Bocconi

WORKING PAPER SERIES
ISSN 2785-1079

Number 15
MAY 2021

MOSQUITOES AND POTATOES: HOW GLOBAL HEALTH CRISES IMPEDE DEVELOPMENT

Maurizio Malpede
Giacomo Falchetta
Soheil Shayegh

Università Bocconi
GREEN
Centre for Geography, Resources, Environment, Energy and Networks

Mosquitoes and Potatoes: How Global Health Crises Impede Development

Maurizio Malpede[a,1], Giacomo Falchetta[b,c], and Soheil Shayegh[d]

[a]GREEN, Bocconi University, Milan, Italy
[b]Fondazione Eni Enrico Mattei, Milan, Italy
[c]Free University of Bozen-Bolzano, Bolzano, Italy
[d]RFF-CMCC European Institute for Economics and the Environment, 20144 Milan, Italy

May 5, 2021

Abstract

The socioeconomic fallout of the COVID-19 pandemic has been felt globally and across many sectors and population groups. While the long-term impact of the pandemic on economic growth is yet to be observed and assessed, history may provide some evidence on how public health emergencies have negatively affected socioeconomic development pathways well beyond the crisis duration. Here we examine the impact of malaria on the contribution of potato to the Old World's development during the 18^{th} and 19^{th} centuries. We exploit local variations in land suitability for potato cultivation and malaria transmission to estimate and compare the impact of growing potato on urbanization and population growth in highly endemic and non-endemic areas. We show that local weather conditions ideal for malaria transmission counteracted the potential benefits of introducing the potato in the Old World. Robustness checks from geographic variations in malaria stability, suitability for potato cultivation, and placebo treatments reinforce the positive effects of eradicating malaria on urbanization and population growth in potato-suitable areas after 1900. Our results highlight the interplay between technological change, public health, and development outcomes.

Key words: Public health, Population, Urbanization, Malaria, Development

1 Introduction

Global public health crises such as the COVID-19 pandemic take a toll on human lives and well-being and have a lasting impact on economic growth and productivity [28, 20, 19]. This is particularly the case in labor-intensive sectors like agriculture, where lower labor productivity due to farmers' poor health conditions may counteract the positive productivity gains from technological innovations [7, 24]. This calls for a systematic assessment of the interaction of these two opposite forces (i.e., improving productivity due to technological innovation and deteriorating development outcomes due to health crises). Here we take a first step in this direction by providing empirical evidence on the interaction of the improved socioeconomic conditions brought by the introduction of the potato to the Old World between 1700 and 1900 [30] and the adverse outcomes of the malaria outbreak [25, 34] as schematically depicted in Figure 1. We analyze the socioeconomic impact of the potato in areas suitable for the cultivation of the crop and with climatic conditions ideal for the diffusion of malaria and associated vector-borne diseases. We compare the urbanization and population growth in these areas with those not affected by malaria. Our analysis thus builds on previous studies linking rural development and migration (to urban areas) with malaria outbreaks [1, 23, 18].

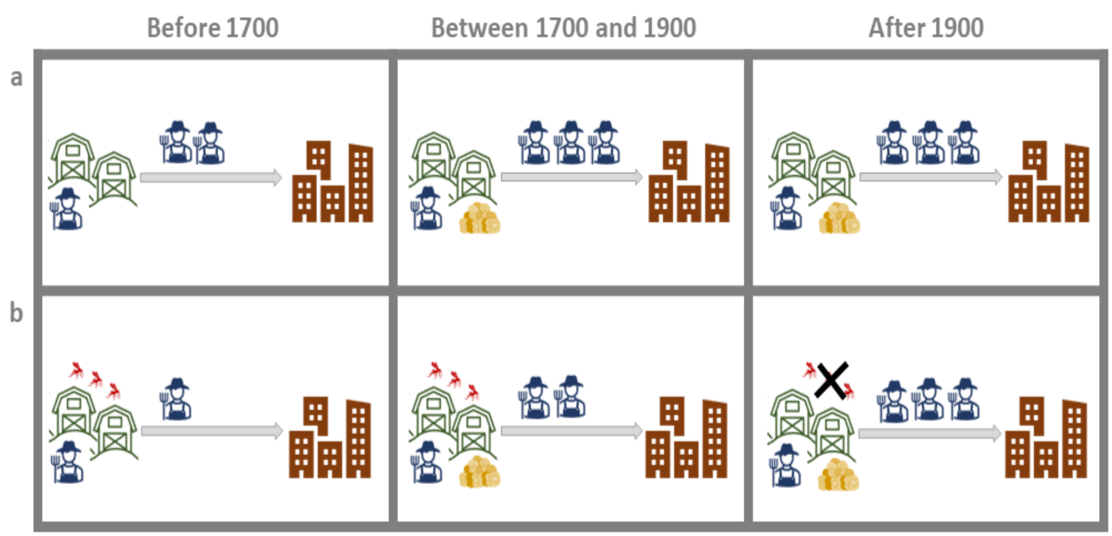

Figure 1: *Historical trends in the introduction of the potato in the Old World from pre-1700 to post-1900 in two areas with high suitability for potato cultivation: (a) without the risk of malaria transmission, and (b) with a high risk of malaria transmission.*

Throughout human history, the introduction of new crops (e.g., maize [5] or potato [30]) or agricultural innovations (e.g., tractors [31] or fertilizers [10]) has been the leading force behind the boost in agricultural productivity and the socioeconomic development that follows it

[4]. Potatoes, in particular, have contributed to one-quarter of the growth in the Old World's population and urbanization between 1700 and 1900 [30, 6]. This has been achieved mainly through two channels. The first channel is *nutrition effect* where the introduction of a new, and more nutrient crop, has resulted in a healthier and more productive population. The second channel is *productivity effect* where increased agricultural productivity has allowed more workers to migrate to the cities and work in industry [13]. Hence, countries with larger areas suitable for growing potatoes have experienced higher population and urbanization growth between 1700 and 1900 as depicted in Figure 1a.

In contrast to the positive impacts of new crops and technological innovations on population growth and development, deteriorating health conditions due to infectious disease outbreaks such as malaria infection have negatively affected economic productivity [14, 8, 3]. Therefore, as panel b in Figure 1 demonstrates, any potential gain from introducing new crops can be offset by the emergence of infectious disease threats that undermine the health and availability of farmers and their productivity [15, 32]. On the contrary, public health interventions can improve the well-being of the farmers and make a positive impact on their productivity [37]. In the case of malaria, widespread eradication programs in the early 1900s in the United States contributed to approximately one-fifth of the agricultural productivity growth at the county level [26]. A more detailed discussion of the positive impacts of agricultural innovations and the adverse effects of malaria outbreaks on socioeconomic development is presented in Appendix A in the SI.

Based on this evidence, we hypothesize that from 1700 to 1900, the positive impact of the potato on the Old World's development was hampered by poor initial health conditions. To evaluate this hypothesis, we assemble a novel grid-cell dataset of climate, population, and urbanization data, covering 31,108 grid cells of the Old World (at a 0.5 x 0.5 degree of resolution) from 1100 to 1900. We use two indexes to measure the suitability of any given area for potato cultivation and malaria transmission. The potato suitability index (PSI) represents the soil conditions required for sustained growth of potato-based on regional weather and land conditions [11]. Higher values of this index indicate that the soil is more suitable for potato cultivation. Similarly, the malaria stability index (MSI) captures the interaction of climatic conditions and malaria prevalence by providing a global indicator of malaria transmission based on regionally dominant vector mosquitoes, temperature, and precipitation data set [21]. We also use additional suitability indexes for other crops [11] and additional geographical factors [29] as control variables in our econometric model. Figure 2 plots the spatial heterogeneity of PSI and MSI. Appendix B in the SI contains detailed description and summary statistics of the data used in our analysis.

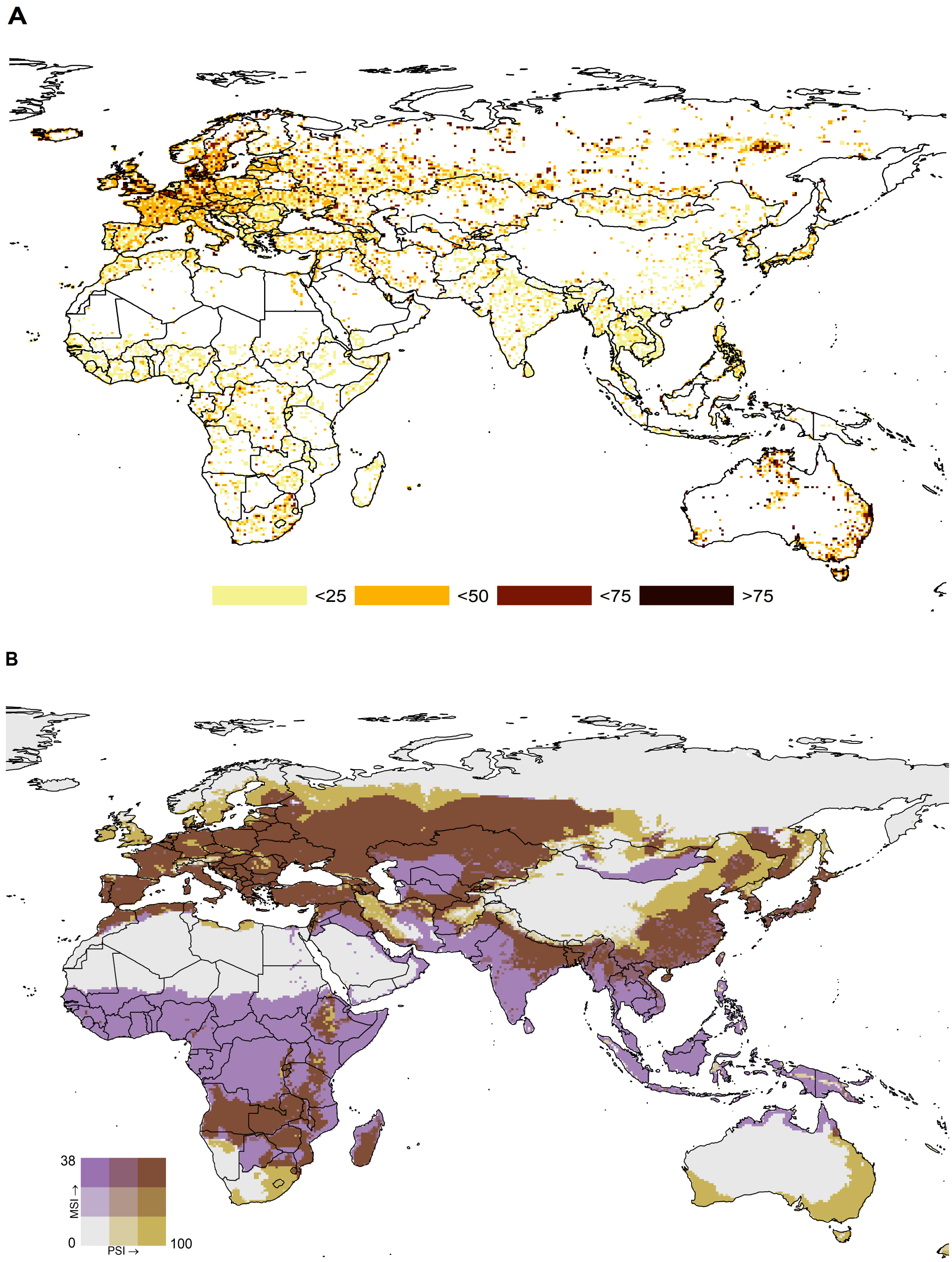

Figure 2: *Maps of the spatial distribution of (A) urbanisation rate change (1700-1900) and (B) the MSI and PSI intersected with country borders.*

We match PSI and MSI with a public long-run grid-cell panel dataset of population and urbanization data [22] (see Methods and Appendix B in the SI). This dataset collects population and urbanization data at the most local level available in existing archives. This allows us to compare the socioeconomic benefits of potato cultivation in areas with high potato suitability and low malaria prevalence with areas equally suitable for cultivating potato but with higher malaria prevalence.

2 Data and Methods

Time invariant grid-cell level data on rain-fed potato suitability (PSI) are retrieved from the Food and Agriculture Organization (FAO)'s Global Agro-Ecological Zones (GAEZ) 2011 database [11]. Malaria Ecology Index [21], indicates weather conditions more or less ideal for the reproduction of mosquito larvae. Both the PSI and the MSI are available at a spatial resolution of 0.5° latitude by 0.5° longitude, which is approximately fifty-six kilometers by fifty-six kilometers (measured at the equator). The PSI ranges from 0 to 1, with one being totally suitable for potato cultivation. In comparison, MSI ranges from 0 to 39, with 39 being extremely suitable for a stable malaria transmission throughout the year. To control for additional time and grid-specific characteristics, which might bias our results, we also consider the grid suitability for growing other crops (i.e., maize, cassava, barley, wheat, rice, and white potato) and additional country-specific characteristics, such as terrain ruggedness, elevation, the presence of a tropical climate, the distance to the equator and the nearest coast [30]. Crop suitability gridded data are retrieved from the Global Agro-Ecological Zones (GAEZ) 2011 database of the FAO [11], while additional country-specific control variables are from the original paper [30]. Finally, to avoid measurement error from changes over time in irrigation intensity and technologies, we have used rain-fed crop suitability measures from (GAEZ).

Grid-cell level data on historical population and urbanization rates are retrieved from the History Database of the Global Environment (HYDE) [22]. This is a very disaggregated dataset with a spatial resolution of 5 arc minute pixels (about 10 kilometers at the equator) covering the whole world from 10,000 BC to 2017 AD. We use century data on population and urbanization from 1100 until 1900. We use GIS software to first convert the spatial resolution of the HYDE database to the same of Potato Suitability Index and Malaria Ecology Index, and second to match the spatial data (Appendix B). To control for additional variables that might have affected the evolution of population and urbanization at grid-level and correlated to land suitability of growing potatoes, we consider crop suitability for other crops (namely cassava, wheat, maize, barley, and sweet potatoes). We intersect the final dataset with historical country boundaries using the GADM (Global Administrative Unit Layer) [16] to include country by time fixed effects and continent dummies. We also control for country-specific characteristics, such as the proportion of cropland or total country area, the ratio of tropical area, the distance from the equator, and the nearest coast [30].

Defining Potato Suitable Areas.

The FAO provides a discrete version of the database that reports the proportion of each grid's land classified under five mutually exclusive categories describing how suitable the environment is for growing each crop. The categories are based on the calculated percentage of the maximum yield that can be attained in each grid cell. The five categories and their corresponding yields are: (i) very suitable land ($0.8 - 1$), (ii) suitable land ($0.6 - 0.8$), (iii) moderately suitable land ($0.4 - 0.6$), (iv) marginally suitable land($0.2 - 0.4$), and (v) unsuitable land ($0 - 0.4$). Following the definition of land suitability for growing potato of the FAO, we define a land to be suitable for cultivation if it is classified in the database as being either very suitable, suitable, or moderately suitable (i.e., with $PSI > 0.4$) [30]. In addition to the categorical version of the index, we also use the continuous version of the PSI (Appendix D in the SI).

Classification of Malaria Endemicity Classes.

We consider different classes of malaria endemicity according to the scientific literature [21]. Specifically: i) non malaria areas are defined as those showing a $MSI < 0.06$; ii) areas with low malaria endemicity ($0.06 < MSI < 1$); iii) mild endemic areas ($1 < MSI < 5$); iv) highly endemic areas ($5 < MSI < 10$); vi) very highly endemic areas ($MSI > 10$). We use the aforementioned classification to assess the impact of potato cultivation on population and urbanization in areas with different malarious conditions. In addition to the categorical version of the index, we also use the continuous version (Appendix D).

Evaluating the Impact of Potato Diffusion on Endemic Areas.

We use a difference-in-difference strategy where the control group is composed of grids not suitable for the cultivation of the potato (i.e., $PSI < 0.4$). We consider two treatment groups. The first treatment group comprises grids suitable for growing potatoes and non-malarious (i.e., $PSI > 0.4$ and $MSI < 1$). The second treatment group is instead composed of grids suitable for growing potatoes and malarious (i.e., $PSI > 0.4$ and $MSI > 1$).

To compute the effects of the introduction of the potato we first estimate the following equation:

$$\mathrm{y}_{i,c,t} = \alpha + \beta(\mathrm{PotMal})_i \times (\mathrm{Post})_t + \gamma\,\mathbf{X}_i^{'} + \omega_i + \delta_t + \sigma_{2,ctrend} + \epsilon_{i,c,t} \tag{1}$$

where index i represents each Old World grid in country c and t indexes time periods considered in the analysis (i.e., 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1750, 1800, 1850, and 1900). The dependent variable is represented as $y_{i,c,t}$ which is either the natural log of population or city population share for each grid i in country c at time t. $(\text{PotMal})_i$ is an indicator variable which takes value zero if the grid is not suitable for the cultivation of the potato (i.e., $PSI < 0.4$) and value of one if the grid is suitable for the cultivation of the potato (i.e., $PSI > 0.4$). Variable $(\text{PotMal})_i$ is interacted with a post indicator variable which takes value of zero for years from 1000 to 1600, which are part of the pre-potato introduction period [30]; and value of one for the post-intervention period ranging from 1700 until 1900. Estimates of equation 1 using alternative cutoff values for potato and malaria suitability thresholds are reported in Appendix F in the SI.

The equation also includes grid and year-fixed effects ω_i and δ_t along with country-specific linear trends $+\sigma_{2,ctrend}$. Moreover, to ensure that the effect of introducing the potato in the Old World is not confounded by other changes in the suitability of crops over the same time period, we include the grid-specific agricultural suitability index [12]. We also include the land suitability for Old World crops such as maize, barley, cassava, and white potato. Furthermore, since potatoes can be successfully cultivated on rugged terrain at high altitudes and are less suitable to be cultivated in tropic areas [30, 29], we control for the natural log of a country's average elevation, the natural log of its ruggedness, and the natural log of a country's land that is defined as being tropical. Standard errors are clustered at the grid-cell level.

The coefficient of interest in equation 1 is β, which is the estimated impact of potato suitability on either population or urbanization. β measures the additional growth in population or urbanization experienced by grid-cells that are suitable for potatoes and not malarious (relative to those that are not) after potatoes were introduced in 1700 (relative to before). A positive coefficient indicates that countries with a geographic environment more suitable for growing potatoes and not malarious witnessed a more significant increase in population after 1700 relative to before 1700.

Table 1 reports the results of equation 1. We find that the diffusion of potato contributed to urbanization and population growth when considering non-malarious grid-cells only. These results further confirm conclusions drawn in recent literature [30]. However, when restricting the sample to only malarious grids (i.e., $MSI > 1$), we find no effects of the diffusion of potato on urbanization and population.

Robustness of the Estimates.

To ensure the validity of our estimates, we employ a series of sensitivity checks. First, equation 1 builds on the assumption of the linearity of the effects of the diffusion of the potato on urbanization and population with the level of endemicity of malaria. We further relax this assumption by employing a lag model that allows us to compute the effects of introducing the potato in different malaria endemicity classes.

The inclusion of fixed effects allows isolating the effects of introducing the potato in the Old World from all other time-invariant factors or common time-varying factors that could be correlated with both suitability for potato cultivation, endemicity of malaria, and economic growth. Specifically, we estimate the following regression:

$$\mathrm{y}_{i,c,t} = \alpha + \sum_{b} \beta_b (\mathrm{PSI})_i \times (\mathrm{Post})_t + \sum_{b} \beta_b \times (\mathrm{PSI})_i + \gamma\, \mathbf{X}_i^{'} + \delta_t + \sigma_{2,ctrend} + \epsilon_{i,c,t} \qquad (2)$$

for $b \in \{0 - 0.06, 0.06 - 1, 1 - 5, 5 - 39\}$.

Results of equation 2 reported in Figure 5 confirm that the potato effects on urbanization and population are found only after 1700 in non-malarious areas, sharply declining in more endemic ones. For locations with a $MSI > 1$, we see few effects on urbanization and no effects on the population, with the estimates approaching zero for higher endemicity levels.

Flexible Estimates.
Another crucial assumption of the empirical strategy in equation 1 is that the cutoff date used for the introduction of potato cultivation in the Old World (i.e., 1700) is correct. The historical evidence suggests that the adoption of the potato began in the late seventeenth century and spread significantly by the early eighteenth century [30, 35, 27]. However, before taking this cutoff date as given, we estimate a fully flexible estimating equation that takes the following form:

$$\mathrm{y}_{i,c,t} = \alpha + \sum_{j=1100}^{1900} \beta_j (\mathrm{PotMal})_i \times \mathrm{I}_t^j + \gamma\, \mathbf{X}_i^{'} + \omega_i + \delta_t + \sigma_{2,ctrend} + \epsilon_{i,c,t} \qquad (3)$$

where all variables are defined as in equation 1. Here we interact PotMal_i with each of the time-period fixed effects rather than a single post-adoption indicator. To be valid, the β_j coefficients must be close to zero for the years before the adoption of potato (i.e., before 1700) and positive for the years after its adoption. Moreover, we expect the coefficients of interest to be positive in magnitude only for non-malarious areas, while no or little effects should be found in more malarious areas. Estimates of equation (3) are reported in Figure 4.

In addition, we consider the continuous version of the Potato Suitability Index (Appendix D in the SI). In Appendix F, we present the results of the analysis considering different cutoff values for the definition of potato suitable areas. Finally, we evaluate the effects of the introduction of the potato in more or less malaria-endemic countries. Results are shown in Appendix G.

3 Results

First, we perform statistical analysis at the grid-cell level to confirm the positive demographic impact of the adoption of potato after its diffusion in the Old World. We estimate the coefficient of the potato suitability interaction term, PSI × Post (first row in Table 1). According to the estimates in column (1), grid-cells that were suitable for potato cultivation (i.e., those with $PSI > 0.4$; see SI Appendix for sensitivity analysis over different thresholds) experienced a 1.8 percentage point increase in urban population compared to non-suitable grid-cells. In addition, the estimated coefficient of the total population in column (4) suggests that the population in areas suitable for potato cultivation increased by 9.5 percentage points relative to areas not suitable for growing potatoes. The Methods section and Appendix C in the SI provide details of the underlying empirical framework. Areas more suitable for potato cultivation provide rural communities with a boost in calorie intake [9] and an increase in agricultural productivity [30, 27]. As a result, these regions have experienced a relatively higher population and urbanization growth compared to areas not suitable for the cultivation of the potato as shown in Figure 3.

However, not all people living in potato-suitable areas could benefit from the positive effects of potato. Infectious diseases such as malaria and associated vector-borne diseases threaten farmers' health and productivity [17, 2]. Malaria is generally more prevalent in rural areas than in urban ones, with farmers being the most affected [33, 36].

We therefore proceed by analysing the mediating role of malaria suitability for the impact of the cultivation of potato on the urbanization rate and population growth. As shown in Figure 3, the agricultural boost brought by the diffusion of the potato led to different socioeconomic development trajectories in potato-suitable areas with higher or lower suitability for malaria transmission. A detailed description of the empirical framework is found in the Methods section and Appendix D in the SI.

Table 1: The impact of potato suitability on urbanization and population: Old World, Pre Malaria Eradication, Gridded Analysis

	Dependent Variable					
	Urbanization Share			Log Population		
	(1)	(2)	(3)	(4)	(5)	(6)
PSI x Post	0.018***			0.095***		
	(0.002)			(0.010)		
PSI x No Mal. x Post		0.030***			0.157***	
		(0.004)			(0.021)	
PSI x Mal. x Post			0.003			-0.044***
			(0.003)			(0.016)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	522197	301909	113190	522197	301909	113190

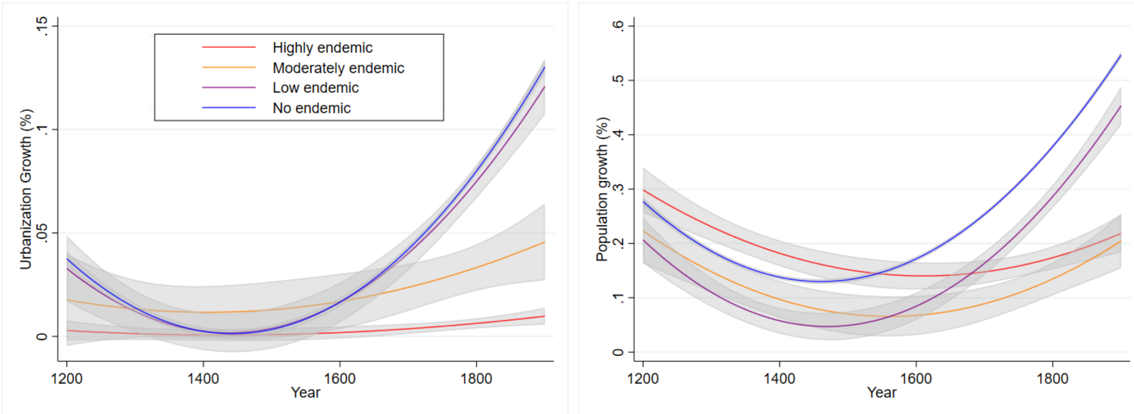

Figure 3: *Historical urbanization (left) and population (right) growth rates in potato suitable areas under different malaria endemicity areas. Gray areas show the 95% confidence intervals.*

The heterogeneous impact of potato cultivation on areas with different initial health conditions can be further evaluated in Figure 4 where the historical association between potato and urbanization rate and population growth has been established for areas with different levels of malaria endemicity. First, we exclude the existence of a preexisting relationship between our outcome variables and potato suitability before the diffusion of the potatoes (i.e., the left-hand side of the vertical red line in all panels). Second, only after 1700 the urbanization and population of potato-suitable locations started to increase relative to locations that were not suitable for growing potato. However, this was only the case for potato-suitable locations with no or low malaria endemicity (i.e., $MSI < 1$). In contrast, the diffusion of the newly discovered staple crop had no impact on equally potato-suitable but highly endemic locations.

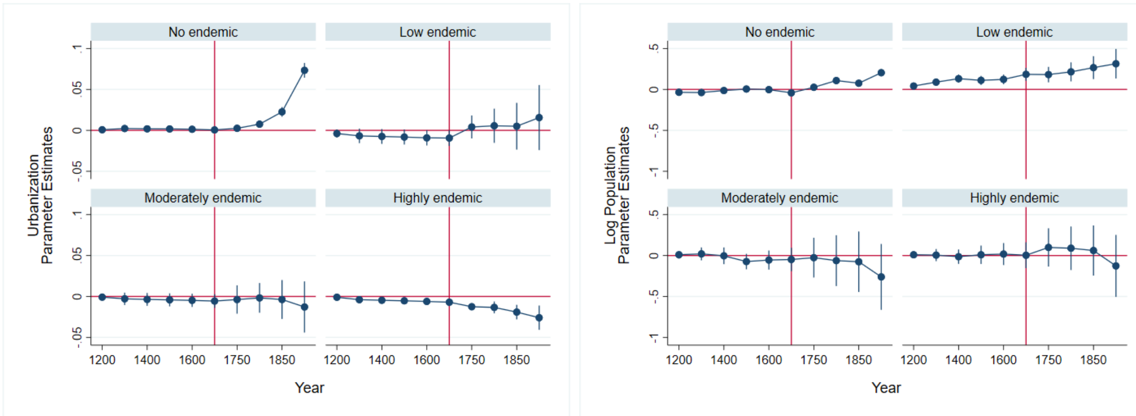

Figure 4: *Flexible estimates of the relationship between potato-suitable areas and urbanization (left) and population (right). Blue vertical lines show the 95% confidence intervals and the vertical red lines indicate the year 1700 when potato was introduced to the Old World.*

These results are confirmed in Table 1. Columns (2) and (5) show that only *non-malarious* ($MSI < 1$) areas have benefited from the positive impacts of the diffusion of the potato. According to estimates in columns (2) and (5), potato-suitable areas increased their urbanization share by three percentage points and total population by 0.157 percentage points relative to non-suitable locations. On the other hand, the suitability for cultivating potato had no statistically significant impact on malarious areas ($MSI > 1$) as estimated coefficients in columns (3) and (6) suggest.

These results become even more robust when evaluating the non-linear effects of potato suitability in areas with different malaria endemicity levels. Figure 5 shows that the diffusion of the potato increased urbanization and population only in the potato-suitable grid-cells with no or low endemicity of malaria.

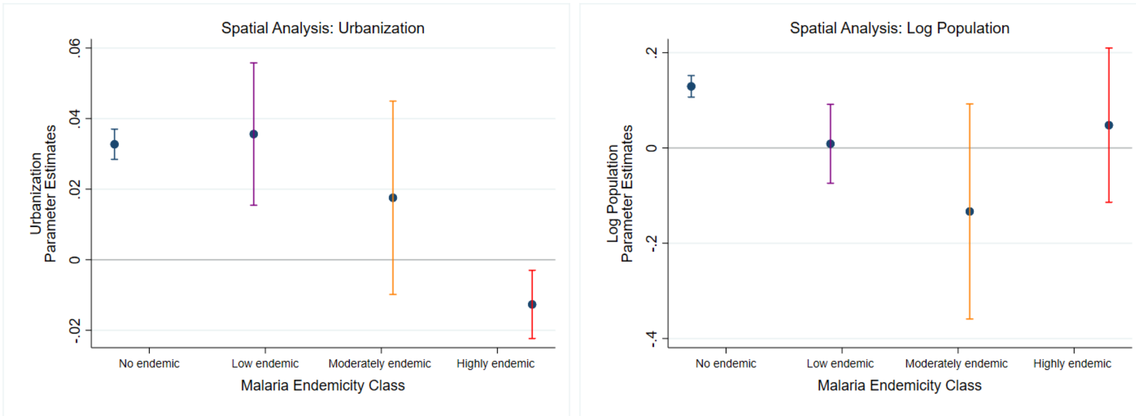

Figure 5: *Effects of the introduction of potato in the Old World on urbanization (left) and on population (right) for each category of malaria endemicity. 95% confidence intervals.*

To validate our findings, we perform several sensitivity and robustness checks. First, we adopt different definitions of potato suitability (in line with previous studies on the impact of potato on urbanization and population [30]). Results of the estimation of equation (1) with different definitions of potato-suitable locations are shown in Appendix F in the SI. Second, and as it is shown in Appendix G in the SI, we run the grid-cell analysis at the country level. Finally, and seen in Appendix H in the SI, we ensure that no residual statistically significant effect is detected after eradicating malaria and associated vector-borne diseases. The outcome of all these assessments further confirms the results of our baseline analysis.

4 Conclusions

While the positive impact of introducing new crops and improvements in agricultural practices have been well established in literature, less attention has been paid to differential impact of emerging public health crises on development pathways in agricultural communities.

Introducing new crops and innovative technologies can boost agricultural productivity and improve food security and other development outcomes. However, it is not clear to what extent public health emergencies (such as the COVID-19 pandemic) may harm this trend by negatively impacting farmer's health and well-being. In this paper we use historical evidence from 1700 to 1900 in the Old World to examine the impact of potato cultivation, a new staple food crop of that time, on urbanization and population growth, conditioned on the prevalence of malaria. Our results indicate that malaria outbreaks had reduced, and in some cases offset, the positive socioeconomic changes caused by the introduction of the potato, underlying an essential link between public health and development.

Focusing on the historical evidence of urbanization and population growth in the areas most suitable for growing potatoes, we have shown how locations that shared the same potato-suitability index had different development trajectories depending on their vulnerability to the endemicity of malaria. While the findings confirm the overall positive role of the potato in urbanization and population growth, they indicate the negative impact of malaria on farmers' health and productivity. This highlights the interplay between health and technological progress, especially in labor-intensive sectors like agriculture, where there is high exposure to the effects of infectious disease outbreaks and other health emergencies.

Our results provide a new foundation for assessing the global and local impacts of emerging health threats on the potential benefits of technological innovations. Our estimates also highlight the dangers from the emergence of infectious diseases due to climate change or other anthropogenic interventions, not only to public health and well-being but also to the socio-economic gains from technological innovations and productivity improvements in the past few decades.

References

[1] Kwadwo Asenso-Okyere et al. "A review of the economic impact of malaria in agricultural development". In: *Agricultural economics* 42.3 (2011), pp. 293–304.

[2] Ousmane Badiane and John Ulimwengu. "Malaria incidence and agricultural efficiency in Uganda". In: *Agricultural Economics* 44.1 (2013), pp. 15–23.

[3] Hoyt Bleakley. "Economic effects of childhood exposure to tropical disease". In: *American Economic Review* 99.2 (2009), pp. 218–23.

[4] Paula Bustos, Bruno Caprettini, and Jacopo Ponticelli. "Agricultural productivity and structural transformation: Evidence from Brazil". In: *American Economic Review* 106.6 (2016), pp. 1320–65.

[5] Shuo Chen and James Kai-sing Kung. "Of maize and men: the effect of a New World crop on population and economic growth in China". In: *Journal of Economic Growth* 21.1 (2016), pp. 71–99.

[6] C Justin Cook. "Potatoes, milk, and the Old World population boom". In: *Journal of Development Economics* 110 (2014), pp. 123–138.

[7] Andre Croppenstedt and Christophe Muller. "The impact of farmers health and nutritional status on their productivity and efficiency: Evidence from Ethiopia". In: *Economic Development and Cultural Change* 48.3 (2000), pp. 475–502.

[8] David Cutler et al. "Early-life malaria exposure and adult outcomes: Evidence from malaria eradication in India". In: *American Economic Journal: Applied Economics* 2.2 (2010), pp. 72–94.

[9] Stanley Davidson and Reginald Passmore. *Human nutrition and dietetics*. E. and S. Livingstone, Ltd., 1959.

[10] Kyle Emerick. "Agricultural productivity and the sectoral reallocation of labor in rural India". In: *Journal of Development Economics* 135 (2018), pp. 488–503.

[11] Günther Fischer et al. *Global agro-ecological zones (GAEZ v3. 0)-model documentation*. Tech. rep. IIASA and FAO, 2012.

[12] Oded Galor and Ömer Özak. *Land productivity and economic development: Caloric suitability vs. agricultural suitability*. preprint. https://ssrn.com/abstract=2625180: SSRN, 2015.

[13] Oded Galor and David N Weil. "Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond". In: *American economic review* 90.4 (2000), pp. 806–828.

[14] Elizabeth Gooch. "The impact of reduced incidence of malaria and other mosquito-borne diseases on global population". In: *Journal of Development Economics* 124 (2017), pp. 214–228.

[15] Casper Worm Hansen and Lars Lønstrup. "The rise in life expectancy and economic growth in the 20th century". In: *The Economic Journal* 125.584 (2015), pp. 838–852.

[16] Robert Hijmans, N. Garcia, and John Wieczorek. *GADM database of Global Administrative Areas*. Tech. rep. 2010.

[17] Margaret Humphreys. *Malaria: poverty, race, and public health in the United States*. JHU Press, 2001.

[18] Nigoon Jitthai. "Migration and malaria." In: *The Southeast Asian journal of tropical medicine and public health* 44 (2013), pp. 166–200.

[19] Òscar Jordà, Sanjay R Singh, and Alan M Taylor. *Longer-run economic consequences of pandemics*. Tech. rep. National Bureau of economic research, 2020.

[20] Marcus Richard Keogh-Brown et al. "The macroeconomic impact of pandemic influenza: estimates from models of the United Kingdom, France, Belgium and The Netherlands". In: *The European Journal of Health Economics* 11.6 (2010), pp. 543–554.

[21] Anthony Kiszewski et al. "A global index representing the stability of malaria transmission". In: *The American journal of tropical medicine and hygiene* 70.5 (2004), pp. 486–498.

[22] Kees Klein Goldewijk et al. "The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years". In: *Global Ecology and Biogeography* 20.1 (2011), pp. 73–86.

[23] Steve W Lindsay and Martin Birley. “Rural development and malaria control in sub-Saharan Africa”. In: *EcoHealth* 1.2 (2004), pp. 129–137.

[24] Maria L Loureiro. “Farmers’ health and agricultural productivity”. In: *Agricultural Economics* 40.4 (2009), pp. 381–388.

[25] Adrienne M Lucas. “The impact of malaria eradication on fertility”. In: *Economic Development and Cultural Change* 61.3 (2013), pp. 607–631.

[26] Maurizio Malpede. “Vector-Borne Diseases and Economic Activity: Evidence from Historical Farmer Productivity in the US”. `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3694321`. 2020.

[27] William H McNeill. “How the potato changed the world’s history”. In: *Social Research* 66.1 (1999), pp. 67–83.

[28] Martin I Meltzer, Nancy J Cox, and Keiji Fukuda. “The economic impact of pandemic influenza in the United States: priorities for intervention.” In: *Emerging infectious diseases* 5.5 (1999), p. 659.

[29] Nathan Nunn and Diego Puga. “Ruggedness: The blessing of bad geography in Africa”. In: *Review of Economics and Statistics* 94.1 (2012), pp. 20–36.

[30] Nathan Nunn and Nancy Qian. “The potato’s contribution to population and urbanization: evidence from a historical experiment”. In: *The Quarterly Journal of Economics* 126.2 (2011), pp. 593–650.

[31] Alan L Olmstead and Paul W Rhode. “Reshaping the landscape: the impact and diffusion of the tractor in American agriculture, 1910-1960”. In: *Journal of Economic history* 61 (2001), pp. 663–698.

[32] Marco Percoco. “The Fight Against Disease: Malaria and Economic Development in Italian Regions”. In: *Economic Geography* 89.2 (2013), pp. 105–125.

[33] Vincent Robert et al. “Malaria transmission in urban sub-Saharan Africa”. In: *The American journal of tropical medicine and hygiene* 68.2 (2003), pp. 169–176.

[34] Jeffrey Sachs and Pia Malaney. “The economic and social burden of malaria”. In: *Nature* 415.6872 (2002), pp. 680–685.

[35] Redcliffe N Salaman, William Glynn Burton, et al. *The history and social influence of the potato*. Cambridge University Press, 1985.

[36] Andrew J Tatem et al. “Human population, urban settlement patterns and their impact on Plasmodium falciparum malaria endemicity”. In: *Malaria Journal* 7.1 (2008), pp. 1–17.

[37] Fleur Wouterse and Ousmane Badiane. "The role of health, experience, and educational attainment in agricultural production: Evidence from smallholders in Burkina Faso". In: *Agricultural Economics* 0.0 (). DOI: 10.1111/agec.12500. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/agec.12500. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/agec.12500.

Supplementary Information for: Mosquitoes and Potatoes: How Global Health Crises Impede Development

Maurizio Malpede[a,1], Giacomo Falchetta[b,c], and Soheil Shayegh[d]

[a]GREEN, Bocconi University, Milan, Italy
[b]Fondazione Eni Enrico Mattei, Milan, Italy
[c]Free University of Bozen-Bolzano, Bolzano, Italy
[d]RFF-CMCC European Institute for Economics and the Environment, 20144 Milan, Italy

May 5, 2021

This PDF file includes:
Supplementary Info Text
Figures S1 to S8
Tables S1 to S8
SI References

A. Technological Innovation, Public Health, and Development

A.1. Agricultural Innovation and Economic Development

A broad stream of literature argued that increases in agricultural productivity are an essential condition for economic development. In particular, the experience of England during the industrial revolution has been the object of the first studies [22, 21, 23].

More recent papers analyze the mechanisms underlying the productivity growth in agriculture and the consequent socio-economic growth. [15] investigated the role of fertilizer in boosting yields and GDP per capita as a result. [4]

Finally, [3] use the introduction of genetically engineered soy in Brazil as an agricultural productivity shock to provide empirical evidence on the structural transformation effects of agrarian technical change. They show that when the agricultural innovation is strongly labor-saving, as in the case of genetically engineered soy, it can foster industrialization.

A.2 Diffusion of Potatoes in the Old World

Particular emphasis is given to the role of crop diffusion in the Old World. [20] have investigated the impact of the diffusion of potatoes from the New World, where they were first discovered and cultivated in the Old World. Finally, [10] focuses on the effects of the introduction of sweet potatoes on peasant uprisings in China. He finds that its diffusion significantly lowered the number of peasants uprisings due to periods of droughts. In fact, if before the introduction of sweet potatoes in China there was a significant negative relationship between precipitations (which represent a positive income shock for peasants) and peasant uprisings, after the diffusion of the sweet potato, which was more resistant to droughts, the uprisings significantly declined in areas more suitable to the cultivation of the newly introduced staple crop.

[8] and [19] focused on other New World crops, which were also introduced along with potatoes to the Old World after the discovery of the Americas. These include maize, cassava, tomatoes, chili and bell peppers, cacao, sunflowers, and sweet potato. Of these, the two crops that became high-caloric staples in the Old World were maize and cassava. Maize is unable to rival potatoes in terms of both nutrients and calories. It produces significantly fewer calories per acre of land. In comparison, cassava has an essential deficiency in terms of protein intake. Therefore, making the latter less nutritional than potatoes for human beings. It is necessary to highlight that for sweet potatoes, which show similar nutritional properties than those of the potato, there is historical evidence of their introduction, and consequently, their diffusion in the Old World already in 1000 [20].

A.3 Epidemic Diseases and Economic Development: The example of Malaria

Macroeconomic effects of health conditions have been investigated in the literature [1]. This study is the first to consider an exogenous measure of life expectancy. Using an IV strategy, the authors proxy life expectancy with mortality rates from the most common and deadly diseases during the 1940-1980 period. They conclude that although the reduction in mortality rates has increased the population levels, the same does not hold for income per capita. In particular, the increase of health conditions does have a positive impact on the population levels in the short run. However, since the land is fixed in the short run, the effects on income per capita are slightly adverse.

Malaria has been one of the primary infectious diseases throughout the 16-20th century. It was not until the 20^{th} century when technological advancements and the discovery of new drugs and modern chemical components made it possible to prevent malaria transmission effectively. An example of the effectiveness of early control policies was the construction of the Panama Canal in the 1910s. Malaria was a major cause of death and illness among workers in the area. According to the Center for Disease Control Prevention (CDC), in 1906, over 26,000 employees were working on the Canal. Of these, over 21,000 (i.e., more than 80%) were hospitalized for malaria at some time during their work. By 1912, there were more than 50,000 workers, and the number of hospitalized workers had decreased to approximately 5,600 (i.e., 11%).

B. Data and Summary Statistics

The primary data sources used in the empirical framework are summarised in Table 1.

B.1 Potato Suitability Index and Malaria Stability Index

The MSI captures the potential stability of malaria transmission based on regionally dominant vector mosquitoes, temperature, and precipitation data set [11].

The PSI captures the potential suitability for cultivating potatoes based on regional weather and land conditions. [5]. Additional sources include suitability indexes for other crops, used as control variables and estimated with the same methodology of the PSI, and a measure of terrain ruggedness [18]. Specifically, the construction of the FAO's GAEZ database occurred in two stages. The FAO first collected information on the characteristics of 154 different crops. These data were used to determine what environmental conditions are required for the cultivation of each crop. The FAO then compiled data on the physical environment of 2.2 million grid cells, spanning the entire globe. Each cell is 0.5 degrees by 0.5 degrees, approximately fifty-six kilometers by fifty-six kilometers (measured at the equator). The primary characteristics used are climatic and are taken from a global climatic database that has been compiled by the Climate Research Unit at the University of East Anglia. In total, nine variables from the global climatic database are used by the FAO: precipitation, frequency of wet days, mean temperature, diurnal (i.e., daily) temperature range, vapor pressure, cloud cover, sunshine, ground-frost frequency, and wind speed. The second set of characteristics are land characteristics and are taken from the FAO's Digital Soil Map of the World. The last characteristic is the slope of soils from the GTOPO30 Database, developed at the U.S. Geological Survey (USGS) EROS Data Center. Combining the information on the constraints for the cultivation of each crop with the data on the physical environment of each grid cell, the FAO calculated an estimate of the potential yield of each crop in each grid cell, given an assumed level of crop management and input use.

The FAO also constructed a discrete version of the database that reports the proportion of each grid's land classified under five mutually exclusive categories describing how suitable the environment is for growing each crop. The categories are based on the calculated percentage of the maximum yield that can be attained in each grid cell. The five categories and their corresponding yields are: (i) very suitable land (0.8 - 1), (ii) suitable land (0.6 - 0.8), (iii) moderately suitable land (0.4 - 0.6),(iv) marginally suitable land(0.2 - 0.4),and (v) unsuitable land (0 - 0.3). To approximate historical conditions as closely as possible, we use variables constructed under the assumption that cultivation occurs under rain-fed conditions and medium input intensity.

Figure 2 plot the spatial heterogeneity in the two main variables of interest, namely the potato suitability index (PSI) and the malaria stability index (MSI), respectively. We show a slight correlation between suitability for potato cultivation and endemicity of malaria (i.e., in the order of 0.48). In particular, some areas extremely suitable for potato[1] shares high endemicity of malaria as well.

Non-gridded control variables, such as the country-level extent of cropland and tropical area, the elevation, the distance from the equator, and the nearest coast, are drawn from [20] 's original replication data. For the country-level regressions, zonal statistics for the mean of spatially-explicit datasets, i.e., malaria Stability index and potato suitability, as well as for the suitability of an array of control crops, namely cassava, wheat, maize, and barley (again, with medium-input, rainfed, historical average parameters), are produced for each grid cell.

B.2 Grid-level Data on Population and Urbanization

Grid-level data on population and urbanization rates for the years between 1100 and 1980 are retrieved from the HYDE database[2]. HYDE is an internally consistent combination of updated historical population (gridded) estimates and land use for the past 12,000 years. Categories include cropland, with a new distinction into irrigated and rain-fed crops (other than rice) and irrigated and rain-fed rice. Also, grazing lands are provided, divided into more intensively used pasture, converted rangeland, and non-converted natural (less intensively used) rangeland. The total population is represented by maps of the urban, rural population, population density, and built-up area. HYDE was developed by the PBL Netherlands Environmental Assessment Agency and is used by the World Climate Research Program Coupled Model Intercomparison Project (CMIP6). The process of estimating population and urbanization at very disaggregated units is explained in the documentation of the research project [12] and involves a series of steps outlined as follows: First, estimated population totals at the state or province-level by using a variety of historical sources. The basis for the state-level population data is the United Nations World Populations Prospects (2008 Revision) for the 1950–2017 period. The pre-1950

[1] Areas highly suitable for potato cultivation have a potato suitability index > 0.8 as defined by the FAO

[2] Population and Urbanization data are available at time steps of 100 years until 1700, 50 years from 1700 to 1900, and 10 years from 1900 to 1980

historical estimates were largely taken from the first Atlas of World Population History [16], [13], and [14]. [12] supplemented these sources with the sub-national population numbers of Populstat (Lahmeyer, personal communication, 2004) and many other country-specific sources. The country-specific sources of population estimates are outlined in the supplementary information of their methodology. Second, they use interpolation techniques to fill in gaps between population sources.

We use version 3.2 of the HYDE database. This is a combination of gridded historical population and land use estimates. [12] used historical records to model population at the province level and land use areas at the national level through time; then used algorithms to spatially distribute the total population and land use areas to 5 arc minute pixels (about 10 kilometers at the equator). We extract data on population and urbanization from the HYDE dataset into a global regular grid shapefile with a resolution of 0.5° latitude and 0.5° longitude, using GIS software. A spatial join is performed between the regular grid and the GADM shapefile to create a country attribute in the regular grid shapefile matching the underlying country name.

A possible relevant issue arising with the use of grid-level estimated historical population data from HYDE is the measurement error of the latter. For instance, estimated historical population and urbanization data for country A might be very highly inaccurate, thus having adverse effects on the performance of the indicators [17, 6]. We address this concern by including country-specific time trends in the grid-level empirical strategy. This allows for a comparison of each grid within the same country. Second, both the potato suitability and malaria ecology indexes are void of this problem, being them time-invariant and computed using weather and soil variables.

C. The Impact of Potato Cultivation on Socio-Economic Outcomes in Potato-Suitable Areas

First, we evaluate the impact of potato suitability on the Old World at a grid level. The empirical model that we use is similar to that of [20] expect that we use more disaggregated spatial data.

$$\mathrm{y}_{i,c,t} = \alpha + \beta_1 (\mathrm{Post})_t \times (\mathrm{Potato})_i + \gamma\, \mathbf{X}_i' + \delta_t + \sigma_{2,ctrend} + \epsilon_{i,c,t} \tag{1}$$

Equation 1 controls for the same fixed effects used in [20]. Moreover, we add grid-specific agricultural characteristics (i.e. the suitability for cultivating other types of crop). Table 2 shows the results of this model.

D. The Impact of Potato Cultivation on Socio-Economic Outcomes, Differentiated by Malaria Endemicity

Using equation 2, we compare population and urbanization levels between grids measuring 0.5° altitude by 0.5° longitude. This specification not only allows for more robust results but also for understanding if the cultivation of potato after its diffusion in the Old World had a local impact on population and urbanization levels as well, and if the presence of weather conditions suitable for the transmission of malaria counteracted the positive effects of potato.

$$y_{i,c,t} = \beta \ \mathrm{PSI}_i \times \mathrm{MSI}_i \cdot I_t^{Post} + \sum_{t=1000}^{1900} \Omega \mathbf{X}'_{i,c} I_t + \sum_{t=1000}^{1900} \delta I_t + \sum_p \gamma_p I_t^p + \epsilon_{i,c,t} \tag{2}$$

The variable PSI_i x MSI_i is the combination of the time invariant suitability index for potato cultivation (PSI) and its degree of stability of malaria (MSI) for each grid i. Our variable of interest PSI_i x MSI_i shows the impact on population and urbanization levels of the suitability of potato in non malaria-endemic grids versus malaria-endemic grids. If malaria effectively counteracted the positive effects of potato on the development of Old World countries, we would expect to find a positive effect of the PSI in non-malaria-endemic grids while disappearing in grids with high MSI.

Again, the variable I_t^{Post} is a post-adoption dummy variable which takes value 1 for years after the exogenous spreading of potato in the Old World (i.e., 1700, 1750, 1800, 1850, and 1900) while value 0 for years ex-ante the exogenous introduction of the potato in the Old World (i.e., 1000, 1100, 1200, 1300, 1400, 1500 and 1600). Equation 2 also includes grid fixed effects $\sum_p \gamma_p I_t^p$, where p indicates the set of Old World grids; time period fixed effects $\sum_{t=1000}^{1900} \delta I_t$ and country-specific fixed effects which allow for a comparison of each grid within the same country. As before $\mathbf{X}_{i,s}$ represents vectors of time-invariant grid-specific controls included in the regression. As grid-level controls, we use a set of relevant geographical and historical characteristics which might have affected population and urbanization: i.e., total cropland area per grid, ruggedness level, and total areas of other crops, i.e., maize, silage, sweet potato, and cassava. Table 3 shows the results of this model.

E. Flexible Estimates

Empirical Framework

The baseline equation (1) examines the average effect of the potato on population and urbanization after its introduction on more or less malaria-endemic areas. The primary assumption of this strategy is that the cutoff date used for the introduction of potato cultivation in the

Old World is correct. The historical evidence presented in [20] suggests that the adoption of the potato began in a few locations in the late seventeenth century and spread significantly by the early eighteenth century. Given the evidence, reasonable cutoff dates range between 1700 and 1750, and therefore 1750 is the first postadoption time period. Before taking this cutoff as given, we use several strategies to examine whether the patterns in the data are consistent with this assumption. The first strategy is to estimate a fully flexible estimating equation that takes the following form:

$$\mathrm{y}_{i,c,t} = \alpha + \sum_{j=1100}^{1900} \beta_j (\mathrm{PotMal})_i \times \mathrm{I}_t^j + \gamma\, \mathbf{X}_i^{'} + \omega_i + \delta_t + \sigma_{2,ctrend} + \epsilon_{i,c,t} \tag{3}$$

where all variables are defined as in equation (1). The only difference from equation (1) is that in equation(3), rather than interacting PotMal_i with a postadoption indicator variable, we interact the suitability measure with each of the time-period fixed effects. With this procedure we expect the coefficients to be close to zero in magnitude for the years before the adoption of potato (i.e., before 1700), and positive for the years after the adoption. We also expect the coefficients of interest to be positive in magnitude only for non-malarious areas while no or little effects should be found in more malarious areas.

F. Different Definitions of Potato Suitability

The baseline equation (1) examines the average effect of the potato on population and urbanization after its introduction on more or less malaria-endemic areas. The main assumption of this strategy is that we define as potato suitable those areas with a Potato Suitability Index ¿ 0.4 [20]. This section shows results obtained using the same specification of equation (1) with different definitions of potato suitable areas.

Results are shown in Figures 3, 4, 5, 6 which show the effects of the estimation of equation (1) using cutoff values of potato suitability levels of 0.5, 0.3, 0.2 and 0.1, respectively. Results confirm the decreasing impact of the cultivation of the potato on urbanization and population of potato-suitable locations relative to non-potato suitable ones after 1700 as the level of malaria endemicity grows.

G. Country-Level Analysis

For country-level modeling, data is extracted with the following procedure. The potato suitability index layer is downloaded from the FAO Global Agro-Ecological Zones database [5] with medium-input, rainfed, historical average parameters. Also, the malaria stability index is retrieved. The two files are imported in Google Earth Engine [7], where the area of (i) zones with potato suitability greater than 40% (as in [20]) and (ii) zones with potato suitability greater than 40% and a malaria Stability index lower than 1, 0.1, 0.06, 0.05, 0.01, and 0 is calculated. The GADM (Global Administrative Unit Layer, [9]) is used as the reference shapefile for national borders.

Estimation Strategy

$$y_{i,t} = \beta_1 \; \textit{ln Potato Area}_i \cdot I_t^{Post} + \beta_2 \; \textit{ln Potato NoMal Area}_i \cdot I_t^{Post} + \\ + \beta_3 \; \textit{ln Potato Mal Area}_i \cdot I_t^{Post} + \sum_{j=1000}^{1900} \Omega_j \mathbf{X'}_i I_t^j + \sum_{j=1000}^{1900} \delta I_t^j + \sum_{c=1}^{130} \gamma_c I_i^c + \epsilon_{i,t} \tag{4}$$

where index i represents each Old World country. A total of 130 countries were considered in the analysis, indexed with c. Index t represents time periods which are part of the analysis (i.e. 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1750, 1800, 1850 and 1900). Years from 1000 to 1600 are part of the pre-potato introduction period as in [20]. Since potato became widely diffused in the Old World soon after 1700, 1700 is considered the start of the post-intervention period, which ranges from 1700 until 1900[3]. The dependent variable is represented as $y_{i,t}$, which is the natural log of one of the two proxies of historical, socio-economic development, i.e., population and city population share for each Old World country i at time t.

The variable $\textit{ln Potato Area}_i$ is the time-invariant area suitable for potato cultivation for each country i. As in [20] β_1 shows the estimated impact of potato-suitable areas on population and city population share after its adoption in the Old World. A positive and statistically significant value of β_1 is interpreted as an additional increase in population and urbanization of Old World countries with higher potato suitable land between 1700 and 1900 compared to countries with less amount of potato suitable land.

Besides computing the impact of the amount of the area suitable for potato cultivation for each country to population and urbanization, our procedure allows us to disentangle the impact of the amount of the country area suitable for the cultivation of potato *but not malaria endemic* with the country are suitable for potato *and malaria endemic.* To this regard, the variable $\textit{ln Potato NoMal Area}_i$ is the time-invariant area suitable for potato cultivation but not malaria-endemic for each country i. The coefficient of our interest is β_2, which is the estimated impact of the adoption of potato in non-malaria-endemic and potato suitable land rather than on mere potato suitable land. On the other hand, the variable $\textit{ln Potato Mal Area}_i$ is the time-invariant area suitable for potato cultivation for each country i and at the same time also suitable for stable transmission of malaria. We expect β_2 to be positive and greater in magnitude then

[3]In the original paper, [20] further test 1700 as the only valid cut off date by performing the empirical analysis using alternative cut-offs.

respectively β_1 and β_3. In particular we would expect that $\beta_2 > \beta_1 > \beta_3$. For concreteness, such a situation would mean that the cultivation of the potato after its introduction in 1700 had a greater impact on population and urbanization on countries that, besides having a geographic environment more suitable for growing potatoes, were not malaria-endemic as well. Moreover, a coefficient β_3 being the smallest in magnitude would mean that the adoption of potatoes had a small impact on those countries that had a large amount of potato-suitable land, which at the same time was malaria-endemic.

Variable I_t^{Post} is a post treatment dummy variable which takes value 1 for years after 1700 (i.e. 1700, 1750, 1800, 1850 and 1900) while value 0 for years before the exogenous introduction of potato in the Old World (i.e. 1000, 1100, 1200, 1300, 1400, 1500 and 1600). This specification also includes country fixed effects $\sum_{c=1}^{130} \gamma_c I_t^c$, where c indicates the set of Old World countries constituted by a total of 130 countries and time period fixed effects $\sum_{t=1000}^{1900} \delta I_t$.

As argued in [20] European countries have, on average, experienced higher growth in population and urbanization between 1700 and 1900 for reasons related to socio-economic and technological progress, which are beyond the cultivation of potato. However, European countries were also naturally more suitable for growing potatoes. It would be, therefore, necessary to estimate the effects of introducing the cultivation of potatoes within continent variation only[4]. Hence, to allow for a comparison of each country within the same continent, we add to equation 4 continent fixed effects interacted with time-period fixed effects. The introduction of *Continent x Year FE* to equation 4 changes the interpretation of our coefficients of interest β_1, β_2 and β_3, which are now identified from within-continent variation only.
$\sum_{j=1000}^{1900} \Omega_j \mathbf{X'}_i I_t^j$ represents time-invariant country-specific characteristics interacted with time period fixed effects to take into account other country-specific characteristics aside from the cultivation of potato that might have affected population and urbanization growth between 1700 and 1900. As country-level controls, we use the set of relevant geographical and historical country-specific characteristics which might have affected population and urbanization between 1000 and 1900 used in [20]: i.e., country elevation, total cropland area, ruggedness, tropical area, distance from the equator and the nearest coast, an indicator variable taking value one if a country is an exporter of potatoes along with total areas of other crops, i.e., maize, silage, sweet potato, and cassava.

Results

Tables 4 and 5 show the estimates from our main estimating equation 4 respectively without the inclusion of additional controls and with the inclusion of all country and geographical specific

[4] In other words we do not want to compare a European country such as Italy with an African country, but comparing a European country with another European country makes the estimation strategy more robust and allows to take into account all the other factors that may have caused European divergence

control variables. All columns include country and time-period fixed effects along with all control variables used in the analysis, i.e., controls for land suitable for Old World staple crops interacted with time-period fixed effects and controls for ruggedness, elevation, and tropics, each interacted with the time-period fixed effects

Results are shown in Tables 4, and 5 confirm our prediction that suggested that overall the introduction of potatoes increased total population and urbanization. However, weather conditions particularly favorable to the transmission of malaria counteracted the positive benefits of potatoes.

This is illustrated by the estimated coefficient of the potato suitability interaction term, $ln\ Potato\ Area_i \cdot I_t^{Post}$, which displays the average increase in population and urbanization levels respectively arising from the wide diffusion of potato in the Old World after 1700. According to the estimates in columns (1) and (4), a 1 percent increase in the amount of land suitable for the cultivation of the potato increased the population by 0.014 percent on average and the urban population by 0.002 percentage points after 1700. This result is with no surprise very similar to what shown in [20][5]. However, the following two rows of tables S1 and S2 show that the positive impact of the introduction of the potato in the Old World was clustered in non-malaria endemic areas. In contrast, the diffusion of the potato did not have any positive effect on malaria-endemic locations.

The results are shown in Tables S1 and S2 and emphasize the beneficial effects of the introduction of potatoes on the increase of population and urbanization of countries with greater land naturally suitable to growing potatoes. This first strategy we employ, however, does not allow us to estimate the eventual impact of higher suitability for growing potatoes on population and urbanization at a very disaggregated level. In the next section, we employ a second econometric strategy that permits us to estimate not only the effects of higher suitability for potato cultivation on population and urbanization after 1700 *locally* but also if the relative level of stability of malaria transmission counteracted the positive effects of cultivating potatoes.

H. The Effects of Malaria after its Eradication (Post 1900)

Here we address the concern that malarious weather conditions might be correlated with other unobservable factors. If this is the case, we will overestimate the negative impact of malaria in the diffusion of potato cultivation in the Old World. One possible solution to mitigate such a concern would be to consider the malaria eradication period as a cut-off date. In the early 1900s, technological and medical advancements made it possible to eradicate malaria [2]. Therefore, if malaria was the only factor absorbing the positive effects of the cultivation of the potato in

[5]The slight differences are due to the fact that [20] use a version of potato suitability index dated 2002 while the version used to perform the following analysis is the most recent available, dated 2011.

malarious areas, then the adverse effects in malarious grids should disappear after its eradication. In other words, the positive impacts of potato diffusion in malarious and not malarious areas should have no difference after eradicating malaria[6].

$$y_{i,c,t} = \beta \text{ PSI}_i \text{ x MSI}_i \cdot I_t^{Post} + \sum_{t=1000}^{1980} \Omega \mathbf{X'}_{i,c} I_t + \sum_{t=1000}^{1980} \delta I_t + \sum_p \gamma_p I_t^p + \sum_{t=1000}^{1980} \sum_q \omega_q I_t^q + \epsilon_{i,c,t} \tag{5}$$

The specification expressed in equation 5 is identical to equation 2 except that in here time periods, indexed with t, range from 1000 to 1980[7]. Years from 1000 to 1900 are part of the pre-malaria eradication period, while years from 1910 to 1980 are part of the post-malaria eradication period. Therefore, in equation 5, variable I_t^{Post} is a post-treatment dummy variable which takes value 1 for years after 1900 (i.e. from 1910 to 1980) while value 0 for years before the exogenous intervention against malaria (i.e. 1000, 1100, 1200, 1300, 1400, 1500 and 1600). This specification also includes grid fixed effects $\sum_p \gamma_p I_t^p$, where p indicates the set of Old World countries; time period fixed effects $\sum_{t=1000}^{1900} \delta I_t$ and country-specific fixed effects which allow for a comparison of each grid within the same country. As before $\mathbf{X}_{i,s}$ represents vectors of time-invariant grid-specific controls included in the regression. As grid-level controls, we use a set of relevant geographical and historical characteristics which might have affected population and urbanization: i.e., total cropland area per grid, ruggedness level, and total areas of other crops, i.e., maize, silage, sweet potato, and cassava.

Results

Tables 6, 7, and 8 the effects of potato diffusion in non-endemic and highly endemic areas after 1900. Results show that the cultivation of the staple crops resulted in positive effects on the population and urbanization of potato-suitable locations. Moreover, there is no statistically significant difference between more or less malaria-endemic areas. In other words, all potato-suitable locations enjoyed higher population and urbanization growth levels regardless of their climate suitability for the transmission of vector-borne diseases.

[6] We also check that the difference in the positive effects of potato in more or less malarious areas decreases in the early 1900s until disappearing at around 1950s

[7] Time periods are respectively 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1750, 1800, 1850, 1900, 1910, 1920, 1930, 1940, 1950, 1960, 1970 and 1980

Figures

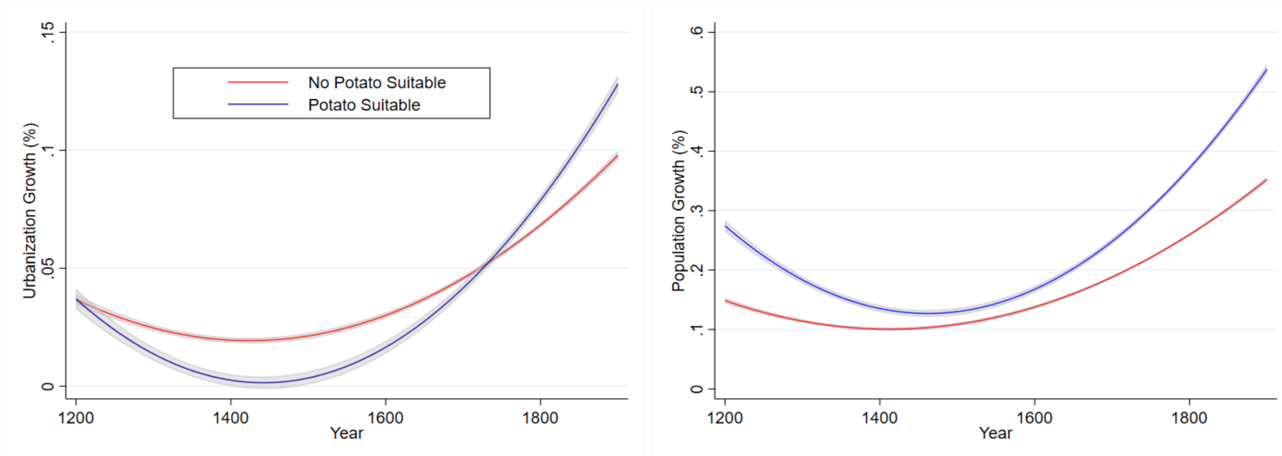

Figure 1: Historical urbanization (left) and population (right) growth rates in potato-suitable location vs non potato-suitable ones from 1200 to 1900.

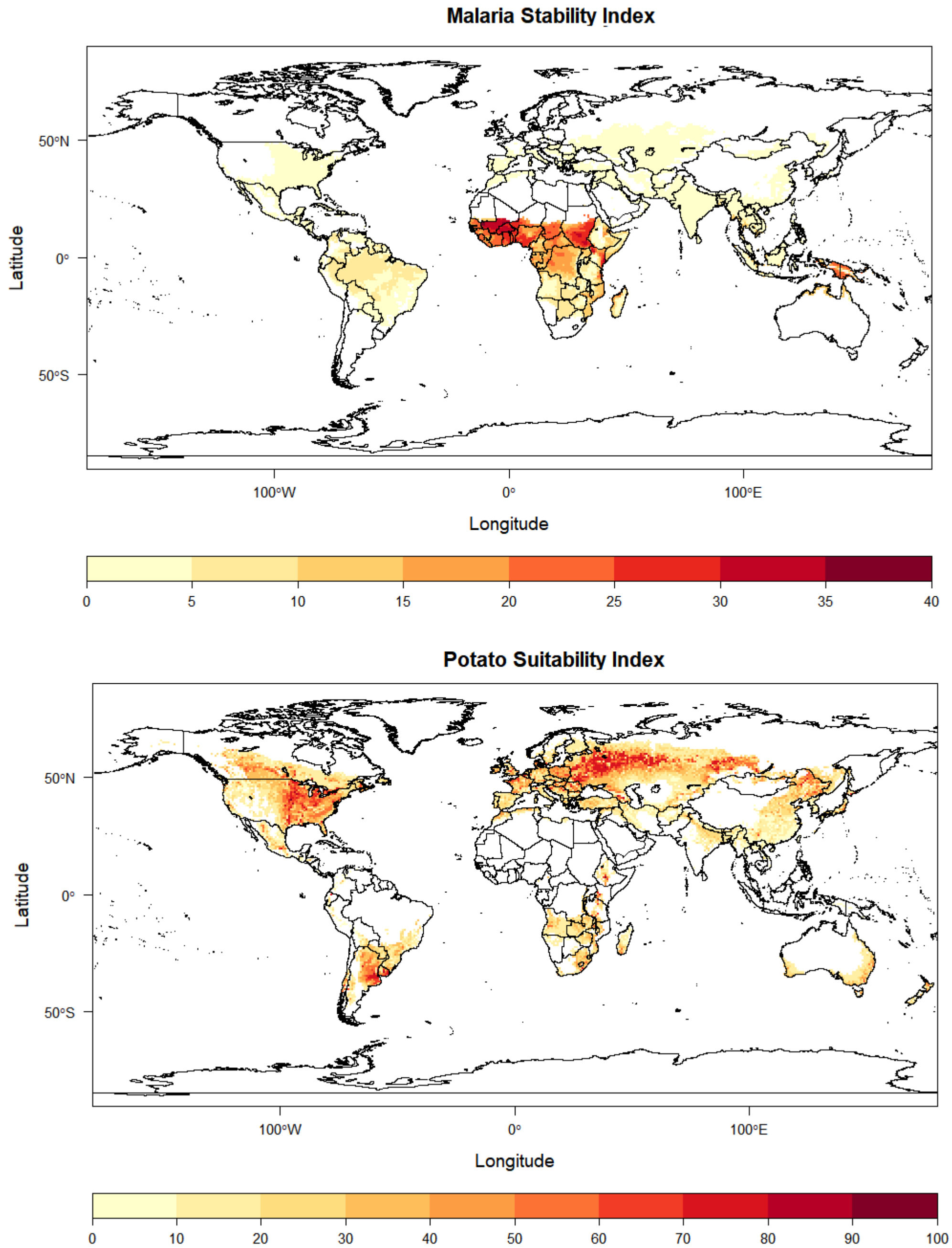

Figure 2: Maps of the spatial distribution of the MSI and PSI intersected with country borders.

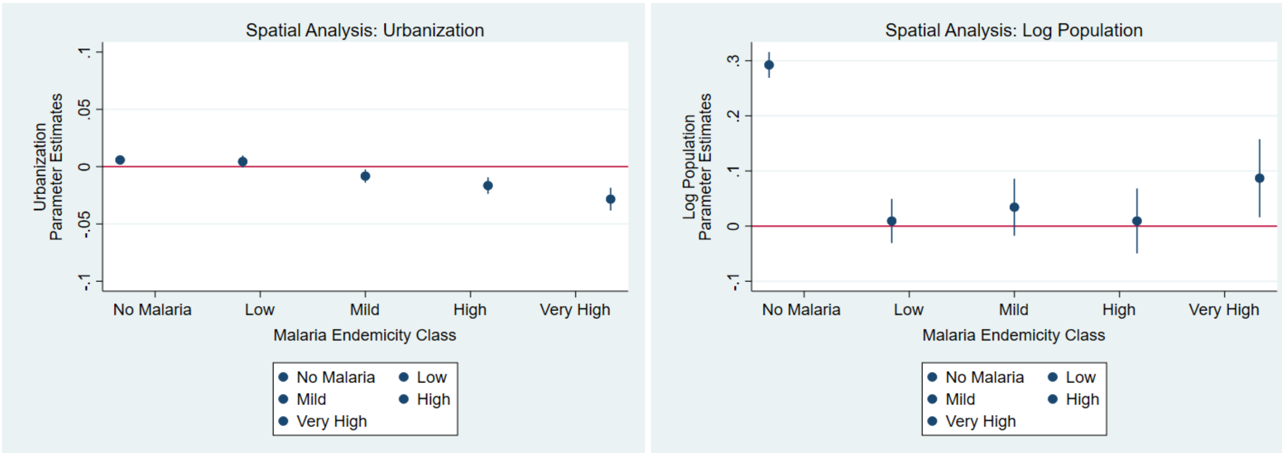

Figure 3: Effects of the introduction of potato in the Old World on urbanization (left) and on population (right) for each category of malaria endemicity. Potato suitable areas defined as $PSI > 0.5$. 95% confidence intervals.

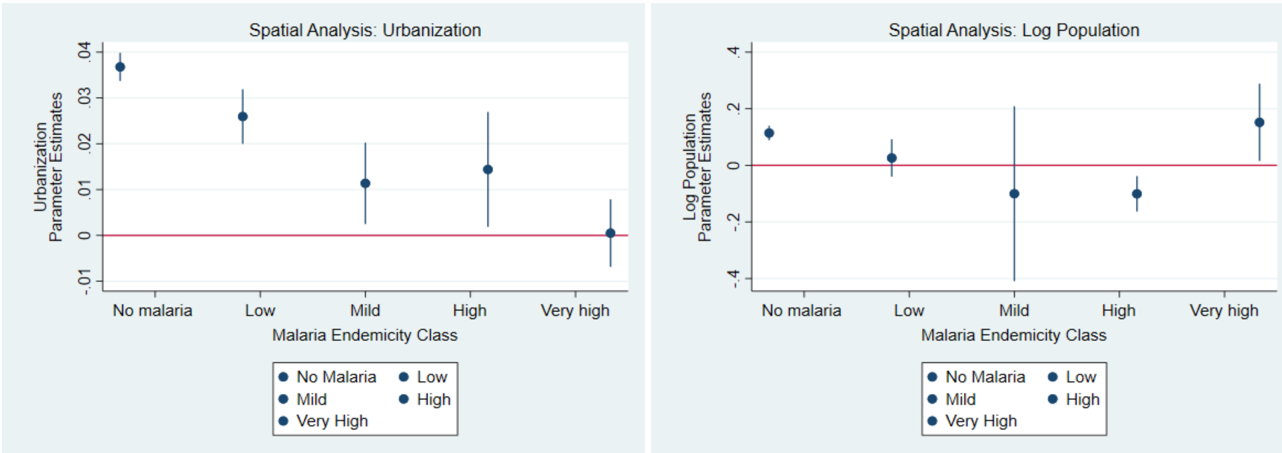

Figure 4: Effects of the introduction of potato in the Old World on urbanization (left) and on population (right) for each category of malaria endemicity. Potato suitable areas defined as $PSI > 0.3$. 95% confidence intervals.

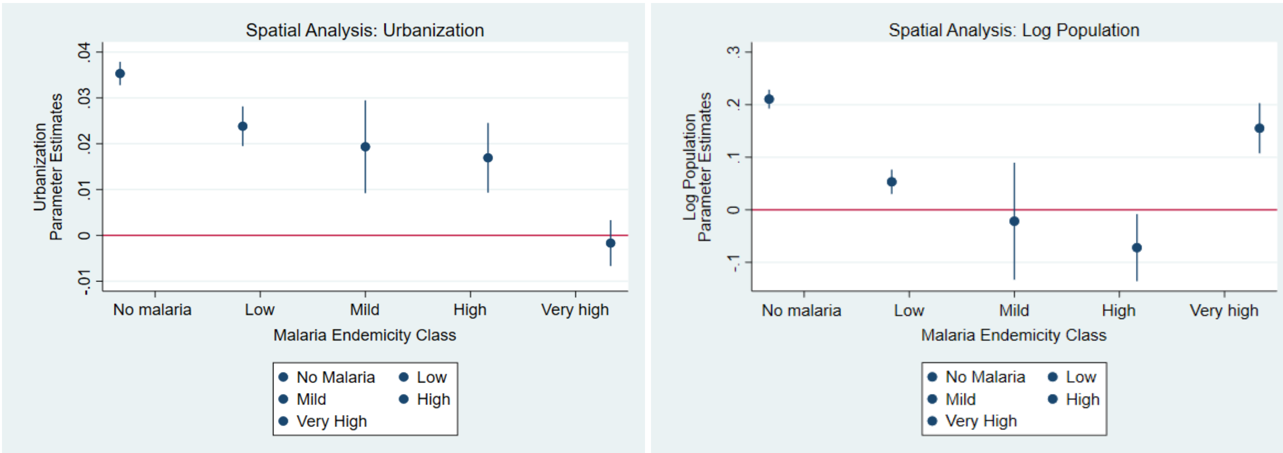

Figure 5: Effects of the introduction of potato in the Old World on urbanization (left) and on population (right) for each category of malaria endemicity. Potato suitable areas defined as $PSI > 0.2$. 95% confidence intervals

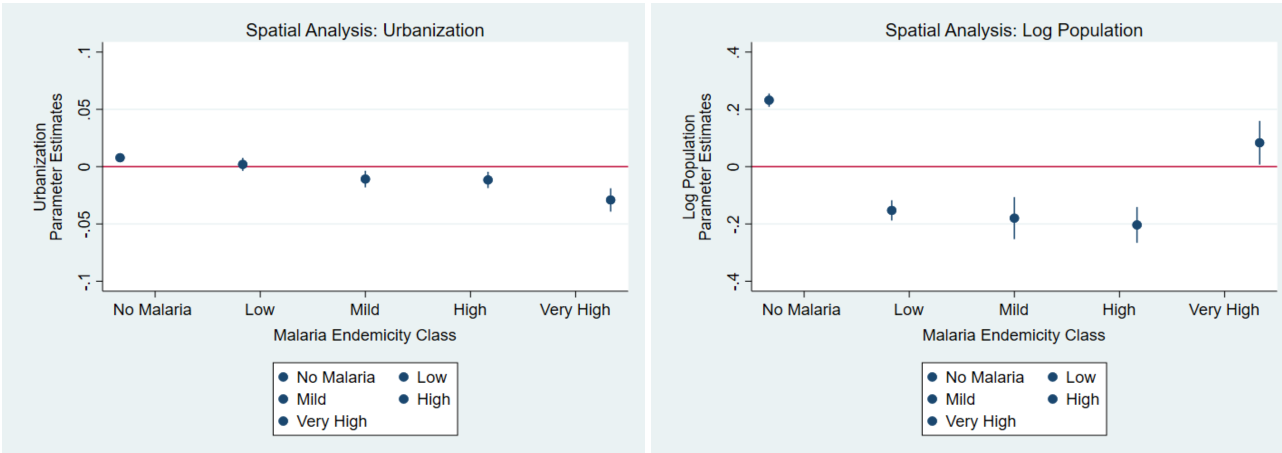

Figure 6: Effects of the introduction of potato in the Old World on urbanization (left) and on population (right) for each category of malaria endemicity. Potato suitable areas defined as $PSI > 0.1$. 95% confidence intervals.

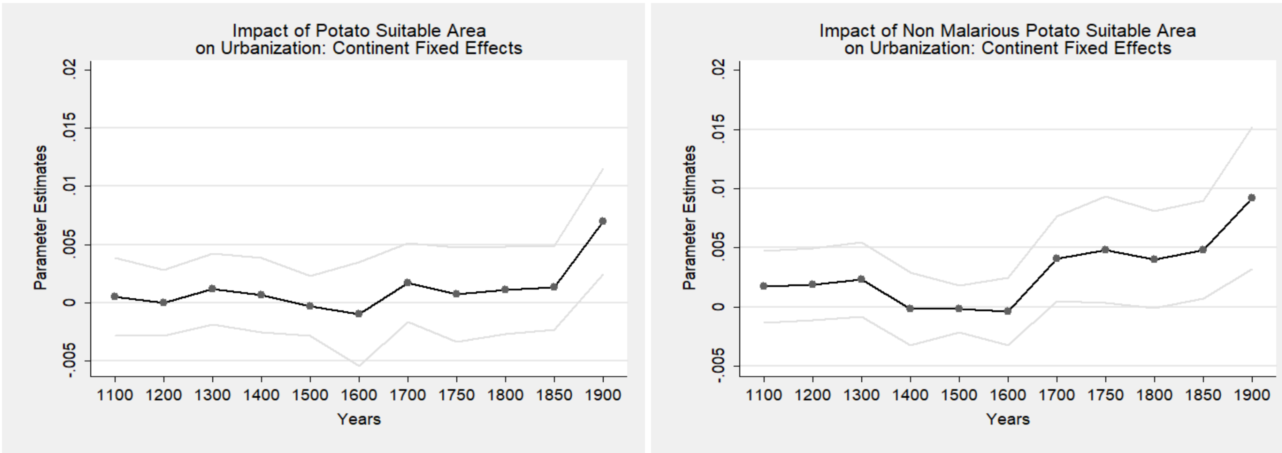

Figure 7: Effects of the introduction of the potato on urbanization of Old World countries. Country Analysis.

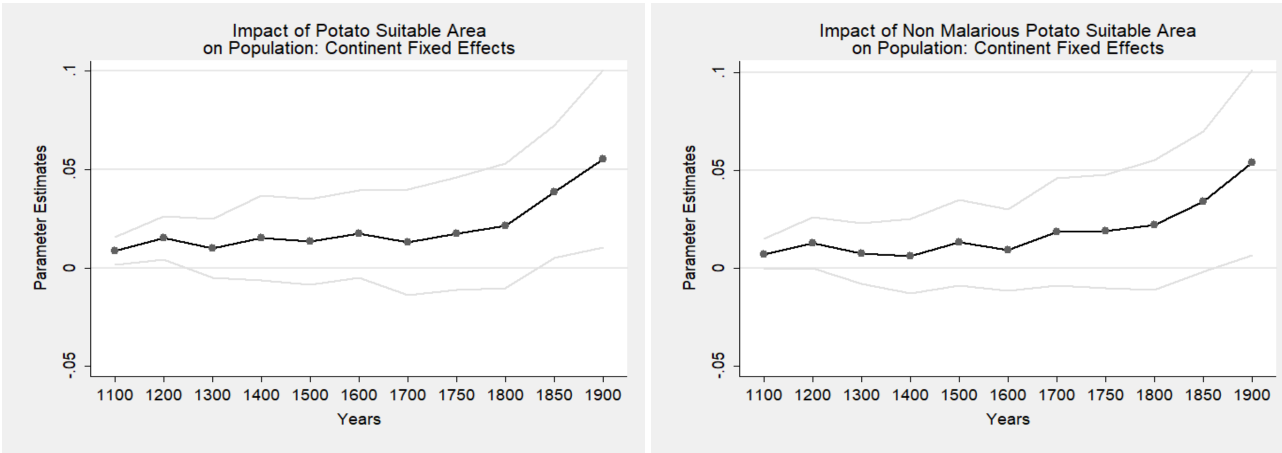

Figure 8: Effects of the introduction of the potato on population of Old World countries. Country Analysis.

Tables

Table 1: Table of data inputs for the econometric identification

Dataset	**Source**	**Time resolution**	**Spatial resolution**
Malaria stability index	[11]	Time-invariant	0.5°
Potato suitability index	[5]	Time-invariant	0.08°
Crop suitability indexes	[5]	Time-invariant	0.08°
Historical population and urbanisation level	[12]	100/50/10 years	
Terrain ruggedness index	[18]		
Country shapefiles	[9]	-	-

Table 2: The impact of potato suitability on urbanization and population: Old World, Pre Malaria Eradication, Gridded Analysis

	Urb b/se	Urb b/se	Urb b/se	Pop b/se	Pop b/se	Pop b/se
Potato Suitable x Post	0.046***	0.034***	0.028***	2.732***	0.680***	0.118***
	(0.007)	(0.007)	(0.002)	(0.853)	(0.249)	(0.010)
Barley	No	Yes	Yes	No	Yes	Yes
Maize	No	Yes	Yes	No	Yes	Yes
Cassava	No	Yes	Yes	No	Yes	Yes
Wheat	No	Yes	Yes	No	Yes	Yes
Tropical	No	Yes	Yes	No	Yes	Yes
Ruggedness	No	Yes	Yes	No	Yes	Yes
Elevation	No	Yes	Yes	No	Yes	Yes
Dist. Coast	No	Yes	Yes	No	Yes	Yes
Observations	522197	522197	522197	457356	457356	457356

Table 3: The impact of potato suitability: Increasing malarious conditions

	Population	Urbanization
	b/se	b/se
Potato Suitable x MSI Post	-0.012***	-0.001**
	(0.002)	(0.000)
Barley	Yes	Yes
Maize	Yes	Yes
Cassava	Yes	Yes
Wheat	Yes	Yes
Tropical	Yes	Yes
Ruggedness	Yes	Yes
Elevation	Yes	Yes
Dist. Coast	Yes	Yes
Observations	40847	40847

Table 4: The impact of non malarious potato suitable area on population and urbanization: Continent FE estimates

	Pop b/se	Pop b/se	Pop b/se	Urb b/se	Urb b/se	Urb b/se
ln Potato Area x Post	0.014			0.002**		
	(0.011)			(0.001)		
ln Non Malarious Potato Area x Post		0.025*			0.002*	
		(0.015)			(0.001)	
ln Malarious Potato Area x Post			-0.008			0.001
			(0.015)			(0.001)
Old World	Yes	Yes	Yes	Yes	Yes	Yes
Elevation	Yes	Yes	Yes	Yes	Yes	Yes
Tropical	Yes	Yes	Yes	Yes	Yes	Yes
Rugged	Yes	Yes	Yes	Yes	Yes	Yes
Continent X Year	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1528	1528	1528	1528	1528	1528
Adjusted R^2	0.990	0.990	0.990	0.400	0.399	0.397

Table 5: The impact of non malarious potato suitable area on population and urbanization: Continent FE estimates with additional controls

	Pop	Pop	Pop	Urb	Urb	Urb
	b/se	b/se	b/se	b/se	b/se	b/se
ln Potato Area x Post	0.010			0.004***		
	(0.013)			(0.001)		
ln Non Malarious Potato Area x Post		0.019			0.004***	
		(0.015)			(0.001)	
ln Malarious Potato Area x Post			-0.001			-0.000
			(0.017)			(0.001)
Old World	Yes	Yes	Yes	Yes	Yes	Yes
Elevation	Yes	Yes	Yes	Yes	Yes	Yes
Tropical	Yes	Yes	Yes	Yes	Yes	Yes
Rugged	Yes	Yes	Yes	Yes	Yes	Yes
Coast distance	Yes	Yes	Yes	Yes	Yes	Yes
Equator distance	Yes	Yes	Yes	Yes	Yes	Yes
Exporter	Yes	Yes	Yes	Yes	Yes	Yes
Continent X Year	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1528	1528	1528	1528	1528	1528
Adjusted R^2	0.991	0.991	0.991	0.456	0.456	0.447

Table 6: The impact of potato suitability on urbanization and population: Old World, Post Malaria Eradication, Gridded Analysis

	Urb b/se	Urb b/se	Urb b/se	Pop b/se	Pop b/se	Pop b/se
Potato x Post 1900	0.048***	0.089***	0.042***	0.589***	2.590***	0.652***
	(0.007)	(0.019)	(0.011)	(0.075)	(0.489)	(0.078)
Barley	No	Yes	Yes	No	Yes	Yes
Maize	No	Yes	Yes	No	Yes	Yes
Cassava	No	Yes	Yes	No	Yes	Yes
Wheat	No	Yes	Yes	No	Yes	Yes
Tropical	No	Yes	Yes	No	Yes	Yes
Ruggedness	No	Yes	Yes	No	Yes	Yes
Elevation	No	Yes	Yes	No	Yes	Yes
Dist. Coast	No	Yes	Yes	No	Yes	Yes
Observations	362490	362490	362490	362490	362490	362490

Table 7: The impact of potato suitability on urbanization and population in increasing malarious conditions

	$MSI = 0$	$0 < MSI < 1$	$1 < MSI < 5$	$5 < MSI < 10$	$MSI > 10$
	b/se	b/se	b/se	b/se	b/se
Potato x Post 1900	0.081***	0.027	0.090***	0.039**	0.045***
	(0.017)	(0.017)	(0.026)	(0.016)	(0.013)
Barley	Yes	Yes	Yes	Yes	Yes
Maize	Yes	Yes	Yes	Yes	Yes
Cassava	Yes	Yes	Yes	Yes	Yes
Wheat	Yes	Yes	Yes	Yes	Yes
Tropical	Yes	Yes	Yes	Yes	Yes
Ruggedness	Yes	Yes	Yes	Yes	Yes
Elevation	Yes	Yes	Yes	Yes	Yes
Dist. Coast	Yes	Yes	Yes	Yes	Yes
Observations	143069	74109	26739	22988	95585

Table 8: The impact of potato suitability on population: Increasing malarious conditions

	$MSI = 0$	$0 < MSI < 1$	$1 < MSI < 5$	$5 < MSI < 10$	$MSI > 10$
	b/se	b/se	b/se	b/se	b/se
Potato x Post 1900	1.173***	0.214***	0.271	0.298	0.543***
	(0.113)	(0.069)	(0.312)	(0.234)	(0.050)
Barley	Yes	Yes	Yes	Yes	Yes
Maize	Yes	Yes	Yes	Yes	Yes
Cassava	Yes	Yes	Yes	Yes	Yes
Wheat	Yes	Yes	Yes	Yes	Yes
Tropical	Yes	Yes	Yes	Yes	Yes
Ruggedness	Yes	Yes	Yes	Yes	Yes
Elevation	Yes	Yes	Yes	Yes	Yes
Dist. Coast	Yes	Yes	Yes	Yes	Yes
Observations	143069	74109	26739	22988	95585

References

[1] Daron Acemoglu and Simon Johnson. “Disease and development: the effect of life expectancy on economic growth”. In: *Journal of political Economy* 115.6 (2007), pp. 925–985.

[2] Hoyt Bleakley. “Malaria eradication in the Americas: A retrospective analysis of childhood exposure”. In: *American Economic Journal: Applied Economics* 2.2 (2010), pp. 1–45.

[3] Paula Bustos, Bruno Caprettini, and Jacopo Ponticelli. “Agricultural productivity and structural transformation: Evidence from Brazil”. In: *American Economic Review* 106.6 (2016), pp. 1320–65.

[4] Kyle Emerick. “Agricultural productivity and the sectoral reallocation of labor in rural India”. In: *Journal of Development Economics* 135 (2018), pp. 488–503.

[5] Günther Fischer et al. *Global agro-ecological zones (GAEZ v3. 0)-model documentation.* Tech. rep. IIASA and FAO, 2012.

[6] Elizabeth Gooch. “The impact of reduced incidence of malaria and other mosquito-borne diseases on global population”. In: *Journal of Development Economics* 124 (2017), pp. 214–228.

[7] Noel Gorelick et al. “Google Earth Engine: Planetary-scale geospatial analysis for everyone”. In: *Remote Sensing of Environment* 202 (2017), pp. 18–27.

[8] Thomas Grennes. “The Columbian exchange and the reversal of fortune”. In: *Cato J.* 27 (2007), p. 91.

[9] Robert Hijmans, N. Garcia, and John Wieczorek. *GADM database of Global Administrative Areas.* Tech. rep. 2010.

[10] Ruixue Jia. “Weather shocks, sweet potatoes and peasant revolts in historical China”. In: *The Economic Journal* 124.575 (2014), pp. 92–118.

[11] Anthony Kiszewski et al. “A global index representing the stability of malaria transmission”. In: *The American journal of tropical medicine and hygiene* 70.5 (2004), pp. 486–498.

[12] Kees Klein Goldewijk et al. “The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years”. In: *Global Ecology and Biogeography* 20.1 (2011), pp. 73–86.

[13] M Livi-Bacci. *A Concise History of World Population. Forth Ed.* 2006.

[14] Angus Maddison. *The world economy.* OECD publishing, 2006.

[15] John W McArthur and Gordon C McCord. “Fertilizing growth: Agricultural inputs and their effects in economic development”. In: *Journal of development economics* 127 (2017), pp. 133–152.

[16] Colin McEvedy, Richard Jones, et al. *Atlas of world population history.* Penguin Books Ltd, Harmondsworth, Middlesex, England., 1978.

[17] Daniel L Millimet. “The elephant in the corner: a cautionary tale about measurement error in treatment effects models”. In: *Missing Data Methods: Cross-Sectional Methods and Applications.* Emerald Group Publishing Limited, 2011, pp. 1–39.

[18] Nathan Nunn and Diego Puga. “Ruggedness: The blessing of bad geography in Africa”. In: *Review of Economics and Statistics* 94.1 (2012), pp. 20–36.

[19] Nathan Nunn and Nancy Qian. “The Columbian exchange: A history of disease, food, and ideas”. In: *Journal of Economic Perspectives* 24.2 (2010), pp. 163–88.

[20] Nathan Nunn and Nancy Qian. “The potato’s contribution to population and urbanization: evidence from a historical experiment”. In: *The Quarterly Journal of Economics* 126.2 (2011), pp. 593–650.

[21] Ragnar Nurkse. *Problems of capital formation in underveloped countries.* Oxford University Press, 1953.

[22] Paul N Rosenstein-Rodan. “Problems of industrialisation of eastern and south-eastern Europe”. In: *The economic journal* 53.210/211 (1943), pp. 202–211.

[23] Walt Whitman Rostow. *The process of economic growth.* Tech. rep. 1960.

GREEN
Centre for Geography, Resources, Environment, Energy and Networks
via Röntgen, 1
20136 Milano - Italia

www.green.unibocconi.eu