Social Adaptation to Diseases and Inequality: Historical Evidence from Malaria in Italy

Paolo Buonanno
Elena Esposito
Giorgio Gulino

November 2020

Barcelona GSE Working Paper Series

Working Paper nº 1220

Social Adaptation to Diseases and Inequality:

Historical Evidence from Malaria in Italy*

Paolo Buonanno
University of Bergamo

Elena Esposito
University of Lausanne

Giorgio Gulino
UPF and Barcelona GSE

November, 2020

Abstract

Disease and epidemics have been a constant presence throughout the history of humanity. In order to mitigate the risks of contagion, societies have long "adapted" to diseases, implementing an array of coping strategies that, in the long run, have had considerable economic and social consequences. This article advances the hypothesis, and documents empirically, that the need to alleviate the dangers of malaria shaped all aspects of life in agricultural communities, from where and how people settled, to how and what they could farm. As larger farms were better equipped to adopt these risk-mitigating strategies, centuries of exposure to malaria had important implications for inequality and wealth distribution.

Keywords: Land Concentration, Inequality, Malaria, Diseases, Human Capital, Long-Run Development
JEL Classification: O40, O13, O15, N30

*Paolo Buonanno: University of Bergamo, paolo.buonanno@unibg.it; Elena Esposito: University of Lausanne, elena.esposito.1@unil.ch; Giorgio Gulino: Universitat Pompeu Fabra and Barcelona GSE, giorgio.gulino@upf.edu. Giorgio Gulino acknowledges financial support from the European Union's Horizon 2020 research and innovation program [Grant 759885]. Giorgio also acknowledges financial support from the Spanish Ministry of Science and Innovation, through the Severo Ochoa Programme for Centres of Excellence in R&D (CEX2019-000915-S). Elena Esposito gratefully acknowledges the support of the Monique de Meuron programme for Academic Promotion.

"Malaria is the basis of all social life; it constrains the production processes and the distribution of wealth more than anything else."

Francesco Saverio Nitti

Prime Minister of Italy from 1919 until 1920

1 Introduction

From the bubonic plague to the Spanish flu, infectious diseases have long played a significant role in human history, in a "continuous battle between us and the microbes" (Crawford, 2007). The most visible consequences of disease outbreaks and epidemics are illness and death. Indeed, an estimated 50 million people died from the Spanish flu (Belser and Tumpey, 2018), whereas malaria alone "may have killed half of all the people that ever lived" (Whitfield, 2002). Beyond these evident outcomes, the responses that societies adopted to reduce the risk of contagion also had considerable and long-lasting effects. Yet, despite its potential relevance, the ramifications and costs of this "social adaptation" to infectious diseases has been much less studied. Nonetheless, efforts to reduce the dangers of contagion indubitably shaped the social and economic structures of societies. In agricultural communities, for example, the constant need to avoid infection affected a range of important aspects of the economy, including where people settled over a given territory, and how and what they farmed. Such strategies profoundly affected production, with potential important consequences for the distribution of income and wealth.

This paper explores the implications of the social adaptation to diseases. Specifically, we examine the effects of behavioral reactions to the deadliest disease known to humankind: malaria. We present evidence from Italy at the turn of the twentieth century, where malaria was responsible for about 20,000 deaths and 2 million infections per year. The need to reduce the risk of contagion and death shaped all aspects of life in Italy, including the strcture of towns, crop choices, and agrarian labor arrangements. Larger farms were better able to cope with the dangers of malaria, implementing risk-mitigating strategies. The threat of sickness and death, in fact, increased the hazards faced by credit-constrained small land-owners. Large farms could also better absorb health shocks to their labor force by hiring temporary external workers, utilizing management practices that minimized labor-related transaction costs. Finally, large farms could more easily diversify production to limit adverse health shocks, opting for labor-saving crops and those that could be harvested before the peak of the malaria season. Centuries of exposure to malaria thus generated an extremely unequal distribution of land in municipalities ravaged by the disease.

We begin our analysis by documenting the relationship between historical malaria prevalence and land inequality across Italian municipalities at the turn of the 20th century. To this end, we assemble unique information on historical land concentration and malaria prevalence at a very high spatial resolution for 7,830 municipalities. Since malaria prevalence may have been higher in poorer, more remote places, historically less subject to drainage and public investments, we rely on an instrumental variable strategy for identification. Specifically, we instrument historical malaria prevalence with an index of *predicted* malaria risk, building on the fact that transmission is possible only under a very specific combination of weather conditions. We base our prediction measures on epidemiological studies that model the relationship between climate and key aspects of the malaria transmission cycle, and reconstruct these using historical data on seasonal temperatures at a high spatial resolution.

Results consistently indicate historical malaria prevalence to be a strong and relevant driver of land inequality and landownership concentration. Being in a municipality with malaria implies an increase in the Gini index of 0.22, which corresponds to about one-third of the mean land inequality in the sample. Further, historical malaria prevalence reduced the share of land of small farms by 39 percentage points, corresponding to a 77% reduction with respect to the mean. In parallel, wherever malaria reigned, large estates prevailed, doubling the average share of land they occupied. These findings are confirmed when using several alternative predicted measures of malaria prevalence, when flexibly accounting for non-linear effects of climate using the LASSO method, and when exploiting only within province-variation by including province fixed effects.

Indubitably, malaria put agricultural workers at constant risk of illness and death. This health threat added a new crucial source of uncertainty to traditional agricultural hazards, such as droughts and crop failures. An extensive literature in economics argues that large farms are better able to cope with general agrarian risk, due to their greater capacity to absorb shocks and diversify production. Likewise, we contend that such farms were similarly better equipped to manage malaria-related risks. Indeed, they could absorb health shocks to their labor force by hiring temporary external workers, as they were organized to minimize labor-related transaction costs. In addition, large farms could more easily diversify production, including labor-saving farming activities.

We apply these arguments to the data, showing how the pervasive risk of malaria and the need to strategically adapt social and economic life had a variety of consequences. First, exploiting unique data on the spatial location of houses within a municipality, we demonstrate that malaria prevalence impeded farmers from locating their houses near to their fields (i.e., contrary to the agricultural landscape of malaria-free areas). This meant they often concentrated in villages and towns, far from agricultural lands. Second, we show that while malaria-infested municipalities were not less cultivated, they were cultivated differently. Labor-saving crops, with short harvest seasons, possibly preceding the peak malaria season, were much more common in such areas. Third, we present data on labor arrangements, documenting that municipalities where malaria was prevelant had a sizably lower share of family farms that owned their lands or cultivated them under share-cropping agreements. These lands were instead predominantly cultivated by daily laborers, recruited to work in malaria-infested fields only for the key phases of production.

Malaria was eradicated in Italy more than half a century ago. While transmission of the disease ceased, the economic and political consequences of a history of inequality and the extreme land concentration it fostered could well have had persistent consequences. The final part of this study reveals that, indeed, historical malaria prevalence hampered long-term development: historically affected municipalities remain poorer and continue to have comparatively lower average levels of human capital.

In the extensive scholarship looking at diseases and development,[1] this article joins the small, but growing, literature that investigates the long-term consequences and "hidden costs" of *social adaptation* to diseases (Fogli and Veldkamp, 2020; Cervellati et al., 2020). In particular, Cervellati et al. (2019) document, theoretically and empirically, that a history of exposure to malaria increased historical ethnic diversity in Africa, as separation into smaller isolated groups mitigated the burden of the disease. It also contributes to a recent body of work showing how a history of disease exposure can affect long-term development trajectories through its impact on deeply rooted features such as racial hatred (Voigtlander and Voth, 2012), pre-colonial population (Alsan, 2015), distrust in medicine (Lowes and Montero, 2018), and slavery (Esposito, 2019).[2] Inequality, the major long term consequence of the social adaptation to diseases documented here, represents a pressing concern for economists and is a key factor for understanding the process of development (Engermann and Sokoloff, 1997 and 2000; and Galor, Moav and Vollrath, 2002). As such, this paper enriches investigations of the historical and geographical drivers of wealth inequality (Engerman and Sokoloff, 1997 and 2000; Galor, Moav and Vollrath, 2002; Ramcharan, 2010; and Cinnirella and Hornung, 2016).

The remainder of this paper is organized as follows. In Section 2, we summarize the key elements of historical malaria prevalence in Italy and its consequences on farm size and land inequality. We then present the original data collected for this project in Section 3. Section 4 presents the empirical strategy and main findings: baseline results are summarized in Section 4.1, and Section 4.2 looks at risk-coping strategies. The final section concludes.

[1]Inaugurated by the contributions of Bloom and Sachs (1998) and Gallup and Sachs (2001), and subsequently revisited by Acemoglu and Johnson (2001) and Weil (2007), among many others.

[2]Deptris-Chauvin and Weil (2017), Cervellati et al. (2019), and Esposito (2019) focus specifically on the effect of historical malaria exposure.

2 Conceptual Framework

2.1 Historical Background

Malaria in Italy Italy has a long and dramatic history of malaria exposure.[3] The disease was, in fact, such a fundamental part of the Italian agricultural landscape that is was referred to as the "Italian national disease." In fact, the name of the disease itself comes from two medieval Italian words, *mal* and *aria*, meaning "bad air" (Snowden, 2008). No other European country suffered a comparable death toll. In 1882, the Italian government registered 21,000 deaths caused by malaria, while the Netherlands, another traditionally affected country, recorded only 250 deaths per year.[4] Equally grim was the extremely high morbidity and debilitation from the disease (Dobson, 1989). According to an official government survey, in 1882, out of the 25 million inhabitants, 11 million were at constant risk of infection while another 2 million contracted the disease annually (Snowden, 2008).

Notably, the greatest losses were in rural areas and among farmers.[5] As Braudel (1972) points out "The height of the annual malaria epidemic in the summer coincided exactly with the peak of the agricultural season, when vast outlays of heavy outdoor labor were required for the tasks of scything, harvesting, and threshing." Moreover, malaria was a constant risk, like no other disease. Again in the words of Braudel (1972): "the plague... altought greatly to be feared, is only a passing visitor to the Mediterranean. Malaria is permanently installed there. It constitutes the background of the Mediterranean pathology."

Landownership Concentration and Malaria Prevalence Perhaps the second greatest affliction in Italy was land inequality, where small landowners were few and there was a remarkable concentration of territory in large estates. The prevalence of sizable and sometimes even immense farms, consisting of hundreds or even thousands of hectares, has puzzled historians for decades. Indeed, this pattern set Mediterranean Europe apart from Northern Europe, where a free and more equal approach to land distribution prevailed, similar to the northern areas of the United States (Braudel, 1972). As Braudel (1972) puts it: "One of the problems of the Mediterranean, and one of the causes of its traditionalism and rigidity, was that ... land remained under the control of the wealthy."[6]

With the reforms implemented by enlightened monarchs and the Napoleonic administrations,

[3]The first literary sources suggest that in Sicily, malaria was an issue as early as the fifth century BC.

[4]Giovan Battista Grassi (1885). Source: *http://archivio.unict.it/sites/default/files/SECONDA.PDF*

[5]"The life expectancy of farmworkers in non-malarial regions of Italy was 35.7 [...] In malarial zones, by contrast, their life expectancy at birth was only 22.5 years" (Snowden, 2008).

[6]Many of the very large estates had a feudal origin or belonged to the Church. There were, however, also examples of more recent formation, such as the estate of the Baracco family in the Abruzzo region.

the emergence of a land market did create the conditions for a potential redistribution of land and the possibility for agriculatural workers to purchase terrain. Yet, this redistribution did not materialize in malaria-infested areas. Politicians (and, later on, historians) were conscious of the clear geographical connection between land inequality and malaria prevalence. In the words of Franscesco Saverio Nitti, prime minister of Italy from 1919-1920, "Malaria is the basis of all social life; it constrains production processes and the distribution of wealth more than anything else. Malaria is at the root of most important demographic and economic facts; the distribution of property, the distribution of crops, the distribution of population..."

2.2 Malaria Risk and Farm Size

Working in malaria-infested areas exposed agricultural workers to the constant threat of illness and death. As Snowden (2006) observes, the malaria season "coincided exactly with the peak of the agricultural season, when vast outlays of heavy outdoor labor were required," so that "to survive, farmworkers and peasants had to expose themselves to disease. But disease in turn entailed suffering, days of absence, and low productivity." The total health burden for agricultural laborers is difficult to quantify. According to certain estimates, for every malaria death, about 1,300 episodes of illness occurred per year.[7]

A large literature in economics has attempted to estimate the average productivity loss related to malaria prevalence.[8] By far the greatest challenge posed by the disease was the constant uncertainty and risk. An influential literature in economics argues that larger farms were generally more capable of managing agricultural risk (Rosenzweig and Binswanger, 1993; Binswanger, Deininger, and Feder, 1995; Eastwood, Lipton, and Newell, 2010; and Ramcharan, 2010).[9] Indeed, they were more equipped to efficiently absorb shocks and diversify risk in general, but also malaria risk in particular. Labor-related transaction costs for hiring external laborers were lower and large estates could thus better react to health shocks affecting their labor force.

Whenever a worker fell sick, especially during the peak of the agricultural season, he/she had to be promptly replaced. For small farms, the transaction costs of replacement, hiring, and management could be cumbersome. This was further complicated by the protective measure of locating their houses far from malaria-infested fields. In contrast, large farms were organized precisely in such a way as to meet this "inextinguishable demand for labor" (Snowden, 2008), and

[7] Source: Chicago Tribune, May 31, 1914.

[8] Conly (1972) for Paraguay and Bonilla Castro et al. (1991) for Colombia. Experimental evidence from Zambia estimates that malaria-protective technology increases average productivity by 14.7% (Zambia, Fink and Masiye, 2015).

[9] For reviews, see Eastwood, Lipton, and Newell (2010) and Binswanger, Deininger, and Feder (1995). Ramcharan (2010) provides empirical evidence by looking at weather risk and farm size across United States counties.

manage the large turnover by relying on hired labor.[10]

Moreover, large farms could more easily diversify production, allowing to hedge against particularly unhealthy years. Labor-saving crops and farming activities, such as cereals and sheep husbandry, could minimize worker-inputs. Of particular appeal were plants requiring little effort during the malaria season, which could then be harvested before the peak of the fevers (Rossi Doria, 1946).[11]

To date, no quantitative analysis - in Italy or elsewhere - has explored the role of "social adaptation" to disease for the emergence and persistence of land inequality. Efforts to detect a causal relationship are complicated by two main empirical issues. First, malaria-infested areas might have differed from other areas due to geographic and climatic characteristics that could, in principle, have also driven landownership concentration. Secondly, the effects of malaria might have been even more persistent in remote places with a history of dysfunctional governments, limited drainage projects, and inadequate public investments to counter the disease.[12] In what follows, we describe the data and the strategy employed to tackle these empirical challenges.

3 Data

3.1 Historical Land Ownership

We collected new data on land ownership distribution at the municipality level for the years 1929 and 1947.[13] For each municipality, the census provides information on the numbers of farms by size range. We carefully mapped the historical municipalities to present-day municipalities, accounting for changes in their borders over time (see Appendix Section A.1 for details), and built several measures of land inequality.

First, we computed the share of land in the municipality occupied by small, and large estates respectively. SMALL LANDOWNERS is the share of land occupied by farms between 0 and 10

[10]In some cases, large estates relied on a mixture of permanent and temporary laborers (Rossi Doria, 1946; "*Il lavoro fornito da salariati fissi o avventizi*"), where the former were mostly responsible for the management and supervision of the latter. In other cases, absent landowners instead "entrusted [estate management] to a speculative farmer on a short-term lease" (Snowden, 2008). All such arrangements, if with some variation, aimed to allow the landlord and his family to be absent, and to implement a supervision system that could facilitate the rapid hiring of workers during peak season, while minimizing labor-related transaction costs.

[11]In Rossi Doria's (1946) words: ""Under the threat of ... malaria ... the agrarian technique best suited for these lands was little more than natural use of pasture ..., and a precarious and occasional cultivation of cereals ... leaving the countryside as empty as possible of men during the summer months, when the malaria epidemics dominated." (own translation).

[12]Braudel (1972), for example, mentions such efforts on the part of various grand dukes, from Cosimo in Maremma and Val di Chiana; the Duke of Ferrara in the Po Delta; and the viceroy in Naples.

[13]We digitized data for 1929 from the Agrarian Cadastre (*Catasto Agrario*). Part of this information is also available for 1947 from *INEA*. See Appendix Section A.1 for details.

hectares (included) over the total share of land occupied by farms; and LANDMARKS measures the share of land occupied by farms larger than 100 hectares over the total share of land occupied by farms in the municipality. We then further measured land inequality as the Gini index of land inequality in the municipality.

To shed light on the different aspects of social adaptation to the disease, we collected data on SCATTERED HOUSES, which measures the share of all houses located in the countryside as opposed to within the village or town. We also gathered census information on the historical population, crops, and cultivated land, as well as on the contract/labor arrangements prevalent in the municipality.

3.2 Historical Malaria Prevalence

Next, we collected data on historical malaria prevalence and combined this primary data with measures of predicted malaria risk.

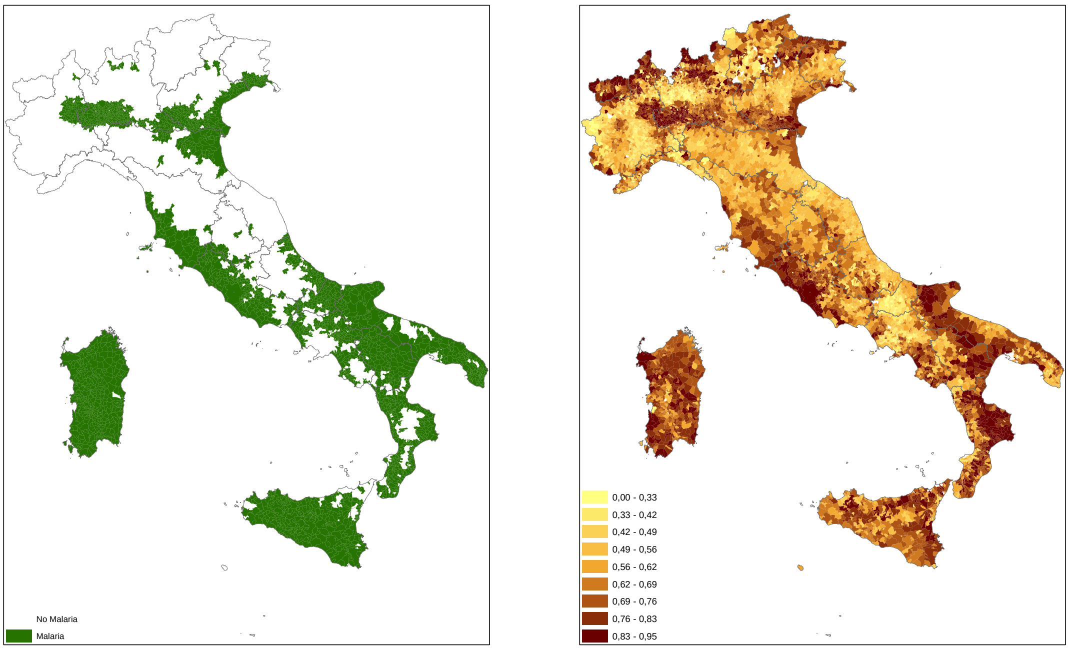
Malaria Prevalence in 1922 To measure the actual historical incidence of malaria, we digitized maps of the disease's prevalence derived from major governmental investigations conducted at the time. As baseline variable, we rely on information from a study led by the Ministry of the Interior dating to 1922. The report contains a detailed map of malaria prevalence across historical municipalities, which we digitized and linked to municipality borders in 2011. A municipality was considered as affected by malaria whenever simultaneous and repeated malaria cases were recorded (within a season) that followed from a local infection. Figure 1 portrays - on the left - the severity of malaria prevalence in 1922 and - on the right - land inequality in 1929. The very strong spatial correspondence between the two is quite apparent.

Malaria Prevalence 1882 According to the map, in 1922 malaria affected 37% of the municipalities in Italy. While some malaria eradication had by then been attempted, the disease's distribution still closely reflected the pattern of prevalence that had characterized the country for centuries. To show this more directly, we also collected data from the first systematic government investigation on the prevalence of malaria in Italy, produced by the Parliamentary Railway Commission led by Senator Luigi Torelli in 1882 (see Appendix Figure A5).

3.3 Predicted Malaria Risk

Predicted Malaria Risk While malaria prevalence was very persistent over the centuries, some areas managed to reduce the disease's presence thanks to drainage and other public infrastructure

FIGURE 1: MALARIA AND LAND INEQUALITY



(a) Malaria Prevalence 1922 (b) Land Inequality 1929

The map on the left portrays historical malaria prevalence as measured by a government investigation in 1922, where green areas indicate malaria-infested municipalities. The map on the right depicts the spatial distribution of land inequality in 1929, measured as the Gini coefficient of agricultural farms. Appendix Section A.6 provides details regarding data sources, data construction, and summary statistics for all the variables discussed in the present Section.

works.[14] Places with less fragmented property may also have benefitted more from anti-malarial strategies such as "*bonifica*" (land reclamation or sanitation). Moreover, it is possible that omitted factors, such as the historical wealth of the municipality, might have characterized malaria-infested areas. However, since malaria transmission is strictly related to bio-climatic characteristics, it is possible to predict the intrinsic higher suitability for transmission, and thus solely exploit this component for identification (thus - ignoring the malaria burden connected to particular institutional, agricultural, or development histories). To this end, we constructed several indexes of *predicted* malaria prevalence, borrowing from decades of research in malaria epidemiology.

For our baseline index of predicted malaria prevalence, we follow the epidemiological literature and use the index proposed by Mordecai et al. (2013).[15] The index models the fundamental link between temperature and malaria transmission by accounting for the effect of temperature on the key elements of disease transmission, such as the parasite development rate, number of eggs

[14]See the examples of Gualtieri (Bonifica Bentivoglio) between the XV and the XIX centuries, Canale della Muzza (XII-XIII centuries), Val di Chiana (XVIII century).

[15]See Fluckiger and Ludwig (2017) for an insightful discussion.

laid per female per day, probability that a mosquito egg survives to become an adult, and larval mosquito development rate (detailed information presented in Appendix Section A.3).

We reconstruct the index adopting three main methodological innovations. First, to better proxy for actual historical malaria risk, we use historical data on seasonal temperatures.[16] Second, given that historical temperatures are only available at a spatial resolution of $0.5C^{\circ}$ x $0.5C^{\circ}$ (larger than our municipalities), we exploit the climatological relationship between temperature and elevation to reconstruct a highly disaggregated measure of temperature.[17] Third, in order to obtain a measure that is not affected by historical population and development, following Fluckiger and Ludwig (2017), we set to zero all model parameters that could in principle be affected by historical human population patterns.[18] For robustness purposes, we produce several versions of the main index so as to verify that our results do not rely on any of these three methodological innovations. Namely, we replicate results using monthly predicted risk instead of seasonal predicted risk with data from Hay et al. (2002), and using temperature data at the original resolution. Furthermore, in order to assess the robustness of our findings to model parameters, we replicate all results with an alternative malaria risk constructed with slightly different parameters and modelling choices: the Malaria Stability index devised by Kiszewski et al. (2004).

4 Malaria and Land Inequality: Empirical Analysis

4.1 Instrumental Variable Estimates

Empirical Strategy To uncover an effect of malaria prevalence on land inequality, we start by presenting cross-municipality correlations that account for an extremely rich set of geo-climatic covariates. Formally, we estimate the following baseline specification:

$$Y_{cr} = \beta_0 + \beta_1 M_{cr} + \gamma \mathbf{X_{cr}} + \theta_\mathbf{r} + \varepsilon_\mathbf{cr}$$

where Y_{cr} measures land inequality in municipality c in region r. M_{cr} measures historical malaria prevalence, a dummy variable taking on the value of 1 if malaria affected the municipality, 0 otherwise. θ_r stands for 20 region fixed effects. X_c includes a rich set of exogenous controls.

To deal with potential measurement error in estimating historical malaria prevalence, as well

[16] As baseline, we use seasonal averages from 1500 to 1922, produced by Xoplaki et al. (2005). As a robustness check, we use monthly data from 1900 to 1922 using CRU CL 2.0 data from Hay et al., (2002).

[17] Due to the so-called environmental lapse rate (reviewed in Appendix Section A.3), every additional meter of elevation brings about a decrease in temperature (generally assumed to be around 0.006 degrees Celsius). This information allows to produce an extremely disaggregated measure of historical average seasonal temperature in the municipality.

[18] See the parameters for the constuction of the index in Appendix Section A.3.

as potential omitted variable bias and reverse causality, we instrument actual historical malaria prevalence with an index of predicted malaria risk.[19] The index, and its variations, will capture only the exogenous component of malaria exposure, as it is constructed in order to predict malaria risk based on a set of specific highly non-linear seasonal climatic conditions, as detailed in Section 3.2. To avoid our predicted index of malaria risk affecting agricultural production patterns through correlation with omitted municipal climatic/geographical features, we control for an extremely rich set of covariates. This includes average temperatures and precipitation for the four seasons (winter, spring, summer, and fall) given that a particular climate might be salient, especially during the growing season or close to harvest.[20] Further, we add a set of controls aiming to proxy for the agricultural suitability of the municipality. Following Galor and Ozak (2016), we include pre-1500 potential caloric yield, which measures the yield potential under a low level of inputs and rain-fed agriculture.[21] In order to measure the multiple dimensions of soil suitability, we also include a widely used measure of suitability that abstains from considering caloric yield: land suitability by Ramankutty (2002). In addition to this, we account for wheat suitability, one of the main crops produced across the sample of municipalities. We then control for the share of sandy soils in the unit, demonstrated to be a predictor of large estates (see Boserup, 1965; and Cinnirella and Hornung, 2016).

We also include a wealth of additional geographic controls that proxy for remoteness, accessibility, flatness of the land, and other locational characteristics, including: latitude and longitude, whether the municipality borders the sea, a river and/or a lake, distance from the sea, access to Roman roads, elevation, standard deviation of elevation, and ruggedness.[22] Since malaria might have been prevalent in flatlands and alluvial plains, we control for the presence of swamps, and for the intrinsic risk of floods and landslides.

We account for spatial correlation of the errors in two ways. First, given the historical relevance of provinces in our setting, we cluster standard errors at the provincial level. Our sample comprises 107 provinces, with an average of 71 municipalities per province. Second, we account for spatial correlation across nearby units using Conley standard errors. Given that the climatic variables,

[19] Our first stage regression will be specified as follows:

$$M_{cr} = \beta_0 + \beta_1 \hat{M}_{cr} + \gamma \mathbf{X_{cr}} + \theta_\mathbf{r} + \varepsilon_\mathbf{cr}$$

where $\hat{M}_{cr}$ is our index of predicted malaria risk.

[20] In extended specifications, we account for both the linear and squared effect of seasonal temperature and precipitation, higher order terms and the full of set of interactions.

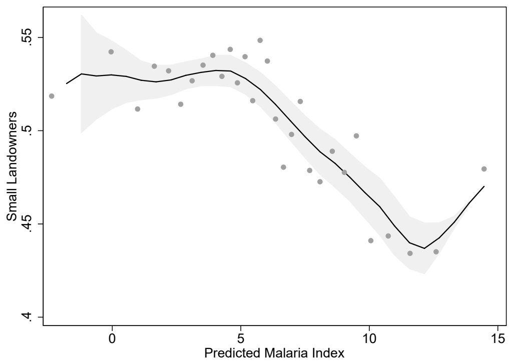
[21] Accounting for post-1500 caloric yield would provide virtually the same results.

[22] Protective settlement patterns could also have emerged for defensive purposes, as shown by Dincecco and Onorato (2016) and Accetturo et al. (2019). We do not include in the specification either the surface area or population of the municipality as these are potentially outcomes and thus bad controls. That said, results are not affected by the inclusion of population, population density, and surface in Km^2 of the municipality.

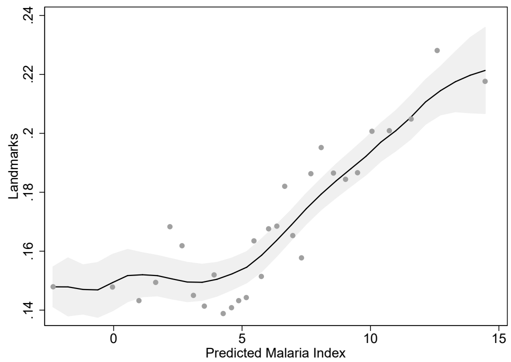
seasonal temperature, and precipitation are available at a resolution that broadly corresponds to 50 square km, we choose 50 km as our baseline cutoff threshold, and present robustness checks for larger thresholds.

Our index of predicted malaria risk, measured in our baseline specifications with the index by Mordecai et al. (2013), is standardized in order to have a 0 mean and a unitary standard deviation. First stage regressions document that the instrument is an relevant predictor of historical malaria prevalence (Appendix Section B.1.2).

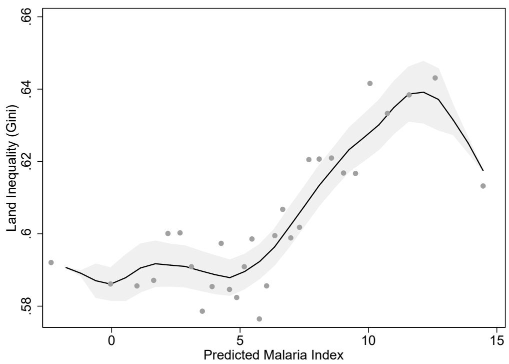
FIGURE 2: MALARIA RISK AND LAND INEQUALITY - LOCAL POLYNOMIAL REGRESSIONS



(a) Small Landowners 1929



(b) Landmarks 1929



(c) Land Inequality 1929

Local polynomial regressions. The dependent variable is: a) the share of small farms, b) the share of large farms and c) the Gini index of land inequality in the municipality in the bottom panel (Source: Agricultural Census, 1929). *Predicted Malaria Risk* measures the predicted risk of malaria transmission constructed following Mordecai et al. (2013). Specifications include region fixed effect and the full set of baseline controls, as detailed in Section 4.1.

Main Results Figure 2 allows to visualize the non-parametric relationship between our Predicted Malaria Risk index and the share of small landowners, large landowners, and land inequality respectively, by local polynomial regressions. We observe a positive relationship between malaria risk and land inequality, as malaria reduces (increases) the share of land occupied by small (large) farms.

These results are confirmed when looking at OLS and IV estimates, summarized in Table 1. In Panel A, the dependent variable is the share of land of small farms, in Panel B the share of land of large farms, and in Panel C the land inequality in the municipality. Columns 1 and 2 look at the OLS coefficient for historical malaria prevalence, columns 3 and 4 at the OLS coefficient for the reduced form effect of predicted malaria risk, and columns 5 and 6 report instrumental variable estimates for historical malaria prevalence. The odd columns do not include the controls, while the even columns include the full set.

All in all, the IV coefficients tend to be larger than OLS ones, suggesting that the OLS is biased towards zero. This potentially implies a large role for measurement error in terms of the inaccuracy of historical malaria prevalence.[23] Reverse causality might also bias the OLS coefficient downward, if places with more fractioned property, and consequently greater coordination problems, had a lower probability of implementing public infrastructure, *bonifica*, and drainage. We take the coefficient estimates of column 6 as our preferred specification. The size of the coefficient suggests a very important quantitative effect. Specifically, being in a muncipality with malaria reduced the share of land with small farms by 39 percentage points, which corresponds to a 77% reduction with respect to the mean. On the other hand, historical malaria increases the share of land with large farms by 32 percentage points, more than doubling the average share. When looking at the full structure of inequality in Panel C, we observe about a 0.22 increase in the Gini index, which corresponds to about one-third of the average land inequality in the sample.

Robustness Appendix Section B.1.3 confirms the results with all versions of our baseline index of predicted malaria risk, as well as with the Malaria Stability Index by Kiszewski et al. (2004). Section B.1.4 presents findings for malaria measured in 1882, whereas Section B.1.5 shows that the effect is larger the stronger the severity of the disease and the lengthier the malaria season. Additional results with variations of the dependent variable are presented in (Section B.1.6) and (Section B.1.7): respectively land inequality measured in 1947 and share of land occupied by farms of different sizes. We then verify the findings with an array of alternative specifications, including province fixed effects (Section B.1.9), examining sub-samples with heterogeneous geography (B.1.10), and exploiting the high-dimensional LASSO method presented by Belloni, Chernozhukov, and Hansen (2014) to more flexibly account for the effect of geography (Section B.1.11). Figure 3 provides a summary of the main robustness results, showing alternative indexes of predicted malaria risk and specifications with and without province fixed effects.

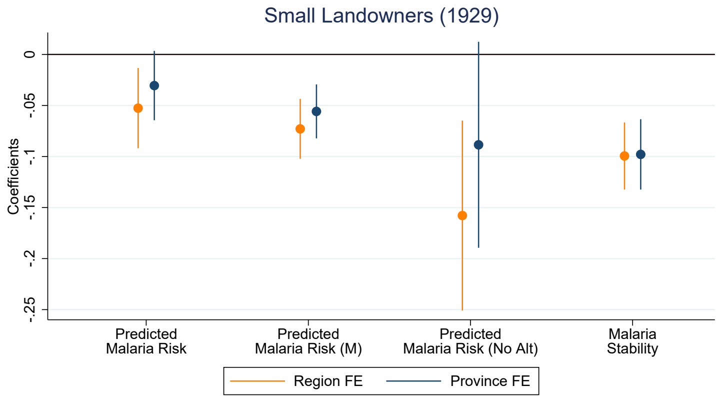
[23] Indirect evidence of this comes from the fact that when malaria is measured using even older sources, i.e., malaria prevalence in 1882, the distance between the IV and OLS coefficient further widens.

Table 1: LAND INEQUALITY AND MALARIA PREVALENCE

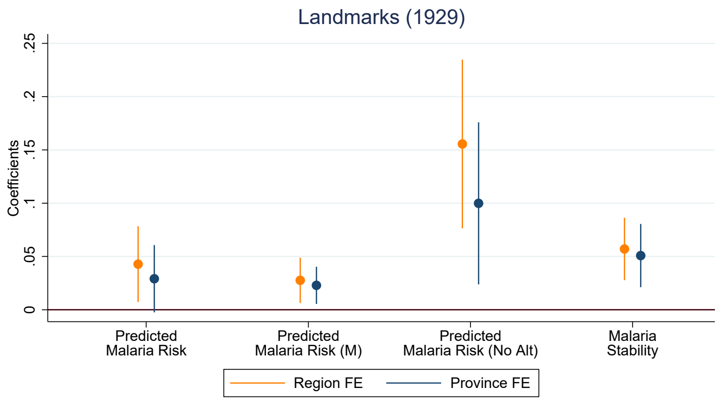
Panel A	SMALL LANDOWNERS (1929)					
	(1) OLS	(2) OLS	(3) ITT	(4) ITT	(5) IV	(6) IV
MALARIA (1922)	-0.184***	-0.169***			-0.208***	-0.388***
	(0.031)	(0.032)			(0.058)	(0.124)
	[0.056]	[0.061]			[0.084]	[0.143]
PREDICTED MALARIA RISK			-0.052***	-0.053**		
			(0.018)	(0.020)		
			[0.027]	[0.029]		
Mean Dependent	0.50	0.50	0.50	0.50	0.50	0.50
F-statistic Instrument					112	19
Observations	7,830	7,830	7,830	7,830	7,830	7,830
R-Square	0.11	0.34	0.04	0.30	0.10	0.26
Panel B	LANDMARKS (1929)					
MALARIA (1922)	0.086***	0.085***			0.006	0.315***
	(0.024)	(0.027)			(0.046)	(0.107)
	[0.033]	[0.038]			[0.054]	[0.114]
PREDICTED MALARIA RISK			0.002	0.043**		
			(0.012)	(0.018)		
			[0.014]	[0.023]		
Mean Dependent	0.17	0.17	0.17	0.17	0.17	0.17
F-statistic Instrument					112	19
Observations	7,830	7,830	7,830	7,830	7,830	7,830
R-Square	0.03	0.32	0.00	0.31	0.00	0.20
Panel C	LAND INEQUALITY (1929)					
MALARIA (1922)	0.105***	0.080***			0.129***	0.218***
	(0.015)	(0.018)			(0.024)	(0.068)
	[0.027]	[0.033]			[0.034]	[0.074]
PREDICTED MALARIA RISK			0.032***	0.030**		
			(0.008)	(0.011)		
			[0.012]	[0.016]		
Mean Dependent	0.60	0.60	0.60	0.60	0.60	0.60
F-statistic Instrument					112	19
Observations	7,830	7,830	7,830	7,830	7,830	7,830
R-Square	0.12	0.30	0.05	0.27	0.12	0.19
Geographical Characteristics	×	√	×	√	×	√
Land Suitability Indexes	×	√	×	√	×	√
Temperature (c°) & Rain (mm)	×	√	×	√	×	√
Region FE	×	√	×	√	×	√

OLS and IV estimates. The unit of observation is the municipality. The *dependent variable* is the share of small farms in Panel A, the share of large farms in Panel B and the Gini index of land inequality in the municipality in Panel C, based on the 1929 agricultural census. *Malaria 1922* is a dummy variable equal to one in municipality with malaria transmission, and zero otherwise. *Predicted Malaria Risk* measures the predicted risk of malaria transmission constructed following Mordecai et al. (2013). Complete data descriptions, data sources and summary statistics are presented in Appendix Section A.6.1 and A.6.2. Robust standard errors clustered at the provincial level are in round parentheses and Conley standard errors (50 km cutoff) are reported in square brackets. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

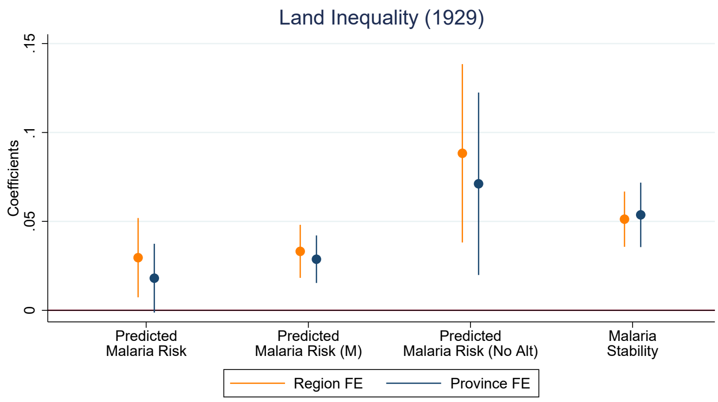
FIGURE 3: ALTERNATIVE MALARIA INDEXES AND PROVINCE FIXED EFFECTS



(a) Small Landowners 1929



(b) Landmarks 1929



(c) Land Inequality 1929

The dependent variable is the share of small farms (a), the share of large farms (b), and the Gini index of land inequality in the municipality (c) (Source: Agricultural Census, 1929). *Predicted Malaria Risk* measures the predicted risk of malaria transmission constructed following Mordecai et al. (2013). *Predicted Malaria Risk (M)* is the same index constructed following Mordecai et al. (2013), using monthly data on temperature from Hay et al., (2002). *Predicted Malaria Risk (No Alt)* is the same index constructed following Mordecai et al. (2013) using seasonal data on temperature with no adjustments based on elevation (i.e., environmental lapse rate, see Appendix A.3) *Malaria Stability* measures the stability of malaria transmission by Kiszewski et al. (2004). Specifications include the full list of baseline controls, as detailed in Section 4.1. Coefficients in orange correspond to specifications with region fixed effects, coefficients in blue to those with province fixed effects. Complete data descriptions, data sources and summary statistics are presented in Appendix Section A.6.1 and A.6.2.

4.2 Malaria Risk and Coping Strategies

We now turn to the coping strategies implemented to mitigate the risk of contagion, which in turn fostered the the observed land inequality.

Living in places with high malaria prevalence was dangerous. Adults frequently fell ill, and death was a likely possibility for children (Dobson, 1989). Farmers consequently relocated to safer ground, far from their mosquitoe-infested fields. Results summarized in the top panel of Figure 4 shows how malaria-affected municipalities were not, in fact, less-densely inhabited, but were occupied differently. Indeed, housing location had the specific purpose of reducing proximity between humans and mosquitoes. Buildings were all concentrated in small centers and towns - where drainage could be more easily performed - as opposed to being dispersed across the fields. Although the role of mosquitoes in the transmission process wasn't discovered until the late nineteenth century, the connection between marshes and stagnant waters was made much earlier (see, among others, Webb, 2009). In the average Italian municipality, about 29 percent of the households were scattered in the countryside. A one standard deviation increase in malaria risk brought about a 14% reduction with respect to this mean share. Yet, while this type of settlement pattern offered protection from contagion, and could be life-saving for children and pregnant women, it increased the distance from the fields for farmers and agricultural laborers.

Similarly, malaria-infested municipalities were not less cultivated, but were cultivated differently, privileging labor-saving farming activities. The middle graph in Figure 4 shows that, when looking at predicted malaria risk, malaria presence does not bring about any sizable or significant change in the share of cultivated land or in the share of unproductive land. The disease did, however, drastically change the way these lands were used, favoring crops that required as little labor input as possible. Among the different crops produced in Italy, the most labor-intensive were fruits, the notable example being vineyards. At the other end of the spectrum, pastures were the most labor-saving option. As expected, malaria is a strong predictor of the presence of pastures and is associated with a lower share of land devoted to vineyards and other fruits.[24]

Under constant returns to scale, the family farm would be the more efficient labor unit (Binswanger, 1995), but could it not thrive in malaria-infested areas. Indeed, such farms were unable to provide the required labor, as children and minors were particularly at risk of infection and because, as a protection measure against contagion, they lived miles away from the fields. In addition, diversification of production to include labor-saving crops was often not the most profitable choice and could not guarantee subsistence. The bottom graph in Figure 4 shows that municipalities with

[24] In terms of magnitudes, a one standard deviation increase in predicted malaria risk brought about a 0.22 standard deviation increase in the land devoted to pastures and a 0.10 decrease in the land devoted to vineyards and fruits, although the coefficient is not precisely estimated.

higher malaria risk had a larger share of landless farmers,[25] and a lower share of landowners and families cultivating the land under share-cropping agreements. Larger farms were more easily able to manage malaria-related risks, since they could absorb health shocks to their labor force by hiring temporary external workers during the peak agricultural season. As shown in Figure 4, a one standard deviation increase in predicted malaria risk sees a 0.36 standard deviation increase in the share of daily workers employed.

5 Conclusion and Final Remarks

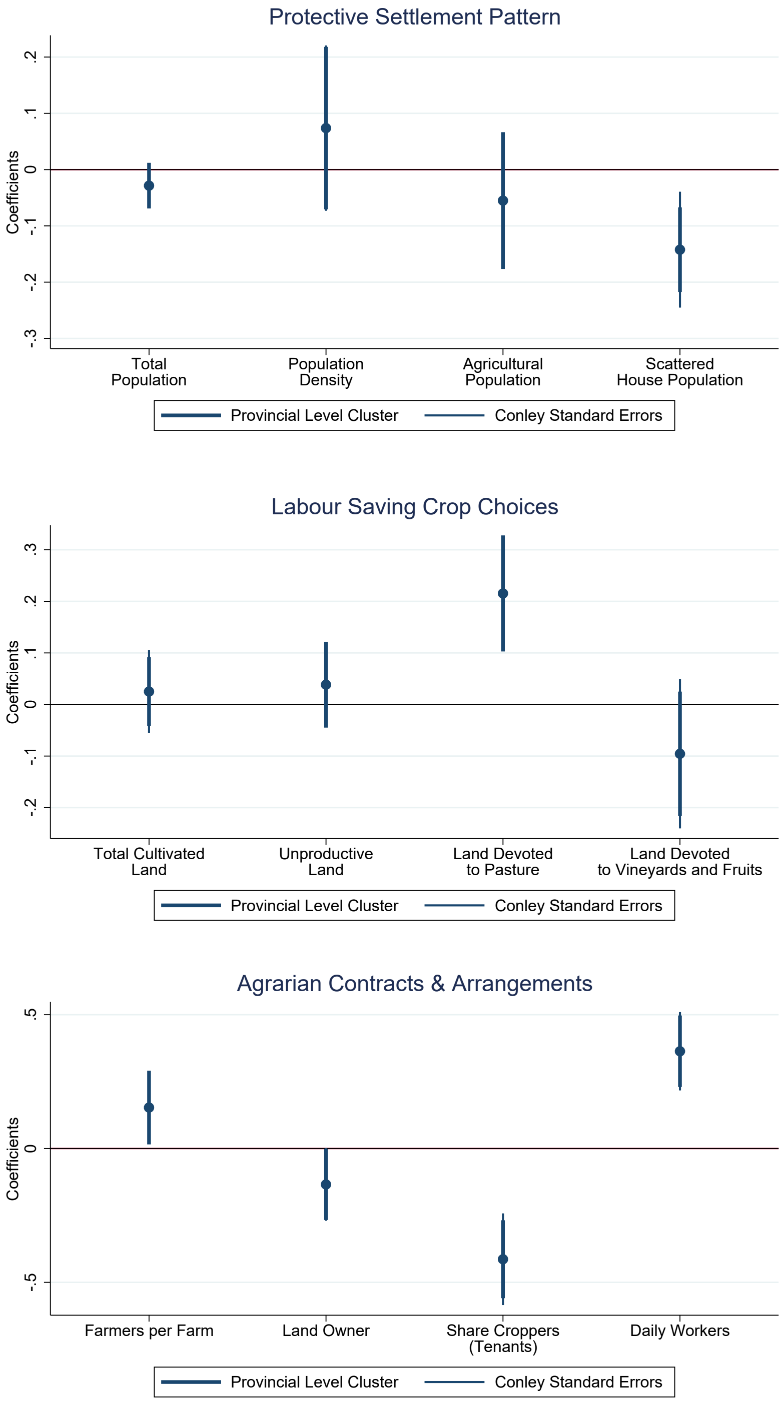
Recent medical developments and experimental evidence suggest that a malaria vaccine granting full protection against the disease might soon become available. The potential social gains in terms of reduced mortality and morbidity, higher human capital accumulation, and productivity of affected areas are enormous. Yet, the vaccine won't heal all the wounds caused by a history of malaria exposure. In fact, malaria - the deadliest killer in human history - has represented such a burden that entire economic and societal structures have been built in order to cope with the disease.

This paper shows that malaria profoundly shaped affected areas, constraining production processes and the historical distribution of wealth. Thanks to new data on historical malaria prevalence, predicted malaria risk, and the size of land estates, our results reveal the disease to be a major driver of historical land inequality. Using micro-data, we demonstrate that settlement patterns, cultivated land, and crop choices were all responses to a need to reduce the dangers of malaria. In particular, these mitigative strategies favored the emergence of substantial land inequality.

Malaria was eradicated in Italy more than half a century ago. While transmission of the disease ceased, the economic and political consequences of a history of inequality and extreme land concentration could well have persistent consequences. The results presented in Appendix Section B.2 suggest that historical malaria prevalence has cast a long shadow on the development of affected areas, which has endured long after complete erasure of the disease.

[25] Measured as the number of farmers per farm, following Vollrath (2007).

FIGURE 4: MALARIA RISK AND COPING STRATEGIES



The graphs report coefficient estimates of the effect of *Predicted Malaria Risk* on a host of outcomes in fully conditioned empirical specifications that account for all fixed effects and baseline covariates. *Predicted Malaria Risk* measures the predicted risk of malaria transmission constructed following Mordecai et al. (2013). Complete data descriptions, data sources and summary statistics are presented in Appendix Section A.6.1 and A.6.2 and extended specifications presented in Appendix Section B.1.12.

References

Accetturo, A., M. Cascarano, G. de Blasio, et al. (2019): "Pirate Attacks and the Shape of the Italian Urban System," Discussion paper, Department of Economics and Management.

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

Alsan, M. (2015): "The effect of the tsetse fly on African development," *American Economic Review*, 105(1), 382–410.

Belloni, A., V. Chernozhukov, and C. Hansen (2014): "High-dimensional methods and inference on structural and treatment effects," *Journal of Economic Perspectives*, 28(2), 29–50.

Binswanger, H. P., K. Deininger, and G. Feder (1995): "Power, distortions, revolt and reform in agricultural land relations," *Handbook of Development Economics*, 3, 2659–2772.

Bloom, D. E., J. D. Sachs, P. Collier, and C. Udry (1998): "Geography, demography, and economic growth in Africa," *Brookings Papers on Economic Activity*, 1998(2), 207–295.

Bonilla-Castro, E., and P. R. Sehk (2005): *Más allá del dilema de los métodos: la investigación en ciencias sociales*. Editorial Norma.

Boserup, E. (1965): *The conditions of agricultural progress*. Aldine Publishing Company, Chicago.

Chiovelli, G., M. Cervellati, and E. Esposito (2019): "Bite and divide: malaria and ethnic diversity," Discussion paper, Working Paper.

Cinnirella, F., and E. Hornung (2016): "Landownership concentration and the expansion of education," *Journal of Development Economics*, 121, 135–152.

Conly, G. N. (1972): "The impact of malaria on economic development," *The American Journal of Tropical Medicine and Hygiene*, 21(5_Suppl), 668–674.

Depetris-Chauvin, E., and D. N. Weil (2017): "Malaria and Early African Development: Evidence from the Sickle Cell Trait," *The Economic Journal*, 128(610), 1207–1234.

Dincecco, M., and M. G. Onorato (2016): "Military conflict and the rise of urban Europe," *Journal of Economic Growth*, 21(3), 259–282.

Dobson, M. J. (1989): "History of malaria in England.," *Journal of the Royal Society of Medicine*, 82(Suppl 17), 3.

Eastwood, R., M. Lipton, and A. Newell (2010): "Farm size," *Handbook of Agricultural Economics*, 4, 3323–3397.

Esposito, E. (2019): "Side effects of immunities: the African slave trade," Discussion paper.

Fink, G., and F. Masiye (2015): "Health and agricultural productivity: Evidence from Zambia," *Journal of Health Economics*, 42, 151–164.

Flückiger, M., and M. Ludwig (2017): "Malaria suitability, urbanization and persistence: Evidence from China over more than 2000 years," *European Economic Review*, 92, 146–160.

Fogli, A., and L. Veldkamp (Forthcoming): "Germs, social networks and growth," Discussion paper, Review of Economic Studies.

GALLUP, J. L., AND J. D. SACHS (2001): "The economic burden of malaria," *The American Journal of Tropical Medicine and Hygiene*, 64(1_suppl), 85–96.

GALOR, O., O. MOAV, AND D. VOLLRATH (2009): "Inequality in Landownership, the Emergence of Human-Capital Promoting Institutions, and the Great Divergence," *Review of Economic Studies*, 76(1), 143–179.

GALOR, O., AND Ö. ÖZAK (2016): "The agricultural origins of time preference," *American Economic Review*, 106(10), 3064–3103.

KISZEWSKI, A., A. MELLINGER, A. SPIELMAN, P. MALANEY, S. E. SACHS, AND J. SACHS (2004): "A global index representing the stability of malaria transmission," *The American Journal of Tropical Medicine and Hygiene*, 70(5), 486–498.

LOWES, S. R., AND E. MONTERO (2018): "The legacy of colonial medicine in Central Africa," Discussion paper.

MORDECAI, E. A., K. P. PAAIJMANS, L. R. JOHNSON, C. BALZER, T. BEN-HORIN, E. DE MOOR, A. MCNALLY, S. PAWAR, S. J. RYAN, T. C. SMITH, ET AL. (2013): "Optimal temperature for malaria transmission is dramatically lower than previously predicted," *Ecology Letters*, 16(1), 22–30.

NEW, M., D. LISTER, M. HULME, AND I. MAKIN (2002): "A high-resolution data set of surface climate over global land areas," *Climate Research*, 21(1), 1–25.

RAMANKUTTY, N., J. A. FOLEY, J. NORMAN, AND K. MCSWEENEY (2002): "The global distribution of cultivable lands: current patterns and sensitivity to possible climate change," *Global Ecology and Biogeography*, 11(5), 377–392.

RAMCHARAN, R. (2010): "Inequality and redistribution: evidence from US counties and states, 1890–1930," *The Review of Economics and Statistics*, 92(4), 729–744.

ROSENZWEIG, M. R., AND H. P. BINSWANGER (1993): "Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments," *The Economic Journal*, 103(416), 56–78.

ROSSI-DORIA, M. (1946): "La terra: il Latifondo e il Frazionamento," *Centro permanente pei problemi del Mezzogiorno, Dati storici e prospettive attuali. Atti del convegno di studi meridionalistici (Bari, 3-5 dicembre 1944), Canfora e &, Bari.*

SNOWDEN, F. (2006): "The Conquest of Malaria–Italy, 1900–1962. 2006," *New Haven (CT): Yale University PressGoogle Scholar.*

SNOWDEN, F. (2008): *The conquest of malaria: Italy, 1900-1962*. Yale University Press.

VOIGTLÄNDER, N., AND H.-J. VOTH (2012): "Persecution perpetuated: the medieval origins of anti-Semitic violence in Nazi Germany," *The Quarterly Journal of Economics*, 127(3), 1339–1392.

VOLLRATH, D. (2007): "Land distribution and international agricultural productivity," *American Journal of Agricultural Economics*, 89(1), 202–216.

WEBB, J. L. (2009): *Humanity's burden: a global history of malaria*. Cambridge University Press New York.

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

Xoplaki, E., J. Luterbacher, H. Paeth, D. Dietrich, N. Steiner, M. Grosjean, and H. Wanner (2005): "European spring and autumn temperature variability and change of extremes over the last half millennium," *Geophysical Research Letters*, 32(15).

ONLINE APPENDIX - NOT FOR PUBLICATION

Social Adaptation to Diseases and Inequality: Historical Evidence from Malaria in Italy

Abstract

This Online Appendix accompanies the paper "Social Adaptation to Diseases and Inequality: Historical Evidence from Malaria in Italy". Section A present the historical data collection exercise, the construction of the predicted malaria indexes, data sources and data construction details for all the variable of interest. Section B presents additional results complementing the empirical analysis section in the main text.

Contents

A Data

A.1 Historical Land Inequality

We digitize archival data from the Census of Agriculture for 1929 and 1947. Data for 1929 come from the Agrarian Cadastre (*Catasto Agrario*), while the information for 1947 are from INEA (*Istituto Nazionale di Economia Agraria*), 1946-1948*, La distribuzione della proprieta fondiaria in Italia.* Figures A1 and A2 show the original sources of pages containing information on land ownership. To give a concrete sense of the available information, we use as an example the municipality of Ragusa in 1929 (Figure A2). Ragusa had 112 farms large between 0 and 0.24 hectares, 148 farms between 0.26 and 0.5 hectares, 329 between 0.51 and 1 and so forth until the last class that is farms larger than 1,000 hectares. Exploiting this data we construct 3 measures of distribution of land: land inequality, Landmarks and Small Landowners. Unlike the Census of Agriculture of 1929, which contains a rich array of data regarding land distribution and agricultural practices, the Census of Agriculture of 1947 contains only the distribution of farms by size.

Figure A1: Land Ownership in 1929 - the Municipality of Sala Monferrato

— 155 —

A

116. - SALA MONFERRATO

I. – Dati generali

DELLE COLLI

2. *Popolazione* (Censim. 21-IV-1931-IX)	
Presente:	
In complesso	1.263
Per km² di superf. territoriale . .	*165*
Per km² di superf. agr. e forest. .	*172*
Dei centri.	952
Delle case sparse	311
Residente:	
In complesso	1.282
Per km² di superf. territoriale. .	*167*
Per km² di superf. agr. e forest. .	*174*

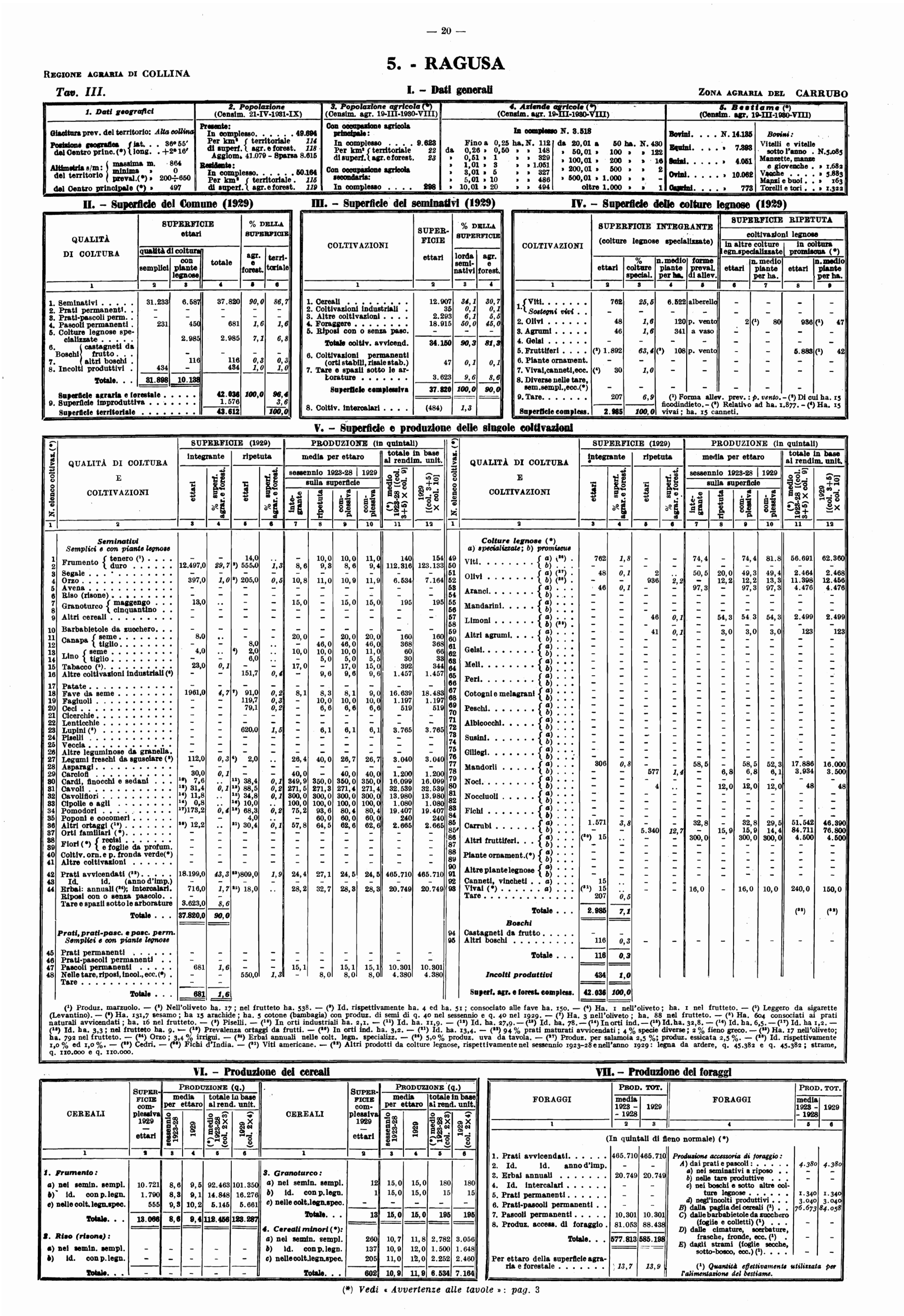
3. *Popolazione agricola* (*) (Censim. 21-IV-1931-IX)		
Posiz. profess. capo famiglia	Famiglie	Componenti
Conduc. terreni propri	186	859
Fittavoli. . . .	3	23
Coloni.	10	40
Giornalieri. . .	9	30
Altri addetti. .	20	85
Totale. . .	228	(¹) 1.037
(¹) Per km.² di superficie territoriale.		*135*
(¹) Per km.² di superficie agr. e forest.		*141*

4. *Aziende agricole* (*) (Censim. agr. 19-III-1930-VIII)					
Classi di ampiezza	Num.	Superf. ha.	Sistema di conduz.	Num.	Superf. ha.
1. Fino a 0,50 ha.	30	9			
2. da 0,51 » 1 »	41	32	Economia diretta.	201	650
3. » 1,01 » 3 »	93	171			
4. » 3,01 » 5 »	59	231			
5. » 5,01 » 10 »	31	213	Affitto. .	5	13
6. » 10,01 » 20 »	3	43			
7. » 20,01 » 50 »	1	22	Colonia .	5	9
8. » 50,01 » 100 »	–	–			
9. » 100,01 » 500 »	–	–	Mista . .	47	149
10. oltre 500 »	–	–			
Totale. . .	258	721	Totale. . .	258	721

di coltura **III. – Repartizione superficie seminativi** **IV. – Superficie e densità dell**

Notes: The picture portrays the scanned copy of one extract from the Agricultural Census of 1929. The page list farm sizes in the Municipality of Sala Monferrato.

Figure A2: Land Ownership in 1929 - Ragusa

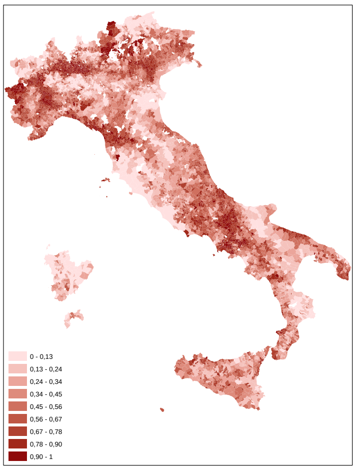


Notes: The picture portrays the scanned copy of one page from the Agricultural Census of 1929 for the city of Ragusa.

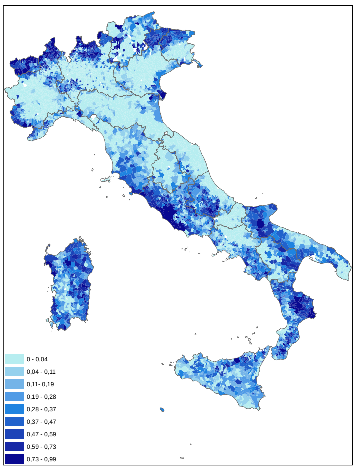
Landmarks and Small Landowners LANDMARKS measure the share of land occupied by farms larger than 100 hectares over the total share of land occupied by farms in the municipality; SMALL LANDOWNERS is the share of land occupied by farms between 0 and 10 hectares (included) over the total share of land occupied by farms.

Land inequality We reconstruct a complete distribution of estates attributing to each farm the mean land size of the category (i.e. for estates from 0 to 0.5 hectares, we associate 0.25 to all farms in this category). For the largest categories, i.e. estates of 500 hectares or more, we assign 500 hectares to all farms of this category. We perform a wide array of checks to verify robustness of results to these choices (using the largest values for each category, etc...). Next, we compute a standard Gini index of inequality based on this reconstructed distribution of estates. Figure A3 maps the spatial distribution of land inequality across municipality in 1929 (graph c), the spatial distribution of large estates (graph b) and small landowners (graph a) in 1929.

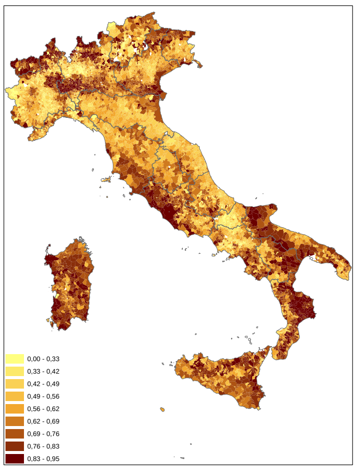
Figure A3: Land Inequality in 1929



(a) Small Landowners



(b) Landmarks



(c) Land Inequality (Gini Index)

The picture portrays the spatial distribution of land inequality in 1929, measured as the share of the municipality occupied by small estates, large estates and the Gini index of land inequality in the municipality (Source: Agricultural Census, 1929).

Mapping to 2011 ISTAT Municipality We map original municipality to municipality borders in 2011 as mapped by ISTAT (National Institute of Statistics).[1] A large effort was devoted to the mapping of historical municipalities with contemporary ones. The database of 2011 contains 8,094 municipalities, about 6500 of them found a perfect match with the names reported on the Census of Agriculture of the two different years. From 1929 and 2011 (and also from 1947 to 2011), several municipalities changed their names, some were formed by the dissolution of large municipalities and others were born from the merging of smaller municipalities. We reconstructed the history of municipalities' borders and their transformation to merge the information as reported in the Census of Agriculture with those contained in the map of 2011. When we had to merge several municipalities into one, we simply summarized the number of farms in each class-size that the various municipalities contained. On the other hand, when an original municipality dissolved into many smaller municipalities, we divided the number of firms in each class-size by the number of new municipalities created, weighed for the surface of the municipality. The final database contains 7,952 municipality in 1929 and 7,292 in 1947.[2]

A.2 Historical Malaria Prevalence

As baseline measure of historical malaria prevalence, we use the report of the Ministry of the Interior (Directorate General of Health) entitled "Malaria in Italy and the results of the anti-malarial fight", drafted in December 1922 and published in 1924 by Libreria dello Stato, Rome. This report attempted to gauge the spread of the disease in the country. In order to construct a measure of historical malaria prevalence relating to municipalities in 2011, we used a GIS software to intersect the digitized polygons (of the original map) with the 2011 municipalities' borders. Wherever a municipality contained at least a share of land with malaria, according to the original map, we considered the municipality to be affected by malaria in 1922. Figure A4, shows the original map and our baseline variable of historical malaria prevalence in 1922 constructed as described.

[1] *Source*: https://www.istat.it/it/archivio/124086.

[2] Note that the Census of Agriculture of 1947 contains missing values for a large number of municipalities. After the second world war, several administrations were not active and the coverage of the census is not complete.

Figure A4: Malaria Prevalence - Ministry of the Interior

(a) Malaria Prevalence in 1922 - Original Map

(b) Malaria Prevalence in 1922 - Mapped to 2001 borders

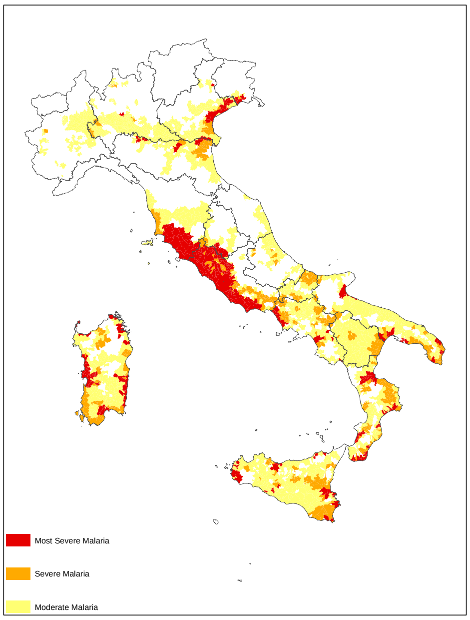
The left panel portrays the original map of Malaria Prevalence in 1922. The right panel portrays our baseline measure of Malaria Prevalence in 1922, retrieved intersecting the original map with municiaplity borders in 2011. The units of observation are municipalities according to the borders in 2011. If a municipality contained at least a share of land with malaria according to the 1922 original map, we considered the municipality affected by malaria transmission.

To have information on malaria distribution before any eradication attempts was ever implemented, we rely on the oldest systematic investigation of the prevalence of malaria in the new kingdom of Italy, which comes from a report of the Parliamentary Railway Commission lead by the senator Luigi Torelli.[3] Soliciting reports on fever from the provincial health councils, Torelli managed to build the first comprehensive map on the spread of malaria in Italy. As Figure A5 shows, the map indicates three main levels of malaria severity by municipality- in three different colors: “moderate”, “severe” and “most severe” (green, yellow and red respectively). As for the other map, in order to construct a measure of historical malaria prevalence relating to municipalities in 2011, we used a GIS software to intersect the digitized polygons (of the original map) with the 2011 municipalities’ borders. Wherever a municipality contained at least a share of land with malaria, according to the original map, we considered the municipality to be affected by malaria in 1882. We attributed to each municipality the degree of severity of the largest polygon. The map (a) in Figure A5 shows the original map, the map (b) shows the measure we derived for municipalities’ borders in 2011.

[3]A systematic investigation was deemed necessary for the problems malaria was creating for the construction of the rail road, because of the appalling rate at which its employees fell ill in vast swathes of the country.

Figure A5: Malaria Prevalence - Luigi Torelli's Map (1882)

(a) Malaria Prevalence in 1882 - Original Map



(b) Malaria Prevalence in 1882 - Mapped to 2011 borders

The left panel portrays the original map of Malaria Prevalence in 1882. The right panel portrays the measure of Malaria Prevalence in 1882 retrieved intersecting the original map with municiaplity in 2011. The units of observation are municipalities according to the borders in 2011. If a municipality contained at least a share of land with malaria according to the 1922 original map, we considered the municipality affected by malaria transmission.

A.3 Predicted Measures of Malaria Exposure

A.3.1 Predicted Malaria Risk: Mordecai et al. (2013)

Our baseline index follows Mordecai et al. (2013), who build an index of malaria risk accounting for accurate nonlinear thermal responses of transmission risk. We reconstruct their index operating two main methodological innovations: i) given the question of interest, we reconstruct the same index using historical data on temperature; ii) we use fine-grained data on altitude to map lower resolution temperature data (0.5x 0.5 degrees) into higher resolution temperatures.

Historical Temperature at High Spatial Resolution Historical seasonal temperature comes from Xoplaki et al. (2005) and are available at a resolution of $0.5° \times 0.5°$ (for data and methods see Luterbacher et al., 2004).[4] To create historical averages of temperature at a high spatial resolution, we follow 2 steps: 1) compute the average historical temperature for spring, summer, autumn and winter using data from 1500 to 1929; 2) use climatological evidence to map data at $0.5° \times 0.5°$

[4]European temperature fields: $25°$ W $-$ $40°$ E; $35°$ N $-$ $70°$ N.

resolution to a more fine-grained resolution, exploiting the high resolution of elevation maps.

Temperature and altitude are stricly related. More precisely, temperature tends to decrease with increasing altitudes, regularity known as the **environmental lapse rate**. An extensive literature from multiple fields has been modelling the relationship between temperature and elevation (see, among many others, reports retrived from the NASA archives by Brombacher, 1963 and Justus and Woodrum, 1972). Consequently, epidemiologists have been exploiting estimates on the relationship between altitude and longitude to better predict malaria incidence, see for instance Bodker et al. (2003). Estimates range around 0.006 lower degrees for every additional meter of elevation, which is now a standard assumed value in the literature (see, for instance, Rumpf et al., 2018). We therefore adjust our original temperature data re-scaling municipality-level temperature by the median elevation of the municipality.

Basic Reproductive Number Since mosquito and parasite vital rates are fundamentally linked to temperature, Mordecai et al. (2013) compute malaria risk as a non-linear function of temperature. Malaria risk is often measured through the Basic Reproductive Number (R_0), which defines the number of cases that arise in a population of susceptible hosts once a new case is introduced.

Mordecai et al. (2013) models R_0 as:

$$R_0(T) = \left(\frac{a(T)^2 bc(T) e^{\frac{-\mu(T)}{PDR(T)}} EFD(T) p_{EA}(T) MDR(T)}{Nr\mu^3(T)} \right) \tag{1}$$

where (T) denotes a temperature-sensitive parameter response fitted from the data, PDR is the parasite development rate, EFD is the number of eggs laid per female per day, p_{EA} is the probability that a mosquito egg survives to become an adult, and MDR is the larval mosquito development rate. Moreover, a is the per-mosquito biting rate, bc is vector competence (the product of the proportion of the bites by infective mosquitoes that infect susceptible humans and the bites by susceptible mosquitoes on infectious humans that infect mosquitoes), μ is the adult mosquito mortality rate, N is human density and r is the rate at which infected humans recover and acquire immunity.

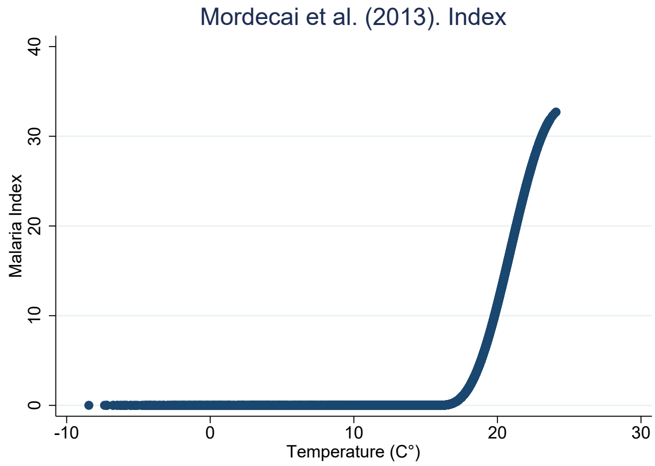
To ensure that our measure is not endogenous with respect to the human population, we adopt the same strategy as Fluckinger and Ludwig (2017) and hold N, r, a, b and c constant ($N = r = a = b = c = 1$). The equation simplifies into:

$$R_0(T) = \left(\frac{e^{\frac{-\mu}{PDR}} \times EFD \times P_{EA} \times MDR}{\mu^3} \right)^{\frac{1}{2}} \tag{2}$$

For the parametrization of the remaining parameters, we follow Mordecai et al. (2013).

Mosquito Development Rate: $MDR = 0.000111 * T(T - 14.7)(34 - T)^{\frac{1}{2}}$

Figure A6: Malaria Risk and Temperature



The graph plots the relationship between *Predicted Malaria Risk* constructed following Mordecai et al. (2013) and average yearly temperature.

Egg-to-adult survival probability:$p_{EA} = [-0.00924T^2 + 0.453T - 4.77]$

Egg laid per adult female per day:$EFD = [-0.153T^2 + 8.61T - 97.7]$

Parasite development rate: $PDR = 0.000111 * T(T - 14.7)(34.4 - T)^{\frac{1}{2}}$

Daily adult survival probability: $p = [-0.000828T^2 + 0.0367T + 0.522]$ and its logarithmic transformation: $\mu = log(p)$.

Predicted Malaria Risk and Temperature Figure A6 depicts the relationship between malaria risk and average summer temperature. In Italy, historically, malaria risk was absent in areas with average temperature lower than 18° and then it was rapidly increasing. In the left panel Figure A7 we portrays the Predicted Malaria Risk across Italian municipality.

A.3.2 Malaria Stability

As an alternative predicted measure of malaria exposure, we employ the Malaria Stability index devised by Kiszewski et al. (2004). The index measure predicted stability of malaria transmission by combining long-average climatic information on temperature and precipitation with the biology of prevalent mosquito vectors. In Italy we find two prevalent mosquito vectors: *Anopheles Labranchiae* in the Center and South, and *Anopheles Atroparvus* in the North.

The predicted stability of malaria infection is therefore predicted as a function of:

- **Mosquitoes**' characteristics:
  - proportion biting people p (0-1)
  - daily survival rate a (0-1)

According to the formula:

$$\sum_{m=1}^{12} \frac{a_{im}^2 p_{im}^E}{-lnp_{im}}$$

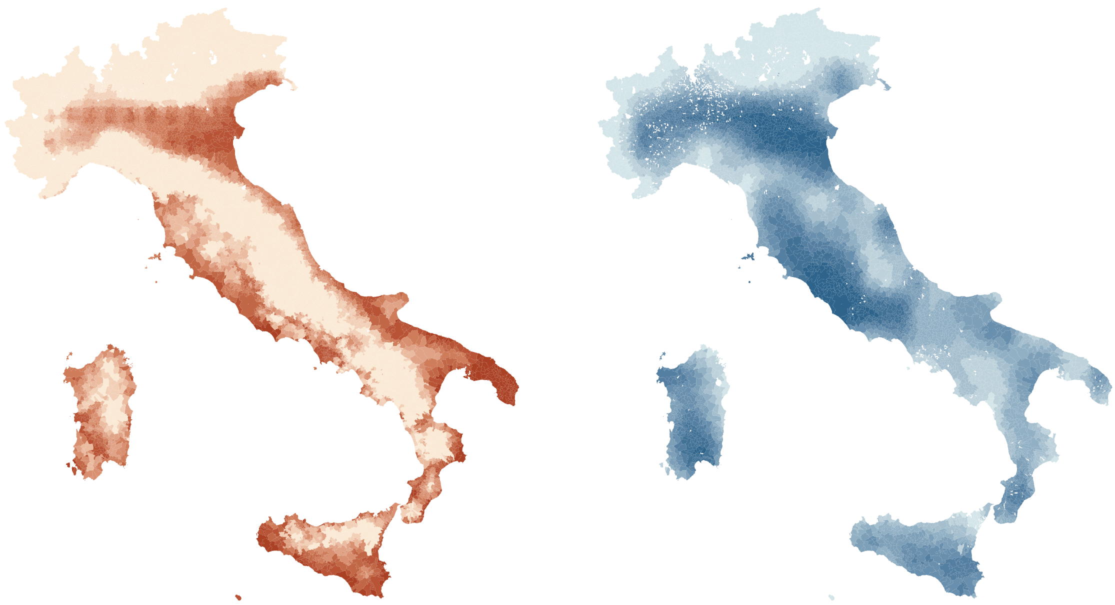
where m stands for month and i for prevalent mosquito vector in the region

- **Temperature** t: trough E, which measures the length of extrinsic incubation period ($E = \frac{111}{t-16}$ for falciparum and $E = \frac{105}{t-14.5}$ for vivax)

- **Precipitation**: a minimum required precipitation of 10 mm per month (threshold for the previous month)

The original index is available at a resolution of 0.5° × 0.5°. Through a bilinear interpolation, we resample the index to a resolution 0.1° × 0.1°. Next, we associate to each municipality the malaria stability measured in the closer 0.1° × 0.1° cell. The right panel of Figure A7 portrays the Malaria Stability index across Italian municipalities.

Figure A7: Predicted Malaria Risk - Indexes



(a) Predicted Malaria Risk - Mordecai et al. (2013)

(b) Malaria Stability - Kiszewski et al. (2004)

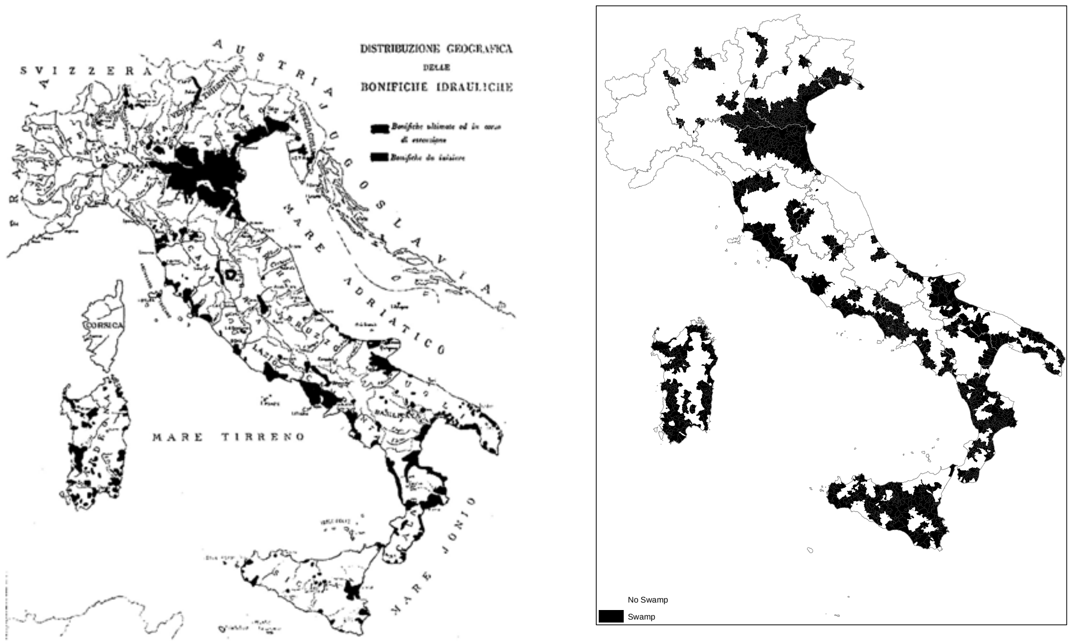
The picture portrays our two main predicted indexes of malaria risk. Predicted Malaria Risk based on Mordecai et al. (2013) on the left, and Malaria Stability Index by Kiszewski et al. (2004) on the right.

A.4 Agnana Calabra and Carfizzi: a Case Study

As a way of more concretely illustrating the variables presented above, we focus on two municipalities located in the same region that appear very similar but for their histories of malaria prevalence. The municipalities of Agnana Calabra and Carfizzi are both located in the southern region of Calabria at a distance of less than 200 km from one another, have similar populations, and are found at comparable elevations.[5] And yet, while there was malaria in Carfizzi in 1899, the disease was not present in Agnana Calabra. Indeed, Carfizzi had a higher intrinsic risk of malaria transmission (the Predcted Malaria Risk index for the two locations are 2.3 and 1.2 respectively). Interestingly for our analysis, the two municipalities also differed greatly in terms of land inequality: equal to 0.78 in Carfizzi and 0.48 in Agnana. In Agnana there were no large estates, while in Carfizzi about 38% of land was occupied by estates greater than 100 hectares. On the contrary, in Agnana about 50% of the land belonged to small farms, which in Carfizzi occupied less than 30% of the terrain.

A.5 Other Data: Covariates

Figure A8: Swamps and Marshes



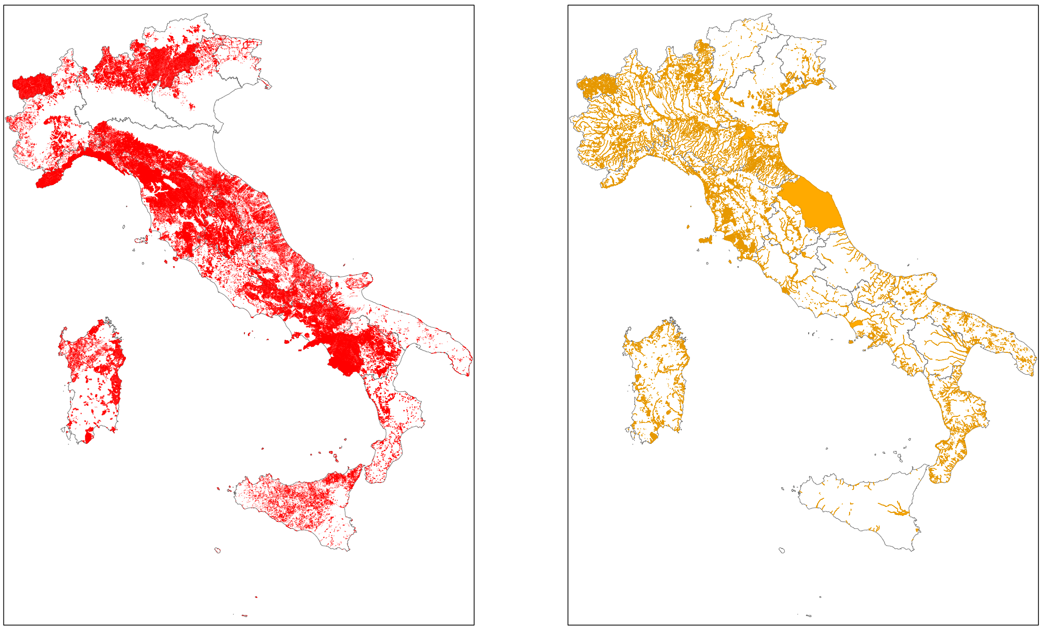
(a) Original Map

(b) 2011 Municipality Borders

The left panel portrays the original map of swamps in Italy at the begnninning of the 20th century. The right panel portrays the measure of swamps retrieved intersecting the original map with municiaplity in 2011. The units of observation are municipalities according to the borders in 2011. If a municipality contained at least a share of land with swamps according to the original map, we considered the municipality to have swamps.

[5]Agnana Calabra and Carfizzi have elevations of 300 and 358 respectively, and in 1921 had populations of 1,297 and 1,123 inhabitants respectively.

Figure A9: Landsslides and Hydreogeological Danger



(a) Landslides Danger

(b) Hydreogeological Danger

The picture portrays original maps: the left figure maps areas subject to landslide danger, the right figure maps areas subject to high and very high hydrogeological danger. Source: ISPRA.

A.6 Extended Data Sources and Summary Statistics

A.6.1 Data Sources

This section provides definition and sources for all variables in the paper: Table A1 focuses on malaria measures, Table A2 on land data, Table A3 describe all other outcomes of interest, Table A5 presents detail information on climatic variables, and Table A4 describe all other covariates in the analysis.

Table A1: Variables' Description and Data Sources: Malaria

Malaria (1922). Dummy variable equal to one if in 1922 there was malaria in the municipality and zero otherwise. *Source*: The presence of malaria was detected by health investigation published by the Ministry of the Interior (Directorate General of Health), in 1922. The volume is entitled "Malaria in Italy and the results of the anti-malarial fight".

Malaria (1882). Dummy variable equal to one if in 1882 there was malaria in the municipality and zero otherwise. *Source*: The presence of malaria was detected by health investigation solicited by Luigi Torelli, Minister of Agriculture, Industry and Commerce of the Kingdom of Italy.

Predicted Malaria Risk. Average malaria (falciparum) transmission index in the municipality built using the adjusted seasonal temperature from 1500 to 1929. *Source*: Mordecai, Paaijmans, Johnson, Balzer, Ben-Horin, de Moor, McNally, Pawar, Ryan, Smith, et al. (2013). See section A.3 for details.

Malaria Stability. Average Malaria Stability in the municipality. Source: Kiszewski et al. (2004).

Predicted Malaria Risk (M). Average malaria (falciparum) transmission index in the municipality built using the adjusted monthly temperature from 1901 to 1929. *Source*: Mordecai, Paaijmans, Johnson, Balzer, Ben-Horin, de Moor, McNally, Pawar, Ryan, Smith, et al. (2013). See section A.3 for details.

Predicted Malaria Risk (No Alt). Average malaria (falciparum) transmission index in the municipality built using the seasonal temperature from 1500 to 1929. *Source*: Mordecai, Paaijmans, Johnson, Balzer, Ben-Horin, de Moor, McNally, Pawar, Ryan, Smith, et al. (2013). See section A.3 for details.

Table A2: Variables' Description and Data Sources: Land Inequality

Land Inequality (1929). The Gini index of land inequality in the municipality. *Source*: The 1929 Census of Agriculture. The census provides data on the number of farms by size category. See Table A8 for information on the categories and their respective shares.

Land Inequality (1947). The Gini index of land inequality in the municipality. *Source*: The 1947 Census of Agriculture. The census provides data on the number of farms by size category. See Table A8 for information on the categories and their respective shares.

Small Landowners 1929. The share of land occupied by farms smaller than 10 hectares over the total share of land occupied by farms in the municipality. *Source*: The 1929 Census of Agriculture.

Medium Landowners 1929. The share of land occupied by farms larger than 10 hectares and smaller than 100 hectare over the total share of land occupied by farms in the municipality. *Source*: The 1929 Census of Agriculture.

Landmarks 1929. The share of land occupied by farms larger than 100 hectares over the total share of land occupied by farms in the municipality. *Source*: The 1929 Census of Agriculture.

Small Landowners 1947. The share of land occupied by farms smaller than 10 hectares over the total share of land occupied by farms in the municipality. *Source*: The 1947 Census of Agriculture.

Landmarks 1947. The share of land occupied by farms larger than 100 hectares over the total share of land occupied by farms in the municipality. *Source*: The 1947 Census of Agriculture.

Table A3: Variables' Description and Data Sources: Alternative Outcomes

Municipality Surface (Km^2). The municipality area in Km^2. *Source*: Italian Institute of Statistics - shape file of 2011.

Population Density 1929. The total population over the total area of the municipality in square kilometres. *Source*: The 1929 Census of Agriculture.

Share of Farmer. The share of population who works in the agriculture sector over the total population in the municipality. *Source*: The 1929 Census of Agriculture.

Scattered Houses Population. The share of population that is scattered in the countryside and not located within the town over the total population in the municipality. *Source*: The 1929 Census of Agriculture.

Farmer per Farm. The share of population who works in the agriculture sector over the total number of farms in the municipality. *Source*: The 1929 Census of Agriculture.

Land Owner. The share of farmers owning their land. *Source*: The 1929 Census of Agriculture.

Share Croppers (Tenants). The share of farmers with share-cropping agreements. *Source*: The 1929 Census of Agriculture.

Daily Workes. The share of farmers employed as daily laborers. *Source*: The 1929 Census of Agriculture.

Total Cultivated Land. The total share of land occupied by farms over the total area of the municipality. *Source*: The 1929 Census of Agriculture.

Unproductive Land. The share of land occupied by unproductive land over the total land occupied by farms in the municipality. *Source*: The 1929 Census of Agriculture.

Land Devoted to Pastures. The share of land occupied by pastures over the total land occupied by farms in the municipality. *Source*: The 1929 Census of Agriculture.

Land Devoted to Vineyards and Fruits. The share of land occupied by viticulture crops and fruits over the total land occupied by farms in the municipality. *Source*: The 1929 Census of Agriculture.

University Rate. The share of population having at least an university degree over the school population. *Source*: Istat.

Average Income. The total income at the municipality level divided by the number of tax payers in the municipality. *Source*: Finance Department. Ministry of Economy and Finance.

Table A4: Variables' Description and Data Sources: Covariates

Latitude. The latitude of the centroid of the municipality.

Longitude. The longitude of the centroid of the municipality.

Seaside City. Dummy variable equal to one if the municipality is on the coast and zero otherwise.

River. Dummy variable equal to one if the municipality is crossed by a river and zero otherwise.

Sea Distance. The distance between the municipality and the sea, in Km.

Lake. Dummy variable equal to one if the municipality is crossed by a lake and zero otherwise.

Average Altitude. Average altitude of the municipal area.

Elevation Std Deviation. Standard deviation altitude of the municipal area.

Ruggedness. Average 1x1 degree cell ruggedness - Terrain Ruggedness Index, 100 m. *Source*: Terrain Ruggedness Index originally devised by Riley, DeGloria, and Elliot (1999), obtained through http://diegopuga.org.

Swamp. Indicator variable taking on value one if a share of the land of the municipality had swamps in 1930. *Source*: Serpieri, Arrigo. "*La bonifica integrale.*" Annali di Economia (1937): 125-141.

Landslides Danger. Dummy variable equal to one if the municipality is categorized as an area subject to landslide danger by ISPRA and zero otherwise. *Source*: ISPRA - Istituto Superiore per la Protezione e la Ricerca Ambientale (*Research Institute for Environmental Protection*).

Hydrogeological Danger. Dummy variable equal to one if the municipality is categorized as an area subject to high and very high hydrogeological danger by ISPRA and zero otherwise. *Source*: ISPRA - Istituto Superiore per la Protezione e la Ricerca Ambientale (*Research Institute for Environmental Protection*).

Sand Soil. Sand content mass fraction in percentage % 15 cm dept, average in the municipality. Source: International Soil Reference and Information Centre (ISRIC) database.

Galor & Ozak (AER 2016). The log of the Caloric Suitability Index built by Galor and Özak (2015) and Galor and Özak (2016). The index measures the average potential agricultural output (measured in calories) across productive crops in each cell 5'×5' for the World.

Soil Suitability. Average land suitability in the municipality. Source: Ramankutty (2002).

Wheat, Rice and Tobacco Suitability. Estimated suitability index (value) for cultivating wheat, rice and tobacco with Low input in a rainfed agriculture. Source: FAO/IIASA, 2011. Global Agro-ecological Zones (GAEZv3.0). FAO Rome, Italy and IIASA, Laxenburg, Austria. *http://gaez.fao.org/Main.html.*

Source 1: Geo-morphological controls are available from the Italian Institute of Statistics. See https://www.istat.it/it/archivio/156224.

Table A5: Variables' Description and Data Sources: Temperature and Rain

Temperature Summer (c°). The average temperature, measured in centigrade, of the summer season over 429 years from 1500 to 1929. *Source:* The seasonal temperature come from Xoplaki, Luterbacher, Paeth, Dietrich, Steiner, Grosjean, and Wanner (2005).

Temperature Autumn (c°). The average temperature, measured in centigrade, of the autumn season over 429 years from 1500 to 1929. *Source:* The seasonal temperature come from Xoplaki, Luterbacher, Paeth, Dietrich, Steiner, Grosjean, and Wanner (2005).

Temperature Winter (c°). The average temperature, measured in centigrade, of the winter season over 429 years from 1500 to 1929. *Source:* The seasonal temperature come from Xoplaki, Luterbacher, Paeth, Dietrich, Steiner, Grosjean, and Wanner (2005).

Temperature Spring (c°). The average temperature, measured in centigrade, of the spring season over 429 years from 1500 to 1929. *Source:* The seasonal temperature come from Xoplaki, Luterbacher, Paeth, Dietrich, Steiner, Grosjean, and Wanner (2005).

Monthly Temperature (c°). The average monthly temperature, measured in centigrade, over 29 years from 1900 to 1929. *Source*: CRU CL 2.0 data from Hay et al. (2002).

Rain Summer (mm). The summer precipitation, measured in millilitres of rain, over 429 years from 1500 to 1929. *Source:* The seasonal rain come from Pauling, Luterbacher, Casty, and Wanner (2006).

Rain Autumn (mm). The autumn precipitation, measured in millilitres of rain, over 429 years from 1500 to 1929. *Source:* The seasonal rain come from Pauling, Luterbacher, Casty, and Wanner (2006).

Rain Winter (mm). The winter precipitation, measured in millilitres of rain, over 429 years from 1500 to 1929. *Source:* The seasonal rain come from Pauling, Luterbacher, Casty, and Wanner (2006).

Rain Spring (mm). The spring precipitation, measured in millilitres of rain, over 429 years from 1500 to 1929. *Source:* The seasonal rain come from Pauling, Luterbacher, Casty, and Wanner (2006).

Monthly Rain. The monthly precipitation, measured in millilitres of rain, over 29 years from 1900 to 1929. *Source*: CRU CL 2.0 data from Hay et al. (2002).

A.6.2 Summary Statistics

Table A6: Summary statistics: Malaria

	N	Mean	Std. Dev.	Min.	Max.
Malaria (1922)	7830	.37	.48	0	1
Malaria (1882)	7830	.41	.49	0	1
Predicted Malaria Risk	7830	6	8	0	33
Malaria Stability	7173	.058	.043	0	.2
Predicted Malaria Risk (M)	7830	4.2	2.9	0	11
Predicted Malaria Risk (No Alt)	7830	13	10	0	33

Table A7: Summary statistics: Monthly Malaria

	N	Mean	Std. Dev.	Min.	Max.
Monthly Predicted Malaria Risk:					
Predicted Malaria Risk (January)	7830	0	0	0	0
Predicted Malaria Risk (February)	7830	0	0	0	0
Predicted Malaria Risk (March)	7830	0	0	0	0
Predicted Malaria Risk (April)	7830	9.7e-07	.000015	0	.00028
Predicted Malaria Risk (May)	7830	.52	.78	0	4
Predicted Malaria Risk (June)	7830	8	6.7	0	24
Predicted Malaria Risk (July)	7830	18	11	0	32
Predicted Malaria Risk (August)	7830	18	11	0	32
Predicted Malaria Risk (September)	7830	5.6	7	0	30
Predicted Malaria Risk (Octuber)	7830	.27	.98	0	11
Predicted Malaria Risk (November)	7830	.0038	.02	0	.29
Predicted Malaria Risk (December)	7830	0	0	0	0

Table A8: Summary statistics: Land Inequality

	N	Mean	Std. Dev.	Min.	Max.
Land Inequality (1929)	7830.00	0.60	0.14	0.14	0.95
Land Inequality (1947)	7163.00	0.75	0.11	0.08	0.99
Small Landowners 1929	7830.00	0.50	0.27	0.00	1.00
Medium Landowners 1929	7830.00	0.33	0.23	0.00	1.00
Large Landowners 1929	7830.00	0.17	0.23	0.00	1.00
Small Landowners 1947	7163.00	0.46	0.25	0.00	1.00
Medium Landowners 1947	7163.00	0.28	0.18	0.00	0.95
Large Landowners 1947	7163.00	0.26	0.25	0.00	0.99
Difference in the Number of Small Landowners	7163.00	0.05	0.07	-0.38	0.98
Difference in the Number of Medium Landowners	7163.00	-0.05	0.07	-0.98	0.26
Difference in the Number Landmarks	7163.00	-0.00	0.02	-0.44	0.30
Total Number of Farms	7830.00	502.03	577.65	6.00	9335.83
Complete Distribution of 1929:					
>0.50	7830.00	0.20	0.36	0.00	8.42
0.50-1	7830.00	0.35	0.47	0.00	5.70
1-3	7830.00	19.18	17.66	0.00	96.90
3-5	7830.00	13.01	8.85	0.00	68.45
5-10	7830.00	17.75	11.43	0.00	73.88
10-20	7830.00	14.18	11.54	0.00	72.73
20-50	7830.00	12.21	12.04	0.00	93.28
50-100	7830.00	6.22	9.70	0.00	96.76
100-500	7830.00	11.65	18.73	0.00	99.83
<500	7830.00	5.25	13.15	0.00	97.26
Complete Distribution of 1947:					
>0.50	7163.00	5.33	5.59	0.01	56.13
0.50-2	7163.00	15.66	11.36	0.00	63.47
2-5	7163.00	14.65	8.66	0.00	48.83
5-10	7163.00	10.32	6.34	0.00	49.36
10-25	7163.00	12.12	8.29	0.00	48.64
25-50	7163.00	8.25	7.32	0.00	70.63
50-100	7163.00	7.51	8.30	0.00	74.28
100-200	7163.00	6.65	8.70	0.00	83.64
200-500	7163.00	7.41	11.44	0.00	97.55
<500	7163.00	12.11	21.64	0.00	97.57

Table A9: SUMMARY STATISTICS: ALTERNATIVE OUTCOMES

	N	Mean	Std. Dev.	Min.	Max.
Municipality Surface (Km^2)	7830.00	3778.30	5058.12	96.00	130771.00
Population Density	7830.00	168.98	294.67	2.02	11347.55
Share of Farmers	7830.00	0.51	0.26	0.00	4.30
Scattered Houses Population	7748.00	0.29	0.29	0.00	7.90
Farmers per Farm	7830.00	3.63	2.49	0.00	59.75
Land Owner	4593.00	0.51	0.27	0.00	1.00
Share Croppers (Tenants)	4593.00	0.12	0.16	0.00	0.90
Daily Worker	4593.00	0.18	0.17	0.00	0.88
Total Cultivated Land	7748.00	0.94	0.16	0.00	1.00
Share Unproductive Land	7748.00	0.08	0.09	0.00	0.95
Land Devoted to Pastures	7748.00	0.01	0.04	0.00	0.71
Land Devoted to Vineyards and Fruits	7748.00	0.24	0.22	0.00	0.97
University Rate	7825.00	0.07	0.03	0.00	0.29
Average Income	7824.00	20427.16	3052.53	11998.20	53589.29

Table A10: SUMMARY STATISTICS: COVARIATES AND SUITABILITY INDEXES

	N	Mean	Std. Dev.	Min.	Max.
Latitude	7830.00	43.45	2.55	36.67	47.05
Longitude	7830.00	11.46	2.77	6.73	18.46
Seaside City	7830.00	0.07	0.25	0.00	1.00
River	7830.00	0.46	0.50	0.00	1.00
Sea Distance	7830.00	71.29	55.38	0.00	230.34
Lake	7830.00	0.37	0.48	0.00	1.00
Share of Mountains	7830.00	47.77	48.33	0.00	100.00
Average Altitude	7830.00	553.44	499.24	0.00	3072.50
Elevation Std Deviation	7830.00	147.36	153.03	1.12	906.97
Ruggedness	7830.00	224.29	216.53	0.89	1151.45
Swamp	7830.00	0.16	0.33	0.00	1.00
Landslides Danger	7830.00	0.24	0.43	0.00	1.00
Hydrogeological Danger	7830.00	0.16	0.36	0.00	1.00
Sand Soil	7830.00	43.73	18.96	11.00	110.13
Galor & Ozak (AER 2016) (log)	7830.00	7.58	0.92	0.00	8.00
Soil Suitability	7830.00	0.78	0.24	0.00	1.00
Wheat Suitability	7830.00	2290.39	471.10	0.00	3031.06
Tobacco Suitability	7830.00	1742.05	1104.15	0.00	5112.70
Rice Suitability	7830.00	470.72	1109.54	0.00	6255.41
Temperature Summer (c°)	7830.00	16.09	5.77	-8.45	24.08
Temperature Autumn (c°)	7830.00	8.72	5.87	-14.44	19.59
Temperature Winter (c°)	7830.00	-0.30	5.88	-22.89	11.54
Temperature Spring (c°)	7830.00	6.41	5.64	-17.48	14.92
Rain Summer (mm)	7830.00	250.95	182.35	12.92	810.19
Rain Autumn (mm)	7830.00	319.74	82.99	164.12	614.16
Rain Winter (mm)	7830.00	245.33	73.17	150.56	530.31
Rain Spring (mm)	7830.00	260.88	109.14	97.66	579.76

Figure A10: Distribution of Land Inequality in 1929

(a) Land Inequality 1929

The histogram portrays the distribution of land inequality in 1929 across the municipalities in the sample.

B Additional Results and Extended Specifications

B.1 Malaria and Land Inequality: Robustness

B.1.1 Conley Standard Errors - Alternative Threshold

Table B1 replicates baseline results with Conley standard errors computed with a 80 km cutoff threshold.

B.1.2 First Stage Regressions

Table B2: First Stage - Predicted Malaria Risk and Historical Malaria Prevalence

	Malaria Prevalence (1922)				Malaria Prevalence (1882)			
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS	(8) OLS
Predicted Malaria Risk	0.252***	0.136***			0.242***	0.157***		
	(0.024)	(0.031)			(0.020)	(0.025)		
	[0.030]	[0.041]			[0.023]	[0.028]		
Malaria Stability			0.180***	0.138***			0.175***	0.133***
			(0.021)	(0.031)			(0.020)	(0.027)
			[0.030]	[0.036]			[0.026]	[0.031]
Mean Dependent	0.37	0.37	0.37	0.37	0.41	0.41	0.41	0.41
F-statistic Instrument	112	19	72	20	150	38	75	25
Observations	7,830	7,830	7,173	7,173	7,830	7,830	7,173	7,173
R-Square	0.27	0.52	0.14	0.52	0.24	0.40	0.12	0.40
Geographical Characteristics	×	√	×	√	×	√	×	√
Land Suitability Indexes	×	√	×	√	×	√	×	√
Temperature (c°) & Rain (mm)	×	√	×	√	×	√	×	√
Region FE	×	√	×	√	×	√	×	√

OLS estimates. The unit of observation is the municipality. The dependent variable is malaria prevalence in the municipality in 1922 and 1882. *Malaria 1922, 1882* is a dummy variable equal to one in municipality with malaria transmission, and zero otherwise. *Predicted Malaria Risk* measures the predicted risk of malaria transmission constructed following Mordecai et al. (2013). *Malaria Stability* measures the stability of malaria transmission by Kiszewski et al. (2004). Complete data descriptions, data sources and summary statistics are presented in Appendix Section A.6.1 and A.6.2. Robust standard errors clustered at the provincial level are in round parentheses and Conley standard errors (50 km cutoff) are reported in square brackets. Significance at the 10% level is represented by , at the 5% level by *, and at the 1% level by ***.

B.1.3 Alternative Indexes of Predicted Malaria Risk

Results are confirmed with all indexes of predicted malaria risk. Appendix Table B3 reports baseline results using all the alternative indexes of predicted malaria prevalence that we constructed. First, we replicate the Predicted Malaria Risk by Mordecai et al. (2013) using monthly temperatures (1900-1929, CRU CL 2.0 data from Hay et al., 2002) instead of seasonal temperatures from 1500. Second, we compute the index without adjusting temperature based on the climatological relation between altitude and temperature. Third, we borrow from another influential measure of predicted malaria risk, the Malaria Stability Index by Kiszewski et al. (2004).

Table B3: Alternative Malaria Indexes

Panel A	Small Landowners (1929)			
	(1) ITT	(2) ITT	(3) ITT	(4) ITT
Predicted Malaria Risk	-0.053** (0.020) [0.024]			
Predicted Malaria Risk (M)		-0.073*** (0.015) [0.024]		
Predicted Malaria Risk (No Alt)			-0.158*** (0.047) [0.056]	
Malaria Stability				-0.099*** (0.017) [0.024]
Mean Dependent	0.50	0.50	0.50	0.50
Observations	7,830	7,830	7,830	7,173
R-Square	0.30	0.30	0.31	0.30
Panel B	Landmarks (1929)			
Predicted Malaria Risk	0.043** (0.018) [0.019]			
Predicted Malaria Risk (M)		0.028** (0.011) [0.012]		
Predicted Malaria Risk (No Alt)			0.156*** (0.040) [0.040]	
Malaria Stability				0.057*** (0.015) [0.018]
Mean Dependent	0.17	0.17	0.17	0.17
Observations	7,830	7,830	7,830	7,173
R-Square	0.31	0.31	0.32	0.32
Panel C	Land Inequality (1929)			
Predicted Malaria Risk	0.030** (0.011) [0.015]			
Predicted Malaria Risk (M)		0.033*** (0.008) [0.011]		
Predicted Malaria Risk (No Alt)			0.088*** (0.026) [0.030]	
Malaria Stability				0.051*** (0.008) [0.011]
Mean Dependent	0.60	0.60	0.60	0.60
Observations	7,830	7,830	7,830	7,173
R-Square	0.27	0.27	0.28	0.28
Geographical Characteristics	✓	✓	✓	✓
Land Suitability Indexes	✓	✓	✓	✓
Temperature (c°) & Rain (mm)	✓	✓	✓	✓
Region FE	✓	✓	✓	✓

OLS estimates in Columns (1-4) and two stage least squares in Columns (5-6). The unit of observation is the municipality. The dependent variables are: the share of land occupied by farms smaller than 10 hectares over the total share of land occupied by farms in the municipality in Panel (A); the share of land occupied by farms larger than 100 hectares over the total share of land occupied by farms in the municipality in Panel (B); the Gini index of 1929 land inequality in the municipality in Panel (C). See Table A2 for details on the dependent variables and Table A8 for summary statistics. *Predicted Malaria Risk* is the average malaria (falciparum) transmission index in the municipality built using the adjusted seasonal temperature from 1500 to 1929. *Predicted Malaria Risk (M)* is the average malaria (falciparum) transmission index in the municipality built using the adjusted monthly temperature from 1901 to 1929. *Predicted Malaria Risk (No Alt)* is the average malaria (falciparum) transmission index in the municipality built using the seasonal temperature from 1500 to 1929. *Malaria Stability* is the Average Malaria Stability in the municipality. Source: Kiszewski et al. (2004). Covariate descriptions and data sources are reported in Tables A4 and A5, and summary statistics are reported in Table A10. Robust standard errors clustered at the provincial level are in round parentheses and Conley standard errors (50 km cutoff) are reported in square brackets. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

B.1.4 Historical Malaria Prevalence as per 1882

Table B4: Land Inequality and Malaria Prevalence in 1882

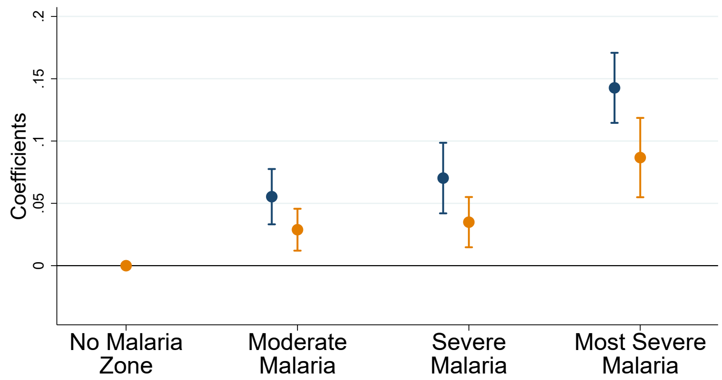
Panel A	Small Landowners (1929)					
	(1) OLS	(2) OLS	(3) ITT	(4) ITT	(5) IV	(6) IV
Malaria (1882)	-0.102***	-0.075***			-0.216***	-0.336***
	(0.023)	(0.019)			(0.065)	(0.129)
	[0.036]	[0.025]			[0.088]	[0.148]
Predicted Malaria Risk			-0.052***	-0.053**		
			(0.018)	(0.020)		
			[0.024]	[0.024]		
Mean Dependent	0.50	0.50	0.50	0.50	0.50	0.50
F-statistic Instrument					150	38
Observations	7,830	7,830	7,830	7,830	7,830	7,830
R-Square	0.03	0.30	0.04	0.30	-0.01	0.17
Panel B	**Landmarks (1929)**					
Malaria (1882)	0.028	0.034***			0.007	0.273**
	(0.020)	(0.012)			(0.048)	(0.116)
	[0.025]	[0.018]			[0.056]	[0.111]
Predicted Malaria Risk			0.002	0.043**		
			(0.012)	(0.018)		
			[0.013]	[0.019]		
Mean Dependent	0.17	0.17	0.17	0.17	0.17	0.17
F-statistic Instrument					150	38
Observations	7,830	7,830	7,830	7,830	7,830	7,830
R-Square	0.00	0.31	0.00	0.31	0.00	0.15
Panel C	**Land Inequality (1929)**					
Malaria (1882)	0.063***	0.032***			0.134***	0.189***
	(0.012)	(0.010)			(0.029)	(0.072)
	[0.017]	[0.013]			[0.040]	[0.086]
Predicted Malaria Risk			0.032***	0.030**		
			(0.008)	(0.011)		
			[0.011]	[0.015]		
Mean Dependent	0.60	0.60	0.60	0.60	0.60	0.60
F-statistic Instrument					150	38
Observations	7,830	7,830	7,830	7,830	7,830	7,830
R-Square	0.05	0.27	0.05	0.27	-0.01	0.09
Geographical Characteristics	×	√	×	√	×	√
Land Suitability Indexes	×	√	×	√	×	√
Temperature (c°) & Rain (mm)	×	√	×	√	×	√
Region FE	×	√	×	√	×	√

OLS estimates in Columns (1-4) and two stage least squares in Columns (5-6). The unit of observation is the municipality. The dependent variables are: the share of land occupied by farms smaller than 10 hectares over the total share of land occupied by farms in the municipality in Panel (A); the share of land occupied by farms larger than 100 hectares over the total share of land occupied by farms in the municipality in Panel (B); the Gini index of 1929 land inequality in the municipality in Panel (C). See Table A2 for details on the dependent variables and Table A8 for summary statistics. *Malaria 1882* is a dummy variable equal to one if in 1882 there was malaria in the municipality and zero otherwise. *Predicted Malaria Risk* measures the predicted risk of malaria transmission constructed following Mordecai et al. (2013). Covariate descriptions and data sources are reported in Tables A4 and A5, and summary statistics are reported in Table A10. Robust standard errors clustered at the provincial level are in round parentheses and Conley standard errors (50 km cutoff) are reported in square brackets. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

B.1.5 Malaria Severity and Seasonality

The harsher the burden of malaria, the more concentrated the land. Figure B1 documents that land inequality is monotonically increasing with the level of severity of malaria in the municipality. In the same vein, the longer the malaria season, the more concentrated the land. In the left panel

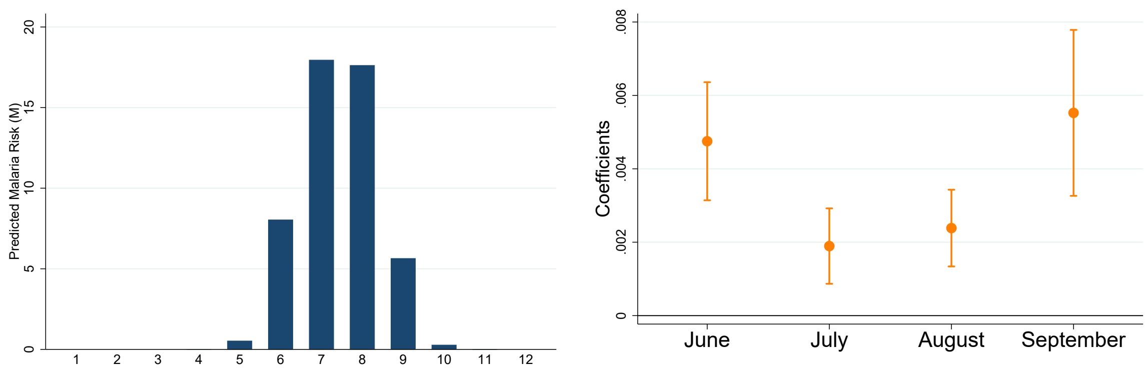
Figure B1: Land Inequality and Historical Malaria Severity



The graph plots regression estimates using the categorical dimensions of historical malaria prevalence as measured by the investigation from the Torelli report from 1882. Orange coefficients correspond to specification not including geographical controls, blue coefficients correspond to specification including geographical controls.

of Appendix Figure B2, we document that malaria risk in Italy was especially high in June, July, August, and September. Looking separately at the effect of malaria risk in each month, we find that it is stronger for municipalities that have a longer malaria season, i.e., that experienced the risk of malaria as early even as June and as late as September.

Figure B2: Land Inequality (1929) and Monthly Predicted Malaria Prevalence



(a) Predicted Malaria Risk (b) Malaria & Inequality

B.1.6 Land Inequality in 1947

Table B5 replicates baseline results with inequality measured in 1947. The effect on inequality of a history of malaria exposure after the II World War are still present, suggesting a persistence of the effect over time.

Table B5: LAND INEQUALITY IN 1947 AND HISTORICAL MALARIA PREVALENCE

Panel A	SMALL LANDOWNERS (1947)					
	(1) OLS	(2) OLS	(3) ITT	(4) ITT	(5) IV	(6) IV
MALARIA (1922)	-0.116***	-0.106***			-0.120**	-0.340***
	(0.031)	(0.025)			(0.057)	(0.090)
	[0.044]	[0.036]			[0.076]	[0.111]
PREDICTED MALARIA RISK			-0.030*	-0.047***		
			(0.016)	(0.014)		
			[0.021]	[0.017]		
Mean Dependent	0.46	0.46	0.46	0.46	0.46	0.46
F-statistic Instrument					112	19
Observations	7,163	7,163	7,163	7,163	7,163	7,163
R-Square	0.05	0.47	0.01	0.46	0.05	0.37
Panel B	LANDMARKS (1947)					
MALARIA (1922)	0.068***	0.094***			-0.049	0.444***
	(0.025)	(0.020)			(0.052)	(0.098)
	[0.035]	[0.027]			[0.072]	[0.117]
PREDICTED MALARIA RISK			-0.012	0.061***		
			(0.012)	(0.014)		
			[0.017]	[0.016]		
Mean Dependent	0.26	0.26	0.26	0.26	0.26	0.26
F-statistic Instrument					112	19
Observations	7,163	7,163	7,163	7,163	7,163	7,163
R-Square	0.02	0.43	0.00	0.42	-0.03	0.19
Panel C	LAND INEQUALITY (1947)					
MALARIA (1922)	0.047***	0.041***			0.030	0.196***
	(0.012)	(0.011)			(0.021)	(0.046)
	[0.016]	[0.014]			[0.028]	[0.059]
PREDICTED MALARIA RISK			0.007	0.027***		
			(0.006)	(0.006)		
			[0.007]	[0.007]		
Mean Dependent	0.75	0.75	0.75	0.75	0.75	0.75
F-statistic Instrument					112	19
Observations	7,163	7,163	7,163	7,163	7,163	7,163
R-Square	0.04	0.39	0.00	0.39	0.04	0.14
Geographical Characteristics	×	✓	×	✓	×	✓
Land Suitability Indexes	×	✓	×	✓	×	✓
Temperature (c°) & Rain (mm)	×	✓	×	✓	×	✓
Region FE	×	✓	×	✓	×	✓

OLS estimates in Columns (1-4) and two stage least squares in Columns (5-6). The unit of observation is the municipality. The dependent variables are: the share of land occupied by farms smaller than 10 hectares over the total share of land occupied by farms in the municipality in Panel (A); the share of land occupied by farms larger than 100 hectares over the total share of land occupied by farms in the municipality in Panel (B); the Gini index of 1929 land inequality in the municipality in Panel (C). See Table A2 for details on the dependent variables and Table A8 for summary statistics. *Malaria 1882* is a dummy variable equal to one if in 1882 there was malaria in the municipality and zero otherwise. *Predicted Malaria Risk* measures the predicted risk of malaria transmission constructed following Mordecai et al. (2013). Covariate descriptions and data sources are reported in Tables A4 and A5, and summary statistics are reported in Table A10. Robust standard errors clustered at the provincial level are in round parentheses and Conley standard errors (50 km cutoff) are reported in square brackets. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table B1: Land Inequality and Malaria Prevalence

Panel A	Small Landowners (1929)					
	(1) OLS	(2) OLS	(3) ITT	(4) ITT	(5) IV	(6) IV
Malaria (1882)	-0.102***	-0.075***			-0.216***	-0.336***
	(0.023)	(0.019)			(0.065)	(0.129)
	[0.042]	[0.029]			[0.095]	[0.152]
Predicted Malaria Risk			-0.052***	-0.053**		
			(0.018)	(0.020)		
			[0.027]	[0.029]		
Mean Dependent	0.50	0.50	0.50	0.50	0.50	0.50
F-statistic Instrument					150	38
Observations	7,830	7,830	7,830	7,830	7,830	7,830
R-Square	0.03	0.30	0.04	0.30	-0.01	0.17
Panel B	Landmarks (1929)					
Malaria (1882)	0.028	0.034***			0.007	0.273**
	(0.020)	(0.012)			(0.048)	(0.116)
	[0.029]	[0.017]			[0.057]	[0.116]
Predicted Malaria Risk			0.002	0.043**		
			(0.012)	(0.018)		
			[0.014]	[0.023]		
Mean Dependent	0.17	0.17	0.17	0.17	0.17	0.17
F-statistic Instrument					150	38
Observations	7,830	7,830	7,830	7,830	7,830	7,830
R-Square	0.00	0.31	0.00	0.31	0.00	0.15
Panel C	Land Inequality (1929)					
Malaria (1882)	0.063***	0.032***			0.134***	0.189***
	(0.012)	(0.010)			(0.029)	(0.072)
	[0.021]	[0.015]			[0.041]	[0.080]
Predicted Malaria Risk			0.032***	0.030**		
			(0.008)	(0.011)		
			[0.012]	[0.016]		
Mean Dependent	0.60	0.60	0.60	0.60	0.60	0.60
F-statistic Instrument					150	38
Observations	7,830	7,830	7,830	7,830	7,830	7,830
R-Square	0.05	0.27	0.05	0.27	-0.01	0.09
Geographical Characteristics	×	√	×	√	×	√
Land Suitability Indexes	×	√	×	√	×	√
Temperature (c°) & Rain (mm)	×	√	×	√	×	√
Region FE	×	√	×	√	×	√

OLS and IV estimates. The unit of observation is the municipality. the *dependent variable* is the share of small farms in the top panel, the share of large farms in the middle panel and the Gini index of land inequality in the municipality in the bottom panel, based on the 1929 agricultural census. *Malaria 1922* is a dummy variable equal to one in municipality with malaria transmission, and zero otherwise. *Predicted Malaria Risk* measures the predicted risk of malaria transmission constructed following Mordecai et al. (2013). Complete data descriptions, data sources and summary statistics are presented in Appendix Section A.6.1 and A.6.2. Robust standard errors clustered at the provincial level are in round parentheses and Conley standard errors (80 km cutoff) are reported in square brackets. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

B.1.7 Malaria and Farms of Different Size

Baseline results do not rest on any specific definition of Small Farms and Landmarks. To prove so, Table B6 replicates baseline findings with 10 different threshold of farm size. Baseline patterns are confirmed.

Table B6: Malaria and Sizes of Land Estates

Panel A	> 0.50 ha	$0.50-1$ ha	$1-3$ ha	$3-5$ ha	$5-10$ ha	$10-20$ ha	$20-50$ ha	$50-100$ ha	$100-500$ ha	< 500 ha
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Predicted Malaria Risk	-0.003	0.038	-0.033	-1.795***	-3.465***	-2.007***	1.129*	1.858***	2.349	1.929***
	(0.021)	(0.032)	(1.263)	(0.505)	(0.631)	(0.623)	(0.679)	(0.546)	(1.548)	(0.718)
	[0.022]	[0.033]	[1.404]	[0.608]	[0.817]	[0.774]	[0.690]	[0.750]	[1.690]	[0.755]
Mean Dependent	0.20	0.35	19.18	13.01	17.75	14.18	12.21	6.22	11.65	5.25
Observations	7,830	7,830	7,830	7,830	7,830	7,830	7,830	7,830	7,830	7,830
R-Square	0.12	0.18	0.28	0.26	0.27	0.37	0.24	0.21	0.18	0.30
Geographical Characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Land Suitability Indexes	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Temperature (c°) & Rain (mm)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Temperature^2 (c°) & Rain^2 (mm)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Region FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

OLS estimates. The unit of observation is the municipality. The *dependent variable* are: > 0.50 *ha* is the share of land occupied by farms smaller than 0.5 hectares over the total share of land occupied by farms in the municipality. $0.50 - 1$ *ha* is the share of land occupied by farms larger than 0.5 hectares and smaller than 1 hectare over the total share of land occupied by farms in the municipality. $1 - 3$ *ha* is the share of land occupied by farms larger than 1 hectare and smaller than 3 hectares over the total share of land occupied by farms in the municipality. $3 - 5$ *ha* is the share of land occupied by farms larger than 3 hectares and smaller than 5 hectares over the total share of land occupied by farms in the municipality. $5 - 10$ *ha* is the share of land occupied by farms larger than 5 hectares and smaller than 10 hectares over the total share of land occupied by farms in the municipality. $10 - 20$ *ha* is the share of land occupied by farms larger than 10 hectares and smaller than 20 hectares over the total share of land occupied by farms in the municipality. $20 - 50$ *ha* is the share of land occupied by farms larger than 20 hectares and smaller than 50 hectares over the total share of land occupied by farms in the municipality. $50 - 100$ *ha* is the share of land occupied by farms larger than 50 hectares and smaller than 100 hectares over the total share of land occupied by farms in the municipality. $100 - 500$ *ha* is the share of land occupied by farms larger than 100 hectares and smaller than 500 hectares over the total share of land occupied by farms in the municipality. < 500 *ha* is the share of land occupied by farms larger than 500 hectares over the total share of land occupied by farms in the municipality. See Table A2 for details on the dependent variables and Table A8 for summary statistics. *Predicted Malaria Risk* measures the predicted risk of malaria transmission. *Malaria Stability* measures the average malaria suitability in the municipality, as from Kiszewski, Mellinger, Spielman, Malaney, Sachs, and Sachs (2004). Covariates description and data sources are reported in Tables A4 and A5, and summary statistics are reported in Table A10. Robust standard errors clustered at the provincial level are in round parentheses and Conley standard errors (50 km cutoff) are reported in square brackets. Significance at the 10% level is represented by , at the 5% level by *, and at the 1% level by ***.

B.1.8 Other Geographical Drivers of Land Inequality

Appendix Table B7 reports coefficient estimates for the other main predictors of land inequality, which are - as expected - sandy soils, altitude, and soil suitability. Of these, sandy soil is confirmed as a driver. In order to gain a better sense of the effect of malaria, it is thus useful to compare its standardardized coefficient with that of sand soil. While a standard deviation of predicted malaria risk is associated with an increase in land inequality of 0.30 standard deviation, a standard deviation in the share of sandy soil is associated with a 0.11 standard deviation increase in land inequality. The effect of land suitabilities disappers after the full inclusion of controls. In principle, it may thus be suitabilities for crops with economies of scale in production, and not general land suitability, that matters. We verify whether this is the case by looking at the only two crops cultivated in Italy that might present some economies of scale in cultivation, i.e., rice and tobacco. Results in Appendix Table B8 show that they are not related to land inequalities.

Table B7: Other Main Determinants of Land Inequality

Panel A	Land Inequality (1929)					
	(1) OLS	(2) OLS	(3) ITT	(4) ITT	(5) IV	(6) IV
Malaria (1922)	0.105***	0.080***			0.129***	0.218***
	(0.015)	(0.018)			(0.024)	(0.068)
	[0.022]	[0.026]			[0.033]	[0.081]
Altitude (Average)		0.084***		0.086***		0.071**
		(0.030)		(0.031)		(0.032)
		[0.029]		[0.031]		[0.031]
Wheat Suitability		-0.006		-0.008		0.008
		(0.009)		(0.010)		(0.010)
		[0.011]		[0.011]		[0.012]
Sand Soil		0.018***		0.011*		0.019***
		(0.006)		(0.006)		(0.007)
		[0.009]		[0.008]		[0.009]
Soil Suitability		-0.006		-0.002		0.006
		(0.016)		(0.015)		(0.016)
		[0.018]		[0.018]		[0.019]
Galor & Ozak (AER 2016)		0.003		0.004		0.002
		(0.004)		(0.004)		(0.004)
		[0.005]		[0.005]		[0.005]
Predicted Malaria Risk			0.032***	0.030**		
			(0.008)	(0.011)		
			[0.011]	[0.015]		
Mean Dependent	0.60	0.60	0.60	0.60	0.60	0.60
F-statistic Instrument					112	19
Observations	7,830	7,830	7,830	7,830	7,830	7,830
R-Square	0.12	0.30	0.05	0.27	0.12	0.19
Geographical Characteristics	×	√	×	√	×	√
Land Suitability Indexes	×	√	×	√	×	√
Temperature (c°) & Rain (mm)	×	√	×	√	×	√
Region FE	×	√	×	√	×	√

OLS estimates in Columns (1-4) and two stage least squares in Columns (5-6). The unit of observation is the municipality. In all specifications, the dependent variable is the Gini index of 1929 land inequality in the municipality. See Table A2 for details on the dependent variables and Table A8 for summary statistics. *Malaria 1922* is a dummy variable equal to one if in 1992 there was malaria in the municipality and zero otherwise. *Average Altitude* is the average altitude of the municipal area. *Soil Suitability* is the average land suitability in the municipality. Source: Ramankutty (2002). *Galor & Ozak (AER 2016)* is the log of the Caloric Suitability Index built by Galor and Özak (2015) and Galor and Özak (2016). *Predicted Malaria Risk* measures the predicted risk of malaria transmission constructed following Mordecai et al. (2013). Covariate descriptions and data sources are reported in Tables A4 and A5, and summary statistics are reported in Table A10. Robust standard errors clustered at the provincial level are in round parentheses and Conley standard errors (50 km cutoff) are reported in square brackets. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table B8: Crop Suitability and Land Inequality

Panel A	Land Inequality (1929)					
	(1) OLS	(2) OLS	(3) ITT	(4) ITT	(5) IV	(6) IV
Malaria (1922)	0.105***	0.074***			0.129***	0.212***
	(0.015)	(0.016)			(0.024)	(0.077)
	[0.022]	[0.022]			[0.033]	[0.087]
Altitude (Average)		0.075**		0.077**		0.068**
		(0.029)		(0.030)		(0.030)
		[0.029]		[0.031]		[0.031]
Wheat Suitability		-0.015		-0.018*		-0.001
		(0.009)		(0.010)		(0.010)
		[0.009]		[0.010]		[0.012]
Sand Soil		0.012**		0.006		0.014**
		(0.006)		(0.006)		(0.006)
		[0.008]		[0.008]		[0.008]
Soil Suitability		0.004		0.006		0.016
		(0.015)		(0.015)		(0.016)
		[0.015]		[0.015]		[0.019]
Galor & Ozak (AER 2016)		0.004		0.005		0.003
		(0.004)		(0.004)		(0.004)
		[0.005]		[0.005]		[0.006]
Tobacco Suitability		-0.015		-0.018		0.001
		(0.009)		(0.011)		(0.010)
		[0.014]		[0.017]		[0.009]
Rice Suitability		-0.027***		-0.026**		-0.029***
		(0.009)		(0.010)		(0.007)
		[0.011]		[0.014]		[0.008]
Predicted Malaria Risk			0.032***	0.023**		
			(0.008)	(0.010)		
			[0.011]	[0.012]		
Mean Dependent	0.60	0.60	0.60	0.60	0.60	0.60
F-statistic Instrument					112	15
Observations	7,830	7,830	7,830	7,830	7,830	7,830
R-Square	0.12	0.32	0.05	0.29	0.12	0.22
Geographical Characteristics	×	√	×	√	×	√
Land Suitability Indexes	×	√	×	√	×	√
Temperature (c°) & Rain (mm)	×	√	×	√	×	√
Region FE	×	√	×	√	×	√

OLS estimates in Columns (1-4) and two stage least squares in Columns (5-6). The unit of observation is the municipality. In all specifications, the dependent variable is the Gini index of 1929 land inequality in the municipality. See Table A2 for details on the dependent variables and Table A8 for summary statistics. *Malaria 1922* is a dummy variable equal to one if in 1992 there was malaria in the municipality and zero otherwise. *Wheat, Rice and Tobacco Suitability* are estimated suitability index (value) for cultivating wheat, rice and tobacco with low input in a rain-fed agriculture. *Predicted Malaria Risk* measures the predicted risk of malaria transmission constructed following Mordecai et al. (2013). Covariate descriptions and data sources are reported in Tables A4 and A5, and summary statistics are reported in Table A10. Robust standard errors clustered at the provincial level are in round parentheses and Conley standard errors (50 km cutoff) are reported in square brackets. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

B.1.9 Accounting for Province Fixed Effect

Italy was declared a nation-state in 1861, with 1870 marking the final unification of the country. Prior to the Risorgimento, Italy had experienced extremely heterogeneous histories of self-government, authoritarian ruling, backward administrations, and foreign domination. Given that municipalities located in the same region often experienced very different institutional histories, we include province fixed effects.

Italy, according to 2011 borders, has 110 provinces. While provinces' borders and their role in the Italian administration changed over time (they were just 75 in 1923 and 93 in 1945), 2011 province borders reflect more homogeneous areas than regions. The inclusion of province FE effects can in fact better account for different cultural and institutional histories. In Table B9 and B10, we report estimates accounting for province fixed effects. Importantly, our results are confirmed.

Table B9: Accounting for Province FE

Panel A	Small Landowners (1929)					
	(1) OLS	(2) OLS	(3) ITT	(4) ITT	(5) IV	(6) IV
Malaria (1922)	-0.184***	-0.108***			-0.208***	-0.252**
	(0.031)	(0.030)			(0.058)	(0.124)
	[0.046]	[0.037]			[0.077]	[0.147]
Predicted Malaria Risk			-0.052***	-0.030*		
			(0.018)	(0.017)		
			[0.024]	[0.021]		
Mean Dependent	0.50	0.50	0.50	0.50	0.50	0.50
F-statistic Instrument					112	18
Observations	7,830	7,830	7,830	7,830	7,830	7,830
R-Square	0.11	0.48	0.04	0.47	0.10	0.45
Panel B	Landmarks (1929)					
Malaria (1922)	0.086***	0.050**			0.006	0.241**
	(0.024)	(0.022)			(0.046)	(0.112)
	[0.028]	[0.023]			[0.053]	[0.112]
Predicted Malaria Risk			0.002	0.029*		
			(0.012)	(0.016)		
			[0.013]	[0.017]		
Mean Dependent	0.17	0.17	0.17	0.17	0.17	0.17
F-statistic Instrument					112	18
Observations	7,830	7,830	7,830	7,830	7,830	7,830
R-Square	0.03	0.44	0.00	0.43	0.00	0.38
Panel C	Land Inequality (1929)					
Malaria (1922)	0.105***	0.051***			0.129***	0.149**
	(0.015)	(0.016)			(0.024)	(0.068)
	[0.022]	[0.019]			[0.033]	[0.083]
Predicted Malaria Risk			0.032***	0.018*		
			(0.008)	(0.010)		
			[0.011]	[0.012]		
Mean Dependent	0.60	0.60	0.60	0.60	0.60	0.60
F-statistic Instrument					112	18
Observations	7,830	7,830	7,830	7,830	7,830	7,830
R-Square	0.12	0.41	0.05	0.40	0.12	0.37
Geographical Characteristics	×	√	×	√	×	√
Land Suitability Indexes	×	√	×	√	×	√
Temperature (c°) & Rain (mm)	×	√	×	√	×	√
Province FE	×	√	×	√	×	√

OLS estimates in Columns (1-4) and two stage least squares in Columns (5-6). The unit of observation is the municipality. The dependent variables are: the share of land occupied by farms smaller than 10 hectares over the total share of land occupied by farms in the municipality in Panel (A); the share of land occupied by farms larger than 100 hectares over the total share of land occupied by farms in the municipality in Panel (B); the Gini index of 1929 land inequality in the municipality in Panel (C). See Table A2 for details on the dependent variables and Table A8 for summary statistics. *Malaria 1922* is a dummy variable equal to one if in 1992 there was malaria in the municipality and zero otherwise. *Predicted Malaria Risk* measures the predicted risk of malaria transmission constructed following Mordecai et al. (2013). Covariate descriptions and data sources are reported in Tables A4 and A5, and summary statistics are reported in Table A10. Robust standard errors clustered at the provincial level are in round parentheses and Conley standard errors (50 km cutoff) are reported in square brackets. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

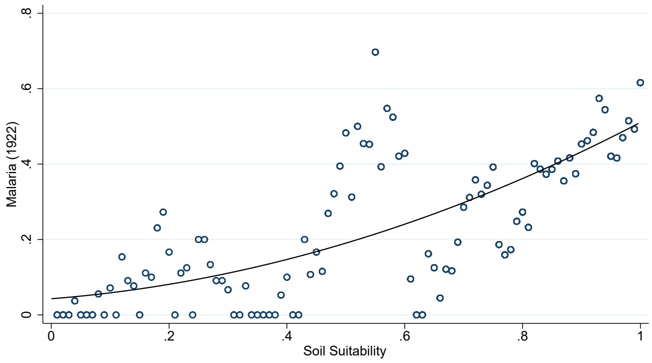
Table B10: Land Inequality and Predicted Malaria Risk Accounting for Province FE

Panel A	Small Landowners (1929)			
	(1) ITT	(2) ITT	(3) ITT	(4) ITT
Predicted Malaria Risk	-0.030*			
	(0.017)			
	[0.021]			
Predicted Malaria Risk (M)		-0.056***		
		(0.013)		
		[0.018]		
Predicted Malaria Risk (No Alt)			-0.089*	
			(0.051)	
			[0.052]	
Malaria Stability				-0.098***
				(0.018)
				[0.023]
Mean Dependent	0.50	0.50	0.50	0.50
Observations	7,830	7,830	7,830	7,173
R-Square	0.47	0.47	0.47	0.47
Panel B	Landmarks (1929)			
Predicted Malaria Risk	0.029*			
	(0.016)			
	[0.017]			
Predicted Malaria Risk (M)		0.023**		
		(0.009)		
		[0.011]		
Predicted Malaria Risk (No Alt)			0.100**	
			(0.039)	
			[0.038]	
Malaria Stability				0.051***
				(0.015)
				[0.018]
Mean Dependent	0.17	0.17	0.17	0.17
Observations	7,830	7,830	7,830	7,173
R-Square	0.43	0.43	0.44	0.45
Panel C	Land Inequality (1929)			
Predicted Malaria Risk	0.018*			
	(0.010)			
	[0.012]			
Predicted Malaria Risk (M)		0.029***		
		(0.007)		
		[0.009]		
Predicted Malaria Risk (No Alt)			0.071***	
			(0.026)	
			[0.029]	
Malaria Stability				0.054***
				(0.009)
				[0.012]
Mean Dependent	0.60	0.60	0.60	0.60
Observations	7,830	7,830	7,830	7,173
R-Square	0.40	0.40	0.41	0.41
Geographical Characteristics	√	√	√	√
Land Suitability Indexes	√	√	√	√
Temperature (c°) & Rain (mm)	√	√	√	√
Province FE	√	√	√	√

OLS estimates in Columns (1-4) and two stage least squares in Columns (5-6). The unit of observation is the municipality. The dependent variables are: the share of land occupied by farms smaller than 10 hectares over the total share of land occupied by farms in the municipality in Panel (A); the share of land occupied by farms larger than 100 hectares over the total share of land occupied by farms in the municipality in Panel (B); the Gini index of 1929 land inequality in the municipality in Panel (C). See Table A2 for details on the dependent variables and Table A8 for summary statistics.*Predicted Malaria Risk* is the average malaria (falciparum) transmission index in the municipality built using the adjusted seasonal temperature from 1500 to 1929. *Predicted Malaria Risk (M)* is the average malaria (falciparum) transmission index in the municipality built using the adjusted monthly temperature from 1901 to 1929. *Predicted Malaria Risk (No Alt)* is the average malaria (falciparum) transmission index in the municipality built using the seasonal temperature from 1500 to 1929. *Malaria Stability* is the Average Malaria Stability in the municipality. Source: Kiszewski et al. (2004). Covariate descriptions and data sources are reported in Tables A4 and A5, and summary statistics are reported in Table A10. Robust standard errors clustered at the provincial level are in round parentheses and Conley standard errors (50 km cutoff) are reported in square brackets. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

B.1.10 Malaria and Land Inequality - Heterogeneous Effects

Figure B3: Malaria and Land Suitability



Binscatter for the relationship between malaria prevalence, measured in 1922, and soil suitability across Italian municipalities.

Further, findings are not driven by structural differences between the plains and the mountains. In Appendix Table B11, we replicate baseline results focusing on municipalities located below 400 meters in elevation. Importantly, Appendix Table B12 shows that the documented effect is also not related to infertile places. On the contrary, malaria-infested areas are, in fact, generally more fertile (see Appendix Figure B3). Appendix Table B12 shows that the effect is stronger when focusing only on the municipalities with higher potential agricultural yields. In other words, in unsuitable locations - with limited scope for agricultural production - malaria does not affect the distribution of wealth as it is not a constraint to the already limited agricultural productivity of these areas. On the contrary, in highly suitable municipalities, the presence of malaria crucially affects the structure of land distribution.

Table B11: Heterogeneous Effects wrt Elevation

Panel A	Small Landowners (1929)											
	(1) OLS	(2) OLS	(3) ITT	(4) ITT	(5) IV	(6) IV	(7) OLS	(8) OLS	(9) ITT	(10) ITT	(11) IV	(12) IV
	Altitude ≤ Median						Altitude > Median					
Malaria (1922)	-0.305***	-0.231***			-0.352***	-0.199	-0.064	-0.007			-0.110**	0.937
	(0.035)	(0.034)			(0.068)	(0.164)	(0.039)	(0.020)			(0.052)	(4.336)
	[0.057]	[0.047]			[0.096]	[0.158]	[0.048]	[0.018]			[0.056]	[4.478]
Predicted Malaria Risk			-0.093***	-0.045					-0.055**	0.014		
			(0.025)	(0.042)					(0.027)	(0.028)		
			[0.033]	[0.042]					[0.030]	[0.032]		
Mean Dependent	0.52	0.52	0.52	0.52	0.52	0.52	0.49	0.49	0.49	0.49	0.49	0.49
F-statistic Instrument					112	17					46	0
Observations	3,886	3,886	3,886	3,886	3,886	3,886	3,944	3,944	3,944	3,944	3,944	3,944
R-Square	0.27	0.51	0.12	0.44	0.27	0.51	0.01	0.31	0.01	0.31	0.01	-0.87
Panel B	Landmarks (1929)											
Malaria (1922)	0.171***	0.114***			0.210***	0.380***	0.044	0.006			0.074	-0.981
	(0.025)	(0.027)			(0.039)	(0.119)	(0.031)	(0.021)			(0.048)	(4.084)
	[0.025]	[0.028]			[0.041]	[0.133]	[0.039]	[0.020]			[0.056]	[4.083]
Predicted Malaria Risk			0.056***	0.087***					0.037	-0.015		
			(0.013)	(0.032)					(0.023)	(0.021)		
			[0.014]	[0.033]					[0.027]	[0.021]		
Mean Dependent	0.11	0.11	0.11	0.11	0.11	0.11	0.23	0.23	0.23	0.23	0.23	0.23
F-statistic Instrument					112	17					46	0
Observations	3,886	3,886	3,886	3,886	3,886	3,886	3,944	3,944	3,944	3,944	3,944	3,944
R-Square	0.18	0.36	0.09	0.33	0.17	0.15	0.01	0.33	0.00	0.33	0.00	-1.05
Panel C	Land Inequality (1929)											
Malaria (1922)	0.156***	0.115***			0.190***	0.215***	0.057***	-0.003			0.123***	-0.152
	(0.017)	(0.017)			(0.026)	(0.081)	(0.020)	(0.012)			(0.031)	(1.185)
	[0.025]	[0.023]			[0.037]	[0.080]	[0.026]	[0.012]			[0.033]	[1.290]
Predicted Malaria Risk			0.050***	0.049**					0.061***	-0.002		
			(0.010)	(0.021)					(0.013)	(0.015)		
			[0.014]	[0.024]					[0.015]	[0.017]		
Mean Dependent	0.60	0.60	0.60	0.60	0.60	0.60	0.61	0.61	0.61	0.61	0.61	0.61
F-statistic Instrument					112	17					46	0
Observations	3,886	3,886	3,886	3,886	3,886	3,886	3,944	3,944	3,944	3,944	3,944	3,944
R-Square	0.29	0.44	0.14	0.37	0.27	0.38	0.03	0.31	0.03	0.31	-0.01	0.22
Geographical Characteristics	×	√	×	√	×	√	×	√	×	√	×	√
Land Suitability Indexes	×	√	×	√	×	√	×	√	×	√	×	√
Temperature (c°) & Rain (mm)	×	√	×	√	×	√	×	√	×	√	×	√
Region FE	×	√	×	√	×	√	×	√	×	√	×	√

OLS estimates in Columns (1-4) & (7-10) and two stage least squares in Columns (5-6) & (11-12). The unit of observation is the municipality. The dependent variables are: the share of land occupied by farms smaller than 10 hectares over the total share of land occupied by farms in the municipality in Panel (A); the share of land occupied by farms larger than 100 hectares over the total share of land occupied by farms in the municipality in Panel (B); the Gini index of 1929 land inequality in the municipality in Panel (C). See Table A2 for details on the dependent variables and Table A8 for summary statistics. *Malaria 1922* is a dummy variable equal to one if in 1992 there was malaria in the municipality and zero otherwise. *Predicted Malaria Risk* measures the predicted risk of malaria transmission constructed following Mordecai et al. (2013). Covariate descriptions and data sources are reported in Tables A4 and A5, and summary statistics are reported in Table A10. Robust standard errors clustered at the provincial level are in round parentheses and Conley standard errors (50 km cutoff) are reported in square brackets. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table B12: Heterogeneous Effects wrt Soil Suitability

Panel A	Small Landowners (1929)											
	(1) OLS	(2) OLS	(3) ITT	(4) ITT	(5) IV	(6) IV	(7) OLS	(8) OLS	(9) ITT	(10) ITT	(11) IV	(12) IV
	Land Suitability ≤ Median						Land Suitability > Median					
Malaria (1922)	-0.117***	-0.080***			0.012	-0.354	-0.248***	-0.221***			-0.523***	-0.415***
	(0.040)	(0.029)			(0.084)	(0.656)	(0.044)	(0.039)			(0.052)	(0.084)
	[0.043]	[0.023]			[0.091]	[0.585]	[0.074]	[0.056]			[0.084]	[0.077]
Predicted Malaria Risk			0.003	0.010					-0.157***	-0.136***		
			(0.018)	(0.019)					(0.020)	(0.034)		
			[0.019]	[0.017]					[0.036]	[0.040]		
Mean Dependent	0.47	0.47	0.47	0.47	0.47	0.47	0.54	0.54	0.54	0.54	0.54	0.54
F-statistic Instrument					57	1					62	47
Observations	3,862	3,862	3,862	3,862	3,862	3,862	3,968	3,968	3,968	3,968	3,968	3,968
R-Square	0.05	0.32	0.00	0.31	-0.01	0.21	0.16	0.47	0.20	0.42	-0.04	0.41
Panel B	Landmarks (1929)											
Malaria (1922)	0.012	0.019			-0.142**	0.021	0.096**	0.124***			0.063	0.383***
	(0.031)	(0.022)			(0.072)	(0.493)	(0.038)	(0.033)			(0.066)	(0.066)
	[0.037]	[0.019]			[0.078]	[0.474]	[0.039]	[0.034]			[0.072]	[0.057]
Predicted Malaria Risk			-0.030**	-0.001					0.019	0.125***		
			(0.013)	(0.014)					(0.022)	(0.031)		
			[0.015]	[0.013]					[0.024]	[0.033]		
Mean Dependent	0.26	0.26	0.26	0.26	0.26	0.26	0.08	0.08	0.08	0.08	0.08	0.08
F-statistic Instrument					57	1					62	47
Observations	3,862	3,862	3,862	3,862	3,862	3,862	3,968	3,968	3,968	3,968	3,968	3,968
R-Square	0.00	0.28	0.02	0.28	-0.09	0.28	0.06	0.23	0.01	0.23	0.06	-0.03
Panel C	Land Inequality (1929)											
Malaria (1922)	0.061***	0.019			0.026	0.301	0.140***	0.121***			0.257***	0.304***
	(0.018)	(0.014)			(0.037)	(0.409)	(0.025)	(0.020)			(0.029)	(0.048)
	[0.019]	[0.011]			[0.042]	[0.405]	[0.036]	[0.026]			[0.042]	[0.047]
Predicted Malaria Risk			0.006	-0.009					0.077***	0.100***		
			(0.008)	(0.010)					(0.013)	(0.019)		
			[0.010]	[0.009]					[0.021]	[0.023]		
Mean Dependent	0.63	0.63	0.63	0.63	0.63	0.63	0.57	0.57	0.57	0.57	0.57	0.57
F-statistic Instrument					57	1					62	47
Observations	3,862	3,862	3,862	3,862	3,862	3,862	3,968	3,968	3,968	3,968	3,968	3,968
R-Square	0.05	0.27	0.00	0.27	0.03	-0.14	0.19	0.40	0.18	0.36	0.06	0.21
Geographical Characteristics	×	√	×	√	×	√	×	√	×	√	×	√
Land Suitability Indexes	×	√	×	√	×	√	×	√	×	√	×	√
Temperature (c°) & Rain (mm)	×	√	×	√	×	√	×	√	×	√	×	√
Region FE	×	√	×	√	×	√	×	√	×	√	×	√

OLS estimates in Columns (1-4) & (7-10) and two stage least squares in Columns (5-6) & (11-12). The unit of observation is the municipality. The dependent variables are: the share of land occupied by farms smaller than 10 hectares over the total share of land occupied by farms in the municipality in Panel (A); the share of land occupied by farms larger than 100 hectares over the total share of land occupied by farms in the municipality in Panel (B); the Gini index of 1929 land inequality in the municipality in Panel (C). See Table A2 for details on the dependent variables and Table A8 for summary statistics. *Malaria 1922* is a dummy variable equal to one if in 1992 there was malaria in the municipality and zero otherwise. *Predicted Malaria Risk* measures the predicted risk of malaria transmission constructed following Mordecai et al. (2013). Covariate descriptions and data sources are reported in Tables A4 and A5, and summary statistics are reported in Table A10. Robust standard errors clustered at the provincial level are in round parentheses and Conley standard errors (50 km cutoff) are reported in square brackets. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

B.1.11 Accounting for Non-linear Effect of Geography through LASSO Method

Another potential local confounder might arise from the fact that our measures of predicted malaria prevalence are non-linear functions of temperature (Mordecai et al. 2013), and of temperature and precipitation (Malaria Stability). While baseline regressions account for the main and the squared effects of temperature and precipitation, climate could affect inequality in a even more non-linear and interacted way. In Appendix Table B13, we exploit the high-dimensional LASSO method presented by Belloni, Chernozhukov, and Hansen (2014) to more flexibly account for the effect of geography. This method allows for a principled search of controls, whenever the potential universe of controls is very large. In our case, LASSO allows to account for the impact of first, second, and third order effects of seasonal temperatures and precipitations, and a full set of interactions between all these various order terms.

Table B13: Accounting for Geography through LASSO Method

Panel A	Small Landowners (1929)					
	(1) All	(2) Lasso	(3) All	(4) Lasso	(5) All	(6) Lasso
Predicted Malaria Risk	-0.119**	-0.046	-0.071***	-0.079***	-0.089*	-0.066**
	(0.054)	(0.039)	(0.023)	(0.022)	(0.046)	(0.026)
Observations	7,830	7,830	7,830	7,830	7,830	7,830
R-Square	0.405	0.245	0.327	0.275	0.494	0.254
Panel B	Landmarks (1929)					
	(1) All	(2) Lasso	(3) All	(4) Lasso	(5) All	(6) Lasso
Predicted Malaria Risk	0.076*	0.038	0.040**	0.044**	0.034	0.023
	(0.040)	(0.027)	(0.017)	(0.017)	(0.033)	(0.021)
Observations	7,830	7,830	7,830	7,830	7,830	7,830
R-Square	0.381	0.300	0.348	0.322	0.438	0.302
Panel B	Land Inequality (1929)					
	(1) All	(2) Lasso	(3) All	(4) Lasso	(5) All	(6) Lasso
Predicted Malaria Risk	0.090***	0.043**	0.033***	0.038***	0.075***	0.037***
	(0.027)	(0.019)	(0.013)	(0.012)	(0.024)	(0.014)
Observations	7,830	7,830	7,830	7,830	7,830	7,830
R-Square	0.369	0.252	0.294	0.257	0.437	0.237
Temperature Varibles and Interactions (c°)	√	√	×	×	√	√
Rain^2 Variables and Interactions (mm)	×	×	√	√	√	√
Geographical Characteristics	√	√	√	√	√	√
Land Suitability Indexes	√	√	√	√	√	√
Region FE	√	√	√	√	√	√

OLS estimates. We exploit the high-dimensional LASSO method presented by Belloni, Chernozhukov, and Hansen (2014). Columns 1, 3 and 5 contain all geographical measures, second and third order term with related interacted terms. Columns 2, 4, and 6 contain the remaining terms as selected by the LASSO method. The dependent variables are: the share of land occupied by farms smaller than 10 hectares over the total share of land occupied by farms in the municipality in Panel (A); the share of land occupied by farms larger than 100 hectares over the total share of land occupied by farms in the municipality in Panel (B); the Gini index of 1929 land inequality in the municipality in Panel (C). See Table A2 for details on the dependent variables and Table A8 for summary statistics. *Predicted Malaria Risk* measures the predicted risk of malaria transmission. Covariates description and data sources are reported in Tables A4 and A5, and summary statistics are reported in Table A10. Robust standard errors clustered at the provincial level are in round parentheses and Conley standard errors (50 km cutoff) are reported in square brackets. Significance at the 10% level is represented by , at the 5% level by *, and at the 1% level by ***.

B.1.12 Malaria Risk and Coping Strategies

Table B14: Protective Settlement Patterns

Panel A	Protective Settlement Pattern			
	Total Population	Population Density	Agricultural Population	Scattered House Population
	(1) ITT	(2) ITT	(3) ITT	(4) ITT
Predicted Malaria Risk	-0.028	0.074	-0.055	-0.142***
	(0.021)	(0.074)	(0.062)	(0.038)
	[0.019]	[0.075]	[0.057]	[0.053]
Mean Dependent	4,948.23	168.98	0.51	0.29
Observations	7,830	7,830	7,830	7,748
R-Square	0.05	0.16	0.32	0.27
Geographical Characteristics	√	√	√	√
Land Suitability Indexes	√	√	√	√
Temperature (c°) & Rain (mm)	√	√	√	√
Region FE	√	√	√	√

OLS estimates in Columns (1-4) and two stage least squares in Columns (5-6). The unit of observation is the municipality. The dependent variables are: the municipality area in Km^2 in Column (1); the total population over the total area of the municipality in square kilometres in Column (2); the share of population who works in the agriculture sector over the total population in the municipality in Column (3); the share of population that is scattered in the countryside and not located within the town over the total population in the municipality in Column (4). See Table A2 for details on the dependent variables and Table A8 for summary statistics. *Predicted Malaria Risk* measures the predicted risk of malaria transmission constructed following Mordecai et al. (2013). Covariate descriptions and data sources are reported in Tables A4 and A5, and summary statistics are reported in Table A10. Robust standard errors clustered at the provincial level are in round parentheses and Conley standard errors (50 km cutoff) are reported in square brackets. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table B15: Labor-Saving Crop Choices

Panel A	Labour Saving Crop Choices			
	Total Cultivated Land	Unproductive Land	Land Devoted to Pasture	Land Devoted to Vineyards and Fruits
	(1) ITT	(2) ITT	(3) ITT	(4) ITT
Predicted Malaria Risk	0.025	0.038	0.215***	-0.096
	(0.034)	(0.042)	(0.057)	(0.062)
	[0.041]	[0.041]	[0.053]	[0.074]
Mean Dependent	0.94	0.08	0.01	0.24
Observations	7,748	7,748	7,748	7,748
R-Square	0.10	0.29	0.17	0.45
Geographical Characteristics	√	√	√	√
Land Suitability Indexes	√	√	√	√
Temperature (c°) & Rain (mm)	√	√	√	√
Region FE	√	√	√	√

OLS estimates in Columns (1-4) and two stage least squares in Columns (5-6). The unit of observation is the municipality. The dependent variables are: the total share of land occupied by farms over the total area of the municipality in hectares in the municipality in Column (1); the share of land occupied by unproductive land over the total share of land occupied by farms in the municipality in Column (2); the share of land occupied by pastures over the total share of land occupied by farms in the municipality in Column (3); the share of land occupied by viticulture crops over the total share of land occupied by farms in the municipality in Column (4). See Table A2 for details on the dependent variables and Table A8 for summary statistics. *Predicted Malaria Risk* measures the predicted risk of malaria transmission constructed following Mordecai et al. (2013). Covariate descriptions and data sources are reported in Tables A4 and A5, and summary statistics are reported in Table A10. Robust standard errors clustered at the provincial level are in round parentheses and Conley standard errors (50 km cutoff) are reported in square brackets. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table B16: Agrarian Contracts and Arrangements

Panel A	Agrarian Contracts & Arrangements			
	Farmers per Farm	Land Owner	Share Croppers (Tenants)	Daily Workers
	(1) ITT	(2) ITT	(3) ITT	(4) ITT
Predicted Malaria Risk	0.153**	-0.134*	-0.414***	0.364***
	(0.070)	(0.068)	(0.074)	(0.068)
	[0.063]	[0.069]	[0.087]	[0.075]
Mean Dependent	4.32	0.51	0.12	0.18
Observations	4,593	4,593	4,593	4,593
R-Square	0.32	0.54	0.25	0.49
Geographical Characteristics	√	√	√	√
Land Suitability Indexes	√	√	√	√
Temperature (c°) & Rain (mm)	√	√	√	√
Region FE	√	√	√	√

OLS estimates in Columns (1-4) and two stage least squares in Columns (5-6). The unit of observation is the municipality. The dependent variables are: the share of population who works in the agriculture sector over the total number od farms in the municipality in Column (1); the share of population who works in the agriculture sector and own his own land over the total number od farms in the municipality in Column (2); the share of population who works in the agriculture as a tenant farmer who gives a part of each crop as rent over the total number od farms in the municipality in Column (3); the share of population who works in the agriculture and is paid by the day over the total number od farms in the municipality in Column (4). See Table A2 for details on the dependent variables and Table A8 for summary statistics. *Predicted Malaria Risk* measures the predicted risk of malaria transmission constructed following Mordecai et al. (2013). Covariate descriptions and data sources are reported in Tables A4 and A5, and summary statistics are reported in Table A10. Robust standard errors clustered at the provincial level are in round parentheses and Conley standard errors (50 km cutoff) are reported in square brackets. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table B17: Malari and Land Inequality: sub-sample of Municipalities with Info on Contracts

Panel A	Small Landowners (1929)					
	(1) OLS	(2) OLS	(3) ITT	(4) ITT	(5) IV	(6) IV
Malaria (1922)	-0.191***	-0.146***			-0.144**	-0.444**
	(0.036)	(0.040)			(0.070)	(0.192)
	[0.046]	[0.042]			[0.093]	[0.168]
Predicted Malaria Risk			-0.036*	-0.040*		
			(0.021)	(0.023)		
			[0.029]	[0.023]		
Mean Dependent	0.51	0.51	0.51	0.51	0.51	0.51
F-statistic Instrument					49	6
Observations	4,593	4,593	4,593	4,593	4,593	4,593
R-Square	0.12	0.41	0.02	0.38	0.11	0.27
Panel B	Landmarks (1929)					
Malaria (1922)	0.091***	0.076**			0.000	0.529***
	(0.029)	(0.034)			(0.061)	(0.169)
	[0.038]	[0.033]			[0.074]	[0.168]
Predicted Malaria Risk			0.000	0.048**		
			(0.015)	(0.020)		
			[0.018]	[0.020]		
Mean Dependent	0.22	0.22	0.22	0.22	0.22	0.22
F-statistic Instrument					49	6
Observations	4,593	4,593	4,593	4,593	4,593	4,593
R-Square	0.03	0.38	0.00	0.38	0.00	-0.03
Panel C	Land Inequality (1929)					
Malaria (1922)	0.106***	0.070***			0.103***	0.291***
	(0.020)	(0.023)			(0.035)	(0.100)
	[0.024]	[0.024]			[0.045]	[0.108]
Predicted Malaria Risk			0.025**	0.026*		
			(0.011)	(0.013)		
			[0.015]	[0.015]		
Mean Dependent	0.61	0.61	0.61	0.61	0.61	0.61
F-statistic Instrument					49	6
Observations	4,593	4,593	4,593	4,593	4,593	4,593
R-Square	0.12	0.33	0.03	0.31	0.12	0.08
Geographical Characteristics	×	√	×	√	×	√
Land Suitability Indexes	×	√	×	√	×	√
Temperature (c°) & Rain (mm)	×	√	×	√	×	√
Region FE	×	√	×	√	×	√

OLS estimates in Columns (1-4) and two stage least squares in Columns (5-6). The unit of observation is the municipality. The dependent variables are: the share of land occupied by farms smaller than 10 hectares over the total share of land occupied by farms in the municipality in Panel (A); the share of land occupied by farms larger than 100 hectares over the total share of land occupied by farms in the municipality in Panel (B); the Gini index of 1929 land inequality in the municipality in Panel (C). See Table A2 for details on the dependent variables and Table A8 for summary statistics. *Malaria 1922* is a dummy variable equal to one if in 1992 there was malaria in the municipality and zero otherwise. *Predicted Malaria Risk* measures the predicted risk of malaria transmission constructed following Mordecai et al. (2013). Covariate descriptions and data sources are reported in Tables A4 and A5, and summary statistics are reported in Table A10. Robust standard errors clustered at the provincial level are in round parentheses and Conley standard errors (50 km cutoff) are reported in square brackets. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

B.2 The Long-Term Effects of Land-Inequality

The eradication of malaria in Italy is regarded as a tremendous success story: in 1952, only 113 cases of malaria were registered throughout the country, 19 in 1954, and 15 in 1954.[6] Indeed, these results inspired further elimination campaigns promoted by the World Health Organization around the world (1957-1969).

Yet while the successful eradication of the disease liberated the country from the burden of sickness and deaths, it did not heal all the social and economic wounds caused by a history of malaria exposure. In constraining agricultural production and shaping the historical distribution of wealth, the disease cast a long shadow on the development of affected areas even after it was completely expunged.

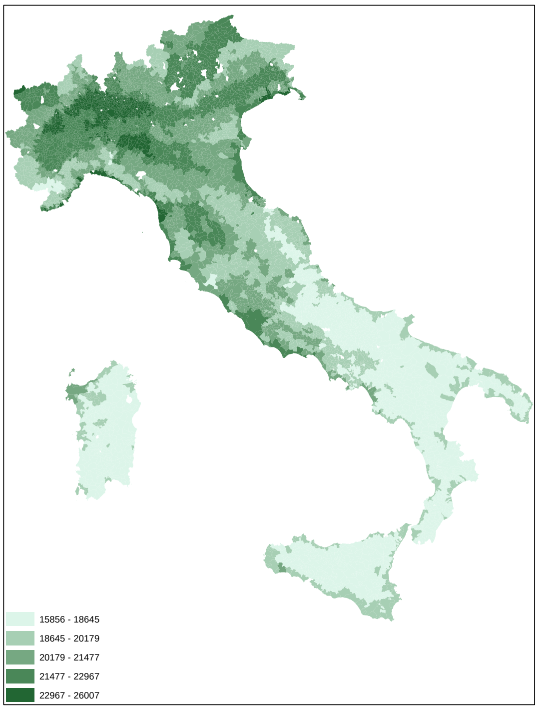
To conclude, we show that historical malaria prevalence continues to haunt those Italian municipalities with a history of malaria exposure. The consequences of centuries of land inequalities fostered by the disease persisted after eradication, with important ramifications for human capital and the economic development of municipalities more historically affected by the disease. While the malaria elimination campaigns rapidly reduced inequality in land ownership, the consequences of a *history* of inequality could not be eradicated.

The negative legacy of malaria may have endured due to a host of channels: historical low agricultural productivity and poverty can indubitably persist over time, with long-term consequences. malaria exposure operated through an history of inequalities in landownership.

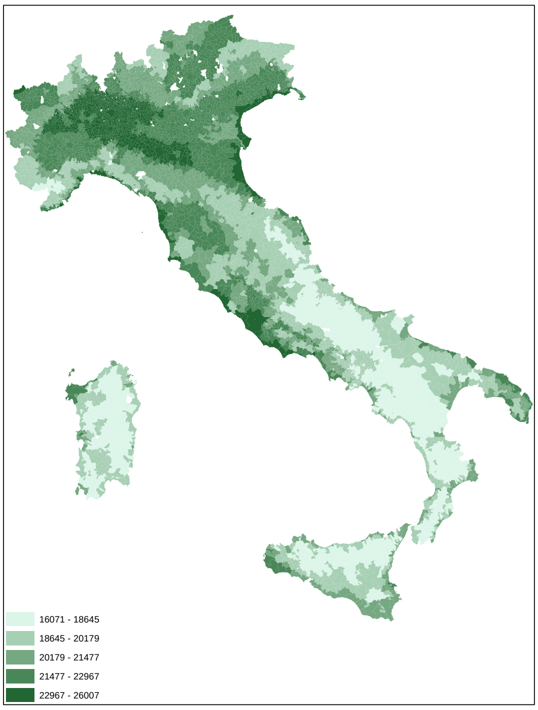
Long-term results are summarized in Table B18. We find that municipalities exposed to malaria in 1922 today have a lower share of individuals with a university degree, higher unemployment and a sizable lower average income. We argue that part of the disease's legacy has operated through land inequalities. As a way to visualize the historical burden of malaria, we attempt to quantify potential municipality income in the absence of a history of malaria exposure. Following the baseline specification of Column 4 - Table B18, we predict average income in the municipality comparing the full model with that where we artificially shut down malaria predicted risk. Figure B4 shows that several southern regions, particularly Sicily, Puglia, and locations along the coasts, would today have a substantially higher average income.

[6] This last remaining cases were mostly due to relapse in the most remote areas, challenging to reach for health authorities. Source: *Cenni storici sulla campagna di eradicazione della malaria in Italia*, Squarciane et al., 1998)

Figure B4: Italy with and without a History of Malaria Exposure - A THOUGHT EXPERIMENT



(a) With Historical Malaria



(b) Without Historical Malaria

The figure on the left portrays the predicted income in the municipality in our full model, in the figure on the right we pretend malaria risk to be absent throughout the country.

Table B18: Historical Malaria Prevalence and Long-Term Development

Panel A	Average Income					
	(1) OLS	(2) OLS	(3) ITT	(4) ITT	(5) IV	(6) IV
Malaria (1922)	-0.094***	-0.042***			-0.091***	-0.087**
	(0.016)	(0.008)			(0.025)	(0.038)
	[0.022]	[0.009]			[0.034]	[0.053]
Predicted Malaria Risk			-0.023***	-0.012**		
			(0.006)	(0.005)		
			[0.009]	[0.007]		
Mean Dependent	9.91	9.91	9.91	9.91	9.91	9.91
F-statistic Instrument					112	19
Observations	7,824	7,824	7,824	7,824	7,824	7,824
R-Square	0.10	0.48	0.03	0.47	0.10	0.47
Panel C	Unemployment Rate					
Malaria (1922)	0.057***	-0.004			0.098***	0.060**
	(0.008)	(0.003)			(0.018)	(0.024)
	[0.011]	[0.003]			[0.025]	[0.030]
Predicted Malaria Risk			0.025***	0.008***		
			(0.004)	(0.002)		
			[0.005]	[0.003]		
Mean Dependent	0.10	0.10	0.10	0.10	0.10	0.10
F-statistic Instrument					112	19
Observations	7,825	7,825	7,825	7,825	7,825	7,825
R-Square	0.19	0.73	0.16	0.73	0.10	0.61
Panel C	University Rate					
Malaria (1922)	-0.002	-0.007***			0.004	-0.016*
	(0.002)	(0.002)			(0.003)	(0.009)
	[0.002]	[0.002]			[0.004]	[0.012]
Predicted Malaria Risk			0.001	-0.002*		
			(0.001)	(0.001)		
			[0.001]	[0.001]		
Mean Dependent	0.07	0.07	0.07	0.07	0.07	0.07
F-statistic Instrument					112	19
Observations	7,825	7,825	7,825	7,825	7,825	7,825
R-Square	0.00	0.15	0.00	0.15	-0.01	0.14
Geographical Characteristics	×	√	×	√	×	√
Land Suitability Indexes	×	√	×	√	×	√
Temperature (c°) & Rain (mm)	×	√	×	√	×	√
Region FE	×	√	×	√	×	√

OLS estimates. The unit of observation is the municipality. In all specifications, the *dependent variable* is the Gini index of 1929 land inequality in the municipality. See Table A2 for details on the dependent variables and Table A8 for summary statistics. *Malaria 1882 (1922)* is a dummy variable equal to one if in 1882 (1992) there was malaria in the municipality and zero otherwise. *Predicted Malaria Risk* measures the predicted risk of malaria transmission constructed following Mordecai et al. (2013). *Malaria Stability* measures the average malaria suitability in the municipality, as from Kiszewski et al. (2004). Covariate descriptions and data sources are reported in Tables A4 and A5, and summary statistics are reported in Table A10. Robust standard errors clustered at the provincial level are in round parentheses and Conley standard errors (50 km cutoff) are reported in square brackets. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

References

Belloni, A., V. Chernozhukov, and C. Hansen (2014): "High-dimensional methods and inference on structural and treatment effects," *Journal of Economic Perspectives*, 28(2), 29–50.

Bødker, R., J. Akida, D. Shayo, W. Kisinza, H. Msangeni, E. Pedersen, and S. Lindsay (2003): "Relationship between altitude and intensity of malaria transmission in the Usambara Mountains, Tanzania," *Journal of Medical Entomology*, 40(5), 706–717.

Brombacher, W. G. (1936): "Altitude-pressure tables based on the United States standard atmosphere," .

Flückiger, M., and M. Ludwig (2017): "Malaria suitability, urbanization and persistence: Evidence from China over more than 2000 years," *European Economic Review*, 92, 146–160.

Galor, O., and Ö. Özak (2015): "Land productivity and economic development: Caloric suitability vs. agricultural suitability," *Agricultural Suitability (July 12, 2015)*.

——— (2016): "The agricultural origins of time preference," *American Economic Review*, 106(10), 3064–3103.

Justus, C. G., and A. Woodrum (1972): "Atmospheric pressure, density, temperature and wind variations between 50 and 200 km," .

Kiszewski, A., A. Mellinger, A. Spielman, P. Malaney, S. E. Sachs, and J. Sachs (2004): "A global index representing the stability of malaria transmission," *The American Journal of Tropical Medicine and Hygiene*, 70(5), 486–498.

Luterbacher, J., D. Dietrich, E. Xoplaki, M. Grosjean, and H. Wanner (2004): "European seasonal and annual temperature variability, trends, and extremes since 1500," *Science*, 303(5663), 1499–1503.

Mordecai, E. A., K. P. Paaijmans, L. R. Johnson, C. Balzer, T. Ben-Horin, E. de Moor, A. McNally, S. Pawar, S. J. Ryan, T. C. Smith, et al. (2013): "Optimal temperature for malaria transmission is dramatically lower than previously predicted," *Ecology Letters*, 16(1), 22–30.

New, M., D. Lister, M. Hulme, and I. Makin (2002): "A high-resolution data set of surface climate over global land areas," *Climate Research*, 21(1), 1–25.

Pauling, A., J. Luterbacher, C. Casty, and H. Wanner (2006): "Five hundred years of gridded high-resolution precipitation reconstructions over Europe and the connection to large-scale circulation," *Climate Dynamics*, 26(4), 387–405.

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

Rumpf, S. B., K. Hülber, G. Klonner, D. Moser, M. Schütz, J. Wessely, W. Willner, N. E. Zimmermann, and S. Dullinger (2018): "Range dynamics of mountain plants decrease with elevation," *Proceedings of the National Academy of Sciences*, 115(8), 1848–1853.

Xoplaki, E., J. Luterbacher, H. Paeth, D. Dietrich, N. Steiner, M. Grosjean, and H. Wanner (2005): "European spring and autumn temperature variability and change of extremes over the last half millennium," *Geophysical Research Letters*, 32(15).