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‖

February 2024

Abstract

This paper studies the causal effect of transport infrastructure on the spatial distribution of economic activity within subnational regions across a large number of developing countries. To do so, we introduce a new global dataset of geolocated Chinese grant- and loan-financed development projects from 2000 to 2014 and combine it with measures of spatial concentration based on remotely sensed data. We find that Chinese-financed transportation projects decentralize economic activity within regions, as measured by a spatial Gini coefficient, by 2.2 percentage points. The treatment effects are particularly strong in regions that are less developed, more urbanized, and located closer to cities.

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

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

*Institute of Economics and Law, University of Stuttgart; email: *richard.bluhm@ivr.uni-stuttgart.de.*
†Alfred-Weber-Institute for Economics, Heidelberg University; CefES; KOF Swiss Economic Institute; CEPR; CESifo; email: *mail@axel-dreher.de.*
‡Department of Economics and Centre for Modern East Asian Studies, University of Göttingen; Kiel Institute for the World Economy; email: *mail@andreas-fuchs.net.*
§AidData, Global Research Institute, William & Mary; Center for Global Development; email: *bparks@aiddata.wm.edu.*
¶Department of Politics and Public Administration, The University of Hong Kong; email: *astrange@hku.hk.*
‖Department of Government, William & Mary; email: *mjtier@wm.edu.*

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, Joseph Dobbels, Isabel Docampo, 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 projects DR 640/5-1&3 and FU 997/1-1&3), and William & 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 William & 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 any of our funders.

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), and 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.

JEL classifications: F35, R11, R12, P33, O18, O19

1 Introduction

In 2009, the Export-Import Bank of China (China Eximbank) approved a loan to the Kenyan government to substantially widen and improve the Nairobi-Thika Highway—a 50.4 km dual carriageway that extends from the center of Nairobi to the town of Thika. The project, locally known as the "Thika Super-Highway," sought to reduce congestion and travel times between Nairobi and a set of satellite towns along a critically important transportation corridor (African Development Fund 2007). Upon completion in 2012, traffic flows increased by 45 percent, 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,b, 2016, 2019). Economic activity spread out along the transport corridor and became substantially less concentrated in the core of Nairobi (see Figure 1). The case of the Nairobi-Thika Highway fits within a broader pattern: Starting with Baum-Snow (2007), a series of studies in a variety of countries show that major transport infrastructure investments can decentralize economic activity.[1]

[Figure 1 about here.]

The key contribution of this study is to examine whether and to what extent Chinese-financed infrastructure projects are decentralizing economic activity within developing countries. We conduct analysis at the level of first-order administrative regions. First-order regions are one layer below the national level and correspond, for example, to provinces, states, oblasts, governorates, or emirates, depending on the administrative divisions in a given country.[2] Our analysis focuses specifically on the provision of transport infrastructure financing from China's government, which 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 concentration[3] of economic activity within regions, but we also present results relating to infrastructure financing and concentration across regions. Our study tests whether the results from the existing, usually country-specific, literature can be generalized across a large sample of developing countries that host projects provided by the world's largest provider of infrastructure financing.

[1] See, for example, Baum-Snow et al. (2017) and Banerjee et al. (2020) on China, Bayes (2007) on Bangladesh, Bird and Straub (2014) on Brazil, Donaldson (2018) on India, Henderson and Kuncoro (1996) on Indonesia, Garcia-Lopez et al. (2015) on Spain, Gibbons et al. (2019) on the United Kingdom, and Duranton and Turner (2012) on the United States. Redding and Turner (2015) as well as Baum-Snow and Turner (2017) provide surveys of this literature.

[2] In our sample, the average region's size is 37,644 square kilometers, which roughly corresponds to the land area of South Carolina. It has 2.1 million inhabitants, which roughly corresponds to New Mexico.

[3] We use the terms *spatial concentration* and *spatial centralization* of economic activity interchangeably, as they both refer to changes in the distribution of people and output across space.

Since 2000, China's government has financed many of the largest transport infrastructure projects in the Global South. 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 studies 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). 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.[4]

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 swiftly finance and build roads, bridges, railways, and ports at a time when Western donors and lenders are not (Swedlund 2017).[5] 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 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).

We introduce the first global dataset of geo-located Chinese government-financed projects that were undertaken in developing countries between 2000 and 2014.[6] The dataset includes 3,485 projects in 6,184 subnational locations across 138 countries during these 15 years. For our analysis, we focus on 269 Chinese government-financed transportation infrastructure projects undertaken in 1,215 subnational locations across 86 countries. The lower bound for the total financial value of these projects is US$64 billion. We estimate the effects of these projects on the spatial concentration of economic

[4] 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).

[5] 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).

[6] Though the Belt and Road Initiative (BRI) was not officially launched until late 2013, the Chinese government had already begun providing a significant number of 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.

activity—both within and across subnational jurisdictions—with satellite data on the geographical dispersion of nighttime light output (similar to Henderson et al. 2018).

To identify the causal effect of Chinese government-financed projects on the spatial concentration of economic activity, we construct an instrumental variable that relies on the availability of resources for the construction of infrastructure projects. 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.[7] Specifically, 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.[8] 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.[9] 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 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.

Our results show that regions which are frequent recipients of projects receive larger amounts of Chinese government financing in years of overproduction than subnational localities that infrequently receive Chinese government-financed transport projects. This difference presumably occurs because existing local capacity and relationships make it easier to implement additional projects. This estimate can be interpreted as a difference-in-differences estimate, similar to those reported in the "China shock" or aid and conflict literatures (e.g., Autor et al. 2013, Nunn and Qian 2014). 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.

[7]The literature typically uses historical transport networks or other country-specific historical circumstances (such as minimum spanning trees connecting the largest cities).

[8]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. (2021) which exclusively used the level of steel production.

[9]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. 2021).

We find that Chinese-financed infrastructure projects reduce spatial concentration within first-order regions and accelerate the diffusion of economic activity around cities (in line with Figure 1 and Baum-Snow 2007). Specifically, we find that the Gini coefficient measuring the spatial concentration of economic activity is reduced by about 2.2 percentage points. Similar specifications for concentration between regions suggest effects of similar magnitude. However, these effects are less precisely estimated, suggesting that Chinese-financed infrastructure cannot be robustly linked to changes in the concentration of economic activity across regions. The absence of a "between effect" is in line with several recent studies which emphasize that transport infrastructure has heterogeneous and context-specific impacts on the distribution of economic activity across regions (Fajgelbaum and Redding 2018, Baum-Snow et al. 2020, Jedwab and Storeygard 2022, He et al. 2020). Our main results hold under various perturbations, such as the choice of control variables to strengthen identification or variations of the instrument.

Consistent with predictions from land use theory, we find that transport projects shift activity from densely developed locations to less densely developed ones, that is, from the highest quintile of the light distribution to lower quintiles. We find no evidence that infrastructure projects increase the fraction of illuminated pixels in a recipient region or that a region's light per capita increases. However, our results show a notable increase in overall light intensity, defined as the sum of light emissions within a region relative to its geographic area. This suggests that regions experience positive population growth in response to Chinese-financed infrastructure projects. The results also show that the impact of these projects on the concentration of activity within regions is heterogeneous. Chinese-financed transport infrastructure reduces concentration more strongly in regions with more urban areas, low travel time to cities, and higher road density. We take this as indirect evidence suggesting that our results are driven by a relocation of workers to the outskirts of cities rather than an increase in economic activity in peripheral cities of a region. We also provide evidence that these effects are largest in African countries and poorer regions within developing countries, which tend to experience rapid population growth and have a high 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 a 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 to reduce their commuting times. Subsequent investments in transportation infrastructure increase transportation speed, 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 in. 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.[10]

Reality is not as stylized and most cities are not monocentric but rather characterized by a CBD that coexists with other employment subcenters. Polycentric cities arise when the location of employment and commercial activity are determined endogenously within the city (for early contributions, see Ogawa and Fujita 1980, Fujita and Ogawa 1982, Henderson and Mitra 1996). In polycentric cities, agglomeration economies are smaller relative to the monocentric benchmark, but commuting costs—hence, dispersion forces—are also lower. New quantitative spatial models allow for complex patterns of agglomeration and dispersion forces, which can lead to various urban forms (see Redding and Rossi-Hansberg 2017, for a review). In these models, locations are not independent but connected via trade, commuting, and migration flows. Firms and residents make optimal choices based on these connections. The predictions of the frameworks depend on these spatial interactions and, hence, the data. Infrastructure improvements can reduce the concentration of activity by making peripheral areas more accessible, but they can also increase concentration by expanding the labor market for central firms. For example, early infrastructure improvements, such as the steam railway, permitted a specialization of city centers into workplaces and their surroundings into residences (Heblich et al. 2020).

A large and growing body of empirical evidence supports the basic prediction of the monocentric city model that new or upgraded transportation infrastructure disperses

[10]Firms have different incentives than residents since they face a more complex set of costs when leaving city centers. They trade agglomeration benefits against various costs (transportation costs of output, interaction costs, labor accessibility, etc.). This gives rise to a pattern where firms that benefit less from face-to-face interaction and knowledge spillovers, such as manufacturing firms, decentralize more than other firms (Rossi-Hansberg et al. 2009, Baum-Snow 2014).

economic activity away from urban agglomerations. Consider suburbanization in 20th-century America, where researchers have documented urban sprawl and strong population growth in cities with more developable surroundings (Burchfield et al. 2006, Saiz 2010). The construction of highways in the United States dramatically lowered commuting times and increased demand for suburban relative to urban residential space (Baum-Snow 2007). Congestion also plays an important role, particularly in contemporary settings. Allen and Arkolakis (2022) highlight that stronger congestion not only reduces the market access of central locations and shifts activity away from the core but that this effect increases with city size.

There is also evidence for similar processes of diffusion around urban areas in developing countries (e.g., Bayes 2007, Zárate 2022). 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. As Chinese-financed infrastructure projects in developing countries often represent a substantial proportion of local infrastructure investment in a given year, we expect them to decentralize economic activity around urban areas similarly. We also expect these projects to spur the formation of new sub-centers as businesses, workers, and other economic actors relocate into the periphery.

Transport projects may also affect the concentration of economic activity across regions, a central concern of economic geography research. The classic core-periphery model stresses the role of 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 a bell shape for regional inequalities in relation to trade costs. The advantage of being in a central location 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 concentration reflects this heterogeneity. Brülhart et al. (2019) show that the advantages of market potential are shrinking in the developed world but remain 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.[11] 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 these mixed findings and the cross-national scope of our study, we do not have strong reasons to believe that Chinese-financed transportation projects will uniformly increase or decrease concentration between regions.

Developing countries are an important and useful application of these theories. 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). 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). Major infrastructure financing gaps make it difficult for developing countries to overcome the spatial bottlenecks created by high levels of urban concentration and rural neglect.

Moreover, subnational regions within developing countries are often defined by dense central cities surrounded by underdeveloped hinterlands.[12] 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). In these settings, high levels of urban congestion in developing economies are a 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 aid expenditures or activities. Nor does it systematically publish project-

[11] In a related 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.

[12] 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.

level data on its less concessional and more commercially-oriented financial expenditures and activities in developing countries. To overcome this challenge, Dreher et al. (2021) collaborated with AidData, a research lab at William & Mary, to build a global dataset of Chinese government-financed projects committed between 2000 and 2014. This project-level dataset uses 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.[13] Economists, political scientists, and computational geographers have for the most part used these data to explain the nature, allocation, and effects of Chinese government-financed projects in Africa and in several country-specific studies outside of Africa (e.g., Hernandez 2017, Dreher et al. 2018, Isaksson and Kotsadam 2018a,b, Anaxagorou et al. 2020, Isaksson 2020, Martorano et al. 2020, Dreher et al. 2021, Eichenauer et al. 2021, Horn et al. 2021, Gehring et al. 2022, Baehr et al. 2023).

In this paper, we build on these project-level data to create a first-of-its-kind dataset of Chinese grant- and loan-financed development project locations around the globe. In contrast to previous versions, our new data enable subnational analyses of Chinese-financed projects in five regions of the world (Africa, the Middle East, Asia and the Pacific, Latin America and the Caribbean, and Central and Eastern Europe) over 15 years (2000–2014). Our dataset takes all projects from Dreher et al. (2021) that secured financial commitments from China and entered implementation or reached completion as a starting point.[14] We then subjected all of these projects to a double-blind geocoding process (Strandow et al. 2011), in which two trained coders independently employ a defined hierarchy of geographic terms and 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

[13]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).

[14]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. 2019). Note that we exclude all suspended and canceled 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).

exact project site; precision code 3 corresponds to a second-order region; and precision code 4 corresponds to a first-order region.[15] 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. This double-blind coding process aims to minimize the risk of missed or incorrect locations.[16] In total, the resulting dataset covers 3,485 projects (worth at least US$273.6 billion in constant 2014 dollars) in 6,184 discrete locations across 138 countries.[17]

To merge these geocoded project data with our outcome measures of spatial concentration within and across subnational regions, we aggregate all projects with precision codes 1–4 to first-order regions. Figure 2 shows the locations of projects that can be placed within first-order regions over the 2000–2014 period. The resulting subsample includes 2,140 Chinese government-financed projects at 4,420 discrete locations (collectively worth US$201 billion) that were completed or being implemented in 883 first-order regions within 129 countries between 2000 and 2014.[18] Our data can be disaggregated by financial flow type and sector. With respect to the former, we distinguish between Official Development Assistance (ODA) and other forms of concessional and non-concessional financing from Chinese government institutions.[19] For the purposes of the latter, we use the OECD's three-digit sector classification scheme, which categorizes projects according to their primary objectives.[20]

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

[16]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.

[17]For comparison, the Africa-specific data provided in Dreher et al. (2019) include 1,650 projects across 2,969 locations in the 2000–2012 period. Note that, in contrast to our dataset, they also cover projects that have not (yet) reached implementation stage.

[18]We only focus on low-income and middle-income countries. More precisely, we include countries that 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, 2023). We also exclude small states with a population size below a threshold of 1,000,000 inhabitants. Table A-1 in the Online Appendix lists all countries included in the analysis.

[19]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 are financed by Chinese government institutions with development intent and a minimum level of concessionality (a 25 percent or higher grant element). Chinese OOF projects are financed by Chinese government institutions with 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 clearly determine 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).

[20]There are 24 of these OECD sector codes (see www.oecd.org/development/financing-sustainable-development/development-finance-standards/purposecodessectorclassification.htm for details).

269 projects were assigned to the transport and storage sector and implemented in 1,215 locations. The combined value of these projects was at least US$64.1 billion (when counting those projects for which financial values are available) and amounts to an average of US$224.9 million per project.[21] This is sizable in light of an average GDP of about US$4.4 billion per first-order region. The vast majority of these projects focused on building transportation infrastructure, such as roads, railways, bridges, seaports, and airports. With 651 project locations, long-distance roads are most frequent, followed by long-distance railways (245), and urban roads (123) (see Table A-2 in the Online Appendix).[22]

In our dataset, the average financial commitment from China for a long-distance road project was US$231.2 million compared to US$979.3 million for a long-distance railway project and US$148.3 million for an urban roads project. To better understand the nature of these projects, consider three projects from three different regions that are broadly indicative of the types of "treatments" that we analyze in our statistical analysis. First, in October 2014, China Eximbank provided a US$943.9 million loan to Montenegro's Ministry of Finance to help finance the construction of a 169 km highway between Bar (the country's main seaport in the south) and Boljare on the Montenegrin-Serbian border in the country's north.[23] The loan was worth approximately a quarter of the country's GDP at the time that it was contracted. Upon completion, the highway was expected to reduce travel time between the capital of Podgorica and the northern city of Kolašin from 90 minutes to 30 minutes and facilitate economic development alongside the transport corridor. Second, in May 2013, China Eximbank issued a US$491.7 million loan to Djibouti's Ministry of Finance for the construction of a 100 km segment of a 756 km railway that runs from Addis Ababa, the capital of Ethiopia, to the Doraleh seaport in Djibouti. At the time of its issuance, the loan represented approximately 40% of Djibouti's GDP. Upon completion, the railway was expected to reduce travel time between Addis Ababa and Doraleh seaport from 7 days (on roads) to 10 hours.[24] Third, in 2009, China Eximbank provided a US$44.2 million loan to the Government of Tonga for phase 3 of the National Road Improvement Project. The loan was worth roughly 16% of GDP at the time that it was contracted. The purpose of the project was to improve the 1 km Alipate Road on the island of Tongatapu, the 8 km road that runs from Kolonga to Talasiu on the island of Tongatapu, and the 4 km road that runs from Haveluliku to

[21] Many of these projects are implemented in multiple locations, for example, when they connect two cities with a railway line. The average amount committed per project location is 49.8 million.

[22] Transport 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).

[23] See `https://china.aiddata.org/projects/42330/` for details on the project.

[24] See `https://china.aiddata.org/projects/46183/` for details.

Lavengatonga on the island of Tongatapu.[25]

[Figure 2 about here.]

Figure 2 illustrates the global reach of China's overseas development program in the 21st century. Consistent with earlier periods of Chinese aid giving (Dreher and Fuchs 2015), Chinese projects cover almost all developing countries (with countries recognizing the Chinese government in Taiwan as a notable exception).[26] Chinese-financed development projects are densely concentrated in African and Asian countries. The figure also illustrates that many Chinese government-financed projects are situated in coastal regions, including some of the highest-value transportation projects.

Measuring concentration 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 previous work that uses nighttime light intensity as a proxy for local economic activity (Henderson et al. 2012, Hodler and Raschky 2014, Michalopoulos and Papaioannou 2014). While nighttime lights were initially proposed as a measure of income for countries with weak statistical capacity, they were quickly adopted more broadly as a measure of subnational economic activity in developing countries. Subsequent studies have demonstrated that changes in light emissions correlate strongly with traditional welfare measures down to the village level (Weidmann and Schutte 2017, Bruederle and Hodler 2018).

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. While we are primarily interested in whether and to what extent infrastructure investments relocate economic activity, we also investigate below whether such investments increase output per capita.

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

[25] For details, see `https://china.aiddata.org/projects/39199/`.

[26] For recent work that studies the allocation of China's development finance across countries, see Dreher et al. (2022) or Hoeffler and Sterck (2022).

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.[27] Second, we intersect this grid with the global first-order administrative boundaries, which creates "squiggly" cells along the regional borders.[28] 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), and the light intensity in the cell ($x_i = s_i/a_i$).[29] 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 cells (including cells with zero light intensity) within an administrative region as

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

where $w_i = \frac{a_i}{\sum_{i=1}^{n} a_i}$ is an area-based weight and n is the total number of cells in a region. We also construct Gini coefficients for concentration *between* first-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. Geometrically, it is the area under the Lorenz curve plotting the cumulative distribution of weighted light intensities against the cumulative distribution of cell areas (in km^2). Including cells with zero light intensity means the Lorenz curve will remain at zero before sloping up to one, but ensures that the Gini coefficient is a proper measure of economic concentration, which not only decreases when the distribution of light becomes more equal among already illuminated cells but also when new cells become illuminated. As can be

[27] Although the nominal resolution of the DMSP-OLS system is 30 arc seconds, geolocation errors and on-board processing of fine-resolution pixels lead to a true ground footprint of 5 km by 5 km (Elvidge et al. 2013). Taking about twice this resolution reduces the influence of this mechanical spatial autocorrelation, reduces the influence of top coding and bottom coding, and limits the computational burden. The newer Visible Infrared Imaging Radiometer Suite (VIIRS) data, which have superior technical properties, do not span a significant portion of our sample.

[28] We obtained the regional borders from the Database of Global Administrative Areas (GADM) vector dataset (version 2.8). The same data were used to geocode the Chinese-financed projects.

[29] Dividing by the land area adjusts for the fact that 6 arc minute cells do not have a uniform area across the globe and may be covered by water. We calculate the land area of each cell using the Gridded Population of the World (v4) land/water raster.

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, both within and across countries (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.[30] Henderson et al. (2018) show that the cross-sectional variation in population density across administrative regions is substantially larger than the variation in income per capita. If this holds across time, then a significant proportion of observed changes in the within-region distribution of light intensities should 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.

We prefer using nighttime lights over data for population density as our main outcome measure. 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 that host many Chinese-financed development projects.

4 Empirical strategy

We are interested in changes in the spatial concentration of economic activity caused by Chinese infrastructure investments. Denoting first-order administrative regions by j, countries by i,[31] and years by t, our main equation relates our luminosity-based measure of spatial concentration, $\textsc{Gini}_{jit}$, to the total number of years in which transportation projects have been committed to a region up to $t-2$, denoted $N_{ji,t-2}$.

We chose $N_{ji,t-2}$ as the baseline treatment because transportation projects can vary widely in type and size, ranging from small bridges to (segments of) extensive

[30] 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.

[31] Going below the first-order level would change many sample characteristics. We would lose a significant share of projects that have only been accurately coded to first-order regions. The exposure variable we introduce below would also be based on substantially fewer projects per region. What is more, in countries with a smaller land mass, second-order regions often correspond to administrative boundaries of cities, which would effectively exclude their surroundings from a within-unit analysis.

multiregional highways.[32] To assess the impact of *any* infrastructure project or bundle of projects in a region in a given year, this variable aggregates over all years in which at least one project was committed (see the discussion of the differenced version below). In robustness checks, we use the number of project locations or financial values (where available). By aggregating over all years, our specification assumes that the *level* of concentration depends on the entire history of projects. Of course, the effect of projects could fade over time so that projects committed in the distant past would no longer affect concentration. Since our sample only covers the period since 2000, and Chinese development finance was comparatively low before that time, we focus on the medium-term effects of new projects rather than those committed in the distant past. We lag this variable by two years to account for the difference between the commitment date and the expected completion date of a project.[33]

We begin with a flexible specification that allows the effects of Chinese-financed projects to be arbitrarily correlated with region-specific fixed effects and region-specific time trends:

$$\text{G}\textsc{ini}_{jit} = \beta N_{ji,t-2} + \mu_{ji} + \theta_{ji} \times t + \lambda_{it} + \epsilon_{jit}, \tag{2}$$

where μ_{ji} are region-fixed effects, $\theta_{ji} \times t$ are region-specific linear time trends, and λ_{it} are country-year-fixed effects that absorb a variety of potential shocks to all regions of a country in a particular year.

A more tractable and intuitive version of this model can be estimated in first differences:

$$\Delta\text{G}\textsc{ini}_{jit} = \beta \Delta N_{ji,t-2} + \theta_{ji} + \tau_{it} + \Delta\epsilon_{jit}, \tag{3}$$

where $\tau_{it} = \lambda_{it} - \lambda_{i,t-1}$ is a new set of country-year-fixed effects and the region-fixed effects in first differences now capture region-specific trends in levels, $\theta_{ji} = \theta_{ji} \times t - \theta_{ji} \times (t-1)$.

The specification in differences shows that β estimates the persistent effect of a new transport project (or a bundle of transport projects committed in the same year) on spatial concentration two years later.[34] The model nests less flexible approaches with a

[32]Another technical reason is that projects are often co-located, e.g., different sections of a highway, but may have distinct project IDs in our data because they are financed through different financial tranches. This does not necessarily capture the intensive margin of infrastructure investments but reflects the definitions adopted during the geocoding process.

[33]Note that start and end dates are available for a subset of the projects in our dataset. Across these approximately 1,100 Chinese government projects, the average time from start to completion 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. 2021, based on data from Bartke 1989). The two-year lag we use for the main analysis thus allows projects to register effects directly after (expected) completion. As we show below, the effects of Chinese transport projects materialize quickly. When longer lags are used, estimates become smaller and less precisely estimated.

[34]The event-study equivalent would be a permanent change of size β starting in $t = 2$ for a project

strict parallel-trends assumption since all θ_{ji} could be zero.

Our preferred measure of transportation infrastructure, $\Delta N_{ji,t-2}$, is thus a binary variable indicating that at least one new project is committed to a particular region in a year. Since the size of the projects is not homogeneous across locations, the effects of projects on spatial concentration 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. 2021), which is why we prefer the binary indicator and present additional results using (logged) aggregate dollar values for comparison. Moreover, 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 data after 2013 (from the DMSP-OLS system), our sample effectively covers the period from 2002 until 2013.[35]

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.[36] While our "treatment-control" strategy does not account for general equilibrium effects, spillovers to untreated regions would have to systematically raise or lower concentration there (which has not been established in the literature) to introduce significant bias into our within-region estimates.

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. An official goal of Beijing's overseas development program is to make "great efforts to ensure its aid benefits as many needy people as possible" (State Council 2011). Contrary to this claim, previous studies demonstrate that the allocation of Chinese development finance 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

committed in $t = 0$.

[35]The original DMSP-OLS data are only available through 2013. Nighttime observations by NPP-VIIRS have replaced this system, but they record light intensities with different sensors and on a different scale. Though researchers have aimed to harmonize the two series, we only lose one year in our analysis and therefore prefer to use data from a single, consistent source.

[36]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 level of regions or at the level of regions *and* years. None of this affects our qualitative results.

to exploit agglomeration effects or located in suburban areas out of cost concerns.

Apart from reverse causality, our parsimonious specification omits a number of variables that are likely to be correlated with Chinese infrastructure funding and spatial concentration. Some of these covariates vary across regions and over time in a non-linear fashion, so our battery of fixed effects and region-specific trends does not capture them. 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. 2022), 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." The same holds for population or population density, which we do not include as a control (although doing so hardly affects our results).

[Figure 4 about here.]

To address these endogeneity concerns, 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 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.[37] We calculate this (endogenous) probability as the fraction of years over the 2000–2014 period in which at least one Chinese government-financed transport infrastructure project has been committed, $\sum_{p=1}^{T} \Delta N_{ijp}/T$, and denote this variable by $\bar{N}_{ji}$. 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.[38] Given that the production of these raw

[37] Our description of the instrument draws on Dreher et al. (2021), 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 Humphrey and Michaelowa (2019), Brazys and Vadlamannati (2021), Iacoella et al. (2021), Marchesi et al. (2021), Zeitz (2021), and Ping et al. (2022). Since introducing our instrument variant with multiple detrended inputs, it has been used in Che et al. (2021), Dreher et al. (2021, 2022), Gehring et al. (2022), and Wellner et al. (2024), among others.

[38] 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

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 of receiving transport projects to form a single instrumental variable (and later report robustness checks probing this choice).

We lag this series by one year (relative to the timing of project commitments), arguing 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 concentration within any particular region), while the probability of receiving projects varies only across regions. This is how our instrument resembles those commonly used in trade and development economics, such as the literature on the impact of the rise of Chinese manufacturing on local US labor markets (e.g., Autor et al. 2016) or studies on aid and civil conflict (e.g., Nunn and Qian 2014).

Putting these elements together, we estimate the following first-stage regression:

$$\Delta N_{ji,t-2} = \delta(F_{t-3} \times \bar{N}_{ji}) + \omega_{ji} + \phi_{i,t-2} + \nu_{ji,t-2}, \tag{4}$$

where F_{t-3} is the first common factor of the (detrended and logged) raw materials produced in China, and $\phi_{i,t-2}$ are country-year-fixed effects. Equations (3) and (4) are estimated using Two-Stage Least Squares (2SLS).[39] When we include controls for various robustness checks, they always enter both equations in first differences (i.e., $\Delta\mathbf{x}_{jit}$).

The intuition behind this identification approach resembles that of a generalized difference-in-differences design. Intuitively, we effectively compare the effects of Chinese transport infrastructure projects on spatial concentration induced by changes in China's domestic production of potential project inputs across two groups: regions that are regular and irregular recipients of Chinese transport infrastructure financing. Put another way, we use differences in local exposure to the common overproduction shock originating in China to identify potential effects of transport infrastructure projects on the spatial distribution of economic activity. We expect China to provide more projects when the inputs going into these projects are more easily available. Moreover, we expect those regions to disproportionally benefit from additional projects if they receive similar

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).

[39] Note that we do not 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 that could be used in Equation (2).

financing from China more often. Like other bilateral donors, China engages with a limited number of recipient countries and regions at any point in time, and extending additional resources to countries and sub-national regions with established contacts is arguably easier than establishing new ones. We thus expect regions that have more often received Chinese projects over the entire sample period to benefit disproportionally from the availability of additional resources.[40] While it is an empirical question whether the marginal project goes to a region that has previously received support or to a new region, research on other bilateral donors backs this expectation.[41]

We investigate the validity of our approach in several steps. First, the identifying assumptions inherent in this approach could be violated if other unobserved variables drive the allocation of Chinese-financed infrastructure projects *and* these variables have heterogeneous effects on spatial concentration that coincide with our distinction between regular and irregular recipient regions. Hence, compared to a standard panel difference-in-differences setting, our instrument ensures that the timing 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 an active financier of transport projects in 2000 between those regions that would subsequently become future recipient regions and those that would not. It shows that spatial concentration among regions that later receive a project runs parallel to concentration in regions that do not subsequently receive a project from 2000 to 2014. There is some narrowing towards the end of the pre-period, which is why we allow linear departures from parallel trends in our specifications. However, there is no evidence in favor of different non-linear trends between treated and non-treated regions. After 2000, we observe a steeper fall in spatial concentration in regions with transport projects as Chinese government financing becomes increasingly active towards the middle of the 2000s. In any case, this type of graphical analysis cannot entirely 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, which is why we pursue an instrumental variables strategy.

In Figure A-1 in the Online Appendix, we visually examine the variation in the transport project indicator and spatial concentration for different terciles of the indicator during our period of interest. The results also give little reason to believe that the parallel pre-trends assumption is violated. There are notable global trends—a secular decline in

[40] Dreher et al. (2022) show that even when variables that researchers typically include in aid allocation regressions are controlled for, those regions that have historically received aid more often receive larger amounts in any year.

[41] Examples include food aid, total aid from the U.S., and bilateral aid from the group of DAC donors (Nunn and Qian 2014, Dreher et al. 2019, Dreher and Langlotz 2020). Results differ for multilateral donors like the World Bank and IMF (see Lang 2021, Dreher et al. 2021), arguably because their recipients at the same time are their members and extending the benefits of IO membership to more of its members is in the interest of their bureaucrats and management.

spatial concentration and a rise in the number of projects—but the probability-specific trends in project numbers and concentration appear broadly parallel across terciles of $\bar{N}_{ji}$. Importantly for our identification strategy, there is no obvious *non-linear* trend in a particular tercile that resembles the trend in Chinese production of input materials—shown in Figure 4—more than in another (see Christian and Barrett 2017).

[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 $Cov(F_{t-3} \times \bar{N}_{ji}, \Delta\epsilon_{jit}) = 0$ conditional on region-specific time trends, as well as region- and country-year-fixed effects. This leaves few sources of confounding variation with heterogeneous non-linear effects. 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 instead 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 would we 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 (FDI). 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). It could be that 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 or 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 FDI, interacted with a set of variables that makes it more or less likely that a region is affected by changes of these variables.

Finally, 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. Absent parallel trends, there are two ways to achieve identification in such settings under an alternative set of assumptions. On the one hand, local industry shares can be interpreted as instruments, provided that they are exogenous (Goldsmith-Pinkham et al. 2020). On the other hand, identification can also be purely based on exogenous variation in (single or multiple) time-series shocks, even when variation in local exposures is endogenous (Borusyak et al. 2020)—as in the case of our instrument.[42] For the panel case, Borusyak et al. (2020) establish that our estimator is consistent when the conditional 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. The fixed effects play an important role in this setting. The region-fixed effects purge all time-invariant unobserved heterogeneity from the residuals and remove the time-invariant component from the (already detrended) shocks. Similarly, country-year-fixed effects remove all unobserved heterogeneity that affects all regions of a country in a specific year and isolate the within-country-period shock variation.

5 Results

Baseline results

Table 1 reports our main results on the relationship between Chinese-financed transport infrastructure projects and economic concentration, focusing on 2,406 first-order regions in 122 countries. We consider two different types of concentration—within regions and between regions—which we estimate in two different ways. Columns 1 and 2 show results when concentration is measured within regions. Column 1 focuses on the binary project indicator, column 2 on (logged) annual dollar amounts.[43] Columns 3 and 4 turn to between-regional concentration and report results with concentration computed over average light intensities in first-order regions, resulting in an analysis at the level of countries, which is the level of analysis typically presented in studies of between-region concentration (e.g., Lessmann and Seidel 2017).

[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. The coefficient estimates are negative in two of four regressions but are imprecisely estimated and small in magnitude.

Panel b reports the reduced-form estimates for the same specifications. Here, we regress the change in the spatial Gini coefficient on our instrumental variable and the complete 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, we should observe a strong reduced-form effect as well. Indeed, columns 1 and 2 show that the

[42] Borusyak et al. (2020) highlight that this includes the design of Nunn and Qian (2014), which is very similar to our approach. The asymptotics can rely on the number of periods, the number of industries (different exposure shares), or both. T is only moderately large in our setting, which is why we conduct many robustness checks modifying the time-series dimension of the instrument.

[43] Project amounts have the advantage that we can account for the size of projects, but come with the drawback that we lack information on the financial amounts for more than a third of these projects. Note that we have added a value of one before taking logs.

instrument has a significant and negative effect on changes in spatial concentration within regions. This effect will be passed through with the same sign in our 2SLS regressions below 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 (owed mainly to higher standard errors).

Panel c in Table 1 presents our main results where the two-period lag of the project variables ($\Delta N_{ji,t-2}$) is instrumented by the detrended project input series (F_{t-3}) times our local exposure variable ($\bar{N}_{ij}$). 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 concentration between regions. Our results are in line with these expectations. For concentration within regions, the 2SLS coefficients are negative, statistically significant at conventional levels, and of substantial magnitude.

The point estimate in column 1 indicates that the Gini coefficient is permanently reduced by 2.2 percentage points in regions where at least one Chinese government-financed transport project has been committed two years before.[44] Column 2 shows results using monetary values. 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), implying a similar estimated decrease in the Gini coefficient.[45] While we have not yet tested if urban areas drive these results, the results are in line with the notion that building or upgrading transport infrastructure allows economic activity to decentralize around congested cities (e.g., Baum-Snow 2007, Baum-Snow et al. 2017). 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.

The difference between the 2SLS estimate and the OLS counterpart suggests that the latter is subject to measurement error (biasing the coefficient toward zero) and/or simultaneous causality bias (biasing it upwards). Both of these are plausible in our setting.[46] Our geolocated Chinese project data track underreported financial flows and

[44]Recall that the average project duration is about 2.1 years, suggesting that the effects we find here result from projects that have, on average, just been completed.

[45]The average Gini coefficient within regions is 0.54. Table A-3 in the Online Appendix reports summary statistics.

[46]Another source of discrepancy between the OLS and 2SLS results is that the former are two-way fixed effects estimates of a staggered difference-in-differences design. The OLS coefficient is a weighted sum of the underlying average treatment effects (ATEs). If treatment effects are heterogeneous, these weights may turn negative, implying that the coefficient is a non-convex combination of the ATEs in each group and year (de Chaisemartin and d'Haultfoeuille 2020). Our strategy differs from the designs studied in this literature in several ways. First, we have a non-absorbing and repeated treatment that is unity only in the period a project is committed (recall that we regress changes on changes in eq. (3)). Second, several regions might have been treated before they were included in our sample, so "never treated" or "not yet treated" regions within the sample do not comprise adequate control groups. Third, and most importantly, we argue that the assumption of parallel trends is unlikely to hold since the timing

use trained coders to assign these projects to subnational regions. This process is rigorous but imperfect. Some projects may be missing or misallocated in space or time. As discussed above, simultaneous causality will induce an upward bias in the OLS estimate if China invests more in regions that are becoming more spatially concentrated than the average region in the country. This is not testable since we only have an identification strategy for the opposite causal direction. However, it fits with narratives that China invests more strongly in regions and cities developing fast and becoming particularly congested (Dreher et al. 2019, 2022).

Turning to the results in columns 3 and 4, we observe point estimates consistent in magnitude and direction with those obtained from our within-region regression analysis. However, the standard errors are significantly larger, resulting in imprecise estimates. Given the similarity in coefficient magnitudes, we cannot rule out the possibility that Chinese-funded infrastructure projects are associated with a decline in interregional inequality in the developing countries in our sample (the first-stage F-statistics remain strong, but the reduced-form estimates also have large standard errors). For the remainder, we focus on the within-region concentration of activity, as the between-region effect of such investments is theoretically ambiguous and appears to be imprecisely estimated in our setting (not least due to the smaller number of observations at the country-year level).

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. They remain strong when computing F-statistics that are robust to heteroskedasticity, autocorrelation, and clustering (Olea and Pflueger 2013).[47] 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 8.8 percentage points ($0.4 \times \frac{7}{14} \times 0.44$) 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 1.3 percentage points in a region which received a Chinese project in only one year ($0.4 \times \frac{1}{14} \times 0.44$).

of Chinese infrastructure projects is not exogenous but can be predicted by domestic overproduction and variation in local exposures. When we nevertheless estimate the weights underlying the coefficient reported in column 1 of Table 1 using the method from de Chaisemartin and d'Haultfoeuille (2020), we find that none of the ATEs receive negative weights. This underscores the importance of simultaneity, omitted variables bias and measurement error as alternative explanations.

[47] The Montiel-Pflueger effective F-statistic is 30, which is significantly above the corresponding critical value for a 10% "worst-case" bias of 23.1 in column 1 (Olea and Pflueger 2013).

Our main results do not depend on the choice of country-clustered standard errors or Conley errors, although the latter tend to be a little smaller and the corresponding first-stage F-statistics (not reported) somewhat higher. Adão et al. (2019) point out that inference in shift-share designs may be biased downwards if we ignore that regions with the same probability of receiving a Chinese-funded transport project have similar regression residuals (no matter which cluster they are located in). While we cannot use their variance estimator in our application with one sector, the estimated standard errors barely change if we cluster them on distinct values of $\bar{N}_{ji}$ to account for some of this bias.

Discussion

According to our results, an additional Chinese transport project reduces the Gini coefficient measuring the spatial concentration of economic activity within first-order regions by 2.2 percentage points. Given an average within-region Gini coefficient of 53.5, a single transport project reduces the Gini coefficient by more than four percent relative to this baseline.

One way to put these estimates into perspective is to compare them with other work in the literature. Henderson et al. (2018) estimate Gini coefficients of light intensities around the globe and relate them to initial country-level differences in education, urbanization, and income. They show that variation in economic activity within late-developing countries is strongly correlated with variables related to trade, while that of early-developing countries is primarily driven by agriculture. Developed countries have a more balanced population distribution because they have already undergone structural change, and agglomeration occurred at different times (before and after the fall in transport costs). Henderson et al. (2018) estimate that an additional year of education in 1950 reduces overall spatial inequality among grid cells in a country by about 0.0263–0.0325 Gini points. Similarly, the partial effects of an additional percentage point of urbanization or a one percent increase in income (both in 1950) on spatial inequality are in the range of two to four Gini points. Although these estimates are not available separately for inequality between and within regions and are based on cross-sections, their magnitudes illustrate that the impact of Chinese-funded infrastructure projects on the within-region distribution of activity is within the range of other fundamental determinants of the spatial allocation of activity. The effect sizes in our analysis are also broadly in line with more detailed estimates investigating the effects of railroads and highways on the spatial distribution of activity around US or Chinese cities (such as Baum-Snow et al. 2007, 2017).[48]

We conduct additional, less restrictive specifications allowing us to compare the

[48] However, they cannot be compared directly, as we do not (directly) observe industrial GDP and residential populations in first-order regions around the globe.

coefficient of Chinese-financed transport projects to those of other variables. Table A-4 in the Online Appendix presents regressions where we include additional explanatory variables in the baseline specification. Given that we would like to include variables that do not vary over time, we omit fixed effects for regions and report conditional correlations with spatial concentration. Similar to Henderson et al. (2018), we include a range of variables related to market access and the level of development, one at a time. Our estimate of the effect of transport projects is hardly affected by the inclusion of these additional variables. Our results show that the effect of one additional transport project is often an order of magnitude larger than that of the controls. For example, it is about eight times larger than a region going from zero to complete urbanization and about 83% of the effect of doubling lights per capita in a region. These partial correlations are, of course, only indicative of the relative effect sizes, as these variables are themselves endogenous to spatial concentration.

Placebo tests and robustness checks

Our robustness checks focus on spatial concentration *within* regions; we do not report results for concentration *between* regions as we find no significant association with Chinese-financed transport projects. Table 2 presents alternative measures of our variable of interest. Column 1 uses the annual number of new project locations within a region rather than the binary project indicator as the variable of interest. Considering that the average project has 2.4 locations, the estimated effect corresponds roughly to our baseline estimate in column 1 of Table 1, where the coefficient is about two-and-a-half times larger. The corresponding first stage weakens noticeably but is still above the rule-of-thumb value of ten.

[Table 2 about here.]

Column 2 uses a binary indicator for projects that are known to have been completed, which pertains to 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 3, 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.[49] The estimated coefficient is smaller and less precisely estimated, but stays significant at the ten-percent level. It is not surprising that our result holds given that our LATE loads heavily on physical infrastructure.

[49]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).

Next, we narrow our definition of a transport project. We expect decentralization effects to be particularly pronounced for roads and railways. Specifically, we include only road projects (urban roads and long-distance roads) in column 4, both road and rail projects (where the latter includes urban railways, tram lines, and long-distance railways) in column 5, and all urban transport projects (urban roads, urban railways, and tram lines) in column 6. The estimated coefficients are larger in all three cases compared to the baseline.[50] In line with theory, it appears that the expansion of the road and rail network (in and around cities) leads to a relocation of activity from the center and generates more sprawl. Finally, column 7 of Table 2 replicates our baseline regression with an additional control variable for the construction phase. This binary indicator $\Delta N_{ji,t-1}$ takes a value of one if a Chinese government-financed transport project has been committed in the previous year. The resulting estimate of $\Delta N_{ji,t-2}$ is only slightly smaller and remains statistically significant at conventional levels. This suggests that our finding is not primarily driven by light emitted during the construction process of new transport infrastructure.

Table 3 presents the results of tests that 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 where 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. Most importantly, it is also possible that the allocation of Chinese transport projects could be highly correlated with the allocation of other Chinese projects, so that our LATE reflects the effect of all projects rather than just transport projects.

[Table 3 about here.]

Columns 1 to 3 focus on three time-varying shocks: Chinese FDI, imports from China, and exports to China.[51] While we do not have location-specific estimates for FDI and trade, they are typically linked to China's demand for natural resources or increasing trade integration with China, 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

[50]Note, however, the much-reduced first-stage F-statistic associated with the results we report in column 6.

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

mines and oil fields).[52] Columns 1 to 3 in Table 3 show that our results hold when we control for these location-specific shocks to Chinese influence.[53]

Column 4 tests whether the presence of a Chinese non-transport project in a region affects our results. To this end, we make use of a second instrument and run instrumental-variable regressions with two endogenous variables (transport projects, $\Delta N_{ji,t-2}$, and non-transport projects, $\Delta M_{ji,t-2}$) and two instruments. We construct the instrument for non-transport projects in analogy to our baseline, i.e., the time-varying component is the first factor of the six production inputs and the cross-sectional component is the fraction of years over the 2000–2014 period in which a Chinese non-transport project has been committed ($F_{t-3} \times \bar{M}_{ji}$). We find that our results for transport projects hardly change once we control for non-transport projects. This implies that our results do not reflect the effect of all projects but those of transport projects specifically.[54]

The final column 5 of Table 3 includes the three time-varying shocks and the indicator for non-transport projects in concert. Again, our results remain unchanged. The coefficient of transport projects remains significant and hardly varies in magnitude. While this does not rule out that other shocks that correlate with our instrument and spatial concentration bias our coefficient, given that the arguably most obvious sources of bias hardly change our results, we expect such bias to be negligible.

The next set of tests focus on the instrument itself. Christian and Barrett (2017) suggest tests to probe the validity of the assumptions underlying our instrumental variable approach. In addition to visual inspection of trends in Figure 4 and Figure A-1, we conduct a randomization inference test where we reassign the transport project indicator and instrumental variable to different countries and years in the sample. As can be seen from the results of 999 Monte Carlo simulations shown in Figure A-2, the resulting coefficient estimates center around zero. According to an exact Fisher test, the coefficient from our main estimate (introduced below and indicated by the vertical dashed line) significantly differs from the randomized coefficients (p-value = 0.016). The same holds when we break the timing structure required for identification less radically and instead randomize (i) the entire time series between regions, (ii) years within regions, and (iii) regions within years (also shown in Figure A-2). Thus, omitted variables are unlikely to correlate with our key variables in a way that spuriously brings about our main result.

[Table 4 about here.]

Table 4 presents tests of robustness to modifications in the time-series ('shock') and

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

[53] Table A-5 shows that our results also hold when we flexibly allow for non-linear trends that vary with our four proxies for differential exposure to Chinese influence.

[54] For brevity, we only report the first-stage equation for transport projects but note that the first-stage F-statistic remains relatively high (for an equation with two endogenous variables), the estimated coefficient on our primary instrumental variable hardly changes, while the instrument for non-transport projects only predicts those types of projects.

cross-sectional ('share') components of our instrument. Since our identification strategy leverages exogenous shocks, we first perturb the time series. Column 1 modifies our detrending method and residualizes each input series via a regression on the log of GDP at constant local currency units. Intuitively, this results in a measure of input growth that is faster or slower than GDP growth, as opposed to deviating from a linear trend. Column 2 ('All inputs') uses detrended Chinese aluminum, cement, glass, iron, steel, and timber production rather than their first principal component, each interacted with the same share component as before, as separate instruments.[55] This modification delivers estimates that are one standard error smaller but, more importantly, does not reject a J-test that tests whether these instruments identify the same LATEs. In the same spirit, column 3 includes the second factor we derived from our factor analysis of China's input materials (interacted with the original share) as part of a second instrument. The first-stage result suggests some information is contained in the second component, but the strength of the relationship drops and the confidence intervals of the 2SLS estimate overlap with that of our baseline.

In columns 4 and 5, we test the robustness of alternative definitions of the share component of our instrument—a region's probability of receiving a project. While we follow the approach of Nunn and Qian (2014), researchers typically use shares calculated before the shock in the broader shift-share literature. A concern might be that computing the probability of receiving a development project using all available years might bias our estimates (even though the fixed effects control for the endogenous probability). We approach this concern in two ways. First, in column 4, we re-estimate our baseline model but exclude contemporaneous projects when we calculate probabilities, excluding $\Delta N_{ji,t-2}$ from the cross-sectional average, and denote this the 'leave-one-out' instrument.[56] Second, column 5 replaces the probability of receiving a project with data from the Cold War period. We take these data from Dreher et al. (2021), who provide geocoded information on 688 Chinese project locations completed in 47 African countries over the 1956–1987 period. Column 4 of Table 4 shows that our results remain similar when we instrument projects with the 'leave-one-out' instrument. The results are still broadly in line with our main findings when we use Cold War projects in a subsample of African countries as the 'share' part of our instrument (in column 5), but the first stage is weaker. This is not surprising as pre-sample probabilities have a less direct relationship with the number of projects in any given year, making them a less potent predictor.

Column 6 reports the results of a placebo regression, using (detrended) United States

[55] We do not report the six first-stage coefficients in the table to reduce clutter.

[56] While this removes the contemporaneous correlation between the instrument and project commitments, the probability is now time-varying and is thus no longer absorbed by the region-fixed effects. Instead, we control for it in both stages of the regression. We have also calculated the probability based on time periods before the respective commitment and using projects committed to neighboring regions rather than a region itself. In both cases, first-stage F-statistics are low.

steel production rather than Chinese production values. This presents a falsification test as US steel production should be unrelated to Chinese-financed infrastructure projects. To facilitate comparison, we show the analogous regression using Chinese steel production as part of our instrument in column 7.[57] As expected, the first stage of the placebo regression collapses and is extremely weak, with a Kleibergen-Paap F-statistic of just 0.76, while the 2SLS estimate is positive and imprecisely estimated. In short, US steel production does not help to predict the commitment of Chinese-financed transport projects. On the contrary, as can be seen from column 7, first-stage and second-stage estimates are similar to our baseline results when we use China's steel production as shift component of our instrument.

Taken together, these estimates suggest that our results do not hinge on the specific choice of how we define Chinese production shocks and the probability to receive Chinese projects and that spurious trends do not drive the time-series component.

[Table 5 about here.]

Table 5 investigates the timing of the diffusion effects of Chinese transport projects in more detail. Recall that we have determined the two-year lag duration for our analysis based on Dreher et al. (2021), who provide start and end dates for 300 projects.[58] Panel a fixes the one-year lag between the instrument and Chinese transport projects but shifts both forward and backward in time. Column 1 starts with the effects of projects in the year of commitment ($s = 0$) instrumented by the first lag of project inputs. Column 5 ends with the effect of projects committed four years earlier ($s = 4$), instrumented by the fifth lag of project inputs. To allow for a meaningful analysis of the effect of the lag structure, we hold the sample constant across all regressions, resulting in fewer observations compared to the baseline regression. Column 3 reports results analogous to our baseline specification. The results are somewhat weaker compared to those reported above but remain qualitatively similar. Estimates become smaller for longer lags (see columns 4 and 5) and are imprecisely estimated, even though the first-stage relationship between inputs and commitments is strong for deeper lags. Column 2 shows the effects on concentration one year after commitment. The coefficients are similar to our baseline regression in magnitude but are estimated less precisely, and the first-stage relationship becomes weak. When we turn to contemporary effects in the year of commitment in column 1, the first stage turns out even weaker and the estimates are no longer significant at conventional levels.[59] Taken together, these results support our choice of a two-year

[57] Note that none of these results control for the other "China shocks" from the previous table, but the results hardly change if we include them.

[58] More precisely, the average observed project duration in this subsample is 664 days. Dreher et al. (2021) point out that this lag choice aligns with the prevailing belief held by development practitioners and government officials in host countries that Chinese development projects are executed swiftly.

[59] We have also tested whether future projects predict past concentration (not reported). Significant

lag and show the robustness to using a one-year lag, in line with the notion that much of the impact occurs early.

Panel b of Table 5 examines the lag between our instrument and project commitments. We focus on transport projects committed in $t-2$, but vary the lag of the instrument between one and four years (corresponding to $s=0$ to $s=4$). We do not claim that variations of our instrumental variable are excludable for different years but consider this as a predictive exercise to find the lag structure for which the relationship is strongest. For example, projects committed in 2010 are likely to be driven by input materials not only in 2009, but also in other adjacent years. When we vary the lag between projects and input materials, the first-stage relationship is strongest for our preferred one-year lag between inputs and commitments, and declines for all other timings. The first-stage relationship between inputs and projects completely breaks down when we predict projects with inputs produced one year in the future (column 1). Overall, these result support our choice of a one-year lag between inputs and commitments.

Extensions

Our finding that Chinese-financed transport projects reduce the concentration of economic activity within subnational regions raises the question of where exactly this diffusion takes place. The monocentric city model implies that we should observe a shift in activity from cities to their immediate periphery (and, hence, expect some heterogeneity with respect to the level of urbanization). 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 less densely developed areas.

Other frameworks offer some guidance. Heblich et al. (2020) study how transportation technology affects the specialization of locations within a city into a workplace and residence in a large range of quantitative urban models. A key finding is that improvements in transport technology lead to faster growth of suburbs relative to the central city by developing open fields and building out preexisting villages. Core-periphery models, in turn, offer predictions on overall activity as regions become better connected. For example, Faber (2014) shows that the effect of infrastructure improvements is negative for peripheral regions but heterogeneous in the level of pre-existing trade integration and relative market size between the core and periphery. We take these results as motivation to study the effects of Chinese infrastructure investment for different moments of the distribution of nighttime lights, different regions of the world, and different levels of urbanization and trade integration.

estimates of future projects on past concentration would substantially weaken the credibility of our estimation strategy. As expected, the "effect" of projects one or two years in the future on today's concentration is estimated very imprecisely, with first-stage F-statistics below one.

[Table 6 about here.]

Table 6 examines different moments of the distribution. Columns 1 and 2 show a strong effect of transport projects on overall economic activity but not on our proxy for per capita incomes. A new transport project increases the average light density in a region by about 15 percent (column 1), which is both economically and statistically significant. When we instead focus on light per capita (column 2), we cannot reject the null hypothesis that Chinese transport projects have no effect on changes in light per capita and estimate a coefficient close to zero. This aligns with the results of Dreher et al. (2022), who also report a null effect of Chinese development finance in general on lights per capita in a global sample. The insignificant coefficient for the world sample stands in contrast to results for the African continent, where previous work finds positive effects of aid on development (Dreher et al. 2021, 2022). Since the per capita data use interpolated population data in the denominator, we cannot rule out that the imprecise estimate occurs due to added noise in the dependent variable. Column 3 uses the fraction of illuminated pixels as the dependent variable. The insignificant result suggests that economic activity does not seem to expand primarily into previously undeveloped areas. To summarize, columns 1 to 3 show that Chinese transport projects increase economic activity in the receiving region. This may represent an increase in population rather than welfare and appears to be primarily occurring in already somewhat developed areas.

The remaining columns of Table 6 focus on relative changes in economic activity across quintiles of the light distribution. This allows us to directly trace which type of changes reduce the Gini coefficient. The pattern is consistent with predictions from urban land use theory. 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 lower quintiles (though estimated imprecisely for the second quintile and with borderline significance for the first). It thus appears that Chinese transport projects gradually redistribute activity from the most densely developed parts of regions, that is, the city centers, to less densely developed places. The magnitudes of the estimated coefficients suggest that this process benefits the higher quintiles more relative to the least developed parts of a region.

[Table 7 about here.]

Table 7 presents a more direct approach to measuring 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 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 above-median proximity to the coast. The estimated effects are substantial in these sub-samples. By contrast, they are imprecisely estimated and typically of the opposite sign or lower in magnitude in the other sub-samples (the exception being below-median proximity to the coast, with similar estimated magnitudes and standard errors in both samples). Last but not least, we also find that the effect seems to be driven by relatively poor regions as measured by below-median light per capita. This 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 aligns 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 concentration were larger within coastal and richer central regions. Similarly, although they do not focus on decentralization within regions per se, studies focusing on the spatial impact of the Belt and Road Initiative (BRI) typically estimate that coastal regions, border crossings, and urban hubs will benefit more (Lall and Lebrand 2020).[60]

Next, we investigate major world regions separately. China's global infrastructure footprint is uneven and most of its transportation projects are located in Africa and Asia (recall Figure 2). Urban population growth is rapid and infrastructure constraints are most severe in these regions. As can be seen from the sub-samples in columns 1 to 3 of Table 8, our main findings are driven by regions in African countries, where the effect is larger than our baseline estimates. The coefficient on Chinese transport projects is insignificant or substantially smaller 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 8 about here.]

Column 4 restricts the sample to countries classified by the World Bank as low-income economies in 2000. It highlights that the diffusion effects of Chinese transport projects also 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 that would arise if regions that received any development-related project from

[60] 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.

China experience different non-linear trends than those which did not. Our results become substantially stronger.

Finally, Table A-6 in the Online Appendix 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 (also recall that our results are robust to including any non-transport project from China, as shown in column 4 of Table 3 above).

6 Conclusion

This article examined whether and to what extent transport infrastructure projects decentralize economic activity in recipient regions across the Global South. We overcome the challenge of missing geolocalized data on comparable infrastructure projects across countries and how to estimate their causal effect by focusing on infrastructure projects financed by the Chinese government—a single but massive source of infrastructure financing across developing countries. 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 focus of China's BRI since its announcement in 2013, and the Chinese government similarly focused on financing and building connective infrastructure projects before the BRI. Transport projects such as roads, highways, railways, harbors, and airports are at the heart of this approach, and China's government has financed hundreds of them in developing countries since 2000 (Dreher et al. 2022).

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 and between recipient regions. 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 that 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. While we find similar effect sizes for concentration between regions, these effects are estimated less precisely. Our within-region results imply that the Gini coefficient measuring the spatial concentration of economic activity is reduced by 2.2 percentage points within first-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 African countries, which most need infrastructure financing. In line with urban land use theory, we find that our results are driven by changes in economic activity in and around urban areas.

In financing major transport projects, China's government appears to be helping cities and regions in developing countries transform from dense, crowded, and unproductive places into productive hubs. While these results are encouraging, they do not imply that Chinese government-financed transport infrastructure only have 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 regions, often at the expense of poorer regions with greater material need (Dreher et al. 2019). There are many other concerns about the consequences of China's development finance, ranging from their impact on the environment to questions of debt sustainability (Horn et al. 2021, Baehr et al. 2023). 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.

References

Adão, R., M. Kolesár, and E. Morales (2019). Shift-Share Designs: Theory and Inference. *Quarterly Journal of Economics 134*(4), 1949–2010.

African Development Bank (2014a). Kenya: Nairobi-Thika Highway Improvement Project. Abidjan, Côte d'Ivoire: African Development Bank. Available at https://www.afdb.org/fileadmin/uploads/afdb/Documents/Project-and-Operations/Presidential_awards_2014_-_Kenya_-_Nairobi_-_Thika_Highway_Improvement_Project.pdf (accessed 3 April 2020).

African Development Bank (2014b). Study on Quality of Bank Financed Road Projects. Abidjan, Côte d'Ivoire: African Development Bank. Available at https://www.afdb.org/fileadmin/uploads/afdb/Documents/Events/ATFforum/Study_on_the_quality_of_Bank_Financed_Road_Projects_-_AfDB.pdf (accessed 3 April 2020).

African Development Bank (2016). PCR Evaluation Note for Public Sector Operations: Nairobi-Thika Highway Improvement Project. Abidjan, Côte d'Ivoire: African Development Bank. Available at https://evrd.afdb.org/documents/docs/EN_PN10706.pdf (accessed 3 April 2020).

African Development Bank (2019). Nairobi-Thika Highway Improvement Project: Project Completion Report. Abidjan, Côte d'Ivoire: African Development Bank. Available at https://www.afdb.org/sites/default/files/documents/projects-and-operations/kenya_-_thika_highway_improvement_project.pdf (accessed 3 April 2020).

African Development Fund (2007). Appraisal Report: Nairobi–Thika Highway Improvement Project. Tunis, Tunisia: African Development Fund. Available at https://www.afdb.org/fileadmin/uploads/afdb/Documents/Project-and-Operations/Kenya_-_Nairobi-Thika_Highway_Improvement_Project_-_Appraisal_Report.PDF (accessed April 2020).

Allen, T. and C. Arkolakis (2022). The Welfare Effects of Transportation Infrastructure Improvements. *Review of Economic Studies 89*(6), 2911–2957.

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

Anaxagorou, C., G. Efthyvoulou, and V. Sarantides (2020). Electoral Motives and the Subnational Allocation of Foreign Aid in Sub-Saharan Africa. *European Economic Review 127*, 103430.

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

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 (2013). The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American Economic Review 103*(6), 2121–68.

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.

Baehr, C., A. BenYishay, and B. Parks (2023). Highway to the Forest? Land Governance and the Siting and Environmental Impacts of Chinese Government-Funded Road Building in Cambodia. *Journal of Environmental Economics and Management 122*, 102898.

Bandiera, L. and V. Tsiropoulos (2020). A Framework to Assess Debt Sustainability under the Belt and Road Initiative. *Journal of Development Economics 146*, 102495.

Banerjee, A., E. Duflo, and N. Qian (2020). On the Road: Access to Transportation Infrastructure and Economic Growth in China. *Journal of Development Economics 145*, 102442.

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. et al. (2007). Suburbanization and Transportation in the Monocentric Model. *Journal of Urban Economics 62*(3), 405–423.

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*, 102441.

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 #6964.

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

Borusyak, K., P. Hull, and X. Jaravel (2020). Quasi-experimental Shift-Share Research Designs. NBER Working Paper 24997. Cambridge, MA: 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 (2021). Aid Curse with Chinese Characteristics? Chinese Development Flows and Economic Reforms. *Public Choice 188*, 407–430.

Bruederle, A. and R. Hodler (2018). Nighttime Lights as a Proxy for Human Development at the Local Level. *PloS one 13*(9), e0202231.

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.

Che, Y., X. He, and Y. Zhang (2021). Natural Resource Exports and African Countries' Voting

Behaviour in the United Nations: Evidence from the Economic Rise of China. *Canadian Journal of Economics/Revue canadienne d'économique 54*(2), 712–759.

Christian, P. and C. Barrett (2017). Revisiting the Effect of Food Aid on Conflict: A Methodological Caution. *Policy Research Working Paper 8171.*

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

de Chaisemartin, C. and X. d'Haultfoeuille (2020). Two-way Fixed Effects Estimators With Heterogeneous Treatment Effects. *American Economic Review 110*(9), 2964–2996.

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*, 102415.

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 (2019). 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 (2021). Is Favoritism a Threat to Chinese Aid Effectiveness? A Subnational Analysis of Chinese Development Projects. *World Development 139*, 105291.

Dreher, A., A. Fuchs, and S. Langlotz (2019). The Effects of Foreign Aid on Refugee Flows. *European Economic Review 112*, 127–147.

Dreher, A., A. Fuchs, B. Parks, A. Strange, and M. Tierney (2021). Aid, China, and Growth: Evidence from a New Global Development Finance Dataset. *American Economic Journal: Economic Policy 13*(2), 135–174.

Dreher, A., A. Fuchs, B. Parks, A. Strange, and M. J. Tierney (2022). *Banking on Beijing: The Aims and Impacts of China's Overseas Development Program.* Cambridge University Press.

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.

Dreher, A. and S. Langlotz (2020). Aid and Growth: New Evidence Using an Excludable Instrument. *Canadian Journal of Economics 53*(3), 1162–1198.

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 (2021). The Effects of Trade, Aid and Investment on China's Image in Latin America. *Journal of Comparative Economics 49*(2).

Elvidge, C. D., K. E. Baugh, M. Zhizhin, and F.-C. Hsu (2013). Why VIIRS data are superior to DMSP for mapping nighttime lights. *Proceedings of the Asia-Pacific Advanced Network 35*, 62–69.

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.

Garcia-Lopez, M.-A., A. Holl, and E. Viladecans-Marsal (2015). Suburbanization and Highways in Spain When the Romans and the Bourbons Still Shape its Cities. *Journal of Urban Economics 85*, 52–67.

Gehring, K., L. Kaplan, and M. H. Wong (2022). China and the World Bank: How Contrasting Development Approaches Affect the Stability of African States. *Journal of Development Economics 158*, 102902.

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.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020). Bartik Instruments: What, When, Why, and How. *American Economic Review 110*(8), 2586–2624.

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*, 101265.

He, G., Y. Xie, and B. Zhang (2020). Expressways, GDP, and the Environment: The Case of China. *Journal of Development Economics 145*, 102485.

Heblich, S., S. J. Redding, and D. M. Sturm (2020). The Making of the Modern Metropolis: Evidence from London. *Quarterly Journal of Economics 135*(4), 2059–2133.

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, J., A. Storeygard, and D. Weil (2012). Measuring Economic Growth from Outer Space. *American Economic Review 102*, 994–1028.

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.

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

Hoeffler, A. and O. Sterck (2022). Is Chinese Aid Different? *World Development 156*, 105908.

Horn, S., C. M. Reinhart, and C. Trebesch (2021). China's Overseas Lending. *Journal of International Economics 133*, 103539.

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

Iacoella, F., B. Martorano, L. Metzger, and M. Sanfilippo (2021). Chinese Official Finance and Political Participation in Africa. *European Economic Review 136*, 103741.

Isaksson, A.-S. (2020). Chinese Aid and Local Ethnic Identification. *International Organization 74*(4), 833–852.

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.

Jedwab, R. and A. Storeygard (2022). The Average and Heterogeneous Effects of Transportation Investments: Evidence from Sub-Saharan Africa 1960-2010. *Journal of the European Economic Association 20*(1), 1–38.

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 146*, 102496.

Lang, V. (2021). The Economics of the Democratic Deficit: The Effect of IMF Programs on Inequality. *Review of International Organizations 16*(3), 599–623.

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.

Marchesi, S., T. Masi, and S. Paul (2021). Project Aid and Firm Performance. University of Milan Bicocca Department of Economics, Management and Statistics Working Paper No. 479.

Martorano, B., L. Metzger, and M. Sanfilippo (2020). Chinese Development Assistance and Household Welfare in Sub-Saharan Africa. *World Development 129*, 104909.

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).

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.

Olea, J. L. M. and C. Pflueger (2013). A Robust Test for Weak Instruments. *Journal of Business & Economic Statistics 31*(3), 358–369.

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).

Ping, S.-N., Y.-T. Wang, and W.-Y. Chang (2022). The Effects of China's Development Projects on Political Accountability. *British Journal of Political Science 52*(1), 65–84.

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

Redding, S. J. and E. Rossi-Hansberg (2017). Quantitative Spatial Economics. *Annual Review of Economics 9*(1), 21–58.

Redding, S. J. and M. A. Turner (2015). Transportation Costs and the Spatial Organization of Economic Activity. In G. Duranton, J. V. Henderson, and W. C. Strange (Eds.), *Handbook* *of Regional and Urban Economics*, Volume 5 of *Handbook of Regional and Urban Economics*, Chapter 20, pp. 1339–1398. Elsevier.

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

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

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

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.

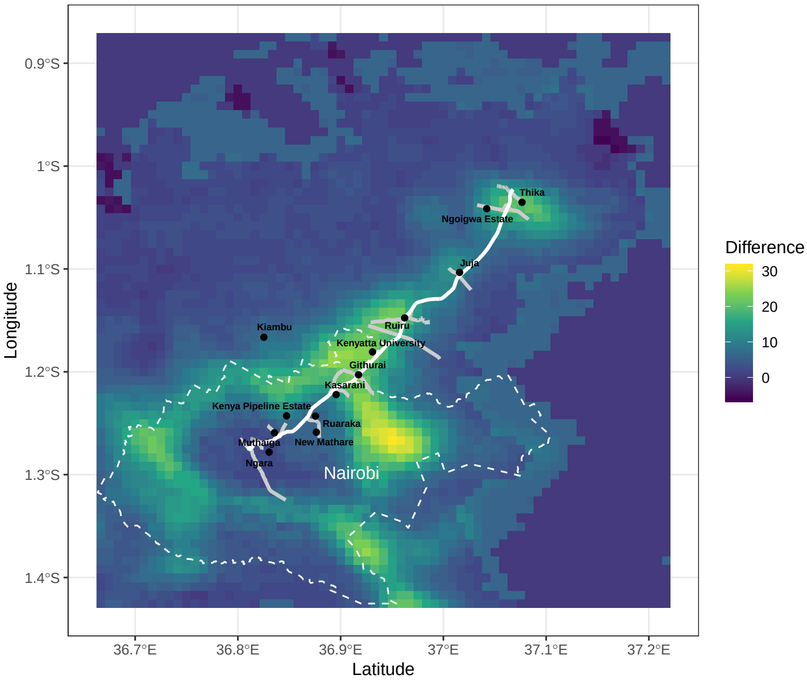
Wellner, L., A. Dreher, A. Fuchs, B. C. Parks, and A. Strange (2024). Can Aid Buy Foreign Public Support? Evidence from Chinese Development Finance. *Economic Development and Cultural Change forthcoming.*

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.

Zárate, R. D. (2022). Spatial Misallocation, Informality, and Transit Improvements: Evidence from Mexico City. Policy Research Working Paper Series #9990, World Bank.

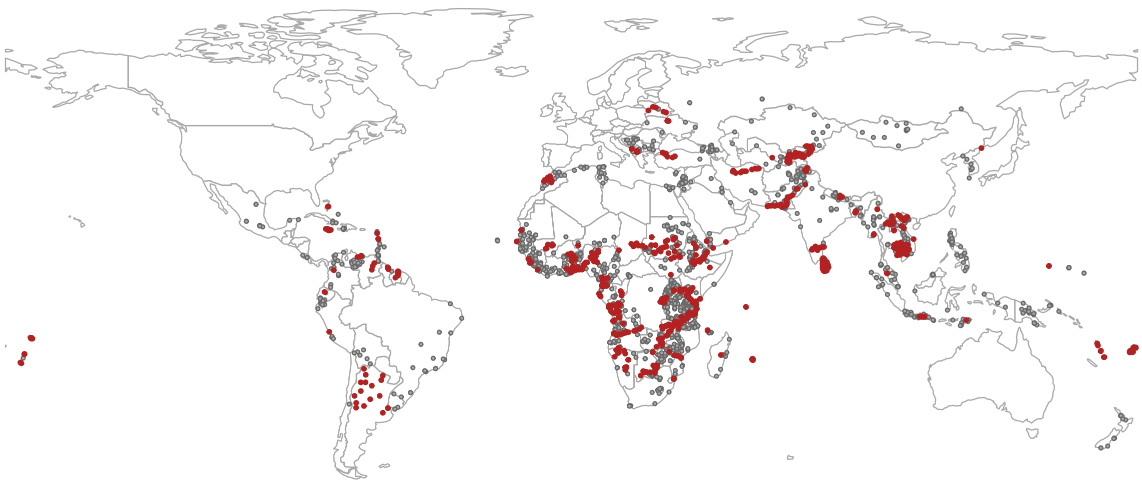
Zeitz, A. (2021). Emulation or Differentiation? China's Development Finance and Traditional Donor Aid in Developing Countries. *Review of International Organizations 16*, 265–292.

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



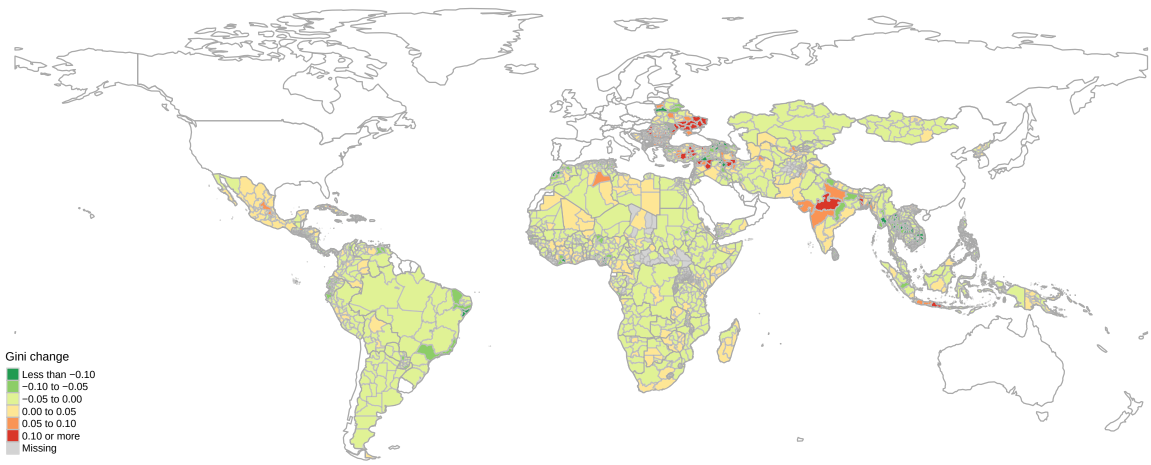
Notes: The figure illustrates the change in nighttime lights from 2008 to 2013 along the route of the Nairobi-Thika Highway in Kenya, which was constructed from January 2009 until October 2012. Major intersections and points of interest are highlighted along the highway. The change in nighttime lights 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. 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. 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. 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 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).

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



Notes: The figure identifies 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.

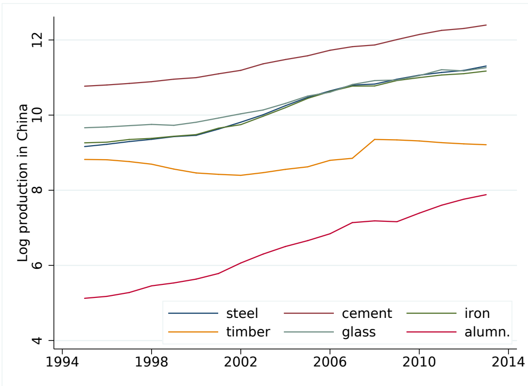
Figure 3 – Long differences in spatial concentration, within first-order regions, 2000–2013



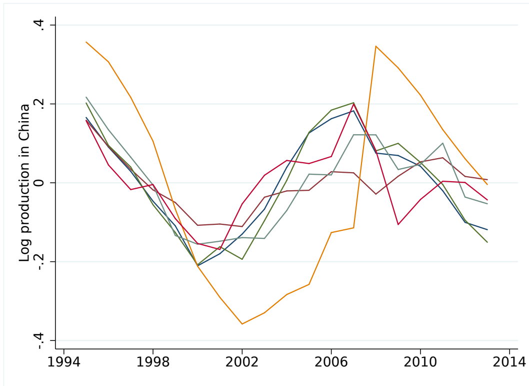
Notes: The figure illustrates the cross-regional and temporal variation in spatial concentration. It shows long differences in the Gini coefficient for spatial concentration within first-order regions, that is, a region's value in 2013 minus the value in 2000. Only countries not classified as high-income countries 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 physical project inputs, 1995–2013

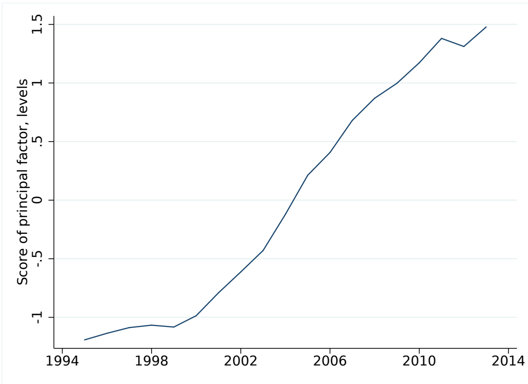
(a) Log levels



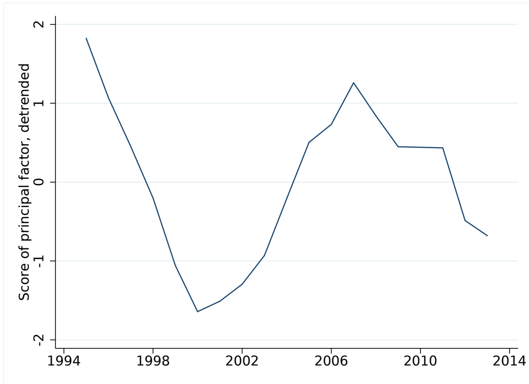
(b) Detrended (a)



(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 and after 2000, i.e., before and after 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, first-order regions, 2002–2013

	Spatial concentration, $\Delta\textsc{Gini}_{jit}$, measured...			
	within first-order regions		between first-order regions	
	Projects	Values	Projects	Values
	(1)	(2)	(3)	(4)
	Panel a) OLS estimates			
Projects ($\Delta N_{ji,t-2}$)	0.0026	0.0002	-0.0042	-0.0002
	(0.0020)	(0.0001)	(0.0043)	(0.0002)
	[0.0021]	[0.0001]	[0.0043]	[0.0002]
	Panel b) Reduced-form estimates			
IV ($F_{t-3} \times \bar{N}_{ji}$)	-0.0096	-0.0096	-0.0075	-0.0075
	(0.0049)*	(0.0049)*	(0.0073)	(0.0073)
	[0.0032]***	[0.0032]***	[0.0073]	[0.0073]
	Panel c) 2SLS estimates			
Projects ($\Delta N_{ji,t-2}$)	-0.0218	-0.0014	-0.0224	-0.0012
	(0.0097)**	(0.0006)**	(0.0215)	(0.0011)
	[0.0073]***	[0.0005]***	[0.0216]	[0.0011]
	Panel d) First-stage estimates			
IV ($F_{t-3} \times \bar{N}_{ji}$)	0.4400	6.8244	0.3361	6.0497
	(0.0747)***	(1.1363)***	(0.0638)***	(1.1467)***
	[0.0688]***	[1.1546]***	[0.0636]***	[1.1359]***
Level of analysis	ADM1	ADM1	ADM0	ADM0
First-stage F-Stat	34.70	36.07	27.73	27.83
Observations	27,162	27,162	1,386	1,386
Regions	2,406	2,406	–	–
Countries	122	122	122	122

Notes: The table reports regression results within and between first-order administrative regions. 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 and 2 include region-fixed effects and country-year-fixed effects; columns 3 and 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. *** p<0.01, ** p<0.05, * p<0.1.

Table 2 – Variants of baseline regression, within first-order regions, 2002–2013

	Variations of the variable of interest ($\Delta N_{ji,t-2}$)						
	Location count (1)	Completed (2)	Economic infra. (3)	Roads (4)	Roads & rail (5)	Urban transport (6)	Transport (7)
	Panel a) 2SLS estimates						
Projects ($\Delta N_{ji,t-2}$)	-0.0086	-0.0412	-0.0165	-0.0274	-0.0230	-0.0355	-0.0155
	(0.0037)**	(0.0201)**	(0.0092)*	(0.0087)***	(0.0131)*	(0.0118)***	(0.0091)*
	[0.0031]***	[0.0143]***	[0.0095]*	[0.0118]**	[0.0083]***	[0.0153]**	[0.0065]**
	Panel b) First-stage estimates						
IV ($F_{t-3} \times \bar{N}_{ji}$)	1.1205	0.2329	0.3600	0.4723	0.4086	0.4886	0.4895
	(0.3243)***	(0.0545)***	(0.0468)***	(0.1067)***	(0.0912)***	(0.2555)*	(0.0879)***
	[0.2782]***	[0.0591]***	[0.0555]***	[0.0988]***	[0.0857]***	[0.2390]**	[0.0798]***
Construction phase ($\Delta N_{ji,t-1}$)	–	–	–	–	–	–	✓
First-stage F-stat	11.94	18.26	59.16	19.59	20.06	3.66	31.03
Observations	27,162	27,162	27,162	27,162	27,162	27,162	27,162
Regions	2,406	2,406	2,406	2,406	2,406	2,406	2,406
Countries	122	122	122	122	122	122	122

Notes: The table reports regression results within first-order administrative regions. 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 the corresponding first-stage where the dependent variable is indicated in the column header. '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). 'Roads' narrows the definition of our base measure by including only road projects (urban roads and long-distance roads). 'Roads & rail' narrows the definition of our base measure by including only road and railroad projects (urban roads, long-distance roads, urban railways and tram lines, and long-distance railways). 'Urban transport' narrows the definition of our base measure by including only urban transport projects (urban roads and urban railways and tram lines). Whenever we change the dependent variable, the cross-sectional component of the instrument follows the same definition of projects. The last column ('Transport') uses the baseline measure of transport projects and adds a control variable for the construction phase, which is a binary indicator $\Delta N_{ji,t-1}$ that takes a value of one if a Chinese government-financed transport project has been committed in the previous year. 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. *** p<0.01, ** p<0.05, * p<0.1.

Table 3 – Identification: Other "China shocks," within first-order regions, 2002–2013

	Accounting for other "China shocks"				
	FDI	Imports	Exports	Projects	All
	(1)	(2)	(3)	(4)	(5)
	Panel a) 2SLS estimates				
Projects ($\Delta N_{ji,t-2}$)	-0.0218	-0.0208	-0.0211	-0.0205	-0.0195
	(0.0097)**	(0.0096)**	(0.0096)**	(0.0106)*	(0.0105)*
	[0.0073]***	[0.0071]***	[0.0071]***	[0.0083]**	[0.0083]**
	Panel b) First-stage estimates				
IV ($F_{t-3} \times \bar{N}_{ji}$)	0.4400	0.4571	0.4604	0.4374	0.4570
	(0.0746)***	(0.0747)***	(0.0748)***	(0.0749)***	(0.0750)**
	[0.0715]***	[0.0726]***	[0.0726]***	[0.0727]***	[0.0740]***
IV ($F_{t-3} \times \bar{M}_{ji}$)				0.0040	0.0009
				(0.0099)	(0.0095)
				[0.0096]	[0.0095]
Δ (Shock $\times$ Distance)	✓	✓	–	–	✓
Δ (Shock $\times$ Urbanization)	✓	✓	–	–	✓
Δ (Shock $\times$ Large mines)	–	–	✓	–	✓
Δ (Shock $\times$ Oil fields)	–	–	✓	–	✓
First-stage(s) F-stat	34.79	37.43	37.84	8.74	7.44
Observations	27,162	26,491	26,491	27,162	26,491
Regions	2,406	2,348	2,348	2,406	2,348
Countries	122	116	116	122	116

Notes: The table reports regression results within first-order administrative regions. 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 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. 'FDI' are Chinese FDI outflows (in logs of current US$) from UNCTAD. 'Distance' is the average "as-the-crow-flies" distance from the region to the nearest coastline from Natural Earth. 'Urbanization' is the fraction of land which is defined as an urban cluster or urban center in 2000 by the Global Human Settlement Layer (Pesaresi et al. 2019). 'Imports' are the imports of the recipient country from China; 'Exports' are the exports of the recipient country to China (both in logs of current US$ and from the IMF Direction of Trade Statistics). 'Large mines' indicate that 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). 'Projects' is binary and indicates the presence of a Chinese non-transport project in the region in $t-2$. The instrument is constructed in analogy to our baseline, i.e., the cross-sectional component is the fraction of years over the 2000–2014 period in which a Chinese non-transport project has been committed ($\bar{M}_{ji}$). 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. *** p<0.01, ** p<0.05, * p<0.1.

Table 4 – Identification: Altering the instrument, within first-order regions, 2002–2013

	Overprod. (1)	All inputs (2)	Both factors (3)	Leave-one-out shares (4)	Cold War shares Africa (5)	US steel placebo (6)	Det. steel (7)
	Panel a) 2SLS estimates						
Projects ($\Delta N_{ji,t-2}$)	-0.0319	-0.0170	-0.0133	-0.0243	-0.0563	0.0952	-0.0262
	(0.0118)***	(0.0109)	(0.0067)**	(0.0099)**	(0.0288)*	(0.1188)	(0.0111)**
	[0.0103]***	[0.0081]**	[0.0066]**	[0.0077]***	[0.0508]	[0.1112]	[0.0089]***
	Panel b) First-stage estimates						
Instrument 1	0.3360	–	0.4819	0.4881	0.0778	-0.0658	0.3495
	(0.0597)***		(0.0853)***	(0.0860)***	(0.0350)**	(0.0752)	(0.0642)***
	[0.0612]***		[0.0772]***	[0.0785]***	[0.0301]***	[0.0749]	[0.0631]***
Instrument 2			0.1748				
			(0.1026)*				
			[0.0778]**				
Additions/Modifications	Modified shock	All six instruments	Modified shock	Modified shares	Modified shares	Placebo instrument	Modified shock
J-test (p-value)	–	0.339	0.090	–	–	–	–
First-stage F-stat	29.82	18.57	19.44	38.27	6.612	0.762	30.40
Observations	27,167	27,167	27,167	27,167	8,401	27,167	27,167
Regions	2,406	2,406	2,406	2,406	729	2,406	2,406
Countries	122	122	122	122	48	122	122

Notes: The table reports regressions results within first-order administrative regions. 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 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. 'Overprod.' implies that the factor inputs were residualized by running a regression of each input on the log of Chinese GDP at constant local currency units before the first factor was extracted. 'Det. steel' uses the linearly detrended log of Chinese steel production as the time-series shock. 'All inputs' uses interactions with detrended Chinese aluminum, cement, glass, iron, steel, and timber production time series as separate instruments. 'Both factors' uses the second principal factor of the detrended inputs as a second instrument. 'Leave-one-out shares' subtracts the value of $\Delta N_{i,t-2}$ from the cross-sectional average in each period. The probability is now time-varying, so we control for it in both stages of the regression. 'Cold War shares Africa' uses the probability of receiving a project with data from the Cold War period from Bartke (1989) geocoded for Africa by Dreher et al. (2021). 'US placebo' uses a US steel production index from FRED (`IPN3311A2BS`) as part of a placebo instrument. 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. *** p<0.01, ** p<0.05, * p<0.1.

Table 5 – Timing of effects, within first-order regions, 2004–2013

	Lag structure for $\Delta N_{i,t-s}$ and F_{t-s-1}				
	$s=0$	$s=1$	$s=2$	$s=3$	$s=4$
	(1)	(2)	(3)	(4)	(5)
	Panel a) 2SLS estimates, changes in s for project and IV				
Projects ($\Delta N_{i,t-s}$)	-0.0223	-0.0188	-0.0181	-0.0045	0.0026
	(0.0187)	(0.0128)	(0.0114)	(0.0067)	(0.0079)
	[0.0167]	[0.0110]*	[0.0081]**	[0.0073]	[0.0088]
	Panel b) 2SLS estimates, only changes in s for IV				
Projects ($\Delta N_{i,t-2}$)	-0.0563	-0.0195	-0.0181	-0.0043	0.0026
	(0.0501)	(0.0131)	(0.0114)	(0.0062)	(0.0082)
	[0.0568]	[0.0113]*	[0.0081]**	[0.0069]	[0.0089]
First-stage F-stat panel a	6.081	6.737	16.56	20.43	25.79
First-stage F-stat panel b	2.396	12.55	16.56	15.86	11.81
Observations	22,445	22,445	22,445	22,445	22,445
Regions	2,389	2,389	2,389	2,389	2,389
Countries	121	121	121	121	121

Notes: The table reports 2SLS regressions where the dependent variable is the first difference of the Gini coefficient of light intensity within first-order administrative regions. Both panels show two-stage least-squares fixed-effects regressions where the dependent variable is the Gini coefficient of light intensities within first-order administrative regions. Panel a fixes the lag structure for the first difference of projects and the instrument but varies years of commitment as indicated in the column header. For example, column 1 estimates the effects of projects committed in t, instrumented with the interaction including inputs produced in $t-1$. Panel b fixes the timing of projects but varies the lag between projects and input production. For example, column 1 reports the effect of projects committed in $t-2$, instrumented with the interaction including inputs produced in $t-2$. 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. *** p<0.01, ** p<0.05, * p<0.1.

Table 6 – Light intensity and quintile shares, within first-order regions, 2002–2013

	Moments of spatial concentration							
	Light	Light	Extensive	Quintile shares				
	density	per capita	margin	0–20%	20–40%	40–60%	60–80%	80–100%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel a) 2SLS estimates							
Projects ($\Delta N_{ji,t-2}$)	0.1462	-0.0005	0.0122	0.0028	0.0032	0.0102	0.0191	-0.0353
	(0.0476)***	(0.0028)	(0.0173)	(0.0015)*	(0.0027)	(0.0040)**	(0.0075)**	(0.0114)***
	[0.0516]***	[0.0028]	[0.0104]	[0.0019]	[0.0024]	[0.0043]**	[0.0078]**	[0.0116]***
	Panel b) First-stage estimates							
IV ($F_{t-3} \times \bar{N}_{ji}$)	0.4505	0.4498	0.4505	0.4394	0.4394	0.4394	0.4394	0.4394
	(0.0732)***	(0.0732)***	(0.0732)***	(0.0749)***	(0.0749)***	(0.0749)***	(0.0749)***	(0.0749)***
	[0.0688]***	[0.0688]***	[0.0688]***	[0.0716]***	[0.0716]***	[0.0716]***	[0.0716]***	[0.0716]***
First-stage F-stat	37.86	37.72	37.86	34.44	34.44	34.44	34.44	34.44
Observations	28,037	28,025	28,037	26,877	26,877	26,877	26,877	26,877
Regions	2,440	2,439	2,440	2,379	2,379	2,379	2,379	2,379
Countries	122	122	122	122	122	122	122	122

Notes: The table reports 2SLS regressions where the dependent variable is the first difference of the Gini coefficient of light intensity within first-order administrative regions. Panel a shows two-stage least-squares fixed-effects regressions where the dependent variable is indicated in the column header. Panel b 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. We use the inverse hyperbolic sine transformation in columns 1 and 2, which is defined as $ihs(z) = log(z + \sqrt{z^2 + 1})$, for light density and light per capita to include regions with zero light and retain an interpretation similar to logs. The extensive margin in column 3 is defined as the (untransformed) 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. *** p<0.01, ** p<0.05, * p<0.1.

Table 7 – Sample splits, within first-order regions, 2002–2013

	Splitting at the median of ...				
	Travel time to cities (1)	Urbanization rate (2)	Road density (3)	Distance to coast (4)	Light per capita (5)
	Panel a) Below median, 2SLS estimates				
Projects ($\Delta N_{ji,t-2}$)	-0.0306	0.0044	-0.0101	-0.0207	-0.0276
	(0.0147)**	(0.0087)	(0.0099)	(0.0154)	(0.0089)***
	[0.0104]***	[0.0103]	[0.0107]	[0.0090]**	[0.0056]***
	Panel b) Above median, 2SLS estimates				
Projects ($\Delta N_{ji,t-2}$)	0.0003	-0.0222	-0.0262	-0.0232	-0.0288
	(0.0105)	(0.0180)	(0.0121)**	(0.0096)**	(0.0287)
	[0.0097]	[0.0129]*	[0.0112]**	[0.0098]**	[0.0312]
First-stage F-stat a	11.74	20.15	22.60	23.23	21.01
First-stage F-stat b	39.68	12.11	16.90	16.33	13.76
Observations a)	13,016	13,642	13,527	13,509	13,142
Observations b)	13,875	13,254	13,451	13,545	13,672

Notes: The table reports 2SLS regressions where the dependent variable is the first difference of the Gini coefficient of light intensity within first-order administrative regions. 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. 'Travel time to cities' 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. *** p<0.01, ** p<0.05, * p<0.1.

Table 8 – Regional variation, within first-order regions, 2000–2013

	Regional subsets and related sample perturbations				
	Africa	Asia	Americas	Low income	$\bar{N}^{\text{all}} > 0$
	(1)	(2)	(3)	(4)	(5)
	Panel a) 2SLS estimates				
Projects ($\Delta N_{ji,t-2}$)	-0.0252	-0.0134	-0.0084	-0.0204	-0.0301
	(0.0076)***	(0.0244)	(0.0079)	(0.0117)*	(0.0073)***
	[0.0100]**	[0.0112]	[0.0019]***	[0.0058]***	[0.0100]***
	Panel b) First-stage estimates				
IV ($F_{t-3} \times \bar{N}_{ji}$)	0.4413	0.4171	0.7981	0.4327	0.4348
	(0.1123)***	(0.1020)***	(0.2285)***	(0.0897)***	(0.0753)***
	[0.0997]***	[0.1096]***	[0.3355]**	[0.0826]***	[0.0744]***
First-stage F-Stat	15.45	16.72	12.20	23.28	33.31
Observations	8,401	9,191	4,954	11,357	8,639
Regions	729	791	430	982	735
Countries	48	34	22	60	92

Notes: The table reports 2SLS regressions. 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 that have received any transport or non-transport project 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. *** p<0.01, ** p<0.05, * p<0.1.

Online Appendix

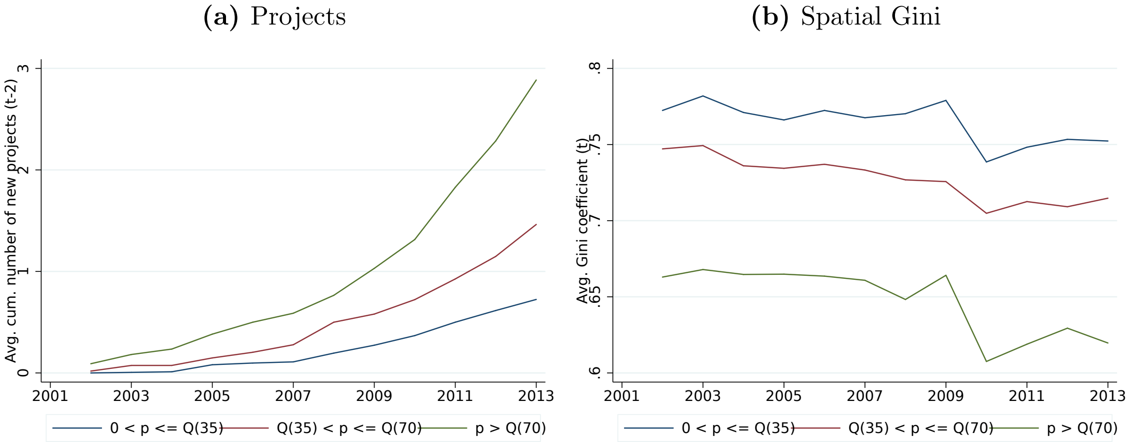
for

"Connective Financing: Chinese Infrastructure Projects and the Diffusion of Economic Activity in Developing Countries"

by

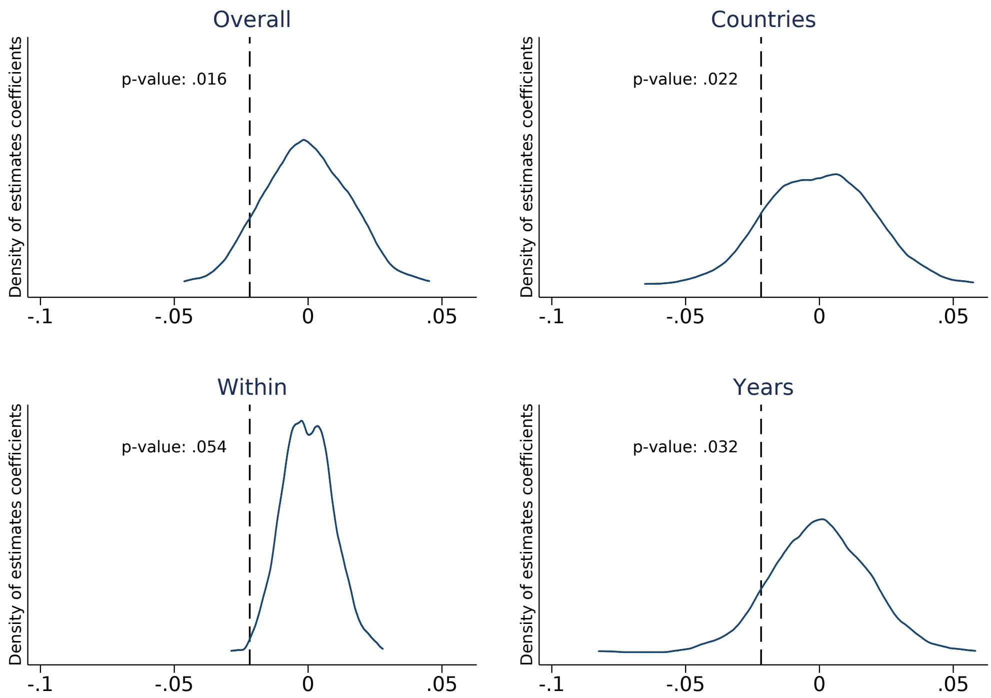
Bluhm, Dreher, Fuchs, Parks, Strange, and Tierney

Figure A-1 – Trends in the count of project-years and spatial Gini by terciles of $\bar{N}_{ji}$



Notes: The figure shows the number of transport projects in tandem with the regions' spatial Gini coefficient for first-order regions with different probabilities of receiving projects over the sample period.

Figure A-2 – Randomization inference



Notes: The figure shows the distribution of point coefficients of Chinese-financed transport projects based on 999 Monte Carlo replications under different randomization inference tests. Panel *'Overall'* swaps the project dummy and instrument for all observations, panel *'Countries'* swaps the entire time series between countries, panel *'Within'* swaps years within countries, and panel *'Years'* swaps countries within years. Dashed vertical lines indicate the original estimate from column 1 of Table 1. The p-values are calculated as the proportion of times that the absolute value of the t-statistics in the simulated data exceed the absolute value of the original t-statistic.

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 – Number of project locations by sector

Project sector	No. of project-locations
Transport and Storage, of which...	1,215
...long-distance roads	651
...long-distance railways	245
...urban roads	123
...airports	38
...bridges	38
...urban railways and tram lines	35
...several categories	24
...ports and waterways	16
...vehicles	16
...other project types not listed above	29
Health	676
Education	538
Energy Generation and Supply	350
Communications	337
Government and Civil Society	273
Emergency Response	238
Other Social Infrastructure and Services	205
Agriculture, Forestry and Fishing	193
Water Supply and Sanitation	150
Industry, Mining, Construction	107
Other Multisector	68
Developmental Food Aid/Food Security Assistance	18
General Environmental Protection	10
Other sectors not listed above	54
Total	4,432

Notes: The table lists project sectors and project types in the sector Transport & Storage together with their frequency in terms of project locations.

Table A-3 – Descriptive statistics

	N	*Mean*	*SD*	*Min*	*Max*
		Panel a) Dependent variables			
GINI_{jit} (within first-order regions)	26,901	0.534	0.160	0.000	0.849
GINI_{jit} (between first-order regions)	22,762	0.452	0.194	0.000	0.985
Light density	27,162	6.146	12.886	0.000	152.19
Light per capita	27,162	0.043	0.062	0.000	1.207
Extensive margin	27,162	0.460	0.345	0.001	1.000
Quintile share (0-20%)	25,123	0.026	0.031	0.000	0.330
Quintile share (20-40%)	25,123	0.058	0.037	0.000	0.232
Quintile share (40-60%)	25,123	0.109	0.046	0.000	0.318
Quintile share (60-80%)	25,123	0.198	0.053	0.000	0.536
Quintile share (80-100%)	25,123	0.609	0.134	0.168	1.000
		Panel b) Variables of interest			
Projects ($N_{i,t-2}$)	27,162	0.047	0.272	0.000	6.000
Projects (log 1 + financial values)	27,162	0.178	1.769	0.000	21.61
Projects (location count)	27,162	0.027	0.316	0.000	10.00
Projects (completed)	27,162	0.029	0.211	0.000	5.000
Projects (economic infrastructure)	27,162	0.098	0.416	0.000	8.000
Projects (roads)	27,162	0.007	0.085	0.000	1.000
Projects (roads & rail)	27,162	0.009	0.097	0.000	1.000
Projects (urban transport)	27,162	0.002	0.040	0.000	1.000
Projects (construction phase)	27,162	0.011	0.105	0.000	1.000
		Panel c) Instruments			
IV ($F_{t-3} \times \bar{N}_{ji}$)	27,162	-0.003	0.044	-0.821	0.630
IV ($F_{t-3} \times \bar{M}_{ji}$)	27,162	-0.009	0.118	-1.525	1.170
IV (Overproduction)	27,162	-0.001	0.045	-0.883	0.633
IV (Steel)	27,162	0.000	0.048	-0.889	0.773
IV (Cement)	27,162	-0.005	0.039	-0.816	0.395
IV (Iron)	27,162	-0.000	0.048	-0.785	0.770
IV (Timber)	27,162	-0.005	0.047	-0.736	0.718
IV (Glass)	27,162	-0.004	0.043	-0.703	0.555
IV (Aluminum)	27,162	-0.002	0.050	-0.906	1.075
IV (Second factor)	27,162	-0.005	0.045	-0.767	0.362
IV (Leave-one-out)	24,811	-0.002	0.043	-0.821	0.630
IV (Cold war Africa)	8,401	-0.004	0.058	-0.616	0.472
IV (US placebo)	27,162	1.223	4.116	0.000	55.49

continued

Table A-3 (continued)

	N	*Mean*	*SD*	*Min*	*Max*
		Panel d) Other variables			
Log FDI	27,162	10.041	1.300	7.831	11.59
Log imports	26,491	6.891	2.108	0.000	11.12
Log exports	26,491	5.455	2.843	0.000	10.74
Distance to coast	27,162	0.317	0.388	0.001	2.455
Urbanization	27,162	0.072	0.150	0.000	1.000
Large mines	27,162	0.260	0.439	0.000	1.000
Oil fields	27,162	0.303	0.459	0.000	1.000
Travel time to cities	27,150	329	384	4.418	4,984
Road density	27,162	0.144	0.296	0.000	8.805
Africa	27,162	0.309	0.462	0.000	1.000
Asia	27,162	0.338	0.473	0.000	1.000
Americas	27,162	0.182	0.386	0.000	1.000
Low income	26,506	0.827	0.378	0.000	1.000
World Bank trans.	27,162	0.096	0.294	0.000	1.000
World Bank any	27,162	0.261	0.439	0.000	1.000
China social	27,162	0.030	0.172	0.000	1.000
China prod.	27,162	0.006	0.080	0.000	1.000
China energy	27,162	0.006	0.079	0.000	1.000
$\bar{N}^{\text{all}} > 0$	27,162	0.321	0.467	0.000	1.000

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

Table A-4 – Additional covariates, within first-order regions, no region FE, 2002–2013

	Controlling for ...				
	Travel time to cities (1)	Urbanization rate (2)	Road density (3)	Distance to coast (4)	Light per capita (5)
	Panel a) 2SLS estimates				
Projects ($\Delta N_{i,t-2}$)	-0.0269	-0.0284	-0.0276	-0.0263	-0.0265
	(0.0155)*	(0.0155)*	(0.0154)*	(0.0156)*	(0.0155)*
	[0.0139]*	[0.0139]**	[0.0138]**	[0.0137]*	[0.0138]*
Added variable	0.0016	-0.0034	-0.0033	0.0001	-0.0318
	(0.0005)***	(0.0010)***	(0.0014)**	(0.0001)	(0.0111)***
	[0.0006]***	[0.0018]*	[0.0018]*	[0.0001]	[0.0083]***
	Panel b) First-stage estimates				
IV ($F_{t-3} \times \bar{N}_{ji}$)	0.2430	0.2481	0.2461	0.2415	0.2412
	(0.0714)***	(0.0714)***	(0.0722)***	(0.0713)***	(0.0713)***
	[0.0811]***	[0.0807]***	[0.0817]***	[0.0813]***	[0.0812]***
Added variable	-0.0144	0.0425	0.0451	-0.0006	-0.0010
	(0.0053)***	(0.0107)***	(0.0190)**	(0.0004)	(0.0061)
	[0.0039]***	[0.0092]***	[0.0160]***	[0.0003]*	[0.0056]
First-stage F-stat	11.58	12.06	11.62	11.46	11.43
Observations	27,157	27,169	27,169	27,169	27,169
Regions	1,386	1,386	1,386	1,386	1,386
Countries	122	122	122	122	122

Notes: The table reports 2SLS regressions. 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 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. 'Travel time to cities' is measured as the travel time (in days) 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 (between 0 and 1) 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 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 (in 100s of km) 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). We use the inverse hyperbolic sine transformation, which is defined as $ihs(z) = log(z + \sqrt{z^2 + 1})$, for road density and light per capita to include regions with zero observations and retain an interpretation similar to logs. 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. *** p<0.01, ** p<0.05, * p<0.1.

Table A-5 – Non-linear trends and instrument, within first-order regions, 2002–2013

	Non-linear trends, i.e., year dummies times ...				
	Distance to coast	Urbanization rate	Large mines	Oil fields	All
	(1)	(2)	(3)	(4)	(5)
	Panel a) 2SLS estimates				
Projects ($\Delta N_{ji,t-2}$)	-0.0218	-0.0145	-0.0218	-0.0218	-0.0149
	(0.0102)**	(0.0107)	(0.0098)**	(0.0097)**	(0.0108)
	[0.0076]***	[0.0073]**	[0.0072]***	[0.0073]***	[0.0074]**
	Panel b) First-stage estimates				
IV ($F_{t-3} \times \bar{N}_{ji}$)	0.4393	0.4389	0.4400	0.4400	0.4390
	(0.0748)***	(0.0745)***	(0.0744)***	(0.0747)***	(0.0741)***
	[0.0716]***	[0.0718]***	[0.0714]***	[0.0715]***	[0.0716]***
First-stage F-Stat	34.53	34.75	34.99	34.68	35.07
Observations	27,162	27,162	27,162	27,162	27,162
Regions	2,406	2,406	2,406	2,406	2,406
Countries	122	122	122	122	122

Notes: The table reports 2SLS regressions. 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 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 to 5 add interactions of a cross-sectional variable (either a dummy for above-median value or a dummy that indicates the presence of something) with year dummies to allow general non-linear "China shocks" to affect some regions more than others. 'Distance to coast' is the average "as-the-crow-flies" distance from the region to the nearest coastline from Natural Earth. 'Urbanization rate' is 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). '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). 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. *** p<0.01, ** p<0.05, * p<0.1.

Table A-6 – Controlling for co-location, within first-order regions, 2002–2013

	Controlling for other projects committed at $t-2$				
	World Bank	World Bank	China	China	China
	(1)	(2)	(3)	(4)	(5)
	Panel a) 2SLS estimates				
Projects ($\Delta N_{ji,t-2}$)	-0.0215	-0.0214	-0.0211	-0.0218	-0.0219
	(0.0092)**	(0.0097)**	(0.0100)**	(0.0098)**	(0.0099)**
	[0.0071]***	[0.0073]***	[0.0074]***	[0.0072]***	[0.0073]***
	Panel b) First-stage estimates				
IV ($F_{t-3} \times \bar{N}_{ji}$)	0.4400	0.4396	0.4353	0.4400	0.4381
	(0.0747)***	(0.0746)***	(0.0759)***	(0.0747)***	(0.0753)***
	[0.0716]***	[0.0715]***	[0.0722]***	[0.0715]***	[0.0718]***
World Bank Transport	✓	–	–	–	–
World Bank Any	–	✓	–	–	–
China Social	–	–	✓	–	–
China Production	–	–	–	✓	–
China Energy	–	–	–	–	✓
First-stage F-Stat	34.68	34.70	32.86	34.65	33.81
Observations	27,162	27,162	27,162	27,162	27,162
Regions	2,406	2,406	2,406	2,406	2,406
Countries	122	122	122	122	122

Notes: The table reports 2SLS regressions. 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 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 and 2 control for World Bank projects using the World Bank Geocoded IBRD-IDA Projects (v1.4.2) dataset. Columns 3 to 5 include binary variables 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. *** p<0.01, ** p<0.05, * p<0.1.