Connective Financing

Chinese Infrastructure Projects and the Diffusion of Economic Activity in Developing Countries

Richard Bluhm* Axel Dreher† Andreas Fuchs‡
Bradley C. Parks§ Austin M. Strange¶ Michael J. Tierney‖

May 25, 2020

Abstract

This paper studies the causal effect of transport infrastructure on the spatial concentration of economic activity. Leveraging a new global dataset of geo-located Chinese government-financed projects over the period from 2000 to 2014 together with measures of spatial inequality based on remotely-sensed data, we analyze the effects of transport projects on the spatial distribution of economic activity within and between regions in a large number of developing countries. We find that Chinese-financed transportation projects reduce spatial concentration within but not between regions. In line with land use theory, we document a range of results which are consistent with a relocation of activity from city centers to their immediate periphery. Transport projects decentralize activity particularly strongly in regions that are more urbanized, located closer to the coast, and less developed.

Keywords: transport costs, infrastructure, development finance, foreign aid, spatial concentration, China

JEL classification: F15, F35, R11, R12, P33, O18, O19

*Institute of Macroeconomics, Leibniz University Hannover; Department of Political Science, University of California San Diego, email: *rbluhm@ucsd.edu*

†Alfred-Weber-Institute for Economics, Heidelberg University; KOF Swiss Economic Institute; CEPR; Georg-August University Goettingen; and CESifo; email: *mail@axel-dreher.de.*

‡Department of Economics and Centre for Modern East Asian Studies, University of Goettingen; and Kiel Institute for the World Economy; email: *mail@andreas-fuchs.net.*

§AidData, Global Research Institute, William & Mary; and Center for Global Development; email: *bparks@aiddata.wm.edu.*

¶Department of Government, Harvard University; email: *strange@g.harvard.edu.*

‖Department of Government, William & Mary; email: *mjtier@wm.edu.*

1 Introduction

The Nairobi-Thika Highway is a 50.4 km dual carriageway that extends from Nairobi City Center to the town of Thika in Kiambu County. China Eximbank approved a loan in 2009 to the Kenyan government to substantially widen and improve it. The project, which was locally known as the "Thika Super-Highway Project," sought to reduce congestion and travel times between Nairobi and a set of satellite towns—including Ruaraka, Kasarani, Kiambu Town, Githurai, Ruiru, Juja, and Thika—along a critically important transportation corridor (African Development Fund 2007). Kenyans increasingly rely on employment, educational opportunities, and services in Nairobi, but prefer to live in less crowded and expensive peri-urban and suburban areas within commuting distance of the country's capital. Consequently, these areas have experienced high levels of population growth, which has placed an ever-growing strain on the roads that households and firms use to travel to and from Nairobi (KARA and CSUD 2012, Maina et al. 2018).

Work on the project began in January 2009 and was completed between February and October 2012 (African Development Bank 2014b, p. 43). Upon completion, traffic flows increased from 85,000 vehicles per day to 123,000 vehicles per day (a 45 percent increase), journey speeds rose from 8 km per hour to at least 45 km per hour in sections with the highest registered traffic, and average commuting times from Thika to Nairobi fell from 2-3 hours to 30-45 minutes (KARA and CSUD 2012, African Development Bank 2014a, 2016, 2019).

Figure 1 shows the route of the highway as well as the crossroads and interchanges that were constructed or rehabilitated as a result of the project. It also uses remotely-sensed nighttime light data to visualize changes in the spatial distribution of economic activity that took place between 2008 and 2013. Economic activity spread out along the transport corridor and became substantially less concentrated in the core of Nairobi during and after the implementation of the project. Many of the largest increases in economic activity took place near the crossroads and interchanges that were constructed or rehabilitated to better connect satellite towns to the highway.[1] It also led to a substantial reduction in the concentration of economic activity in a narrow stretch along the highway as people and businesses moved to new locations. Between 2008 and 2013, the geographical areas within a 4 km buffer of the highway experienced a 27 percent reduction in the spatial concentration of nighttime light intensity.[2] At the same time, land values doubled in Thika and rose even faster in areas closer to Nairobi (like Kasarani, where land values increased from US$46,000 per acre to US$500,000 per acre), farmgate

[1]Between 2008 and 2013, the geographical areas that fall within a 4 km buffer of the Nairobi-Thika Highway experienced an approximately 53 percent increase in average nighttime light output.

[2]In this paper, we use the terms *spatial concentration*, *spatial inequality*, and *spatial decentralization* of economic activity interchangeably, as they all refer to changes in the distribution of people and output across geographic space. Spatial concentration, as measured by the Gini coefficient introduced later in this paper, fell from 0.425 in 2008 to 0.31 in 2013 in the 4 km buffer.

prices for dairy products and horticulture rose due to increased access to markets, and trade and investment alongside the road corridor expanded (KARA and CSUD 2012, African Development Bank 2014b, 2016, 2019).[3]

[Figure 1 about here.]

The Nairobi-Thika Highway is one example of a more general relationship between transportation infrastructure and the spatial distribution of economic activity. This relationship differs across and within regions. Across regions, a large body of literature suggests that investments in transport infrastructure have heterogeneous and context-specific impacts on the distribution of economic activity. Fajgelbaum and Redding (2018) show that the construction of a national railway network in Argentina contributed to the development of interior regions and increased population density in both rural and urban areas. They interpret this pattern as evidence of reduced inland transport costs which allowed the interior to participate in Argentina's rapid rise in the late 19th century. Faber (2014) and Baum-Snow et al. (2020) provide evidence that the extension of China's national highway system from 1990 to 2010 had the opposite effect. These new highways increased economic output and population in regional centers at the expense of the hinterland.

The concentration of economic activity within regions, however, responds more uniformly to changes in transport costs. Baum-Snow (2007) demonstrates that construction of highways in the United States reduced the populations of the cities through which they passed by 18 percent by enabling people to move to the suburbs. Moreover, firms often follow populations, setting in motion a broad decentralization of activity. Similar results for China suggest that new roads replaced substantial parts of central city population and industrial GDP to its surroundings (Baum-Snow et al. 2017). This line of research has generated a series of related studies across a range of countries.[4] To identify causal effects, these studies typically rely on historical transport networks or other country-specific historical circumstances. Comparable instrumental variables are typically not available for a broad set of countries. As a result, the conclusions that the literature derives are based on evidence from a selected set of countries where large expansions of the transport network coincide with opportunities for identification. Developing countries—where the geographic concentration of economic activity is most severe—are particularly understudied (Baum-Snow et al. 2017).

[3] According to African Development Bank (2014a, 5), "[t]he road has [...] spurred the establishment of new manufacturing, food processing, and small and medium enterprises" and it also "brought employment opportunities to the people along the road corridor."

[4] See, for example, Bayes (2007) on Bangladesh, Bird and Straub (2014) on Brazil, Donaldson (2018) on India, Henderson and Kuncoro (1996) on Indonesia, and Gibbons et al. (2019) on the United Kingdom. For another study on the United States, see Duranton and Turner (2012); on China, also see Banerjee et al. (2012). Baum-Snow and Turner (2017) provide a recent survey of this literature.

The key contribution of this study is to examine whether and to what extent the results from the existing literature can be generalized across a large sample of developing countries. We focus specifically on the provision of transport infrastructure financing from China's government, as it has assumed a dominant role in the construction and rehabilitation of transportation infrastructure around the world during the 21st century. Most of our analysis centers on the decentralization of economic activity within regions, as China's local footprint within a region is often sizable, but we also present results relating to infrastructure financing and inequality across regions.

Since 2000, China's government has financed many of the largest transport infrastructure projects in developing countries. The short- and long-run consequences of China's infrastructure financing activities—including the US$1 trillion Belt and Road Initiative (BRI)—are the subject of considerable debate in the media and within policy circles. A growing number of articles in economics focus on the expected impact of the BRI in different regions (e.g., Perlez and Huang 2017, Bandiera and Tsiropoulos 2020, Bird et al. 2020, de Soyres et al. 2020, Lall and Lebrand 2020).[5] Beijing's critics claim that it finances poorly-designed and hastily-executed projects that provide few economic benefits, while Western donors and lenders have learned through decades of experience to design and implement infrastructure projects in more careful and sustainable ways. In response to mounting criticism that it finances politically motivated and economically unsustainable projects, the Chinese government has doubled down on its leadership role in the market for global infrastructure finance.[6]

Many developing countries have unmet infrastructure financing needs, and the leaders of these countries are quick to point out that China is willing and able to finance roads, bridges, railways, and ports at a time when Western donors and lenders are not (Swedlund 2017).[7] For example, during his tenure as the President of Senegal, Abdoulaye Wade admonished traditional donors and creditors for their cumbersome bureaucratic procedures, noting that: "[w]ith direct aid, credit lines and reasonable contracts, China

[5]As explained by Mauk (2019), "China has never released any official map of Belt and Road routes nor any list of approved projects, and it provides no exact count of participating nations or even guidelines on what it means to be a participant." Beijing's unwillingness to share detailed information about BRI projects has aroused suspicion and speculation about its intentions due to the sheer scale and scope of the initiative, which has "little precedent in modern history, promising more than US$1 trillion in infrastructure and spanning more than 60 countries"(Perlez and Huang 2017).

[6]At the 2017 Belt and Road Forum for International Cooperation, President Xi emphasized that "[i]nfrastructure connectivity is the foundation of development through cooperation. We should promote land, maritime, air and cyberspace connectivity, concentrate our efforts on key passageways, cities and projects and connect networks of highways, railways and sea ports [...]" (Xi 2017).

[7]An important reason for these infrastructure financing gaps follows from the fact that "Western donors have by and large gotten out of hard infrastructure sectors [...] and [t]hey [instead] channel their assistance overwhelmingly to social sectors or to infrastructure sectors such as water supply and sanitation that have direct effects on household health" (Dollar 2008). Indeed, Western aid agencies and multilateral development banks have become significantly more risk averse about bankrolling large-scale infrastructure projects because of the environmental, social, and financial risks that they pose (Nielson and Tierney 2003, Hicks et al. 2008).

has helped African nations build infrastructure projects in record time. [...] I have found that a contract that would take five years to discuss, negotiate and sign with the World Bank takes three months when we have dealt with Chinese authorities" (Wade 2008).

In order to estimate the potential effects of Chinese infrastructure financing on the geographic concentration of economic activity, we introduce a new dataset of geo-located Chinese government-financed projects that were undertaken in developing countries between 2000 and 2014.[8] The larger dataset includes 3,485 projects that took place in 6,184 subnational locations across 138 countries during this fifteen-year period. For the purposes of our analysis, we focus on 269 Chinese government-financed transportation infrastructure projects that were undertaken in 1,211 subnational locations across 86 countries. The total financial value of these projects amounts to US$64 billion. We estimate the effects of these projects on the spatial concentration of economic activity—both within and across subnational jurisdictions—with satellite data on the geographical dispersion of nighttime light output (as in Henderson et al. 2018).

To estimate the causal effect of Chinese government-financed projects on spatial concentration, we circumvent the problem of missing instrumental variables for infrastructure for a large number of countries by going one step back, and focusing on the availability of *resources* for transport infrastructure. This approach has the advantage that comparable data are available for a large number of countries, and plausibly exogenous instruments can be applied across these diverse empirical settings. We introduce an instrumental variable that uses an exogenous supply push variable interacted with a local exposure term: China's domestic production of potential project inputs interacted with each recipient region's probability of receiving projects. We use China's annual production of aluminum, cement, glass, iron, steel, and timber to proxy its capacity to provide physical project inputs.[9] The intuition behind this approach is that the Chinese government has long considered these production materials as strategic commodities and therefore produced them in excess of domestic demand. This policy results in large surpluses, some of which China redirects to overseas infrastructure projects.[10] We therefore expect China to be more lenient towards countries that request financing for transport infrastructure projects in the years when such inputs are abundant

[8]Though the Belt and Road Initiative (BRI) was not officially launched until late 2013, the Chinese government had already begun providing large-scale financing for transport infrastructure in developing countries by the turn of the century. These pre-2014 projects share most of the characteristics of transport infrastructure projects that are now formally part of the BRI.

[9]Exporting excess capacity in a variety of materials through infrastructure investments abroad is one of the secondary motives often ascribed to the Chinese government's BRI initiative. For example, the Economist writes "Mr Xi [...] hopes to [...] export some of his country's vast excess capacity in cement, steel and other metals" (see `www.economist.com/the-economist-explains/2017/05/14/what-is-chinas-belt-and-road-initiative`). Our approach extends the strategy proposed in Dreher et al. (2019b) which exclusively used the level of steel production.

[10]Chinese infrastructure projects usually require construction inputs that are oversupplied in China, and Chinese state-owned banks usually obligate their borrowers to import these inputs on a preferential basis (Dreher et al. 2020).

and less lenient in the years when such inputs are scarce. We also expect subnational localities that frequently receive Chinese government-financed transport projects to be more severely affected by year-to-year fluctuations in the supply of project inputs. We show that they receive larger amounts of Chinese government financing in years of overproduction than subnational localities that infrequently receive Chinese government-financed transport projects, which presumably occurs because existing local capacity and relationships make it easier to implement additional projects. This approach is similar to the identification strategies used in the "China shock" literature, which analyzes the effects of Chinese import competition on US labor markets, or the literature on US food aid and conflict (e.g., Nunn and Qian 2014), and can be interpreted as a difference-in-differences estimate. We essentially compare the effects of Chinese transport projects induced by annual changes in the production of raw materials in subnational localities with a high probability of receiving such projects and subnational localities with a low probability of receiving such projects.

Our results show that Chinese-financed infrastructure projects reduce spatial concentration within first-order regions and accelerate the diffusion of economic activity around cities (in line with the case study presented in Figure 1). Specifically, we find that the Gini coefficient measuring the spatial concentration of economic activity is reduced by about 3.4 percentage points in first-order regions, and 3.6 percentage points in second-order regions. Similar specifications for inequality between regions suggest that Chinese-financed infrastructure cannot be robustly linked to changes in the concentration of economic activity across regions. Our main results are robust to a variety of perturbations, such as the choice of control variables or variations of the instrumental variable. In line with the predictions of land use theory, we find that transport projects shift activity from dense locations to their surroundings, that is, from the highest quintile of the light distribution to lower quintiles. What is more, the impact of Chinese-financed infrastructure projects on concentration within regions is very heterogeneous. Our results show that these projects particularly reduce concentration in regions with more urban areas, with low travel-time to cities, and in those that are located closer to the coast. We also provide evidence that these effects are largest in Africa and poorer regions within developing countries, which tend to experience rapid population growth and have a large demand for infrastructure.

The remainder of the paper proceeds as follows. Section 2 briefly discusses what theory and the existing empirical literature suggest about the relationship between transport projects and the spatial concentration of economic activity within and across regions. Section 3 introduces a subnationally georeferenced dataset of Chinese government-financed projects around the world, and discusses the remotely-sensed measure of spatial concentration. Section 4 describes the empirical strategy. Section 5 presents and discusses the results. Section 6 concludes.

2 Transport infrastructure and the concentration of economic activity

Urban land use theory suggests that transport infrastructure should reduce spatial concentration *within* subnational regions if these jurisdictions primarily consist of urban areas and their surroundings. This is a key prediction of the canonical monocentric city model (Alonso et al. 1964, Mills 1967, Muth 1969), in which all workers commute to a single location in the central business district (CBD). In this model, agglomeration benefits and rents are highest in the city center but decline with distance from the CBD. Initially, many people choose to live near the center and pay higher rents in order to reduce their commuting times. Subsequent investments in transportation infrastructure increase the speed of transportation, reduce commuting costs, and increase the supply of readily accessible land, shifting this gradient outwards. Transportation infrastructure thus facilitates urban sprawl—the flow of people out of the city center—by turning a city's agricultural surroundings into valuable locations to live. The model also implies that people should spread out along newly created highways (Baum-Snow 2007), just as we document above for the case of Nairobi. Firms also have incentives to move out of cities in response to new transportation infrastructure but they depend on a number of factors, not least how much a particular industry benefits from agglomeration economies.[11] Going beyond the monocentric city model, as people and firms move out of existing urban centers, new sub-centers emerge and create replicas of CBDs that lure even more people and firms away from existing urban centers (Ogawa and Fujita 1980, Fujita and Ogawa 1982, Henderson and Mitra 1996).

Even though the model is highly stylized, the prediction that new or upgraded transportation infrastructure disperses economic activity away from urban agglomerations to their immediate hinterland is supported by a large and growing body of empirical evidence. Consider suburbanization in 20th-century America, where urban sprawl and strong population growth in cities with more developable surroundings are well documented (Burchfield et al. 2006, Saiz 2010). The construction of highways in the United States dramatically lowered commuting times, which increased demand for suburban residential space relative to urban one (Baum-Snow 2007). There is also some evidence of similar processes of diffusion around urban areas in developing countries. Baum-Snow et al. (2017) examine the effect of road and railway infrastructure on the spatial distribution of economic activity in China, and find that ring road investments displaced 50 percent of industrial GDP from central cities to outlying areas.[12]

[11]Firms have weaker incentives than individuals since they face a more complex set of costs when leaving city centers. They trade agglomeration benefits off against a variety of costs. This gives rise to a pattern where industries relying on manufacturing and less-skilled jobs decentralize more than those with more high-skilled and managerial employment (Rossi-Hansberg et al. 2009, Baum-Snow 2014).

[12]They also find that railway investments have similar, but quantitatively smaller, impacts.

Similarly, Bayes (2007) provides evidence that a US$1 billion bridge investment in Bangladesh, which connected farmers and firms in the underdeveloped, northwestern division of Rajshahi to the country's more economically developed eastern divisions, expanded market access, reduced input prices, facilitated diversification into higher-value crops, and ultimately reduced the level of income inequality within Rajshahi (a first-order region). As Chinese-financed infrastructure projects often represent a substantial proportion of local infrastructure investment in a given year, we expect that they lead to a decentralization of economic activity around urban areas in the developing world.

Economic geography is concerned with the concentration of economic activity across regions. The classic core-periphery model stresses the role played by increasing returns to scale when economic activity starts to concentrate in a particular region. When trade costs are high or prohibitive, firms are spread out evenly across regions to locate themselves close to consumer demand. When transport projects increase connectivity between leading and lagging regions, labor and capital should move from the periphery to the better connected core, creating a core-periphery split until there is almost complete specialization (Krugman 1991). However, some of these forces reverse at high levels of concentration. Puga (1999), for example, shows how a lack of migration with low trade costs implies that firms will again locate closer to final demand. This gives rise to the characteristic bell-shape for regional inequalities in relation to trade costs. The advantage of being in a central location that is well-connected with other markets erodes when very low trade costs make it easy to reach the periphery. While the bell-shaped curve is a robust prediction, it requires additional heterogeneity in agricultural trade costs, urban congestion or migration decisions that make the overall relationship difficult to identify.

Empirical evidence on how transport costs shape regional inequalities reflects this heterogeneity. Brülhart et al. (2019) show that the advantages of market potential are shrinking in the developed world where trade costs are almost negligible but are still an important determinant of employment growth in developing countries. In terms of infrastructure investments, Bird and Straub (2014) find that investments in Brazil's road network increased economic agglomeration in the already prosperous population centers of the South, while also facilitating economic agglomeration in less developed areas of the North. On balance, these investments reduced spatial inequality across the country's municipalities (second-order regions).[13] However, Faber (2014) provides evidence that China's National Trunk Highway System—a major inter-regional transportation infrastructure project—reduced levels of economic activity in the newly connected peripheral regions relative to non-connected peripheral regions. Given this heterogeneity, the cross-national scope of our study and our focus on Chinese-financed

[13] Relatedly, in their study of Argentina's steam railroad network and the agricultural sector, Fajgelbaum and Redding (2018) suggest that lower transport costs can enable economic actors located in remote, interior regions to participate in structural transformation.

projects only, we do not have strong reasons to believe that the transportation projects in our data will uniformly increase or decrease concentration between regions.

Regardless of whether one considers inequalities within or across subnational jurisdictions, the distributional consequences of transportation infrastructure are especially important for developing countries. Most developing countries face major transportation infrastructure gaps in both urban and rural regions. Internal transport costs are four to five times higher within Ethiopia or Nigeria than within the United States (Atkin and Donaldson 2015). What is more, the total length of the road network per 1,000 people is roughly 10 times lower for South Asia, East Asia and the Pacific, Sub-Saharan Africa, and the Middle East and North Africa than for North America (Andrés et al. 2014). Many of these countries have both rapidly expanding populations and underfunded, poorly designed transportation systems (Cervero 2013). This confluence of problems has severely taxed the existing urban transportation networks of developing economies (Pucher et al. 2005). It has also exacerbated neglect of rural hinterlands: The rural access index (RAI), or the percentage of people who live within two kilometers of an "all-season" road, is close to 100 percent for most European countries but below 51 percent for every Sub-Saharan African country (Rozenberg and Fay 2019). Major infrastructure financing gaps make it extremely difficult for developing countries to overcome the spatial bottlenecks created by high levels of urban concentration and rural neglect.

Subnational regions within developing countries are often defined by dense central cities surrounded by underdeveloped hinterlands.[14] Large cities in many African countries, for example, tend to be highly congested relative to overall levels of infrastructure, industry, and economic opportunity (Lall et al. 2017), while Africa's secondary cities tend to be isolated from world markets (Gollin et al. 2016).[15] The Dar es Salaam administrative region, for example, was home to 8% of the Tanzanian population and 53% of manufacturing value-added in 2002 (Storeygard 2016). This concentration is in part due to historical legacies. Colonial-era investments were highly localized in many developing countries, setting in motion powerful forces of economic agglomeration and creating spatial inequalities that have persisted over long periods of time (Bonfatti and Poelhekke 2017, Roessler et al. 2018). Regardless of its origin, a high concentration of

[14]As Baum-Snow et al. (2017) point out, the urban distribution of economic activity in many developing countries today largely resembles that of early 20th century America, in which industry was initially overwhelmingly concentrated in urban centers. At the same time, the historical development of coastal urban areas in many developing countries occurred under different conditions than those of many developed countries. In particular, coastal cities emerged in part to take advantage of comparative trading advantages for manufacturers (Henderson et al. 2018).

[15]Moreover, individuals often have fewer alternatives to public transportation in developing countries. For instance, per capita automobile usage is especially low in Sub-Saharan Africa. Many city dwellers use the minibus or walk to work in absence of alternatives (Lall et al. 2017). When individuals have fewer private transportation options, the effect of public transportation infrastructure may be particularly pronounced since ex ante commuting costs will be relatively high (Glaeser and Kahn 2004, Duranton and Puga 2015).

people and economic activity in primate cities is often considered to be far from optimal (see, for example, Ades and Glaeser 1995). In our empirical setting, high levels of urban congestion in developing economies represent a powerful latent force for spatial dispersion as new transportation options become available.

3 Data

New geocoded dataset of Chinese government-financed projects

The Chinese government considers the details of its overseas development program to be a "state secret" (Bräutigam 2009, p. 2). It does not publish a country-by-country breakdown of its expenditures or activities. Nor does it systematically publish project-level data on its less concessional and more commercially-oriented financial expenditures and activities in developing countries. In order to overcome this challenge, we collaborated with AidData, a research lab at the College of William & Mary, to build a first-of-its-kind dataset of the subnational locations where Chinese government-financed projects took place around the globe between 2000 and 2014. The underlying project-level data are from Dreher et al. (2020), who use a publicly documented method called Tracking Underreported Financial Flows (TUFF) to facilitate the collection of detailed and comprehensive financial, operational, and locational information about Chinese government-financed projects (Strange et al. 2017, 2018). The TUFF method triangulates information from four types of open sources—English, Chinese and local-language news reports; official statements from Chinese ministries, embassies, and economic and commercial counselor offices; the aid and debt information management systems of finance and planning ministries in counterpart countries; and case study and field research undertaken by scholars and non-governmental organizations (NGOs)—in order to minimize the impact of incomplete or inaccurate information.[16] Economists, political scientists, and computational geographers have used this dataset and earlier versions of it—capturing fewer countries and years—to explain the nature, allocation and effects of Chinese government-financed projects (e.g., BenYishay et al. 2016, Hernandez 2017, Dreher et al. 2018, Eichenauer et al. 2018, Martorano et al. 2020, Isaksson and Kotsadam 2018a,b).[17]

[16] The method is organized in three stages: two stages of primary data collection (project identification and source triangulation) and a third stage to review and revise individual project records (quality assurance). The TUFF data collection and quality assurance procedures are described at length in Strange et al. (2017, 2018).

[17] We build on earlier georeferenced datasets that cover Africa, the Tropical Andes, and the Mekong Delta for fewer years only (BenYishay et al. 2016, Dreher et al. 2019a). Note that we exclude all suspended and cancelled projects as well as projects that reached the (non-binding) pledge stage or (binding) official commitment stage but never reached implementation or completion during the period of study (2000-2014).

Our data captures all officially committed projects that entered implementation or reached completion between 2000 and 2014 in five regions of the world (Africa, the Middle East, Asia and the Pacific, Latin America and the Caribbean, and Central and Eastern Europe) and were supported by Chinese official financing—i.e., foreign aid and other forms of concessional and non-concessional financing from Chinese government institutions.[18] In total, we identify 3,485 projects (worth US$273.6 billion in constant 2014 dollars) in 6,184 discrete locations across 138 countries. All of these projects were subjected to a double-blind geocoding process (Strandow et al. 2011), in which two trained coders independently employ a defined hierarchy of geographic terms and independently assign uniform latitude and longitude coordinates and standardized place names to each location where the project in question was active. Coders also specify a precision code for each location. Precision code 1 corresponds to an exact location; precision code 2 corresponds to locations within 25 kilometers of the exact project site; precision code 3 corresponds to a second-order region; and precision code 4 corresponds to a first-order region.[19] If the coordinates and precision codes do not match, a senior "arbitrator" identifies the source of the discrepancy and assigns a final set of geocodes for all sites. The purpose of this double-blind coding process is to minimize the risk of missed or incorrect locations.[20]

In order to merge these geocoded project data with our outcome measures of spatial inequality within and across subnational regions, we aggregate all projects with precision codes 1-4 to first-order regions and all projects with precision codes 1-3 to second-order regions. Figure 2 shows the locations of projects that can be placed at least within second-order regions, over the 2000-2014 period. The resulting subsample includes 2,142 Chinese government-financed projects at 4,432 discrete locations (collectively worth US$201 billion) that were completed or being implemented in 883 first-order regions and 1,319 second-order regions within 129 countries between 2000 and 2014.[21] 269

[18]More precisely, we code all Chinese government-financed projects as Official Development Assistance ("ODA-like"), Other Official Flows ("OOF-like"), or "Vague Official Finance." Chinese ODA-like projects refer to projects financed by Chinese government institutions that have development intent and a minimum level of concessionality (a 25 percent or higher grant element). Chinese OOF projects refer to projects financed by Chinese government institutions that have commercial or representational intent and/or lack a grant element of 25 percent or more. Projects assigned to the Vague Official Finance category are Chinese government-financed projects where there is insufficient information in the public domain about concessionality and/or intent to make a clear determination as to whether the flows are more akin to ODA or OOF. Total Chinese Official Finance (OF) is therefore the sum of all projects coded as ODA-like, OOF-like, or Vague (Official Finance). For more detailed discussion of the distinction between these types of Chinese development finance, see Dreher et al. (2018).

[19]We exclude all projects with precision codes between 5 and 9 from the regression analysis below. Such projects (e.g., country-wide projects) were not able to be geocoded with a sufficient level of spatial precision to be included in the regional-level data.

[20]Note that the point-based method used to geocode these projects is not designed to measure the exact linear path of transportation infrastructure. This implies that one cannot 'connect the dots' and look for effects alongside the roads, railways etc. However, it is useful for measuring the effects within treated subnational regions as we do in the present paper.

[21]We only focus on low-income and middle-income countries. More precisely, we include countries that

of these projects were assigned to the "transport and storage" sector, implemented in 1,211 different locations (and with a combined value of about US$64 billion). The vast majority of these projects focused on building transportation infrastructure, such as roads, railways, bridges, seaports, and airports. These projects are the ones we exploit for most of our analyses. We also use a larger sample of projects that supported economic infrastructure and services, which includes roads, railways, bridges, seaports and airports but also power grids, power lines, cell phone towers, and fiber optic cable lines (514 projects at 1,897 locations with a value of about US$165 billion).

[Figure 2 about here.]

Figure 2 illustrates the global reach of Chinese official finance in the 21st century. Consistent with earlier periods of Chinese aid giving (Dreher and Fuchs 2015), Chinese transport and other projects alike are densely concentrated in African and Asian countries. The figure also calls attention to the fact that many Chinese government-financed projects are situated in coastal regions, including some of the highest-value transportation projects.

Our data can be disaggregated by financial flow type (Official Development Assistance, Other Official Flows, Vague Official Finance), or sector. For the purposes of the latter, we use the OECD's three-digit sector classification scheme, which categorizes projects according to their primary objectives.[22]

Measuring inequality within and across subnational regions

Reliably measuring local economic activity across the globe with official data is difficult. Few countries collect and report comprehensive data at the individual or plant/establishment level at regular intervals and subnational GDP data are generally only available in highly developed countries. To circumvent this problem, we follow recent literature that uses nighttime light intensity as a proxy for local economic activity (Hodler and Raschky 2014, Michalopoulos and Papaioannou 2014, Dreher et al. 2019a,b). While

the World Bank does not classify as high-income countries in a given year (see `https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lendinggroups`, last accessed September 13, 2017). We also exclude small states with a population size that falls below a threshold of 1,000,000 inhabitants. Table A-1 in the Appendix lists all countries included in the analysis.

[22]There are 24 of these OECD sector codes: education (110), health (120), population policies/programs and reproductive health (130), water supply and sanitation (140), government and civil society (150), other social infrastructure and services (160), transport and storage (210), communications (220), energy generation and supply (230), banking and financial services (240), business and other services (250), agriculture, forestry, and fishing (310), industry, mining, and construction (320), trade and tourism (330), general environmental protection (410), women in development (420), other multisector (430), general budget support (510), developmental food aid/food security assistance (520), non-food commodity assistance (530), action relating to debt (600), emergency response (700), support to NGOs and government organizations (920), and unallocated/unspecified (998).

nighttime lights were initially proposed as a measure of income for countries with weak statistical capacity, they were quickly adopted as a measure of subnational economic activity in developing countries more broadly. Subsequent studies have demonstrated that changes in light emissions correlate strongly with traditional measures of welfare down to the village level (Bruederle and Hodler 2017, Weidmann and Schutte 2017).

One of the main purposes of infrastructure investments is to enable the relocation of economic activity. Hence we are interested in shifts in total activity, rather than per capita measures. We follow Henderson et al. (2018), who use nighttime light intensity at the grid-cell level as a measure of aggregate economic activity—i.e., the product of population and light output per capita—and then calculate a spatial Gini coefficient based on the distribution of this proxy for total GDP.

We obtain data on nighttime light intensity from the Defense Meteorological Satellite Program's (DMSP) Operation Line Scan satellites. The DMSP satellites circle the earth in sun-synchronous orbit and record evening lights between 8:30 and 9:30 pm on a 6-bit scale ranging from 0 to 63. The National Oceanic and Atmospheric Administration (NOAA) processes these data, creates annual composites of the daily images at a resolution of 30 arc seconds, and makes them available to the general public. We use the so-called "stable lights" product, which filters out most background noise, forest fires, and stray lights. Even though there are well-known issues in these data with bottom and top coding (see Jean et al. 2016, Bluhm and Krause 2018), nighttime lights are measured in a consistent manner around the globe and avoid many of the measurement errors involved in more traditional survey data.

We proceed in four steps to calculate our measure of spatial concentration. First, we divide the entire world into a grid of 6 arc minute cells (i.e., an area of about 9.3 km by 9.3 km at the equator) and align the grid with lights data. Second, we intersect this grid with the global first-order and second-order administrative boundaries, which creates "squiggly" cells along the regional borders.[23] Third, for all squiggly cells in this grid and all years in the nightlights data, we compute the sum of light (s_i), the land area of each cell in km^2 (a_i),[24] and the light intensity in the cell ($x_i = s_i/a_i$). We average the resulting light intensities whenever more than one satellite is available and turn off all pixels that do not fall on land before aggregating the lights to the grid level. Finally, we compute the Gini coefficient of light intensities over all lit cells within an administrative unit as

$$\textsc{Gini} = \frac{\sum_{i=1}^{n} w_i \sum_{j=1}^{n} w_j |x_i - x_j|}{2 \sum_{i=1}^{n} w_i \sum_{i=1}^{n} w_i x_i}, \tag{1}$$

[23]We obtained the regional borders from the GADM vector dataset (version 2.8). The same data were used to geocode the Chinese-financed projects.

[24]We calculate the land area of each cell using the Gridded Population of the World (v4) land and water area raster.

where $w_i = \frac{a_i}{\sum_{i=1}^{n} a_i}$ is an area-based weight and n is the total number of lit cells in a region.[25] We also construct Gini coefficients for inequality *between* first-order and second-order regions. The formula remains the same only that it is based on the average light intensity of a region (swapping $\bar{x}_i$ for x_i) and w_i is then defined as the land area of the entire region.

Our spatial Gini coefficient can be interpreted as the average (weighted) difference between the light intensities of all possible pairs of cells within an administrative region. Alternatively, we may think of it geometrically as the area under the Lorenz curve plotting the cumulative distribution of weighted light intensities against the cumulative distribution of cell areas (in km^2). As can be seen from the long differences in the spatial Gini coefficient presented in a world map of first-order regions in Figure 3, our dependent variable shows considerable variation over the time period under analysis (2000–2013).

[Figure 3 about here.]

It is important to emphasize that the Gini coefficient captures the overall dispersion of economic activity, which is a product of the population distribution and the distribution of light per capita.[26] Henderson et al. (2018) show that the temporal variation in population density across administrative regions is substantially larger than the variation in income per capita. In our context, this implies that, given a reasonable degree of mobility, a significant proportion of observed changes in the within-region distribution of light intensities may be attributable to shifts in the population distribution rather than differences in per-capita income. This is precisely the type of variation we are interested in and expect to be affected by transport infrastructure investments.[27]

4 Empirical strategy

We are interested in permanent changes in spatial concentration as a result of infrastructure investments. Denoting regions by j, countries by i, and years by t, our main equation relates our luminosity-based measure of spatial concentration, $\textsc{Gini}_{jit}$, to

[25]Note that we exclude unlit cells when computing Gini coefficients; otherwise all larger countries would have Gini coefficients near unity. This is equivalent to assuming that these pixels are unpopulated. Bruederle and Hodler (2017) suggest that focusing only on lit pixels improves the correlation with local welfare.

[26]To see this, consider that x_i is defined as $\frac{p_i}{a_i} \times \frac{s_i}{p_i}$, where $\frac{p_i}{a_i}$ is population density and $\frac{s_i}{p_i}$ is light per capita in each cell. We prefer using nighttime lights over alternative data sources for population density. Population data at comparable resolutions (such as the Global Human Settlement Layer, Gridded Population of the World, or Landscan) are based on rarely available censuses, which are then disaggregated in space and interpolated over time. They would not allow us to exploit annual variation in the commitment of transport projects and changes in economic activity, which are the basis of our identification strategy. Census data are also less frequently available in poorer developing countries which are the primary target of China's investment activities.

[27]Since the time period of our study is relatively short, a relocation of people is likely to dominate a somewhat slower relocation of firms.

the total number of years in which transportation projects have been committed to a region until $t-2$, $N_{ji,t-2}$. We lag this variable by two years to account for the difference between the commitment date and the expected finalization of a project.[28] We begin with a flexible specification which allows the effect of a Chinese-financed project to be arbitrarily correlated with region-specific fixed effects and time trends:

$$\textsc{Gini}_{jit} = \beta N_{ji,t-2} + \mu_{ji} + \theta_{ji} \times t + \lambda_{it} + \mathbf{x}'_{jit}\boldsymbol{\gamma} + \epsilon_{jit}, \tag{2}$$

where $\mathbf{x}_{jit}$ is a vector of control variables (mainly included in tests for robustness), μ_{ji} are region-fixed effects, $\theta_{ji} \times t$ are region-specific linear time trends, and λ_{it} are country-year-fixed effects, which absorb a variety of potential shocks hitting all regions of a country in a particular year.

A more tractable version of this model can be obtained by taking first differences:

$$\Delta\textsc{Gini}_{jit} = \beta \Delta N_{ji,t-2} + \theta_{ji} + \tau_{it} + \Delta\mathbf{x}'_{jit}\boldsymbol{\gamma} + \Delta\epsilon_{jit}, \tag{3}$$

where $\tau_{it} = \lambda_{it} - \lambda_{i,t-1}$ is a new set of country-year-fixed effects and the region-specific trends in levels are now captured by the region-fixed effects in first differences, $\theta_{ji} = \theta_{ji} \times t - \theta_{ji} \times (t-1)$. In both equations, β measures the effect of committing at least one new project two years prior on contemporary changes in spatial concentration. The model nests less flexible approaches with a strict parallel-trends assumption since all θ_{ji} could be zero. We allow a wide range of dependency structures to occur in $\Delta\epsilon_{jit}$. Transportation infrastructure projects often connect more than one administrative region. Clustering standard errors on the country level permits arbitrary spatial and temporal correlation among all regions within a country. To account for connections across countries, we also report Conley errors with a spatial cutoff of 500 km and a heteroskedasticity- and autocorrelation-consistent (HAC) structure with a lag cutoff of 1,000 years in the time-series dimension.[29]

Our preferred measure of transportation infrastructure projects, $\Delta N_{ji,t-2}$, is thus a binary variable indicating that at least one new project is committed to a particular

[28] This lag duration corresponds to the difference between actual project start and end dates for a subset of projects where both data points are available. In our data set, approximately 1,100 Chinese government project records had enough information to calculate the average time between project start and finish. The average time from start to completion within this subset is about 2.1 years; historical data on Chinese development projects also reveal a median of two years between project start and completion (Dreher et al. 2020, based on data from Bartke 1989).

[29] Using a long lag cutoff in the HAC part of the errors implies that the weight of a time-series shock is almost constant, which is equivalent to clustering on regions in the time-series dimension. We also used higher spatial cutoffs to estimate Conley errors but found no substantive changes beyond 500 km (and a very small subset of standard errors could not be computed beyond 500 km). Alternatively, we clustered at the region level or country-year and region level. The results are qualitatively invariant to these choices.

region in a year.[30] Clearly, the size of the projects is not necessarily homogeneous across locations, and thus the effects of projects on spatial inequality might differ along the intensive margin. Unfortunately, we lack comprehensive information on the financial values for more than a third of these projects (see Dreher et al. 2020), so we prefer the binary indicator and only present additional results using (logged) aggregate dollar values for comparison. We define our dependent variable based on commitment years rather than actual disbursement dates as comprehensive data on disbursements are not available and virtually impossible to estimate through open-source data collection. With first differences, a two-year lag, and a lack of nighttime lights after 2013 (from the DMSP-OLS system), our sample effectively covers the period from 2002 until 2013.

The subnational allocation of Chinese development projects is almost certainly endogenous to spatial concentration. For example, China may allocate more resources to poorer and less connected regions. After all, an official goal of Chinese development financing is to make "great efforts to ensure its aid benefits as many needy people as possible" (Information Office of the State Council 2011). Previous studies also demonstrate that the allocation of Chinese funding is correlated with per-capita income and population size (Dreher and Fuchs 2015, Dreher et al. 2018). Reverse causality could also stem from commercially-motivated projects that either get sited in economic centers to exploit agglomeration effects, or locate in suburban areas out of cost concerns.

Apart from reverse causality, our parsimonious specification omits a number of variables which are likely to be correlated with Chinese infrastructure funding as well as with spatial concentration. Some of these covariates vary across regions and over time in a non-linear fashion, so that they are not captured by our battery of fixed effects and region-specific trends. Such unobserved variation could, for example, arise when the decision to finance an infrastructure project in a particularly abundant region is driven by an increased demand for natural resources (or other commodities) in China (Guillon and Mathonnat 2020). More broadly, Chinese development projects have been linked to deteriorating political institutions and higher levels of corruption at the local level (Brazys et al. 2017, Isaksson and Kotsadam 2018a). Chinese development financing also directly affects subnational and national development in Africa (Dreher et al. 2019b, 2020) but how this relates to the spatial distribution of economic activity is not clear ex ante. Greater local growth could lead to a reduction of spatial concentration within regions—both directly and indirectly through positive spillovers—or it could increase the within-region concentration of economic activity at the expense of poorer cities and villages in the region. In the context of this paper, local growth is a mediating factor and therefore a "bad control" (Angrist and Pischke 2008). The same holds for population or

[30] This measure does not count the number of projects or project locations in a particular year. This is because projects are often co-located, e.g., different sections of a highway, and are coded with multiple locations per project. Both do not necessarily capture the intensive margin of infrastructure investments but rather reflect the definitions adopted during the geocoding process.

population density, which we do not include as a control (although this hardly affects our results).

[Figure 4 about here.]

To address concerns about endogeneity we use an instrumental-variables strategy. We instrument Chinese infrastructure projects with the interaction of two variables. The first is China's production of raw materials that are typically used in transport infrastructure projects—aluminum, cement, glass, iron, steel, and timber—which proxy China's capacity to provide physical project inputs. The second is the regional probability of receiving a Chinese-financed transport infrastructure project in a given year.[31] We calculate this probability as the fraction of years over the 2000-2014 period in which a Chinese government-financed transport infrastructure project has been committed and denote this variable by $\overline{p}_{ji}$. We measure China's production of aluminum, cement, iron, and steel in 10,000 tons, glass in 10,000 weight cases, and timber in 10,000 cubic meters.[32] Given that the production of these raw materials trends upwards over time, we detrend the individual time series. We then extract the first common factor from these six inputs, F_t, resulting in one variable that maximizes the variation of the underlying components. Figure 4 reports the corresponding graphs, including the original and detrended input materials, and the first factors of both to emphasize the non-linear time variation that we exploit for identification. Rather than using six separate interactions as instruments, which are strongly correlated, we interact the first common factor of the detrended (logged) inputs with the probability to receive aid to form a single instrumental variable.

We lag this series by one year (relative to the timing of project commitments) so that domestic overproduction in China translates into transport infrastructure projects abroad approximately one year later. In this setup, the production of Chinese raw materials only varies over time (and is exogenous to spatial inequality within any particular region), while the probability of receiving projects varies only across regions. This is how our instrument resembles the supply-shock instruments commonly used in trade and development economics, such as the recent literature on the impact of the rise of Chinese manufacturing on local US labor markets (Autor et al. 2016) or studies on US food aid and civil conflict (e.g., Nunn and Qian 2014). Putting these elements together,

[31] Our description of the instrument draws in part on Dreher et al. (2019b), where the level of steel production in China was first introduced as a supply shock in the time-series dimension. The original instrument has been used in a number of studies, including Brazys and Vadlamannati (2018), Humphrey and Michaelowa (2019), and Zeitz (2020). Since we have introduced our variant of the instrument with multiple detrended inputs, it has been used in Gehring et al. (2019) and Dreher et al. (2020).

[32] We use USGS data on annual production of aluminum (`https://www.usgs.gov/centers/nmic/aluminum-statistics-and-information`, last accessed 12 October 2019). We have retrieved the annual production volumes of cement, glass, pig iron, steel, and timber via Quandl and complemented them with information from the website of the National Bureau of Statistics of China (`http://www.stats.gov.cn/english/statisticaldata/yearlydata/YB1999e/m12e.htm`; last accessed 12 October 2019).

we estimate the following first-stage regression:

$$\Delta N_{ji,t-2} = \delta(F_{t-3} \times \bar{\rho}_{ji}) + \omega_{ji} + \phi_{i,t-2} + \Delta \mathbf{x}'_{jit}\boldsymbol{\zeta} + \nu_{ji,t-2}, \tag{4}$$

where F_{t-3} is the first common factor of the (detrended and logged) raw materials produced in China, $\Delta \mathbf{x}_{jit}$ are the controls from the main equation, ω_{ji} are region-fixed effects, and $\phi_{i,t-2}$ are country-year-fixed effects. Equations (3) and (4) are estimated using Two-Stage Least Squares (2SLS). Note that we cannot estimate the level equation directly, as our instrument is linked to new project commitments, i.e., differences in $N_{ji,t-2}$, and has no counterpart which could be used in eq. (2).

The intuition behind this identification approach resembles that of a difference-in-differences design. Simplifying somewhat, we essentially compare the effects of Chinese transport infrastructure projects on spatial concentration induced by changes in domestic production of potential project inputs in China across two groups: regions that are regular and irregular recipients of Chinese transport infrastructure financing. Or, in other words, we use differences in the local exposure to the common overproduction shock originating in China to identify the effects of transport infrastructure projects on the spatial distribution of economic activity.

We investigate the validity of this approach in several steps. First, the identifying assumptions inherent in this approach could be violated if changes in the production volumes of raw materials in China led to heterogeneous changes in the propensity of receiving a project in regular recipient regions as opposed to irregular recipient regions, and if the changes produced different effects on spatial concentration in these regions. Hence, in comparison to a standard panel difference-in-differences setting, our instrument ensures that the timing of the intervention is exogenous but still requires parallel pre-treatment trends across regions that are regular versus irregular recipients of Chinese transport projects. Figure 5 compares the trends in spatial concentration before China became active in 2000 among future recipient regions. It shows that spatial concentration among regions that will ultimately receive a project is on a remarkably parallel path to those that will not receive a project in the period from 2000 to 2014. Nevertheless, this does not rule out that there is dynamic selection in the period after 2000, that is, once China's government had become a major global infrastructure supplier.

[Figure 5 about here.]

Second, allowing for correlated random trends implies that we do not need to assume parallel pre-treatment trends. The key identification assumption is $Cov(F_{t-3} \times \bar{\rho}_{ji}, \Delta\epsilon_{jit}) = 0$ conditional on region-specific time trends, as well as region- and country-year-fixed effects. This leaves few sources of confounding variation. Commodity price shocks and commodity cycles are known to affect local incomes heterogeneously (e.g., Berman and Couttenier 2015). The detrended input series might be correlated with the production volumes and prices of other commodities. If their time-varying effect on spatial concentration is uniform across regions in a country, then it is fully captured by detrending the raw series and including country-year-fixed effects. If their time-varying effect is heterogeneous across regions but linear, then it is captured by the region-specific trends. Only if their time-varying effect is non-linear and heterogeneous across regions, we would need to control for these shocks (which we do in the robustness checks).

Third, the production of physical project inputs could also be correlated with overall trade volumes or foreign direct investment. China's share of world manufacturing value added rose steadily over the sample period and this rise coincided with a large demand shock for raw materials (Autor et al. 2016). Quite plausibly, frequent recipients of Chinese transport infrastructure projects are also frequent host regions of investment projects and have close trade ties with China. If this is the case, then the differences in the spatial concentration of economic activity might be the result of trade and investment, rather than transport infrastructure projects. To address this concern, we later present robustness checks where we control for the yearly volume of exports to China, imports from China, and Chinese foreign direct investment, interacted with a set of variables that makes it more or less likely that a region is affected by changes of these variables.

Last but not least, our empirical strategy is related to a large shift-share literature, in which instruments are usually constructed as sums of shocks to a variety of industries with varying local exposures. There are two ways to achieve identification in such settings. Local industry shares can be interpreted as instruments, provided that they are exogenous (Goldsmith-Pinkham et al. 2020). As Borusyak et al. (2018) demonstrate, identification can also be purely based on exogenous variation in the time-series shocks, even when variation in local exposures is endogenous.[33] In fact, our design is most similar to that of Nunn and Qian (2014), which Borusyak et al. (2018) highlight as one of the cases where their result applies.

5 Results

Baseline results

Table 1 reports our main results on the relationship between Chinese-financed transport infrastructure investments and economic concentration. Throughout this table we emphasize the two different types of concentration—within regions and between regions—

[33]For the panel case, Borusyak et al. (2018) establish that our estimator is consistent when the covariance between the detrended input series and a weighted average of the within location time-variation in unobserved factors affecting spatial concentration approaches zero in large samples. This is likely to work with reasonably large T, together with a battery of fixed effects, and can be supported by including proxies for the remaining unobserved variation.

which we estimate in four different ways. Column 1 shows results when concentration is measured as inequality within second-order regions (i.e., districts). This is the most granular level at which the project-level data can be used while retaining comprehensive coverage.[34] Column 2 moves one level of spatial aggregation upwards and reports results for spatial concentration within first-order regions. Column 3 turns to between-regional inequality and reports results with inequality computed over average light intensities in second-order regions, resulting in an analysis at the level of first-order regions.[35] Finally, column 4 shows results at the country level, which is the level of analysis typically presented in studies of between-region inequality (e.g., Lessmann and Seidel 2017). The Gini coefficient is computed over average light intensity at the first-order level.

[Table 1 about here.]

We report four specifications in each of the four columns. Panel a shows the results from least-squares fixed-effects regressions. It demonstrations that a combination of measurement errors, simultaneous causality, and the omission of potential confounders biases the estimates toward zero, no matter the type of concentration measure or level of spatial aggregation. Although the coefficient estimates are negative in three of four regressions, they are imprecisely estimated in all of them and small in magnitude.

Panel b reports the reduced-form estimates for the same set of regressions. Here we regress the change in the spatial Gini coefficient on our instrumental variable and the full set of fixed effects. If our identification strategy works and there is an effect of transportation infrastructure on spatial concentration at any of these levels, then we should observe a strong reduced-form effect as well. Indeed, columns 1 to 2 show that there is a sizable and negative effect of the instrument on changes in spatial concentration within first- and second-order regions. This effect will be passed through with the same sign if the corresponding first-stage regression is sufficiently strong and the coefficient on the instrument in those regressions is positive. Columns 3 and 4 show that we do not find a significant reduced-form relationship for the between-region regressions.

Panel c in Table 1 presents our main results where the two-period lag of the dummy for new project commitments ($\Delta N_{ji,t-2}$) is instrumented by the detrended project input series (F_{t-3}) times our local exposure variable ($\bar{\rho}_{ij}$). Recall that we expect to find a negative effect on spatial concentration within regions and have no clear prior on the effect of transport projects on inequality between regions. Our results are in line with these expectations. For inequality within first- and second-order regions, the 2SLS coefficients

[34] Going below the second-order level would remove almost half of all transport projects. 55 percent of project-location pairs have a precision code referring to an exact location, very few are within 25 kilometers of an exact location, 26 percent have a precision code at the level of a second-order region and 15 percent have a precision code at the level of a first-order region.

[35] Note that column 3 has fewer observations than the preceding column because we do not have second-order level data for all countries.

are negative, statistically significant at conventional levels, and of substantial magnitude. It appears that measurement error, simultaneous causality, and/or omitted variables caused an upward bias in the OLS estimates which could be addressed by instrumenting Chinese project locations. According to the results in columns 3 and 4, we find no evidence in favor of the hypothesis that these projects affect inequality between regions.

The point estimates in columns 1 and 2 indicate that the Gini coefficient is permanently reduced by 3.4 (3.6) percentage points in first-order (second-order) regions where at least one Chinese government-financed transport project has been committed two years before.[36] While we have not yet tested how much of this result is driven by urban areas, it is in line with the notion that building or upgrading transport infrastructure allows economic activity to decentralize around congested cities. The estimated effect sizes are plausible—the literature typically finds large displacements away from city centers (e.g., Baum-Snow 2007, Baum-Snow et al. 2017)—and even moderate compared to the decreases observed in the case of Nairobi. Moreover, our Local Average Treatment Effect (LATE) uses variation induced by the production of physical input factors in China and will thus have a greater impact on big infrastructure projects requiring large volumes of steel, cement and other physical inputs.

Panel d in Table 1 reports the associated first-stage regressions. Reassuringly, none of the above estimates suffer from a weak-instrument problem. The coefficients are highly significant and all associated first-stage F-statistics are considerably larger than the conventional rule-of-thumb value of 10. Note that this also holds for the between-region results where the F-statistics are around 30. As expected, we observe a positive relationship between the supply-push instrument and the probability of hosting a Chinese transport project. Domestic production of aluminum, cement, glass, iron, steel, and timber within China translates into more transport projects abroad at a meaningful rate. While it does not map directly into the growth rates of the underlying inputs, a typical change in F_{t-3} is about 0.4 in either direction. Such an annual increase raises the probability of receiving a Chinese-funded transport project by about 7.1 percentage points ($0.4 \times \frac{7}{14} \times 0.437$) for a region which has been getting at least one new project location in half of all years—the maximum we observe in the data—but only by about 1 percentage point in a region which received a Chinese project in only one year ($0.4 \times \frac{1}{14} \times 0.437$).

Table 2 adds a second set of baseline results and presents alternative measures of our variable of interest. The remainder of our analysis focuses exclusively on spatial concentration within first-order regions.[37] This allows us to capture a greater number of projects, as more projects can be geocoded with a precision of states or provinces

[36] Note that the average Gini coefficient within first-order regions is 0.54, while it is 0.36 within second-order regions. Table A-2 in the Online Appendix reports summary statistics for first-order regions.

[37] We do not report results for concentration between regions as we find no robust association with Chinese-financed transport projects but report qualitatively and quantitatively similar results for second-order regions in Table A-3 of the Online Appendix.

compared to second-order regions. Column 1 takes the (logged) annual financial amount rather than the binary indicator to analyze the impact of Chinese government-financed transport projects. This has the advantage that we can account for the size of projects, but comes with the drawback that we lack information on the financial amounts for more than a third of these projects. Though the coefficient initially appears small, note that the average value of a project at a particular location is about US$7.6 million (or 15.84 log points). This implies that such an increase leads to a decrease in the Gini coefficient by 3.5 percentage points. This is very close to our baseline estimates using the dummy variable. In Column 2, we use the annual number of new project locations within a first-order region rather than the binary project indicator as variable of interest. The coefficient is two-and-a-halve times larger than in the baseline regression in column 2 of Table 1, which also recovers our baseline estimate once we consider that the average project has 2.4 locations. However, the first stage weakens noticeably (but is still above the rule-of-thumb value of ten). The results in columns 1 and 2 suggest that the effect varies at the intensive margin, that is, larger projects or projects with more locations have a greater effect on the decentralization of economic activity.[38]

[Table 2 about here.]

Column 3 uses a binary indicator for projects where it is known that they have been completed (which holds true for about 60 percent of the projects in our sample). It shows that the effect of finished projects is almost twice as large. However, these results are based on substantially fewer observations and the confidence interval includes our baseline estimate. In column 4, we broaden the definition of what constitutes an infrastructure project by including all projects that are defined as "economic infrastructure and services" according to OECD definitions.[39] The results are similar to our baseline estimates, which is not surprising given that our LATE loads heavily on physical infrastructure. Finally, column 5 of Table 2 replicates our baseline regression with an additional control variable for the construction phase. This binary indicator takes a value of one if a Chinese government-financed transport project has been committed in the previous year. Controlling for this variable is an imperfect way of accounting for the construction phase but it suggests that most of the diffusion occurs after the (assumed) completion of a project. In summary, our main result is confirmed across these variants of our baseline regression.

[38] We prefer not to overemphasize these effects as financial values are often missing and evenly split over locations. Similarly, the number of locations is in part a product of the geocoding rules, which might be one of the reasons why the first-stage relationship is weaker.

[39] In addition to transportation infrastructure, i.e., projects such as roads, railways, and airports, this category includes energy production and distribution projects, and information and communication technology (ICT) projects (e.g., broadband internet and mobile phone infrastructure).

Robustness tests

Table 3 probes the robustness of our results in two dimensions. First, we control for other potentially important shocks that could influence project allocation. For example, rising Chinese exports or commodity demand shocks may have influenced the regions to which China allocated projects. As we have discussed above, commodity price shocks and commodity cycles heterogeneously affect local incomes and to the extent that such time-varying effects on spatial concentration move systematically with the incidence of Chinese transport projects, they might bias our estimates. Second, we alter the instrument and construct a falsification test to further probe the validity of our approach.

[Table 3 about here.]

Columns 1 to 4 focus on three time-varying shocks: Chinese foreign direct investments (FDI) to a country, imports from China, and exports to China.[40] While we do not have location-specific estimates of these shocks, they are typically linked to China's demand for natural resources or increasing trade integration with China, both of which are uneven across regions. We construct time-varying and location-specific proxies for these shocks similar to our instrument by interacting these variables with two proxies for openness and market size (distance to the coast and urbanization) and two proxies for the existence of natural resources in a region (large mines or oil fields).[41] Columns 1-4 show that our results are robust to controlling for various combinations of these interacted variables. The coefficients remain significant and hardly vary in magnitude.

The next set of tests focus on the instrument itself. Column 5 modifies our detrending method and smoothes the input series using a Hodrick-Prescott-Filter (as opposed to linear detrending). The estimate is robust to this perturbation, and, if anything, becomes slightly larger. Column 6 shows results using the detrended production of Chinese steel as the time-series shock rather than the first factor of all six project inputs. The first-stage and second-stage estimates are similar to our baseline results. The main purpose of using steel only is to facilitate a comparison with the next column, which reports results from a placebo regression using United States steel production rather than Chinese production values. This presents a falsification test, as US steel production should be unrelated to Chinese-financed infrastructure projects around the world. As expected, the first-stage collapses and is extremely weak, with a Kleibergen-Paap F-statistic of just 0.38, and the 2SLS estimate becomes insignificant by a wide margin. US steel production does not help to predict the commitment of Chinese-financed transport projects. Taken together, these estimates give us confidence that our results do not hinge on the specific choice of how

[40] More specifically, we use Chinese FDI outflows (in logs of current USD) from UNCTAD. We measure imports and exports to and from China using bilateral trade flows (in logs of current USD) from the IMF Direction of Trade Statistics.

[41] All variables and sources are defined in the table notes.

we define Chinese production shocks, and that the time-series component is not driven by spurious trends.[42]

[Table 4 about here.]

Table 4 investigates the timing of the diffusion-effects of Chinese transport projects in more detail. We fix the one-year lag between the instrument and the key variable of interest, but then shift both forward and backward in time. Column 1 starts with the effects of projects in the year of commitment ($p = 0$) instrumented by the first lag of project inputs ($q = 1$). Column 5 ends with the effect of projects committed four years earlier ($p = 4$), instrumented by the fifth lag of project inputs ($q = 5$). The estimated effect is largest for the specification that corresponds to our baseline regression above (shown again in column 3 for reference). It is the only specification where the coefficient is statistically significant at the ten- or five-percent level, depending on the standard errors. The estimates become smaller for longer and shorter lags, and are imprecisely estimated (with the exception of the third lag, which is significant at the ten-percent level according to standard errors clustered at the country level). The first-stage relationship between inputs and commitments is relatively stable throughout. Taken together, this supports the notion that much of the diffusion effect occurs early on when projects are completed.

Extensions

Our finding that Chinese-financed transport projects result in a reduced concentration of economic activity within subnational regions raises the question where exactly this diffusion takes place. The monocentric city model implies that we should observe a shift in activity from central cities to their immediate periphery. The model has little to say about whether this should also increase overall activity in a region or whether this kind of development occurs by leap-frogging into undeveloped areas or integrating its suburban and peri-urban surroundings.

[Table 5 about here.]

Table 5 examines different moments of the distribution. Column 1 shows that Chinese transport projects are not robustly linked to changes in lights per capita.[43] Column 2 uses the fraction of illuminated pixels as the dependent variable. It shows that economic

[42] We also replaced the second component of our instrument—the probability to receive a project—with the analogous measure calculated with data from the Cold War period, taken from Dreher et al. (2019b). However, the first-stage F-statistic is too low, so we do not report these results here.

[43] This stands in some contrast to the results in Dreher et al. (2019b). Note however that Dreher et al. focus on all Chinese government-financed development projects rather than just transport projects, restrict their sample to Africa only, and use a different identification strategy. Moreover, their results are strongest for second-order regions. At this level, we also find positive and significant effects of transport projects on light per capita in Africa, suggesting that they have very localized benefits on development.

activity does not seem to primarily expand into previously undeveloped areas. The remaining columns of Table 5 focus on changes in economic activity across quintiles of the light distribution. This allows us to directly trace which type of changes create the overall reduction in the Gini coefficient. The pattern is fully in line with what urban land use theory would lead us to expect. We find that Chinese transport projects significantly reduce nighttime lights in the highest quintile, while they raise the share of activity taking place in the second, third and fourth quintile. This implies that Chinese transport projects gradually redistribute activity from the most densely developed parts of regions, that is, the city centers, to their immediate surroundings. That this process stops before the lowest quintile suggests that it does not reach the most rural parts of a region.

[Table 6 about here.]

Table 6 presents a more direct approach to studying from where to where the relocation of activity takes place. We report a series of regressions that split the sample along the median of several variables typically linked with rapid urban growth. The results provide further support for the conjecture that these effects occur exclusively around cities. We find a sizable decentralization of activity in regions with below-median travel time to cities, high urbanization rates, high road density, and below-median proximity to the coast. The estimated effects are substantial in those sub-samples but absent, imprecisely estimated and typically of the opposite sign in the other sub-samples. Last but not least, we also find that the effect seems to be driven by relatively poor regions (as indicated by below median light per capita), which is not surprising given that some of the poorest regions have some of the highest population growth rates and are home to many of the fastest growing cities. The evidence presented here is in line with the literature focusing on individual countries or regions discussed above. For example, in their study of the expansion of China's highway system, Baum-Snow et al. (2017) find that reductions in spatial inequality were larger within coastal and richer central regions. Similarly, although they do not focus on decentralization within regions per se, studies focusing on modelling the spatial impact of the Belt and Road Initiative typically estimate that it will benefit coastal regions, border crossings and urban hubs more (Lall and Lebrand 2020).[44]

As a final test, we investigate major world regions separately. China's global footprint is uneven. Most of its transportation infrastructure financing occurs in Africa and Asia (recall Figure 2), where urban population growth is rapid and infrastructure constraints are most severe; substantially less is devoted to other regions. Columns 1 to 3 in Table 7 split the sample by major world regions. Our main findings are driven by Africa

[44]This literature suggests that BRI projects will lead to an increasing specialization among regions and hence more concentration of economic activity in regions with better access to world markets but does not consider the distribution of activity within regions.

where the effect almost doubles compared to our baseline estimates. The coefficient on Chinese transport projects is insignificant for Asia and the Americas, although the first-stage remains about equally powerful in all three regions. This is not surprising given that Africa lags behind the other two world regions in terms of infrastructure development. It is also the region where urban primacy is most pronounced and where deficiencies in urban infrastructure have been linked to slower economic growth at the national level (Castells-Quintana 2017). Chinese-financed projects in Africa therefore appear to mitigate congestion which, eventually, could enable cities to reap the benefits of agglomeration economies.

[Table 7 about here.]

In column 4, we restrict the sample to countries classified as low income economies in 2000. It highlights that the diffusion effects of Chinese transport projects occur in the poorest countries of the world. Finally, in column 5, we restrict our analysis to only those subnational regions that have received at least one transport project from China over the entire sample period. This addresses one last identification challenge. Perhaps regions that received any development-related project are on different non-linear trends than those who did not. Our results become substantially stronger and mirror the findings for Africa once we introduce this limitation—not least because most of the projects are located in Africa.[45]

6 Conclusion

The monocentric city model predicts that transport infrastructure decentralizes economic activity within subnational regions, at least to the extent that such regions primarily consist of urban areas and their surroundings. The theory has previously been tested for single countries only, relying on identification strategies that make use of historical transport networks or other country-specific circumstances. Whether this process occurs in developing countries more broadly has not been tested.

We overcome the challenge of missing data on comparable infrastructure projects across countries and how to estimate their causal effect by focusing on infrastructure projects financed by China—a single but increasingly important source of infrastructure financing across the developing world since 2000. While many scholars and policymakers are skeptical about the quality and effects of China's development projects, its commitment to financing infrastructure is unambiguous. Connectivity has been a central

[45] Table A-4 further explores the issue of co-location with other types of projects. Our results remain similar when we control for the presence of World Bank projects in the transport sector (or in any sector) and for Chinese-financed projects in other sectors.

focus of China's Belt and Road Initiative (BRI) from its announcement in 2013 and projects financed prior to BRI similarly focus on connective infrastructure.

One of our key contributions is to provide a new geocoded dataset of China's emerging footprint around the world, much of which comes in the form of large scale infrastructure investments but extends across a variety of sectors. While our data cover the period from 2000 to 2014, and thus mostly precede the BRI, the projects we focus on share many of the characteristics of the more recent initiative. Using these data, we test whether infrastructure projects influence the spatial concentration of economic activity within recipient countries at different levels of aggregation. Our identification strategy relies on commodity inputs produced in China that affect the availability of projects over time in tandem with a variable that measures the likelihood whether countries receive a smaller or larger share of China's projects.

Our results show that Chinese government-financed transportation projects reduce the concentration of economic activity within regions in developing countries. More specifically, our results imply that the Gini coefficient measuring the spatial concentration of economic activity within first-order regions is reduced by 3.4 percentage points and 3.6 percentage points in second-order regions. These results are robust in a large number of different specifications, to the choice of control variables, and variations of the instrumental variable. The effect increases for completed projects, holds for projects financing economic infrastructure more broadly, and is largest in poor regions and low-income countries, who are most in need of infrastructure financing. In line with urban land use theory, we find that our results are driven by urban areas.

In financing major transport projects, China's government appears to be helping cities and regions in developing countries in their transformation from dense, crowded and unproductive places towards hubs of productivity. While these results are encouraging, they do not imply that Chinese government-financed transport infrastructure has only positive effects. There is growing evidence that Chinese development projects also produce negative externalities. For example, in related work, we have shown that China's "aid on demand" approach is vulnerable to domestic political capture wherein incumbent government leaders steer Chinese development projects toward their home and ethnic regions. There are many other concerns about the consequences of China's investments, ranging from their impact on the environment to questions of debt sustainability. In short, Chinese-financed transportation projects may help deal with congestion in developing countries, but our study should not be read as a comprehensive assessment of their costs and benefits. This leaves considerable scope for future research.

Acknowledgements: We owe a debt of gratitude to the large team of research assistants—including Melanie Aguilar-Rojas, Bilal Asad, Zach Baxter, Rachel Benavides, Ellie Bentley, Liliana Besosa, Allison Bowers, Ariel Cadby-Spicer, Emma Cahoon, Bree Cattelino, Alex Chadwick, Ava Chafin, Tina Chang, Yuning Chen, Mengfan Cheng, Tiffanie Choi, Miranda Clarke, Kate Connors, McKay Corbett, Graeme Cranston-Cuebas, Catherine Crowley, Alex DeGala, Hannah Dempsey, Rohan Desai, Justin DeShazor, Isabel Docampo, Joseph Dobbels, Weiwei Du, Ashton Ebert, Caleb Ebert, Aili Espigh, Claire Etheridge, Jordan Fox, Robert Francis, Ze Fu, Melanie Gilbert, Sara Gomez, Liz Hall, Thompson Hangen, Lauren Harrison, Michael Hathaway, Collin Henson, Jasmine Herndon, Elizabeth Herrity, Keith Holleran, Weijue Huang, Daniel Hughes, Torey Jackson, Jiaorui Jiang, Qi Jiang, Emmaleah Jones, Amar Kakirde, Rachel Kellogg, Connor Kennedy, Ciera Killen, Ian Kirkwood, Warren Kirkwood, Emily Koerner, Dylan Kolhoff, Lidia Kovacevic, Mirian Kreykes, Isabella Kron, Karthik Kumarappan, Daniel Lantz, Caroline Lebegue, Jade Li, Xiao Liu, Steven Livingston, Yaseen Lofti, Adriane Lopez, Flynn Madden, Dominick Margiotta, Sarah Martin, Emily McLenigan, Marie Mullins, Will Nelson, Qiuyan Ni, Jack Nicol, Alexandra Pancake, Carol Peng, Grace Perkins, Sophia Perrotti, Victor Polanco, Laura Preszler, Emily Qiu, Kamran Rahman, Sarah Reso, David Rice, Sara Rock, Ann Rogers, Elizabeth Saccoccia, Natalie Santos, Dominic Sanzotta, Faith Savaiano, Dominic Scerbo, Rebecca Schectman, Leigh Seitz, Ryan Septon, Lu Sevier, Kaitlan Shaub, Andrea Soleta, Lauren Su, Joanna Tan, Emily Tanner, Nate Tanner, Brittany Tennant, Rebecca Thorpe, Austin Trotta, Anna Umstead, Jessica Usjanauskas, Julia Varoutsos, Emily Walker, Yale Waller, Katherine Walsh, Xinyi Wang, Matt Westover, Tom Westover, James Willard, (Jiacheng) Jason Xi, Hanyang Xu, Darice Xue, Erya Yang, Antonio Tianze Ye, Jack Zhang, Yue Zhang, Echo Zhong, Joana Zhu, and Junrong Zhu—who helped to assemble and geo-locate the dataset of Chinese development projects used in this study. We also thank Seth Goodman and Miranda Lv for their spatial data quality assurance and integration assistance, as well as Gerda Asmus for sharing code for aggregating the project-level data.

This study was made possible with generous financial support from John D. and Catherine T. MacArthur Foundation, Humanity United, the William and Flora Hewlett Foundation, the Academic Research Fund of Singapore's Ministry of Education, the United Nations University World Institute for Development Economics Research (UNU-WIDER), the German Research Foundation (DFG), and the College of William and Mary. We also acknowledge that this study was indirectly made possible through a cooperative agreement (AID-OAA-A-12-00096) between USAID's Global Development Lab and AidData at the College of William and Mary under the Higher Education Solutions Network (HESN) Program, as it supported the creation of a spatial data repository and extraction tool which we used to execute our data analysis. The views expressed here do not necessarily reflect the views of the John D. and Catherine T.

MacArthur Foundation, Humanity United, the William and Flora Hewlett Foundation, Singapore's Ministry of Education, the United Nations University World Institute for Development Economics Research (UNU-WIDER), the German Research Foundation, the College of William and Mary, USAID, or the United States Government.

Last but not least, this paper received many helpful comments at research seminars, workshops and conferences. We thank Bruce Bueno de Mesquita, Vivien Foster, Erik Haustein, Robert Inklaar, Mathilde Lebrand, Ferdinand Rauch, seminar audiences at New York University (New York City, USA, September 2018), the Graduate Institute of International and Development Studies (Geneva, Switzerland, April 2019), the University of Milano-Bicocca (Milan, Italy, May 2019), the University of Hohenheim (Stuttgart, Germany, July 2019), the European Stability Mechanism (Luxembourg, November 2019), the University College Dublin (Dublin, Ireland, November 2019), the University of Groningen (Groningen, Netherlands, January 2020), the University of British Columbia (Vancouver, Canada, February 2020), as well as conference participants of the HSU-IfW-Workshop in Development and Environmental Economics (Hamburg, Germany, November 2018), the TEDE Workshop "Topics in Development and Environmental Research" at the University of Birmingham (Birmingham, UK, February 2019), the Annual Meeting of the European Public Choice Society at the Hebrew University (Jerusalem, Israel, April 2019), the GeoData in Economics Workshop at the University of Hamburg (Hamburg, Germany, May 2019), the European Meeting of the Urban Economics Association at the Vrije Universiteit (Amsterdam, Netherlands, June 2019), the Annual Conference of the Verein für Socialpolitik Research Committee Development Economics at the DIW (Berlin, Germany, June 2019), the Annual Bank Conference on Development Economics at the World Bank (Washington DC, USA, June 2019), the China Economics Summer Institute at Peking University (Beijing, China, August 2019), the Annual Economic Research Southern Africa Workshop on "Structural Constraints on the Economy, Growth and Political Economy" at the University of the Witwatersrand (Johannesburg, South Africa, September 2019), the Biennial Conference of the Economic Society of South Africa (Johannesburg, South Africa, September 2019), and the Annual Meeting of the International Political Economy Society at the University of California (San Diego, USA, November 2019) for comments on earlier versions of this paper.

References

Ades, A. F. and E. L. Glaeser (1995). Trade and Circuses: Explaining Urban Giants. *Quarterly Journal of Economics 110*(1), 195–227.

African Development Bank (2014a). Kenya: Nairobi-Thika Highway Improvement Project. Technical report, African Development Bank, Abidjan.

African Development Bank (2014b). Study on Quality of Bank Financed Road Projects. Technical report, African Development Bank, Abidjan.

African Development Bank (2016). PCR Evaluation Note for Public Sector Operations: Nairobi-Thika Highway Improvement Project. Technical report, African Development Bank, Abidjan.

African Development Bank (2019). Nairobi-Thika Highway Improvement Project: Project Completion Report. Technical report, African Development Bank, Abidjan.

African Development Fund (2007). Appraisal Report: Nairobi–Thika Highway Improvement Project. Technical report, African Development Fund, Tunis.

Alonso, W. et al. (1964). *Location and Land Use. Toward a General Theory of Land Rent.* Cambridge, Mass.: Harvard University Press.

Andrés, L., D. Biller, and M. H. Dappe (2014). *Infrastructure Gap in South Asia: Infrastructure Needs, Prioritization, and Financing.* The World Bank.

Angrist, J. D. and J.-S. Pischke (2008). *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press.

Atkin, D. and D. Donaldson (2015). Who's Getting Globalized? The Size and Implications of Intra-national Trade Costs. Working Paper 21439, National Bureau of Economic Research.

Autor, D. H., D. Dorn, and G. H. Hanson (2016). The China Shock: Learning from Labor-Market Adjustment to Large Changes in Trade. *Annual Review of Economics 8*, 205–240.

Bandiera, L. and V. Tsiropoulos (2020). A Framework to Assess Debt Sustainability under the Belt and Road Initiative. *Journal of Development Economics 146*, forthcoming.

Banerjee, A., E. Duflo, and N. Qian (2012). On the Road: Access to Transportation Infrastructure and Economic Growth in China. Working Paper 17897, National Bureau of Economic Research.

Bartke, W. (1989). *The Economic Aid of the PR China to Developing and Socialist Countries* (2nd ed.). K. G. Saur.

Baum-Snow, N. (2007). Did Highways Cause Suburbanization? *Quarterly Journal of Economics 122*(2), 775–805.

Baum-Snow, N. (2014). Urban Transport Expansions, Employment Decentralization, and the Spatial Scope of Agglomeration Economies. *Unpublished Manuscript, Brown University, Providence*.

Baum-Snow, N., L. Brandt, J. V. Henderson, M. Turner, and Q. Zhang (2017). Roads, Railroads, and Decentralization of Chinese Cities. *Review of Economics and Statistics 99*(3), 435–448.

Baum-Snow, N., J. V. Henderson, M. A. Turner, Q. Zhang, and L. Brandt (2020). Does Investment in National Highways Help or Hurt Hinterland City Growth? *Journal of Urban Economics 115*, 103–124.

Baum-Snow, N. and M. A. Turner (2017). Transport Infrastructure and the Decentralization of Cities in the People's Republic of China. *Asian Development Review 34*(2), 25–50.

Bayes, A. (2007). *Impact Assessment of Jamuna Multipurpose Bridge Project (JMBP) on Poverty.* Dhaka, Bangladesh: Japan Bank for International Cooperation.

BenYishay, A., B. Parks, D. Runfola, and R. Trichler (2016). Forest Cover Impacts of Chinese Development Projects in Ecologically Sensitive Areas. AidData Working Paper #32, AidData at William & Mary.

Berman, N. and M. Couttenier (2015). External Shocks, Internal Shots: The Geography of Civil Conflicts. *Review of Economics and Statistics 97*(4), 758–776.

Bird, J., M. Lebrand, and A. J. Venables (2020). The Belt and Road Initiative: Reshaping Economic Geography in Central Asia? *Journal of Development Economics 144*, forthcoming.

Bird, J. and S. Straub (2014). The Brasilia Experiment: Road Access and the Spatial Pattern of Long-Term Local Development in Brazil. *World Bank Policy Research Working Paper*.

Bluhm, R. and M. Krause (2018). Top Lights: Bright Cities and Their Contribution to Economic Development. AidData Working Paper #67, AidData at William & Mary.

Bonfatti, R. and S. Poelhekke (2017). From Mine to Coast: Transport Infrastructure and the Direction of Trade in Developing Countries. *Journal of Development Economics 127*, 91–108.

Borusyak, K., P. Hull, and X. Jaravel (2018). Quasi-experimental Shift-Share Research Designs. Working Paper 24997, National Bureau of Economic Research.

Brazys, S., J. A. Elkink, and G. Kelly (2017). Bad Neighbors? How Co-located Chinese and World Bank Development Projects Impact Local Corruption in Tanzania. *Review of International Organizations 12*(2), 227–253.

Brazys, S. and K. C. Vadlamannati (2018). Aid Curse with Chinese Characteristics? Chinese Development Flows and Economic Reforms. AidData Working Paper #52, AidData at William & Mary.

Bruederle, A. and R. Hodler (2017). Nighttime Lights as a Proxy for Human Development at the Local Level. CESifo Working Paper 6555. CESifo Group Munich.

Brülhart, M., K. Desmet, and G.-P. Klinke (2019). The Shrinking Advantage of Market Potential. CEPR Discussion Paper No. 14157, CEPR.

Bräutigam, D. (2009). *The Dragon's Gift: The Real Story of China in Africa.* Oxford, UK: Oxford University Press.

Burchfield, M., H. G. Overman, D. Puga, and M. A. Turner (2006). Causes of Sprawl: A Portrait from Space. *Quarterly Journal of Economics 121*(2), 587–633.

Castells-Quintana, D. (2017). Malthus Living in a Slum: Urban Concentration, Infrastructure and Economic Growth. *Journal of Urban Economics 98*, 158–173.

Cervero, R. (2013). Linking Urban Transport and Land Use in Developing Countries. *Journal of Transport and Land Use 6*(1), 7–24.

CIESIN and ITOS (2013). Global Roads Open Access Data Set, Version 1 (gROADSv1). NASA Socioeconomic Data and Applications Center (SEDAC).

de Soyres, F., A. Mulabdic, and M. Ruta (2020). Common Transport Infrastructure: A Quantitative Model and Estimates from the Belt and Road Initiative. *Journal of Development Economics 143*, forthcoming.

Dollar, D. (2008). Supply Meets Demand: Chinese Infrastructure Financing in Africa. World Bank Blog. 10 July 2008. Accessed at https://blogs.worldbank.org/eastasiapacific/supply-meets-demand-chinese-infrastructure-financing-in-africa.

Donaldson, D. (2018). Railroads of the Raj: Estimating the Impact of Transportation Infrastructure. *American Economic Review 108*(4-5), 899–934.

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

Dreher, A., A. Fuchs, R. Hodler, B. C. Parks, P. A. Raschky, and M. J. Tierney (2019a). African Leaders and the Geography of China's Foreign Assistance. *Journal of Development Economics 140*, 44–71.

Dreher, A., A. Fuchs, R. Hodler, B. C. Parks, P. A. Raschky, and M. J. Tierney (2019b). Is Favoritism a Threat to Chinese Aid Effectiveness? A Subnational Analysis of Chinese Development Projects. *CEPR Discussion Paper 13840*.

Dreher, A., A. Fuchs, B. Parks, A. Strange, and M. Tierney (2020). Aid, China, and Growth: Evidence from a New Global Development Finance Dataset.

Dreher, A., A. Fuchs, B. C. Parks, A. Strange, and M. J. Tierney (2018). Apples and Dragon Fruits: The Determinants of Aid and Other Forms of State Financing from China to Africa.

International Studies Quarterly 62, 182–194.

Duranton, G. and D. Puga (2015). Urban Land Use. In *Handbook of Regional and Urban Economics*, Volume 5, pp. 467–560. Elsevier.

Duranton, G. and M. A. Turner (2012). Urban Growth and Transportation. *Review of Economic Studies 79*(4), 1407–1440.

Eichenauer, V. Z., A. Fuchs, and L. Brückner (2018). The Effects of Trade, Aid and Investment on China's Image in Latin America. Department of Economics Discussion Paper 646, Heidelberg University.

Faber, B. (2014). Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System. *Review of Economic Studies 81*(3), 1046–1070.

Fajgelbaum, P. and S. Redding (2018). Trade, Structural Transformation and Development: Evidence from Argentina 1869-1914. *Revised Version of NBER Working Paper 20217*.

Fujita, M. and H. Ogawa (1982). Multiple Equilibria and Structural Transition of Non-monocentric Urban Configurations. *Regional Science and Urban Economics 12*(2), 161–196.

Gehring, K., L. Kaplan, and M. H. Wong (2019). China and the World Bank: How Contrasting Development Approaches Affect the Stability of African States. AidData Working Paper #70, AidData at William & Mary.

Gibbons, S., T. Lyytikäinen, H. G. Overman, and R. Sanchis-Guarner (2019). New Road Infrastructure: The Effects on Firms. *Journal of Urban Economics 110*, 35–50.

Glaeser, E. L. and M. E. Kahn (2004). Sprawl and Urban Growth. In *Handbook of Regional and Urban Economics*, Volume 4, pp. 2481–2527. Elsevier.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020). Bartik Instruments: What, When, Why, and How. *American Economic Review* (forthcoming).

Gollin, D., R. Jedwab, and D. Vollrath (2016). Urbanization With and Without Industrialization. *Journal of Economic Growth 21*(1), 35–70.

Guillon, M. and J. Mathonnat (2020). What Can We Learn on Chinese Aid Allocation Motivations from Available Data? A Sectorial Analysis of Chinese Aid to African Countries. *China Economic Review 60*, forthcoming.

Henderson, J. and A. Kuncoro (1996). Industrial Centralization in Indonesia. *World Bank Economic Review 10*(3), 513–540.

Henderson, J., T. Squires, A. Storeygard, and D. Weil (2018). The Global Distribution of Economic Activity: Nature, History, and the Role of Trade. *Quarterly Journal of Economics 133*, 357–406.

Henderson, V. and A. Mitra (1996). The New Urban Landscape: Developers and Edge Cities. *Regional Science and Urban Economics 26*(6), 613–643.

Hernandez, D. (2017). Are "New" Donors Challenging World Bank Conditionality. *World Development 96*, 529–549.

Hicks, R., B. Parks, T. Roberts, and M. Tierney (2008). *Greening Aid? Understanding the Environmental Impact of Development Assistance*. Oxford, UK: Oxford University Press.

Hodler, R. and P. A. Raschky (2014). Regional Favoritism. *Quarterly Journal of Economics 129*(2), 995–1033.

Humphrey, C. and K. Michaelowa (2019). China in Africa: Competition for Traditional Development Finance Institutions? *World Development 120*(C), 15–28.

Information Office of the State Council (2011). *White Paper on China's Foreign Aid*. Xinhua/Information Office of the State Council, People's Republic of China, Beijing, China.

Isaksson, A.-S. and A. Kotsadam (2018a). Chinese Aid and Local Corruption. *Journal of Public Economics 159*, 146–159.

Isaksson, A.-S. and A. Kotsadam (2018b). Racing to the Bottom? Chinese Development Projects and Trade Union Involvement in Africa. *World Development 106*, 284–298.

Jean, N., M. Burke, M. Xie, M. Davis, D. Lobell, and S. Ermon (2016). Combining Satellite Imagery and Machine Learning to Predict Poverty. *Science 353*, 790–794.

KARA and CSUD (2012). The Social/Community Component of the Analysis of the Thika Highway Improvement Project. Kenya Alliance of Resident Associations (KARA) and the Center for Sustainable Urban Development (CSUD). Accessed at http://csud.ei.columbia.edu/files/2012/11/KARA-report˙FINAL.pdf.

Krugman, P. (1991). Increasing Returns and Economic Geography. *Journal of Political Economy 99*(3), 483–499.

Lall, S. V., J. V. Henderson, and A. J. Venables (2017). *Africa's Cities: Opening Doors to the World*. Washington, DC: World Bank.

Lall, S. V. and M. Lebrand (2020). Who Wins, Who Loses? Understanding the Spatially Differentiated Effects of the Belt and Road Initiative. *Journal of Development Economics*, forthcoming.

Lessmann, C. and A. Seidel (2017). Regional Inequality, Convergence, and its Determinants—A View from Outer Space. *European Economic Review 92*, 110–132.

Lujala, P., J. Ketil Rod, and N. Thieme (2007). Fighting Over Oil: Introducing a New Dataset. *Conflict Management and Peace Science 24*(3), 239–256.

Maina, E., A. Gachanja, M. Gatari, and H. Price (2018). Demonstrating PM2.5 and Road-Side Dust Pollution by Heavy Metals along Thika Superhighway in Kenya. *Environmental Monitoring and Assessment 190*, 251.

Martorano, B., L. Metzger, and M. Sanfilippo (2020). Chinese Development Assistance and Household Welfare in Sub-Saharan Africa. *World Development 129*, forthcoming.

Mauk, B. (2019). Can China Turn the Middle of Nowhere Into the Center of the World Economy? *New York Times*.

Michalopoulos, S. and E. Papaioannou (2014). National Institutions and Subnational Development in Africa. *Quarterly Journal of Economics 129*(1), 151–213.

Mills, E. S. (1967). An Aggregative Model of Resource Allocation in a Metropolitan Area. *American Economic Review 57*(2), 197–210.

Muth, R. F. (1969). *Cities and Housing: The Spatial Pattern of Urban Residential Land Use.* University of Chicago. Graduate School of Business.

Nelson, A. (2008). Estimated Travel Time to the Nearest City of 50,000 or More People in Year 2000. European Commission, Joint Research Centre (JRC).

Nielson, D. and M. Tierney (2003). Delegation to International Organizations: Agency Theory and World Bank Environmental Reform. *International Organization 57*(2), 241–276.

Nunn, N. and N. Qian (2014). U.S. Food Aid and Civil Conflict. *American Economic Review 104*(6), 1630–1666.

Ogawa, H. and M. Fujita (1980). Equilibrium Land Use Patterns in a Nonmonocentric City. *Journal of Regional Science 20*(4), 455–475.

Perlez, J. and Y. Huang (2017). Behind China's $1 Trillion Plan to Shake Up the Economic Order. *New York Times*.

Pesaresi, M., A. Florczyk, M. Schiavina, M. Melchiorri, and L. Maffenini (2019). GHS Settlement Grid, Updated and Refined REGIO Model 2014 in Application to GHS-BUILT R2018A and GHS-POP R2019A, Multitemporal (1975-1990-2000-2015) R2019A. European Commission, Joint Research Centre (JRC).

Pucher, J., N. Korattyswaropam, N. Mittal, and N. Ittyerah (2005). Urban Transport Crisis in India. *Transport Policy 12*(3), 185–198.

Puga, D. (1999). The Rise and Fall of Regional Inequalities. *European Economic Review 43*(2), 303–334.

Roessler, P., Y. Pengl, R. Marty, K. Titlow, and N. van de Walle (2018). On the Origins of Spatial Inequality in Africa. Working Paper.

Rossi-Hansberg, E., P.-D. Sarte, and R. Owens III (2009). Firm Fragmentation and Urban Patterns. *International Economic Review 50*(1), 143–186.

Rozenberg, J. and M. Fay (2019). *Beyond the Gap: How Countries Can Afford the Infrastructure They Need While Protecting the Planet*. The World Bank.

Saiz, A. (2010). The Geographic Determinants of Housing Supply. *Quarterly Journal of Economics 125*(3), 1253–1296.

Storeygard, A. (2016). Farther on Down the Road: Transport Costs, Trade and Urban Growth in Sub-Saharan Africa. *Review of Economic Studies 83*(3), 1263–1295.

Strandow, D., M. Findley, D. Nielson, and J. Powell (2011). The UCDP-AidData Codebook on Geo-Referencing Foreign Aid. Version 1.1, Uppsala Conflict Data Program.

Strange, A. M., A. Dreher, A. Fuchs, B. Parks, and M. J. Tierney (2018). Tracking Underreported Financial Flows: China's Development Finance and the Aid–Conflict Nexus Revisited. *Journal of Conflict Resolution 61*(5), 935–963.

Strange, A. M., S. Ghose, B. Russel, M. Cheng, and B. Parks (2017). AidData's Methodology for Tracking Underreported Financial Flows. Version 1.3. *AidData*.

Swedlund, H. J. (2017). *The Development Dance: How Donors and Recipients Negotiate the Delivery of Foreign Aid*. Ithaca, NY: Cornell University Press.

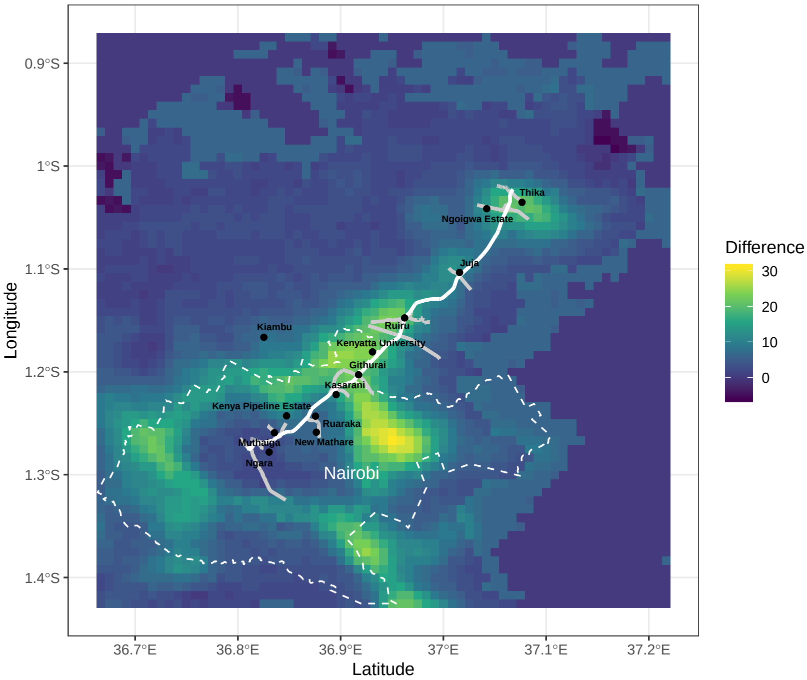
Wade, A. (2008). Time for the West to Practise What It Preaches. *Financial Times January 23*.

Weidmann, N. B. and S. Schutte (2017). Using Night Lights for the Prediction of Local Wealth. *Journal of Peace Research 54*(2), 125–140.

Xi, J. (2017). Work Together to Build the Silk Road Economic Belt and the 21st Century Maritime Silk Road. A Keynote Speech at the Opening Ceremony of the Belt and Road Forum (BRF) for International Cooperation in Beijing, China.

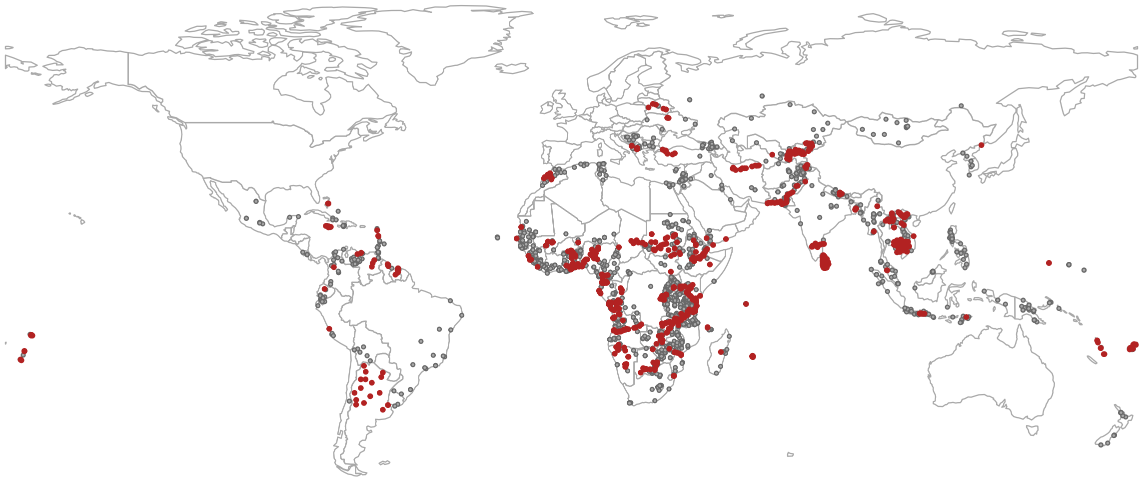
Zeitz, A. (2020). Emulation or Differentiation? China's Development Finance and Traditional Donor Aid in Developing Countries. *Review of International Organizations*. forthcoming.

Figure 1 – Nairobi-Thika Highway, change in nighttime lights, 2008–2013



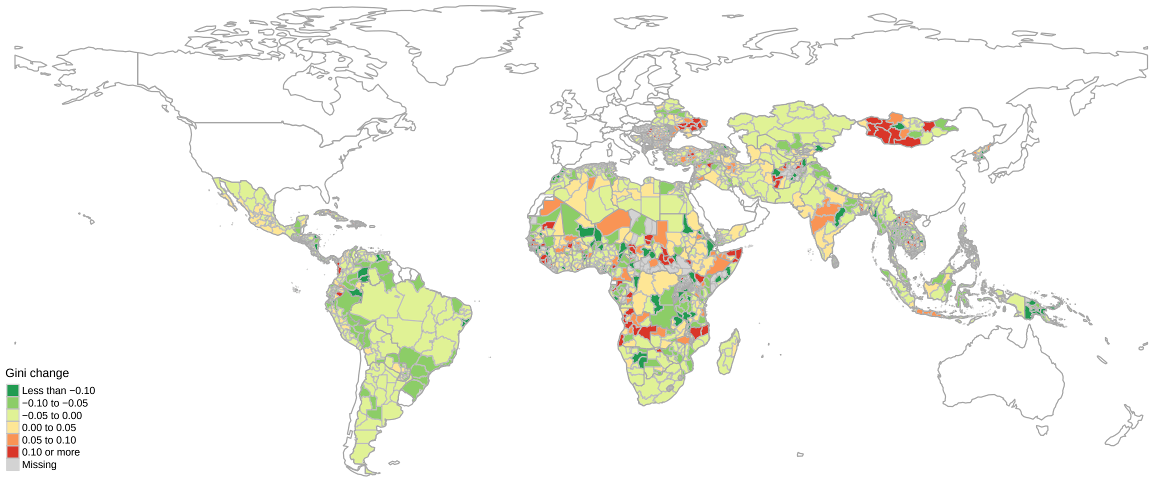
Notes: The figure illustrates the change in nighttime lights from 2008 to 2013 along the route of the Nairobi-Thika Highway, which was constructed from January 2009 until October 2012. Major intersections and points of interest are highlighted along the highway. The change in nighttime lights is the difference between the F18 2013 image (in DN from 0 to 63) and the F16 2008 image (in the same units). The differences have a range from -6 to 31 DN. The expansion of light around Nairobi is related to other infrastructure projects, many of which are also Chinese-financed but not highlighted here.

Figure 2 – Locations of Chinese-financed projects, transport and non-transport, 2000–2014



Notes: The figure illustrates all Chinese-financed transport (red) and non-transport projects (grey) which were committed and implemented in the period from 2000 to 2014. It shows a total of 2,140 projects in 4,420 discrete locations which have a precision accuracy of (at least) a first-order administrative division. 1,345 projects have a precision accuracy less than the first-order region (not shown). Although there are "only" 269 transportation projects, 1,211 of the 4,420 locations shown in the figure have directly received (some part) of a larger transportation project.

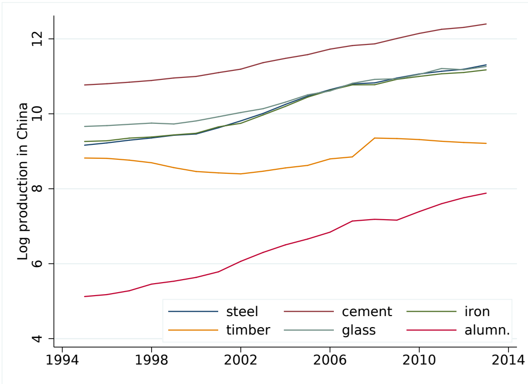
Figure 3 – Long differences in spatial concentration, within ADM1, 2000–2013



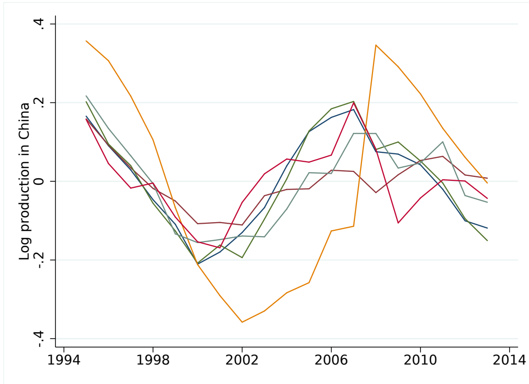
Notes: The figure illustrates the cross-regional and temporal variation in spatial concentration. It shows long differences in the Gini coefficient of inequality within first-order regions, that is, a region's value in 2013 minus the value in 2000. Only countries that are not classified as high-income by the World Bank are shown. Missing values occur when there were too few lit cells to compute the Gini coefficient in the initial or final period.

Figure 4 – Variations in overcapacity of project inputs, 1995–2013

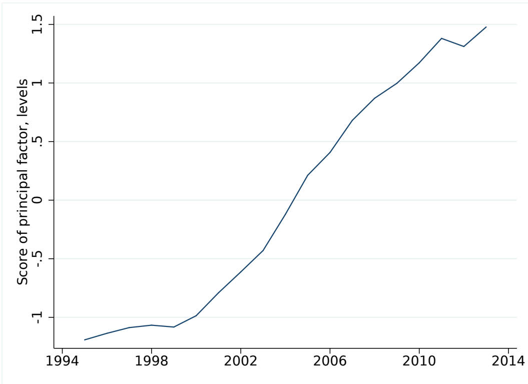
(a) Log levels



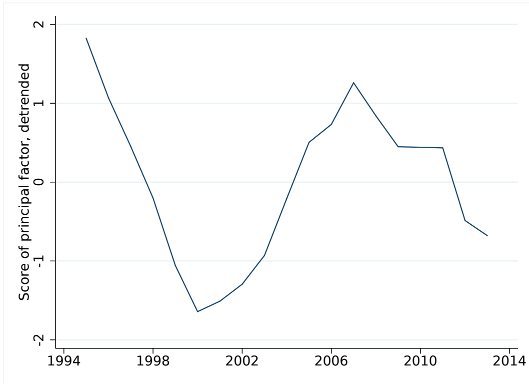
(b) Detrended



(c) First common factor of (a)



(d) First common factor of (b)



Notes: The figure illustrates the time variation in the production of physical project inputs in China. Panel a shows the raw data over time (in logarithms). Panel b shows the linearly detrended series. Panel c shows the first common factor of all level series in panel a. Panel d shows the first common factor of all detrended series in panel b. The annual data for steel, cement, pig iron, timber and glass have been obtained from the National Bureau of Statistics of China. The time series for aluminum has been obtained from the Minerals Yearbook by the US Geological Survey.

Figure 5 – Parallel pre-trends, 1992–1999

Notes: The figure illustrates the average Gini coefficient of light intensity within first-order administrative regions over time in the period before China became increasingly active in funding transport projects in other countries. The time series is reported separately for regions which will eventually receive a project in the 2000–2014 period and those regions which will not.

Table 1 – Transport projects and concentration within and between regions, 2002–2013

	Spatial concentration, $\Delta\textsc{Gini}_{jit}$, measured ...			
	Within second-order regions (1)	Within first-order regions (2)	Between second-order regions (3)	Between first-order regions (4)
	Panel a) OLS estimates			
Projects ($\Delta N_{i,t-2}$)	-0.0021	0.0031	-0.0061	-0.0030
	(0.0040)	(0.0050)	(0.0056)	(0.0038)
	[0.0053]	[0.0056]	[0.0089]	[0.0038]
	Panel b) Reduced-form estimates			
IV ($F_{t-3} \times \bar{\rho}_{ir}$)	-0.0161	-0.0149	-0.0101	-0.0054
	(0.0075)**	(0.0078)*	(0.0085)	(0.0068)
	[0.0083]*	[0.0080]*	[0.0075]	[0.0068]
	Panel c) 2SLS estimates			
Projects ($\Delta N_{i,t-2}$)	-0.0363	-0.0342	-0.0215	-0.0162
	(0.0182)**	(0.0158)**	(0.0169)	(0.0201)
	[0.0187]*	[0.0178]*	[0.0147]	[0.0201]
	Panel d) First-stage estimates			
IV ($F_{t-3} \times \bar{\rho}_{ir}$)	0.4431	0.4367	0.4712	0.3361
	(0.1186)***	(0.0772)***	(0.0855)***	(0.0638)***
	[0.1023]***	[0.0688]***	[0.0737]***	[0.0636]***
Level of analysis	ADM2	ADM1	ADM1	ADM0
First-stage F-Stat	13.95	32.04	30.36	27.73
Observations	343,682	26,759	22,452	1,386
Regions	31,428	2,385	2,015	–
Countries	112	122	111	122

Notes: The table reports regression results. Panel a shows least squares fixed effects regressions where the dependent variable is indicated in the column header. Panel b shows reduced-form regressions where the dependent variable is indicated in the column header. Panel c shows two-stage least squares fixed effects regressions where the dependent variable is indicated in the column header. Panel d shows the corresponding first-stage regressions where the dependent variable is a binary indicator for new project commitments ($\Delta N_{ji,t-2}$) in a region. Columns 1–3 include region fixed effects and country-year fixed effects; column 4 includes country fixed effects and year fixed effects. Standard errors clustered at the country level are reported in parentheses. Conley errors with a spatial cutoff of 500 km and a time-series HAC with a lag cutoff of 1,000 years are reported in brackets.

Table 2 – Variants of baseline regression, within ADM1, 2002–2013

	Variations of the variable of interest				
	Financial values	Location count	Completed	Economic infra.	Transport
	(1)	(2)	(3)	(4)	(5)
	Panel a) 2SLS estimates				
Δ Projects $(t-2)$	-0.0022	-0.0136	-0.0659	-0.0312	-0.0279
	(0.0010)**	(0.0065)**	(0.0349)*	(0.0112)***	(0.0147)*
	[0.0011]*	[0.0076]*	[0.0346]*	[0.0128]**	[0.0166]*
	Panel b) First-stage estimates				
IV $(F_{t-3} \times \bar{\rho}_{ir})$	6.7896	1.1171	0.2287	0.1983	0.4863
	(1.1521)***	(0.3277)***	(0.0542)***	(0.0364)***	(0.0889)***
	[1.2044]***	[0.2783]***	[0.0593]***	[0.0864]***	[0.0803]***
Construction phase	–	–	–	–	✓
First-stage F-stat	33.30	10.91	17.05	29.47	28.50
Observations	27,013	27,013	27,013	27,013	27,013
Regions	2,399	2,399	2,399	2,399	2,399
Countries	122	122	122	122	122

Notes: The table reports regression results. Panel a shows two-stage least squares fixed effects regressions where the dependent variable is the first difference of the Gini coefficient of light intensity within first-order administrative regions. Panel b shows least squares fixed effects regressions where the dependent variable is indicated in the column header. 'Financial values' refers to the (logged) financial value of project commitments in constant 2014 US dollar. 'Location count' is the number of project locations of newly committed projects in a region. 'Completed' is a binary indicator for completed projects in a region. 'Economic infrastructure' broadens the definition of our base measure by including all economic infrastructure projects (transportation, energy production and distribution, and information and communication technology). The last column shows the baseline measure with an added control variable for the construction phase. All specifications include region-fixed effects and country-year-fixed effects. Standard errors clustered at the country level are reported in parentheses. Conley errors with a spatial cutoff of 500 km and a time-series HAC with a lag cutoff of 1,000 years are reported in brackets.

Table 3 – Identification: Other "China shocks" and instrument variations, within ADM1, 2002–2013

	Accounting for other "China shocks"				*Altering the instrument*		
	Log FDI (1)	Log imports (2)	Log exports (3)	All (4)	HP Filtered (5)	Det. Steel (6)	US Placebo (7)
	Panel a) 2SLS estimates						
Projects ($\Delta N_{i,t-2}$)	-0.0338	-0.0333	-0.0325	-0.0328	-0.0391	-0.0375	-0.0358
	(0.0158)**	(0.0155)**	(0.0153)**	(0.0154)**	(0.0189)**	(0.0192)*	(0.2097)
	[0.0178]*	[0.0175]*	[0.0173]*	[0.0175]*	[0.0213]*	[0.0212]*	[0.2214]
	Panel b) First-stage estimates						
IV ($F_{t-3} \times \bar{\rho}_{ir}$)	0.4369	0.4545	0.4578	0.4550	0.3832	0.3611	0.0059
	(0.0771)***	(0.0773)***	(0.0774)***	(0.0774)***	(0.0712)***	(0.0666)***	(0.0097)
	[0.0719]***	[0.0730]***	[0.0730]***	[0.0730]***	[0.0680]***	[0.0645]***	[0.0105]
Other China shock × Distance to coast	✓	✓	–	✓	✓	✓	✓
Other China shock × Urbanization	✓	✓	–	✓	✓	✓	✓
Other China shock × Large mines	–	–	✓	✓	✓	✓	✓
Other China shock × Oil fields	–	–	✓	✓	✓	✓	✓
First-stage F-stat	32.11	34.55	34.97	34.59	28.95	29.37	0.38
Observations	26,759	26,098	26,098	26,098	26,098	26,098	26,098
Regions	2,385	2,327	2,327	2,327	2,327	2,327	2,327
Countries	122	116	116	116	116	116	116

Notes: The table reports regression results. Panel a shows two-stage least squares fixed effects regressions where the dependent variable is the first difference of the Gini coefficient of light intensity within first-order administrative regions. Panel b shows least squares fixed effects regressions where the dependent variable is a binary indicator for new project commitments ($\Delta N_{ji,t-2}$) in a region. 'FDI' are Chinese FDI outflows (in logs of current USD) from UNCTAD. 'Distance to coast' is the average "as-the-crow-flies" distance from the region to the nearest coastline (from Natural Earth). 'Urbanization' is measured as the fraction of land which is defined as an urban cluster or urban center in 2000 by the Global Human Settlement Layer (Pesaresi et al. 2019). 'Imports' are bilateral imports from China (in logs of current USD) from the IMF Direction of Trade Statistics. 'Exports' is the value of donor country exports to China (in logs of current USD) from the IMF Direction of Trade Statistics. 'Large mines' indicate if the region has at least one major mineral deposit in 2005 according to the United States Geological Survey. 'Oil fields' indicate if the region has at least one major on-shore oil or gas field (Lujala et al. 2007). 'HP Filtered' implies that the factor inputs were non-linearly detrended using a Hodrick-Prescott high-pass filter before the first factor was extracted. 'Det. Steel' uses the linearly detrended log of Chinese steel production from the National Bureau of Statistics of China as the time-series shock. We standardize this variable before multiplying it with the exposure term so that the coefficient is comparable with that using the first common factor of all inputs. 'US Placebo' uses a US raw steel production index from FRED hosted by the Federal Reserve Bank of St. Louis (Series IPN3311A2RN) as a placebo instrument. These data have not been filtered, as there is no strong underlying trend. All specifications include region fixed effects and country-year fixed effects. Standard errors clustered at the country level are reported in parentheses. Conley errors with a spatial cutoff of 500 km and a time-series HAC with a lag cutoff of 1,000 years are reported in brackets.

Table 4 – Timing of effects, within ADM1

	Lag structure for $\Delta N_{i,t-p}$ *and* F_{t-q}				
	$p = 0$	$p = 1$	$p = 2$	$p = 3$	$p = 4$
	$q = 1$	$q = 2$	$q = 3$	$q = 4$	$q = 5$
	(1)	(2)	(3)	(4)	(5)
	Panel a) 2SLS estimates				
Projects ($\Delta N_{i,t-p}$)	-0.0145	-0.0222	-0.0342	-0.0214	-0.0203
	(0.0209)	(0.0153)	(0.0158)**	(0.0128)*	(0.0168)
	[0.0199]	[0.0190]	[0.0178]*	[0.0136]	[0.0129]
	Panel b) First-stage estimates				
IV ($F_{t-q} \times \bar{\rho}_{ir}$)	0.4583	0.4682	0.4369	0.4047	0.3878
	(0.0884)***	(0.0844)***	(0.0772)***	(0.0784)***	(0.0788)***
	[0.0733]***	[0.0724]***	[0.0719]***	[0.0727]***	[0.0732]***
First-stage F-stat	26.85	30.76	32.04	26.64	24.24
Observations	31,563	29,076	26,754	24,430	22,126
Regions	2,464	2,386	2,385	2,371	2,368
Countries	123	122	122	121	121

Notes: The table reports regression results. Panel a shows two-stage least squares fixed effects regressions where the dependent variable is the Gini coefficient of light intensities within first-order administrative regions. Panel b shows least squares fixed effects regressions where the dependent variable is a binary indicator for new project commitments ($\Delta N_{ji,t-2}$) in a region. The lag structure for the first difference of projects and the instrument is indicated in the column header. All specifications include region fixed effects and country-year fixed effects. Standard errors clustered at the country level are reported in parentheses. Conley errors with a spatial cutoff of 500 km and a time-series HAC with a lag cutoff of 1,000 years are reported in brackets.

Table 5 – Light intensity and quintile shares, within ADM1, 2002–2013

	Moments of spatial concentration						
	Log light	Extensive	Quintile shares				
	per capita	margin	0–20%	20–40%	40–60%	60–80%	80–100%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel a) 2SLS estimates						
Projects ($\Delta N_{i,t-2}$)	0.0233	0.0122	-0.0028	0.0055	0.0134	0.0215	-0.0375
	(0.0443)	(0.0173)	(0.0031)	(0.0027)**	(0.0047)***	(0.0118)*	(0.0162)**
	[0.0487]	[0.0104]	[0.0043]	[0.0042]	[0.0060]**	[0.0104]**	[0.0182]**
	Panel b) First-stage estimates						
IV ($F_{t-3} \times \bar{\rho}_{ir}$)	0.4498	0.4505	0.4488	0.4488	0.4488	0.4488	0.4488
	(0.0732)***	(0.0732)***	(0.0877)***	(0.0877)***	(0.0877)***	(0.0877)***	(0.0877)***
	[0.0688]***	[0.0688]***	[0.0728]***	[0.0728]***	[0.0728]***	[0.0728]***	[0.0728]***
First-stage F-stat	37.72	37.86	26.16	26.16	26.16	26.16	26.16
Observations	28,025	28,037	24,639	24,639	24,639	24,639	24,639
Regions	2,439	2,440	2,242	2,242	2,242	2,242	2,242
Countries	122	122	121	121	121	121	121

Notes: The table reports regression results. Panel a shows two-stage least squares fixed effects regressions where the dependent variable is indicated in the column header. Panel b shows least squares fixed effects regressions where the dependent variable is a binary indicator for new project commitments ($\Delta N_{ji,t-2}$) in a region. The extensive margin is defined as the fraction of pixels with a non-zero light density. All specifications include region fixed effects and country-year fixed effects. Standard errors clustered at the country level are reported in parentheses. Conley errors with a spatial cutoff of 500 km and a time-series HAC with a lag cutoff of 1,000 years are reported in brackets.

Table 6 – Sample splits, within ADM1, 2002–2013

	Splitting at the median of ...				
	City access	Urbanization rate	Road density	Distance to coast	Light per capita
	(1)	(2)	(3)	(4)	(5)
	Panel a) Below median, 2SLS estimates				
Projects ($\Delta N_{i,t-2}$)	-0.0681	0.0516	-0.0002	-0.0404	-0.0571
	(0.0214)***	(0.0498)	(0.0327)	(0.0248)	(0.0204)***
	[0.0394]*	[0.0448]	[0.0289]	[0.0186]**	[0.0246]**
	Panel b) Above median, 2SLS estimates				
Projects ($\Delta N_{i,t-2}$)	0.0150	-0.0708	-0.0391	-0.0328	0.0045
	(0.0305)	(0.0259)***	(0.0132)***	(0.0221)	(0.0376)
	[0.0297]	[0.0302]**	[0.0263]	[0.0312]	[0.0452]
First-stage F-stat a)	9.86	17.80	19.61	22.94	21.06
First-stage F-stat b)	38.19	11.52	15.40	15.35	13.72
Observations a)	12,819	13,457	13,320	13,312	13,007
Observations b)	13,660	13,017	13,251	13,352	13,396

Notes: The table reports regression results. Panel a shows two-stage least squares fixed effects regressions for first-order regions with below median values of the variable indicated in the column header. Panel b shows two-stage least squares fixed effects regressions for first-order regions with above median values of the variable indicated in the column header. 'City access' is measured as the travel time to the nearest city of 50,000 or more people in the year 2000 (Nelson 2008). The 'urbanization rate' is measured as the fraction of land in the region which is defined as an urban cluster or urban center in 2000 by the Global Human Settlement Layer (Pesaresi et al. 2019). 'Road density' is measured as the total road length over the area of the region where road length is derived from the gROADS data set (CIESIN and ITOS 2013). 'Distance to coast' is the average "as-the-crow-flies" distance to the nearest coastline (from Natural Earth). 'Light per capita' is the sum of light in a region divided by its population in 2000 (from the Global Human Settlement Layer). All specifications include region fixed effects and country-year fixed effects. Standard errors clustered at the country level are reported in parentheses. Conley errors with a spatial cutoff of 500 km and a time-series HAC with a lag cutoff of 1,000 years are reported in brackets.

Table 7 – Regional variation, within ADM1

	Regional subsets and related sample perturbations				
	Africa	Asia	Americas	Low income	$\bar{\rho}^{\text{all}} > 0$
	(1)	(2)	(3)	(4)	(5)
	Panel a) 2SLS estimates				
Projects ($\Delta N_{i,t-2}$)	-0.0563	-0.0132	0.0048	-0.0463	-0.0548
	(0.0212)**	(0.0268)	(0.0168)	(0.0183)**	(0.0162)***
	[0.0287]*	[0.0154]	[0.0157]	[0.0225]**	[0.0187]***
	Panel b) First-stage estimates				
$F_{t-3} \times \bar{\rho}_{ir}$	0.4343	0.4183	0.7981	0.4285	0.4316
	(0.1148)***	(0.1071)***	(0.2285)***	(0.0933)***	(0.0781)***
	[0.1006]***	[0.1093]***	[0.3355]**	[0.0836]***	[0.0752]***
First-stage F-Stat	14.31	15.25	12.20	21.10	30.54
Observations	8,110	9,092	4,941	10,988	8,481
Regions	713	787	429	962	728
Countries	48	34	22	60	92

Notes: The table reports regression results. Panel a shows two-stage least squares fixed effects regressions where the dependent variable is the first difference of the Gini coefficient of light intensity within first-order administrative regions. Panel b shows least squares fixed effects regressions where the dependent variable is a binary indicator for new project commitments ($\Delta N_{ji,t-2}$) in a region. Columns 1 to 3 report regional subsets as indicated in the column header. Column 4 uses only countries classified as low income economies by the World Bank in 2000. Column 5 uses only regions which have received any transport or non-transport financing from China over the entire period. All specifications include region fixed effects and country-year fixed effects. Standard errors clustered at the country level are reported in parentheses. Conley errors with a spatial cutoff of 500 km and a time-series HAC with a lag cutoff of 1,000 years are reported in brackets.

Online Appendix

Table A-1 – List of countries

Afghanistan	Ghana	North Macedonia
Albania	Guatemala	Oman
Algeria	Guinea	Pakistan
Angola	Guinea-Bissau	Panama
Argentina	Haiti	Papua New Guinea
Armenia	Honduras	Paraguay
Azerbaijan	Hungary	Peru
Bangladesh	India	Philippines
Belarus	Indonesia	Poland
Benin	Iran, Islamic Rep.	Romania
Bolivia	Iraq	Russian Federation
Bosnia and Herzegovina	Jamaica	Rwanda
Botswana	Jordan	Saudi Arabia
Brazil	Kazakhstan	Senegal
Bulgaria	Kenya	Serbia
Burkina Faso	Korea, Dem. People's Rep.	Sierra Leone
Burundi	Kyrgyz Republic	Slovak Republic
Cambodia	Lao PDR	Somalia
Cameroon	Latvia	South Africa
Central African Republic	Lebanon	South Sudan
Chad	Lesotho	Sri Lanka
Chile	Liberia	Sudan
Colombia	Libya	Syrian Arab Republic
Congo, Dem. Rep.	Lithuania	Tajikistan
Congo, Rep.	Madagascar	Tanzania
Costa Rica	Malawi	Thailand
Cote d'Ivoire	Malaysia	Togo
Croatia	Mali	Trinidad and Tobago
Cuba	Mauritania	Tunisia
Czech Republic	Mauritius	Turkey
Dominican Republic	Mexico	Turkmenistan
Ecuador	Moldova	Uganda
Egypt, Arab Rep.	Mongolia	Ukraine
El Salvador	Morocco	Uruguay
Eritrea	Mozambique	Uzbekistan
Estonia	Myanmar	Venezuela, RB
Eswatini	Namibia	Vietnam
Ethiopia	Nepal	Yemen, Rep.
Gabon	Nicaragua	Zambia
Gambia, The	Niger	Zimbabwe
Georgia	Nigeria	

Notes: The table lists all 122 countries included in the regression analysis.

Table A-2 – Descriptive statistics (first-order regions)

	N	*Mean*	*SD*	*Min*	*Max*
Panel a) Dependent variables					
GINI_{jit} (within first-order regions)	26,759	0.535	0.159	0.000	0.849
GINI_{jit} (between first-order regions)	22,664	0.453	0.194	0.000	0.985
Log light per capita	26,759	-3.301	0.942	-4.604	1.793
Extensive margin	26,759	0.466	0.344	0.001	1.000
Quintile share (0-20%)	25,092	0.026	0.031	0.000	0.330
Quintile share (20-40%)	25,092	0.058	0.037	0.000	0.232
Quintile share (40-60%)	25,092	0.109	0.046	0.000	0.318
Quintile share (60-80%)	25,092	0.198	0.053	0.000	0.536
Quintile share (80-100%)	25,092	0.609	0.134	0.168	1.000
Panel b) Variables of interest					
Projects ($N_{i,t-2}$)	26,759	0.047	0.272	0.000	6.000
Projects (log 1 + financial values)	26,759	0.177	1.769	0.000	21.61
Projects (location count)	26,759	0.027	0.317	0.000	10.000
Projects (completed)	26,759	0.029	0.211	0.000	5.000
Projects (economic infrastructure)	26,759	0.099	0.417	0.000	8.000
Projects (construction phase)	26,759	0.013	0.112	0.000	1.000
Panel c) Instruments					
IV ($F_{t-3} \times \bar{p}_{ir}$)	26,759	-0.003	0.044	-0.821	0.630
IV (Det. steel)	26,759	0.000	0.048	-0.889	0.773
IV (US placebo)	26,759	1.224	4.100	0.000	55.48
Panel d) Other variables					
Log FDI	26,759	10.04	1.300	7.831	11.59
Log imports	26,098	6.924	2.093	0.000	11.12
Log exports	26,098	5.492	2.835	0.000	10.74
Distance to coast	26,759	0.313	0.385	0.001	2.455
Urbanization	26,759	0.072	0.151	0.000	1.000
Large mines	26,759	0.262	0.440	0.000	1.000
Oil fields	26,759	0.306	0.461	0.000	1.000
City access	26,747	326.5	384.3	4.418	4,984
Road density	26,759	0.143	0.297	0.000	8.805
Light per capita	26,759	0.044	0.062	0.000	1.207
Africa	26,759	0.303	0.460	0.000	1.000
Asia	26,759	0.340	0.474	0.000	1.000
Americas	26,759	0.185	0.388	0.000	1.000
Developing	26,108	0.825	0.380	0.000	1.000
$\bar{\rho}^{\text{all}} > 0$	26,759	0.320	0.466	0.000	1.000

Notes: The table provides descriptive statistics for the regressions using data at the level of first-order regions.

Table A-3 – Variants of baseline regression, within ADM2, 2002–2013

	Variations of the variable of interest				
	Financial values	Location count	Completed	Economic infra.	Transport count
	(1)	(2)	(3)	(4)	(5)
	Panel a) 2SLS estimates				
Δ Projects $(t-2)$	-0.0023	-0.0246	-0.0681	-0.0268	-0.0315
	(0.0011)**	(0.0125)*	(0.0367)*	(0.0169)	(0.0165)*
	Panel b) First-stage estimates				
IV $(F_{t-3} \times \bar{\rho}_{ir})$	7.2081	0.6643	0.2474	0.4160	0.4927
	(1.6976)***	(0.1690)***	(0.0534)***	(0.0717)***	(0.1317)***
Construction phase	–	–	–	–	✓
First-stage F-stat	16.10	13.86	18.93	30.41	12.53
Observations	343,514	343,514	343,514	343,514	343,514
Regions	31,428	31,428	31,428	31,428	31,428
Countries	112	112	112	112	112

Notes: The table reports regression results. Panel a shows two-stage least squares fixed effects regressions where the dependent variable is the first difference of the Gini coefficient of light intensity within second-order administrative regions. Panel b shows least squares fixed effects regressions where the dependent variable is indicated in the column header. 'Financial values' refers to the (logged) financial value of project commitments in constant 2014 US dollar. 'Location count' is the number of project locations of newly committed projects in a region. 'Completed' is a binary indicator for completed projects in a region. 'Economic infrastructure' broadens the definition of our base measure by including all economic infrastructure projects (transportation, energy production and distribution, and information and communication technology). The last column shows the baseline measure with an added control variable for the construction phase. All specifications include region-fixed effects and country-year-fixed effects. Standard errors clustered at the country level are reported in parentheses.

Table A-4 – Controlling for co-location, within ADM1, 2002–2013

	Controlling for other projects committed at $t-2$				
	World Bank (1)	World Bank (2)	China (3)	China (4)	China (5)
	Panel a) 2SLS estimates				
Projects ($\Delta N_{i,t-2}$)	-0.0348	-0.0342	-0.0349	-0.0341	-0.0352
	(0.0159)**	(0.0158)**	(0.0164)**	(0.0159)**	(0.0155)**
	[0.0178]*	[0.0178]*	[0.0182]*	[0.0178]*	[0.0179]**
	Panel b) First-stage estimates				
$F_{t-3} \times \bar{\rho}_{ir}$	0.4371	0.4366	0.4319	0.4369	0.4350
	(0.0772)***	(0.0771)***	(0.0785)***	(0.0772)***	(0.0779)***
	[0.0720]***	[0.0719]***	[0.0726]***	[0.0719]***	[0.0723]***
World Bank Transport	✓	–	–	–	–
World Bank Any	–	✓	–	–	–
China Social	–	–	✓	–	–
China Production	–	–	–	✓	–
China Energy	–	–	–	–	✓
First-stage F-Stat	32.02	32.04	30.26	31.99	31.21
Observations	26,759	26,759	26,759	26,759	26,759
Regions	2,385	2,385	2,385	2,385	2,385
Countries	122	122	122	122	122

Notes: The table reports regression results. Panel a shows two-stage least squares fixed effects regressions where the dependent variable is the first difference of the Gini coefficient of light intensity within first-order administrative regions. Panel b shows least squares fixed effects regressions where the dependent variable is a binary indicator for new project commitments ($\Delta N_{ji,t-2}$) in a region. Columns 1 to 2 control for World Bank projects using the World Bank Geocoded Aid Data v1.4.2. Column 3 to 5 include dummies for projects in different sectors based on the geocoded China data presented in this paper. All specifications include region fixed effects and country-year fixed effects. Standard errors clustered at the country level are reported in parentheses. Conley errors with a spatial cutoff of 500 km and a time-series HAC with a lag cutoff of 1,000 years are reported in brackets.