**Brautigam, Deborah.** 2009. *The Dragon's Gift.* New York:Oxford University Press.

**Busse, M., C. Erdogan, and H. Mühlen.** 2016. "China's Impact on Africa – The Role of Trade, FDI and Aid." *Kyklos*, 69(2): 228–262.

**Cervellati, Matteo, and Uwe Sunde.** 2011. "Disease and development: The role of life expectancy reconsidered." *Economics Letters*, 113(3): 269–272.

**Cervellati, Matteo, Elena Esposito, and Uwe Sunde.** 2017. "Long Term Exposure to Malaria and Development: Disaggregate Evidence for Contemporaneous Africa." *Journal of Demographic Economics*, 83(1): 129–148.

**Cervellati, Matteo, Elena Esposito, and Uwe Sunde.** 2021. "Epidemic Shocks and Civil Violence." *Review of Economics and Statistics, forthcoming.*

**Cervellati, Matteo, Elena Esposito, Uwe Sunde, and Simona Valmori.** 2018. "Long-term exposure to Malaria and Violence in Africa." *Economic Policy*, 33(95): 403–446.

**Chen, Chuan, Andrea Goldstein, and Ryan J Orr.** 2009. "Local operations of Chinese construction firms in Africa: An empirical survey." *International Journal of Construction Management*, 9(2): 75–89.

**Chen, William, David Dollar, and H. Tang.** 2018. "Why is China investing in Africa? Evidence from the firm level." *World Bank Economic Review*, 32(3): 610–632.

**Cheung, Yin-Wong, Jakob de Haan, Xingwang Qian, and Shu Yu.** 2012. "China's Outward Direct Investment in Africa." *Review of International Economics*, 20(2): 201–220.

**Colella, Fabrizio, Rafael Lalive, Seyhun Orcan Sakalli, and Mathias Thoenig.** 2019. "Inference with Arbitrary Clustering." *IZA Discussion Paper*, 12584.

**Conley, Timothy G.** 1999. "GMM Estimation with Cross-sectional Dependence." *Journal of Econometrics*, 92: 1–45.

**Cook, Seth, Jixia Lu, Henry Tugenthat, and Dawit Alemu.** 2016. "Chinese migrants in Africa: Facts and Fictions from the Agri-Food Sector in Ethiopia and Ghana." *World Development*, 81(1): 61–70.

**Corkin, Lucy, and Christopher Burke.** 2006. "China's interest and activity in Africa's construction and infrastructure sectors." In *Report prepared for DFID China.* Vol. 2006. Centre for Chinese Studies, Stellenbosch, South Africa.

**Crost, Benjamin, Joseph Felter, and Patrick Johnston.** 2014. "Aid under Fire: Development Projects and Civil Conflict." *American Economic Review*, 104(6): 1833–1856.

**Cutler, David, Winnie Fung, Michael Kremer, Monica Singhal, and Tom Vogl.** 2010. "Early-life malaria exposure and adult outcomes: Evidence from malaria eradication in India." *American Economic Journal: Applied Economics*, 2(2): 72–94.

**Depetris-Chauvin, Emilio, and David N Weil.** 2018. "Malaria and Early African Development: Evidence from the Sickle Cell Trait." *The Economic Journal*, 128(610): 1207–1234.

**Desmet, Klaus, and Vernon J. Henderson.** 2015. "The Geography of Development Within Countries." In *Handbook of Regional and Urban Economics.* Vol. 5, , ed. Gilles Duranton, J. Vernon Henderson and William C. Strange, Chapter 22, 1457–1517. Elsevier.

**Desmet, Klaus, Dávid Krisztián Nagy, and Esteban Rossi-Hansberg.** 2018. "The Geography of Development." *Journal of Political Economy*, 126(3): 903–983.

**Dollar, David.** 2016. *China's Engagement with Africa: From Natural Resources to Human Resources.* Washington, D.C.:John L. Thornton China Center, Brookings Institution.

**Donou-Adonsou, Ficawoyi, and Sokchea Lim.** 2018. "On the importance of Chinese investment in Africa." *Review of Development Finance*, 8(1): 63–73.

**Dreher, Axel, and Andreas Fuchs.** 2015. "Rogue Aid? An Empirical Analysis of China's Aid Allocation." *Canadian Journal of Economics*, 48: 988–1023.

**Dreher, Axel, Andreas Fuchs, Bradley C. Parks, Austin Strange, and Michael J. Tierney.** 2018. "Apples and Dragon Fruits: The Determinants of Aid and Other Forms of State Financing from China to Africa." *International Studies Quarterly*, 62: 182–194.

**Dreher, Axel, Andreas Fuchs, Roland Hodler, Bradley C. Parks, Paul A. Raschky, and Michael J. Tierney.** 2019. "African Leaders and the Geography of China's Foreign Assistance." *Journal of Development Economics*, 140: 44–71.

**Dreher, Axel, Andreas Fuchs, Roland Hodler, Bradley C. Parks, Paul A. Raschky, and Michael J. Tierney.** 2021*a*. "Is Favoritism a Threat to Chinese Aid Effectiveness? A Subnational Analysis of Chinese Development Projects." *World Development*, 139: 105291.

**Dreher, Axel, Andreas Fuchs, Strange Austin Parks, Bradley C., and Michael J. Tierney.** 2021*b*. "Aid, China, and Growth: Evidence from a New Global Development Finance Dataset." *American Economic Journal: Economic Policy*, 13(2): 135–174.

**Dreher, Axel, and Sarah Langlotz.** 2020. "Aid and growth: New evidence using an excludable instrument." *Canadian Journal of Economics*, 53(3): 1162–1198.

**Enke, Benjamin.** 2019. "Kinship, cooperation, and the evolution of moral systems." *Quarterly Journal of Economics*, 134(2): 953–1019.

**ENR.** 2019. *ENR 2019 Top 250 International Contractors.* Engineering News Record.

**Ferreira, Ana, Ivo Marguti, Ingo Bechmann, Viktória Jeney, Ângelo Chora, Nuno R Palha, Sofia Rebelo, Annie Henri, Yves Beuzard, and Miguel P Soares.** 2011. "Sickle hemoglobin confers tolerance to Plasmodium infection." *Cell*, 145(3): 398–409.

**Galiani, Sebastian, Stephen Knack, Lixin Colin Xu, and Ben Zou.** 2017. "The effect of aid on growth: evidence from a Quasi-experiment." *Journal of Economic Growth*, 22(1): 1–33.

**Gallup, John Luke, and Jeffrey D Sachs.** 2001. "The economic burden of malaria." *American Journal of Tropical Medicine and Hygiene*, 64(1_suppl): 85–96.

**Galor, Oded, and Ömer Özak.** 2016. "The agricultural origins of time preference." *American Economic Review*, 106(10): 3064–3103.

**Gennaioli, Nicola, Rafael La Porta, Florencio Lopez-de Silanes, and Andrei Shleifer.** 2013. "Human Capital and Regional Development." *Quarterly Journal of Economics*, 105: 164.

**Gething, Peter W, Anand P Patil, David L Smith, Carlos A Guerra, Iqbal RF Elyazar, Geoffrey L Johnston, Andrew J Tatem, and Simon I Hay.** 2011. "A new world malaria map: Plasmodium falciparum endemicity in 2010." *Malaria Journal*, 10(1): 378.

**Ghosh, Sucharita, and Francesco Renna.** 2015. "The relationship between communicable diseases and FDI flows: an empirical investigation." *The World Economy*, 38(10): 1574–1593.

**Grogger, Jeffrey, and Gordon H. Hanson.** 2011. "Income maximization and the selection and sorting of international migrants." *Journal of Development Economics*, 95(1): 42–57.

**Guindo, Aldiouma, Rick M Fairhurst, Ogobara K Doumbo, Thomas E Wellems, and Dapa A Diallo.** 2007. "X-linked G6PD deficiency protects hemizygous males but not heterozygous females against severe malaria." *PLoS Medicine*, 4(3): e66.

**Habiyaremye, Alexis.** 2013. "'Angola-mode' Trade Deals and the Awakening of African Lion Economies." *Journal of Economic Perspectives*, 25(4): 636–647.

**Henderson, Vernon J., Zmarak Shalizi, and Anthony J. Venables.** 2001. "Geography and development." *Journal of Economic Geography*, 1(1): 81–105.

**Hong, Sok Chul.** 2011. "Malaria and economic productivity: a longitudinal analysis of the American case." *The Journal of Economic History*, 71(3): 654–671.

**Howes, Rosalind E, Anand P Patil, Frédéric B Piel, Oscar A Nyangiri, Caroline W Kabaria, Peter W Gething, Peter A Zimmerman, Céline Barnadas, Cynthia M Beall, Amha Gebremedhin, et al.** 2011. "The global distribution of the Duffy blood group." *Nature communications*, 2: 266.

**Howes, Rosalind E, Frédéric B Piel, Anand P Patil, Oscar A Nyangiri, Peter W Gething, Mewahyu Dewi, Mariana M Hogg, Katherine E Battle, Carmencita D Padilla, J Kevin Baird, et al.** 2012. "G6PD deficiency prevalence and estimates of affected populations in malaria endemic countries: a geostatistical model-based map." *PLoS Medicine*, 9(11): e1001339.

**Huang, Meibo, and Peiqiang Ren.** 2013. "A study on the employment effect of Chinese investment in South Africa." *mimeo, Stellenbosch University, Centre for Chinese Studes.*

**Huang, Zhengli, and Xiangming Chen.** 2016. "Is China Building Africa?" *European Financial Review*, 2016(6): 7–13.

**ICA.** 2015. *Infrastructure Financing Trends in Africa – 2015 Report.* The Infrastructure Consortium for Africa.

**ICA.** 2018. *Infrastructure Financing Trends in Africa – 2018 Report.* The Infrastructure Consortium for Africa.

**Jansson, Johanna.** 2013. "The Sicomines agreement revisited: prudent Chinese banks and risk-taking Chinese companies." *Review of African Political Economy*, 40(135): 152–162.

**Kelly, Morgan.** 2019. "The Standard Errors of Persistence." *CEPR Discussion Paper*, 13783.

**Kiszewski, Anthony, Andrew Mellinger, Andrew Spielman, Pia Malaney, Sonia Ehrlich Sachs, and Jeffrey Sachs.** 2004. "A global index representing the stability of malaria transmission." *The American Journal of Tropical Medicine and Hygiene*, 70(5): 486–498.

**Kodongo, Odongo, and Kalu Ojah.** 2016. "Does infrastructure really explain economic growth in Sub-Saharan Africa?" *Review of Development Finance*, 6(2): 105–125.

**Kudamatsu, Masayuki, Torsten Persson, and David Strömberg.** 2012. "Weather and Infant Mortality in Africa." *CEPR Discussion Paper*, 9222.

**Kwiatkowski, Dominic P.** 2005. "How malaria has affected the human genome and what human genetics can teach us about malaria." *The American Journal of Human Genetics*, 77(2): 171–192.

**Lai, Shengjie, Nicola A. Wardrop, Zhuojie Huang, Claudio Bosco, Junling Sun, Tomas Bird, Amy Wesolowski, Sheng Zhou, Qian Zhang, Canjun Zheng, Zhongjie Li, Andrew J. Tatem, and Hongjie Yua.** 2016. "Plasmodium falciparum malaria importation from Africa to China and its mortality: an analysis of driving factors." *Scientific Reports*, 6(39524).

**Lell, Bertrand, Benjamin Mordmüller, Jean-Claude Dejon Agobe, Josiane Honkpehedji, Jeannot Zinsou, Juliana Boex Mengue, Marguerite Massinga Loembe, Ayola Akim Adegnika, Jana Held, Albert Lalremruata, et al.** 2018. "Impact of sickle cell trait and naturally acquired immunity on uncomplicated malaria after controlled human malaria infection in adults in Gabon." *The American journal of tropical medicine and hygiene*, 98(2): 508–515.

**Lucas, Adrienne M.** 2010. "Malaria eradication and educational attainment: evidence from Paraguay and Sri Lanka." *American Economic Journal: Applied Economics*, 2(2): 46–71.

**Lyons, Candice L, Maureen Coetzee, and Steven L Chown.** 2013. "Stable and fluctuating temperature effects on the development rate and survival of two malaria vectors, Anopheles arabiensis and Anopheles funestus." *Parasites & vectors*, 6(1): 104.

**Mayshar, Joram, Omer Moav, Zvika Neeman, and Luigi Pascali.** 2021. "The Origin of the State: Land Productivity or Appropriability?" *Journal of Political Economy*, forthcoming.

**Michalopoulos, Stelios, Alireza Naghavi, and Giovanni Prarolo.** 2018. "Trade and Geography in the Spread of Islam." *The Economic Journal*, 128(616): 3210–3241.

**Moraga, Juan F. H.** 2013. "Understanding different migrant selection patterns in rural and urban Mexico." *Journal of Development Economics*, 103(2): 182–201.

**Murdock, George Peter.** 1959. *Africa: Its Peoples and Their Culture History.* New York:McGraw-Hill.

**Murdock, George Peter.** 1967. *Ethnographic Atlas.* Pittsburgh:University of Pittsburgh Press.

**Nissanke, Machiko, and Marie Söderberg.** 2011. "The Changing Landscape in Aid Relationships in Africa: Can China's engagement make a difference to African development?" *Swedish Institute for International Affairs, UI Papers*, 2011/2.

**Nunn, Nathan.** 2008. "The Long-Term Effects of Africa's Slave Trades." *Quarterly Journal of Economics*, 123(1): 139–176.

**Nunn, Nathan, and Nancy Qian.** 2014. "US Food Aid and Civil Conflict." *American Economic Review*, 104(6): 1630–1666.

**Piel, Frederic B, Simon I Hay, Sunetra Gupta, David J Weatherall, and Thomas N Williams.** 2013. "Global burden of sickle cell anaemia in children under five, 2010–2050: modelling based on demographics, excess mortality, and interventions." *PLoS Medicine*, 10(7): e1001484.

**Purohit, Prasanta, Pradeep K Mohanty, Siris Patel, Padmalaya Das, Jogeswar Panigrahi, and Kishalaya Das.** 2018. "Comparative study of clinical presentation and hematological indices in hospitalized sickle cell patients with severe Plasmodium falciparum malaria." *Journal of infection and public health*, 11(3): 321–325.

**Qin, Bei, and Yanhui Wu.** 2017. "Why Does China Allow Freer Social Media? Protests Versus Surveillance and Propaganda." *Journal of Economic Perspectives*, 31(1): 117–140.

**Sachs, Jeffrey, and Pia Malaney.** 2002. "The economic and social burden of malaria." *Nature*, 415(6872): 680.

**Sanghi, Apurva, and Dylan Johnson.** 2016. *Deal or no deal: strictly business for China in Kenya?* World Bank.

**Shu, Wei Sheng.** 1932. "Medical and Health Reports: A Preliminary Survey of the Endemicity of Malaria in Nanking, Soochow, Hangchow, Wukong and Districts along the Yangtze Valley." *Chinese Medical Journal*, 46: 719–726.

**Tang, Lin-Hua, HL Qian, and SH Xu.** 1991. "Malaria and its Control in the People's Republic of China." *Southeast Asian Journal of Tropical Medicine and Public Health*, 22(4): 467–476.

**Tang, Xiaoyang.** 2010. "Bulldozer or locomotive? The impact of Chinese enterprises on the local employment in Angola and the DRC." *Journal of Asian and African Studies*, 45(3): 350–368.

**Tang, Xiaoyang.** 2016. "Does Chinese Employment Benefit Africans? Investigating Chinese Enterprises and their Operations in Africa." *African Studies Quarterly*, 16: 107–128.

**Tanser, Frank C, Brian Sharp, and David Le Sueur.** 2003. "Potential effect of climate change on malaria transmission in Africa." *The Lancet*, 362(9398): 1792–1798.

**Taylor, S. M., C. M. Parobek, and R. M. Fairhurst.** 2012. "Haemoglobinopathies and the clinical epidemiology of malaria: a systematic review and meta-analysis." *Lancet Infectious Diseases*, 12: 457–468.

**UNCTAD.** 2016. *World Inversment Report 2016.* World Health Organization.

**Uyoga, Sophie, Alex W. Macharia, Carolyne M. Ndila, Gideon Nyutu, Mohammed Shebe, Kennedy O. Awuondo, Neema Mturi, Norbert Peshu, Benjamin Tsofa, J. Anthony G. Scott, Kathryn Maitland, and Thomas N. Williams.** 2019. "The indirect health effects of malaria estimated from health advantages of the sickle cell trait." *Nature Communications*, 10(856).

**Weibo, Sina.** 2015. "Weibo User Development Report 2015." , (https://data.weibo.com/report/reportDetail?id=297).

**Weil, David N.** 2007. "Accounting for the effect of health on economic growth." *Quarterly Journal of Economics*, 122(3): 1265–1306.

**WHO.** 2015. *Achieving the malaria MDG target: reversing the incidence of malaria 2000-2015.* World Health Organization.

**WHO.** 2017. *World malaria report 2017.* World Health Organization.

**WHO.** 2019. *World malaria report 2019.* World Health Organization.

**World Bank.** 2017. "Africa's Pulse: An Analysis of Issues Shaping Africa's Economic Future." *World Bank Report*, 15(Office of the Chief Economist for the Africa Region).

**Zhao, Zhen Yu, and Li Yin Shen.** 2008. "Are Chinese contractors competitive in international markets?" *Construction Management and Economics*, 26(3): 225–236.

**Zheng, Siqi, Jianghao Wang, Cong Sun, Xiaonan Zhang, and Matthew E Kahn.** 2019. "Air pollution lowers Chinese urbanites' expressed happiness on social media." *Nature Human Behaviour*, 3(3): 237.

# Supplementary Appendix
# Malaria and Chinese Economic Activities in Africa*


Matteo Cervellati† University of Bologna CEPR, IZA

Elena Esposito‡ HEC University of Lausanne

Uwe Sunde§ LMU Munich CEPR, IZA

Song Yuan¶ University of Warwick



**Abstract**

This file contains additional information on data sources, variable construction, descriptive statistics, and additional tables and figures referenced in the main text.


## Contents




*The authors are grateful to Axel Dreher, Kevin Donovan, Roland Hodler, Jonas Hjort, Omer Moav, Sharun Mukand and especially James Fenske, as well as seminar participants at the German Institute of Global and Area Studies, Hamburg, University of Warwick, University of Lausanne, Jinan University, Sichuan University, and the audiences at the CSAE Annual Conference (Oxford), the Royal Economic Society Annual Conference, The European Winter Meeting of the Econometric Society (Nottingham) and the Chinese Young Economist Society (Guangzhou) for helpful comments. The authors would like to thank Jianghao Wang for sharing the Weibo Data and Andrea Tesei for sharing the 2G signal coverage data.



†Department of Economics, University of Bologna. Email: m.cervellati@unibo.it

‡Department of Business and Economics, University of Lausanne. Email: elena.esposito.1@unil.ch

§Department of Economics, LMU Munich. Email: Uwe.Sunde@econ.lmu.de

¶Department of Economics, University of Warwick. Email: s.yuan.2@warwick.ac.uk

# A Data Sources

Table D1: Data Sources and Description of Main Variables of Interest

**Variable Description and Data Sources**

**Construction projects**

*Construction project value*: the contracted value of infrastructure construction projects undertaken by Chinese companies in 2013. Source: *China Overseas Engineering Projects And labour Service Yearbook 2014*, the Minister of Commerce of China.

*Construction project number*: the number of contracted infrastructure construction projects undertaken by Chinese companies in 2013.

*Number of building projects*:the number of contracted public building projects undertaken by Chinese companies in 2013.

*Number of hydro projects*: the number of contracted hydro construction projects undertaken by Chinese companies in 2013.

*Number of transport projects*: the number of contracted transportation construction projects including bridge, road, railway and harbour undertaken by Chinese companies in 2013.

*Number of utility projects*: the number of contracted utility construction projects undertaken by Chinese companies in 2013.

**Aid data**

*Number of aid Projects*: The number of all types of Chinese aid projects to Africa at or below the administrative 1 (province or region) level from 2010 to 2014. Source: Bluhm et al. (2020).
*Number of infrastructure aid projects*: The number of Chinese aid projects that are devoted to infrastructure construction such as road, dams, or public buildings and so on at or below the administrative 1 level in Africa from 2010 to 2014. Source: Bluhm et al. (2020).
*Number of aid projects involving Chinese*: The number of Chinese aid projects classified as infrastructure projects or with the flow type of "free-standing technical assistance" at or below the administrative 1 level in Africa from 2010 to 2014. Source: Bluhm et al. (2020).
*Number of ODA-like aid projects*: The number of Chinese aid projects aimed at development as defined by AidData at or below the administrative 1 level in Africa from 2010 to 2014. Source: Bluhm et al. (2020).
*Number of aid projects in donations*: The number of Chinese aid projects in the form of monetary and material donations, scholarships, and training in China at or below the administrative 1 level in Africa from 2010 to 2014. Source: Bluhm et al. (2020).

**Weibo posts**

*ln Weibo User (worker)*: the natural logarithm of the number of Weibo users who are identified as workers in 2014. Source: Weibo.com, SINA Corporation.

*ln Weibo User (unidentified)*: the natural logarithm of the number of Weibo users whose occupations are unidentified.

*ln Weibo User (worker) with(out) leisure*: the natural logarithm of the number of Weibo users who are identified as workers and (don't) have posts unrelated to work in 2014.

*ln Weibo User (worker) in mining*: the natural logarithm of the number of Weibo users who are identified as workers in mining sector in 2014.

Continued on next page

**Variable Description and Data Sources**

*ln Weibo User (worker) in construction*: the natural logarithm of the number of Weibo users who are identified as workers in construction sector in 2014.

*ln Weibo User (worker) in manufacture*: : the natural logarithm of the number of weibo users who are identified as workers in manufacture sector in 2014.

*ln Weibo User (worker) in business*: the natural logarithm of the number of weibo users who are identified as workers in manufacture sector in 2014.

*ln Weibo User (worker) in hydropower*: the natural logarithm of the number of weibo users who are identified as workers in hydro power sector in 2014.

*ln Weibo User (worker) in telecom*: the natural logarithm of the number of Weibo users who are identified as workers in telecommunication sector in 2014.

**Malaria**

*Malaria stability index*: a index measuring the force and stability of malaria transmission based on biological characteristics of diverse vector mosquitoes and their interaction with climate. Source: Kiszewski et al. (2004).

*Malaria endemic*: a binary index for high malaria exposure taking value one for cells with an average malaria ecology index equal to or larger than 15. Source: Kiszewski et al. (2004).

*Malaria epidemic*: a binary index for low malaria exposure taking value one for cells with an average malaria ecology index greater than 0 and smaller than 15. Source: Kiszewski et al. (2004).

*Malaria suitable months*: average number of months which are suitable for malaria transmission of each year from 1998 to 2012, based on an algorithm requiring determined threshold of temperature and precipitation in up to 12 months before. The algorithm was computed with temperature and precipitation data from ECMWF ERA-Interim dataset ECMWF ERA-Interim dataset.

*Malaria temperature stability*: the average temperature suitability for Plasmodium falciparum transmission globally. It is a relative measure of number of infectious mosquitos in an environment defined by the temperature profile at that location. Source: Gething et al. (2011*b*).

*Malaria climate stability 2000*: average Falciparum Suitability in the 1x1 degree grid cell predicted using temperature. Source: Gething et al. (2011*a*).

*McCord Falciparum CRU 2000*: average Malaria Stability Index in the 1x1 degree grid cell predicted using long-term averages of temperature and precipitation and ignoring mosquito characteristics. Source: Cervellati, Chiovelli and Esposito (2019).

**Cell specific controls**

*Temperature*: average monthly temperature for year 2013 or 2014. Source: Terrestrial Air Temperature: 1900-2014 Gridded Monthly Time Series (Version 4.01) Matsuura and Willmott (2001).

*Precipitation*: total monthly precipitation for year 2013 or 2014. Source: Terrestrial Precipitation: 1900-2014 Gridded Monthly Time Series (Version 4.01) Matsuura and Willmott (2001).

*Absolute latitude*: absolute latitude of the centroid of the cell.

Continued on next page

Table D1 – continued from previous page

| Variable Description and Data Sources |
|---|
| *Longitude*: longitude of the centroid of the cell. |
| *Ruggedness*: the average degree of terrain ruggedness of grid cell. Source: Global Agro-ecological Zones (GAEZ v3.0). FAO Rome, Italy and IIASA, Laxenburg, Austria. |
| *Elevation*: average cell elevation. Source: National Oceanic and Atmospheric Administration (NOAA) and U.S. National Geophysical Data Center, TerrainBase, release 1.0 (CD-ROM), Boulder, Colo. |
| *Water area*: total area occupied by water in the cell (seas, oceans, lakes and rivers). Source: constructed with Digital Chart of the World in water shape file and the Digital Chart of the World oceans and sea shape file. |
| *Agriculture land*: percentage area of the cell covered by agricultural area in 2009. Source: GlobCover project, European Space Agency. |
| *Main crop area*: the sum of the harvested area (given in hectares) for the cell's main crop. Source: Portmann, Siebert and Döll (2010). |
| *Forest land*: the percentage area of the cell covered by forested area in 2009. Source: GlobCover project, European Space Agency. |
| *Barren land*: the percentage area of the cell covered by barren area in 2009. Source: GlobCover project, European Space Agency. |
| *Mountains area*: the proportion of mountainous terrain within the cell. Source: UNEP-WCMC Biodiversity Series 12. |
| *Herbaceous land*: the percentage area of the cell covered by herbaceous vegetation and lichens/mosses in 2009. Source: GlobCover project, European Space Agency. |
| *Dummy desert*: the variable takes on value 1 if the cells intersect the polygon of a desert as mapped by the World Land-Based Polygon Features, 1:10 million (2012). |
| *ln Dist. to coast*: natural logarithm of the average cell distance to closest coast. Source: constructed with coastline shape file from Global Self-consistent Hierarchical High-resolution Geography Version 4.2.2 January 1, 2013. |
| *ln Dist. to capital*: natural logarithm of the average cell distance to the country capital. Source: constructed with country shape file from DIVA-GIS. http://www.diva-gis.org. |
| *ln Dist. to China*: natural logarithm of the distance of the cell centroid to Shanghai, China. Source: constructed with the country shape file from DIVA-GIS. http://www.diva-gis.org. |
| *Land suitability*: average land suitability for agriculture in the cell. Source: Ramankutty et al. (2002). |
| *Petroleum*: binary variable taking value 1 if at least one petrol field is located in the cell, 0 otherwise. Source: Lujala, Ketil Rod and Thieme (2007). |
| *Mines*: a binary index taking value 1 if at least a mineral facilities or mineral deposit is located in the cell, 0 otherwise. Source: U.S. Geological Survey, U.S. Department of the Interior. |

Continued on next page

Table D1 – continued from previous page

**Variable Description and Data Sources**

*Population*: log(population count + 1) in a cell in 2010. Source: Center for International Earth Science Information Network - CIESIN - Columbia University. 2017. Gridded Population of the World, Version 4 (GPWv4): Population Count, Revision 10. Palisades, NY: NASA Socioeconomic Datand Applications Center (SEDAC). https://doi.org/10.7927/H4PG1PPM

*Night lights*: log(average night light intensity + 0.01) in a cell in 2013. Constructed with the data from NOOA National Geophysical Data Centre.

*Number of roads*: the variable counts the number of roads that intersect the cell. Source: Center for International Earth Science Information Network (CIESIN)/Columbia University, and Information Technology Outreach Services (ITOS)/University of Georgia. 2013. Global Roads Open Access Data Set, Version 1 (gROADSv1). Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC).

*Ethnic group*: number of ethnic groups in the cell. Source: Geo-referencing of ethnic groups dataset (GREG). https://icr.ethz.ch/data/greg/.

*Conflict*: fraction of years with at least one violent event for each grid cell in Africa over the observation period 1998-2012. Source: ACLED Version 4 (1997-2013), ACLED-Armed Confict Location and Event Data Project.

*Travel time to nearest urban center*: the travel time to the nearest major city. Source: Uchida and Nelson (2010).

*Number of sea ports*: the number of sea ports within the border of the cell. Source: Global ports (WFP SDI-T - Logistics Database)

**User specific characteristics from Weibo**

*Weibo User (worker) location*: a binary variable taking value 1 if a Weibo user is located in malaria-endemic cell, 0 otherwise. Source: Weibo.com, SINA Corporation.

*Malaria exposure (malaria ecology, China)*: the average malaria ecology index in a user's home prefecture in China. Source: Kiszewski et al. (2004).

*Genetic immunity (G6PD, China)*: the average level of the median predicted allele frequency for G6PD deficiency in a user's home prefecture in China. Source: Howes et al. (2012).

*Genetic immunity (sickle cell, China)*:the average level of sickle cell trait in a user's home prefecture in China in % of the population. Source: Piel et al. (2013).

*Male*: a bianray variable taking value one if a Weibo user is male.

*Number of followers*: the number of followers of a user on Weibo.

*Number of friends*: the number of followers who are also followed by the user on Weibo.

*User level*: the level of a Weibo user determined by his active days on Weibo.

*Number of posts*: the total number of Weibo posts a user posted in Africa in 2014.

**Chinese prefecture level controls**

*Population*: total population of a Weibo user's home prefecture in China in 2013. Source: *China Statistical Yearbook for Regional Economy 2014*, National Bureau of Statistics of China.

*Share of tertiary sector*: percentage of GDP from tertiary sector of a Weibo user's home prefecture in China in 2013.

Continued on next page

Table D1 – continued from previous page

| Variable Description and Data Sources |
|---|
| *GDP per capita*: GDP per capita of a Weibo user's home prefecture in China in 2013. |
| *GDP growth*: GDP growth rate of a Weibo user's home prefecture in China in 2013. |
| **Internet coverage** |
| *2G signal coverage*: the average share of a grid cell covered by 2G signal in 2012. Source: the Global System for Mobile Communications Association (GSMA). |
| *Fibre cable area (10km radius)*: the buffering area of fibre optic cables in a grid cell at 10km radius. Constructed with the shapefile from https://afterfiber.nsrc.org/. We remove the fibre optic cables which are under construction. |
| **African locals' immunity** |
| *Sickle cell*: the average level of sickle cell trait in a cell in % in population. Source: Piel et al. (2013). |
| *Duffy negativity*: the average proportion of the population predicted to be Duffy negative stratified into 12 classes of prevalence. Source: Howes et al. (2011). |
| **Anti-malaria policies** |
| *Therapy*: the average coverage of artemisinin-based combination therapy for the treatment of fever symptoms in a cell in % of the population in 2014. Source: Malariatlas Project https://map.ox.ac.uk/. |
| *Bednet*: the average modelled coverage of insecticide-treated bednets in a cell in % of the population in 2014.. Source: Malariatlas Project https://map.ox.ac.uk/. |

# B Summary Statistics and Additional Descriptives

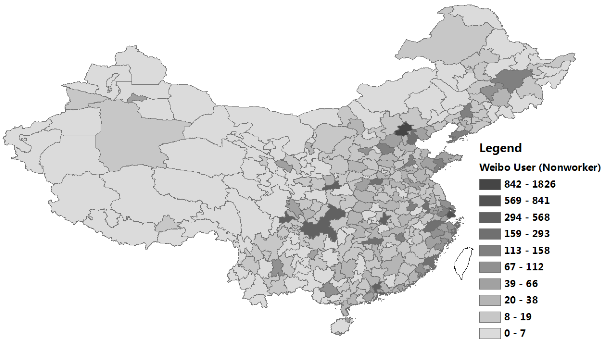
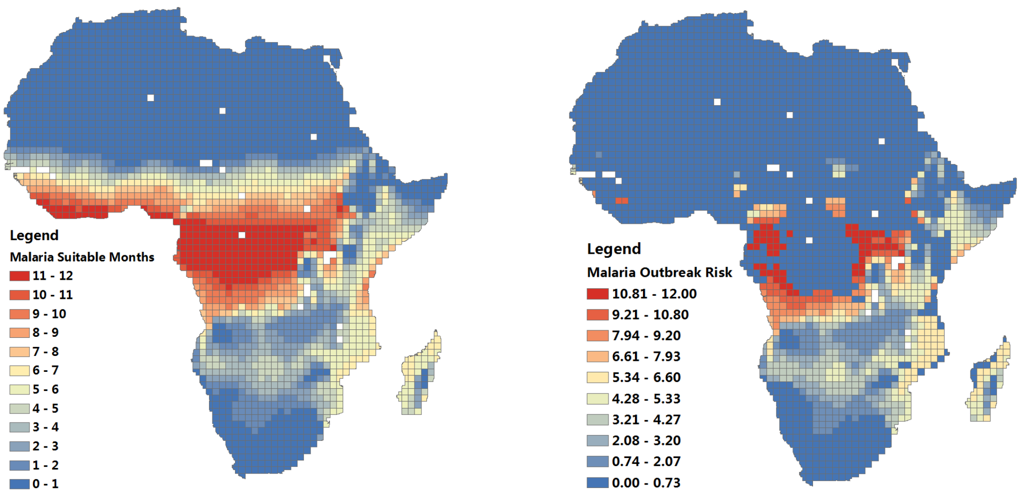



Figure C1: Chinese Prefectures of Origin of Weibo Users (Unidentified) in Africa

*Note.* The figure shows the prefectures of origin of Weibo users in Africa with no unambiguous relation to work (unidentified) and the frequency for each Chinese prefecture.



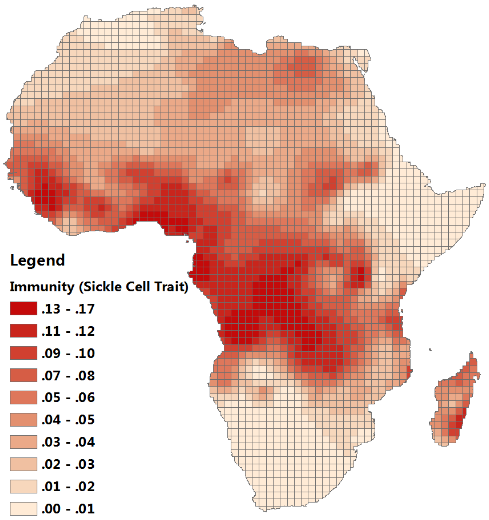

(a) Malaria Suitable Months (Mean) (b) Malaria Outbreak Risk

Figure C2: Malaria Exposure in Africa (extended)

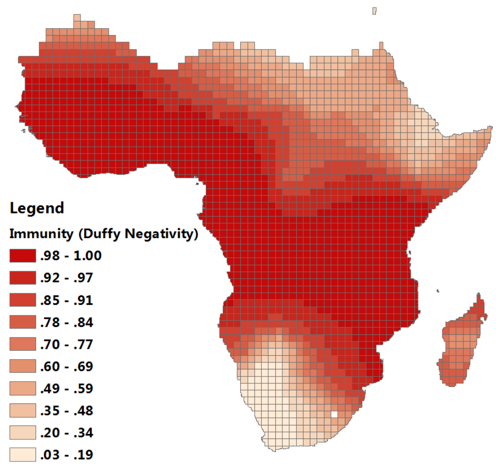
*Note.* Panel (a) displays the average number of months per year in which the climatic conditions were suitable for malaria transmission for the period 1998 to 2012. Red shadow indicates a large number of malaria suitable months, and blue shadow indicates a small number. Panel (b) shows the malaria outbreak risk measured by malaria epidemic (a malaria ecology index larger than zero and smaller than 15) × the average number of months per year in a given cell in which climatic conditions were suitable for malaria transmission between 1998 and 2012. Red indicates high risk of malaria outbreak, and blue the low risk.

Figure C3: Immunity Degree in Africa

(a) Sickle Cell Trait




(b) Duffy Negativity Phenotype Frequency




*Note.* Panel (a) shows the average level of sickle cell disease in a cell in % of the population. Panel (b) shows the proportion of the population predicted to be Duffy negative.

Table C1: Summary Statistics for Variables in Main Results

| Variable | Mean | Min. | Max. | S.D. | N |
|---|---|---|---|---|---|
| **Construction projects** | | | | | |
| Project value | 1464.984 | 0.00 | 168842.00 | 9263.20 | 2542 |
| Project number | 0.262 | 0.00 | 31.00 | 1.24 | 2542 |
| Number of public building projects | 0.083 | 0.00 | 16.00 | 0.63 | 2542 |
| Number of hydro projects | 0.013 | 0.00 | 3.00 | 0.13 | 2542 |
| Number of transport projects | 0.115 | 0.00 | 13.00 | 0.53 | 2542 |
| Number of utility projects | 0.052 | 0.00 | 12.00 | 0.36 | 2542 |
| **Aid projects** | | | | | |
| Value of aid projects | 1280.844 | 0.00 | 219426.72 | 9513.44 | 2542 |
| Number of aid projects | 0.327 | 0.00 | 29.00 | 1.49 | 2542 |
| Number of infrastructure aid projects | 0.186 | 0.00 | 14.00 | 0.84 | 2542 |
| Number of aid projects involving Chinese | 0.238 | 0.00 | 19.00 | 1.02 | 2542 |
| Number of ODA-like aid projects | 0.196 | 0.00 | 17.00 | 0.92 | 2542 |
| Number of aid projects in donations | 0.057 | 0.00 | 14.00 | 0.51 | 2542 |
| **Weibo** | | | | | |
| Weibo user (worker) | 2.302 | 0.00 | 473.00 | 18.09 | 2542 |
| Weibo user (unidentified) | 6.131 | 0.00 | 4898.00 | 105.51 | 2542 |
| Weibo user (worker) in construction | 0.490 | 0.00 | 68.00 | 3.06 | 2542 |
| Weibo user (worker) in manufacture | 0.699 | 0.00 | 107.00 | 5.01 | 2542 |
| Weibo user (worker) in business | 0.644 | 0.00 | 140.00 | 5.15 | 2542 |
| Weibo user (worker) in hydropower | 0.170 | 0.00 | 41.00 | 1.54 | 2542 |
| Weibo user (worker) in telecommunication | 0.088 | 0.00 | 16.00 | 0.70 | 2542 |
| **Malaria** | | | | | |
| Malaria ecology | 8.271 | 0.00 | 36.42 | 9.44 | 2542 |
| Malaria epidemic | 0.376 | 0.00 | 1.00 | 0.48 | 2542 |
| Malaria suitable months | 3.233 | 0.00 | 12.00 | 3.93 | 2542 |
| **Geographic controls** | | | | | |
| Temperature in 2013 | 24.507 | 11.19 | 31.90 | 3.76 | 2542 |
| Temperature in 2014 | 24.705 | 11.58 | 32.51 | 3.76 | 2542 |
| Precipitation in 2013 | 629.115 | 0.20 | 3417.53 | 606.34 | 2542 |
| Precipitation in 2014 | 633.329 | 0.10 | 3792.90 | 604.09 | 2542 |
| Absolute latitude | 16.296 | 0.50 | 37.25 | 9.26 | 2542 |
| Roughness | 0.062 | 0.00 | 0.69 | 0.08 | 2542 |
| ln Dist. to coast | 13.050 | 4.82 | 14.40 | 1.10 | 2542 |
| ln Dist. to capital | 1.690 | 0.00 | 2.88 | 0.65 | 2542 |
| ln Dist. to China | 9.383 | 9.01 | 9.65 | 0.13 | 2542 |
| Land suitability | 0.245 | 0.00 | 1.00 | 0.27 | 2542 |
| Petroleum | 0.100 | 0.00 | 1.00 | 0.30 | 2542 |
| Mines | 0.050 | 0.00 | 1.00 | 0.22 | 2542 |
| **Economic controls** | | | | | |
| ln Population | 11.123 | 0.00 | 16.99 | 2.14 | 2542 |
| ln Night lights | -3.037 | -4.61 | 3.81 | 1.85 | 2542 |
| 2G signal coverage | 0.256 | 0.00 | 1.00 | 0.33 | 2542 |
| Fibre cable area (10km radius) | 0.094 | 0.00 | 0.80 | 0.15 | 2542 |

Table C2: Summary Statistics for Variables in Weibo User Level Analysis

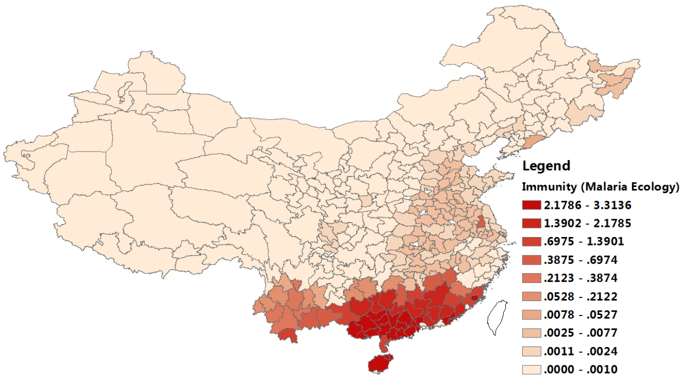
| Variable | Mean | Min. | Max. | S.D. | N |
|---|---|---|---|---|---|
| **Weibo user location** | | | | | |
| Location is malaria endemic | 0.19 | 0.00 | 1.00 | 0.40 | 3825 |
| **Weibo user immunity** | | | | | |
| Immunity(malaria ecology) | 0.24 | 0.00 | 3.31 | 0.74 | 3825 |
| Immunity(G6PD) | 0.02 | 0.00 | 0.11 | 0.02 | 3825 |
| Immunity(principal component) | -0.01 | -0.83 | 4.57 | 1.26 | 3825 |
| Immunity(sickle cell) | 0.00 | 0.00 | 0.00 | 0.00 | 3825 |
| **User account characteristics** | | | | | |
| Male | 0.50 | 0.00 | 1.00 | 0.50 | 3825 |
| # of Followers | 5845.72 | 0.00 | 3.12e+06 | 65192.86 | 3825 |
| # of Friends | 297.11 | 0.00 | 3000.00 | 310.69 | 3825 |
| User level | 2.67 | 1.00 | 10.00 | 2.61 | 3825 |
| # of Posts | 15.66 | 1.00 | 394.00 | 24.53 | 3825 |
| **Prefecture characteristics** | | | | | |
| ln Population | 6.91 | 2.29 | 8.00 | 0.76 | 3825 |
| Tertiary sector(%) | 0.54 | 0.13 | 0.77 | 0.15 | 3825 |
| GDP per capita | 79788.66 | 8440.00 | 196728.00 | 29285.42 | 3825 |
| GDP growth | 109.70 | 103.90 | 122.30 | 1.84 | 3825 |

Table C3: Summary Statistics for Variables in Mechanisms and Robustness

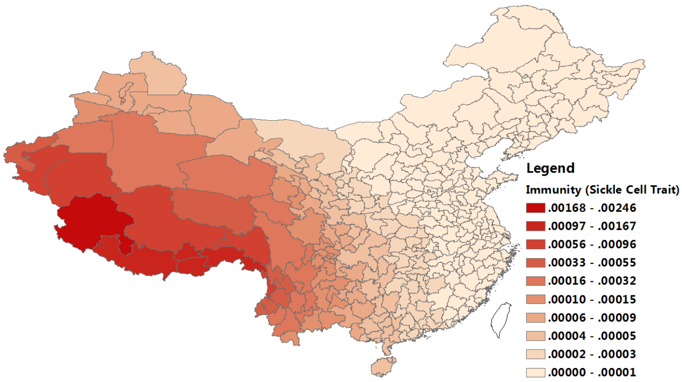
| Variable | Mean | Min. | Max. | S.D. | N |
|---|---|---|---|---|---|
| **Other diseases** | | | | | |
| HIV | 0.05 | 0.00 | 0.33 | 0.06 | 728 |
| Tsetse | 0.20 | 0.00 | 1.00 | 0.27 | 2003 |
| **African locals' immunity** | | | | | |
| African immunity (sickle cell) | 0.04 | 0.00 | 0.17 | 0.04 | 2542 |
| African immunity (duffy negativity) | 0.82 | 0.03 | 1.00 | 0.24 | 2035 |
| **Additional controls** | | | | | |
| Elevation | 0.62 | -0.16 | 2.51 | 0.42 | 2542 |
| Area | 11543.62 | 2466.00 | 12391.32 | 1258.78 | 2542 |
| Longitude | 18.07 | -16.50 | 50.58 | 15.53 | 2542 |
| Water area | 0.29 | 0.00 | 8.10 | 0.66 | 2542 |
| Number of roads | 131.49 | 0.00 | 5175.00 | 368.72 | 2542 |
| Ethnic group | 1.99 | 1.00 | 7.00 | 1.17 | 2542 |
| Conflict | 0.17 | 0.00 | 1.00 | 0.25 | 2542 |
| Travel time to nearest urban center | 721.20 | 44.02 | 5651.77 | 702.79 | 2542 |
| Number of sea ports | 0.03 | 0.00 | 1.00 | 0.16 | 2542 |
| Agriculture land | 15.81 | 0.00 | 98.80 | 22.99 | 2542 |
| Crop's harvest area | 5809.37 | 0.00 | 112088.70 | 12022.80 | 2542 |
| Forest land | 25.09 | 0.00 | 99.94 | 30.23 | 2542 |
| Barren land | 34.20 | 0.00 | 100.00 | 44.59 | 2542 |
| Mountains area | 0.14 | 0.00 | 1.00 | 0.24 | 2542 |
| Herbaceous land | 11.40 | 0.00 | 99.89 | 23.12 | 2542 |
| Dummy desert | 0.03 | -0.92 | 1.09 | 1.00 | 2542 |
| **Anti-malaria policies** | | | | | |
| Therapy in 2014 | 0.11 | 0.00 | 0.49 | 0.11 | 1668 |
| Bednet in 2014 | 0.41 | 0.00 | 1.00 | 0.23 | 1668 |
| **Other measures of malaria risk** | | | | | |
| Malaria temperature stability | 35111.92 | 2054.62 | 73407.59 | 16381.30 | 2541 |
| Malaria climate stability 2000 | 185.27 | 0.04 | 322.65 | 76.92 | 2542 |
| McCord Falciparum CRU 2000 | 1.42 | 0.00 | 3.61 | 0.94 | 2531 |

Figure C4: Immunity Degree in China

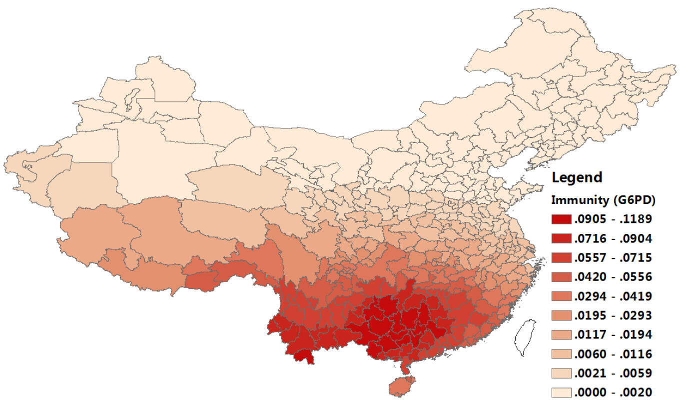
(a) Immunity (Malaria Ecology)




(b) Immunity (Sickle Cell Trait)




(c) Immunity (G6PD)




*Note.* Panel (a) shows the level of immunity to malaria measured by malaria ecology index in Chinese prefectures. Red shadow indicates high level of immunity, blue shadow the low level. Panel(b) shows the average level of sickle cell disease in Chinese prefectures in % in population. Panel (c) depicts the average level of immunity to malaria measured by G6PD deficiency in Chinese prefectures.

# C Additional Results

Table A1: Malaria Exposure and Chinese Projects in Africa (Cells with Malaria Ecology > 0)

| Panel A: Aid Projects | | | | | | |
|---|---|---|---|---|---|---|
| Dependent Variables | Aid Projects | | Number of aid projects | | | |
| | Value | Number | Infrastructure | Involving Chinese | ODA-like | Donation |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Malaria ecology | -1417.258 | -0.252* | -0.096 | -0.142 | -0.229*** | -0.053* |
| | (1374.620) | (0.130) | (0.088) | (0.105) | (0.082) | (0.031) |
| | [1113.035] | [0.093] | [0.057] | [0.075] | [0.064] | [0.022] |
| Observations | 1609 | 1609 | 1609 | 1609 | 1609 | 1609 |
| $R^2$ | 0.590 | 0.689 | 0.609 | 0.619 | 0.705 | 0.809 |
| Panel B: Construction Projects | | | | | | |
| Dependent Variables | Construction Projects | | Number of Projects in | | | |
| | Value | Number | Hydropower | Public Building | Transport | Utility |
| Malaria ecology | -18.288 | -0.211* | -0.042 | -0.144* | -0.016 | -0.009 |
| | (894.923) | (0.147) | (0.028) | (0.083) | (0.061) | (0.037) |
| | [602.110] | [0.102] | [0.030] | [0.062] | [0.044] | [0.023] |
| Observations | 1609 | 1609 | 1609 | 1609 | 1609 | 1609 |
| $R^2$ | 0.714 | 0.705 | 0.463 | 0.727 | 0.613 | 0.556 |
| All Panels: | | | | | | |
| Admin. Area FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Ethnic Group FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Economic Controls | Yes | Yes | Yes | Yes | Yes | Yes |

*Notes.* OLS estimates. All continuous control variables are standardized. Panel A: The dependent variables are number or value of the infrastructure aid projects in various categories, data from AidData. Panel B: are number or value of the infrastructure projects in various sectors contracted by Chinese companies in Africa in 2013 from Yearbook. "Malaria ecology" is a time-invariant index measuring the stability and force of malaria transmission in a cell. "Admin. Area FEs" represent (dummy) controls for level-1 administrative areas at the sub-national level from the Database of Global Administrative Areas (GADM). "Ethnic Group FEs" represent (dummy) controls for culture groups as defined by Murdock (1967) (see also Nunn, 2008). "Geographic Controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic Controls" include population, night lights, 2G internet coverage and Fibre cable area. The unit of observation is a 1 x 1 degree cell. Robust standard errors are reported in parentheses. Conley (1999) standard errors are reported in square brackets (cutoff = 500 km). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A2: Malaria Exposure and Chinese Projects in Africa: Restricted Samples

| Panel A: Aid Projects | | | | | | |
|---|---|---|---|---|---|---|
| Dependent Variables | Aid Projects | | Number of aid projects | | | |
| | Value | Number | Infrastructure | Involving Chinese | ODA-like | Donation |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Panel A1: Cells with population above 10 percentile | | | | | | |
| Malaria ecology | -1684.956** | -0.280*** | -0.153*** | -0.191*** | -0.198*** | -0.049*** |
| | (715.323) | (0.070) | (0.049) | (0.057) | (0.043) | (0.017) |
| Observations | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 |
| $R^2$ | 0.575 | 0.667 | 0.592 | 0.595 | 0.679 | 0.760 |
| Panel A2: Cells with population above 20 percentile | | | | | | |
| Malaria ecology | -1849.443** | -0.308*** | -0.166*** | -0.211*** | -0.213*** | -0.048** |
| | (820.178) | (0.080) | (0.056) | (0.065) | (0.050) | (0.019) |
| Observations | 2002 | 2002 | 2002 | 2002 | 2002 | 2002 |
| $R^2$ | 0.579 | 0.672 | 0.597 | 0.600 | 0.683 | 0.763 |
| Panel B: Construction Projects | | | | | | |
| Dependent Variables | Construction Projects | | Number of Projects in | | | |
| | Value | Number | Hydropower | Public Building | Transport | Utility |
| Panel B1: Cells with population above 10 percentile | | | | | | |
| Malaria ecology | -158.037 | -0.160** | -0.019 | -0.079* | -0.033 | -0.028 |
| | (528.512) | (0.071) | (0.013) | (0.041) | (0.028) | (0.018) |
| Observations | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 |
| $R^2$ | 0.687 | 0.666 | 0.411 | 0.658 | 0.598 | 0.538 |
| Panel B2: Cells with population above 20 percentile | | | | | | |
| Malaria ecology | -143.336 | -0.176** | -0.023 | -0.085* | -0.043 | -0.025 |
| | (637.294) | (0.087) | (0.016) | (0.050) | (0.032) | (0.021) |
| Observations | 2002 | 2002 | 2002 | 2002 | 2002 | 2002 |
| $R^2$ | 0.689 | 0.670 | 0.415 | 0.661 | 0.601 | 0.543 |
| Both Panels: | | | | | | |
| Admin. Area FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Ethnic Group FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Economic Controls | Yes | Yes | Yes | Yes | Yes | Yes |

*Notes.* OLS estimates. All continuous control variables are standardized. The dependent variable is the total number of Chinese aid projects in Africa from 2010 to 2014, data from AidData (Panel A) or number or value of the infrastructure projects in various sectors contracted by Chinese companies in Africa in 2013 from Yearbook (Panel B). "Malaria ecology" is a time-invariant index measuring the stability and force of malaria transmission in a cell. "Admin. Area FEs" represent (dummy) controls for level-1 administrative areas at the sub-national level from the Database of Global Administrative Areas (GADM). "Ethnic Group FEs" represent (dummy) controls for culture groups as defined by Murdock (1967) (see also Nunn, 2008). "Geographic Controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic Controls" include population, night lights, 2G internet coverage and Fibre cable area. The unit of observation is a 1 x 1 degree cell. Robust standard errors are reported in parentheses. Conley (1999) standard errors are reported in square brackets (cutoff = 500 km). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Malaria and Weibo Users Identified as Workers - Restricted Samples

| Dependent Variables | ln Weibo users (workers) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Panel A. Cells with population above 10 percentile | | | | |
| Malaria ecology | -0.063*** | -0.142*** | -0.174*** | -0.213*** |
| | (0.015) | (0.034) | (0.035) | (0.032) |
| Observations | 2277 | 2259 | 2259 | 2259 |
| $R^2$ | 0.007 | 0.602 | 0.631 | 0.714 |
| Panel B. Cells with population above 20 percentile | | | | |
| Malaria ecology | -0.103*** | -0.153*** | -0.184*** | -0.214*** |
| | (0.017) | (0.037) | (0.040) | (0.036) |
| Observations | 2019 | 2002 | 2002 | 2002 |
| $R^2$ | 0.016 | 0.615 | 0.643 | 0.730 |
| Controls: | | | | |
| Geographic Controls | No | No | Yes | Yes |
| Economic Controls | No | No | No | Yes |
| Admin. Area FEs | No | Yes | Yes | Yes |
| Ethnic Group FEs | No | Yes | Yes | Yes |

*Notes.* OLS estimates. All continuous predictor variables are standardized. "Geographic controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic controls" include population, night lights, 2G internet coverage and Fibre cable area. The unit of observation is a 1 x 1 degree cell. Robust standard errors are reported in parentheses. Conley (1999) standard errors are reported in square brackets (cutoff = 500 km). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Mediation Analysis: Malaria Exposure, Chinese Workers and Projects in Africa

| Panel A: Aid Projects | | | | | | |
|---|---|---|---|---|---|---|
| Dependent Variables | Aid Projects | | Number of aid projects | | | |
| | Value | Number | Infrastructure | Involving Chinese | ODA-like | Donation |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Malaria ecology | -851.186 | -0.114** | -0.072* | -0.081* | -0.108*** | -0.013 |
| | (549.390) | (0.053) | (0.039) | (0.045) | (0.034) | (0.015) |
| ln Weibo user (worker) | 2746.076*** | 0.527*** | 0.255*** | 0.349*** | 0.279*** | 0.128*** |
| | (715.008) | (0.091) | (0.054) | (0.066) | (0.059) | (0.034) |
| Observations | 2524 | 2524 | 2524 | 2524 | 2524 | 2524 |
| $R^2$ | 0.598 | 0.700 | 0.616 | 0.627 | 0.703 | 0.776 |
| Panel B: Construction Projects | | | | | | |
| Dependent Variables | Construction Projects | | Number of Projects in | | | |
| | Value | Number | Hydropower | Public Building | Transport | Utility |
| Malaria ecology | 338.940 | -0.062 | -0.015 | -0.038 | 0.008 | -0.018 |
| | (462.355) | (0.050) | (0.010) | (0.028) | (0.025) | (0.015) |
| ln Weibo user (worker) | 1639.386*** | 0.300*** | 0.009 | 0.133*** | 0.125*** | 0.033 |
| | (541.900) | (0.081) | (0.011) | (0.041) | (0.026) | (0.026) |
| Observations | 2524 | 2524 | 2524 | 2524 | 2524 | 2524 |
| $R^2$ | 0.693 | 0.679 | 0.411 | 0.669 | 0.611 | 0.518 |
| Both Panels: | | | | | | |
| Admin. Area FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Ethnic Group FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Economic Controls | Yes | Yes | Yes | Yes | Yes | Yes |

*Notes.* OLS estimates. The unit of observation is a 1 x 1 degree cell. All continuous control variables are standardized. The dependent variable is the total number of Chinese aid projects in Africa from 2010 to 2014, data from AidData (Panel A) or number or value of the infrastructure projects in various sectors contracted by Chinese companies in Africa in 2013 from Yearbook (Panel B). "Malaria ecology" is a time-invariant index measuring the stability and force of malaria transmission globally. "Admin. Area FEs" represent (dummy) controls for level-1 administrative areas at the sub-national level from the Database of Global Administrative Areas (GADM). "Ethnic Group FEs" represent (dummy) controls for culture groups as defined by Murdock (1967) (see also Nunn, 2011). "Geographic Controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic Controls" include population, night lights, 2G internet coverage and Fibre cable area. Robust Standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Robustness – Binary Dependent Variables

| Dependent Variables | Aid Projects | | Number of aid projects | | | | Weibo user(worker) |
|---|---|---|---|---|---|---|---|
| | Value | Number | Infrastructure | Involving Chinese | ODA-like | Donation | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Panel A: Aid Projects | | | | | | | |
| Malaria ecology | -0.033** | -0.055*** | -0.034** | -0.040** | -0.058*** | -0.027*** | -0.066*** |
| | (0.013) | (0.016) | (0.015) | (0.016) | (0.015) | (0.009) | (0.012) |
| Observations | 2524 | 2524 | 2524 | 2524 | 2524 | 2524 | 2542 |
| $R^2$ | 0.572 | 0.592 | 0.577 | 0.587 | 0.568 | 0.528 | 0.415 |
| Panel B: Construction Projects | | | | | | | |
| Dependent Variables | Construction Projects | | Number of Projects in | | | | |
| | Value | Number | Hydropower | Public Building | Transport | Utility | |
| Malaria ecology | -0.025* | -0.025* | -0.009 | -0.026*** | -0.012 | -0.010 | |
| | (0.015) | (0.015) | (0.007) | (0.009) | (0.013) | (0.010) | |
| Observations | 2524 | 2524 | 2524 | 2524 | 2524 | 2524 | |
| $R^2$ | 0.529 | 0.529 | 0.450 | 0.529 | 0.506 | 0.470 | |
| Both Panels: | | | | | | | |
| Admin. Area FEs | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Ethnic Group FEs | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Economic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

*Notes.* OLS estimates. All continuous predictor variables are standardized. All dependent variables are binary and take value 1 if the value or number of projects or Weibo user (worker) in the cell is greater than 0. Geographic controls, Economic controls and Country FEs are included in all models. "Geographic controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic controls" include population, night lights, 2G internet coverage and Fibre cable area. The unit of observation is a 1 x 1 degree cell. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: Malaria and Weibo Users - Linear Specification

| Dependent Variables | ln Weibo users | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Panel A: Full Sample | | | | |
| Malaria ecology | -0.052*** | -0.161*** | -0.202*** | -0.250*** |
| | (0.019) | (0.042) | (0.042) | (0.040) |
| Observations | 2542 | 2524 | 2524 | 2524 |
| $R^2$ | 0.003 | 0.591 | 0.619 | 0.697 |
| Panel B: Cells with Malaria Ecology > 0 | | | | |
| Malaria ecology | -0.248*** | -0.275*** | -0.258*** | -0.195** |
| | (0.032) | (0.075) | (0.090) | (0.079) |
| Observations | 1621 | 1609 | 1609 | 1609 |
| $R^2$ | 0.044 | 0.645 | 0.674 | 0.770 |
| Controls: | | | | |
| Geographic Controls | No | No | Yes | Yes |
| Economic Controls | No | No | No | Yes |
| Admin. Area FEs | No | Yes | Yes | Yes |
| Ethnic Group FEs | No | Yes | Yes | Yes |

*Notes.* OLS estimates. All continuous control variables are standardized. The dependent variable "ln Weibo users" is the logarithm of the Weibo users plus one in Africa in 2014. "Malaria ecology" is a time-invariant index measuring the stability and force of malaria transmission in a cell."Geographic Controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic Controls" include population, night lights, 2G internet coverage and Fibre cable area. "Admin. Area FEs" represent (dummy) controls for level-1 administrative areas at the sub-national level from the Database of Global Administrative Areas (GADM). "Ethnic Group FEs" represent (dummy) controls for culture groups as defined by Murdock (1967) (see also Nunn, 2008). The unit of observation is a 1 x 1 degree cell. Robust standard errors are reported in parentheses. Conley (1999) standard errors are reported in square brackets (cutoff = 500 km). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: Malaria and Historical Precolonial Institutions and Development

| Dependent Variables | Centralization | ln Pop. density | Intensive agriculture | Plow use |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Malaria ecology | -0.005 | 0.015 | 0.005 | 0.001 |
| | (0.004) | (0.018) | (0.004) | (0.001) |
| TseTse | -0.075** | -0.745*** | -0.090*** | -0.057** |
| | (0.035) | (0.229) | (0.028) | (0.023) |
| Tropics | -0.044 | -0.521 | -0.347 | -0.420*** |
| | (0.213) | (0.827) | (0.229) | (0.132) |
| Temperature | 0.000 | -0.375* | -0.006 | -0.037 |
| | (0.047) | (0.192) | (0.047) | (0.028) |
| Humidity | -0.002 | -0.137 | 0.006 | -0.012 |
| | (0.020) | (0.085) | (0.019) | (0.013) |
| Humidity | 0.000 | 0.008** | -0.001 | 0.001 |
| x Temperature | (0.001) | (0.003) | (0.001) | (0.001) |
| Coast | 0.107 | 0.318 | -0.132** | 0.023 |
| | (0.102) | (0.278) | (0.062) | (0.034) |
| River | 0.143*** | -0.146 | -0.066 | -0.019 |
| | (0.043) | (0.186) | (0.040) | (0.017) |
| Longitude | -0.000 | -0.011* | -0.001 | 0.001 |
| | (0.002) | (0.006) | (0.002) | (0.001) |
| Latitude | 0.009 | -0.027 | -0.006 | 0.007* |
| | (0.009) | (0.036) | (0.008) | (0.004) |
| Altitude | 0.175** | -0.621** | -0.094 | 0.026 |
| | (0.082) | (0.306) | (0.101) | (0.050) |
| Agricultural Suitability | -0.005 | 0.753 | -0.248* | -0.091 |
| | (0.146) | (0.532) | (0.144) | (0.056) |
| Observations | 467 | 398 | 485 | 484 |
| $R^2$ | 0.139 | 0.254 | 0.268 | 0.462 |

*Notes.* OLS estimates. The unit of observation is a Murdock ethnic group (Murdock, 1959, 1967, Nunn, 2008). The dataset and empirical framework are from the Table 1 in (Alsan, 2015). The dependent variable "Centralization" is a binary variable equaling 1 if an ethnic group has a value greater than 1 for the jurisdictional hierarchy variable. The data for ln population density are from Murdock's (1959a) book *Africa, Its Peoples and Their Culture History.* "Intensive agriculture" is a binary variable equalling 1 if an ethnic group was characterized by manure for fertilizer and plowing to aerate and loosen the soil. "Plow use" is a binary variable equalling 1 if an ethnic group used plow in agriculture historcally. Tsetse is an Tsetse suitability index created by (Alsan, 2015). The climate variables are from the twentieth century reanalysis for the year 1871. Climate controls refer to temperature, relative humidity, and the first-order interaction between temperature and humidity as well as the proportion of land that is in the tropics. Waterway controls include whether a river was located within the ethnic group boundaries and whether the boundaries included a coast. Geography controls include mean altitude, the FAO's agricultural suitability index, longitude, and absolute latitude. Robust standard errors clustered at the level of cultural province in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: Malaria and Economic Development

| Dependent Variables | ln (0.01 + Light density) | | | | | |
|---|---|---|---|---|---|---|
| | Full sample | | | Malaria ecology > 0 | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Malaria ecology | 0.047 | -0.048 | -0.119* | -0.237** | -0.171* | -0.195** |
| | (0.077) | (0.067) | (0.064) | (0.120) | (0.097) | (0.094) |
| ln Population | | | 0.508*** | | | 0.587*** |
| | | | (0.070) | | | (0.099) |
| 2G signal coverage | | 1.045*** | 0.934*** | | 0.866*** | 0.742*** |
| | | (0.057) | (0.057) | | (0.065) | (0.065) |
| Fibre cable area | | 0.468*** | 0.413*** | | 0.536*** | 0.482*** |
| | | (0.040) | (0.040) | | (0.047) | (0.047) |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Ethnic Group FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Admin. Area FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 2524 | 2524 | 2524 | 1609 | 1609 | 1609 |
| $R^2$ | 0.712 | 0.816 | 0.824 | 0.772 | 0.869 | 0.874 |

*Notes.* OLS estimates. All continuous predictor variables are standardized. 'ln (0.01 + Light density)" is log(average night light intensity + 0.01) in a cell in 2013."Geographic Controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Admin. Area FEs" represent (dummy) controls for level-1 administrative areas at the sub-national level from the Database of Global Administrative Areas (GADM). "Ethnic Group FEs" represent (dummy) controls for culture groups as defined by Murdock (1967) (see also Nunn, 2011). The unit of observation is a 1 x 1 degree cell. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: Mediation Analysis: Economic Development

| Dependent Variables | # Aid projects | | # Infrastr. projects | | ln Weibo user (worker) | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Panel A: Full sample | | | | | | |
| Malaria ecology | -0.269*** | -0.254*** | -0.157** | -0.141** | -0.225*** | -0.205*** |
| | (0.066) | (0.064) | (0.064) | (0.062) | (0.033) | (0.030) |
| ln (0.01 + Light density) | | 0.228*** | | 0.252*** | | 0.308*** |
| | | (0.052) | | (0.050) | | (0.031) |
| ln Population | 0.108** | 0.046 | 0.101* | 0.034 | 0.207*** | 0.123*** |
| | (0.054) | (0.051) | (0.056) | (0.051) | (0.043) | (0.039) |
| 2G signal coverage | 0.137** | 0.025 | 0.141*** | 0.017 | 0.143*** | -0.009 |
| | (0.057) | (0.062) | (0.048) | (0.052) | (0.033) | (0.033) |
| Fibre cable area | 0.205*** | 0.155*** | 0.225*** | 0.170*** | 0.168*** | 0.100*** |
| | (0.054) | (0.051) | (0.041) | (0.038) | (0.025) | (0.024) |
| Observations | 2524 | 2524 | 2524 | 2524 | 2524 | 2524 |
| $R^2$ | 0.661 | 0.665 | 0.655 | 0.662 | 0.676 | 0.704 |
| Panel B: Malaria Ecology > 0 | | | | | | |
| Malaria ecology | -0.297** | -0.252* | -0.265* | -0.211 | -0.240*** | -0.192*** |
| | (0.134) | (0.130) | (0.155) | (0.147) | (0.061) | (0.057) |
| ln (0.01 + Light density) | | 0.470*** | | 0.514*** | | 0.508*** |
| | | (0.113) | | (0.104) | | (0.048) |
| ln Population | 0.271* | 0.125 | 0.280 | 0.121 | 0.243*** | 0.085 |
| | (0.151) | (0.142) | (0.175) | (0.154) | (0.064) | (0.059) |
| 2G signal coverage | 0.194** | 0.010 | 0.121 | -0.081 | 0.150*** | -0.049 |
| | (0.089) | (0.098) | (0.078) | (0.091) | (0.045) | (0.042) |
| Fibre cable area | 0.262*** | 0.143* | 0.319*** | 0.188*** | 0.210*** | 0.082*** |
| | (0.083) | (0.078) | (0.062) | (0.058) | (0.033) | (0.031) |
| Observations | 1609 | 1609 | 1609 | 1609 | 1609 | 1609 |
| $R^2$ | 0.681 | 0.689 | 0.691 | 0.705 | 0.733 | 0.777 |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Ethnic Group FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Admin. Area FEs | Yes | Yes | Yes | Yes | Yes | Yes |

*Notes.* OLS estimates. All continuous predictor variables are standardized. 'ln (0.01 + Light density)" is log(average night light intensity + 0.01) in a cell in 2013. "'Geographic Controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Admin. Area FEs" represent (dummy) controls for level-1 administrative areas at the sub-national level from the Database of Global Administrative Areas (GADM). "Ethnic Group FEs" represent (dummy) controls for culture groups as defined by Murdock (1967) (see also Nunn, 2011). The unit of observation is a 1 x 1 degree cell. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: Roles of Other Diseases: HIV, Tsetse

| Dependent Variables | ln Weibo users (workers) | | | | | Number of infrastructure aid projects | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | DHS Sample: HIV | | | Tsetse | | DHS Sample: HIV | | | Tsetse | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Malaria ecology | -0.146*** | | -0.153* | | -0.179*** | -0.101 | | -0.078 | | -0.136*** |
| | (0.054) | | (0.089) | | (0.031) | (0.089) | | (0.133) | | (0.047) |
| HIV | | 0.064 | 0.069 | | | | 0.024 | 0.026 | | |
| | | (0.045) | (0.045) | | | | (0.082) | (0.082) | | |
| Tsetse | | | | 0.027 | 0.025 | | | | -0.022 | -0.023 |
| | | | | (0.048) | (0.048) | | | | (0.072) | (0.071) |
| Controls: | | | | | | | | | | |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Economic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Admin. Area FEs | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Ethnic Group FEs | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 841 | 641 | 641 | 1991 | 1991 | 841 | 641 | 641 | 1991 | 1991 |
| $R^2$ | 0.796 | 0.803 | 0.804 | 0.713 | 0.720 | 0.704 | 0.713 | 0.713 | 0.585 | 0.588 |

*Notes.* OLS estimates. All continuous predictor variables are standardized. "HIV" indicates the average prevalence rate in a cell. "Tsetse" is Average a index representing the predicted suitability for tsetse flies a cell. "Geographic controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic controls" include population, night lights, 2G internet coverage and Fibre cable area. The unit of observation is a 1 x 1 degree cell. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

# Table A11: Additional Controls for Weibo Users (Workers) - Linear Regression

| Dependent Variables | ln Weibo Users (workers) | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) | (15) | (16) | (17) |
| Malaria ecology | -0.208*** | -0.205*** | -0.207*** | -0.207*** | -0.204*** | -0.204*** | -0.195*** | -0.187*** | -0.206*** | -0.205*** | -0.203*** | -0.205*** | -0.147*** | -0.206*** | -0.171*** | -0.204*** | -0.120*** |
| | (0.031) | (0.030) | (0.030) | (0.030) | (0.030) | (0.030) | (0.030) | (0.030) | (0.030) | (0.030) | (0.031) | (0.030) | (0.040) | (0.030) | (0.033) | (0.030) | (0.041) |
| Elevation | -0.015 | | | | | | | | | | | | | | | | 0.015 |
| | (0.046) | | | | | | | | | | | | | | | | (0.045) |
| Area | | 0.060* | | | | | | | | | | | | | | | 0.100*** |
| | | (0.036) | | | | | | | | | | | | | | | (0.039) |
| Longitude | | | 0.355 | | | | | | | | | | | | | | 0.705** |
| | | | (0.348) | | | | | | | | | | | | | | (0.357) |
| Water area | | | | 0.030 | | | | | | | | | | | | | 0.044** |
| | | | | (0.020) | | | | | | | | | | | | | (0.021) |
| Number of roads | | | | | 0.033* | | | | | | | | | | | | 0.013 |
| | | | | | (0.017) | | | | | | | | | | | | (0.017) |
| Ethnic groups | | | | | | -0.003 | | | | | | | | | | | -0.003 |
| | | | | | | (0.019) | | | | | | | | | | | (0.019) |
| Conflict | | | | | | | 0.169*** | | | | | | | | | | 0.167*** |
| | | | | | | | (0.030) | | | | | | | | | | (0.030) |
| Travel time to nearest urban center | | | | | | | | 0.047*** | | | | | | | | | 0.040*** |
| | | | | | | | | (0.010) | | | | | | | | | (0.010) |
| Number of sea ports | | | | | | | | | 0.066** | | | | | | | | 0.059** |
| | | | | | | | | | (0.028) | | | | | | | | (0.028) |
| Agriculture land | | | | | | | | | | 0.003 | | | | | | | -0.008 |
| | | | | | | | | | | (0.025) | | | | | | | (0.031) |
| Main crop area | | | | | | | | | | | -0.013 | | | | | | -0.002 |
| | | | | | | | | | | | (0.034) | | | | | | (0.033) |
| Forest land | | | | | | | | | | | | 0.027 | | | | | 0.028 |
| | | | | | | | | | | | | (0.027) | | | | | (0.033) |
| Barren land | | | | | | | | | | | | | 0.178*** | | | | 0.135* |
| | | | | | | | | | | | | | (0.060) | | | | (0.072) |
| Mountains area | | | | | | | | | | | | | | -0.014 | | | -0.011 |
| | | | | | | | | | | | | | | (0.033) | | | (0.032) |
| Herbaceous land | | | | | | | | | | | | | | | -0.097*** | | -0.061* |
| | | | | | | | | | | | | | | | (0.026) | | (0.032) |
| Dummy desert | | | | | | | | | | | | | | | | -0.014 | -0.002 |
| | | | | | | | | | | | | | | | | (0.032) | (0.032) |
| Observations | 2524 | 2524 | 2524 | 2524 | 2524 | 2524 | 2524 | 2524 | 2524 | 2524 | 2524 | 2524 | 2524 | 2524 | 2524 | 2524 | 2524 |
| $R^2$ | 0.704 | 0.705 | 0.704 | 0.704 | 0.704 | 0.704 | 0.715 | 0.705 | 0.706 | 0.704 | 0.704 | 0.704 | 0.706 | 0.704 | 0.707 | 0.704 | 0.724 |

*Notes.* OLS estimates. All continuous predictor variables are standardised. "Elevation" is the mean elevation of a cell. "Area" is the area of a cell. "Longitude" is the average longitude of a cell. "Water area" is the total area occupied by water in the cell. "Number of roads" is the number of roads that intersect the cell. "Ethnic group" is the number of ethnic groups in the cell. "Conflict" is the fraction of years with at least one violent event for each grid cell in Africa over 1998-2012. "Travel time to nearest urban center" is the travel time to the nearest major city. "Number of sea ports" indicates the number of sea ports within the border of the cell. "Agriculture land" percentage area of the cell covered by agricultural area in 2009. "Main crop area" is the sum of the harvested area. "Forest land" is the percentage area of the cell covered by forested area in 2009. "Barren land" is the percentage area of the cell covered by barren area in 2009. "Mountains area" is the proportion of mountainous terrain within the cell. "Herbaceous land" is the percentage area of the cell covered by herbaceous vegetation and lichens/mosses in 2009. "Dummy desert" takes on value 1 if the cells intersect the polygon of a desert. Geographic controls, Economic controls and Country fixed effects are included in all models. "Geographic controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic controls" include population, night lights, 2G internet coverage and Fibre cable area. The unit of observation is a 1 x 1 degree cell. "Admin. Area FEs" and
"Ethnic Group FEs" are included in all columns. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: Other Measures of Malaria Risk

| Dependent Variables | ln Weibo users (workers) | | | # Infrastructure aid projects | | | # of Construction projects | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Malaria temperature stability | -0.087** | | | -0.023 | | | -0.054 | | |
| | (0.041) | | | (0.069) | | | (0.076) | | |
| Malaria climate stability 2000 | | -0.070* | | | -0.053 | | | -0.104 | |
| | | (0.039) | | | (0.059) | | | (0.071) | |
| McCord Falciparum CRU 2000 | | | -0.075** | | | -0.082 | | | -0.122 |
| | | | (0.032) | | | (0.054) | | | (0.093) |
| Controls: | | | | | | | | | |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Economic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Admin. Area FEs | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Ethnic Group FEs | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 2523 | 2524 | 2513 | 2523 | 2524 | 2513 | 2523 | 2524 | 2513 |
| $R^2$ | 0.698 | 0.698 | 0.702 | 0.588 | 0.588 | 0.589 | 0.661 | 0.662 | 0.662 |

*Notes.* OLS estimates. All continuous predictor variables are standardised. "Malaria temperature stability" is the average temperature suitability for Plasmodium falciparum transmission. "Malaria climate stability 2000" is average Falciparum Suitability in the 1x1 degree grid cell predicted using temperature. "McCord Falciparum CRU 2000" is average Malaria Stability Index in the 1x1 degree grid cell predicted using long-term averages of temperature and precipitation and ignoring mosquito characteristics. "Geographic controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic controls" include population, night lights, 2G internet coverage and Fibre cable area. The unit of observation is a 1 x 1 degree cell. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A13: Malaria and Weibo Users Identified as Workers - Non-linear Specification

| Dependent Variables | ln Weibo users (workers) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Panel A. Full sample | | | | |
| Malaria endemic | -0.306*** | -0.284*** | -0.284*** | -0.278*** |
| | (0.044) | (0.063) | (0.062) | (0.057) |
| | [0.042] | [0.103] | [0.081] | [0.072] |
| Malaria epidemic × MSM | -0.298*** | -0.268*** | -0.159*** | -0.104** |
| | (0.047) | (0.052) | (0.046) | (0.040) |
| | [0.078] | [0.082] | [0.053] | [0.047] |
| Malaria suitable months | 0.165*** | 0.149*** | -0.008 | 0.022 |
| | (0.022) | (0.043) | (0.052) | (0.046) |
| | [0.030] | [0.057] | [0.060] | [0.055] |
| Observations | 2542 | 2542 | 2542 | 2542 |
| $R^2$ | 0.028 | 0.143 | 0.273 | 0.424 |
| Panel B. Cells with malaria ecology > 0 | | | | |
| Malaria endemic | -0.300*** | -0.304*** | -0.180** | -0.157** |
| | (0.043) | (0.078) | (0.078) | (0.067) |
| | [0.032] | [0.122] | [0.091] | [0.079] |
| Malaria epidemic × MSM | -0.205*** | -0.138*** | -0.075* | -0.084** |
| | (0.043) | (0.048) | (0.044) | (0.039) |
| | [0.086] | [0.063] | [0.042] | [0.044] |
| Malaria suitable months | 0.052** | 0.074* | -0.047 | 0.049 |
| | (0.024) | (0.043) | (0.059) | (0.052) |
| | [0.036] | [0.050] | [0.063] | [0.058] |
| Observations | 1621 | 1621 | 1621 | 1621 |
| $R^2$ | 0.036 | 0.194 | 0.304 | 0.491 |
| All Panels: | | | | |
| Geographic controls | No | No | Yes | Yes |
| Economic controls | No | No | No | Yes |
| Country FEs | No | Yes | Yes | Yes |

*Notes.* OLS estimates. All continuous predictor variables are standardized. The dependent variable "ln Weibo users (workers)" is the logarithm of the Weibo users identified as worker plus one in Africa in 2014. "Malaria endemic" is a binary variable for high malaria exposure taking value one for cells with an average malaria ecology index equal to or larger than 15; "Malaria epidemic" is a binary index for low malaria exposure taking value one for cells with an average malaria ecology index greater than 0 and smaller than 15. "Malaria suitable months" is the average number of months that were suitable for malaria to be transmitted of each year from 1998 to 2012. "Geographic controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic controls" include population, night lights, 2G internet coverage and fibre cable area. The unit of observation is a 1 x 1 degree cell. Robust standard errors are reported in parentheses. Conley (1999) standard errors are reported in square brackets (cutoff = 500 km). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A14: Malaria and Weibo Users Identified as Workers - Panel Specification

| Dependent Variables | Ln(Weibo Users (Workers)) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Panel A: Full Sample | | | | |
| Malaria epidemic × MSM | -0.495* | -0.518* | -0.562* | -0.563* |
| | (0.287) | (0.292) | (0.331) | (0.330) |
| Malaria Suitable Months | 0.293 | 0.364 | 0.432 | 0.441 |
| | (0.283) | (0.301) | (0.377) | (0.366) |
| Precipitation | | -0.009* | | -0.003 |
| | | (0.005) | | (0.006) |
| Temperature | | | 0.411 | 0.379 |
| | | | (0.303) | (0.347) |
| Observations | 4,922 | 4,922 | 4,922 | 4,922 |
| $R^2$ | 0.882 | 0.882 | 0.882 | 0.882 |
| Panel B: Cells with Malaria epidemic =1 | | | | |
| Malaria Suitable Months | -0.215* | -0.181 | -0.186* | -0.189* |
| | (0.114) | (0.112) | (0.111) | (0.111) |
| Precipitation | | -0.007* | | -0.008 |
| | | (0.004) | | (0.005) |
| Temperature | | | 0.133 | -0.054 |
| | | | (0.106) | (0.137) |
| Observations | 1,852 | 1,852 | 1,852 | 1,852 |
| $R^2$ | 0.984 | 0.984 | 0.984 | 0.984 |
| All Panels: | | | | |
| Cell FEs | Yes | Yes | Yes | Yes |
| Half-Year FEs | Yes | Yes | Yes | Yes |
| Country × Half-Year FEs | Yes | Yes | Yes | Yes |

*Notes.* Fixed effects estimates. The dependent variable "ln Weibo users (workers)" is the logarithm of the Weibo users identified as worker plus one in Africa in 2014. "Malaria epidemic" is a binary index for low malaria exposure taking value one for cells with an average malaria ecology index greater than 0 and smaller than 15. "Malaria Suitable Months (MSM)" is the average number of months that were suitable for malaria to be transmitted of each year from 1998 to 2012. The unit of observation is a 1 x 1 degree cell. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A15: Chinese Workers and the Immunity of the Local Population – Country Fixed Effects

| Dependent Variables | ln Weibo users (workers) in | | | | | |
|---|---|---|---|---|---|---|
| | All Sectors | Construction | Manufacturing | Business | Hydropower | Telecom |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Malaria ecology | -0.122*** | -0.079*** | -0.088*** | -0.070*** | -0.030*** | -0.026*** |
| | (0.023) | (0.015) | (0.017) | (0.017) | (0.010) | (0.009) |
| African immunity (prinicipal comp.) | -0.087** | -0.074** | -0.081*** | -0.077** | -0.040** | -0.046*** |
| | (0.041) | (0.029) | (0.031) | (0.030) | (0.017) | (0.017) |
| Observations | 2035 | 2035 | 2035 | 2035 | 2035 | 2035 |
| $R^2$ | 0.474 | 0.381 | 0.381 | 0.357 | 0.257 | 0.219 |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Economic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Country FEs | Yes | Yes | Yes | Yes | Yes | Yes |

*Notes.* OLS estimates. All continuous predictor variables are standardized. The dependent variable "ln Weibo users (workers)" is the logarithm of the Weibo users identified as worker plus one in Africa in 2014. "Malaria ecology" is a time-invariant index measuring the stability and force of malaria transmission in a cell. "African immunity (principal comp.)" is the principal component of "African immunity (sickle cell)" and "African immunity (duffy)". "Geographic controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic controls" include population, night lights, 2G internet coverage and Fibre cable area. The unit of observation is a 1 x 1 degree cell. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A16: Chinese Workers and the Immunity of the Local Population – Country and Cultural Group Fixed Effects

| Dependent Variable | ln Weibo users (workers) in | | | | | |
|---|---|---|---|---|---|---|
| | All Sectors | Construction | Manufacture | Business | Hydropower | Telecom |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Malaria ecology | -0.146*** | -0.088*** | -0.093*** | -0.085*** | -0.035*** | -0.032*** |
| | (0.025) | (0.016) | (0.018) | (0.017) | (0.010) | (0.009) |
| African immunity | -0.079 | -0.056 | -0.084** | -0.084** | -0.040* | -0.051** |
| (principal comp.) | (0.052) | (0.035) | (0.038) | (0.039) | (0.021) | (0.021) |
| Observations | 2022 | 2022 | 2022 | 2022 | 2022 | 2022 |
| $R^2$ | 0.527 | 0.446 | 0.448 | 0.414 | 0.316 | 0.279 |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Economic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Country FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Ethnic Group FEs | Yes | Yes | Yes | Yes | Yes | Yes |

*Notes.* OLS estimates. The unit of observation is a 1 x 1 degree cell. All continuous predictor variables are standardized. The dependent variable "ln Weibo users (workers)" is the logarithm of the Weibo users identified as worker plus one in Africa in 2014. "Malaria ecology" is a time-invariant index measuring the stability and force of malaria transmission in a cell. "African immunity" is the principal component of the shares of the local African population that exhibit resistance in terms of the sickle cell traits or the negative Duffy antigen. "Geographic controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic controls" include population, night lights, 2G internet coverage and Fibre cable area. "Ethnic Group FEs" represent (dummy) controls for culture groups as defined by Murdock (1967) (see also Nunn, 2008). Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A17: Impact on Other Types of Weibo Users

| Dependent Variables | ln Weibo users | ln Weibo users (unidentified) | ln Weibo users (workers) with leisure | ln Weibo users (workers) without leisure |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Panel A. Full sample | | | | |
| Malaria ecology | -0.234*** | -0.194*** | -0.158*** | -0.077*** |
| | (0.031) | (0.027) | (0.022) | (0.016) |
| Observations | 2542 | 2542 | 2542 | 2542 |
| $R^2$ | 0.428 | 0.446 | 0.427 | 0.251 |
| Panel B. Cells with malaria ecology > 0 | | | | |
| Malaria endemic | -0.223** | -0.209*** | -0.152** | -0.079** |
| | (0.090) | (0.079) | (0.063) | (0.040) |
| Malaria epidemic × MSM | -0.134*** | -0.116*** | -0.081** | -0.045** |
| | (0.051) | (0.044) | (0.037) | (0.023) |
| Malaria suitable months | 0.105 | 0.052 | 0.050 | 0.017 |
| | (0.070) | (0.061) | (0.048) | (0.034) |
| Observations | 1621 | 1621 | 1621 | 1621 |
| $R^2$ | 0.496 | 0.498 | 0.490 | 0.297 |
| Geographic controls | Yes | Yes | Yes | Yes |
| Economic controls | Yes | Yes | Yes | Yes |
| Country FEs | Yes | Yes | Yes | Yes |

*Notes.* OLS estimates. All continuous predictor variables are standardized. The dependent variable "ln Weibo user" is the logarithm of all Weibo users (plus one) in Africa in 2014. "ln Weibo User (unidentified)" corresponds to Weibo users that are not unambiguously identified as workers. "ln Weibo users (workers) with leisure" is log of Weibo users who are identified as workers and have posts without any work-related keywords. "ln Weibo users (workers) without leisure" the log of Weibo users who are identified as workers and all of whose posts contain work-related keywords. "Geographic controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic controls" include population, night lights, 2G internet coverage and Fibre cable area. The unit of observation is a 1 x 1 degree cell. Country fixed effects are included in all specifications. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A18: Immunities of Weibo Users (Workers) and Malaria Risk of Destination in Africa - User level Unweighted Regression

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Panel A. Full sample | | | | | | |
| Dependent Variables | Malaria ecology index of Weibo user (worker)'s location in Africa | | | | | |
| Malaria exposure (m. ecol., China) | 0.015** | 0.012 | | | | |
| | (0.007) | (0.007) | | | | |
| Genetic immunity (G6PD, China) | | | 0.014 | 0.001 | | |
| | | | (0.030) | (0.030) | | |
| Immunity index (princ. comp., China) | | | | | 0.025** | 0.018 |
| | | | | | (0.012) | (0.013) |
| Observations | 3825 | 3825 | 3825 | 3825 | 3825 | 3825 |
| $R^2$ | 0.887 | 0.916 | 0.887 | 0.916 | 0.887 | 0.916 |
| Panel B. Cells with malaria ecology > 0 | | | | | | |
| Dependent Variables | Location of Weibo user (worker) in Africa is malaria endemic | | | | | |
| Malaria exposure (m. ecol., China) | 0.013*** | 0.008*** | | | | |
| | (0.003) | (0.003) | | | | |
| Genetic immunity (G6PD, China) | | | 0.030** | 0.016 | | |
| | | | (0.015) | (0.013) | | |
| Immunity index (princ. comp., China) | | | | | 0.022*** | 0.013*** |
| | | | | | (0.005) | (0.005) |
| Observations | 3058 | 3058 | 3058 | 3058 | 3058 | 3058 |
| $R^2$ | 0.841 | 0.883 | 0.840 | 0.883 | 0.841 | 0.883 |
| User characteristics | Yes | Yes | Yes | Yes | Yes | Yes |
| Home prefecture characteristics | Yes | Yes | Yes | Yes | Yes | Yes |
| African country FEs | Yes | No | Yes | No | Yes | No |
| Chinese province FEs | Yes | No | Yes | No | Yes | No |
| Month FEs | Yes | No | Yes | No | Yes | No |
| African country×Month FEs | No | Yes | No | Yes | No | Yes |
| Chinese province×Month FEs | No | Yes | No | Yes | No | Yes |

*Notes.* Unweighted OLS estimates. All continuous predictor variables are standardized. "Malaria exposure (m. ecol., China)" is the average malaria ecology index in a user's home prefecture in China. "Genetic immunity (G6PD, China)" is the average level of the median predicted allele frequency for G6PD deficiency in a user's home prefecture in China. "Immunity index (princ. comp., China)" is the principal component of Malaria exposure (m. ecol., China) and Genetic immunity (G6PD, China). "Genetic immunity (sickle cell, China)" is the average level of sickle cell disease in a user's home Chinese prefecture in % of the population. User characteristic controls include Male, # of Followers. # of Friends, User level, # of Posts. Home prefecture characteristic controls are Population, Share of Tertiary sector, GDP per capita and GDP growth. "African country FEs" controls for the country specific effect a user is located in Africa. "Month FEs" corresponds to the fixed effect of the month a user posted most Weibo posts in a cell. "Chinese province FEs" is the fixed effect of a user's home province in China. The unit of observation is a Weibo user (worker). Standard errors clustered at the user's home Chinese prefecture level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A19: Immunities of Weibo Users (Workers) and Malaria Risk of Destination in Africa - Post Level Weighted Regression

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Panel A. Full sample | | | | | | |
| Dependent Variables | Malaria ecology index of Weibo user (worker)'s location in Africa | | | | | |
| Malaria exposure (m. ecol., China) | 0.028 | 0.038* | | | | |
| | (0.022) | (0.022) | | | | |
| Genetic immunity (G6PD, China) | | | 0.080 | 0.123* | | |
| | | | (0.069) | (0.070) | | |
| Immunity index (princ. comp., China) | | | | | 0.051 | 0.070** |
| | | | | | (0.033) | (0.034) |
| Observations | 24520 | 24520 | 24520 | 24520 | 24520 | 24520 |
| $R^2$ | 0.821 | 0.837 | 0.821 | 0.837 | 0.821 | 0.838 |
| Panel B. Cells with malaria ecology > 0 | | | | | | |
| Dependent Variables | Location of Weibo user (worker) in Africa is malaria endemic | | | | | |
| Malaria exposure (m. ecol., China) | 0.054*** | 0.045*** | | | | |
| | (0.014) | (0.014) | | | | |
| Genetic immunity (G6PD, China) | | | 0.075 | 0.072 | | |
| | | | (0.064) | (0.068) | | |
| Immunity index (princ. comp., China) | | | | | 0.088*** | 0.074*** |
| | | | | | (0.023) | (0.022) |
| Observations | 20274 | 20274 | 20274 | 20274 | 20274 | 20274 |
| $R^2$ | 0.752 | 0.783 | 0.750 | 0.782 | 0.752 | 0.783 |
| User characteristics | Yes | Yes | Yes | Yes | Yes | Yes |
| Home prefecture characteristics | Yes | Yes | Yes | Yes | Yes | Yes |
| African country FEs | Yes | No | Yes | No | Yes | No |
| Chinese province FEs | Yes | No | Yes | No | Yes | No |
| Month FEs | Yes | No | Yes | No | Yes | No |
| African country×Month FEs | No | Yes | No | Yes | No | Yes |
| Chinese province×Month FEs | No | Yes | No | Yes | No | Yes |

*Notes.* Weighted OLS estimates. The weight is the reciprocal of the number of Weibo user (worker) in a cell in Africa. All continuous predictor variables are standardized. "Malaria exposure (m. ecol., China)" is the average malaria ecology index in a user's home prefecture in China. "Genetic immunity (G6PD, China)" is the average level of the median predicted allele frequency for G6PD deficiency in a user's home prefecture in China. "Immunity index (princ. comp., China)" is the principal component of Malaria exposure (m. ecol., China) and Genetic immunity (G6PD, China). "Genetic immunity (sickle cell, China)" is the average level of sickle cell disease in a user's home Chinese prefecture in % of the population. User characteristic controls include Male, # of Followers. # of Friends, User level, # of Posts. Home prefecture characteristic controls are Population, Share of Tertiary sector, GDP per capita and GDP growth. "African country FEs" controls for the country specific effect a user is located in Africa. "Month FEs" corresponds to the fixed effect of the month a user posted most Weibo posts in a cell. "Chinese province FEs" is the fixed effect of a user's home province in China. The unit of observation is a Weibo post. Standard errors clustered at the user's home Chinese prefecture level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A20: Immunities of Weibo Users (Workers) and Malaria Risk of Destination in Africa – Post Level Unweighted Regression

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Panel A. Full sample | | | | | | |
| Dependent Variables | Malaria ecology index of Weibo user (worker)'s location in Africa | | | | | |
| Malaria exposure (m. ecol., China) | 0.004 | 0.004 | | | | |
| | (0.010) | (0.008) | | | | |
| Genetic immunity (G6PD, China) | | | -0.006 | 0.009 | | |
| | | | (0.057) | (0.053) | | |
| Immunity index (princ. comp., China) | | | | | 0.005 | 0.006 |
| | | | | | (0.016) | (0.014) |
| Observations | 24520 | 24520 | 24520 | 24520 | 24520 | 24520 |
| $R^2$ | 0.847 | 0.856 | 0.847 | 0.856 | 0.847 | 0.856 |
| Panel B. Cells with malaria ecology > 0 | | | | | | |
| Dependent Variables | Location of Weibo user (worker) in Africa is malaria endemic | | | | | |
| Malaria exposure (m. ecol., China) | 0.017* | 0.014*** | | | | |
| | (0.005) | (0.005) | | | | |
| Genetic immunity (G6PD, China) | | | 0.032 | 0.040 | | |
| | | | (0.027) | (0.026) | | |
| Immunity index (princ. comp., China) | | | | | 0.029*** | 0.024*** |
| | | | | | (0.008) | (0.007) |
| Observations | 20274 | 20274 | 20274 | 20274 | 20274 | 20274 |
| $R^2$ | 0.792 | 0.798 | 0.791 | 0.798 | 0.792 | 0.798 |
| User characteristics | Yes | Yes | Yes | Yes | Yes | Yes |
| Home prefecture characteristics | Yes | Yes | Yes | Yes | Yes | Yes |
| African country FEs | Yes | No | Yes | No | Yes | No |
| Chinese province FEs | Yes | No | Yes | No | Yes | No |
| Month FEs | Yes | No | Yes | No | Yes | No |
| African country×Month FEs | No | Yes | No | Yes | No | Yes |
| Chinese province×Month FEs | No | Yes | No | Yes | No | Yes |

*Notes.* Unweighted OLS estimates. All continuous predictor variables are standardized. "Malaria exposure (m. ecol., China)" is the average malaria ecology index in a user's home prefecture in China. "Genetic immunity (G6PD, China)" is the average level of the median predicted allele frequency for G6PD deficiency in a user's home prefecture in China. "Immunity index (princ. comp., China)" is the principal component of Malaria exposure (m. ecol., China) and Genetic immunity (G6PD, China). "Genetic immunity (sickle cell, China)" is the average level of sickle cell disease in a user's home Chinese prefecture in % of the population. User characteristic controls include Male, # of Followers. # of Friends, User level, # of Posts. Home prefecture characteristic controls are Population, Share of Tertiary sector, GDP per capita and GDP growth. "African country FEs" controls for the country specific effect a user is located in Africa. "Month FEs" corresponds to the fixed effect of the month a user posted most Weibo posts in a cell. "Chinese province FEs" is the fixed effect of a user's home province in China. The unit of observation is a Weibo post. Standard errors clustered at the user's home Chinese prefecture level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A21: Robustness: Malaria in Africa and Immunity of Chinese - Continuous Measure of Immunity

| Dependent Variables | Chinese immunity (malaria ecology - continous variable) | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Malaria ecology | 0.195*** | 0.185*** | 0.151*** | 0.151*** | 0.140** | 0.144** |
| | (0.064) | (0.064) | (0.058) | (0.058) | (0.058) | (0.059) |
| African immunity | | | -0.087* | -0.091* | -0.072 | -0.072 |
| (principal component) | | | (0.051) | (0.050) | (0.050) | (0.050) |
| Number of projects | | | | -0.005 | -0.008 | -0.005 |
| | | | | (0.007) | (0.007) | (0.009) |
| ln Weibo user (worker) | | | | | 0.031 | 0.033* |
| | | | | | (0.019) | (0.020) |
| Number of aid projects | | | | | | -0.028 |
| | | | | | | (0.037) |
| User characteristics | Yes | Yes | Yes | Yes | Yes | Yes |
| Home prefecture characteristics | Yes | Yes | Yes | Yes | Yes | Yes |
| Economic controls | No | Yes | Yes | Yes | Yes | Yes |
| Chinese province×Month FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| African country×Month FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3889 | 3889 | 2579 | 2579 | 2579 | 2579 |
| $R^2$ | 0.803 | 0.804 | 0.832 | 0.832 | 0.832 | 0.832 |

*Notes.* Weighted OLS estimates. The weight is the reciprocal of the number of Weibo users (workers) in a cell in Africa. All continuous predictor variables are standardized. The dependent variable "Chinese immunity (malaria ecology - continuous variable)" is the average malaria ecology index in a user's home prefecture in China. "Malaria ecology" is a time-invariant index measuring the stability and force of malaria transmission in Africa."Geographic controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic controls" include population, night lights, 2G internet coverage and Fibre cable area. The unit of observation is a 1 x 1 degree cell. User characteristic controls include Male, # of Followers. # of Friends, User level, # of Posts. Home prefecture characteristic controls are Population, Share of Tertiary sector, GDP per capita and GDP growth. "African country FEs" controls for the country specific effect a user is located in Africa. "Month FEs" corresponds to the fixed effect of the month a user posted most Weibo posts in a cell. "Chinese province FEs" is the fixed effect of a user's home province in China. The unit of observation is a Weibo user (worker). Standard errors clustered at the cell level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A22: Robustness: Malaria in Africa and Immunity of Chinese - Controlling for HIV and Tsetse

| Dependent Variables | Chinese immunity (malaria ecology - binary variable) | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Malaria ecology | 0.061** | 0.074*** | 0.074*** | 0.074*** | 0.074*** | 0.073*** |
| | (0.024) | (0.026) | (0.027) | (0.027) | (0.027) | (0.027) |
| HIV | 0.004 | -0.001 | 0.001 | 0.001 | 0.002 | 0.002 |
| | (0.010) | (0.010) | (0.010) | (0.010) | (0.010) | (0.010) |
| African immunity | | | -0.013 | -0.013 | -0.009 | -0.010 |
| (principal component) | | | (0.020) | (0.021) | (0.022) | (0.022) |
| Number of projects | | | | -0.000 | -0.002 | -0.002 |
| | | | | (0.003) | (0.003) | (0.004) |
| ln Weibo user (worker) | | | | | 0.009 | 0.009 |
| | | | | | (0.009) | (0.010) |
| Number of aid projects | | | | | | 0.001 |
| | | | | | | (0.016) |
| User characteristics | Yes | Yes | Yes | Yes | Yes | Yes |
| Home prefecture characteristics | Yes | Yes | Yes | Yes | Yes | Yes |
| Economic controls | No | Yes | Yes | Yes | Yes | Yes |
| Chinese province×Month FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| African country×Month FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1327 | 1327 | 1327 | 1327 | 1327 | 1327 |
| $R^2$ | 0.940 | 0.940 | 0.940 | 0.940 | 0.940 | 0.940 |

*Notes.* Weighted OLS estimates. The weight is the reciprocal of the number of Weibo users (workers) in a cell in Africa. All continuous predictor variables are standardized. The dependent variable "Chinese immunity (malaria ecology - binary variable)" is a binary variable equalling to 1 if the average malaria ecology index in a user's home prefecture in China is greater than zero, 0 otherwise. "Malaria ecology" is a time-invariant index measuring the stability and force of malaria transmission in Africa. "HIV" indicates the average prevalence rate in a cell. "Tsetse" is an index representing the predicted suitability for tsetse flies a cell. "Geographic controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic controls" include population, night lights, 2G internet coverage and Fibre cable area. The unit of observation is a 1 x 1 degree cell. User characteristic controls include Male, # of Followers. # of Friends, User level, # of Posts. Home prefecture characteristic controls are Population, Share of Tertiary sector, GDP per capita and GDP growth. "African country FEs" controls for the country specific effect a user is located in Africa. "Month FEs" corresponds to the fixed effect of the month a user posted most Weibo posts in a cell. "Chinese province FEs" is the fixed effect of a user's home province in China. The unit of observation is a Weibo user (worker). Standard errors clustered at the cell level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A23: Robustness: Malaria in Africa and Immunity of Chinese - Immunity above Median

| Dependent Variables | Chinese immunity (malaria ecology - above median) | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Malaria ecology | 0.091*** | 0.088*** | 0.092*** | 0.092*** | 0.091*** | 0.090*** |
| | (0.024) | (0.025) | (0.027) | (0.027) | (0.027) | (0.027) |
| African immunity | | | -0.031 | -0.030 | -0.028 | -0.028 |
| (principal component) | | | (0.029) | (0.029) | (0.030) | (0.030) |
| Number of projects | | | | 0.001 | 0.001 | -0.000 |
| | | | | (0.004) | (0.004) | (0.004) |
| ln Weibo user (worker) | | | | | 0.003 | 0.002 |
| | | | | | (0.009) | (0.009) |
| Number of aid projects | | | | | | 0.010 |
| | | | | | | (0.016) |
| User characteristics | Yes | Yes | Yes | Yes | Yes | Yes |
| Home prefecture characteristics | Yes | Yes | Yes | Yes | Yes | Yes |
| Economic controls | No | Yes | Yes | Yes | Yes | Yes |
| Chinese province×Month FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| African country×Month FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3889 | 3889 | 2579 | 2579 | 2579 | 2579 |
| $R^2$ | 0.894 | 0.894 | 0.915 | 0.915 | 0.915 | 0.915 |

*Notes.* Weighted OLS estimates. The weight is the reciprocal of the number of Weibo users (workers) in a cell in Africa. All continuous predictor variables are standardized. The dependent variable "Chinese immunity (malaria ecology - above median)" is a binary variable equalling to 1 if the average malaria ecology index in a user's home prefecture in China is above the median, 0 otherwise. "Malaria ecology" is a time-invariant index measuring the stability and force of malaria transmission in Africa."Geographic controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic controls" include population, night lights, 2G internet coverage and Fibre cable area. The unit of observation is a 1 x 1 degree cell. User characteristic controls include Male, # of Followers. # of Friends, User level, # of Posts. Home prefecture characteristic controls are Population, Share of Tertiary sector, GDP per capita and GDP growth. "African country FEs" controls for the country specific effect a user is located in Africa. "Month FEs" corresponds to the fixed effect of the month a user posted most Weibo posts in a cell. "Chinese province FEs" is the fixed effect of a user's home province in China. The unit of observation is a Weibo user (worker). Standard errors clustered at the cell level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A24: Chinese Workers and the Immunity of the Local Population – Sickle Cell Trait

| Dependent Variables | ln Weibo users (workers) in | | | | | |
|---|---|---|---|---|---|---|
| | All Sectors | Construction | Manufacture | Business | Hydropower | Telecom |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Malaria ecology | -0.206*** | -0.120*** | -0.135*** | -0.124*** | -0.052*** | -0.038*** |
| | (0.030) | (0.021) | (0.023) | (0.022) | (0.014) | (0.011) |
| African immunity | 0.007 | -0.032 | -0.023 | -0.018 | -0.045** | -0.022 |
| (sickle cell) | (0.046) | (0.032) | (0.034) | (0.033) | (0.023) | (0.015) |
| Admin. Area FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Ethnic Group FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Economic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 2524 | 2524 | 2524 | 2524 | 2524 | 2524 |
| $R^2$ | 0.704 | 0.665 | 0.668 | 0.645 | 0.607 | 0.601 |

*Notes.* OLS estimates. All continuous predictor variables are standardized. "Geographic controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic controls" include population, night lights, 2G internet coverage and Fibre cable area. The unit of observation is a 1 x 1 degree cell. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A25: Robustness - Anti-malaria Policies

| Dependent Variables | ln Weibo user | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Malaria ecology | -0.185*** | -0.177*** | -0.225*** | -0.218*** |
| | (0.038) | (0.037) | (0.050) | (0.049) |
| Policy (above median) | 0.185*** | 0.175** | 0.226** | 0.212** |
| | (0.070) | (0.070) | (0.088) | (0.085) |
| Value of health aid | | 0.054*** | | 0.044** |
| | | (0.015) | | (0.018) |
| Geographic Controls | Yes | Yes | Yes | Yes |
| Economic Controls | Yes | Yes | Yes | Yes |
| Ethnic Group FEs | Yes | Yes | Yes | Yes |
| Country FEs | Yes | Yes | No | No |
| Admin. Area FEs | No | No | Yes | Yes |
| Observations | 1656 | 1656 | 1656 | 1656 |
| $R^2$ | 0.546 | 0.550 | 0.744 | 0.746 |

*Notes.* OLS estimates. All continuous predictor variables are standardized. "Policy (above median)" is a binary variable equal to 1 if the mean of anti-malaria policies -Therapy and Bednet in 2014 - is above the median. "Value of health aid" is the value of Chinese health aid projects in Africa from 2010 to 2014. "Geographic controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic controls" include population, night lights, 2G internet coverage and Fibre cable area. The unit of observation is a 1 x 1 degree cell. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A26: Malaria and Health Aid Projects

| Dependent Variables | Number of health aid projects | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Malaria ecology | | -0.054** | -0.047* |
| | | (0.024) | (0.024) |
| ln Weibo users (workers) | 0.181*** | | |
| | (0.042) | | |
| African immunity | | | -0.017 |
| (principal component) | | | (0.046) |
| Geographic Controls | Yes | Yes | Yes |
| Economic Controls | Yes | Yes | Yes |
| Ethnic Group FEs | Yes | Yes | Yes |
| Admin. Area FEs | Yes | Yes | Yes |
| Observations | 2524 | 2524 | 2022 |
| $R^2$ | 0.615 | 0.589 | 0.594 |

*Notes.* OLS estimates. The unit of observation is a 1 x 1 degree cell. All continuous predictor variables are standardized. "Number of health aid projects" is the total number of Chinese aid projects related to heath in Africa from 2010 to 2014. "ln Weibo users (workers)" is the logarithm of the Weibo users identified as worker plus one in Africa in 2014. "Malaria ecology" is a time-invariant index measuring the stability and force of malaria transmission globally. "Geographic controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic controls" include population, night lights, 2G internet coverage and Fibre cable area. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A27: Malaria and Health Aid Projects

| Dependent Variables | ln Value of health aid projects | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Malaria ecology | | -38.075 | -36.489 |
| | | (34.746) | (33.325) |
| ln Weibo users (workers) | 71.218 | | |
| | (62.427) | | |
| African immunity | | | -18.613 |
| (principal component) | | | (44.039) |
| Geographic Controls | Yes | Yes | Yes |
| Economic Controls | Yes | Yes | Yes |
| Ethnic Group FEs | Yes | Yes | Yes |
| Admin. Area FEs | Yes | Yes | Yes |
| Observations | 2524 | 2524 | 2022 |
| $R^2$ | 0.223 | 0.215 | 0.221 |

*Notes.* OLS estimates. The unit of observation is a 1 x 1 degree cell. All continuous predictor variables are standardized. "ln Value of health aid projects" is the total number of Chinese aid projects related to heath in Africa from 2010 to 2014. "ln Weibo users (workers)" is the logarithm of the Weibo users identified as worker plus one in Africa in 2014. "Malaria ecology" is a time-invariant index measuring the stability and force of malaria transmission globally. "Geographic controls" contain temperature, precipitation, latitude, land ruggedness, distance to the country capital, to the coast and to China, agriculture suitability, petrol and mines. "Economic controls" include population, night lights, 2G internet coverage and Fibre cable area. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

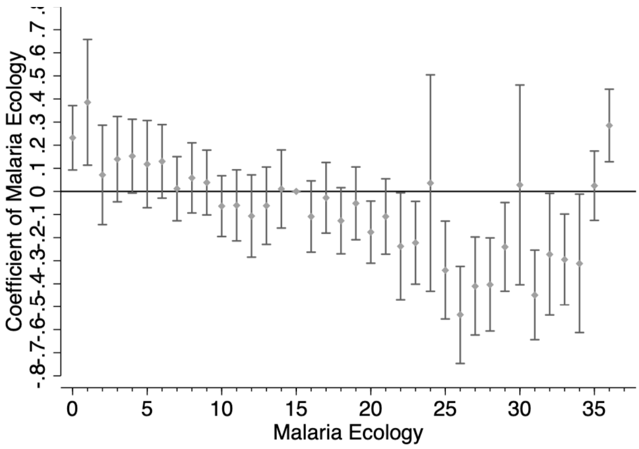



Figure A1: Semi-parametric Regression for Malaria Ecology with 15 as Reference Category

*Note.* Coefficients from the regression of Weibo users (workers) on all binary variables for each malaria ecology index value with 15 as the reference category. Reporting 90% confidence intervals. See text for details.

## References


**Alsan, Marcela.** 2015. "The Effect of the Tsetse Fly on African Development." *American Economic Review*, 105(1): 382–410.

**Bluhm, Richard, Andreas Fuchs, Dreher Axel Strange, Austin, Bradley Parks, and Michael J. Tierney.** 2020. "Connective Financing: Chinese Infrastructure Projects and the Diffusion of Economic Activity in Developing Countries." *CESifo Working Paper*, 8344.

**Cervellati, Matteo, Giorgio Chiovelli, and Elena Esposito.** 2019. "Bite and Divide: Malaria and Ethnolinguistic Diversity." *CEPR Discussion Paper*, 13437.

**Conley, Timothy G.** 1999. "GMM Estimation with Cross-sectional Dependence." *Journal of Econometrics*, 92: 1–45.

**Gething, Peter W, Anand P Patil, David L Smith, Carlos A Guerra, Iqbal RF Elyazar, Geoffrey L Johnston, Andrew J Tatem, and Simon I Hay.** 2011*a*. "A new world malaria map: Plasmodium falciparum endemicity in 2010." *Malaria Journal*, 10(1): 378.

**Gething, Peter W, Thomas P Van Boeckel, David L Smith, Carlos A Guerra, Anand P Patil, Robert W Snow, and Simon I Hay.** 2011*b*. "Modelling the global constraints of temperature on transmission of Plasmodium falciparum and P. vivax." *Parasites & Vectors*, 4(1): 92.

**Howes, Rosalind E, Anand P Patil, Frédéric B Piel, Oscar A Nyangiri, Caroline W Kabaria, Peter W Gething, Peter A Zimmerman, Céline Barnadas, Cynthia M Beall, Amha Gebremedhin, et al.** 2011. "The global distribution of the Duffy blood group." *Nature communications*, 2: 266.

**Howes, Rosalind E, Frédéric B Piel, Anand P Patil, Oscar A Nyangiri, Peter W Gething, Mewahyu Dewi, Mariana M Hogg, Katherine E Battle, Carmencita D Padilla, J Kevin Baird, et al.** 2012. "G6PD deficiency prevalence and estimates of affected populations in malaria endemic countries: a geostatistical model-based map." *PLoS Medicine*, 9(11): e1001339.

**Kiszewski, Anthony, Andrew Mellinger, Andrew Spielman, Pia Malaney, Sonia Ehrlich Sachs, and Jeffrey Sachs.** 2004. "A global index representing the stability of malaria transmission." *The American Journal of Tropical Medicine and Hygiene*, 70(5): 486–498.

**Lujala, Päivi, Jan Ketil Rod, and Nadja Thieme.** 2007. "Fighting over oil: Introducing a new dataset." *Conflict Management and Peace Science*, 24(3): 239–256.

**Matsuura, Kenji, and Cort Willmott.** 2001. "Terrestrial Air Temperature and Precipitation: Monthly and Annual Time Series (1950 - 1999)."

**Murdock, George Peter.** 1959. *Africa: Its Peoples and Their Culture History.* New York:McGraw-Hill.

**Murdock, George Peter.** 1967. *Ethnographic Atlas.* Pittsburgh:University of Pittsburgh Press.

**Nunn, Nathan.** 2008. "The Long-Term Effects of Africa's Slave Trades." *Quarterly Journal of Economics*, 123(1): 139–176.

**Nunn, Nathan, and Leonard Wantchekon.** 2011. "The Slave Trade and the Origins of Mistrust in Africa." *American Economic Review*, 101(7): 3221–3252.

**Piel, Frederic B, Simon I Hay, Sunetra Gupta, David J Weatherall, and Thomas N Williams.** 2013. "Global burden of sickle cell anaemia in children under five, 2010–2050: modelling based on demographics, excess mortality, and interventions." *PLoS Medicine*, 10(7): e1001484.

**Portmann, Felix T, Stefan Siebert, and Petra Döll.** 2010. "MIRCA2000—Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling." *Global Biogeochemical Cycles*, 24(1): 1–24.

**Ramankutty, Navin, Jonathan A Foley, John Norman, and Kevin McSweeney.** 2002. "The global distribution of cultivable lands: current patterns and sensitivity to possible climate change." *Global Ecology and Biogeography*, 11(5): 377–392.

**Uchida, Hirotsugu, and Andrew Nelson.** 2010. "Agglomeration index: Towards a new measure of urban concentration." *UNU-WIDER Working Paper*, 2010/29.