Policy Research Working Paper 9146

The Role of Historical Christian Missions in the Location of World Bank Aid in Africa

Matteo Alpino
Eivind Moe Hammersmark

WORLD BANK GROUP
Development Economics
Knowledge and Strategy Team
February 2020

Abstract

This article documents a positive and sizable correlation between the location of historical Christian missions and the allocation of present-day World Bank aid at the grid-cell level in Africa. The correlation is robust to an extensive set of geographical and historical control variables that predict settlement of missions. The study finds no correlation with aid effectiveness, as measured by project ratings and survey-based development indicators. Mission areas display a different political aid cycle than other areas, whereby new projects are less likely to arrive in years with new presidents. Hence, political connections between mission areas and central governments could be one likely explanation for the correlation between missions and aid.

This paper is a product of the Knowledge and Strategy Team, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at alpino.mtt@gmail.com.

The Role of Historical Christian Missions in the Location of World Bank Aid in Africa

Matteo Alpino and Eivind Moe Hammersmark*

JEL classification: F35, I3, N37, N77, O19
Keywords: development aid; Christian missions; political favoritism; Africa

1 Introduction

Where does foreign development aid go? This question is of central importance in the aid effectiveness debate. The World Bank has an explicit goal to end extreme poverty and to focus on the poorest segment of the population (World Bank Group, 2013). This strategy suggests that aid allocation should be guided by efficiency and need, but several empirical studies seem to indicate that allocation is biased by political and strategic considerations.

*Matteo Alpino (corresponding author) is an Economist in the Regional Economic Research Division at the Bank of Italy, Bari branch, Italy; his email address is alpino.mtt@gmail.com. Eivind Moe Hammersmark is a Senior Economist at Oslo Economics, Oslo, Norway, and an affiliated researcher at the Centre of Equality, Social Organization and Performance (ESOP), Oslo, Norway; his email address is eivindhammers@gmail.com. This work was partly carried out while the authors were at the University of Oslo and was supported by ESOP, which is funded by the Research Council of Norway through its Centres of Excellence funding scheme [179552]. Alpino also acknowledges support from the SFB 884 "Political Economy of Reforms" hosted at the University of Mannheim and funded by the German Research Foundation (DFG). The views expressed in this article are those of the authors and do not necessarily reflect the views of the Bank of Italy. The authors are grateful to Valeria Rueda and Julia Cagé for sharing their data. The authors also thank the editor Aart Kraay and two anonymous referees for a fruitful editorial process, as well as Axel Dreher, Martin Flatø, Nicola Gennaioli, Rune Jansen Hagen, Andreas Kotsadam, Eliana La Ferrara, Edwin Leuven, Jo Thori Lind, Silvia Marchesi, Halvor Mehlum, Anirban Mitra, Kalle Moene, Alexander Moradi, Andreas Müller, Ola Olsson, Johanna Rickne, Anna Tompsett, Gaute Torsvik, and Nina Bruvik Westberg for comments.

Most of the early research on this subject has focused on the cross-country and across-time dimensions.[1] However, the more recent literature on the determinants of within-country aid allocation has come to similar conclusions: various kinds of favoritism are important in explaining the spatial allocation of development aid (Dreher, Fuchs, Hodler, Parks, Raschky, & Tierney, 2019; Jablonski, 2014; Masaki, 2018).

The present article adds to this literature by studying the role of history in shaping present-day within-country allocation of aid. Although development aid in its present form is a post–World War II phenomenon, similar activities implemented by Westerners in developing countries began much earlier.[2] In particular, they can be traced back to the work of Christian missionaries, who were particularly active at the end of the 19th century. The missionary effort was primarily driven by proselytization motives, but it was not restricted to conversion. Missions provided locals with a wide range of education and health services, primarily to boost the odds of conversion. In some ways, mission stations can be considered the ancestors of modern micro-development projects.

The main empirical analysis of this article compares a snapshot of the location of mission stations in Africa in 1903 to the precise locations of projects funded by the World Bank in 1995–2014. The unit of analysis is derived from a grid of 55 km × 55 km square cells covering the African mainland and Madagascar.[3] The results imply that the presence of (at least) one mission station increases the probability that an area is allocated a development project by approximately 50 percent. Several empirical measures are taken to alleviate concerns of omitted variable bias. First, the regressions include country dummies in all specifications, because the first step of aid allocation is at the country level. Second, the empirical strategy addresses the nonrandom selection of missionaries into specific locations. To this end, the sample al-

[1] For example, Alesina and Dollar (2000) found colonial history and co-voting in the United Nations to be major predictors of donor-recipient foreign aid flows. Along the same lines, Dreher, Sturm, and Vreeland (2009) showed that the World Bank allocates disproportionately more development projects to countries during their tenure as temporary members of the UN Security Council.

[2] The beginning of modern development aid coincides with the establishment of the World Bank in 1944 and the launch of the US-sponsored Marshall Plan in 1948, aimed at reconstructing European economies after World War II.

[3] The main results are robust to collapsing the data to administrative levels 1 and 2 (regions and districts, respectively), as well as to ethnic homelands as defined in Murdock (1959). These results are available from the authors upon request.

ways excludes areas covered by desert or dense forest, and regressions always control for historical and geographical factors that guided the missionaries' settlement decisions, according to historical sources. Third, the correlation is robust when the sample is restricted to areas that are more likely to be similar: areas that intersect the coastline or one of the main rivers, and subsamples obtained by propensity score matching. Fourth, the link between historical missions and aid survives also when controlling flexibly for present-day population density. Finally, the test developed by Oster (2019) to assess the extent of omitted variable bias suggests that only a small part of the estimated correlation is likely to be driven by unobservable factors.[4]

The findings of this study relate to the recent literature on the long-lasting effects of Christian missionary activities on development. Several convincing pieces of evidence point to large effects of both Catholic and Protestant missions on present-day education (Caicedo, 2018; Castelló-Climent, Chaudhary, & Mukhopadhyay, 2018; Mantovanelli, 2014; Meier zu Selhausen, 2014; Nunn, 2014; Okoye & Pongou, 2017; Waldinger, 2017), health (Cagé & Rueda, 2017; Calvi & Mantovanelli, 2018), and income (Caicedo, 2018; Chen, Wang, & Yan, 2014).[5] The literature also documents that the effects of missionary interventions are not explained by the persistence of infrastructure (e.g., schools and hospitals). Instead, they seem to be explained by the transmission of new values, the introduction of better practices, and an increase in non-cognitive skills such as collaborative behavior.

The literature described above offers one possible explanation for the correlation between the location of historical missions and the present-day geographical allocation of aid. Aid donors always face a trade-off between need and effectiveness and might decide to channel aid toward areas where the probability of success is higher. Areas that hosted historical missions

[4]In a previous version of this article (Alpino & Hammersmark, 2017), the allocation of Chinese-financed aid was also analyzed. The correlation between historical missions and Chinese aid is unstable across specifications, and it is not consistent across different sources of missionary data. The present article focuses only on projects funded by the World Bank.

[5]Additional articles in this literature include Cagé and Rueda (2016), which finds that the introduction of the printing press by Protestant missionaries facilitated the birth of newspapers and, in turn, the accumulation of social capital, and Kudo (2017), which finds that missionary-educated women marry later and are less likely to marry a polygamous husband. The positive effects of missionaries are also found outside developing countries; for example, Andersen, Bentzen, Dalgaard, and Sharp (2017) documented positive effects of monasteries in medieval England.

could be more suited to successful implementation of aid projects, thanks to higher levels of social and human capital. This would be consistent with recent evidence that more-developed areas receive more aid (Briggs, 2017; Nunnenkamp, Öhler, & Andrés, 2017).

The second part of the article (Section 3) puts this hypothesis to an empirical test, employing two different strategies. The first compares the performance of projects implemented in the vicinity of missions with those further away, using the ratings by the World Bank's Independent Evaluation Group (IEG) as a proxy for project performance. There is no evidence of a correlation between proximity to mission stations and project ratings, but the estimates are rather imprecise and thus do not allow definitive conclusions to be drawn. The second strategy exploits information on start and end dates of the aid projects to implement a triple-differences design. In particular, areas that received aid at different points in time are compared to test whether aid arrival affects the level of development and, more importantly, whether the effect is higher in the vicinity of mission stations. Proxies of development are constructed with georeferenced survey data from the Demographic and Health Survey (DHS) and include measures of wealth and access to public utilities. Under the identifying assumption that development trends (not levels) between areas with and without missions as of 1903 are comparable in the period 1995–2014, there is no evidence that mission presence matters for aid effectiveness. As with the first strategy, it is not possible to precisely estimate the absence of an effect, and so the conclusion drawn from this exercise is suggestive rather than definitive.

The final part of the article (Section 4) investigates whether favoritism can explain the correlation between aid and missions. First, it examines the role of *political* favoritism. There is evidence from Africa that funds from the World Bank and the African Development Bank have been diverted to politically important areas—competitive electoral districts (Masaki, 2018), strongholds of the incumbent regime (Briggs, 2014; Jablonski, 2014), or birth regions of presidents (Dreher, Fuchs, Hodler, Parks, Raschky, & Tierney, 2019). In light of the fact that areas close to historical missions have higher social capital and are more developed, it is reasonable to suspect that they are also politically more important. Mission areas may also have had strong connections with the central government in colonial times, and these connections may have persisted over time. The analysis puts this hypothesis to an empirical test by estimating the existence of a political aid cycle specific to mission areas. Findings from a specification that includes country-year

as well as cell fixed effects on a balanced annual panel suggest that areas in the vicinity of mission stations experience a 40 percent drop in the probability of receiving a new aid project in the year of a presidential turnover. A corresponding increase takes place in election years when the incumbent is re-elected, although this estimate is less precise.

The article also investigates the role of religious favoritism. There is evidence that Christian missionaries have been successful in converting indigenous peoples to Christianity (Nunn, 2010; Waldinger, 2017). As World Bank donors are predominantly Western and Christian countries, they might prefer to channel aid to areas with a large Christian population. To probe this mechanism, the Christian share of the population is included as a control in the baseline specification. Controlling for religion has virtually no effect on the correlations between mission presence and aid. The same empirical strategy is used to investigate the role of education, for which there is extensive evidence that missionary interventions matter. In contrast to religion, the inclusion of education in the baseline specification weakens the correlation between mission presence and aid. In light of this finding, it is not possible to rule out that human capital plays a role.[6]

2 Spatial Correlation Between Aid and Historical Missions

The main aim of this article is to test whether the presence of historical mission stations is correlated with present-day allocation of aid. The empirical strategy exploits within-country variation in missionary activity and aid across Africa. Since the locations of missions are predetermined and do not vary over time, the temporal dimension of aid allocation is collapsed into a cross-sectional dataset. The units of observation are contiguous grid cells at a resolution of $0.5° \times 0.5°$, which at the equator roughly corresponds to $55\,\text{km} \times 55\,\text{km}$. The grid is superimposed on the African mainland and Madagascar.[7] Cells are split by borders in this process, to make sure that aid projects are geographically assigned to the correct country. Cells are assigned to the two highest subnational administrative levels from the GADM

[6]Note that this exercise is problematic, because the regression controls for a covariate that is not predetermined with respect to mission presence. These findings are therefore suggestive and must be interpreted with caution.

[7]Cells in South Sudan are dropped to ensure consistent country fixed effects.

database of global administrative areas—ADM1, corresponding to states or governorates, and ADM2, corresponding to districts, municipalities, or communes—using their center points ("centroids"). The baseline specification to be estimated by ordinary least squares (OLS) is

$$\text{EverAid}_{ik} = \beta \cdot \text{Mission}_{ik} + \delta_k + \mathbf{X}_{ik}\gamma + \varepsilon_{ik}\,, \tag{1}$$

where i and k are indices for cell and country, respectively; EverAid_{ik} is a binary variable that equals 1 if cell i had at least one active aid project in the period of study and equals 0 otherwise; Mission_{ik} is an indicator that equals 1 if at least one historical mission was located in cell i and equals 0 otherwise.

The baseline estimation exploits only within-country variation. The inclusion of country dummies δ_k controls for all time-invariant country-level characteristics, many of which are important determinants of foreign aid.[8] To test for robustness, additional specifications include dummies for subnational administrative divisions (ADM1 or ADM2) instead of country-level dummies.

The vector $\mathbf{X}_{ik}$ consists of control variables at the cell level, including a number of historical and geographical factors described in Section 2.2. The error term ε_{ik} is allowed to be spatially correlated within a radius of 220 kilometers (approximately four times the side length of a cell) from the centroid of each cell. This means that clusters are unique to each cell and that a typical landlocked cluster covers about 45 contiguous cells.[9] Clustered standard errors are calculated using the estimator developed by Conley (1999).[10] This type of standard error is appropriate because missions are highly clustered in the data (see fig. 1).

The coefficient β has a causal interpretation if and only if $\mathbf{X}_{ik}$ and δ_k contain all relevant determinants of mission locations that are also correlated with present-day aid allocation. Failure to include important controls will bias the size of the coefficient. The most obvious threat to a causal interpretation of β is nonrandom selection of mission stations. Missions are

[8] For example, to be eligible for International Development Association (IDA) funds from the World Bank, a country must be below a threshold level of GNP per capita (Galiani, Knack, Xu, & Zou, 2017).

[9] Coastal clusters are obviously smaller.

[10] The standard errors are calculated using the Stata program `x_ols` written by Jean-Pierre Dube, available at http://economics.uwo.ca/faculty/conley/.

predetermined with respect to present-day aid, but they may have been located in areas that were more suitable for missionary work, or areas where missionaries could survive and be self-sustained. If these areas are more or less likely to be selected for aid projects for some reason other than mission station presence, β will be biased. Detailed information is available on the determinants of the location of mission stations from historical sources, notably Johnson (1967) and Robinson (1915). Moreover, increasing amounts of detailed and spatially disaggregated historical data are available, so the regression analysis can plausibly control for most determinants: precolonial ethnic institutions, distance from coast and rivers, water accessibility, malaria prevalence, altitude, terrain characteristics, historical population and cities, and distance from historical routes. In addition to controlling for these factors, Oster (2019) bounds are also calculated to formally assess the extent of the bias due to unobserved controls.

A second source of bias is lack of common support in the distribution of the control variables. Even if the set of controls $\mathbf{X}_{ik}$ fully accounts for the selection problem, the simple OLS estimator will still be biased if the cells hosting a historical mission (the treated observations) are very different in their covariates from cells without a historical mission (the control observations) and if the control function is incorrectly specified (Imbens & Rubin, 2015). Several measures are taken to deal with this issue. First, the sample always excludes cells where more than 90 percent of the surface area was covered by desert or forest in the 18th century. Second, in some specifications the sample is further restricted to coastal cells and cells that intersect one of the main African rivers.[11] Third, in other specifications the analysis is conducted on subsamples obtained using propensity score matching.

Finally, a third source of bias derives from the spatial nature of the data. Since mission stations are spatially clustered, the regression analysis may be overestimating or underestimating the impact of a single mission station. To address this issue, alternative specifications also include spatial lags of mission stations as regressors.

[11] The Nile, the Niger, the Senegal, the Zambezi, and the Congo and its tributaries the Ubangi and Kasai.

2.1 Aid Data

Data on foreign aid are sourced from World Bank projects in the period 1995–2014, geocoded by AidData. Other geocoded datasets on aid exist, but this one covers the whole African continent (and beyond) for the longest period. The World Bank dataset contains projects from both the International Bank for Reconstruction and Development (IBRD) and the International Development Association (IDA), totaling 1,900 projects in Africa split across 16,553 different locations. The IBRD provides low or zero interest rate loans to sufficiently creditworthy countries, whereas the IDA gives loans to poorer and less creditworthy countries; 12 percent of IDA funds are given as grants not to be paid back. Both types of lending are accompanied by technical assistance from the World Bank, and projects are monitored by World Bank staff.[12]

The data contain information on all locations in which a given project has been implemented. Locations are classified into categories 1 to 8 according to the level of geographical disaggregation of their coordinates, with the categories called, somewhat misleadingly, "precision categories." Precision 1 locations correspond to a specific place, that is, a populated place of some kind (such as a village, town, or city), in approximately 80 percent of the cases, or to a third-order administrative division (ADM3, i.e., a neighborhood or suburb) in approximately 15 percent of the cases.[13] Precision 2 locations are similar to precision 1 locations, but their reported coordinates, being within 25 km of the exact location, are not as accurate. Precision 3, 4, and 6 locations correspond to second-order administrative divisions (ADM2), first-order administrative divisions (ADM1), and countries, respectively. Note that precision 3, 4 and 6 categories refer to projects intended to serve the whole administrative division (e.g., training for all public employees in a

[12]The information available at the project level includes the original World Bank identifier, project title, date of approval, expected date of completion, share in different sectors (finance, transportation, energy, health, education, agriculture, water, industry and trade, information communication technology, public administration), lending instrument (development policy lending versus investment), local implementing agency, total committed and disbursed amounts, completion and supervision costs, and independent evaluation rating.

[13]First-order administrative divisions (ADM1) are the largest administrative units (provinces, states, or governorates); second-order administrative divisions (ADM2) are units at the next level (districts, municipalities, or communes); third-order administrative divisions (ADM3) are subdivisions of ADM2s (neighborhoods or suburbs), and so on.

Table 1: Precision of Aid Locations

Precision code	Percentage of locations
Precision 1: specific place	40.9
Precision 2: within 25 km of specific place	2.4
Precision 3: municipality (ADM1)	25.7
Precision 4: province (ADM2)	19.7
Precision 5: imprecise	1.9
Precision 6: country-wide projects	4.4
Precision 7: unclear	0
Precision 8: state or national capitals	5

Source: Authors' summary based on AidData, World Bank geocoded research release, version 1.3.
Note: The sample is composed of 1,900 projects in Africa.

province). They do *not* refer to imprecisely georeferenced point locations. Precision 5 locations are imprecisely geocoded, so only approximate coordinates are reported. Precision 7 locations are "unclear" in the sense that it was only possible to identify the country in which the project is located. Finally, precision 8 locations correspond to capital cities (both national and local) and also include projects aimed at government institutions (such as a ministry or the central bank). The distribution of location precision categories is reported in table 1.

Most projects are implemented across several locations (around 40 on average), often belonging to different precision categories. Consider, for example, a project that aims to build a road connecting two towns in two different provinces. In this case, at least four locations are assigned to the project: the two towns as precision 1 locations and the two provinces as precision 4 locations, plus any province crossed by the road as an additional precision 4 location, and any town crossed by the road as an additional precision 1 location. For the purpose of the analysis, only locations of precision 1 or 2 are retained: precisions 3, 4, and 6 are too coarse to be uniquely assigned to a cell in the grid-level analysis, and precision 8 locations are removed to prevent the results from being driven by capitals. This leaves 768 projects (40 percent) and 12,318 project locations (74 percent) in the analysis.

These sample restrictions raise the question of whether the excluded projects are systematically different from the retained ones. This question is

investigated by checking whether the included projects are also implemented in locations with different precision categories (see appendix). Reassuringly, more than 60 percent of the projects retained in the sample are also assigned to at least one location at the ADM1 (precision 4) or ADM2 (precision 3) level. The projects in the sample are also compared to projects with at least one location at precision 3 or 4 but without any location at precision 1 or 2, in terms of observable characteristics (see appendix). The projects in the two groups are broadly similar, although there are some differences in terms of sectoral composition; the projects in the sample have, on average, smaller shares in agriculture, health, and education, but larger shares in energy, transport, and water sanitation. This suggests that the sample of precise locations has (unsurprisingly) a disproportionate share of projects dedicated to the building of facilities and infrastructure.

2.2 Location of Mission Stations

The analysis relies on two different historical sources for information on the locations of Christian mission stations. The preferred source is *Geography and Atlas of Christian missions* by Beach (1903), digitized by Cagé and Rueda (2016). It includes the locations of Protestant mission stations in Africa as of 1903, together with information on the investment of each mission (school, dispensary, hospital, etc.). An alternative source is *Ethnographic Survey of Africa: Showing the Tribes and Languages* by Roome (1924), digitized by Nunn (2010). This source reports locations of both Protestant and Catholic foreign mission stations in Africa as of 1924.

As is apparent from fig. 1, the two sources do not overlap perfectly. The cell-level correlation between Protestant missions from Beach (1903) and Roome (1924) is 0.31. One reason for the low correlation could be that missionaries started penetrating the African inland only after first settling on the coast. The 1924 data indeed show a higher concentration of mission stations farther from the coast. Therefore, this study takes a conservative approach, and the analysis is conducted using the two sources separately. The results of virtually all the analyses, using either one of the two atlases to construct Mission_{ik}, are quantitatively similar and qualitatively identical. Results from the baseline regression, equation (1), are reported for both sources, but for ease of exposition only results obtained using data from

Beach (1903) are reported for the other regressions.[14]

These two atlases are standard sources for georeferenced mission stations in Africa in the economic literature. However, Jedwab, Meier zu Selhausen, and Moradi (2018) have documented that both sources are subject to measurement error in the exact locations of missions due to geocoding mistakes. They also showed that any issue of *classical* measurement error can be substantially mitigated by using large enough cells. In particular, classical measurement error is almost completely eliminated when the cell size is increased to $0.3° \times 0.3°$. Reassuringly, the present study uses even larger cells, with a resolution of $0.5° \times 0.5°$. Jedwab, Meier zu Selhausen, and Moradi (2018) also collected more complete records of missions in Ghana from multiple sources. They were able to show that for Ghana, the correlation between their geocoded locations and those reported in Beach (1903) and Roome (1924) is very high only for those missions that were established early, but is lower for missions that were established later.[15] As early missions were located in better and more accessible areas, this might induce nonclassical measurement error, which tends to magnify the omitted variable bias induced by self-selection of missionaries. The next section explains how the analysis deals with the selection issue.

2.3 Selection of Missions and Historical Controls

Missionary activity in Africa was not randomly assigned across the continent, as illustrated by the case studies in Johnson (1967) and confirmed empirically in Jedwab, Meier zu Selhausen, and Moradi (2018). If the factors that determine the selection of mission station locations correlate with present-day aid allocation, the coefficient β in equation (1) will be biased.

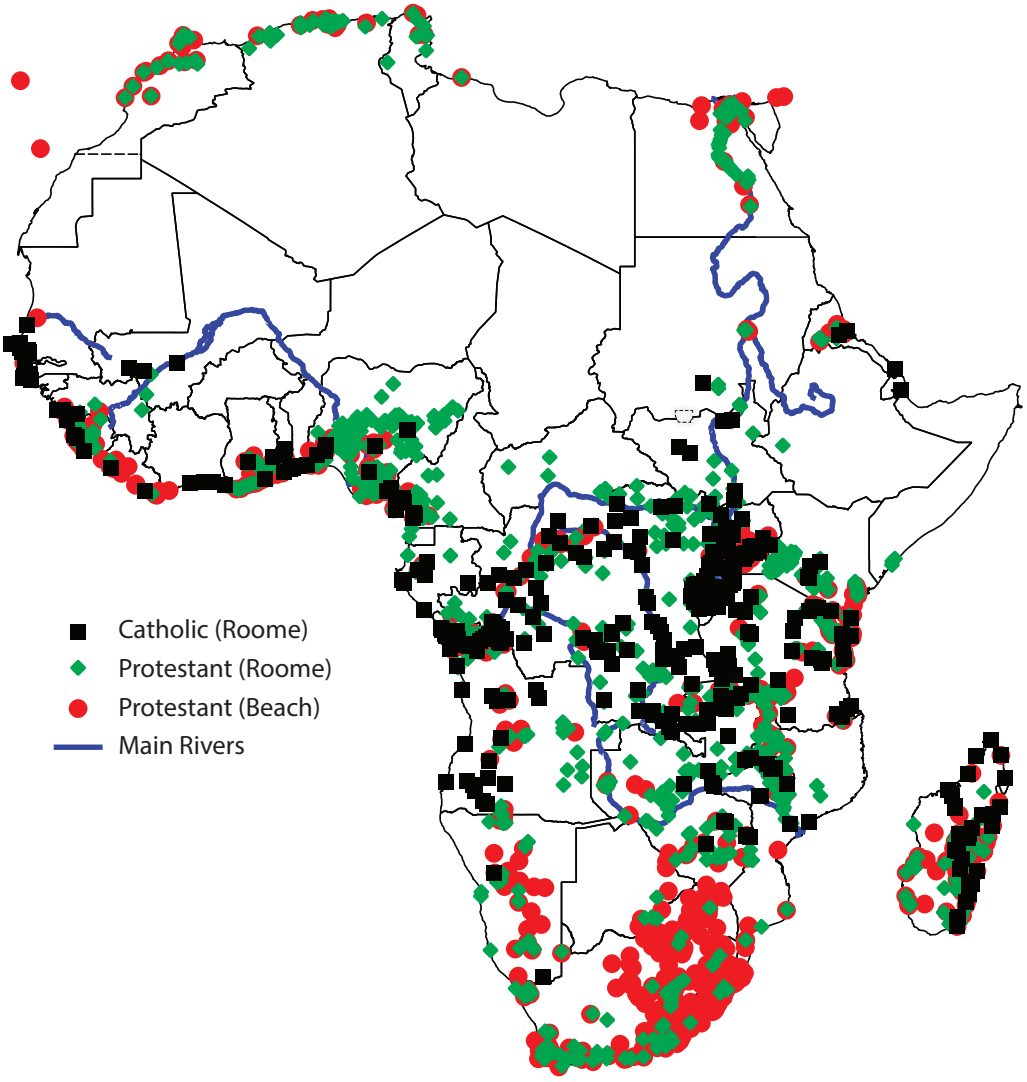
The first factor to consider is accessibility. Missionaries came by sea, and inland penetration was difficult, so they followed the tracks of early European explorers, which partly correspond to the courses of the main rivers. There is evidence that areas along the coast and along large rivers have an advantage in development and are more densely populated (Gallup, Sachs, & Mellinger, 1999), so the regressions control for the (log) distance (in kilometers) from the closest point on the coast and from the closest main river.[16]

[14] Results obtained using the Roome (1924) data are available from the authors upon request.

[15] The correlation increases as the cell size increases.

[16] Logarithms of distances are used because the marginal effect of a unit of proximity

Figure 1: Location of Mission Stations and Main Rivers



Source: Authors' summary of data from Beach (1903) and Roome (1924). *Note*: This map shows the main rivers in Africa and the locations of Catholic and Protestant missions.

All specifications also control for the (log) distance to the closest colonial railway, as the railways had long-lasting effects on urbanization and growth in Africa (Jedwab, Kerby, & Moradi, 2017; Jedwab & Moradi, 2016), and distance to the closest explorer route. Finally, as a measure of accessibility, regressions also include a measure of terrain ruggedness, which in addition to shaping location decisions may have indirectly affected development and therefore aid allocation. There is evidence, for example, that a rugged landscape made it easier to hide from slave traders (Nunn & Puga, 2012) and enables rebel warfare (Fearon & Laitin, 2003).

The second factor to consider is the capacity to keep the settlement self-sustained for a long period of time. Self-sustainability crucially depended on access to water and suitability of land cultivation, which are likely to be important for present-day outcomes as well. Thus the regressions control for both factors, proxied respectively by the *Caloric Suitability Index* (Galor & Özak, 2016) and the share of cell area that is within 10 km of a water source (following Nunn [2010]). Missions were also more likely to be established at high altitudes, partly to avoid diseases such as malaria, but also because of the more comfortable climate (Johnson, 1967). Regressions therefore control for average elevation and for its interaction with a dummy for the tropics. As a further control for disease environment, a measure of malaria prevalence, the *Malaria Ecology Index* developed by Kiszewski, Mellinger, Spielman, Malaney, Sachs, and Sachs (2004), is included in the control set.

In addition, the existence of different ethnic groups may have played an important role in the missionaries' settlement decisions. Unobserved variables at the ethnicity level may introduce biases, in light of research showing that precolonial ethnic institutions had long-lasting effects on development and public goods provision in Africa (Gennaioli & Rainer, 2007; Michalopoulos & Papaioannou, 2013). A separate dummy for each of the more than 800 precolonial ethnic homelands is included to address this concern, constructed using boundaries from Murdock's (1959) ethnolinguistic map. Each cell is assigned to the ethnic polygon that covers the highest percentage of its surface. Regressions therefore exploit only within-ethnicity variation, and the results cannot be driven by factors varying across ethnicities.

It is also necessary to account for the main missionary purpose, namely conversion of Africans to Christianity. In particular, there is a concern that

is likely to approach zero as distance increases. More specifically, the regressors have the form $\log(1 + \text{distance})$ to avoid losing the cells at zero distance.

missionaries might have targeted more populated areas or cities. Regressions therefore control for a fourth-order polynomial in average population density in the 18th century obtained from the History Database of the Global Environment (HYDE). Furthermore, the control set also includes a dummy for the presence of cities at any time before 1800.

According to Robinson (1915), competition with Islam was a deterring factor, because spreading the gospel in predominantly Muslim areas was complicated. Muslim populations may receive less development aid for political and religious reasons, so the (log) distance from the closest Arab medieval trade route (which Michalopoulos, Naghavi, and Prarolo [2018] showed had a strong impact on adherence to Islam) is controlled for.[17] Table 2 presents summary statistics of the controls for cells with and without missions. More details about the data sources are given in appendix A.

2.4 Results

Equation (1) is estimated by least squares, and the results are reported in table 3. The dependent variable is an indicator that equals 1 if the cell ever received a World Bank project between 1995 and 2014.[18] All regressions include the full set of historical and geographical controls described above. The first set of regressions employs different data sources and definitions of the dummy Mission_{ik}. In the first four columns of table 3, the mission data are from Roome (1924). The regression includes only Catholic missions in column 1, only Protestant missions in column 2, and both kinds of mission in column 3; in column 4 the two types of mission are collapsed into a single

[17]Some missions were set up for the purpose of ending the slave trade (Johnson, 1967), a practice that was especially prevalent along the coast of west Africa and had long-lasting detrimental effects on development and social capital (Nunn, 2008; Nunn & Wantchekon, 2011). The authors are not aware of precisely georeferenced measures of slave trade; however, the regressions in this study already control for many of its correlates, such as distance to the coast, terrain ruggedness, and distance to Arab trade routes, as well as ethnic-level dummies.

[18]Here and in the rest of the article the analysis always relies on linear probability models (LPMs) when the dependent variable is binary, rather than using nonlinear models. As the focus is on estimating differences in averages between groups, rather than predicting outcomes, the LPM performs well compared to nonlinear models, and the coefficients are more straightforward to interpret, especially in the presence of interaction terms. Furthermore, the controls often introduce large sets of dummies in the regressions, which regularly cause failure of convergence of the likelihood maximization algorithm for logit or probit models.

Table 2: Difference in Means of Control Variables

	No mission	Mission	Difference
Log(Distance to coast)	5.76	4.01	1.76^{***}
Log(Distance to main river)	5.29	5.89	-0.60^{***}
Percentage of area within 10 km of water	0.07	0.10	-0.03^{***}
Malaria Ecology Index	11.64	6.82	4.83^{***}
Caloric Suitability Index / 1000	1.33	1.57	-0.24^{***}
Terrain Ruggedness Index	17.18	26.42	-9.25^{***}
Mean elevation	714.05	750.26	-36.22
Tropical dummy	0.86	0.56	0.31^{***}
Log(Distance to explorer route)	3.62	4.20	-0.58^{***}
Log(Distance to colonial railway)	5.28	3.59	1.68^{***}
18th-century population	11.46	23.48	-12.02^{***}
Precolonial city	0.01	0.04	-0.04^{***}
Log(Distance to Arab trade)	5.44	6.09	-0.64^{***}
Observations	6,512	380	

Source: Authors' calculations based on mission locations from Beach (1903) and control variables from multiple sources, detailed in appendix A.
Note: $^{*}\, p < .1 \quad ^{**}\, p < .05 \quad ^{***}\, p < .01$

Table 3: Correlation between Aid and Missions

	World Bank aid 1995–2014						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Catholic mission (Roome)	0.20***		0.18***			0.19***	
	(0.03)		(0.03)			(0.03)	
Protestant mission (Roome)		0.15***	0.14***				
		(0.02)	(0.02)				
Any mission (Roome)				0.17***			
				(0.02)			
Protestant mission (Beach)					0.13***	0.11***	
					(0.03)	(0.03)	
Any mission							0.15***
							(0.02)
Ethnic dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean dependent variable	0.25	0.25	0.25	0.25	0.25	0.25	0.25
Oster bound	0.11	0.07		0.08	0.09		0.10
R-squared	0.40	0.40	0.40	0.40	0.39	0.40	0.40
Observations	6,876	6,876	6,876	6,876	6,876	6,876	6,876

Source: Authors' analysis based on mission locations from Beach (1903) and Roome (1924); aid projects from AidData, World Bank geocoded research release, version 1.3; and control variables from multiple sources, detailed in appendix A.

Note: Estimation is by ordinary least squares. Conley (1999) standard errors are given in parentheses, with the cutoff at 2 degrees (220 km). The dependent variable is a dummy for at least one project commitment in the sample period. Control variables include: log distances to coast, explorer route, colonial railway, and Arab trade route; a third-order polynomial in the 18th-century population; presence of a city as of 1800; country dummies; precolonial ethnic dummies; average altitude; Terrain Ruggedness Index; percentage area within 10 km from water source; Caloric Suitability Index; Malaria Ecology Index; and a tropics dummy (which is also interacted with mean altitude). Cells that are more than 90 percent covered by barren land or more than 90 percent covered by forest are excluded from the sample. The lower bound on the coefficient of interest is calculated as in Oster (2019): the R-squared from the hypothetical regression on Mission_{ik} and both observed and unobserved controls is set to 1.3 times the R-squared from the actual regression on Mission_{ik} and the observed control. $^{*}\, p < .1$ $^{**}\, p < .05$ $^{***}\, p < .01$

dummy. In column 5 the regression includes only Protestant missions from Beach (1903), and in column 6 it includes Protestant missions from Beach (1903) and Catholic missions from Roome (1924). Finally, in column 7, the regression includes a dummy for Protestant and Catholic missions collapsed into one using data from both sources together.

The correlation between historical mission presence and Word Bank aid location is positive and significant across the columns of table 3. The estimated coefficients imply that cells with missions are approximately 45–80 percent more likely to host a World Bank project, relative to the sample mean. In order to assess the bias from unobservables, the analysis draws on the procedure developed by Oster (2019) to obtain a lower bound for β. This test formalizes the common practice of inspecting the stability of the coefficient of interest when controls are added. It is a refinement of the approach

of Altonji, Elder, and Taber (2005) in that it takes into account whether controls absorb residual variation. This is important for the credibility of the exercise, as one should not expect to observe coefficient instability upon adding controls that are unrelated to the outcome variable. Under the assumption that selection on observables has the same direction as selection on unobservables, the test produces lower-bound coefficients close to 0.1 (see the row headed "Oster bound" in table 3), corresponding to mission cells having a 40 percent higher likelihood of aid allocation. Importantly, the Oster bound is always significant at the 99 percent level.[19]

The estimated correlation is higher for Catholic missions, but the difference between denominations is significant at the 90 percent level only in column 6, which uses data from two different sources. Furthermore, there is no feasible way to address differential selection of Protestant versus Catholic missionaries, so it is difficult to interpret the difference between the two coefficients. Finally, if the Democratic Republic of the Congo (DRC) is dropped from the sample, the difference between the two disappears in column 3 and halves in column 6, becoming not significant. The DRC covers 7 percent of the sample cells, making it the largest country in the sample, and it has a high concentration of Catholic missions. Hence, the differential in the coefficient size seems to be driven by this heavyweight outlier. To simplify the exposition, the rest of the article reports results from regressions on the mission data from Beach (1903) only.

2.5 Robustness Tests

2.5.1 Subsample Analysis

The first two robustness tests are aimed at restricting estimation of equation (1) to subsamples that are more likely to exhibit common support in the covariate distribution (see table 4). First, the sample is restricted to cells on the coast or along one of the main rivers (C/R column). As discussed above, this subsample is likely to be more homogeneous and would enhance the credibility of the selection on observables strategy, which relies on mission and

[19]The R-squared from the hypothetical regression in which unobserved controls are included is set to 1.3 times the R-squared from the actual regression on Mission_{ik} and observed controls, as suggested by Oster (2019). The procedure is only suitable for models with one treatment, so bounds are not calculated for columns 3 and 6. Confidence intervals are calculated using standard errors from the regressions.

non-mission cells having the same covariate distributions (Imbens & Rubin, 2015). The analysis then proceeds in a more structured way by constructing three subsamples in which observations are balanced on propensity scores, using three different strategies. The first strategy (PS1) has no geographical restrictions in estimation of propensity scores or matching between treatment and control groups. In the second strategy (PS2), propensity scores are estimated separately within each country and the subsample includes treatment-control pairs that are statistical neighbors in the same country. In the third strategy (PS3), propensity scores are estimated on the full sample but the subsample includes only treatment-control pairs that are neighbors within the same country.[20]

The subsample analysis yields results similar to those of the baseline (table 3). Across the different subsamples, the coefficient of interest is always positive, significant at least at the 95 percent level, and sizable (35 percent of the sample means).

2.5.2 Controlling for Present-Day Population

All regressions estimated so far include only controls that are predetermined with respect to the variable of interest, Mission_{ik}. This approach is appropriate when the goal is causal interpretation of β. However, it requires the analysis to rely heavily on historical controls, some of which are likely to be measured with error, in particular population density. If the measurement error in the historical variables is severe, regressions may fail to control credibly for the selection of missions. Furthermore, there is reason to believe that the presence of mission stations and the activities of missionaries have influenced settlement patterns. In that case, the coefficient on Mission_{ik} partly captures a correlation between current population and aid allocation.

To address these concerns, different measures of present-day population

[20]For PS1, a logit model is estimated on the full sample using Mission_{ik} as the dependent variable and with $\mathbf{X}_{ik}$ (except ethnic dummies), its interactions, its squared terms, and country dummies as predictors; then mission cells are matched with their nearest (statistical) neighbor without replacement using the predicted values as propensity scores. For PS2, propensity scores are re-estimated separately country by country (using the logit model but without interactions and squared terms, because the individual country samples are small), and the pool of possible matches is restricted to cells that belong to the same country. For PS3, propensity scores are re-estimated with the same logit model as in PS1 but without country dummies; then the same matching procedure is performed, but country by country.

Table 4: Subsample Analysis

	World Bank aid 1995–2014			
	C/R	PS1	PS2	PS3
Mission	0.11**	0.12***	0.07**	0.10***
	(0.05)	(0.03)	(0.03)	(0.03)
Ethnic dummies	Yes	Yes	Yes	Yes
Mean dependent variable	0.43	0.36	0.29	0.35
R-squared	0.52	0.66	0.66	0.68
Observations	844	799	604	794

Source: Authors' analysis based on mission locations from Beach (1903); aid projects from AidData, World Bank geocoded research release, version 1.3; and control variables from multiple sources, detailed in appendix A.

Note: "C/R" stands for coast or river subsample. "PS" refers to different subsamples obtained by propensity score matching. Estimation is by ordinary least squares. Conley (1999) standard errors are given in parentheses, with the cutoff at 2 degrees (220 km). The dependent variable is a dummy for at least one project commitment in the sample period. Control variables include: log distances to coast, explorer route, colonial railway, and Arab trade route; a third-order polynomial in the 18th-century population; presence of a city as of 1800; country dummies; precolonial ethnic dummies; average altitude; Terrain Ruggedness Index; percentage area within 10 km from water source; Caloric Suitability Index; Malaria Ecology Index; and a tropics dummy (which is also interacted with mean altitude). Cells that are more than 90 percent covered by barren land or more than 90 percent covered by forest are excluded from the sample. The rivers include the Nile, Niger, Senegal, Zambezi, and Congo together with its tributaries the Ubangi and Kasai. $^{*}\,p < .1$ $^{**}\,p < .05$ $^{***}\,p < .01$

are introduced in the control set. Under the assumption that population density is positively autocorrelated, the present-day measures can serve as proxy controls for historical population. Under the additional assumptions that selection on population is positive and that the presence of mission stations increases population, it is possible to interpret the coefficients on Mission_{ik} from these regressions as lower bounds on the true causal effect.[21] The results from regressions that control for present-day population are reported in table 5.

In the first column, the specification includes fourth-order polynomial terms of population in 1995. In the second, it also includes a set of dummies for the presence of populated places of different sizes.[22] In the third column, the sample is restricted to cells containing a provincial capital, because these cities are likely to be large in terms of population and politically important. The regressions survive all three robustness tests. Estimates are still significant at least at the 99 percent level, and coefficients are comparable to those obtained previously (35 percent of the sample means).

2.5.3 Spatial Spillovers

The baseline equation (1) does not account for the possibility that benefits from hosting a mission station in one cell could spill over to surrounding cells. On the one hand, this might induce an overestimation of the effect of missions on aid. In cases where a pair of neighboring cells both hosted missions but only one receives aid, the neighbor's mission may be adding to the aid attraction. On the other hand, spatial spillover could induce an underestimation of the effect of missions on aid, because missionary presence in one cell might increase the probability of attracting aid to surrounding cells, even if those do not host any mission themselves. In both cases, failure to account for the presence of missions in surrounding cells could lead to omitted variable bias, in the first case with a positive sign and in the second case with a negative sign.

[21] See Angrist and Pischke (2008) for a discussion of this point, and Michalopoulos and Papaioannou (2013, footnote 13) for an example.

[22] These are dummies for the presence of at least one a) national capital, b) provincial capital, c) urban agglomeration of at least one million people or city with at least 500,000 people, d) urban agglomeration of at least 250,000 people or city with at least 100,000 people, e) urban agglomeration of at least 100,000 people or city with at least 50,000 people, f) places with at least 10,000 people, and g) places with at least 1,000 people.

Table 5: Present-Day Population Controls

	World Bank aid 1995–2014		
	All	All	PrC
Mission	0.10***	0.07***	0.18***
	(0.03)	(0.03)	(0.05)
Population in 1995 (4th-order polynomial)	Yes	Yes	Yes
Populated place dummy	No	Yes	No
Ethnic dummies	Yes	Yes	Yes
Mean dependent variable	0.25	0.25	0.62
R-squared	0.40	0.43	0.64
Observations	6,876	6,876	698

Source: Authors' analysis based on mission locations from Beach (1903); aid projects from AidData, World Bank geocoded research release, version 1.3; and control variables from multiple sources, detailed in appendix A.

Note: "PrC" denotes a sample of cells that contain the capital of a province. Estimation is by ordinary least squares. Conley (1999) standard errors are given in parentheses, with the cutoff at 2 degrees (220 km). The dependent variable is a dummy for at least one project commitment in the sample period. Control variables include: log distances to coast, explorer route, colonial railway, and Arab trade route; a third-order polynomial in the 18th-century population; presence of a city as of 1800; country dummies; precolonial ethnic dummies; average altitude; Terrain Ruggedness Index; percentage area within 10 km from water source; Caloric Suitability Index; Malaria Ecology Index; and a tropics dummy (which is also interacted with mean altitude). Cells that are more than 90 percent covered by barren land or more than 90 percent covered by forest are excluded from the sample.
$^{*}\, p < .1 \quad ^{**}\, p < .05 \quad ^{***}\, p < .01$

To assess spillover bias from neighboring cells, a regression is run on a specification where the spatial lag of missions is added as a control. The lag variable is an indicator for the presence of at least one mission in one of the (up to) eight cells surrounding cell i, referred to as the inner ring (see fig. 2 for an example).[23] The specification also includes the interaction of the lag variable with the mission dummy. The coefficient on the lag variable captures the effect of having a mission in a neighboring cell, apart from that of hosting a mission in the cell itself. The coefficient on the interaction term captures the additional effect due to the contemporaneous presence of missions both in the cell itself and in the inner ring.

The results are reported in table 6. The coefficients on the first spatial lag and its interaction with Mission_{ik} are very close to zero and insignificant at conventional levels. Furthermore, their inclusion does not affect the estimate of the main coefficient of interest (columns 2 and 3) relative to the baseline (column 1). Inclusion of a second spatial lag (an indicator of mission presence in one of the up to 16 cells surrounding the inner ring) yields similar results (reported in the appendix). This can be taken as evidence that the estimated correlations hold mostly at the local level (i.e., the cell itself), with no discernible role for spatial spillovers to surrounding cells.

2.5.4 Sensitivity Tests

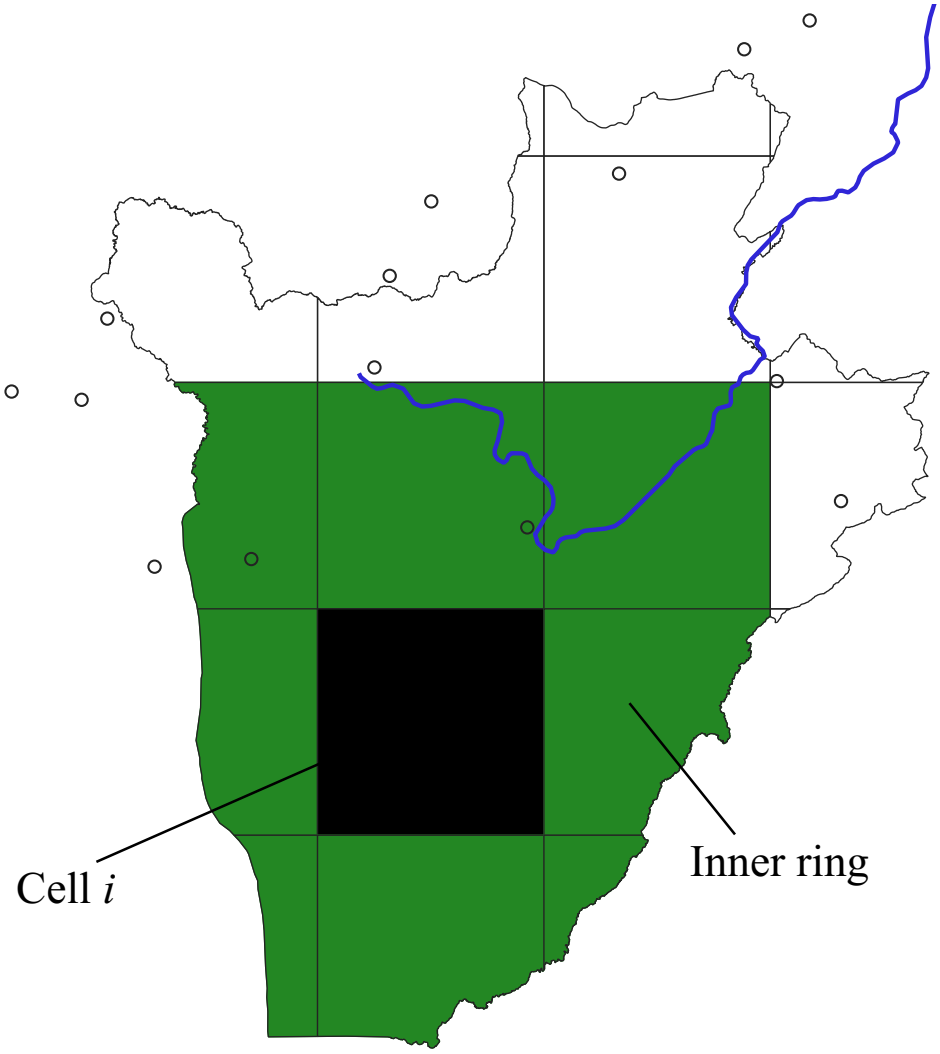
The correlation between World Bank aid and historical missions is largely robust to various other sensitivity tests, the results of which are reported in the appendix. These additional sensitivity tests are described in the following paragraphs.

The first test is of whether the estimated relationships are stable over time; it involves splitting the sample in two, defined by aid projects before and after the 2005 Paris Declaration.[24] The next test is of whether the observed

[23] Because the locations of projects funded by the World Bank are likely determined by each country's government, they are probably not correlated across country borders. Allowing for cross-border correlation would bias the results, as zero correlations between border cells would pull down the overall estimate. Cross-border spatial correlation could be relevant if missions had persistent effects on surrounding areas, especially before current borders were put in place. The inclusion of cross-border cells in the weighting matrix has virtually no impact on the results, so only the preferred specification is presented here.

[24] The Paris Declaration, signed at the Second High Level Forum on Aid Effectiveness organized by the OECD in 2005, was aimed at transferring more management and discretion to recipient countries. See http://www.oecd.org/dac/effectiveness/

Figure 2: Example of Spatial Lags in Burundi



Source: Authors' summary based on mission locations from Roome (1924).
Note: This figure illustrates how the lag variable is defined based on the presence of at least one mission in the (up to) eight neighbors adjacent to a cell.

Table 6: Spatial Lags

	World Bank aid 1995–2014		
	(1)	(2)	(3)
Mission	0.11***	0.11***	0.14***
	(0.03)	(0.03)	(0.05)
Mission lag		0.01	0.01
		(0.02)	(0.02)
Mission × Mission lag			−0.04
			(0.06)
Ethnic dummies	Yes	Yes	Yes
Mean dependent variable	0.26	0.26	0.26
R-squared	0.43	0.43	0.44
Observations	5,840	5,840	5,840

Source: Authors' analysis based on mission locations from Beach (1903); aid projects from AidData, World Bank geocoded research release, version 1.3; and control variables from multiple sources, detailed in appendix A.

Note: Mission lag refers to the (up to) eight neighbors adjacent to each cell. Estimation is by ordinary least squares. Conley (1999) standard errors are given in parentheses, with the cutoff at 2 degrees (220 km). The dependent variable is a dummy for at least one project commitment in the sample period. Control variables include: log distances to coast, explorer route, colonial railway, and Arab trade route; a third-order polynomial in the 18th-century population; presence of a city as of 1800; country dummies; precolonial ethnic dummies; average altitude; Terrain Ruggedness Index; percentage area within 10 km from water source; Caloric Suitability Index; Malaria Ecology Index; and a tropics dummy (which is also interacted with mean altitude). Cells that are more than 90 percent covered by barren land or more than 90 percent covered by forest are excluded from the sample. $^{*}\,p < .1$ $^{**}\,p < .05$ $^{***}\,p < .01$

pattern holds for aid in all sectors (health, transport, etc.). If differences are detected across sectors, this might give some hints as to the underlying mechanisms. Irrespective of the sector considered, coefficients on the mission dummies are always positive, statistically significant at conventional levels, and large relative to the means of the dependent variables (see estimates presented in the appendix); the magnitudes are highly comparable across different sectors.

A further test checks if the use of binary treatment and outcome variables is important for the results. When using the number of missions and the log number of missions plus 1 as treatment variables, results are very similar to the baseline. Another investigation is of whether the relationship between aid and missions also holds at the intensive margin; this is done by replacing the dependent variable with the number of World Bank aid projects. Estimates from OLS, Poisson, and negative binomial regressions confirm the baseline results. Finally, the correlation between aid and missions also survives the inclusion of subnational fixed effects at the ADM1 (states, governorates) or ADM2 (districts, municipalities, communes) level.

Taking stock, this section has documented a robust and sizable spatial relationship between World Bank aid and the historical presence of mission stations. Although it is hard to claim causality with observational data, the estimated relationship survives a vast array of robustness tests, including the most recent procedure suggested in the econometric literature for assessing the bias from unobservable factors (Oster, 2019).

3 Implications for Aid Effectiveness

Having established that mission areas attract a disproportionate share of World Bank aid, it is natural to ask whether this has implications for aid effectiveness. Answering this question is interesting both from a policy perspective and in terms of gaining a better understanding of the mechanism at play. For example, it is possible that missionary interventions paved the ground for aid interventions later on, by providing suitable conditions for effective project implementation. These conditions might include cooperative behavior, trust in foreigners, and specific skills (such as language), all factors that previous research has found to be positively affected by missionary activity. The goal of this section is to test whether aid projects implemented

parisdeclarationandaccraagendaforaction.htm.

in mission areas are more successful at achieving their goals. To this end, two empirical tests are conducted, one using data on project ratings and the other using survey data on development outcomes from the Demographic and Health Survey (DHS).

3.1 Project Ratings

Each World Bank project is headed by a team leader who is also responsible for evaluating its success (with respect to the stated goal) upon completion. After the initial evaluation, the World Bank's Independent Evaluation Group (IEG) performs a second assessment based on available project documentation. Furthermore, the IEG performs an additional in-depth evaluation of approximately 25 percent of the projects, which includes on-site visits and additional analyses (Denizer, Kaufmann, & Kraay, 2013). Each layer of evaluation rates the projects on a six-point scale from "highly satisfactory" to "highly unsatisfactory."

The data contain rating information on 43 percent of the projects included in the grid analysis, thus allowing tests to be run on whether projects implemented in the vicinity of missions display higher ratings. Since rating information is at the project level rather than the location level, here the analysis departs from the grid-level dataset and relies instead on a project-level dataset (recall that each project is implemented across several locations). The baseline specification to be estimated by OLS is

$$\text{Rating}_{pk} = \beta \cdot \text{MissionLocations}_{pk} + \mathbf{X}_p\gamma + \mathbf{W}_k\delta + \varepsilon_{pk}\,, \tag{2}$$

where p and k are indices for project and country, respectively.

Rating_{pk} is a binary variable that equals 1 if the rating of project p is "satisfactory" or better and equals 0 otherwise (the most recent available IEG evaluation is used, which means desk reviews in 87 percent of the cases); $\text{MissionLocations}_{pk}$ is the project's fraction of precision 1 and precision 2 locations that are within 25 km of a mission station (this radius implies observational units that roughly correspond to the size of cells in the grid structure). The selection of relevant covariates follows Denizer, Kaufmann, and Kraay (2013). In terms of country-level variables, the controls include growth in average GDP per capita over the life of the project (from the World Bank) and the sum of Freedom House scores on civil liberties and political rights.[25]

[25]Denizer, Kaufmann, and Kraay (2013) also included CPIA ratings from the World

Table 7: The World Bank's IEG Project Ratings and Missions

	IEG rating: satisfactory (0/1)					
	(1)	(2)	(3)	(4)	(5)	(6)
Fraction of locations with missions	0.02	0.02	−0.02	0.01	0.02	−0.01
	(0.07)	(0.07)	(0.09)	(0.07)	(0.07)	(0.10)
Sector dummies	No	No	No	Yes	Yes	Yes
Mean dependent variable	0.67	0.67	0.66	0.67	0.67	0.66
R-squared	0.00	0.05	0.16	0.06	0.11	0.20
Observations	324	324	188	324	324	188

Source: Authors' analysis based on mission locations from Beach (1903); aid projects from AidData, World Bank geocoded research release, version 1.3; and control variables from multiple sources, detailed in appendix A

Note: Robust standard errors are given in parentheses. The dependent variable equals 1 if the IEG rating is at least moderately satisfactory and equals 0 otherwise. Regressions are done without controls in columns 1 and 4. $^{*}\, p < .1$ $^{**}\, p < .05$ $^{***}\, p < .01$

In terms of project-level variables, the controls include project length (in years), the log of total committed funds, a dummy for new projects (versus follow-ups), a dummy for investment projects, sector dummies, share in the largest sector, and the log of completion and preparation costs relative to total committed funds.

Table 7 presents estimates from six different variants of equation (2). In column 1 the regression does not include any controls, in column 2 it includes all controls except cost variables (which are not available for many projects), and in column 3 it includes the entire control set. Columns 4–6 replicate the same specifications but also include sector fixed effects. The coefficient on the fraction of locations in the vicinity of a mission is small, and its sign is not consistent across specifications. The standard errors are at least three times as large as the coefficient. In short, table 7 shows no evidence that mission presence is correlated with better (or worse) project performance, as measured by IEG ratings.

The results in table 7 should be interpreted with some caution: the sample size is relatively small, and the explanatory variable is subject to measurement error due to the need to keep the dataset at the project level. These factors imply low statistical power to reject the null hypothesis. Furthermore, the IEG rating is an imperfect measure of project performance. It measures performance with respect to a goal, which is not the same across sectors, and it is at least partly based on documentation produced by the team leader.

Bank, but these data for the period before 2005 could not be located for the present study.

Furthermore, although formally independent, it is materially conducted by present and future World Bank employees (Denizer, Kaufmann, & Kraay, 2013). Finally, the nature of the test is descriptive, because it does not account for nonrandom locations of projects to mission areas. This means that it is possible to draw suggestive, but not definitive, conclusions about project effectiveness.

3.2 Survey Data on Development Outcomes

The second empirical strategy attempts to overcome the limitations of the first by relying on a quasi-experimental setup and on direct measures of economic development. To identify effects of World Bank aid, the strategy exploits the longitudinal dimension of the aid data, obtaining a panel of cells-years spanning the period 1995–2014. Equipped with these data, it is possible to compare development in cells that received aid at different points in time, exploiting information on the dates of project approval and completion. The regression equation to be estimated by OLS has the form

$$\begin{aligned} Y_{ikt} = {} & \kappa \cdot \text{Mission}_{ik} + \beta \cdot \text{ActiveAid}_{ikt} \cdot \text{Mission}_{ik} + \gamma \cdot \text{ActiveAid}_{ikt} \\ & + \delta \cdot \text{CompletedAid}_{ikt} \cdot \text{Mission}_{ik} + \theta \cdot \text{CompletedAid}_{ikt} \\ & + \mu \cdot \text{EverAid}_{ik} \cdot \text{Mission}_{ik} + \nu \cdot \text{EverAid}_{ik} + \lambda_{kt} + \varepsilon_{ikt} \,, \end{aligned} \tag{3}$$

where i, k, and t are indices for cell, country, and year, respectively.

The outcome variable Y_{ikt} is a measure of economic development (e.g., electrification or access to water). ActiveAid_{ikt} is a binary indicator of the presence of at least one active project in year t.[26] As the effects of aid will not necessarily materialize immediately in the years of project implementation, regressions also include $\text{CompletedAid}_{ikt}$, which is a binary indicator of the past presence of a project. (The indicator is equal to 1 in every year after completion of the first project.) EverAid_{ik} is a binary indicator that equals 1 if cell i ever received at least one World Bank project in the sample period, and Mission_{ik} is an indicator that equals 1 if at least one historical mission was located in the cell. Fixed effects at the country-year level (λ_{kt}) are always included.

Specification (3) amounts to a triple-differences setup (difference-in-differences-in-differences): EverAid_{ik}, Mission_{ik}, and their interaction control

[26] A project is defined as active if the year t is between commitment and completion.

for time-invariant differences between different categories of cells. The coefficient on ActiveAid_{ikt} captures whether Y_{ikt} is higher when a project is active (relative to periods before the arrival of aid) and β whether it is more so in cells that hosted a historical mission. The coefficient on $\text{CompletedAid}_{ikt}$ captures whether Y_{ikt} is higher after the completion of a project (relative to periods before the arrival of aid) and δ whether it is more so in cells that hosted a historical mission. The coefficients of interest are β and δ, because it is not a priori clear when the effects of aid should materialize.

It is important to stress that the main goal here is to test whether mission cells cause higher or lower aid effectiveness, and not to test whether aid is effective per se, as this is outside the scope of the article. As such, the identifying assumption is that conditional on $\text{EverAid}_{ik} = 1$, cells with and without missions have parallel trends in Y_{ikt}. Identification of β and δ (the coefficients of interest) does not require any assumption about parallel trends between cells that received aid at different points in time or between cells that ever/never received aid. These assumptions would be arguably very demanding, as they amount to saying that the timing of aid implementation is random. In contrast, the identifying assumption used here is much less stringent, and it states that development trends (not levels) between cells with and without missions as of 1903 are the same in the period 1995–2014.

To implement this strategy it is necessary to have a measure of economic development at the cell level observed at several points in time. Proxies of development are constructed from individual-level georeferenced survey data from the DHS. Specifically, this test uses the individual recode of the DHS, which includes women of reproductive age (15–49), as it has the highest country coverage (in some countries the DHS surveys only women). Four questions are selected that cover different dimensions of development and are asked consistently across countries and time. Answers are collapsed at the cell level to obtain four variables measuring the fractions of cell respondents with certain characteristics. The characteristics are the following: having piped water as the main source of drinking water,[27] access to electricity, owning a television and/or radio, and having a floor made of modern material.[28] These measures are not necessarily representative at the cell-year level, but the number of respondents is high (with cell mean 190 and median 105), and the geographical and temporal coverage is also good (1,900 cells, 4,200

[27] Other sources are worse: well water, surface water, or rainwater.

[28] Non-modern floors could be made of, for example, leaves or sand.

cell-year observations, and on average 200 cells per year).[29]

Table 8 presents estimates of equation (3) for each measure of Y_{ikt}. The odd-numbered columns serve as benchmarks, because they correspond to the simple difference-in-differences estimation (no interaction with the mission dummy). Across different outcomes, the coefficient on ActiveAid_{ikt} is positive and significant at the 99 percent level, which means that when a World Bank project is active (between commitment and completion) the outcome is between 30 percent (for electricity and piped water) and 9 percent (for radio/TV owners) higher than in years before the arrival of aid. The coefficient on $\text{CompletedAid}_{ikt}$ is also positive but smaller, and it is significant only in the case of piped water. These estimates have a causal interpretation under the assumption that the timing of aid arrival is unrelated to trends in outcomes. If this assumption does not hold, the estimates conflate the effect of aid with the selection bias (e.g., cells growing faster being able to attract more aid). Finally, the coefficient on EverAid_{ik} is positive and significant. This is consistent with past research, which has found that aid does not go to the least-developed areas. These estimates are encouraging because they suggest that the outcome variables considered here are indeed decent proxies of development at the cell-year level.

The even-numbered columns correspond to the triple differences because they also include the mission dummy and its interaction with the aid indicators. Compared with the odd-numbered columns, inclusion of the interactions does not affect the estimated coefficients on the variables already included in the odd-numbered columns. Furthermore, the coefficients on both interactions of interest do not have consistent signs across different outcomes and are almost always not statistically significant. In particular, the sign of the interaction between ActiveAid_{ikt} and Mission_{ik} is negative and insignificant for the outcome "Radio/TV" and is positive and insignificant for the other outcomes. The size of the positive coefficients is (at least 50 percent) smaller than the coefficient on ActiveAid_{ikt} alone. The sign on the interaction between $\text{CompletedAid}_{ikt}$ and Mission_{ik} is negative for the outcomes "Radio/TV" and "Proper floor" (significant in the former case) and is positive and insignificant for the other two outcomes. As in the previous case, the size of the positive coefficients is (at least 60 percent) smaller than the coefficient on $\text{CompletedAid}_{ikt}$ alone.

[29] Cell-year observations with fewer than 50 respondents are dropped; see the appendix for a map of the geographical coverage.

Table 8: Triple Differences: Aid Effectiveness and Missions

	Electricity		Piped water		Radio/TV		Proper floor	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Active aid	0.11***	0.10***	0.13***	0.12***	0.06***	0.06***	0.11***	0.10***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.01)	(0.01)	(0.02)	(0.02)
Completed aid	0.02	0.03*	0.07***	0.07**	0.02	0.02	0.03	0.04*
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)
Ever had aid	0.06**	0.05***	0.06***	0.05**	0.04***	0.04***	0.08***	0.07***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)
Mission		0.11*		0.11*		0.09***		0.08**
		(0.06)		(0.06)		(0.03)		(0.03)
Mission × Active aid		0.01		−0.03		0.02		0.02
		(0.05)		(0.05)		(0.02)		(0.04)
Mission × Completed aid		−0.04		0.02		0.00		−0.07*
		(0.06)		(0.06)		(0.03)		(0.04)
Mission × Ever had aid		0.02		0.04		−0.06**		0.06**
		(0.07)		(0.06)		(0.03)		(0.03)
Mean outcome	0.29	0.29	0.30	0.30	0.69	0.69	0.44	0.44
Number of cells	1,970	1,970	1,971	1,971	1,970	1,970	1,971	1,971
Observations	4,201	4,201	4,202	4,202	4,201	4,201	4,202	4,202

Source: Authors' analysis based on mission locations from Beach (1903); aid projects from AidData, World Bank geocoded research release, version 1.3; development outcomes from the Demographic and Health Survey; and control variables from multiple sources, detailed in appendix A.

Note: Estimation is by ordinary least squares. Standard errors are clustered at the country level. Country-year fixed effects are always included. The dependent variable is the fraction of respondents whose dwellings include the feature reported above each regression. $^{*}\, p < .1$ $^{**}\, p < .05$ $^{***}\, p < .01$

As a robustness test, a more parsimonious model is also estimated, in which ActiveAid_{ikt} and $\text{CompletedAid}_{ikt}$ are collapsed together into a single treatment variable that equals 1 in the year of commitment of the first project and in all subsequent years. The estimates of the coefficient on the interaction of interest (not reported) are again small and not significant. To sum up, the analysis in this section does not find any evidence of development outcomes being higher in cells with missions than in cells without missions, either after or during the implementation of aid projects.

There are a few caveats in this analysis. First, the exercise does not estimate a precise "zero effect," so it is not possible to conclude definitively that aid does not at all work better in mission areas. Second, using measures of actual development as outcome variables (instead of project ratings) has some disadvantages: although *access* to piped water and to electricity may be direct products of specific aid projects, the outcomes themselves measure ownership of private goods. The time span of the analysis may not be long enough for the effects of development aid to materialize into private consumption or investment. Furthermore, proxies based on DHS data may not capture the dimensions of aid that are most affected by the projects. Finally, note that if the "true" differences in aid success between cells with and without missions are small in magnitude, measurement error in several variables could result in too much attenuation bias to be able to estimate them.

3.3 Discussion

The analyses above are derived from a corollary of a potential mechanism, which is that aid is allocated to mission areas because it is thought to be more effective there. However, the results from two different empirical exercises do not lend support to this hypothesis. This conclusion rests on the assumption that the variables considered are actually able to measure effectiveness, as well as on a number of assumptions about the validity of the test. There is also an implicit assumption that the allocating authorities have a reliable way to observe effectiveness, which may not be the case. If so, the allocation of aid to mission areas could be related to a prior belief that aid will be more effective there, a belief that is not updated because of a lack of evidence.

4 Potential Mechanisms

Section 2 has documented that areas close to historical missions tend to attract more World Bank aid. The previous section tested whether this has implications for aid effectiveness and was unable to find evidence of such in the data. The present section is devoted to exploring mechanisms (or mediators) through which the location of historical missions might have affected present-day allocation of aid. The point of origin for this analysis is the literature on development aid allocation and on effects of historical missions, from which factors are identified that could potentially act as mediators.

There is widespread evidence that aid targeting depends on political considerations. Regions with ties to the party in power tend to receive more development aid and more resources in general. There is also evidence that missionary activities had long-lasting positive effects on social capital, collaborative behavior, and literacy, all factors that likely contribute to higher *de jure* and *de facto* local political power (see Section 1). To investigate the role of political ties, an indirect test is conducted that probes for the existence of a political aid cycle in mission areas.

Next, this section investigates the role of religious favoritism. Western and predominantly Christian countries stand out as the largest shareholders of the World Bank.[30] If shareholders for some reason prefer to allocate more aid to areas that are culturally similar to their own countries, one should expect to find disproportionately more World Bank projects in African regions with a large Christian population, or in areas that have strong ties with Western countries. In both cases, cells with a history of Christian missionary activity are likely to be prominent, as the presence of missions increased both Christian and Western footprints. Finally, this section investigates the specific role of education levels. Among all development dimensions affected by missionaries, human capital stands out as the single outcome for which there is the largest body of convincing causal evidence.

[30]The United States is the largest shareholder (holding 10 percent in the IDA and 16 percent in the IBRD), and, although Japan is second (with 8 percent and 7 percent, respectively), western European countries together form a sizable group: the United Kingdom, France, and Germany together account for 15 percent and 12 percent of IDA and IBRD funds, respectively; these figures are available at https://www.worldbank.org/en/about/leadership/votingpowers/.

4.1 Political Aid Cycle

If mission areas attract more aid because of stronger political connections, these connections might be expected to break down in the event of government turnover. Turnover is a relatively rare phenomenon in African countries, in both autocratic and democratic societies (in the data, the median country has two turnovers in 20 years). Against this backdrop, political connections are both durable and valuable. In order to test for the existence of a differential political aid cycle, an annual balanced panel of cells covering the period 1995–2014 is constructed, and data on national elections are compared with the timings of cell-level aid commitments. The equation to be estimated by OLS is

$$\begin{aligned} \text{Aid}_{ikt} = \beta \cdot \text{Election}_{kt} \cdot \text{Mission}_{ik} + \gamma \cdot \text{Turnover}_{kt} \cdot \text{Mission}_{ik} + \\ \lambda_{kt} + \mu_i + \varepsilon_{ikt} \,, \end{aligned} \tag{4}$$

where i, k, and t are indices for cell, country, and year, respectively.

The variable Aid_{ikt} is an indicator that equals 1 if at least one aid project is committed in the corresponding calendar year with at least one project location in cell i; Election_{kt} is a binary indicator that equals 1 if there is a national election for the office of head of state; and Turnover_{kt} is an indicator that equals 1 if the election results in a change in the person holding office. Turnover is defined as a change in the head of state, irrespective of party affiliation.[31] The data on elections come from the Varieties of Democracy (V-DEM) dataset, version 7.1 (Coppedge, Gerring, Lindberg, Skaaning, Teorell, Altman, Bernhard, Fish, Glynn, Hicken, et al., 2017). Fixed effects at the country-year level (λ_{kt}) and at the cell level (μ_i) are included in all regressions.

Note that the inclusion of country-year fixed effects means that the specification cannot identify a general political aid cycle at the country level, which is beyond the scope of this article. It does, however, allow credible testing for the presence of a political aid cycle specific to mission cells. The coefficient β captures whether mission cells are more likely to receive aid in election years when the incumbent is re-elected, and γ captures the additional effect in case of turnover. The presence of cell fixed effects gives a difference-in-differences interpretation of the parameters of interest, β and γ.[32]

[31] Estimates are virtually identical if turnover is defined as a change of party; the results are available from the authors upon request.

[32] This difference-in-differences setup is admittedly somewhat unconventional, due to the

Table 9 reports estimates from several variants of equation (4) (always including cell and country-year fixed effects). In column 1, the regression includes only the interaction between Election_{kt} and Mission_{ik}; the coefficient is very small and is not significant at conventional levels, suggesting that election years are not different from other years in terms of aid arrival in mission cells. In column 2, the specification includes only the interaction between Turnover_{kt} and Mission_{ik}; the coefficient is negative and significant at the 95 percent level, implying that the probability of receiving a new World Bank aid project is 40 percent lower in years when the election results in a change in the head of state. The same coefficient is again negative and significant in column 3, where the regression includes both interactions together. In this case, the coefficient on $\text{Election}_{kt} \cdot \text{Mission}_{ik}$ becomes larger but is still insignificant—which is to say, there is no evidence of any reduction or increase in aid in years when the election results in a victory of the incumbent head of state.

One remaining threat to identification stems from the cross-sectional correlation of Mission_{ik} with several covariates. To account for this, columns 4–6 of table 9 replicate the first three regressions but also include the same historical and geographical correlates of missions as the baseline regression, equation (1), interacted with Election_{kt} and/or Turnover_{kt}. The coefficient on $\text{Turnover}_{kt} \cdot \text{Mission}_{ik}$ is again negative and significant at least at the 95 percent level. The coefficient on $\text{Election}_{kt} \cdot \text{Mission}_{ik}$ becomes bigger and significant in column 6. The estimates from the last regression imply that the probability of aid arrival in mission cells increases by 40 percent in election years when the incumbent is re-elected, and decreases by approximately the same amount when the election leads to a turnover in the head of state. Both effects are relative to non-election years due to the inclusion of cell fixed effects.

The cyclical pattern exhibited by aid to mission areas suggests that political ties to the central government may be relevant in explaining the correlation uncovered in Section 2. One interpretation consistent with the evidence is the following: Areas close to historical missions are able to develop better ties with the ruling head of state (or with his party) over time. The political connection gives these areas an advantage in the competition to attract aid projects. But when there is political turnover, the connection breaks

inclusion of two treatments in the same equation. However, separate regressions were also run for both treatments.

Table 9: Political Aid Cycle in Mission Areas

	At least one World Bank project committed in cell-year					
	(1)	(2)	(3)	(4)	(5)	(6)
Election × Mission	−0.00		0.01	0.01		0.03**
	(0.01)		(0.01)	(0.01)		(0.01)
Turnover × Mission		−0.03**	−0.04**		−0.03**	−0.05***
		(0.01)	(0.02)		(0.01)	(0.02)
Country-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Cell fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	Yes	Yes	Yes
Mean dep. var. of non-mission cells	0.03	0.03	0.03	0.03	0.03	0.03
Mean dep. var. of mission cells	0.07	0.07	0.07	0.07	0.07	0.07
Number of cells	6,819	6,819	6,819	6,811	6,811	6,811
Observations	136,380	136,380	136,380	136,220	136,220	136,220

Source: Authors' analysis based on mission locations from Beach (1903); aid projects from AidData, World Bank geocoded research release, version 1.3; elections from Varieties of Democracy; and control variables from multiple sources, detailed in appendix A.
Note: Estimation is by ordinary least squares. Results are from a balanced annual panel over 20 years. Standard errors are clustered at the country level. Controls included in columns 4–6 are the following variables interacted with the Election and/or Turnover dummy: log distances to coast, explorer route, colonial railway, and Arab trade route; a third-order polynomial in the 18th-century population; presence of a city as of 1800; country dummies; precolonial ethnic dummies; average altitude; Terrain Ruggedness Index; percentage area within 10 km from water source; Caloric Suitability Index; Malaria Ecology Index; and a tropics dummy (which is also interacted with mean altitude). $^{*}\,p < .1$ $^{**}\,p < .05$ $^{***}\,p < .01$

down and these areas experience a temporary drop in aid. The (less robust) increase in election years without turnover could constitute a reward from the incumbent in exchange for political support, but the instability of the estimate makes it difficult to draw firm conclusions.

The findings in table 9 cannot be explained by a general slowdown of project commitments due to government officials being busy with the electoral campaign, because the regressions include country-year fixed effects. One might argue that the effect is driven by turnover between heads of state who belong to different religions, which would imply religious favoritism as opposed to political favoritism. However, this interpretation is feasible only if all (or most) turnovers are from a Christian to a non-Christian head of state. A qualitative check of this explanation using multiple sources (e.g., Encyclopedia Britannica and Wikipedia) suggests that most turnovers happen between individuals of the same religion, making this explanation implausible. Yet another possible explanation has to do with heads of state coming disproportionately from mission cells (or having had a missionary education). However, to be consistent with the findings, this interpretation would also require (most) turnovers to happen between individuals with different

missionary backgrounds and in one specific direction, which seems unlikely. Furthermore, Dreher et al., 2019 tested and rejected the hypothesis that African presidents channel a disproportionate amount of World Bank aid to their birth regions by using a dataset with approximately the same temporal and geographical coverage as the present study.

4.2 The Role of Christian Religion and of Human Capital

Human capital and religion are potential mediators in the relationship between aid and historical missions, although investigation of this is challenging, as is any mediation analysis. Simply including proxies for the candidate mechanism in the baseline equation (1) is problematic for causal inference; if these variables really are mediators, they are by definition not predetermined with respect to Mission_{ik}, which leads to a bad control problem (Angrist & Pischke, 2008). In the absence of a quasi-experimental strategy like the one in Section 4.1, the analysis has to proceed with caution, as interpretation of results relies on several assumptions. First, an assumption is made on the direction of the selection bias that is introduced by controlling for a "post-treatment" variable M_{ik} (education or religion). Then the equation

$$\text{EverAid}_{ik} = \beta \cdot \text{Mission}_{ik} + \vartheta \cdot \text{M}_{ik} + \delta_k + \mathbf{X}_{ik}\gamma + \varepsilon_{ik} \tag{5}$$

is estimated with or without M_{ik}. When M_{ik} is included, and assuming that Mission_{ik} positively (respectively, negatively) affects M_{ik} and that M_{ik} positively affects aid allocation, it is possible to interpret the coefficient β as a lower (respectively, an upper) bound on the true effect of Mission_{ik} on EverAid_{ik}.[33] If the estimates of β from the regressions with and without M_{ik} are not significantly different from each other, one can conclude that M_{ik} is not the main mechanism of interest.

Individual-level georeferenced data from the individual recode of the DHS (the same data source as used in Section 3) is used to construct measures of education and Christian religion. The DHS asks respondents (women between the ages of 15 and 49) about their highest level of education achieved: no education, primary, secondary, or tertiary. Three measures of education at the cell level are constructed: the average level of education, the fraction of respondents with at least primary education, and the fraction of respondents

[33] This is the same logic used when controlling for present-day population.

with at least secondary education. As in Section 3, cells with fewer than 50 respondents are dropped, so the final sample contains 2,264 cells, each with more than 400 respondents on average (median 185). The DHS also includes a question about respondents' religious beliefs, which is used to obtain the fraction of respondents of Christian religion (any denomination). The geographical coverage is slightly smaller than for education, with a sample of 2,083 cells.[34] Consistent with the literature discussed in Section 1, historical missions are assumed to have had positive effects on both education and prevalence of Christianity.

The first column of table 10 reports results of an estimation of equation (1) on the subsample of cells for which a measure of education from the DHS is available. The coefficient on the mission dummy has similar size and precision to the full-sample results. In relative terms, however, the correlation is smaller, owing to the larger incidence of aid projects in this subsample of cells (50 percent compared to 25 percent in table 3).

Columns 2–4 present results for models with different measures of education: average level, fraction with at least primary education, and fraction with at least secondary education. The coefficients on all three measures of education are positive, significant at the 99 percent level, and large. Inclusion of either measure reduces the size of the coefficient on Mission_{ik}. Secondary education seems to be most important here; it has the largest coefficient, and its inclusion almost halves the mission coefficient. Under the assumption that missionary activities have positive effects on education, the coefficient on Mission_{ik} is a lower bound on the true effect. This lower bound ranges from 60 to 90 percent of the main effect in column 1. This suggests that a heritage of high human capital could be one reason for more aid going to mission areas. However, there seems to be room for other explanations as well, since the coefficient on Mission_{ik} is not reduced to zero.

In column 5, the baseline specification is estimated in the subsample of cells for which a DHS measure of religion is available. Although religion seems to play an independent role in aid allocation, the coefficient on Mission_{ik} does not change when a measure of Christian religion is included in the regression. This means that there is no evidence in the data that Christian favoritism is the mechanism behind the correlation uncovered in Section 2.

[34] The question about religion is not included in all countries for all rounds.

Table 10: Education, Christian Religion and Mission Areas

	World Bank aid 1995–2014					
	(1)	(2)	(3)	(4)	(5)	(6)
Mission	0.10**	0.07*	0.09**	0.06	0.08*	0.08*
	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)
Average education level		0.44***				
		(0.05)				
At least primary education (share)			0.70***			
			(0.09)			
At least secondary (share)				0.92***		
				(0.10)		
Christians (share)						0.14**
						(0.07)
Ethnic dummies	Yes	Yes	Yes	Yes	Yes	Yes
Mean dependent variable	0.50	0.50	0.50	0.50	0.51	0.51
Mean M		0.76	0.54	0.20		0.52
R-squared	0.43	0.46	0.46	0.46	0.45	0.45
Observations	2,264	2,264	2,264	2,264	2,083	2,083

Source: Authors' analysis based on mission locations from Beach (1903); aid projects from AidData, World Bank geocoded research release, version 1.3; measures of human capital and religion from the Demographic and Health Survey (DHS); and control variables from multiple sources, detailed in appendix A.

Note: Estimation is by ordinary least squares. Conley (1999) standard errors are given in parentheses, with the cutoff at 2 degrees (220 km). The dependent variable is a dummy for at least one project commitment in the sample period. Control variables include: log distances to coast, explorer route, colonial railway, and Arab trade route; a third-order polynomial in the 18th-century population; presence of a city as of 1800; country dummies; precolonial ethnic dummies; average altitude; Terrain Ruggedness Index; percentage area within 10 km from water source; Caloric Suitability Index; Malaria Ecology Index; and a tropics dummy (which is also interacted with mean altitude). Cells that are more than 90 percent covered by barren land or more than 90 percent covered by forest are excluded from the sample. $^{*}\, p < .1$ $^{**}\, p < .05$ $^{***}\, p < .01$

4.3 Discussion

Secondary education seems to partly explain why mission areas get more aid, although the evidence is not conclusive. Even if these results are taken at face value, it is not obvious why higher levels of education should make areas attractive targets for aid. It could be that some level of human capital is necessary for successful implementation of projects, in which case this finding would be indirect evidence of the "effectiveness" hypothesis examined in Section 3. On the other hand, if the evidence of a political aid cycle is interpreted as favoritism (as argued here), human capital could play the role of a necessary condition for the formation of political connections. These are certainly questions that could be investigated empirically and which merit further research.

5 Conclusion

This article has documented that 19th- and 20th-century Christian missionary activity in Africa predicts the location of present-day World Bank aid allocation. The findings suggest that the probability of receiving development projects is about 40 percent higher in areas that contain a historical Christian mission station. This result is the main contribution of the article, although attempts have been made to explain the source of the correlation and to explore policy implications.

First, the article tested whether aid effectiveness is higher in areas that hosted historical missions. As missionary activity had long-lasting positive effects on social and human capital, it is possible that donors channel aid to such areas in the hope that these endowments will facilitate successful implementation of the projects. Using data on project ratings and survey-based development indicators, the empirical analysis in this article was unable to find evidence of superior aid effectiveness in areas exposed to early missionary activity. These results are not conclusive, however, due to data limitations.

Second, the article investigated three potential mechanisms: political favoritism, religious favoritism, and education. Education is positively correlated with aid allocation, and its inclusion among the controls in the baseline regression reduces the coefficient on missions. There is no evidence of religious favoritism, but there is evidence of a political aid cycle specific to mission areas: aid commitments are reduced whenever a presidential turnover

occurs. This finding can be interpreted as evidence that mission areas have better connections with the central government, giving them an advantage in the competition to attract development projects. Electoral turnover breaks these ties temporarily, whereas incumbent re-election seems to strengthen them. The conclusion is that political favoritism is likely to play a role, but the analysis is not able to rule out other mediating factors.

The correlations reported in this article generate a number of questions, leaving ample room for further research. The attempts to test for aid effectiveness were not able to establish conclusive evidence, and this issue deserves more attention. Moreover, although the analysis has provided evidence of a political aid cycle, it was only possible to make conjectures about why it exists. Also, how the aid cycle actually matters for long-term patterns of aid allocation is not clearly understood. Lastly, the list of potential mechanisms discussed in this article is not exhaustive, and there may well be other explanations that were not considered, and therefore not tested, in the present study.

References

Alesina, A., & Dollar, D. (2000). Who gives foreign aid to whom and why? *Journal of Economic Growth*, *5*(1), 0005074v1, 33–63. https://doi.org/10.1023/a:1009874203400

Alpino, M., & Hammersmark, E. M. (2017). *Lighting the Path: The Influence of Historical Christian Missions on Modern-day Development Aid Allocation in Africa* (Working paper No. 47). AidData.

Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of Political Economy*, *113*(1), 151–184.

Andersen, T. B., Bentzen, J., Dalgaard, C.-J., & Sharp, P. (2017). Pre-reformation Roots of the Protestant Ethic. *The Economic Journal*, *127*(604), 1756–1793.

Angrist, J. D., & Pischke, J. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ, Princeton University Press.

Beach, H. P. (1903). *A geography and atlas of protestant missions: Their environment, forces, distribution, methods, problems, results and prospects at the opening of the twentieth century* (Vol. 2). New York, NY, Student Volunteer Movement For Foreign Missions.

Brice, W. C. (2001). *An historical atlas of Islam* (H. Kennedy, Ed.). Leiden, Netherlands, Brill Academic Publishers.

Briggs, R. C. (2014). Aiding and abetting: Project aid and ethnic politics in Kenya. *World Development*, *64*, 194–205. https://doi.org/10.1016/j.worlddev.2014.05.027

Briggs, R. C. (2017). Does foreign aid target the poorest? *International Organization*, *71*(1), 187–206.

Cagé, J., & Rueda, V. (2016). The long-term effects of the printing press in sub-Saharan Africa. *American Economic Journal: Applied Economics*, *8*(3), 69–99.

Cagé, J., & Rueda, V. (2017). *Sex and the mission: The conflicting effects of early Christian investments on the HIV epidemic in sub-Saharan Africa* (CEPR Discussion paper No. 12192). Centre for Economic Policy Research.

Caicedo, F. V. (2018). *The Mission*: Human Capital Transmission, Economic Persistence, and Culture in South America. *The Quarterly Journal of Economics*, *134*(1), 507–556.

Calvi, R., & Mantovanelli, F. G. (2018). Long-term effects of access to health care: Medical missions in colonial India. *Journal of Development Economics*, *135*, 285–303. https://doi.org/https://doi.org/10.1016/j.jdeveco.2018.07.009

Castelló-Climent, A., Chaudhary, L., & Mukhopadhyay, A. (2018). Higher education and prosperity: Catholic missionaries to luminosity in India. *The Economic Journal*, *128*(616), 3039–3075.

Chandler, T. (1987). *Four thousand years of urban growth*. Lewiston, NY, Edwin Mellen Press.

Chen, Y., Wang, H., & Yan, S. (2014). *The long-term effects of Protestant activities in China* (MPRA Paper No. 53531). Munich University Library. https://mpra.ub.uni-muenchen.de/53531/1/MPRA_paper_53531.pdf

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, *92*(1), 1–45. https://doi.org/10.1016/S0304-4076(98)00084-0

Coppedge, M., Gerring, J., Lindberg, S. I., Skaaning, S.-E., Teorell, J., Altman, D., Bernhard, M., Fish, M. S., Glynn, A., Hicken, A., Et al. (2017). V-Dem Dataset v7.

Denizer, C., Kaufmann, D., & Kraay, A. (2013). Good countries or good projects? Macro and micro correlates of World Bank project performance. *Journal of Development Economics*, *105*, 288–302.

Dreher, A., Fuchs, A., Hodler, R., Parks, B. C., Raschky, P. A., & Tierney, M. J. (2019). African leaders and the geography of China's foreign assistance. *Journal of Development Economics*, *140*, 44–71.

Dreher, A., Sturm, J. E., & Vreeland, J. R. (2009). Development aid and international politics: Does membership on the UN Security Council influence World Bank decisions? *Journal of Development Economics*, *88*(1), 1–18. https://doi.org/10.1016/j.jdeveco.2008.02.003

Fearon, J. D., & Laitin, D. D. (2003). Ethnicity, insurgency, and civil war. *American Political Science Review*, *97*(1), 75–90.

Galiani, S., Knack, S., Xu, L. C., & Zou, B. (2017). The effect of aid on growth: Evidence from a quasi-experiment. *Journal of Economic Growth*, *22*(1), 1–33. https://doi.org/10.1007/s10887-016-9137-4

Gallup, J. L., Sachs, J. D., & Mellinger, A. D. (1999). Geography and economic development. *International Regional Science Review*, *22*(2), 179–232. https://doi.org/10.1177/016001799761012334

Galor, O., & Özak, Ö. (2016). The agricultural origins of time preference. *American Economic Review*, *106*(10), 3064–3103.

Gennaioli, N., & Rainer, I. (2007). The modern impact of precolonial centralization in Africa. *Journal of Economic Growth*, *12*(3), 185–234. https://doi.org/10.1007/s10887-007-9017-z

Imbens, G. W., & Rubin, D. B. (2015). *Causal inference for statistics, social, and biomedical sciences: An introduction*. Cambridge, England, Cambridge University Press.

Jablonski, R. S. (2014). How aid targets votes: The impact of electoral incentives on foreign aid distribution. *World Politics*, *66*(2), 293–330.

Jedwab, R., Kerby, E., & Moradi, A. (2017). History, path dependence and development: Evidence from colonial railroads, settlers and cities in Kenya. *The Economic Journal*, *137*(603), 1467–1494. https://doi.org/10.1111/ecoj.123

Jedwab, R., Meier zu Selhausen, F., & Moradi, A. (2018). *The Economics of Missionary Expansion: Evidence from Africa and Implications for Development* (Working paper No. 7). Centre for the Study if African Economies.

Jedwab, R., & Moradi, A. (2016). The permanent effects of transportation revolutions in poor countries: Evidence from Africa. *The Review of*

Economics and Statistics, *98*(2)arXiv arXiv:1011.1669v3, 268–284. https://doi.org/10.1162/REST

Johnson, H. B. (1967). The location of Christian missions in Africa. *Geographical Review*, *57*(2), 168–202. https://doi.org/10.2307/213158

Kiszewski, A., Mellinger, A. D., Spielman, A., Malaney, P., Sachs, S. E., & Sachs, J. D. (2004). A global index representing the stability of malaria transmission. *American Journal of Tropical Medicine and Hygiene*, *70*(5), 486–498. https://doi.org/70/5/486[pii]

Kudo, Y. (2017). Missionary influence on marriage practices: Evidence from the Livingstonia mission in Malawi. *Journal of African Economies*, *26*(3), 372–431.

Mantovanelli, F. (2014). *Christian missions, HIV and sexual behavior in sub-Saharan Africa* [Unpublished manuscript]. Retrieved September 12, 2017, from https://sites.google.com/site/fedmantovanelli/home/research

Masaki, T. (2018). The political economy of aid allocation in Africa: Evidence from Zambia. *African Studies Review*, *61*(1), 55–82.

Meier zu Selhausen, F. (2014). Missionaries and female empowerment in colonial Uganda: New evidence from Protestant marriage registers, 1880–1945. *Economic History of Developing Regions*, *29*(1), 74–112. https://doi.org/10.1080/20780389.2014.927110

Michalopoulos, S., Naghavi, A., & Prarolo, G. (2018). Trade and geography in the spread of Islam. *The Economic Journal*, *128*(616), 3210–3241.

Michalopoulos, S., & Papaioannou, E. (2013). Pre-colonial ethnic institutions and contemporary African development. *Econometrica*, *81*(1), 113–152. https://doi.org/10.3982/ECTA9613

Murdock, G. P. (1959). *Africa: Its peoples and their culture history*. New York, NY, McGraw-Hill.

Nunn, N. (2008). The long-term effects of Africa's slave trades. *The Quarterly Journal of Economics*, *123*(1), 139–176.

Nunn, N. (2010). Religious conversion in colonial Africa. *American Economic Review: Papers & Proceedings*, *100*(2), 147–152. https://doi.org/10.1257/aer.100.2.147

Nunn, N. (2014). Gender and missionary influence in colonial Africa. In E. Akyeampong, R. H. Bates, N. Nunn, & J. A. Robinson (Eds.), *Africa's development in historical perspective* (pp. 489–512). Cambridge, England, Cambridge University Press.

Nunn, N., & Puga, D. (2012). Ruggedness: The blessing of bad geography in Africa. *The Review of Economics and Statistics*, *94*(1), 20–36.

Nunn, N., & Wantchekon, L. (2011). The slave trade and the origins of mistrust in Africa. *American Economic Review*, *101*(7), 3221–3252. https://doi.org/10.1257/aer.101.7.3221

Nunnenkamp, P., Öhler, H., & Andrés, M. S. (2017). Need, merit, and politics in multilateral aid allocation: A district-level analysis of World Bank projects in India. *Review of Development Economics*, *21*(1), 126–156. https://doi.org/10.1111/rode.12259

Okoye, D., & Pongou, R. (2017). *Sea changes: The transatlantic slave trade and missionary activity in Africa* [Unpublished manuscript]. Retrieved December 6, 2017, from https://sites.google.com/site/dozieaokoye/home/research

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, *37*(2), 187–204.

Robinson, C. H. (1915). *History of Christian missions*. Edinburgh, Scotland, Clark.

Roome, W. J. W. (1924). *Ethnographic survey of Africa: Showing the tribes and languages*. London, England, Stanford.

Strandow, D., Findley, M., Nielson, D., & Powell, J. (2011). *The UCDP and AidData codebook on georeferencing aid: Version 1.1* (UCDP Paper No. 4). Uppsala University. http://docs.aiddata.org/ad4/files/ucdp_aiddata_codebook_published.pdf

Waldinger, M. (2017). The long-run effects of missionary orders in Mexico. *Journal of Development Economics*, *127*, 355–378. https://doi.org/10.1016/j.jdeveco.2016.12.010

World Bank Group. (2013). *A stronger, connected, solutions World Bank Group: An overview of the World Bank Group strategy* (tech. rep.). World Bank Group. Washington, DC.

Appendix A Data Used for the Analyses

Table A.1: Data Sources

Data	Source	Link	Access date
World Bank projects	AidData	www.aiddata.org/	2016-06-22
World Bank project docs	World Bank	projects.worldbank.org/	2016-06-22
Missions in 1903	Beach (1903), Cagé and Rueda (2016)		
Missions in 1924	Nunn (2010), Roome (1924)	scholar.harvard.edu/nunn/	2016-06-15
Country borders	GADM	www.gadm.org/	2016-12-22
Coastlines	Natural Earth	www.naturalearthdata.com/	2017-02-20
Rivers	Natural Earth	www.naturalearthdata.com/	2017-02-20
Explorer routes	Century Company; Nunn (2010)	scholar.harvard.edu/nunn/	2016-12-13
Colonial railways	Century Company; Nunn (2010)	scholar.harvard.edu/nunn/	2016-12-13
Gridded elevation data	United States Geological Survey	topotools.cr.usgs.gov/	2016-05-23
Caloric Suitability Index	Galor and Özak (2016)	www.omerozak.com/	2016-10-05
Water sources	WorldGeoDatasets (fee)	www.worldgeodatasets.com/	2016-05-20
Malaria Ecology Index	Kiszewski et.al (2004)	www.gordonmccord.com/	2016-12-13
18th-century population	HYDE	www.pbl.nl/hyde/	2016-03-08
Historical cities	Chandler (1987)	www.worldcitypop.com/	2016-04-05
Ethnic groups	Murdock (1959), Nunn (2008)	scholar.harvard.edu/nunn/	2016-03-14
Arab trade routes	Brice (2001)	referenceworks.brillonline.com/	2016-09-03
Population in 1995	SEDAC	sedac.ciesin.columbia.edu	2016-02-06
Populated places	WorldGeoDatasets (fee)	www.worldgeodatasets.com/	2016-06-15
DHS variables	USAID DHS Program	www.dhsprogram.com/	2016-04-08
Elections	V-DEM	www.v-dem.net	2017-11-23
GDP growth	World Bank	data.worldbank.org/	2018-07-18
Rights and liberties	Freedom House	freedomhouse.org	2018-07-18

Source: Authors' summary of the sources of data used for the study.
Note: This table lists the data sources for all the variables considered in the analyses of this article.

Table A.2: Frequencies of Other Locations of Projects Included in the Sample

	World Bank (%)
Projects with only precision 1 or precision 2 locations	15.7
Projects with also precision 3 and precision 4 locations	18.2
Projects with also precision 3 but not precision 4 locations	20.4
Projects with also precision 4 but not precision 3 locations	23.8
Projects with only precision 1, 2, or 6 locations	11.2
Projects with only precision 1, 2, or 8 locations	8.3
Residual category	2.2

Source: Authors' summary of data from World Bank project documents and personal communication with project managers if additional details were required.

Note: The World Bank sample is composed of 768 projects, and the China sample consists of 800 projects. If locations cannot be retrieved from donor documents, AidData checks recipient country documents and aid management systems, or information from the websites of implementing agencies. Locations may be towns, hills, farms, or other geographical features. The coders then search for coordinates in geographical databases such as Geonames and Google Earth. If the name of a specific location cannot be matched with a set of coordinates, coders look for nearby towns or other identifiable features. The geocoding of the World Bank data is based on the same methodology as for the UCDP Georeferenced Event Dataset, described in detail in Strandow, Findley, Nielson, and Powell (2011).

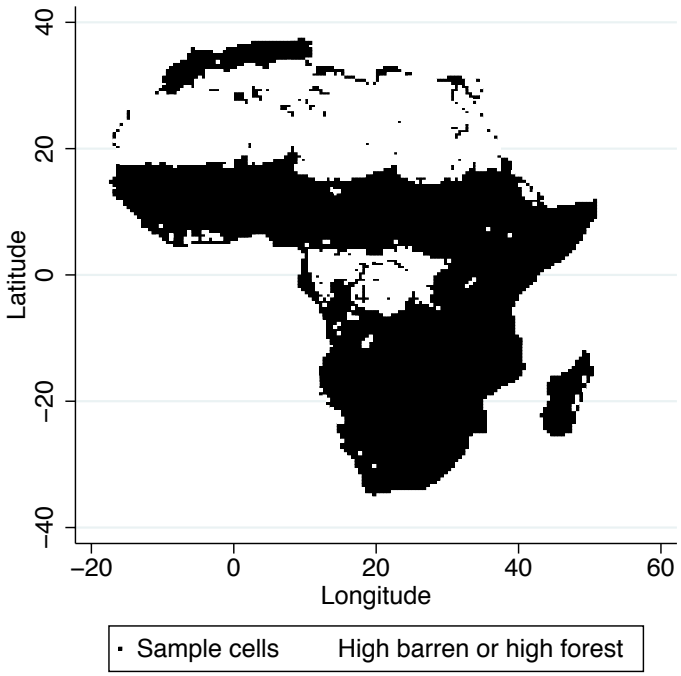
Table A.3: Comparison of World Bank Projects in Sample and Excluded World Bank Projects

	Not in sample	In sample	Difference
Commitments (millions USD)	60.55	71.73	−11.18
Disbursements (millions USD)	23.25	29.83	−6.58
Start year	2006.62	2005.76	0.86*
End year	2011.81	2011.52	0.29
Length (in years)	5.95	6.56	−0.61***
Repeater (0/1)	0.27	0.27	−0.00
Largest sector (%)	73.01	76.07	−3.06*
Completion cost (%)	1.43	1.21	0.22
Supervision cost (%)	2.57	2.52	0.05
IEG: satisfactory (0/1)	0.66	0.67	−0.01
Investment (0/1)	0.98	0.98	−0.00
Agriculture (%)	16.56	7.33	9.24***
Public Admin. (%)	20.33	20.52	−0.18
ICT (%)	0.19	1.95	−1.77**
Education (%)	11.96	6.16	5.80***
Finance (%)	2.41	1.95	0.47
Health (%)	28.46	11.16	17.30***
Energy (%)	4.01	13.76	−9.74***
Transport (%)	6.39	19.18	−12.80***
Water (%)	6.56	14.54	−7.98***
Industry & trade (%)	3.12	3.45	−0.33
No. observations	300	768	

Source: Authors' analysis based on data from World Bank project documents.

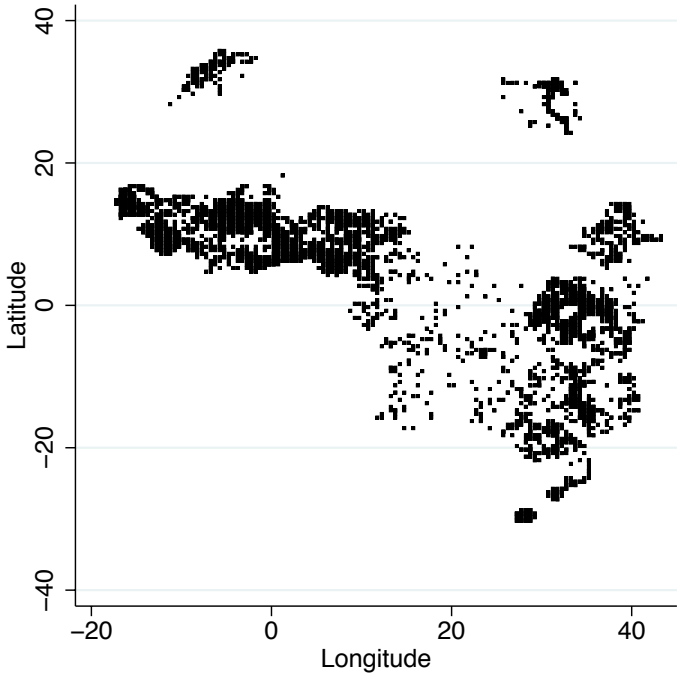
Note: Projects in sample are those with at least one precision 1 or precision 2 location; projects not in sample have at least one precision 3 or precision 4 location but no locations at precision 1 or precision 2. $^{*}\,p < .1$ $^{**}\,p < .05$ $^{***}\,p < .01$

Figure A.1: Gridded Sample



Source: Authors' summary based on information from the HYDE database (www.pbl.nl/hyde/).
Note: The blank spaces represent cells covered by barren land or by forest for more than 90 percent of their surface area in the 19th century.

Figure A.2: DHS Samples



Source: Authors' summary based on data from the USAID DHS Program (www.dhsprogram.com/).
Note: Black cells are those included in the analysis that rely on DHS data.

Appendix B Supplementary Analyses

Table B.1: World Bank Aid and Missions from Beach (1903) in Two Decades

	Ever had World Bank aid in period	
	1995–2004	2005–2014
Mission	0.09***	0.13***
	(0.02)	(0.03)
Ethnic dummies	Yes	Yes
Mean dep. var.	0.15	0.19
R-squared	0.39	0.36
N	6,876	6,876

Source: Authors' analysis based on mission locations from Beach (1903); aid projects from AidData, World Bank geocoded research release, version 1.3; and control variables from multiple sources, detailed in table A.1.

Note: Conley (1999) standard errors are given in parentheses, with the cutoff at 2 degrees (220 km). Estimation is by ordinary least squares. The dependent variable is a dummy for at least one World Bank project commitment in each period. Control variables include: log distances to coast, explorer route, colonial railway, and Arab trade route; a third-order polynomial in the 18th-century population; presence of a city as of 1800; country dummies; precolonial ethnic dummies; average altitude; Terrain Ruggedness Index; percentage area within 10 km from water source; Caloric Suitability Index; Malaria Ecology Index; and a tropics dummy (which is also interacted with mean altitude). Cells that are more than 90 percent covered by barren land or more than 90 percent covered by forest are excluded from the sample. $^{*}\,p < .1$ $^{**}\,p < .05$ $^{***}\,p < .01$

Table B.2: World Bank Aid and Missions from Beach (1903): Non-Binary Treatment and Outcome

	Ever had World Bank aid		Number of World Bank projects			
	OLS	OLS	OLS	OLS	Poisson	NB
	(1)	(2)	(3)	(4)	(5)	(6)
No. of missions	0.05***					
	(0.01)					
ln(No. of missions)		0.13***				
		(0.03)				
Mission dummy			1.39***	2.34***	0.83***	0.73***
			(0.33)	(0.75)	(0.11)	(0.07)
Ethnic dummies	Yes	Yes	Yes	Yes	No	No
Mean dep. var.	0.25	0.25	0.92	3.63	0.92	0.92
R-squared	0.39	0.39	0.48	0.51		
N	6,876	6,876	6,876	1,737	6,884	6,884

Source: Authors' analysis based on mission locations from Beach (1903); aid projects from AidData, World Bank geocoded research release, version 1.3; and control variables from multiple sources, detailed in table A.1.

Note: OLS = ordinary least squares; NB = non-binary. Conley (1999) standard errors are given in parentheses, with the cutoff at 2 degrees ($\approx$ 200 km) in columns 1–4. Robust standard errors are reported in columns 5 and 6. In columns 1 and 2 the dependent variable is a dummy for at least one World Bank project commitment in 1995–2014. In columns 3–6 the dependent variable is the number of aid commitments. In column 4 the sample is restricted to cells with at least one project commitment. Control variables are the following, but without ethnic dummies in columns 5 and 6: log distances to coast, explorer route, colonial railway, and Arab trade route; a third-order polynomial in the 18th-century population; presence of a city as of 1800; country dummies; precolonial ethnic dummies; average altitude; Terrain Ruggedness Index; percentage area within 10 km from water source; Caloric Suitability Index; Malaria Ecology Index; and a tropics dummy (which is also interacted with mean altitude).

* $p < .1$ ** $p < .05$ *** $p < .01$

Table B.3: World Bank Aid and Missions from Beach, 1903: Present-Day Population Controls

	All cells	All cells	Populated place		Provincial capital	
	(1)	(2)	(3)	(4)	(5)	(6)
Mission	0.10***	0.07***	0.07	0.10**	0.18***	0.21***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.05)	(0.04)
Population (1995)	0.01***	0.01***				
	(0.00)	(0.00)				
Population^2	−0.80***	−0.52***				
	(0.11)	(0.10)				
Population^3	0.00***	0.00***				
	(0.00)	(0.00)				
Population^4	−0.00***	−0.00***				
	(0.00)	(0.00)				
Populated place dummy	No	Yes	No	No	No	No
Ethnic dummies	Yes	Yes	Yes	No	Yes	No
Mean dep. var.	0.25	0.25	0.49	0.49	0.62	0.62
R-squared	0.40	0.43	0.59	0.34	0.64	0.32
N	6876	6876	1168	1168	698	698

Source: Authors' analysis based on mission locations from Beach (1903); aid projects from AidData, World Bank geocoded research release, version 1.3; and control variables from multiple sources, detailed in table A.1.

Note: Conley (1999) standard errors are given in parentheses, with the cutoff at 2 degrees (220 km). Estimation is by ordinary least squares. The dependent variable is a dummy for at least one project commitment in the sample period. Control variables include: log distances to coast, explorer route, colonial railway, and Arab trade route; a third-order polynomial in the 18th-century population; presence of a city as of 1800; country dummies; precolonial ethnic dummies; average altitude; Terrain Ruggedness Index; percentage area within 10 km from water source; Caloric Suitability Index; Malaria Ecology Index; and a tropics dummy (which is also interacted with mean altitude). Cells that are more than 90 percent covered by barren land or more than 90 percent covered by forest are excluded from the sample. The estimation samples in columns 3 and 4 are restricted by the presence of a populated place with at least 10,000 inhabitants, and the samples in columns 5 and 6 are restricted to locations with a provincial capital. $^{*}\, p < .1$ $^{**}\, p < .05$ $^{***}\, p < .01$

Table B.4: World Bank Aid and Missions from Beach (1903) and Roome (1924): Specifications without Controls

	World Bank aid 1995–2014			
	(1)	(2)	(3)	(4)
Catholic mission (Roome)	0.46***		0.56***	
	(0.04)		(0.04)	
Protestant mission (Roome)	0.45***			
	(0.03)			
Protestant mission (Beach)		0.44***	0.36***	
		(0.04)	(0.04)	
Any mission				0.46***
				(0.03)
Controls	No	No	No	No
Ethnic dummies	No	No	No	No
Country dummies	No	No	No	No
Mean dep. var.	0.25	0.25	0.25	0.25
R-squared	0.05	0.01	0.03	0.04
N	6892	6892	6892	6892

Source: Authors' analysis based on mission locations from Beach (1903) and Roome (1924); aid projects from AidData, World Bank geocoded research release, version 1.3; and control variables from multiple sources, detailed in table A.1.

Note: Conley (1999) standard errors are given in parentheses, with the cutoff at 2 degrees (220 km). The dependent variable is a dummy for at least one project commitment in the sample period. $^{*}p < .1$ $^{**}p < .05$ $^{***}p < .01$

Table B.5: World Bank Aid and Missions from Beach, 1903: Spatial Lags

	World Bank aid 1995–2014				
	(1)	(2)	(3)	(4)	(5)
Mission	0.11***	0.11***	0.14***	0.11***	0.09
	(0.03)	(0.03)	(0.05)	(0.03)	(0.06)
Mission Lag1		0.01	0.01	0.00	0.01
		(0.02)	(0.02)	(0.02)	(0.03)
Mission × Mission Lag1			−0.04		0.09
			(0.06)		(0.10)
Mission Lag2				−0.02	−0.01
				(0.02)	(0.02)
Mission Lag1 × Mission Lag2					−0.01
					(0.04)
Mission × Mission Lag2					0.10
					(0.09)
Mission × Mission Lag1 × Mission Lag2					−0.19
					(0.12)
Ethnic dummies	Yes	Yes	Yes	Yes	Yes
Mean dep. var.	0.26	0.26	0.26	0.26	0.26
R-squared	0.43	0.43	0.44	0.44	0.44
N	5,840	5,840	5,840	5,840	5,840

Source: Authors' analysis based on mission locations from Beach (1903); aid projects from AidData, World Bank geocoded research release, version 1.3; and control variables from multiple sources, detailed in table A.1.

Note: Conley (1999) standard errors are given in parentheses, with the cutoff at 2 degrees (220 km). The dependent variable is a dummy for at least one project commitment in the sample period. Mission Lag1 refers to the (up to) eight neighbors adjacent to each cell. Mission Lag2 refers to the next (up to) 16 closest outer neighbors. Control variables include: log distances to coast, explorer route, colonial railway, and Arab trade route; a third-order polynomial in the 18th-century population; presence of a city as of 1800; country dummies; precolonial ethnic dummies; average altitude; Terrain Ruggedness Index; percentage area within 10 km from water source; Caloric Suitability Index; Malaria Ecology Index; and a tropics dummy (which is also interacted with mean altitude). Cells that are more than 90 percent covered by barren land or more than 90 percent covered by forest are excluded from the sample. * $p < .1$ ** $p < .05$ *** $p < .01$

Table B.6: Aid and Missions from Beach (1903) and Roome (1924): Specifications with ADM1 Fixed Effects

	World Bank aid 1995–2014						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Catholic mission (Roome)	0.22***		0.20***			0.21***	
	(0.04)		(0.04)			(0.04)	
Protestant mission (Roome)		0.13***	0.11***				
		(0.02)	(0.02)				
Any mission (Roome)				0.16***			
				(0.02)			
Protestant mission (Beach)					0.10***	0.09***	
					(0.03)	(0.03)	
Any mission							0.14***
							(0.02)
Ethnic dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ADM1 dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep. var.	0.25	0.25	0.25	0.25	0.25	0.25	0.25
Oster bound	0.11	0.07		0.09	0.10		0.10
R-squared	0.39	0.40	0.40	0.40	0.39	0.40	0.40
N	6,796	6,796	6,796	6,796	6,796	6,796	6,796

Source: Authors' analysis based on mission locations from Beach (1903) and Roome (1924); aid projects from AidData, World Bank geocoded research release, version 1.3; and control variables from multiple sources, detailed in table A.1.

Note: Conley (1999) standard errors are given in parentheses, with the cutoff at 2 degrees (220 km). The dependent variable is a dummy for at least one project commitment in the sample period. Control variables include: log distances to coast, explorer route, colonial railway, and Arab trade route; a third-order polynomial in the 18th-century population; presence of a city as of 1800; country dummies; precolonial ethnic dummies; average altitude; Terrain Ruggedness Index; percentage area within 10 km from water source; Caloric Suitability Index; Malaria Ecology Index; and a tropics dummy (which is also interacted with mean altitude). Cells that are more than 90 percent covered by barren land or more than 90 percent covered by forest are excluded from the sample. The lower bound on the coefficient of interest is calculated as in Oster (2019): the R-squared from the hypothetical regression on Mission_{ik} and both observed and unobserved controls is set to 1.3 times the R-squared from the actual regression on Mission_{ik} and the observed control. $^{*}\, p < .1$ $^{**}\, p < .05$ $^{***}\, p < .01$

Table B.7: Aid and Missions from Beach (1903) and Roome (1924): Specifications with ADM2 Fixed Effects

	World Bank aid 1995–2014						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Catholic mission (Roome)	0.27***		0.25***			0.26***	
	(0.05)		(0.05)			(0.05)	
Protestant mission (Roome)		0.14***	0.11***				
		(0.03)	(0.03)				
Any mission (Roome)				0.17***			
				(0.03)			
Protestant mission (Beach)					0.09***	0.08***	
					(0.03)	(0.03)	
Any mission							0.15***
							(0.02)
Ethnic dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ADM2 dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Oster bound	0.11	0.07		0.09	0.10		0.10
Mean dep. var.	0.25	0.25	0.25	0.25	0.25	0.25	0.25
R-squared	0.39	0.40	0.40	0.40	0.39	0.40	0.40
N	6796	6796	6796	6796	6796	6796	6796

Source: Authors' analysis based on mission locations from Beach (1903) and Roome (1924); aid projects from AidData, World Bank geocoded research release, version 1.3; and control variables from multiple sources, detailed in table A.1.

Note: Conley (1999) standard errors are given in parentheses, with the cutoff at 2 degrees (220 km). The dependent variable is a dummy for at least one project commitment in the sample period. Control variables include: log distances to coast, explorer route, colonial railway, and Arab trade route; a third-order polynomial in the 18th-century population; presence of a city as of 1800; country dummies; precolonial ethnic dummies; average altitude; Terrain Ruggedness Index; percentage area within 10 km from water source; Caloric Suitability Index; Malaria Ecology Index; and a tropics dummy (which is also interacted with mean altitude). Cells that are more than 90 percent covered by barren land or more than 90 percent covered by forest are excluded from the sample. The lower bound on the coefficient of interest is calculated as in Oster (2019): the R-squared from the hypothetical regression on Mission_{ik} and both observed and unobserved controls is set to 1.3 times the R-squared from the actual regression on Mission_{ik} and the observed control. $^{*}\, p < .1$ $^{**}\, p < .05$ $^{***}\, p < .01$

Table B.8: Aid by Sector and Missions from Beach (1903)

	Ever received World Bank aid with major sector				
	(1)	(2)	(3)	(4)	(5)
	Agriculture	Public admin.	Infrastructure	Education	Finance
Protestant mission (Dennis et al.)	0.093***	0.121***	0.033**	0.053***	0.040***
	(0.023)	(0.025)	(0.013)	(0.016)	(0.014)
Mean dep. var.	0.103	0.221	0.030	0.050	0.024
Adjusted R-squared	0.282	0.318	0.212	0.210	0.299
No. of observations	6,876	6,876	6,876	6,876	6,876

	Ever received World Bank aid with major sector				
	(1)	(2)	(3)	(4)	(5)
	Health	Energy	Transport	Water	Industry
Protestant mission (Dennis et al.)	0.081***	0.057***	0.105***	0.075***	0.049***
	(0.019)	(0.021)	(0.024)	(0.021)	(0.019)
Mean dep. var.	0.133	0.091	0.154	0.105	0.067
Adjusted R-squared	0.288	0.251	0.276	0.249	0.283
No. of observations	6,876	6,876	6,876	6,876	6,876

Source: Authors' analysis based on mission locations from Beach (1903); aid projects from AidData, World Bank geocoded research release, version 1.3; and control variables from multiple sources, detailed in table A.1.

Note: Conley (1999) standard errors are given in parentheses, with the cutoff at 2 degrees (220 km). The dependent variable is a dummy for at least one project commitment in the sample period. In each column, only projects in the sector at the top of the column are considered. Control variables include: log distances to coast, explorer route, colonial railway, and Arab trade route; a third-order polynomial in the 18th-century population; presence of a city as of 1800; country dummies; precolonial ethnic dummies; average altitude; Terrain Ruggedness Index; percentage area within 10 km from water source; Caloric Suitability Index; Malaria Ecology Index; and a tropics dummy (which is also interacted with mean altitude). Cells that are more than 90 percent covered by barren land or more than 90 percent covered by forest are excluded from the sample. $^{*}\, p < .1 \quad ^{**}\, p < .05 \quad ^{***}\, p < .01$

Table B.9: Aid by Main Sector and Missions from Beach (1903)

	Ever received World Bank aid with main major sector				
	(1)	(2)	(3)	(4)	(5)
	Agriculture	Public admin.	Infrastructure	Education	Finance
Protestant mission (Dennis et al.)	0.052***	0.035**	0.017**	0.029**	0.029***
	(0.018)	(0.016)	(0.008)	(0.011)	(0.011)
Mean dep. var.	0.058	0.055	0.014	0.017	0.012
Adjusted R-squared	0.232	0.240	0.149	0.224	0.396
No. of observations	6876	6876	6876	6876	6876

	Ever received World Bank aid with main major sector				
	(1)	(2)	(3)	(4)	(5)
	Health	Energy	Transport	Water	Industry
Protestant mission (Dennis et al.)	0.050***	0.059***	0.065***	0.059***	0.024*
	(0.016)	(0.020)	(0.021)	(0.019)	(0.013)
Mean dep. var.	0.059	0.070	0.129	0.059	0.017
Adjusted R-squared	0.249	0.245	0.263	0.202	0.142
No. of observations	6876	6876	6876	6876	6876

Source: Authors' analysis based on mission locations from Beach (1903); aid projects from AidData, World Bank geocoded research release, version 1.3; and control variables from multiple sources, detailed in table A.1.

Note: Conley (1999) standard errors are given in parentheses, with the cutoff at 2 degrees (220 km). The dependent variable is a dummy for at least one project commitment in the sample period. In each column, only projects in the sector at the top of the column are considered. Control variables include: log distances to coast, explorer route, colonial railway, and Arab trade route; a third-order polynomial in the 18th-century population; presence of a city as of 1800; country dummies; precolonial ethnic dummies; average altitude; Terrain Ruggedness Index; percentage area within 10 km from water source; Caloric Suitability Index; Malaria Ecology Index; and a tropics dummy (which is also interacted with mean altitude). Cells that are more than 90 percent covered by barren land or more than 90 percent covered by forest are excluded from the sample. $^{*}\, p < .1$ $^{**}\, p < .05$ $^{***}\, p < .01$