Lighting the path: The influence of historical Christian missions on modern-day development aid allocation in Africa[a]

Matteo Alpino[b] Eivind M. Hammersmark[c]

October 5, 2017

Abstract

Recent studies suggest that development aid is not always directed towards the poorest areas within countries, contrary to what we would expect if the objective is to end poverty. In this paper we document a positive correlation between the location of historical Christian missions and the within-country allocation of World Bank and Chinese financed development projects in Africa. The correlation is conditioned on observable geographical and historical factors that have shaped missionaries' settlement decisions. Christian missions often included health and education facilities, and we consider them to be ancestors of modern micro-development projects. We interpret our finding as evidence of historical path dependence in the spatial distribution of aid, which can be rationalized by assuming decreasing unit cost of implementing aid at the local level. We do not find evidence that the results are driven by donor preference for Christian or Westernized areas.

Keywords: Development Aid, Christian Missions, Historical Path Dependence.
JEL classification: F35, I3, N37, N77, O19.

1 Introduction

Where does foreign development aid go? This question is of central importance in the aid effectiveness debate. The World Bank has the explicit goal to end extreme poverty, and to focus on the poorest segment on the population (World Bank Group, 2013). This strategy would suggest that aid allocation should be targeted to the poorest regions first, but several empirical

[a]The paper was previously circulated as "On a mission from God: The historical roots of development aid allocation".

[b]Department of Economics. University of Oslo, PO-Box 1095, Blindern, 0317 Oslo, Norway. E-mail: alpino.mtt@gmail.com

[c]Corresponding author. Department of Economics. University of Oslo, PO-Box 1095, Blindern, 0317 Oslo, Norway. E-mail: eivindhammers@gmail.com

studies seem to suggest that it is instead biased by political and strategic considerations. For example, Alesina and Dollar (2000) find colonial history and co-voting in the UN to be major predictors of donor-recipient foreign aid flows. Along the same lines, Dreher, Sturm, and Vreeland (2009) show that the World Bank allocates disproportionately more development projects to countries during their tenure as temporary members of the UN security council. Most of the early research on this subject has focused on the cross-country and across-time dimensions; 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 distribution of development aid (Dreher et al., 2016; Jablonski, 2014; Masaki, 2014).

We add to this literature by studying the role of historical accidents in shaping the present-day within-country allocation of aid. In particular, we argue that the development efforts initiated by 19th and 20th century Christian missionaries laid the foundation for future development aid in Africa, by decreasing the cost of implementing additional aid projects in the same location. Consequently, the post-war emergence of the aid industry has disproportionately benefited areas with a missionary history.

In the empirical analysis we compare a georeferenced snapshot of all mission stations in Africa in 1903 to the precise locations of World Bank funded projects in 1995–2014. The unit of analysis is derived from a grid of 55×55 km square cells covering the African mainland and Madagascar.[1] 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 %. We take several empirical measures to alleviate concerns of omitted variables bias. First, we control for country dummies in all specifications, because the first step of aid allocation is at the country level. Second, we address the non-random selection of missionaries into specific locations. To this end we always exclude areas covered by desert or dense forest, and control for historical and geographical factors that guided the missionaries' settlement decisions according to historical sources. Controls capture cultivation potential, accessibility, 19th century western contact, historical settlement patterns and the pre-colonial presence of Islam. Third, we show that the correlation is robust when restricting the sample to areas that are more likely to be similar: areas that intersect the ocean coastline or one of the main rivers, and subsamples obtained by propensity score matching. Fourth, we show that the link between historical missions and aid survives also when controlling flexibly for present-day population density. Finally, the test developed by Oster (2017) 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.

We interpret our findings as evidence of path dependence, i.e. long-term persistence in the location of specific activities. Although development aid in its present form is a relatively recent phenomenon, similar activities implemented by westerners in developing countries began much earlier.[2] In particular, they can be traced back to the work by Christian missionaries, who

[1] 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). Results available on request.

[2] The beginning of modern development aid coincides with the establishment of the World Bank in 1944, and

were particularly active at the end of the 19th century. The first wave of Christian missionaries reached Africa before the European colonization of the continent, and established a permanent presence in the form of mission stations, which was expanded in Colonial age. The missionary effort was primarily driven by proselytism motives, but it was not restricted to conversion. Missions provided the locals with a wide range of education and health services, primarily to boost the odds of conversion. Missionaries were quite explicit about the instrumental role of these investments. For example Robinson (1915), in his *History of Christian Missions*, devotes a whole chapter to how building schools and hospitals is by far the most effective way to convert locals to Christianity. The belief in the instrumental role of medicine led to the establishment of special medical schools designed to provide missionaries with medical training (R. Johnson, 2010). Many missionaries served instead as teachers in missionary schools, where they taught not only the Bible, but also secular subjects, and provided job training. All in all, we believe that mission stations can be considered as the ancestors of modern micro-development projects. Our interpretation is consistent with the recent literature in economics which has found positive long-term effect of missionary interventions on a variety of development outcomes (discussed in section 1.1.2).

Evidence of path dependence in the spatial distribution of development aid parallels the well-know evidence of path dependence in the location of production activities (e.g. the Manufacturing Belt in the U.S.). The new economic geography theory explains path dependence in the location of manufacturing firms by assuming increasing returns to scale at the firm-level (Krugman, 1991b). Economies of scale force firms to produce in a limited number of locations. When choosing their preferred production location, firms would like to establish in high-demand regions, and serve the rest of the market from there in order to minimize transport costs. In turn, high-demand regions are exactly those where other firms chose to locate. In this set-up, history matters and the initial spatial distribution of firms perpetuates over time.[3]

A similar pattern of geographical path dependence would equivalently be generated by decreasing unit cost. There are reasons to believe that the implementation of development aid is indeed subject to decreasing unit costs at the local level, for example due to the presence of initial fixed costs. Setting up a development project in a village where others have been implemented in the past is likely to be substantially cheaper than doing so in a "greenfield" village. First, returning to the same location allows for exploiting existing material infrastructures (wells, roads, buildings etc.) built during previous interventions. Second, continued interaction between locals and aid workers (missionaries or not) is also likely to enhance mutual trust and effective cooperation. Third, and along the same lines, continued investment in the same location enables the accumulation of region-specific knowledge, so as to decrease information costs. Finally, complex development projects may require that the target population has already acquired particular

the launch of the US-sponsored Marshall Plan in 1948, aimed at reconstructing European economies after WWII.

[3]For example Krugman (1991a) uses this type of model to illustrates the case of the US Manufacturing belt. The first bulk of US manufacturing firms located in the Midwest region stretching from New York to the Great Lakes when this area was the main agriculture center, but this production cluster remained in the same location for decades after farming activities shifted elsewhere.

skills during previous interventions. In this framework, the presence of Christian missions is likely to have permanently shifted downwards the cost of implementing development projects in the regions where missionaries first settled. We provide some indirect evidence of the presence of decreasing unit costs by showing that development projects closer to historical missions have higher absorption capacity, i.e. the disbursement of funds is faster than elsewhere. This finding suggests that projects close to missions proceed faster, and are able to unlock subsequent payments earlier compared to projects elsewhere.

We contrast the path dependence hypothesis with the most likely competing explanation: that the allocation of aid is shaped by favoritism and strategic considerations. The most influent World Bank members are western and predominantly Christian countries; as such, World Bank aid may favor locations close to Christian missions because of their cultural proximity to Western values, and their higher Christian share of population, as shaped by the missionary work. In contrast to this hypothesis we find the same robust correlation when analyzing the distribution of Chinese-financed aid, which is unlikely to be prone to Christian and Western biases. Furthermore, the spatial correlation between aid projects and historical missions is robust when controlling for the Christian share of population in the area, which in itself is a predictor of aid location.

1.1 Literature

Our paper connects the literature on determinants of development aid allocation with studies on the long-term effects of missionary activity, both relatively new branches of development economics. The following review substantiates our argument for path dependence in aid, by presenting evidence that existing conditions and some level of development are prerequisites for receiving aid, and by describing some ways in which the presence of missions has influenced long-term development, and thereby contributed to fulfilling these prerequisites. A secondary objective is to give the reader some knowledge of the observed consequences of Christian missionary activity, in order to contextualize our empirical results. To the best of our knowledge, we are the first to provide evidence that historical events affect present day aid allocation.

1.1.1 Determinants of aid allocation

Most research on aid allocation falls within two categories; papers that investigate whether poverty targeting is successful, and those that explicitly examine whether political targeting occurs. Our preferred mechanism relates to the first category, in that existing conditions reduce the cost of implementing aid projects. Evidence of political targeting supports an alternative hypothesis, in which mission stations play a political (or religious) role in attracting aid projects. This literature has largely pivoted towards within-country studies, thanks to the availability of new granular data at the sub-national level.

Several studies explore whether aid is directed to poor areas, as we should expect if poverty alleviation is the primary objective. Francken, Minten, and Swinnen (2012) study relief aid following a cyclone in Madagascar. They find that while some effort is devoted to the affected areas, some NGO and foreign funded projects are located in less affected, but more accessible areas. Importantly, their findings on the importance of accessibility are consistent with aid displaying decreasing unit costs at the local level. Nunnenkamp, Öhler, and Andrés (2017) provide evidence from India that health, water and transportation aid does target poor areas, whereas aid within energy and agriculture is not related to poverty. They further show that World Bank infrastructure projects are located in areas with higher foreign direct investment, at the cost of not reaching the poorest areas. Briggs (2017) investigates the localization of African Development Bank (AfDB) and World Bank (WB) projects in Africa. The evidence show a near monotonic relationship between wealth percentile and share of total aid allocation. The results from these papers are consistent with our hypothesis about a path dependence in aid, in that more developed areas are likely to provide better conditions for efficient project implementation.

The recent literature also provides evidence that domestic political considerations have considerable weight in the within country allocation of aid. Dreher et al. (2016) find that Chinese aid is disproportionately allocated to the birth region of African presidents, while there is no such effect for World Bank funded projects. A paper by Jablonski (2014) shows that AfDB and WB aid to Kenya in the period 1992–2010 was diverted away from areas with strong support for the opposition, and Briggs (2014) confirms this pattern, using data on project aid from individual donors in 1989–1995. The opposite pattern emerges in Zambia, however; Masaki (2014) shows that Zambian political elites have funneled AfDB and WB aid away from their political base and towards areas where the opposition has been strong. The apparent inconsistencies between these results can likely be explained by the importance of clientelistic ethnic politics in Kenya (Chabal & Daloz, 1999). These findings raise the possibility that a correlation between missionary activity and development aid may be a result of a political or religious bias.

1.1.2 Long-term effects of missionary activity

Economist are becoming increasingly aware of the long-term effects of historical Christian missions. Nunn (2010) shows that Christian missionaries have been successful in their primary target: conversion. Using Afrobarometer survey data and the location of historical missions, he documents that individuals whose ethnic group has been in contact with missions are more likely to be Christian today. There is also ubiquitous evidence that historical Christian missions have had a positive impact on current education attainment of the people in the surrounding areas. Nunn (2014) finds that the locations of Catholic and Protestant missions in Africa is associated with higher average educational attainment, but that the effect from Catholic missions only appears in males. Using historical records from Uganda on women born between 1880 and 1945, Meier zu Selhausen (2014) finds that Protestant mission education improved female literacy, and that being employed at a mission station had a significant positive impact on women's position within the household. Okoye (2015) exploit the fact that missionary activity in Nigeria was partly aimed at ending the slave trade, and find that although slave trade had a total negative effect on schooling, it had a positive indirect effect through attracting missionaries.

Mantovanelli (2014) documents the educational legacy of Protestant missionary work in India. He finds that districts that had at least one Protestant mission at the beginning of the 20th century have a more literate population today. The effect is not present in districts with Catholic missions. Using the same data, Castelló-Climent, Chaudhary, and Mukhopadhyay (2017) use Catholic missions as an instrumental variable for tertiary education in a regression on luminosity. Their first stage displays a positive and significant correlation. Interestingly, the same first-stage relationship does not appear for Protestant missions, which seems inconsistent with previous findings. Castelló-Climent et al. (2017) reconcile the results by pointing to the Protestant concern for basic literacy. Similarly, Chen, Wang, and Yan (2014) document a positive effect of protestant missions on sub-national GDP in China, working partly through education and health. Using historical floods and droughts to instrument for the likelihood of attracting missionaries, they estimate the elasticity of current GDP with respect the 1920 convert share of population to be between 0.1 and 0.2. In a recent paper, Waldinger (2017) exploits the fact that the missionary expansion from Mexico city followed the direction of initial missionaries, giving plausibly exogenous variation in missionary activity. He shows that the historical presence of Mendicant missions has led to persistent increases in current educational attainment and literacy.

In the literature reviewed so far there seems to be consistent denominational differences in the long-term effects of missionary activity. Gallego and Woodberry (2010) argues that part of this differential can be explained by differences in competition; whereas Catholic missions in Africa were protected from competitors, Protestant missions faced competitive pressure, leading them to increase the provision of education.

There is also evidence that missions have improved health outcomes. Calvi and Mantovanelli (2016) investigates the long-run effects of missionary health investment in India, comparing missions with and without hospitals. They find persistent positive effects on current individuals' body mass index. Their findings are not explained by persistence of physical infrastructure, but rather by improved health habits and practices.

Furthermore, several papers have investigated the effects of missions on attitudes and behavioral outcomes. Caicedo (2017) looks at the effect of the Jesuit expansion and subsequent expulsion in southern parts of Latin America. Through surveys and behavioral experiments he finds that people in these areas exert substantially higher collaborative behavior today. The effect seems to be context dependent, however. Using data from Nigeria, Okoye (2015) presents evidence of a negative relationship of missionary activity on collaborative behavior. He argues that the erosion of traditional institutions that missionaries deemed incompatible with Christian values led to societies with lower levels of trust.

There is also evidence that Christian values have changed particular traditional behaviors. For example, Kudo (2017) find that missionary-educated women marry later, and are less likely to marry a polygamous husband. Cagé and Rueda (2017) and Mantovanelli (2014) find that individuals living close to historical missions more often have negative attitudes towards condom use, and that this explains higher HIV prevalence in these areas. Finally, Cagé and Rueda (2016) focus on missionary investments in a specific technology, rather than in education or health. They find that the printing press imported by Protestant missionaries to print Bibles facilitated the birth of newspapers, and is still important in explaining newspaper diffusion today, which in turn results in higher social capital.

2 Empirical strategy and data

2.1 Empirical strategy

Our empirical strategy exploits spatial variation in historical missionary stations and present-day aid allocation across Africa. Since the location of missions is predetermined and does not vary over time, we collapse the temporal dimension of aid allocation to get a cross section dataset. The unit of observation is contiguous grid cells at a resolution of 0.5×0.5 degrees, which at the equator roughly corresponds to 55×55 kilometers. We intersect the grid with the African continent, and keep the part of the grid covering continental Africa. Cells are split at borders, in order to make sure that aid projects are geographically assigned to the correct country.[4] Our baseline specification to be estimated via OLS is

$$EverAid_{ik} = \beta \cdot Mission_{ik} + \delta_k + \mathbf{X}_{ik}\gamma + \varepsilon_{ik}, \tag{1}$$

where i and k are indexes for cell and country, respectively. $EverAid_{ik}$ is an indicator variable equal to one if the grid cell had at least one active aid project in the period of study, and zero otherwise. $Mission_{ik}$ is an indicator equal to one if at least one historical mission was located there, and zero otherwise.

The baseline estimation exploits within country variation only. The inclusion of country dummies δ_k is important, because the first step of the aid allocation procedure is at the country level.[5] The vector $\mathbf{X}_{ik}$ contains control variables at the grid cell level, including a number of historical and geographical factors, described in section 2.2.2. The error term ε_{ik} is allowed to be spatially correlated within a radius of 220 kilometers from the centroid of each cell, using the procedure in Conley (1999).

Coefficient β has a causal interpretation if $\mathbf{X}$ 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. In the following we discuss possible sources of bias, and how we deal with them.

[4] Information on borders is from the Global Administrative Areas dataset (GADM, available at http://www.gadm.org). We further assign the cells to the first two within-country administrative levels from the GADM dataset (region and district), by matching the centroid of each cell with the ADM in which it falls.

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

The most obvious threat to a causal interpretation of β is the non-random selection of mission stations. Missions are predetermined with respect to present-day aid, but they may have located in areas that were more suitable for missionary work, and areas where missionaries could survive and be self-sustaining. If these areas for some reason, other than mission station presence, are more or less likely to be selected for aid projects today, β will be biased, or the correlation may be entirely spurious. For example, we may believe that missionaries were more likely to establish in relatively more populated areas, in order to convert as much people as possible. If these areas are more populated also today, we would expect them to be large recipients of aid, for reasons unrelated to the presence of missionaries. We have quite detailed information about the determinants of the location of mission stations from historical sources, notably from H. B. Johnson (1967) and Robinson (1915). Moreover, there is an increasing amount of detailed and spatially disaggregated historical data available, so we are able to plausibly control for most of these determinants by using existing data sources. Section 2.2.2 discusses in detail the data we use to deal with this source of bias.

A second source of bias is the lack of common support in the distribution of the control variables. Even if our 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 compared to cells without a historical mission (the control observations), and if the control function is misspecified (Imbens & Rubin, 2015). The African continent is geographically and historically diverse, and so are single African countries. Although we use only within country variation, i.e. use cells within a country as control groups for treated cells within the same country, the risk of non-common support remains. We have a number of strategies to tackle this problem.

First, the sample always excludes cells covered by more than 90 % barren land and cells that are more than 90 % covered by forest from our sample. These cells are likely special with respect to both aid allocation and the potential for sustaining a permanent missionary presence.[6]

Second, we experiment with restricting the sample to coastal cells, to those that intersect one of the main African rivers,[7] and both. There is evidence that areas close to the coast and to navigable rivers are more developed and more densely populated (Gallup, Sachs, & Mellinger, 1999), so these cells are likely to form a more homogeneous sample. At the same time, these areas were the most accessible for the missionaries, who came predominantly by sea, and often penetrated inland by navigating rivers upstream (H. B. Johnson, 1967). In addition to consolidating the distributions of observed covariates, these subsamples likely make mission and non-mission cells more comparable in *unobserved* covariates, reducing the potential for selection bias.

Third, we restrict the analysis to subsamples obtained using propensity score matching. In particular, we construct three different balanced subsamples of our data: one where the matching

[6]The specific choice of threshold at 90 % is chosen to exclude a significant mass in the right tail of the distributions.

[7]Nile, Niger, Senegal, Zambezi, Congo and its attributes Ubangi and Kasai.

has no geographical restrictions, and two where we match within countries to account for the country-level nature of World Bank aid. To obtain the first, we estimate a logit model on the full sample using $Mission_{ik}$ as dependent variable, and as predictors: $\mathbf{X}_{ik}$, its interactions, its squared terms, and country dummies.[8] We then match mission cells with their nearest (statistical) neighbor without replacement using the predicted values as propensity scores. For the second sample, we re-estimate propensity scores with the same logit model, but without country dummies; we then perform the same matching procedure, but restricting the pool of possible matches to cells that belong to the same country only. For the third sample, we re-estimate propensity scores separately country-by-country (using the logit model but without interactions and squared terms because the country-samples are smaller), and we again match country-by-country. Imbens (2015) stresses that different matching strategies are equally valid as long as they tackle the lack of common support in the covariate distribution. He suggests a procedure similar to ours for those studies where the number of control observations is way higher than the number of treated ones, which is the case in our application.

Finally, a third source of bias derives from the spatial nature of our data. If mission stations are clustered, we may be overestimating the impact of a single mission station on aid allocation in a given cell, since the density of missions in nearby cells may be similarly important. We deal with this by accounting for spatial lags in mission stations.

2.2 Data

2.2.1 Aid

Data on foreign aid is based on World Bank projects, geocoded by AidData (version 1.3), and on Chinese financed aid projects, geocoded by Strange, Dreher, Fuchs, Parks, and Tierney (2015) and Dreher et al. (2016) (version 1.1.1).[9] These are the only geocoded datasets covering the whole African continent (and beyond) over a pluriannual period of time. Although we collapse the data to a cross-section, substantial temporal coverage ensures that results are not driven by aid in a given year.

The World Bank dataset contains projects from both the International Bank for Reconstruction and Development (IBRD) and International Development Association (IDA), in total 1,092 projects in Africa, split across 5,079 different locations. The IBRD provides low-to-zero interest rate loans to sufficiently credit worthy countries, whereas the IDA gives loans to poorer and less credit worthy countries. 12 % of IDA funds are given as grants not to be paid back. Both types of lending is accompanied by technical assistance from the Bank, and projects are monitored by Bank agents.

The information available at the project-level includes: the original World Bank identifier, project title, date of approval, expected date of completion, total net disbursement, sector (e.g. Finance, Transportation, Health...), lending instrument, implementing agency, total committed

[8]We do not include ethnic dummies, because this prevents the logit model from converging.
[9]Both sources are available from AidData at http://www.aiddata.org.

Table 1: Precision of aid locations

	World Bank aid	Chinese aid
Precision 1 – specific place	41 %	39 %
Precision 2 – within 25km of specific place	2 %	4 %
Precision 3 – municipality (ADM1)	26 %	8 %
Precision 4 – province (ADM2)	20 %	11 %
Precision 5 – imprecise	2 %	2 %
Precision 6 – country-wide projects	4 %	19 %
Precision 8 – state or national capitals	5 %	17 %

amount (in US dollars). The Chinese financed aid dataset contains 1,952 projects in Africa, split across 1,308 different locations, and the information at the project-level is similar to that available for the World Bank dataset.

The data contains information on all sites in which a given project has been implemented. The geocoding of the World Bank data is based on UCDP methodology.[10] The source of location information is World Bank project documents, or personal communication with project managers if additional detail is required. If locations cannot be retrieved from donor documents, AidData checks recipient country documents and aid management systems, or information from implementing agency websites. Locations may be towns, hills, farms, or other geographical features. The coders then search for coordinates in geographical databases like 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 Chinese aid data is similarly geocoded, but the underlying project information is collected based on the "Tracking Underreported Financial Flows" methodology, due to a low degree of transparency for Chinese aid. Coders identify projects using media reports, and supporting information is then collected via aid information management systems operated by recipients, Chinese donor institution websites, web scraping and journal article search.[11]

Each location is categorized according to the precision of its coordinates: precision 1 locations correspond to a specific place (e.g. village, hill, bridge, railroad station...) and precision 2 locations are identified within a 25 kilometer distance from the reported coordinates. The remaining precision codes (3–8) include aid that is either not successfully geocoded, too imprecise, or given to municipalities, provinces or the entire country. These project locations are too coarse to be useful in our grid-level analysis, so we drop them from the sample. Table 1 shows the distribution of location precision codes. On average, a World Bank project is split across 46 locations, while the median project is split across 28 locations.

[10] Described in detail at http://aiddata.org/sites/default/files/ucdp_aiddata_codebook_published.pdf.
[11] Detailed information at http://china.aiddata.org/content/methodology.

2.2.2 Missionary stations

We rely on two different historical sources to retrieve information on the location of Christian missionary stations. Our preferred source is the *Geography and Atlas of Christian Missions* (Beach, 1903), digitized by Cagé and Rueda (2016).[12] It includes the location of Protestant missionary stations in Africa as of 1903, coupled with information on the investment of each mission (school, dispensary, hospital, etc.). In appendix A.2, we replicate the entire analysis using locations from the *Ethnographic Survey of Africa: Showing the Tribes and Languages* (Roome, 1924), digitized by Nunn (2010).[13] This source reports locations of both Protestant and Catholic foreign missionary stations in Africa as of 1924. The results of virtually all the analyses are qualitatively identical using either one of these sources to construct $Mission_{ik}$.

The reason to rely on two different sources is twofold. First, historical sources are likely to be subject to measurement error in the precise location of missions, and we can not rule out that some stations are not reported. The two sources are independent from each other, therefore any non-random measurement error affecting one is not likely to affect the other as well. Second, both sources report locations at a specific point in time; however, new missions are likely to be founded afterwards, and some of those exisfting are abandoned. It is less likely that our results are driven by the location of missions in one specific year if we can replicate the analysis using sources recorded twenty years apart. The cell-level correlation between protestant missions from Beach (1903) and Roome (1924) is 31 %, so they are quite different. One reason for the low correlation is that missionaries started penetrating the African inland after settling on the coast, so the 1924 data has a higher concentration further from the coast (see Figure 1).

2.2.3 Selection of missions and historical controls

Missionary activity in Africa was not randomly located, as illustrated by the case studies in H. B. Johnson (1967). If the factors that determine selection into mission location also correlate with present day aid allocation, coefficient β in equation (1) will be biased.

The first factor to consider is accessibility. Missionaries came by sea, so it is not surprising to find a high concentration of stations along the shore of Africa (see Figure 1). Inland penetration was difficult, so missionaries followed the tracks of early European explorers, some of which were missionaries themselves.[14] The early explorer routes partly correspond to the course of the main rivers, for two reasons. First, upstream navigation was the most effective means of transportation, and would ensure water supply all along the trip. Second, one of the main goals of explorers was to map the geography of Africa, and in particular track rivers from mouth to spring.[15] As there is evidence that areas along coast and rivers have an advantage in development, and are

[12] We thank Julia Cagé and Valeria Rueda for sharing their data with us.

[13] Available at http://scholar.harvard.edu/nunn/pages/data-0

[14] For example, David Livingstone was a missionary at the London Missionary Society (Shepperson, 2017).

[15] For example, in a famous expedition, Henry Morton Stanley discovered first the source of the Nile. Then he established that the Lualaba was not part of the Nile, by navigating it downstream until was clear that was instead connected with the Congo (Encyclopædia Britannica, 2016).

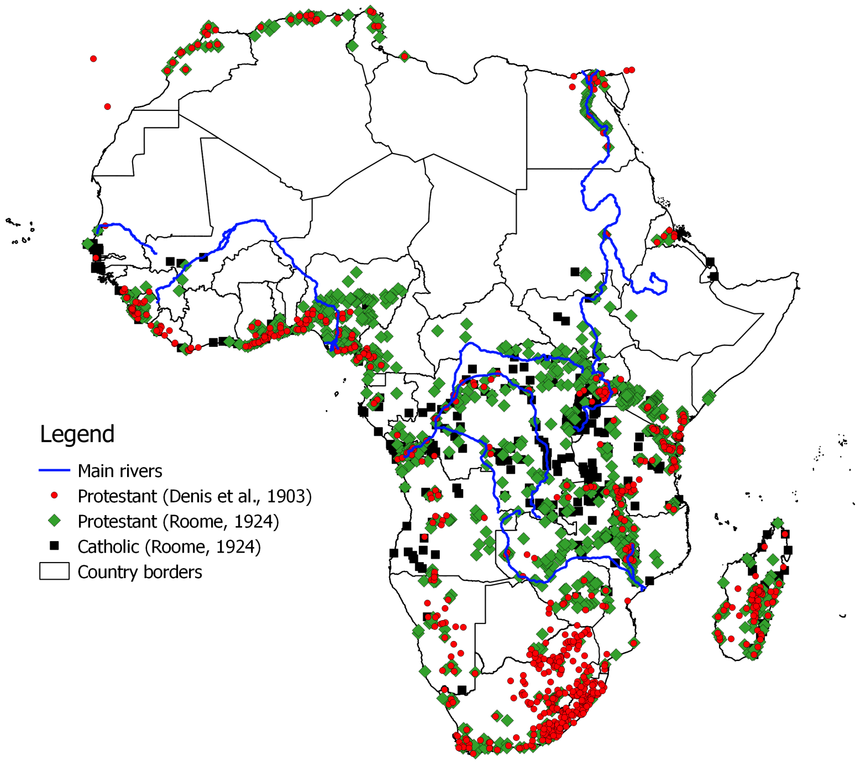

Figure 1: Location of mission stations (Roome, 1924) and (Beach, 1903) and main rivers

more densely populated (Gallup et al., 1999), we control for the (log) kilometer distance from the closest point on the coast and from the closest main river. The data on coastline and rivers is from Natural Earth, and the rivers are: Senegal, Nile, Niger, Zambezi, Congo and its attributes Ubangi and Kasai.[16] We also control for the (log) distance to the closest explorer route, using data from the Century Company Atlas, digitized by Nunn (2010).[17] Distances are in logs of because the marginal effect of a unit of proximity is likely to approach zero as distance increases.[18] Colonial railways built by European powers provided an alternative way of transportation inland. As recent research has documented long-lasting effects of colonial railways on urbanization and growth in Africa (Jedwab, Kerby, & Moradi, 2017; Jedwab & Moradi, 2016), we also control for the (log) kilometer distance from these using data from the Century Company Atlas. Finally, as a (inverse) measure of accessibility, we also include terrain ruggedness.[19] Ruggedness has been found to have both positive and negative effects on development in Africa. Rugged landscape makes it easier to hide, which provided protection from slave traders (Nunn & Puga, 2012), but it also enables rebel warfare (Fearon & Laitin, 2003).

The second factor to consider is the capacity to self-sustain the settlement for a long period of time. Self-sustainability crucially depended on access to water and land cultivation suitability, which are likely to be important for present-day outcomes as well. We control for the access to water by the share of cell area that is within 10 kilometers of a water source, using data from the Seamless Digital Chart of the World.[20] Land cultivation suitability is proxied by the Caloric Suitability Index (CSI) (Galor & Özak, 2016). The CSI builds on information from FAO about the land suitability for a number of crops over the world and computes from this the maximum caloric potential of crops in 5 by 5 arcminute cells, from which we take averages at a 0.5 degree resolution. Missions were also more likely to be established in high altitudes, partly to avoid diseases like malaria, but also because of a more comfortable climate (H. B. Johnson, 1967). We therefore control for average elevation, calculated at the 0.5 degree cell level from the GMTED2010 map by the USGS, and for its interaction with a tropical dummy. As a further control for disease environment, we include a measure of malaria prevalence, the malaria ecology index from Kiszewski et al. (2004).[21]

Finally, we also need to account for the main missionary purpose, namely conversion of Africans to Christianity. Whether missionaries considered conversion more likely in densely populated areas than sparsely populated areas is not clear. Different missionary societies likely had different strategies; some may have preferred sparsely populated areas to convert peoples

[16] Data available at http://www.naturalearthdata.com/features/. We pick all rivers of categories three or above, according to the importance ranking by the source.

[17] Available at http://scholar.harvard.edu/nunn/pages/data-0.

[18] More precisely, we use the log of 1 + the distance measured in kilometers, to avoid losing the cells at zero distance. The average distance is large, so the addition of 1 is inconsequential, besides allowing for zero values.

[19] Calculated with the QGIS raster analysis plug-in, based on USGS GMTED2010 gridded elevation data. Data available at http://topotools.cr.usgs.gov/gmted_viewer/.

[20] 10 kilometers is considered the maximum distance that cattle can travel without stress (Mati, Muchiri, Njenga, de Vries, & Merrey, 2006)

[21] Available at https://sites.google.com/site/gordoncmccord/datasets.

that would not otherwise have been reached. However, some minimum level of settlement was likely a prerequisite for establishing a presence. Ideally we would like to control for population in 18th century, prior to most missionary activity on the African continent. There are no spatially disaggregated censuses from this time, but there have been some attempts to produce population estimates. We employ the population data from the History Database of the Global Environment (HYDE), which provides estimates of population density for the whole world for a resolution of 5 arcminutes for every decade in the 18th century. We take the 0.5 degree (30 arcminutes) average of these estimates, and average the result across the 18th century. We include the population data as a fourth order polynomial to allow for non-linearities. Furthermore, we also employ data on the location of historical cities collected by Chandler (1987).[22] We assign a dummy equal to one to cells that hosted a city any time before 1800. In addition, ethnic considerations may have played an important role in the missionaries' settlement decisions. Some ethnic groups may have been more open to missionaries and foreigners in general, while others were more hostile; alternatively, some tribes may have been particularly backward, thus offering a higher potential for selling the Gospel among them, with the lure of educational and health services. We are concerned that omitting these (and other) unobservable variables at the ethnicity-level may introduce bias, in light of research showing that pre-colonial ethnic institutions had long-lasting effects on development and public good provision in Africa (Gennaioli & Rainer, 2007; Michalopoulos & Papaioannou, 2013). We tackle this concern by including a separate dummy for each of the more than 800 pre-colonial ethnic homelands, whose boundaries are from Murdock (1959)'s ethnolinguistic map. We assign cells to the ethnic polygons that covers the highest percentage of its surface. Our regressions therefore only exploit within ethnicity variation, and results can not be driven by variables varying across ethnicities.

According to Robinson (1915) competition with Islam was a deterring factor, in that spreading the Gospel in predominantly Muslim areas was more complicated. Although speculative, Muslim populations may receive less development aid today for political and religious reasons, so we would like to control for a measure of Muslim penetration at the time of early colonization. Michalopoulos, Naghavi, and Prarolo (2017) show that medieval trade routes to the Arab world had a strong impact on adherence to Islam, persisting until today. We digitize data on Arab trade routes and ports from Brice and Kennedy (2001), and include as controls the (log) kilometer distance from the closest of these routes. Finally, some missions were set up for the purpose of ending slave trade (H. B. 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). We are not aware of geocoded measures of slave trade that are precise enough for our application, but we are already controlling for many of its correlates (distance to coast, terrain ruggedness, and distance to Muslim trade routes), and for ethnic-level dummies.

Table 2 shows the unconditional average values of all historical controls in cells without mis-

[22] Available at http://www.worldcitypop.com.

Table 2: Difference in means of the historical control set

	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***
Malaria ecology	11.64	6.82	4.83***
Caloric mean 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***
Tropical dummy X mean elevation	596.16	367.96	228.20***
Log(Distance to explorer route)	3.62	4.20	−0.58***
Log(Distance to colonial railway)	5.28	3.59	1.68***
18th cent. 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***
No. observations	6512	380	

Mission data from Beach (1903). Each row represents an unconditional mean. Stars * / ** / *** denote significance at the 0.10 / 0.05 / 0.01 level.

sions, and cells with at least one mission, respectively. All variables but elevation are significantly different between mission and non-mission cells. The difference in means of distance to main river and distance to explorer route do not have the expected sign. Note, however, that these are unconditional correlations, and the distance variables are all correlated.[23]

2.2.4 Present-day control variables

We also collect some present-day variables. We do not control for these variables in the baseline estimation, because they may be endogenous to our variable of interest $Mission_{ik}$, and hence constitute bad controls that may cause a bias (Angrist & Pischke, 2008). We do, however, use these variables to probe for mechanisms.

We collect estimates of grid level population data in 1995 from the Gridded Population of the World version 3 (GPW) at 2.5 arc minute resolution, and we aggregate it to the 0.5 degree resolution.[24] To complement the gridded data, we get the location and population of populated places from the Populated Place Merge Project.[25] The data contains information about 16,207 places in Africa, from which we construct dummies for populated places of more than 10,000 inhabitants, as well as province and national capitals.

To test if there is a Christian bias, we collect information on religion from the USAID Demo-

[23] The conditional correlations have the correct sign.

[24] Data from Socioeconomic Data and Applications Center, available at http://sedac.ciesin.columbia.edu/data/collection/gpw-v3. The data contain estimates based on official censuses, and may therefore be noisy at the local level. The results are unchanged if we instead use similarly estimated population data from the ISAM-HYDE project for 2000.

[25] Available for purchase at http://www.worldgeodatasets.com.

graphic and Health Surveys (DHS). The DHS is extensive in both spatial and temporal coverage, and we exploit the fact that the survey relocates geographically between waves to attain a large spatial coverage, approximately one third of the cells in baseline sample. We make the assumption that religious affiliation is fairly stable over the period, or at least that changes in religion is not related to the other variables in the regressions. We use the individual recode of the DHS, which includes women of reproductive age (15–49), as it has the highest country coverage (in some countries, DHS only surveys women). We use the stated religion to compute the population share of each religion in a cell.[26]

3 Empirical results

3.1 Missions and World Bank aid

We estimate baseline equation (1) by least squares. The dependent variable is an indicator equal to one if the cell ever hosted a World Bank project between 1995 and 2015. All regressions include the full set of historical an geographic controls described in section 2.2.2. Results are reported in Table 3. In the first column mission data are from Roome (1924), and we separate missions stations into Protestant and Catholic; in the second column, we use only Protestant missions from Beach (1903); in the third, we use Catholic missions from Roome (1924) and Protestant from Beach (1903); finally, in column 4, we define $Mission_{ik}$ equal to one if the cell hosted at least one mission, either Catholic or Protestant, from any of the two sources.

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 46 % to 80 % more likely to host a World Bank project, compared to the sample mean. In order to assess the bias from unobservables, we draw on the test constructed by Oster (2017) to get a lower bound for β.[27] This test is inappropriate for models with more than one treatment, so we consider only columns 2 and 4. Under the assumption that selection on observables is equal to selection on unobservables, the test produces coefficients close to 0.1, corresponding to a 40 % increase in the likelihood to receive aid.[28]

The estimated correlation is higher for Catholic missions, but we have no way to address

[26] "Protestant" is assigned to respondents who state one of the following (all sic): protestant, protestantism, anglican, sda, Evangelical, anglican church, presbyterian, pentecotist, pentecostal/charismatic, pentecostal, baptist, anglican/protestant, prostestant, salvation army, other christian, armé de salut, evangelical, other christians, ccap, anglican, seventh day advent./baptist, united faith, kimbanguist, kimbanguiste, kibanguist, evangelical/pentecostal, evangelical presbyterian, evangelical, evangelist, evangelica (crente), evangelic. "Catholic" is one of: Catholic, Roman Catholic, catholic, catholique, catolica romana. "Muslim" is one of: muslim, moslem, mulsim, muslem, muslim - alawi, muslim - sunni, muslum, muslman, musulman, musulmane, islam, muslim/islam, islamic, Muslim.

[27] This test is a refinement of Altonji, Elder, and Taber (2005), in that it takes into account the effect of the observable selection on the R-squared.

[28] We calculate the lower bound setting the R-squared from the hypothetical regression on $Mission_{ik}$ and both observed and unobserved controls equal to 1.3 times the R-squared from the actual regression on $Mission_{ik}$ and observed controls, as suggested by Oster (2017). The calculation is based on the assumption that observables and unobservables correlate with $Mission_{ik}$ in the same direction.

Table 3: WB aid and missions from Beach (1903) and Roome (1924)

	World Bank aid 1995–2014			
	(1)	(2)	(3)	(4)
Catholic mission (Roome)	0.184***		0.201***	
	(0.032)		(0.032)	
Protestant mission (Roome)	0.142***			
	(0.022)			
Protestant mission (Beach)		0.135***	0.117***	
		(0.026)	(0.026)	
Any mission				0.160***
				(0.018)
Mean dep. var.	0.253	0.253	0.253	0.253
Oster bound		0.101		0.109
R-sq.	0.400	0.392	0.398	0.399
N	6876	6876	6876	6876

Notes: Conley, 1999 standard errors in parentheses, with cutoff at 2 degrees (≊ 200 Km). The dependent variable is a dummy for whether the grid cell had at least one active World Bank project in the period 1995–2014. Country dummies always included. Historical controls include log distance to closest: explorer route, colonial railway, Arab trade route; a third order polynomial in 18th century population and presence of a city as of 1800; pre-colonial ethnic dummies. Geographical controls include log distance to coast, average altitude, terrain ruggedness index, percentage area area within 10 kilometers from a water source, the caloric suitability index, malaria ecology, and tropics dummy (also interacted with mean altitude). Cells that are more than 90 % covered by barren land or more than 90 % covered by forest are excluded from the sample. The Oster lower bound on the coefficient of interest is calculated as in Oster, 2017: we set the R-squared from the hypothetical regression on $Mission_{ik}$ and both observed and unobserved controls equal to 1.3 times the R-squared from the actual regression on $Mission_{ik}$ and observed control. Stars * / ** / *** denote significance at the 0.10 / 0.05 / 0.01 level.

differences in geographical selection of different confessions, so it is difficult to interpret the difference between the two coefficients. If we drop the Democratic Republic of Congo from the sample the difference between the two is lower than 2 percentage points, and not significant. Congo DRC is the largest country in our sample with 509 cells, and has a large concentration of Catholic missions. Hence, the differential in the coefficient size seems to be driven by this heavy-weight outlier.

In the remaining parts of the paper we only report results obtained constructing $Mission_{ik}$ using data from Beach (1903). The reason is partly to simplify the exposition, and partly because the base map from which the missions were geocoded is of much higher resolution than the one in Roome (1924), so there is likely less measurement error coming from inaccurate geocoding. Nonetheless, in appendix A.2 we report all tables that appear in the paper using data from Roome (1924), to ensure that results are not sensitive to the data source.

Subsample analysis As a first robustness test, we restrict the sample to cells on the coast, or on one of the main rivers. As discussed in section 2, this subsample is likely more homogeneous and enhances the credibility of our selection on observables strategy, that relies on mission and non-mission cells having the same covariate distributions (Imbens & Rubin, 2015). We present estimates both with and without pre-colonial ethnic dummies, because the within-ethnicity variation is very small for the subsamples. The results are reported in Table 4, where the sample is restricted to coastal cells in 1 and 2; to river cells in 3 and 4; and to the union of the two in 5 and 6.

The subsample analysis yields a correlation similar in sign and magnitude to that estimated in the baseline Table 3. When ethnic dummies are included, the point estimate is smaller and marginally insignificant at conventional level in the coastal subsample; it is instead slightly higher in the river subsample. However, even the most conservative specification suggests that mission cells are approximately 20 % more likely to receive World Bank aid.

Table 4: WB aid and missions (Beach, 1903): coastal and river samples

	Coast		River		Coast or river	
	(1)	(2)	(3)	(4)	(5)	(6)
Mission	0.148***	0.107	0.189**	0.175*	0.183***	0.118**
	(0.056)	(0.065)	(0.089)	(0.096)	(0.047)	(0.054)
Ethnic dummies	No	Yes	No	Yes	No	Yes
Mean dep. var.	0.452	0.457	0.407	0.409	0.422	0.425
R-sq.	0.393	0.585	0.295	0.505	0.309	0.519
N	367	363	494	491	851	844

Notes: Conley, 1999 standard errors in parentheses, with cutoff at 2 degrees ($\approxeq$ 200 Km). The dependent variable is a dummy for whether the grid cell had at least one active World Bank project in the period 1995–2014. Country dummies always included. Historical controls include log distance to closest: explorer route, colonial railway, Arab trade route; a third order polynomial in 18th century population and presence of a city as of 1800. Geographical controls include log distance to coast, average altitude, terrain ruggedness index, percentage area area within 10 kilometers from a water source, the caloric suitability index, malaria ecology, and tropics dummy (also interacted with mean altitude). Cells that are more than 90 % covered by barren land or more than 90 % covered by forest are excluded from the sample. Rivers include: Nile, Niger, Senegal, Zambezi, Congo and its attributes Ubangi and Kasai. Stars * / ** / *** denote significance at the 0.10 / 0.05 / 0.01 level.

The next test is again aimed at restricting estimation to a subsample which is more likely to exhibit common support in the covariate distribution. Here, we proceed in a more structured way by constructing three subsamples with propensity score matching: one where the matching has no geographical restrictions (column 1 "Simple" in Table 5), and two where we match within countries (columns 2 "Within 1", and 3 "Within 2"), as detailed in section 2. We then estimate our baseline equation 1 in each of them, with or without ethnic dummies. Across the different subsamples, the coefficient of interest is always positive, significant, and very similar in size to those estimated on the full sample. The coefficients are slightly smaller when ethnic dummies are included, as in Table 4, but still significant at conventional level, and sizable (30 % over the

Table 5: WB aid and missions (Beach, 1903): propensity score matched samples

	Simple		Within 1		Within 2	
	(1)	(2)	(3)	(4)	(5)	(6)
Mission	0.169***	0.111***	0.113***	0.085***	0.135***	0.088**
	(0.025)	(0.033)	(0.028)	(0.032)	(0.026)	(0.034)
Ethnic dummies	No	Yes	No	Yes	No	Yes
Mean dep. var.	0.337	0.338	0.278	0.279	0.347	0.349
R-sq.	0.471	0.678	0.459	0.674	0.466	0.694
N	790	787	593	592	800	797

Notes: OLS on matched subsamples. The dependent variable is a dummy for whether the grid cell had at least one active World Bank project in the period 1995–2014. Robust standard errors in parenthesis. Country dummies always included. Historical controls include log distance to closest: explorer route, colonial railway, Arab trade route; a third order polynomial in 18th century population and presence of a city as of 1800. Geographical controls include log distance to coast, average altitude, terrain ruggedness index, percentage area area within 10 kilometers from a water source, the caloric suitability index, malaria ecology, and tropics dummy (also interacted with mean altitude). Stars * / ** / *** denote significance at the 0.10 / 0.05 / 0.01 level.

sample mean in the most conservative specification).

Controlling for present-day population In all the regressions estimated so far, we have only included controls that are pre-determined with respect to the variable of interest, $Mission_{ik}$. This approach is the most appropriate to attempt a causal interpretation of β. However, it forces us 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, we may fail to control credibly for the selection of missions.[29] Furthermore, there is reason to believe that the presence of mission stations and the activities of missionaries has influenced settlement patterns. In that case, the coefficient on $Mission_{ik}$ partly captures a correlation between current population and aid allocation.

To tackle these concerns, we introduce in the control set different measures of present-day population. Under the assumption that population density is positively auto correlated, the present-day measures may serve as proxy controls for historical population. Under the additional assumptions that selection on population is positive, and that the presence of missions stations increase population, we can interpret the coefficients on $Mission_{ik}$ from these augmented regressions as a lower bound on the true causal effect.[30] The results from these regressions are reported in Table 6.

In column 1 we include fourth order polynomial terms in population in 1995. The coefficients are all highly significant in themselves, suggesting that population has a non-linear effect on aid allocation. The coefficient on $Mission_{ik}$ is slightly smaller than the baseline, but still highly

[29] In the most extreme scenario, we are just controlling for random variables.

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

significant. In column 2 we add dummies for populated places of different size.[31] The coefficient drops to about half of the baseline, but stays highly significant. In column 3 we restrict the sample to cells that contain at least one populated place with more than 10,000 inhabitants. The coefficient stays positive, but drops a bit and loses significance. Note, however, that the inclusion of both country and ethnic dummies in this reduced sample is likely to decrease variation, giving higher standard errors. Dropping the ethnic dummies in column 4 gives a coefficient and standard error closer to the baseline. Column 5 and 6 is estimated on the sample of cells that contain a provincial capital, with and without ethnic dummies, respectively. The coefficient increases sharply compared to the baseline, but the effect relative to the mean of the dependent variable is very similar.

Table 6: WB aid and missions (Beach, 1903): present-day population controls

	All cells	All cells	Pop. place rank 6		Prov. capital cells	
	(1)	(2)	(3)	(4)	(5)	(6)
Mission	0.100***	0.069***	0.072	0.101**	0.190***	0.222***
	(0.026)	(0.025)	(0.044)	(0.041)	(0.052)	(0.038)
Population in 1995	0.011***	0.006***				
	(0.001)	(0.001)				
Population in 1995, quadratic	-0.797***	-0.522***				
	(0.113)	(0.102)				
Population in 1995, cubic	0.002***	0.001***				
	(0.000)	(0.000)				
Population in 1995, quartic	-0.000***	-0.000***				
	(0.000)	(0.000)				
Populated place dummies	No	Yes	No	No	No	No
Ethnic dummies	Yes	Yes	Yes	No	Yes	No
Mean dep. var.	0.253	0.253	0.492	0.492	0.617	0.617
R-sq.	0.403	0.433	0.588	0.334	0.639	0.316
N	6876	6876	1168	1168	698	698

Notes: Conley, 1999 standard errors in parentheses, with cutoff at 2 degrees (≊ 200 Km). Estimation by ordinary least squares. The dependent variable is a dummy for whether the grid cell had at least one active aid project in the period 1995–2014. Country dummies always included. Historical controls include log distance to closest: explorer route, colonial railway, Arab trade route; a third order polynomial in 18th century population and presence of a city as of 1800; pre-colonial ethnic dummies. Geographical controls include log distance to coast, average altitude, terrain ruggedness index, percentage area area within 10 kilometers from a water source, the caloric suitability index, malaria ecology, and tropics dummy (also interacted with mean altitude). Cells that are more than 90 % covered by barren land or more than 90 % covered by forest are excluded from the sample. Stars * / ** / *** denote significance at the 0.10 / 0.05 / 0.01 level.

Sensitivity tests The results in the baseline estimation in Table 3 are largely robust to various sensitivity tests. First, we check if the estimated relationships are stable over time. The

[31]Dummies for each for the presence of at least one: a) national capital; b) province 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; d) urban agglomeration of at least 100,000 people, or city with at least 50,000 people; e) places with at least 10,000 people; f) places with at least 1,000 people.

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 (although casual observation suggests that it has not had much effect).[32] In Table 7 we therefore test if we see the same pattern for aid committed in two different ten year periods, 1995–2004 and 2005–2014. The coefficients are slightly higher for the second period, but so are the total number of cells receiving aid, so in relative terms the coefficients are virtually unchanged.

Table 7: WB aid and missions (Beach, 1903) in two decades

	Ever World Bank aid in period	
	1995-2004	2005-2014
Mission	0.095***	0.125***
	(0.023)	(0.026)
Mean dep. var.	0.154	0.187
R-sq.	0.391	0.363
N	6876	6876

Notes: Conley, 1999 standard errors in parentheses, with cutoff at 2 degrees (≈ 200 Km). The dependent variable is a dummy for whether the cell had at least one active aid project in the period. Baseline control set, i.e. same as in table 3. Stars * / ** / *** denote significance at the 0.10 / 0.05 / 0.01 level.

Next, we check if our use of binary treatment and outcomes is important for the results. In column 1 and 2 of Table 8, we estimate equation (1), but instead use the number of missions and the log number of missions + 1 as the treatment variable, with results very similar to the baseline.

We also investigate whether the relationship between aid and missions is also at work at the intensive margin. We estimate the baseline equation 1 using as dependent variable the number of World Bank aid projects, and report the results on the right panel of Table 8. In column 3 we do so simply by OLS. In column 4, we use OLS, but we restrict the sample to cells that ever received at least one project. In the last two columns, we estimate it using theoretically more appropriate count models: Poisson model in column 5, and Negative Binomial model in column 6.[33] All models deliver a similar result: mission cells host a higher number of aid projects, compared to non mission cells.

Spatial spillovers Missions are highly clustered in our data, as apparent from Figure 1.[34] This leaves open a possibility that we are overestimating the effect of missions on aid; for two neighboring cells that both have missions, but where only one gets aid, it may be argued that

[32] See http://www.oecd.org/dac/effectiveness/parisdeclarationandaccraagendaforaction.htm

[33] The last two models are estimated without ethnic dummies to ensure convergence.

[34] Moran's I statistic for the baseline sample confidently rejects the null hypothesis of zero spatial correlation of mission stations, with a p-value close to zero (not shown).

Table 8: WB aid and missions (Beach, 1903): non-binary variables, intensive margin

	Ever WB aid		Number of WB projects			
	OLS	OLS	OLS	OLS	Poisson	NB2
Number of missions	0.052***					
	(0.013)					
Ln(Number of missions)		0.130***				
		(0.028)				
Mission dummy			1.388***	2.338***	0.834***	0.733***
			(0.333)	(0.752)	(0.110)	(0.074)
Ethnic dummies	Yes	Yes	Yes	Yes	No	No
Mean dep. var.	0.253	0.253	0.916	3.626	0.915	0.915
R-sq.	0.391	0.392	0.478	0.512		
N	6876	6876	6876	1737	6884	6884

Notes: Conley (1999) standard errors in parentheses, with cutoff at 2 degrees ($\approxeq$ 200 Km) in columns 1 to 4. Robust standard errors in column 5 and 6. In columns 1 and 2 the dependent variable is a dummy for whether the cell had at least one active aid project in the period. In columns 3 to 6 the dependent variable is the number of active projects in the period. In column 4 the sample is restricted to cells that received at least one aid project. Baseline control set (same as in table 3), but without ethnic dummies in columns 5 and 6. Stars * / ** / *** denote significance at the 0.10 / 0.05 / 0.01 level.

we should be accounting for the presence of that neighbor mission. Failure to do so may induce omitted variable bias.

To alleviate concerns of spillover bias from neighboring cells, we include spatial lags of missions into the regression. The lag variables for cell i are constructed by multiplying the value of the dummy $Mission_{ik}$ for a given neighbor j with a binary indicator variable for whether the neighbor is within a given cutoff distance from cell i, and summing these products for all $j \neq i$.[35] We include spatial lags for both the inner ring of surrounding 8 neighbors (cutoff distance $\approx$ 70 kilometers) and for the outer ring of 16 next neighbors (cutoff distance $\approx$ 140 kilometers).

Since the location of World Bank funded projects is likely determined by each country's government, it is probably not correlated across country borders. Allowing for cross-border correlations would bias the results, since zero correlations between border cells would pull down the overall estimate.[36] Therefore, we construct the spatial weighting matrix so that it gives zero weight to all neighbors that are not in the same country. Define the $N_k \times N_k$ weighting matrix W_k^r for country k and neighbor ring $r \in \{inner, outer\}$, and let $Mission_k$ be the $1 \times N_k$ column vector, with elements indicating the presence of missions in each cell in country k. The equation to be estimated is then:

[35] There are different ways of incorporating spatial lags of independent variables, including inverse distance weighting matrices. The choice of a binary contiguity is one of convenience, because it makes calculations less intensive, and more natural since the cell-structure of the data implies non-continuous distances. Results are robust to using an inverse distance weighting matrix.

[36] 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, and we present only our preferred specification here.

$$EverAid_{ik} = \beta \cdot Mission_{ik} + \sum_r \zeta_r W_{ik}^r \cdot Mission_k + \delta_k + \mathbf{X}_{ik}\gamma + \varepsilon_{ik}, \tag{2}$$

where W_{ik}^r is row i of the weighting matrix for neighbors in ring r. The estimates, reported in Table 9, reveal that the coefficient of interest is not affected by the inclusion of the spatial lags. Furthermore, there is no evidence of spatial spillovers across cells, suggesting that our estimated correlation holds at a very local level. The results are reassuring, given the high degree of spatial clustering in the data.

Table 9: WB aid and missions (Beach, 1903): spatial lags

	World Bank aid 1995–2014	
	(1)	(2)
Mission	0.170***	0.168***
	(0.028)	(0.028)
Mission spatial lag, inner ring		0.029
		(0.020)
Mission spatial lag, outer ring		-0.006
		(0.016)
Mean dep. var.	0.257	0.257
R-sq.	0.400	0.401
No. of observations	6102	6102

Notes: Conley (1999) standard errors in parentheses, with cutoff at 2 degrees (≊ 200 Km). The dependent variable is a dummy for whether the cell had at least one active aid project. Inner ring refers to the 8 neighbors adjacent to each cell. Outer ring refers to the next 16 closest outer neighbors. Baseline control set (same as in table 3). Stars * / ** / *** denote significance at the 0.10 / 0.05 / 0.01 level.

3.2 Christian religion and Western culture

There is extensive evidence that strategic and political considerations are key factors behind aid allocation, so it is natural consider these as a mechanism behind our results. Western and predominantly Christian countries stand out as the largest shareholders of the World Bank: the US are the largest (holding respectively 10 % in the IDA, and 16 % in the IBRD), and while Japan is second (with respectively 8 % and 7 %), Western European countries together form a sizable group (for example UK, France and Germany together account for, respectively, 15 % and 12 %).[37] If shareholders prefer to allocate more aid to culturally similar areas, then we should expect to find disproportionately more World Bank projects in African regions with a large Christian population, and that have strong ties with Western countries. In both respects, cells

[37] Figure as of 2017 available at the World Bank website: www.worldbank.org/en/about/leadership/votingpowers

that used to host a Christian mission are likely to stand out, as missionary presence increased both Christian and Western footprints.

We explore this channel by replicating our analysis using aid data for China, a donor that is not likely to be biased towards Christian and Western area. If cultural or religious bias is the main explanation behind our results, then we should find no relationship between the presence of a historical mission and present-day Chinese-financed aid. Chinese culture is markedly distinct from Western culture, and the Chinese government is unlikely to give preferential treatment to Christians of any confession. On the other hand, existing research has found China to be roughly comparable to other Western donors in terms of political and economic considerations behind aid allocation (Dreher & Fuchs, 2015; Dreher, Fuchs, Parks, Strange, & Tierney, 2017). We estimate the baseline equation (1) on an indicator for presence of at least one Chinese aid project over the period 2000–2012, and report the coefficient on $Mission_{ik}$ in column 1 of Table 10. The coefficient is smaller relative to the baseline (0.04), but the share of cells in the sample that are allocated Chinese aid is roughly one third of the same share for World Bank aid. This means that in relative terms, the estimates are comparable to the World Bank case, with the mean of Chinese aid in mission cells about 60 % higher than in the sample mean.

The similarity of the results may not be surprising if Chinese aid simply co-locates with World Bank aid; for example, the presence of World Bank aid may provide a signal of the quality of project locations. Although we are not aware of any evidence in this respect, the unconditional correlation between the two outcome variables is non-negligible at 0.25. We tackle this concern by introducing World Bank aid as a control in the regression on Chinese aid, which becomes:

$$EverAidfromChina_{ik} = \beta \cdot Mission_{ik} + \theta \cdot EverAidfromWB_{ik} + \delta_k + \mathbf{X}_{ik}\gamma + \varepsilon_{ik}. \qquad (3)$$

However, $EverAidfromWB_{ik}$ is not predetermined with respect to $Mission_{ik}$, and therefore the interpretation of β requires additional caution. In particular, the coefficient β now captures the difference in the outcome between mission and non-mission cells, conditional on having ever received World Bank aid. The cells that never received World Bank aid despite having a mission must be less suitable for aid due to other unobservables characteristics; the cells that received World Bank aid despite not having a mission must be more suitable due to unobservables characteristics. In other words, conditional on $EverAidfromWB_{ik} = 0$, equation 3 compares "bad" mission cells to regular non-mission cells; conditional on $EverAidfromWB_{ik} = 1$, equation 3 compares "good" non-mission cells to regular mission cells. Thus, controlling for $EverAidfromWB_{ik}$ is likely to bias the estimate of β downward.

Consistent with this logic, the estimate of β (reported in column 2 of Table 10) is 1 percentage point smaller than the coefficient in column 1. However, it is still positive and significant, suggesting that missions do predict Chinese aid, even when the selection on unobservables (introduced by controlling for $EverAidfromWB_{ik}$) is working against finding any effect. Overall, we interpret the results in Table 10 as evidence against the Christian and Western bias hypothesis discussed above.

Table 10: Chinese aid and missions (Beach, 1903)

	Chinese aid 2000–2012	
	(1)	(2)
Mission	0.043**	0.031*
	(0.019)	(0.019)
World Bank aid		0.091***
		(0.011)
Mean dep. var.	0.071	0.071
R-sq.	0.311	0.324
N	7911	7911

Notes: Conley, 1999 standard errors in parentheses, with cutoff at 2 degrees (≊ 200 Km). Estimation by OLS. The dependent variable is a dummy for whether the grid cell had at least one active Chinese aid project in the period 2000–2012. Baseline control set, i.e. same as in table 3. Stars * / ** / *** denote significance at the 0.10 / 0.05 / 0.01 level.

As a second and more direct attempt to test the Christian bias channel, we control for the Christian share of population in our baseline regression. Although there is no comprehensive cell level data on religion, it can be estimated using geocoded surveys from the DHS. In particular, our cell-level Christian share corresponds to the share of DHS respondents in the cell who self-identify as Christians. We can construct this variable for about 3000 cells. Our measure is not likely to be representative at the cell level, and thus subject to (non-systematic) measurement error, which will attenuate coefficients towards zero. If our measure were extremely noisy, we would effectively not be controlling for religion. In column 1 and column 4 of Table 11 we examine whether the Christian share is correlated with aid allocation. Failing to find a statistically significant correlation would be cause for concern, given the hypothesis of a Christian bias. In both cases, however, the correlation is positive, significant and very large.

Next, we re-estimate our equations (1) and (3), and we find that our coefficient of interest remains virtually the same with or without controlling for the Christian share (Table 11).[38] To conclude, we do not find any evidence that the association between missions and aid is motivated by religious or culture bias on the donor side.

3.3 Speed of disbursement

Having found no evidence that the relationship between missions and aid is driven by biased preferences on the donor side, we turn to investigating whether mission cells are somehow more

[38] Also in this case the interpretation of the coefficient is not trivial due to the endogeneity of the Christian share. The coefficients on $Mission_{ik}$ in column 3 and 6 can be interpreted as a lower bound in light of existing evidence that missions increased conversion (Nunn, 2010). The coefficients on $Mission_{ik}$ are not significant in column 5 and 6, but this is a heavily reduced sample, and the coefficient is the same as in the full sample (see Table 10).

Table 11: WB aid and missions (Beach, 1903): Christian share of population

	World Bank aid 1995–2014			Chinese aid 2000–2012		
	(1)	(2)	(3)	(4)	(5)	(6)
Share Christian	0.117**		0.113**	0.113***		0.111***
	(0.054)		(0.054)	(0.035)		(0.035)
Mission		0.099**	0.097**		0.042	0.040
		(0.044)	(0.044)		(0.034)	(0.034)
Mean dep. var.	0.413	0.413	0.413	0.131	0.131	0.131
R-sq.	0.409	0.409	0.410	0.414	0.369	0.414
No. of observations	2895	2895	2895	2792	2792	2792

Notes: Conley, 1999 standard errors in parentheses, with cutoff at 2 degrees (≊ 200 Km). Estimation by ordinary least squares. The dependent variable is a dummy for whether the grid cell had at least one active aid project in the period 1995–2014. Country dummies always included. Historical controls include log distance to closest: explorer route, colonial railway, Arab trade route; a third order polynomial in 18th century population and presence of a city as of 1800; pre-colonial ethnic dummies. Geographical controls include log distance to coast, average altitude, terrain ruggedness index, percentage area area within 10 kilometers from a water source, the caloric suitability index, malaria ecology, and tropics dummy (also interacted with mean altitude). Cells that are more than 90 % covered by barren land or more than 90 % covered by forest are excluded from the sample. Stars * / ** / *** denote significance at the 0.10 / 0.05 / 0.01 level.

suitable to host aid projects, for example due to the possibility of exploiting preexisting material and immaterial infrastructure. Evidence in this direction would provide support for our path dependence hypothesis. This is not a simple task, because we do not have measures of suitability at the cell-level. However, we attempt to test this indirectly using data on financial-flows at the project level.

Initially committed project funds do not get disbursed up-front as a lump sum, but in successive tranches over the course of several years. The structure of disbursements vary between projects, but are usually linked to the successful implementation of some steps in the project. For example, the borrower can ask for reimbursement for eligible expenditures that he has anticipated, or the World Bank can pay third parties (suppliers, NGOs, etc.) directly for implementing some specific activities.[39] As such, a high speed of disbursement can be taken as an indication that the project is proceeding fast and efficiently toward its goals. In order to test if projects implemented in mission cells display a higher speed of disbursement, we collect financial data for each project from the World Bank website.[40] This data records the date and type of all financial transactions over the project lifetime and during the repayment phase.

We adopt the methodology in Kersting and Kilby (2016), who investigate how the speed of disbursements depends on recipient countries' political alignment with the US. They measure this speed by tracking the number of months until 25 % of initial commitments have been

[39] See siteresources.worldbank.org/PROJECTS/Resources/DisGuideEng.pdf.
[40] Available at http://www.worldbank.org/projects?lang=en.

disbursed. We use only initial commitments and disbursements linked to those commitments.[41] The rationale behind the 25 % threshold is that most projects cross it during the sample period, giving a larger sample, but we also report results using 50 % and 75 % thresholds.

Since disbursement data is at the project, rather than location level, we collapse the data to the project level. We also depart from the grid-cell structure, since we want information about specific projects, and there may be several projects in the same cell. We thus match missions to aid by identifying mission stations within a 25 kilometer buffer around each aid location. This distance corresponds approximately to the extent of the 0.5 × 0.5 degree cells in the baseline. Our explanatory variable of interest is the project's share of precision 1 and 2 locations that are within 25 kilometers of a mission station. Since the speed of disbursement may depend on the size of the initial commitment we include it as a control in the regression. We also include dummies for country and year of commitment.

Table 12: WB aid and missions (Beach, 1903): project speed of disbursement

	Months until disbursement reached			Log(months)
	25%	50%	75%	25%
Share of locations with mission	-6.14***	-6.96*	-8.46	-0.28**
	(2.03)	(3.61)	(5.30)	(0.12)
Log initial commitment	0.01	0.84	2.21	-0.07
	(1.04)	(1.34)	(1.46)	(0.05)
Country FE	Yes	Yes	Yes	Yes
Year of committment FE	Yes	Yes	Yes	Yes
R-sq.	0.18	0.19	0.28	0.17
Mean dep. var.	33	47	61	3
N	524	477	404	521

Notes: Standard errors clustered at country level in parentheses. Estimation by OLS. The dependent variable is the number of months until X % (on top of each column) of initial commitments have been disbursed. Stars * / ** / *** denote significance at the 0.10 / 0.05 / 0.01 level.

Table 12 shows that having more project locations in the proximity of Christian missions is positively correlated with speed of disbursement (negatively correlated with months until X % disbursed). The sign is invariably negative no matter what threshold we use, and it is marginally not significant only for the highest threshold (75 %), which corresponds to the smallest sample. The coefficients are small, but still non-negligible: projects with 1 standard deviation higher share of locations in the vicinity of a mission, reach the 25 % threshold more than 2 months earlier than other projects (6 % over the sample mean).

The analysis of disbursement speed indicates that projects with a high share of locations near mission stations tend to receive money slightly faster. We interpret this as evidence that projects in mission cells are implemented at a faster pace, due to favorable conditions on the ground. However, World Bank staff do have some degrees of discretion over the timing of the

[41] Some projects receive additional commitments several years after the initial commitment.

tranches, and Kersting and Kilby (2016) showed that disbursement is sometimes sped up or slowed down in response to political pressure. As such, the results in Table 12 could also be consistent with a Christian favoritism story, for which, however, we have not found any other evidence.

4 Conclusion

We find that the probability of receiving development aid is higher in areas with historical missionary presence. In light of the importance of education and health facilities in the missionary effort, we consider mission stations as the ancestors of present-day development projects. We therefore interpret our findings as evidence of historical path dependence in the spatial distribution of aid. The missionary presence is likely to have shifted permanently downward the cost of implementing development projects in the same locations, due to their heritage of material and immaterial infrastructure. Consistent with this hypothesis, we find that projects implemented in the proximity of mission stations manage to unlock successive disbursements faster than the average project, which suggests a higher absorption capacity. On the contrary, we do not find any evidence that our findings are motivated by donors' biased preference toward Christian or Westernized areas.

Acknowledgments

We are grateful to Valeria Rueda and Julia Cagé for sharing their missionary data with us. We are also grateful to Axel Dreher, Martin Flatø, Nicola Gennaioli, Andreas Kotsadam, Eliana La Ferrara, Rune Jansen Hagen, Edwin Leuven, Jo Thori Lind, Silvia Marchesi, Halvor Mehlum, Anirban Mitra, Kalle Moene, Alexander Moradi, Andreas Müller, Ola Olsson, Anna Tompsett, Gaute Torsvik and Nina Bruvik Westberg for comments. We also thank participants at the 2016 Conference on Development Economics and Policy in Heidelberg, the 2016 Nordic Conference on Development Economics, the 2016 Frisch-Prio Workshop on Foreign Aid, the 2016 African Economic History Network Conference, the 2017 CSAE Conference, and at the 2017 Royal Economic Society meeting. All errors are our own. This work was supported by the Centre of Equality, Social Organization and Performance (ESOP). ESOP is supported by the Research Council of Norway through its Centres of Excellence funding scheme, project number 179552.

References

Alesina, A. & Dollar, D. (2000). Who Gives Foreign Aid to Whom and Why? *Journal of Economic Growth*, *5*(1), 33–63.

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.

Angrist, J. D. & Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion.* Princeton: 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.* New York: Student Volunteer Movement For Foreign Missions.

Brice, W. C. & Kennedy, H. (2001). *An Historical Atlas of Islam.* Leiden: Brill Academic Publishers.

Briggs, R. C. (2014). Aiding and Abetting: Project Aid and Ethnic Politics in Kenya. *World Development*, *64*, 194–205.

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* (Discussion paper No. 12192). Centre for Economic Policy Research. Retrieved September 26, 2017, from http://cepr.org/active/publications/discussion_papers/download.php?dpno=12192

Caicedo, F. V. (2017). Missionaries, human capital transmission and economic persistence in South America. In S. Michalopoulos & E. Papaioannou (Eds.), *The Long Economic and Political Shadow of History. Europe and the Americas* (Vol. 3). London: CEPR Press.

Calvi, R. & Mantovanelli, F. (2016). *Long-Term Effects of Access to Health Care: Medical Missions in Colonial India.* Unpublished manuscript. Retrieved September 26, 2017, from https://sites.google.com/site/fedmantovanelli/CalviMantovanelli2014_Oct2014.pdf

Castelló-Climent, A., Chaudhary, L., & Mukhopadhyay, A. (2017). Higher Education and Prosperity: Catholic Missionaries to Luminosity in India. *The Economic Journal*, (forthcoming).

Chabal, P. & Daloz, J.-P. (1999). *Africa Works: Disorder as Political Instrument.* African issues. Suffolk: James Currey.

Chandler, T. (1987). *Four thousand years of urban growth.* Lewiston: 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.

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, *92*(1), 1–45.

Dreher, A. & Fuchs, A. (2015). Rogue Aid? An Empirical Analysis of China's Aid Allocation. *Canadian Journal of Economics*, *48*(3), 988–1023.

Dreher, A., Fuchs, A., Hodler, R., Parks, B. C., Raschky, P. A., & Tierney, M. J. (2016). *Aid on Demand: African Leaders and the Geography of China's Foreign Assistance* (Development Studies Working Paper No. 400). Centro Studi Luca d'Agliano. Retrieved September 26, 2017, from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2900351

Dreher, A., Fuchs, A., Parks, B. C., Strange, A., & Tierney, M. J. (2017). Apples and Dragon Fruits: the Determinants of Aid and other forms of State Financing from China to Africa. *International Studies Quarterly*, (forthcoming).

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.

Encyclopædia Britannica. (2016). Sir Henry Morton Stanley. *Encyclopædia Britannica Online*.

Fearon, J. D. & Laitin, D. D. (2003). Ethnicity, Insurgency, and Civil War. *American Political Science Review*, *97*(1), 75–90.

Francken, N., Minten, B., & Swinnen, J. F. M. (2012). The Political Economy of Relief Aid Allocation: Evidence from Madagascar. *World Development*, *40*(3), 486–500.

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.

Gallego, F. A. & Woodberry, R. D. (2010). Christian Missionaries and Education in Former African Colonies: How Competition Mattered. *Journal of African Economies*, *19*(3), 294–329.

Gallup, J. L., Sachs, J. D., & Mellinger, A. D. (1999). Geography and economic development. *International Regional Science Review*, *22*(2), 179–232.

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.

Imbens, G. W. (2015). Matching methods in practice: Three examples. *The Journal of Human Resources*, *50*(2), 373–419.

Imbens, G. W. & Rubin, D. B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. New York: 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.

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), 268–284.

Johnson, H. B. (1967). The Location of Christian Missions in Africa. *Geographical Review*, *57*(2), 168–202.

Johnson, R. (2010). Colonial mission and imperial tropical medicine: Livingstone College, London, 1893-1914. *Social History of Medicine*, *23*(3), 549–566.

Kersting, E. K. & Kilby, C. (2016). With a little help from my friends: Global electioneering and World Bank lending. *Journal of Development Economics*, *121*, 153–165.

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.

Krugman, P. (1991a). History and industry location: the case of the manufacturing belt. *American Economic Review: Papers & Proceedings*, *81*(2), 80–83.

Krugman, P. (1991b). Increasing Returns and Economic Geography. *Journal of Political Economy*, *99*(31), 483–499.

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/useful-links/files/MissionsHIVSex.pdf

Masaki, T. (2014). *The Political Economy of Aid Allocation in Africa: Evidence from Zambia* (Working paper No. 5). AidData. Retrieved September 12, 2017, from http://takaakimasaki.com/wp-content/uploads/2014/08/Zambia_Draft_2014-12-3.pdf

Mati, B. M., Muchiri, J. M., Njenga, K., de Vries, F. P., & Merrey, D. J. (2006). *Assessing water availability under pastoral livestock systems in drought-prone Isiolo District, Kenya* (Working paper No. 106). International Water Management Institute. Retrieved September 26, 2017, from http://www.iwmi.cgiar.org/Publications/Working_Papers/working/WOR106.pdf

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.

Michalopoulos, S., Naghavi, A., & Prarolo, G. (2017). Trade and Geography in the Spread of Islam. *The Economic Journal*, (forthcoming).

Michalopoulos, S. & Papaioannou, E. (2013). Pre-Colonial Ethnic Institutions and Contemporary African Development. *Econometrica*, *81*(1), 113–152.

Murdock, G. P. (1959). *Africa: Its Peoples and Their Culture History.* New York: 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.

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: 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.

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.

Okoye, D. (2015). *Things Fall Apart: Missions, Institutions, and Trust*. Unpublished manuscript. Retrieved September 12, 2017, from http://cega.berkeley.edu/assets/miscellaneous_files/120_-_Okoye_Missions_Trust_-_ABCA.pdf

Oster, E. (2017). Unobservable Selection and Coefficient Stability: Theory and Validation. *Journal of Business & Economic Statistics*, (forthcoming).

Robinson, C. H. (1915). *History of Christian missions*. International theological library. Edinburgh: T. & T. Clark.

Roome, W. J. W. (1924). *Ethnographic Survey of Africa: Showing the Tribes and Languages*. London: Edward Stanford.

Shepperson, G. A. (2017). David Livingstone. *Encyclopædia Britannica Online*.

Strange, A., Dreher, A., Fuchs, A., Parks, B. C., & Tierney, M. J. (2015). Tracking Underreported Financial Flows: China's Development Finance and the Aid–Conflict Nexus Revisited. *Journal of Conflict Resolution*, *61*(5), 935–963.

Waldinger, M. (2017). The Long-Run Effects of Missionary Orders in Mexico. *Journal of Development Economics*, *127*, 355–378.

World Bank Group. (2013). *A Stronger, Connected, Solutions World Bank Group: An Overview of the World Bank Group Strategy*. World Bank Group. Washington, DC.

A Appendix

A.1 Maps

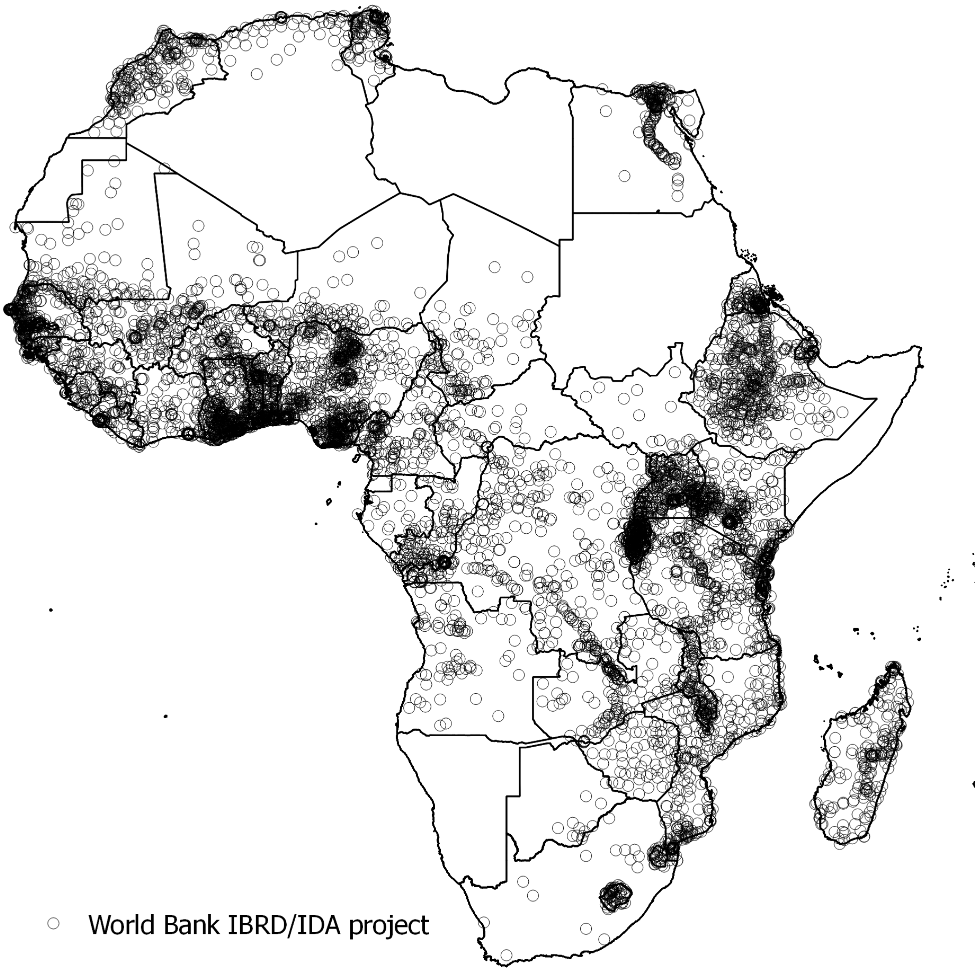

Figure A.1: Location of World Bank IBRD/IDA projects

A.2 Tables using missionary data by Roome (1924)

Table A.1: WB aid and missions (Roome, 1924): coastal and river samples

	Coast		River		Coast or river	
	(1)	(2)	(3)	(4)	(5)	(6)
Catholic mission	0.263***	0.284***	0.088	0.147	0.171***	0.159**
	(0.074)	(0.085)	(0.095)	(0.121)	(0.058)	(0.071)
Protestant mission	0.141**	0.116*	0.181**	0.172**	0.180***	0.155***
	(0.058)	(0.063)	(0.081)	(0.080)	(0.051)	(0.053)
Ethnic dummies	No	Yes	No	Yes	No	Yes
Mean dep. var.	0.452	0.457	0.407	0.409	0.422	0.425
R-sq.	0.417	0.604	0.303	0.514	0.323	0.530
N	367	363	494	491	851	844

Notes: Conley, 1999 standard errors in parentheses, with cutoff at 2 degrees (≊ 200 Km). The dependent variable is a dummy for whether the grid cell had at least one active World Bank project in the period 1995–2014. Country dummies always included. Historical controls include log distance to closest: explorer route, colonial railway, Arab trade route; a third order polynomial in 18th century population and presence of a city as of 1800. Geographical controls include log distance to coast, average altitude, terrain ruggedness index, percentage area area within 10 kilometers from a water source, the caloric suitability index, malaria ecology, and tropics dummy (also interacted with mean altitude). Cells that are more than 90 % covered by barren land or more than 90 % covered by forest are excluded from the sample. Rivers include: Nile, Niger, Senegal, Zambezi, Congo and its attributes Ubangi and Kasai. Stars * / ** / *** denote significance at the 0.10 / 0.05 / 0.01 level.

Table A.2: WB aid and missions (Roome, 1924): propensity score matched samples

	Simple		Within 1		Within 2	
	(1)	(2)	(3)	(4)	(5)	(6)
Mission	0.196***	0.212***	0.212***	0.216***	0.190***	0.186***
	(0.021)	(0.026)	(0.023)	(0.028)	(0.021)	(0.027)
Ethnic dummies	No	Yes	No	Yes	No	Yes
Mean dep. var.	0.385	0.385	0.332	0.332	0.379	0.379
R-sq.	0.328	0.528	0.284	0.515	0.345	0.537
N	1532	1532	1222	1222	1521	1521

Notes: OLS on matched subsamples. The dependent variable is a dummy for whether the grid cell had at least one active World Bank project in the period 1995–2014. Robust standard errors in parenthesis. Country dummies always included. Historical controls include log distance to closest: explorer route, colonial railway, Arab trade route; a third order polynomial in 18th century population and presence of a city as of 1800. Geographical controls include log distance to coast, average altitude, terrain ruggedness index, percentage area area within 10 kilometers from a water source, the caloric suitability index, malaria ecology, and tropics dummy (also interacted with mean altitude). Stars * / ** / *** denote significance at the 0.10 / 0.05 / 0.01 level.

Table A.3: WB aid and missions (Roome, 1924): present-day population controls

	All cells	All cells	Pop. place rank 6		Prov. capital cells	
	(1)	(2)	(3)	(4)	(5)	(6)
Catholic mission	0.165***	0.124***	0.079	0.056	0.127**	0.127***
	(0.033)	(0.034)	(0.076)	(0.054)	(0.064)	(0.046)
Protestant mission	0.114***	0.085***	0.112***	0.106***	0.180***	0.214***
	(0.022)	(0.021)	(0.042)	(0.037)	(0.059)	(0.043)
Population in 1995	0.011***	0.006***				
	(0.001)	(0.001)				
Population in 1995, quadratic	-0.765***	-0.503***				
	(0.113)	(0.102)				
Population in 1995, cubic	0.002***	0.001***				
	(0.000)	(0.000)				
Population in 1995, quartic	-0.000***	-0.000***				
	(0.000)	(0.000)				
Populated place dummies	No	Yes	No	No	No	No
Ethnic dummies	Yes	Yes	Yes	No	Yes	No
Mean dep. var.	0.253	0.253	0.492	0.492	0.617	0.617
R-sq.	0.409	0.437	0.591	0.336	0.645	0.324
N	6876	6876	1168	1168	698	698

Notes: Conley, 1999 standard errors in parentheses, with cutoff at 2 degrees (≊ 200 Km). Estimation by ordinary least squares. The dependent variable is a dummy for whether the grid cell had at least one active aid project in the period 1995–2014. Country dummies always included. Historical controls include log distance to closest: explorer route, colonial railway, Arab trade route; a third order polynomial in 18th century population and presence of a city as of 1800; pre-colonial ethnic dummies. Geographical controls include log distance to coast, average altitude, terrain ruggedness index, percentage area area within 10 kilometers from a water source, the caloric suitability index, malaria ecology, and tropics dummy (also interacted with mean altitude). Cells that are more than 90 % covered by barren land or more than 90 % covered by forest are excluded from the sample. Stars * / ** / *** denote significance at the 0.10 / 0.05 / 0.01 level.

Table A.4: WB aid and missions (Roome, 1924) in two decades

	Ever world bank aid in period	
	(1)	(2)
Catholic mission	0.167***	0.199***
	(0.030)	(0.030)
Protestant mission	0.132***	0.135***
	(0.019)	(0.022)
Mean dep. var.	0.154	0.187
R-sq.	0.403	0.374
N	6876	6876

Notes: Conley, 1999 standard errors in parentheses, with cutoff at 2 degrees (≊ 200 Km). The dependent variable is a dummy for whether the cell had at least one active aid project in the period. Baseline control set, i.e. same as in table 3. Stars * / ** / *** denote significance at the 0.10 / 0.05 / 0.01 level.

Table A.5: WB aid and missions (Roome, 1924): non-binary variables, intensive margin.

	Ever WB aid		Number of WB projects			
	OLS	OLS	OLS	OLS	Poisson	NB2
Number of Catholic missions	0.123***					
	(0.024)					
Number of Protestant missions	0.079***					
	(0.014)					
Ln(Number of Catholic missions)		0.156***				
		(0.029)				
Ln(Number of Protestant missions)		0.171***				
		(0.026)				
Catholic dummy			2.572***	2.833***	0.687***	0.668***
			(0.443)	(0.732)	(0.099)	(0.076)
Protestant dummy			1.279***	2.274***	0.700***	0.624***
			(0.210)	(0.499)	(0.074)	(0.063)
Ethnic dummies	Yes	Yes	Yes	Yes	No	No
Mean dep. var.	0.253	0.253	0.916	3.626	0.915	0.915
R-sq.	0.398	0.400	0.499	0.533		
N	6876	6876	6876	1737	6884	6884

Notes: Conley (1999) standard errors in parentheses, with cutoff at 2 degrees (≊ 200 Km) in columns 1 to 4. Robust standard errors in column 5 and 6. In columns 1 and 2 the dependent variable is a dummy for whether the cell had at least one active aid project in the period. In columns 3 to 6 the dependent variable is the number of active projects in the period. In column 4 the sample is restricted to cells that received at least one aid project. Baseline control set (same as in table 3), but without ethnic dummies in columns 5 and 6. Stars * / ** / *** denote significance at the 0.10 / 0.05 / 0.01 level.

Table A.6: WB aid and missions (Roome, 1924): spatial lags.

	World Bank aid 1995–2014	
	(1)	(2)
Catholic missions	0.204***	0.205***
	(0.034)	(0.034)
Protestant mission	0.131***	0.128***
	(0.024)	(0.024)
Protestant spatial lag - inner ring		0.025*
		(0.015)
Protestant spatial lag - outer ring		-0.013
		(0.014)
Catholic spatial lag - inner ring		0.025
		(0.020)
Catholic spatial lag - outer ring		0.016
		(0.020)
Mean dep. var.	0.257	0.257
R-sq.	0.435	0.436
No. of observations	6102	6102

Notes: Conley (1999) standard errors in parentheses, with cutoff at 2 degrees (≊ 200 Km). The dependent variable is a dummy for whether the cell had at least one active aid project. Inner ring refers to the 8 neighbors adjacent to each cell. Outer ring refers to the next 16 closest outer neighbors. Baseline control set (same as in table 3). Stars * / ** / *** denote significance at the 0.10 / 0.05 / 0.01 level.

Table A.7: Chinese aid and missions (Roome, 1924)

	Chinese aid 2000–2012	
	(1)	(2)
Catholic mission	0.131***	0.116***
	(0.030)	(0.030)
Protestant mission	0.060***	0.049***
	(0.016)	(0.016)
World Bank aid		0.083***
		(0.011)
Mean dep. var.	0.071	0.071
R-sq.	0.320	0.330
N	7911	7911

Notes: Conley, 1999 standard errors in parentheses, with cutoff at 2 degrees (≊ 200 Km). Estimation by OLS. The dependent variable is a dummy for whether the grid cell had at least one active Chinese aid project in the period 2000–2012. Baseline control set, i.e. same as in table 3. Stars * / ** / *** denote significance at the 0.10 / 0.05 / 0.01 level.

Table A.8: WB aid and missions (Roome, 1924): Christian share of population

	World Bank aid 1995–2014			Chinese aid 2000–2012		
	(1)	(2)	(3)	(4)	(5)	(6)
Share of Christians	0.117**		0.108**	0.113***		0.104***
	(0.054)		(0.053)	(0.035)		(0.034)
Catholic mission		0.153***	0.152***		0.162***	0.161***
		(0.036)	(0.035)		(0.039)	(0.039)
Protestant mission		0.088***	0.087***		0.073***	0.071***
		(0.029)	(0.029)		(0.022)	(0.022)
Mean dep. var.	0.413	0.413	0.413	0.131	0.131	0.131
R-squared	0.409	0.414	0.415	0.371	0.383	0.384
No. of observations	2895	2895	2895	2792	2792	2792

Notes: Conley, 1999 standard errors in parentheses, with cutoff at 2 degrees ($\approxeq$ 200 Km). Estimation by ordinary least squares. The dependent variable is a dummy for whether the grid cell had at least one active aid project in the period 1995–2014. Country dummies always included. Historical controls include log distance to closest: explorer route, colonial railway, Arab trade route; a third order polynomial in 18th century population and presence of a city as of 1800; pre-colonial ethnic dummies. Geographical controls include log distance to coast, average altitude, terrain ruggedness index, percentage area area within 10 kilometers from a water source, the caloric suitability index, malaria ecology, and tropics dummy (also interacted with mean altitude). Cells that are more than 90 % covered by barren land or more than 90 % covered by forest are excluded from the sample. Stars * / ** / *** denote significance at the 0.10 / 0.05 / 0.01 level.

Table A.9: WB aid and missions (Roome, 1924): project speed of disbursement

	Months until disbursement reached			Log(months)
	25%	50%	75%	25%
Share of locations with mission	-4.69**	-5.32*	-8.36*	-0.19*
	(1.76)	(2.91)	(4.86)	(0.11)
Log initial commitment	0.09	0.50	2.23	-0.05
	(0.98)	(1.31)	(1.38)	(0.04)
Country FE	Yes	Yes	Yes	Yes
Year of committment FE	Yes	Yes	Yes	Yes
Mean dep. var.	33	47	61	3
R-sq.	0.247	0.251	0.348	0.201
N	524	477	404	521

Notes: Standard errors clustered at country level in parentheses. Estimation by OLS. The dependent variable is the number of months until X % (on top of each column) of initial commitments have been disbursed. Stars * / ** / *** denote significance at the 0.10 / 0.05 / 0.01 level.