

# The Political Geography of Cities\*

Richard Bluhm

Christian Lessmann

Paul Schaudt

September 2022

## Abstract

We study the link between subnational capital cities and urban development around the globe. To identify causal effects, we exploit new data on hundreds of first-order administrative and capital city reforms from 1987 until 2018 in an event study framework. We show that gaining subnational capital status has a sizable effect on city growth in the medium run and that these effects spill over to nearby cities. We provide new evidence that political importance complements locational fundamentals, such as greater internal market access, and is limited by economies of scale and the age of the urban network. Consistent with both an influx of public investments and a private response of individuals and firms, we document an increase of urban built-up, population, skilled migrants, and foreign investments in subnational capitals. Our results imply that the locational choice of capital cities affects agglomeration and that there are limits to how much administrative politics can shift economic activity and promote urban growth in the hinterland.

*Keywords:* Capital cities, administrative reforms, economic geography, urban primacy

*JEL Classification:* H10, R11, R12, O1

---

\*Bluhm: University of Stuttgart, Institute of Economics and Law, e-mail: [richard.bluhm@ivr.uni-stuttgart.de](mailto:richard.bluhm@ivr.uni-stuttgart.de); Lessmann: Technische Universität Dresden, Ifo Institute for Economic Research & CESifo Munich, e-mail: [christian.lessmann@tu-dresden.de](mailto:christian.lessmann@tu-dresden.de); Schaudt: University of Bern, Wyss Academy for Nature, e-mail: [paul.schaudt@unibe.ch](mailto:paul.schaudt@unibe.ch)). Markus Rosenbaum and Jonas Klärchen provided excellent research assistance. Richard Bluhm has benefited from financial support by the Alexander von Humboldt foundation and thanks the University of California San Diego for hospitality during parts of this research. We are grateful for helpful suggestions by Samuel Bazzi, Sascha Becker, Filipe Campante, Beatrix Eugster, Roland Hodler, Ruixue Jia, Michael Peters, Paul Raschky, Tobias Rommel, Mark Schelker, Kurt Schmidheiny, Claudia Steinwender, Michelle Torres, Daniel Treisman, Nick Tsivanidis, and David Weil, as well as comments from seminar participants at UC San Diego, University of St. Gallen, University of Bergen, University of Hannover, Monash University, The College of William & Mary, and participants at APSA, DENS, UEA, and the Harvard Cities and Development workshop.

# 1. Introduction

Subnational capitals are typically the largest city of their administrative region. For example, most capital cities in Brazil, Indonesia, and India are also the primary city and largest agglomeration of their region. However, the same is true in only about half the regions of Kenya, South Africa, or the United States. Similar variation exists among cities that have recently become capitals. Compare Raipur in India, which rose to become the primary city of the new state of Chhattisgarh (founded in 2000), to Kiambu in Kenya, which remained the third-ranked city of Kiambu county (founded in 2010). Why do we observe such differences in relative growth rates among cities that formally occupy the same position in the administrative hierarchy?

This paper studies how the change in the capital status of cities influences the relative growth rates of cities. We exploit the recent wave of decentralization reforms<sup>1</sup> in developing countries and, to a lesser extent, in developed countries as a quasi-experimental setting to estimate what we call the ‘capital city premium.’ Our analysis focuses on city growth in a global sample of cities. We compile data on hundreds of first-order administrative reforms that result in changes of the capital status of cities. Using these reforms and the varied contexts in which they occur, we investigate *i)* whether new capital cities increase density and attract more economic activity to a location and its surroundings, and *ii)* whether the political status of a city complements or substitutes for (a lack of) favorable locational fundamentals. The answers have important policy implications. They shed light on the circumstances under which capitals shift activity towards their locations and bring about local, or even regional, development.

To examine these questions, we compile comprehensive data on cities and data on whether administrative reforms treated them. Using Geographic Information Systems (GIS) and a plethora of more traditional sources, we first catalog all first-order subnational units and the location of their capitals over the period from 1987 until 2018. We then detect the boundaries of all cities with a population above 20,000 people in 1990 (and 2015) using data derived from high-resolution daytime images (in an approach similar to [Rozenfeld et al., 2011](#); [Baragwanath et al., 2019](#); [Eberle et al., 2020](#)) and assign these cities their time-varying capital status. We measure annual variation in economic activity and population density at the city level using nighttime light intensity and urban built-up (similar to [Storeygard, 2016](#); [Campante and Yanagizawa-Drott, 2017](#); [Henderson et al., 2018](#)). To capture how attractive particular locations are, we compile an array of geographic characteristics for the greater area inhabited by each city, ranging from agriculture over internal market access to the ease of external trade. This gives us a globally comparable sample of cities and their characteristics.

---

<sup>1</sup>[Grossman and Lewis \(2014\)](#) document a trend towards administrative unit proliferation in sub-Saharan Africa. Our data show that this pattern holds globally at the highest level of subnational government.

We analyze the short and medium-run effects of capital city reforms on city growth using event studies and difference-in-differences specifications. Our primary source of identifying variation is more than three decades of panel variation in the capital status of cities. While the choice to reform a particular region and promote or demote a city to a subnational capital is seldom random, we document that the timing of these reforms is usually unrelated with pre-reform characteristics of these cities and that unobserved confounders are likely to affect all cities in a reformed region similarly. This is aided by our focus on first-order capital cities. Their importance in the political hierarchy of a country implies that reforming them often requires constitutional changes and includes political considerations which are typically unrelated to local conditions at the city level (see, e.g, [Bai and Jia, 2021](#), or [Düben and Krause, 2021](#), on location choices in Imperial China). To strengthen this approach and minimize the scope for dynamic selection into treatment, we focus on the effect of gaining the status of a subnational capital for the first time, typically when regions are split. Our design allows us to separate the potentially endogenous decision to split a region, which we control for with fixed effects, from the treatment of designating a new regional capital (similar to [Bazzi and Gudgeon, 2021](#), who study district splitting and conflict). Testing for pre-trends suggests that the identifying assumptions hold in the sample of cities in reformed regions, the larger sample of all cities, and matched samples with comparable control cities. The pattern of leads and lags shows no anticipatory increases in activity but a substantial effect following the reform. It is inconsistent with unobserved shocks driving our results and robust to using estimation approaches of the event study model that allow for heterogeneous effects across cohorts of cities treated at different points in time. We also find no evidence suggesting that heterogeneous selection on observed fundamentals is driving our results. In other words, cities that become capitals but are in locations with better market access or in countries that are just urbanizing are not on a different growth path than other cities in the same initial region prior to the reform.

Our analysis establishes two main findings. First, we show that there are sizable advantages for cities that are elevated to subnational capitals. The advantages persist in the medium run and spill over to nearby cities. Economic activity, as measured by light intensity, in new capital cities rises by 15.7–25% five years after a territorial reform, depending on the specification. The event-study estimates show that these effects take about two years to materialize and then gradually increase until five years after the change in status. Both the larger agglomeration and surrounding cities benefit from being proximate to a city with capital status. Our analysis suggests positive but decreasing benefits for cities up to 100 km away from the new capital. We find no evidence of negative spillovers to non-capital cities. Second, we study the limits to how much politics can shift economic activity to certain locations via the administrative status of a city. We find that political status complements good locational fundamentals. Locating subnational capitals

in areas with better fundamentals, in particular, greater internal market access, has a larger effect on economic activity than locating them in areas suitable for agriculture. Moreover, we show that the capital premium depends on the intensive margin of political importance, that is, the size of the territory governed by the new capital city, and we find evidence suggesting that the premium is not uniform across the level of development but primarily occurs in countries which have started to agglomerate later. This suggests that politics exerts a more powerful force on the spatial equilibrium when the speed of urbanization is rapid and the urban network is not fully settled, but also that policymakers are constrained in how much they use administrative reforms to shift economic activity towards the hinterland. Additional results show that the capital premium dissipates only slowly after cities lose their status.

We examine three mechanisms. First, we observe significant increases in housing and infrastructure (urban built-up) and population density. Second, using microdata from recent migrants to cities that become new regional capitals, we provide direct within-city evidence of selective in-migration to new capitals. More educated individuals migrate to capital cities after they gain the status of a subnational capital. Third, we find evidence of increased private investments in finance and insurance, manufacturing, and other productive sectors in capital cities. Public investments are larger in capitals and appear to be concentrated in water and sanitation, infrastructure, and government. Hence, our findings are unlikely to be driven by increases in public employment in capital cities alone. Consistent with the pattern of private and public investments, we document that residents of capital cities accumulate more assets, have better access to electricity, higher educational attainment, and lower child mortality rates than residents of other comparably dense cities in the same region. This pattern is robust to excluding recent migrants to capital cities, except for educational attainment, where selective migration plays a larger role. Together with the pattern of spillovers, these results suggest that the investments associated with cities becoming capitals promote local and regional development more broadly.

Our finding that the political status of a city affects where economic activity takes place relates to a seminal literature in urban economics which has established that (national) capitals and primate cities are more populous in non-democracies than in democracies ([Ades and Glaeser, 1995](#); [Davis and Henderson, 2003](#)). We show that territorial decentralization in democracies and non-democracies alike can shift activity within an initial region and the entire country. To the best of our knowledge, our study is the first to document the contemporary effects of capital cities on the concentration of economic activity over a policy-relevant time frame in a global sample of cities. This differentiates our contribution from the nascent literature on subnational capitals. Isolated state capitals in the US, for example, are associated with more corruption, less accountability, and lower public good provision ([Campante and Do, 2014](#)), but we lack

broader evidence on how location interacts with urban growth.<sup>2</sup> In related studies, [Bai and Jia \(2021\)](#) and [Chambru et al. \(2021\)](#) study the effects of subnational capitals on population growth in Chinese prefectures from 1000 to 2000 CE and French municipalities after the French Revolution. They observe long-run increases in population, public goods, and connectivity, while the latter also document a lack of short-term growth in the new departmental capitals of France. Our focus is very different. Beyond establishing that capitals grow faster, we study why the capital premium is higher or absent in some cities and show that the characteristics of a location matter for whether administrative cities grow in the short and medium run. Moreover, both studies focus on highly centralized countries that have had a remarkable administrative capacity throughout their history, leaving open the question of whether this extends to developing countries today—many of whom are more decentralized but also have limited state capacity.<sup>3</sup>

Our finding that economic fundamentals, such as internal market access, play an important role for capital city growth supports a central premise of economic geography. Increasing returns to scale and path dependence can explain why we observe cities in places that do not seem to have favorable fundamentals today ([Krugman, 1991](#); [Davis and Weinstein, 2002](#); [Miguel and Roland, 2011](#); [Bleakley and Lin, 2012](#); [Michaels and Rauch, 2018](#)). We add that granting capital status to cities in locations with good fundamentals can facilitate agglomeration in more productive locations. However, the capital premium dissipates slowly and does not persist once the status is lost. Hence, choosing particular cities as capitals can improve welfare, provided that the relative importance of fundamentals shifts only slowly (see [Allen and Donaldson, 2020](#), on how temporary shocks can permanently affect aggregate welfare). A related literature is concerned with whether fundamentals, sorting, or learning drive the large productivity gains observed for workers in bigger cities (see, e.g., [Glaeser, 1999](#); [Combes et al., 2008](#); [Roca and Puga, 2016](#)). We provide new evidence that migrants with higher initial educational attainment sort into capital cities (after, not before, they change their status).

Our results are also relevant to debates about what policies can be used to address spatial inequality (see, e.g., [Lessmann, 2014](#); [Kessler et al., 2011](#); [Baum-Snow et al., 2020](#)). [Henderson et al. \(2018\)](#), for example, show that city locations in countries that have developed earlier tend to have been more influenced by agricultural characteristics and exhibit a more balanced distribution of economic activity than late agglomerators. We add a policy-relevant margin to this finding. Although countries that began to agglomerate late exhibit higher spatial concentration today, decentralizing their territorial structure can influence the spatial equilibrium. Our results suggest that a key constraint for policymakers is the size of the administrative regions. Moreover, our finding of

---

<sup>2</sup>[Campante et al. \(2019\)](#) add that isolation can also provide autocrats with shelter from insurrection, which is precisely why they sometimes relocate *national* capitals to the hinterland.

<sup>3</sup>In fact, in additional results, we show that the capital city premium is about twice as large in decentralized (federal) countries than in more centralized countries.

complementarity between politics and fundamentals speaks to the literature on place-based policies (Glaeser and Gottlieb, 2008; Kline, 2010; Neumark and Simpson, 2015). This literature has yielded mixed results empirically but highlights the potential (and challenges) of exploiting agglomeration economies through public policy.

Several aspects of this paper aim to move the current literature forward. First, we offer new global data on first-order administrative and capital city reforms. There is an extensive list of single country studies focusing on the diverse impacts of administrative reforms but sparse global evidence of this phenomenon. Second, leveraging large amounts of remotely-sensed data allows us to focus directly on cities, rather than administrative regions which change as a result of territorial reforms. Third, taking a global perspective enables us to ask a different set of questions over a shorter time frame. The heterogeneity in fundamentals and national contexts allows us to exploit variation that is usually unavailable within a single country.

The paper is organized as follows. [Section 2](#) presents the data on capital city reforms and describes the global sample of cities. [Section 3](#) discusses the empirical strategy. [Section 4](#) presents the results and discusses them. [Section 5](#) investigates treatment heterogeneity, [Section 6](#) household outcomes, and [Section 7](#) mechanisms. [Section 8](#) concludes.

## 2. Data

We start by describing our data, focusing on the construction of our main variables of interest. Other data sources are introduced later when they are used for the first time. A key constraint is that all data need to be available on a global scale, which is why we rely heavily on remotely-sensed data. This is not necessarily a disadvantage. Little to no data are available on the city level in developing countries and, even if more were available, it would be difficult to harmonize measurement across countries. Satellite-based measures are consistently defined for the entire globe and allow us to apply uniform definitions throughout. [Online Appendix A](#) provides a complete overview of the sources, variables, their coverage, and summary statistics.

### A. Capital city reforms

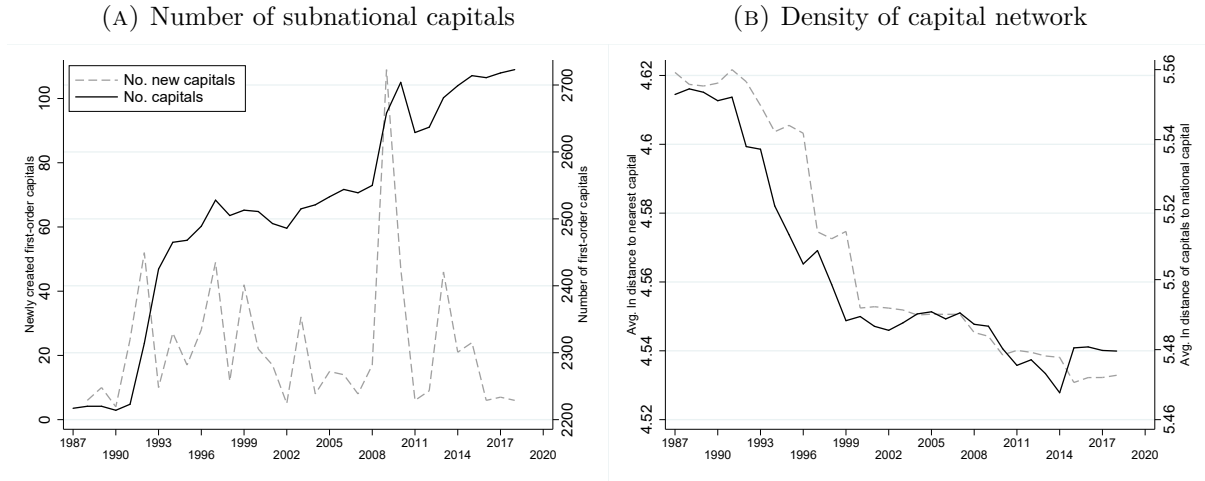
No off-the-shelf data systematically record administrative reforms, the boundaries of administrative units, and the location of capital cities across the world.<sup>4</sup> We compile

---

<sup>4</sup>Two sources come somewhat close: The Global Administrative Unit Layers (GAUL) project of the Food and Agriculture Organization (FAO) tracks the spatial evolution of administrative units between 1990 and 2014 across the world. Drawing on input from a variety of sources, the Statoids project collects (non-spatial) information on capital cities and administrative units (Law, 2010). Unfortunately, both data sets are riddled with errors and omissions, cover different time spans, and do not contain coordinates

new data containing the names and spatial extent of all first-order administrative units from 1987 until 2018, including the names and locations of capital cities over time. The data covers all types of territorial reforms, that is, splits and mergers of regions, area swaps, capital city re-locations and the creation of new countries. Creating this data is a two step process. First, we identify suitable vector data which accurately represents the boundaries of each unit within a country at a particular point in time. This involves a variety of sources (e.g., GAUL, GADM, Digital Chart of the World, United Nations Environment Program, and AidData’s GeoBoundaries project) and an algorithm which re-allocates small differences in boundaries to match those reported in the most accurate data sources. When no suitable data are available, we use international or national atlases, georeference and digitize the corresponding map. Second, we geocode all capital cities. [Online Appendix B](#) contains a detailed explanation of the data construction and provides summary statistics.

FIGURE I  
Subnational capitals: Global trends



*Notes:* The figure illustrates how global trends in territorial reforms change the number and network of first-order administrative capitals. Panel A illustrates the net number of capital cities over time and the number of cities which became capital cities in each year. Newly independent countries are included in the former but not in the latter. We omit the Sudan and South Sudan after their separation in 2011. Panel B plots the average log distance of cities to the nearest capital in gray and the average log distance to the national capital within countries in black.

Panel A of [Figure I](#) illustrates the variation in the number of capital cities over time. We observe a net increase of 506 capitals and new first-order units over the entire period from 1987 to 2018. Note that this understates the variation in our data, as some cities lose their capital status at the same time, some countries become independent over this

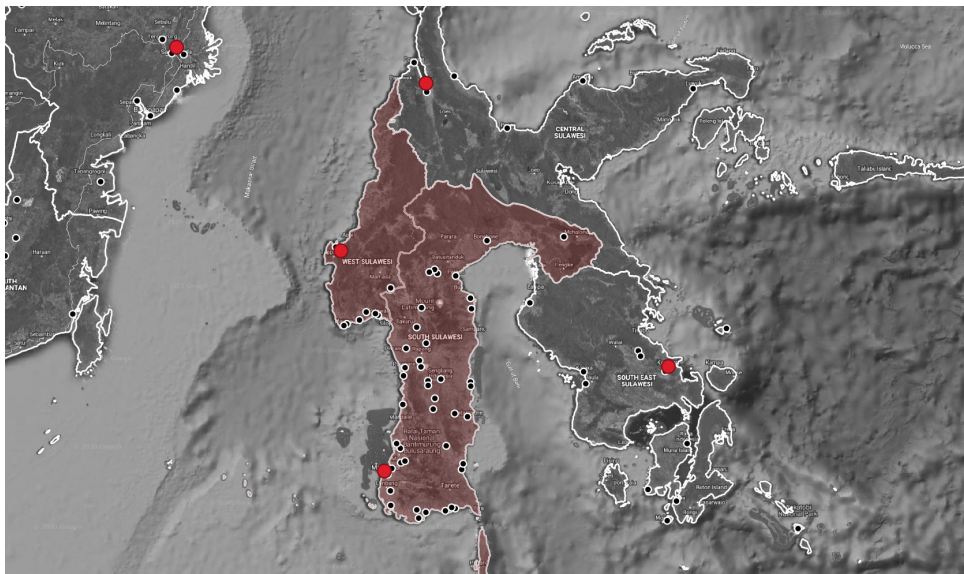
---

of capital cities. Other sources only document the most recent boundaries and contain no information about the relevant time-frame of these administrative borders or their capital cities (such as the Database of Global Administrative Areas, GADM).



period, and in a few rare occasions a capital city is simply moved within the same region.<sup>5</sup> In fact, when we track each city from when it enters our sample, we observe 701 cities which have gained capital city status and 336 cities who have lost this status over the same period. Panel A of [Figure I](#) also shows that a substantial number of new capitals has been created in every decade since 1987 (net of the creation of new countries). Panel B of [Figure I](#) highlights that new capitals are both intensifying and expanding the capital network over time, i.e., reducing the average distance between capitals and the national capital, and the distance of non-capitals to any subnational capital.

FIGURE II  
A provincial split: West and South Sulawesi in Indonesia



*Notes:* The figure illustrates the split of the former province of South Sulawesi into South and West Sulawesi in 2004. Post-2004 boundaries are indicated in white. The pre-reform area of the province is shaded in red. Red circles indicate capital cities. Black circles indicate other cities detected using our approach.

[Figure II](#) illustrates a typical provincial split, which is frequent in our data and will be the basis of our identification strategy. South Sulawesi (Sulawesi Selatan) was the fifth largest province of Indonesia with a population of about 8 million people in 2000. In 2004, West Sulawesi (Sulawesi Barat) was created out of the northwestern segment of the southern province. The new province had a population of little more than one million people and completed the partition of the island into north, south, east and west that was started in 1964. Makassar remained the capital of the south, while the city of Mamuju received the new status of a provincial capital.

---

<sup>5</sup>Our sample includes all countries which have a population of at least 1.5 million people, a land area of at least 22,500 km<sup>2</sup>, and have gained independence before 2000. Smaller states typically only have one administrative layer and are not well captured by our approach. To document that this is the case, we compiled time-varying administrative data on these countries as well.



Online Appendix C provides descriptive statistics about which (static) variables correlate with the probability that a particular city becomes a capital during a territorial reform.

## B. Urban boundaries and economic activity within cities

Our city-level approach requires us to identify the urban footprint of a host of potential control cities in addition to the administrative capitals. We follow a recent literature in urban economics which uses daytime images to accurately delineate city boundaries and nighttime light intensities as a proxy for economic activity within those boundaries (e.g. Baragwanath et al., 2019). Remotely-sensed city footprints diverge from administrative definitions in the sense that they tend to capture larger agglomerations which often run across several smaller cities. Using a globally consistent definition of cities is an important feature of our analysis.

We rely on two products from the Global Human Settlement Layer (GHSL)<sup>6</sup> derived from global moderate resolution (30 m) Landsat images and auxiliary data. The first is a built-up grid at a resolution of 1 km. It indicates the density of buildings and other human structure detected in the underlying high resolution data. The second is a population grid at the same resolution. It takes census estimates of the population at the smallest spatial scale available and distributes them using built-up intensities.<sup>7</sup> We use data from 1990 and 2015 to define the initial and final footprint of a city.<sup>8</sup>

Our definition of a city or an agglomeration follows a recent literature on urban boundaries by applying a city clustering algorithm (Rozenfeld et al., 2011; Baragwanath et al., 2019; Dijkstra et al., 2021). We consider a city to consist of a connected cluster of 1 km pixels with at least 50% built-up content per pixel or a minimum population density of 1,500 people per pixel (as in Dijkstra et al., 2021). Any cluster with an estimated population of at least 20,000 people is a city. While this is lower than the typically employed threshold of 50,000 people, it allows us to capture more secondary cities and towns in initially less urbanized developing countries. In fact, our data represents the global urban population quite well. Our data suggest an urban population of 2.59 billion in 1990 compared to the 2.27 billion reported by the World Bank. The difference becomes smaller still when we use the 2015 boundaries, with which we find 3.95 billion urban dwellers based on our data compared to 3.96 billion reported by the World Bank. We later document the robustness of our results to this parameter.

Our primary level of analysis is the universe of cities in 1990. Figure III shows

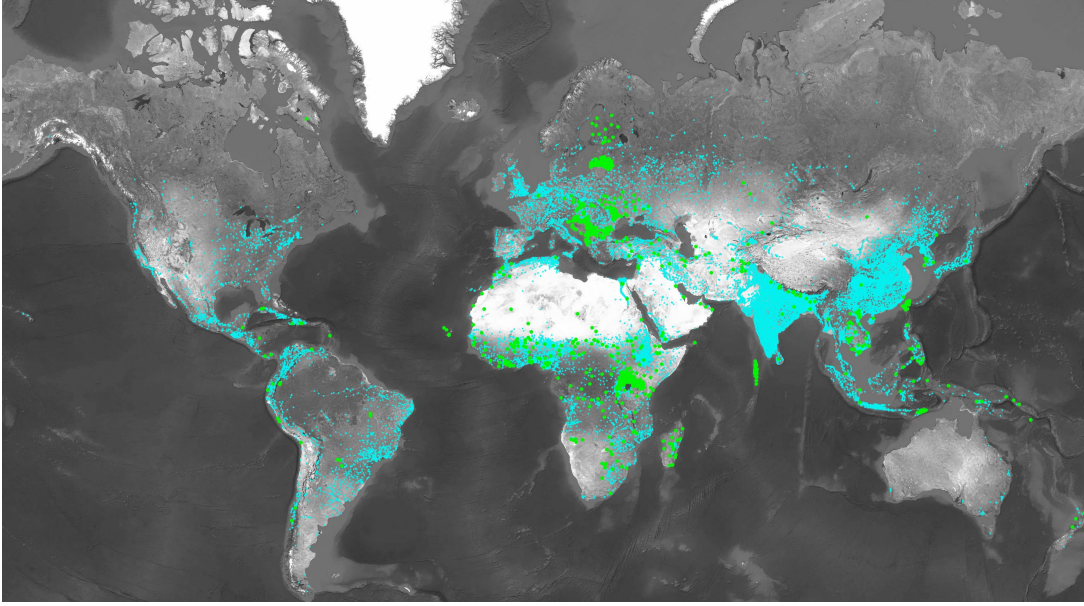
---

<sup>6</sup>The data is constructed by the Joint Research Centre and the Directorate General for Regional and Urban Policy of the European Commission. It can be accessed at <https://ghsl.jrc.ec.europa.eu>.

<sup>7</sup>Both products are available for 1975, 1990, 2000 and 2015.

<sup>8</sup>The GHSL project also provides a pre-classified layer of cities, the GHS settlement model, which is available for the same years. We do not use this layer in order to be able to control every parameter which defines a city, including the population threshold.

FIGURE III  
Locations of capital and non-capital cities in 1990



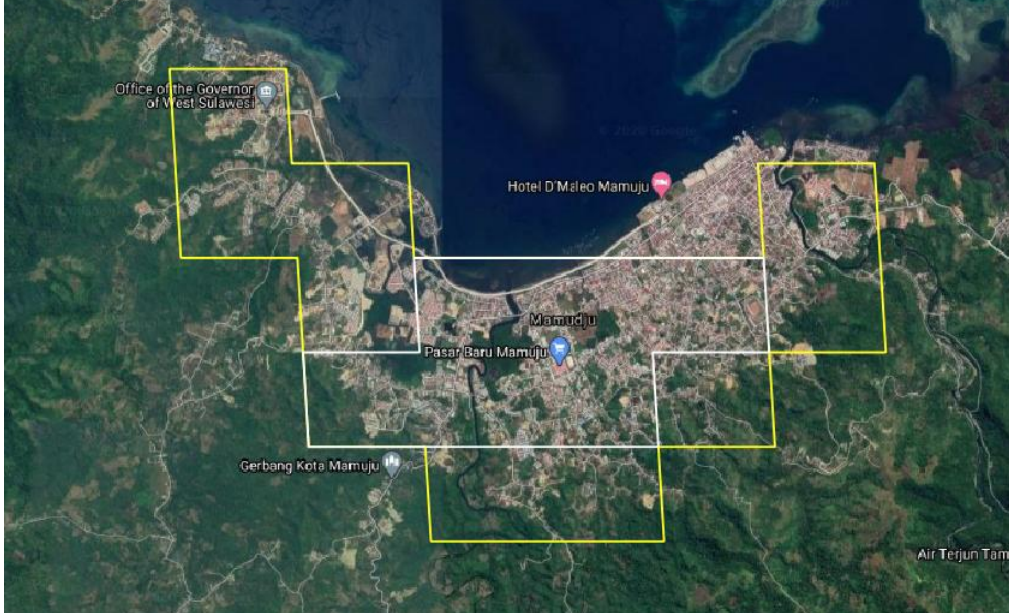
*Notes:* The figure shows the coordinates of 24,315 cities with a population above 20,000 people detected using the clustering algorithm described in the text. All cities are shown in blue. Cities elevated to capitals during the 1987-2019 period are highlighted in green.

the coordinates of about 24,000 cities detected in this manner. We also define larger *agglomerations* as the union of the initial and final boundaries, which will allow us to study overall growth later on. Naturally, we obtain fewer agglomerations than cities when joining the boundaries, as cities expand and merge into one over time. When studying agglomerations, we focus on new parts of a city forming around a 1990 city or cities which become amalgamated and ignore new cities detected only in 2015. The main reason is that the detection probability of a city increases (relative to non-capital cities) when it becomes a capital. We discuss this issue further in [Online Appendix D](#), where we show that gaining capital status over the period from 1990 to 2015 predicts inclusion in the 2015 sample (see [Table D-1](#)).

[Figure IV](#) illustrates this approach using the city of Mamuju, Indonesia. We observe a significant increase in the urban perimeter as the city grew from less than 50,000 people in 1990 to slightly more than 175,000 by 2015. The envelope here corresponds to the 2015 boundaries, as they fully contain the urban area in 1990. The early boundaries, on the other hand, give an accurate indication of the older core of the city.

Our primary outcome is the log of nighttime light intensity from the Defense Meteorological Satellite Program Operational Linescan System (DMSP-OLS). These data have been used in a variety of small scale and city level applications (e.g., [Storeygard, 2016](#); [Baum-Snow et al., 2020](#)) but suffer from sensor saturation in cities which severely understates economic activity in urban centers relative to rural areas ([Henderson et al.,](#)

FIGURE IV  
Urban footprint of Mamuju (Mamudju) in 1990 and 2015



*Notes:* The figure shows the urban area of Mamuju (or Mamudju) in Indonesia, as detected using the algorithms and data described in the text. The white boundaries delineate the 1990 footprint, while the yellow boundaries indicate the 2015 footprint (which coincides with the larger agglomeration). Slight differences in the coast line imply that one urban pixel is missing in both. The background shows a contemporary Google Maps image. Note that some of the urban areas with partial forest cover have a per pixel population density that easily crosses our threshold of 1,500 people. Google images are used as part of their “fair use” policy. All rights to the underlying maps belong to Google.

2018; Bluhm and Krause, 2022). For our main analysis, we use a version of this data which has been corrected for bottom coding<sup>9</sup> and top coding (see Bluhm and Krause, 2022, for details). Correcting for top-coding in the lights data ensures that light intensity is approximately linear in GDP and population density, even for smaller geographies in highly developed countries, such as the United States or Germany (Bluhm and Krause, 2022). We present results varying these adjustments later in the robustness section. We normalize light by the area of the city in 1990 to study increases in density (and henceforth refer to this measure as light density or light intensity).

Our preferred interpretation is that light intensity proxies for population density in the city.<sup>10</sup> We thus view our results in light of a large literature in urban economics which emphasizes the importance of city size and population density for productivity

<sup>9</sup>We use a simple adjustment to remove artificial variation at the bottom. The stable lights detection process carried out by the U.S. National Oceanic and Atmospheric Administration (NOAA) filters our background noise by effectively setting all clusters of pixels with a value of 3 or less equal to zero (Storeygard, 2016). Since we know that all light in our sample originates from a city, we undo this filtering by imposing a lower bound of 3 DN for each city pixel.

<sup>10</sup>Henderson et al. (2018) make a similar claim and use data on subnational regions to show that, conditional on country fixed effects, the  $R^2$  from a regression of lights on population density is 0.775, whereas it falls to 0.128 for income per capita. This correlation is just suggestive, given that local purchasing power parities are not available in most countries.

(see e.g. [Rosenthal and Strange, 2004](#); [Combes et al., 2010](#)). While density is not always synonymous with higher welfare, an emerging literature documents that city dwellers in developing countries are substantially better off than those living in the countryside.<sup>11</sup>

## C. Additional data

To capture how economic fundamentals vary with city locations, we compute a large set of geographic characteristics for a 25 km radius around the centroid of each agglomeration and assign these to the cities constituting the larger agglomeration. While the overwhelming majority occupy an area far smaller than this, the main advantage of focusing on such large areas is that we capture how well suited the area surrounding the city is for different economic activities.

We use three types of fundamentals describing how attractive a particular location is for agriculture, internal trade, or external trade. All of these are time-invariant. The set of agricultural characteristics consists of wheat suitability, temperature, precipitation, and elevation. External trade integration is measured by a set of distances: a dummy if the city is within 25 km of a natural harbor or the coast, and the continuous distance to the coast. Our measures of internal trade are dummies whether a city is within 25 km of a river or lake and a measure of market access in 1990. Market access of each city is defined by the sum of the cost of trading with every other city, the population of that other city and the market access of every other city to others in the same country. [Donaldson and Hornbeck \(2016\)](#), for example, show that such a measure summarizes the direct and indirect effects of changes in trade costs in general equilibrium trade theory.<sup>12</sup> Moreover, we use ruggedness ([Nunn and Puga, 2012](#)) and the estimated malaria burden ([Depetris-Chauvin and Weil, 2018](#)) to proxy for how hospitable a location is for human settlement.

---

<sup>11</sup>[Gollin et al. \(2021\)](#) and [Henderson et al. \(2020\)](#), for example, document that amenities in cities are higher than in the country-side (in addition to high urban wages, consumption and income). They find a positive density gradient in access to public goods and many other outcomes. Moreover, wages appear to rise faster with city size than the cost of living in several developing countries ([Chauvin et al., 2017](#)). [Henderson and Turner \(2020\)](#) argue that this relatively high utility of living in developing country cities could imply that urbanization is too slow.

<sup>12</sup>Since we are not interested in changes in trade costs elsewhere, we do not construct costs using along the actual road or rail network but use geographic distances to create a measure of the initial market access of each city at the start of the sample. Specifically, we define market access for each city  $c$  as  $MA_c = \sum_{d \neq c} pop_{1990} \times dist_{cd}^{-\theta}$  where we set the distance elasticity  $\theta$  to 1.4 following [Baragwanath et al. \(2019\)](#) and  $dist_{cd}$  is the geographic distance from city  $c$  to city  $d$ . We exclude each city  $c$  from the summation to focus only on its relationship to other cities. [Baragwanath et al. \(2019\)](#) find that a non-trivial proportion of market access in India is explained by cities that are close by.



### 3. Empirical strategy

Capital city reforms rarely occur in response to exogenous shocks, such as natural disasters.<sup>13</sup> In the absence of a randomized experiment on the location of subnational capitals, we will use observational data and leverage two aspects of the reform process: *i*) the timing of reforms is often idiosyncratic and, more importantly, *ii*), unobserved confounders are likely to affect all cities in reformed regions similarly. In other words, other cities in the region that will be reformed (i.e., split) were likely candidates to become capitals and were on similar growth trajectories before the reform took place.

#### A. Event-study design

Our base specification tests the role of capital cities in an event-study framework, where we exploit the switching of some cities into status of a subnational capital. We specify a standard event-study specification with an effect window running from  $\underline{j}$  to  $\bar{j}$  for all  $t = \underline{t}, \dots, \bar{t}$

$$\ln \text{LIGHTS}_{cit} = \sum_{j=\underline{j}}^{\bar{j}} \beta_j b_{cit}^j + \mu_c + \lambda_{(i,d)t} + \mathbf{z}'_c \gamma_t + e_{cit} \quad (1)$$

where  $\ln \text{LIGHTS}_{cit}$  is the log of light density in the urban cluster  $c$  in country  $i$  at time  $t$ ,  $b_{cit}^j$  are treatment change indicators, which indicate whether a city became a capital exactly  $j$  periods before or after period  $t$  (with bins at the endpoints)<sup>14</sup>,  $\mu_c$  are city fixed effects and  $\lambda_{(i,d)t}$  are country-year or initial-region times year fixed effects,  $\mathbf{z}_c$  are time-invariant fundamentals and  $\gamma_t$  are time-varying coefficients on the fundamentals. We omit  $b_{cit}^{-1}$  so that all effects are estimated relative to the last pre-treatment period.

Our combination of city and country-year fixed effects implies that we essentially stack many individual country-level event studies. In this setting,  $\lambda_{it}$  nets out all country-wide variation in a specific year. This does not just include business cycle variation but also the national level decision to reform the territorial structure in more than one region at the same time. For most of our specifications, we go one step further and define  $\lambda_{(i,d)t} \equiv \lambda_{dt}$  as initial-region times year fixed effects. Together with our focus on cities which gain the status as a capital, this structure implies a well-defined identification strategy. We

<sup>13</sup>We do observe an instance where the capital city was moved from Rabaul to Kokopo in Papua New Guinea's East New Britain province following the destruction of the former by a volcanic eruption.

<sup>14</sup>We summarize the dynamic treatment effects prior to period  $\underline{j} + 1$  and beyond period  $\bar{j} - 1$  in one estimate for each. More formally, following the notation from Schmidheiny and Siegloch (2019), we define

$$b_{cit}^j = \begin{cases} \sum_{s=\underline{t}-\underline{j}}^{\bar{t}-\underline{j}-1} d_{cis} & \text{for } j = \underline{j} \\ d_{ci,t-\underline{j}} & \text{for } \underline{j} < j < \bar{j} \\ \sum_{s=\underline{t}-\bar{j}+1}^{t-\bar{j}} d_{cis} & \text{for } j = \bar{j} \end{cases}$$

where  $d_{cit}$  is a treatment change indicator.

compare cities who gain the status after an administrative region is partitioned to all other non-capitals in the initial region (i.e., we compare the new capital, Mamuju, to all other non-capital cities in the region highlighted in [Figure II](#)). This allows us to separate the treatment effect of the regional split on all cities from the treatment effect of becoming a capital. Shocks that affect all cities in the initial region within a particular year, such as the decision to reform the territorial structure or common trends, are absorbed. The influence of the fundamentals in the baseline period is absorbed by the city fixed effects. However, allowing time-varying coefficients on the fundamentals accounts for a variety of meaningful patterns, such as a shift towards local density and/or market potential as transport costs fall (as in [Brühlhart et al., 2020](#)).

The event-study design allows us to test for pre-trends and study the dynamics of the estimated treatment effect. When testing for pre-trends, we set  $\underline{j} = -5$  to  $\bar{j} = 5$  for a symmetric window around the treatment date. We rely primarily on visual evidence of the underlying specifications, where we report confidence intervals (clustered on initial regions) together with simultaneous confidence bands (which have the correct coverage probabilities for the entire parameter vector at 95%). We construct sup- $t$  bootstrap confidence bands with block sampling over initial regions to mirror the dependency structure of the errors ([Montiel Olea and Plagborg-Møller, 2019](#)).

For most of the extensions and robustness checks, we collapse the event study to a difference-in-differences specification but keep all other aspects of the design the same.

## B. Identifying variation and identification assumptions

The key identification assumption in these types of event-study designs is that light intensity in cities and the change in capital status are not both driven by some time-varying unobserved factor which affects treatment and control cities differently. We take two steps to make sure this assumption is credible. First, we restrict the estimation sample to the set of cities located within initial regions that are split only once to obtain a comparable control group. Second, we discuss and analyze the timing of events and test for pre-trends in the outcome variable.

**Treatment and control groups:** Our data on subnational capitals and first-order administrative regions contains a wide variety of reforms (splits, mergers, re-locations and wholesale changes in the administrative territorial structure). A potential concern could be that these treatments are very different, in that they imply different pre-treatment trends and subsequent treatment effects, conditional on the chosen control group. For example, losing the status as a regional capital in a merger could be associated with a secular decline in the importance of the city, resulting in pre-treatment trends. Moreover, our data and time frame are not well suited to deal with negative shocks to durable



housing. We exploit the strengths of our setting and focus on the effect of gaining the status as a subnational capital. Hence, we limit the sample to cases where an initial region is split, such that one or several new capitals are created. We defer the issue of capital loss to [Online Appendix F](#), in which we discuss the appropriate comparison groups for different treatments, the effects of capital loss, as well as related issues, such as “mother” capitals.<sup>15</sup>

TABLE I  
Identifying variation

	All admin cities	Matched to urban clusters in 1990	Clusters in 1990 with single changes
<i>Panel A. Event-study period, 1987 – 2018</i>			
Always capitals	2,118	1,721	–
Gained status	701	329	277
Lost status	336	168	117
<i>Panel B. Diff-in-diff period, 1992 – 2013</i>			
Always capitals	2,211	1,798	–
Gained status	592	263	217
Lost status	275	123	86

*Notes:* The table shows summary statistics of the capital cities and urban clusters data. The capital cities data in column 1 covers all administrative centers, no matter if the city footprint is detected by the city clustering algorithm or not. The urban clusters data in column 2 shows how many of these capital cities have been matched to cities which pass the detection thresholds of the city clustering algorithm. Column 3 shows the subset of these which experienced a single reform.

[Table I](#) illustrates the capital city reforms we observe in our data and the subset we use for identification. For the event study, we obtain data on the treatment status from  $\underline{t} = 1987$  to  $\bar{t} = 2018$ . We later collapse the estimates into a difference-in-differences design, for which we only use information on cities that switch their status during the period from  $\underline{t} = 1992$  until  $\bar{t} = 2013$ . The first column shows the total number of cities and their changes in status, no matter if we actually observe them in the satellite-derived data on city footprints or not. The second column indicates how many urban clusters derived from the 1990 satellite data were always capitals or experienced a change in status.<sup>16</sup> While we observe a large share of administrative cities in 1990, not all of them pass the population threshold of 20,000 in 1990 and many close-by capitals are matched to the same cluster. Several administrative cities in developing countries which are heavily

<sup>15</sup>Our identification strategy is not well suited to deal with multiple treatments. There are several instances in our data (about 9% of the ever treated cities) where a capital was moved or a new administrative region was created sometime in the 1990s, followed by another reform in the 2000s. We discard all multiple treatments and focus only on instances where a city received the status as a subnational capital only once during the period of interest.

<sup>16</sup>We match administrative cities to an urban cluster if the centroid of the administrative city is within 3 km of the urban cluster or the names are identical. Note that some clusters contain several administrative cities so that the fraction of matched cities is somewhat higher than implied by the table.

decentralized by the end of the period, such as Uganda, are initially too small. We prefer to focus on the 1990 universe of cities, as this avoids selection problems by which cities pass the detection threshold in later years precisely because they became a subnational capital (discussed in [Online Appendix D](#)). The last column highlights the switches which we effectively use for identification. The event-study design uses 277 cities which become capitals of which 217 switchers are observed during the 1992-2013 period for which we observe our outcomes.

We typically compute our results for two samples: *i*) all cities and *ii*) cities in regions which have been reformed within the period of observation. If we are concerned with potential spillovers and “forbidden comparisons”, where treated units end up being used as controls due to the staggered design (see, e.g., [Borusyak et al., 2021](#)), then we would prefer a large control group since this reduces the weights of these comparisons in a staggered difference-in-differences setting. If we are concerned about obtaining a control group that closely resembles the treatment group, then we would prefer to restrict ourselves to places that are in close proximity. Given that our control group is more than an order magnitude larger than the treatment group in either sample (mitigating the first set of concerns), we have a preference for the latter approach but report both for completeness.

The size of the never treated control group can mitigate biases in staggered difference-in-differences designs ([Borusyak et al., 2021](#)) but is no panacea in event studies. [Sun and Abraham \(2021\)](#) show that the coefficients obtained from the two-way fixed effects estimator of the event study specification in [eq. 1](#) can be contaminated by information from other periods. This occurs when there is heterogeneity in treatment effects across treatment cohorts and this contamination affects leads, lags, and endpoint bins. [Sun and Abraham \(2021\)](#) and [Callaway and Sant’Anna \(2021\)](#) present alternative (and related) estimators that do not suffer from this problem. We compare how fixed-effects estimation performs relative to alternatives in our setting in robustness checks.

**Timing of reforms:** Capital city reforms occur for a variety of circumstances and policy-makers may pursue a range of political and economic objectives (e.g. granting regional autonomy, avoiding conflict, improving service delivery, and more). Our focus on first-order units implies that these reforms are seldom carried out without the influence of national politics. This helps identification in our context, as it makes the timing of reforms less predictable and, therefore, pre-trends at the city level less likely.

Anecdotal evidence supports this conjecture. The 2010 restructuring of Kenya’s provinces illustrates this well. A constitutional reform process was started following the post election violence in 2008. A key objective of this process was to reduce ethnic tensions in the country which was, at least in part, to be achieved by a devolution of power and

territorial reform.<sup>17</sup> Up on to this point, Kenya was organized into eight large provinces. The first attempt at constitutional reform had failed in 2005 and even in January 2010 “it appeared that the political disputes which had undermined previous attempts at constitutional reform were likely to resurface” (Kramon and Posner, 2011, p. 93). There were lengthy debates about how many tiers and counties the new administrative structure should have, which were finally settled when the parliamentary committee “agreed to the least controversial position: a two-tier system with 47 county governments whose boundaries would be congruent to the country’s pre-1992 regions” in April 2010 (Kramon and Posner, 2011, p. 94). The new constitution was adopted by national referendum in August 2010, leaving little scope for anticipation effects.

Even when splitting of regions is driven by local demands, such as in neighboring Uganda, the national parliament is usually involved in approving them, so that the timing of splits becomes difficult to predict. Uganda decentralized its administrative structure from 34 regions in 1990 to 127 by 2018. The reforms were carried out in several waves. While most splits were eventually approved, some were denied by the parliament (Grossman and Lewis, 2014). National involvement in these types of reforms is not limited to Africa. Indonesia created eight new provinces and more than 150 new second-tier regions after the fall of Suharto in 1998. Splitting required parliamentary and presidential approval. India’s national parliament created three new states in 2000. There were local movements in favor of these states for cultural and economic considerations, but previous attempts to carve out new territories had failed repeatedly before their final adoption (Agarwal, 2017).

Although random timing of the reforms is appealing, it is not necessary for identification and likely to be violated in several settings.<sup>18</sup> The parallel trends assumption needed for our strategy to work is substantially weaker. On top of static selection, it allows for time-varying omitted variables to affect the treatment and control group, provided that these two are affected equally. We consider this assumption particularly plausible in the sample of cities in reformed regions with initial-region-by-year fixed effects, as all cities in those regions are indirectly affected by the same territorial reform. To test whether selection is heterogeneous in terms of locational fundamentals and other observed factors, we interact the entire event-time path with these variables.

---

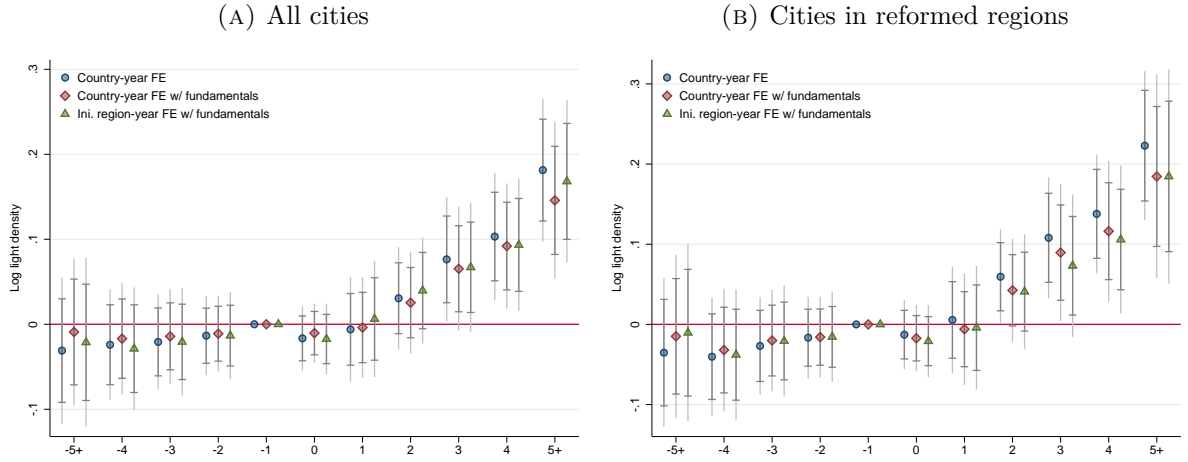
<sup>17</sup>See Bluhm et al. (2021) for a study of the effects of this reform on ethnic voting.

<sup>18</sup>Identification is straightforward if the timing of the intervention is exogenous to city level characteristics (conditional on the fixed effects and observed covariates). If the pre-reform time indexes can be swapped, there cannot be any pre-trends. Figure E-1 in Online Appendix E shows that the timing of capital city reforms is difficult to predict, at least with time-invariant initial city characteristics and especially once we focus only on within country variation.

## 4. Results

**Baseline results:** Figure V reports the results from our main event-study specification based on two different samples. Panel A plots estimates based on a specification using all cities and country-year FEs (circles). This is our baseline estimate for the larger sample where the control group consists of all other (non-capital) cities in the same country. The diamonds report results for a specification that purges the time-varying effects of the fundamentals, and the triangles show estimates obtained by adding initial-region-by-year fixed effects. The latter are our preferred estimates since the control group now only consists of non-capital cities within the same initial region. Panel B repeats this set-up for the sample of cities in reformed regions.

FIGURE V  
Capital cities and light intensity: Event study



*Notes:* The figure illustrates event-study results from fixed effects regressions of the log of light intensity per square kilometer on the binned sequence of treatment change dummies defined in the text. Circles represent point estimates from a regression with city and country-year fixed effects, diamonds represent specifications with additional controls for locational fundamental, and triangles represent specifications with initial-region-by-year fixed effects in addition. Panel A shows estimates for all cities. Panel B shows estimates for cities in reformed regions. All regressions include city fixed effects. 95% confidence intervals based on standard errors clustered on initial regions are provided by the gray error bars. The orange error bars indicate 95% sup- $t$  bootstrap confidence bands with block sampling over initial regions (Montiel Olea and Plagborg-Møller, 2019).

The estimates and their confidence bands strongly support the notion that gaining the status as a capital is exogenous to pre-reform changes in the economic activity of treated cities. The pre-trends are essentially flat. There are no systematic differences in city light intensities prior to a change in capital status. Both the pointwise confidence intervals and the sup- $t$  bands rule out a wide range of positive anticipation effects. We view this as strong evidence for the validity of our identification strategy. Any unobserved confounding factor would have to very closely mimic the timing implied by this observed pattern. Moreover, since both the full sample and the sample of cities in reformed

regions reveal a very similar pattern, and we find no evidence of pre-trends in any of the specifications we examine, we consider it unlikely that pre-testing bias is a major concern in our application (Roth, 2021).<sup>19</sup>

In all dynamic specifications, we do not detect a spike in activity in the first year. This is intuitive in the sense that constructing new buildings, an influx of public and private investment, moving an administration, or a migratory response all take time. More importantly, the medium-run estimates across both panels and all specifications suggest that the light intensity of capital cities is 15.7 to 25.0% higher five years after the gain in status. The estimates are slightly larger in the sample of cities in reformed regions (panel B). Here light intensity begins to increase by approximately 4.2 to 6.1% two years after a city becomes a subnational capital and settles within a narrower band of 20.3 to 25% five years after the reform and beyond.

We investigate how the size of the event-study window influences the estimate of the medium-run effect (endpoints) and compare it to the difference-in-differences version in Online Appendix E. Consistent with the gradual rise of the estimated effects, we find that the estimate of the medium-run effect is closer to 20% for longer event windows (estimated on fewer treatments) but its confidence interval always contains the difference-in-differences estimate (see Figure E-2). The difference-in-differences is around 9.2 to 15.4%, depending on the specification, and utilizes only switchers during the period from 1992 to 2013 (see Table E-1). These estimates are lower than the medium-run estimates from the corresponding event studies, as they average over the first years and all subsequent post-treatment periods.

In Online Appendix E, we also explore whether these results are affected by the estimation method. Figure E-2 compares the estimates using TWFE to those obtained via the interaction-weighted (IW) estimator (Sun and Abraham, 2021), which correctly identifies cohort-average treatment effects in the presence of heterogeneous treatment effects.<sup>20</sup> No matter if we focus on a five-year event window or a 15-year event

---

<sup>19</sup>Using the approach from Roth (2021) to calculate the power of detecting specific linear pre-trends in our setting, we find that the slope of a linear pre-trend which we would detect only half the time is about 0.0119 in the specification with country-year fixed effects in panel A of Figure V and about 0.0149 in the preferred specification with initial-region-by-year fixed effects in panel B. This leads to a mean after pre-testing in period 5 and beyond of 0.0720 and 0.0895, respectively. Power rises quickly against more severe pre-trends. A linear pre-trend which we would expect to detect in 80% of cases has a slope of 0.0183 and 0.0230, implying a conditional expectation after pre-testing in period 5 and beyond of 0.111 and 0.1379, respectively. Hence, in our setting, it seems unlikely that strong linear pre-trends explaining about half our effect size in the final period would have gone undetected. Moreover, we have more than 97% power to detect linear pre-trends that have a conditional expectation after pre-testing which passes through the final period estimate in either specification.

<sup>20</sup>The IW estimator is constructed in three steps. First, we estimate an event study specification where the relative time dummies are interacted with cohort (treatment year) dummies using TWFE. Second, we estimate the probability of first being treated in a particular year as the sample shares of each cohort in the relevant period(s). Finally, we form the IW estimator by taking averages of the estimates by cohort weighted by their probability of treatment. Following Sun and Abraham (2021), we balance the sample in calendar time, which is why the TWFE results differ marginally from Figure V.

window to mitigate the influence of binning, we find that TWFE and IW estimation deliver quantitatively and qualitatively similar results. The estimates of leads, lags, and endpoints bins obtained via the two estimators are typically close and well within each others confidence intervals. We continue to find no indication of pre-trends,<sup>21</sup> a rise from the second or third period onward, and effect sizes that begin to level off to their medium-run value sometime around the fifth post-treatment period.

In summary, our main results suggest that cities which become capitals grow substantially faster than their peers in the subsequent period. There is a build up in economic activity during the first five years after which we observe a medium-term increase around 20%, with some variation across specifications. This is up to half of the within city standard deviation in light intensity. To put this in perspective, consider the results in [Storeygard \(2016\)](#) where an African city which is further away from the primate city than the median city loses about 12% of its economic activity when the oil price is high.<sup>22</sup> For the remainder of the paper, we report difference-in-differences estimates in the main text and relegate the corresponding event-study plots to [Online Appendix E](#).

Unlike the economic advantages documented in [Bleakley and Lin \(2012\)](#), we find evidence that this political premium does not persist. [Online Appendix F](#) provides a detailed analysis of cities that lose their status as a regional capital (usually during a merger of first-order regions). Such an analysis necessitates a somewhat different design, as the relevant comparison group now consists of cities that remain capitals over the entire period. Our results suggest that former capitals lose economic activity, roughly mirroring the initial increase in light intensity, starting around four years after losing their elevated political status.

**Agglomerations and city peripheries:** Urban expansion is an important component of city growth and a function of geography, policies, and the economic structure of the wider area surrounding a city ([Burchfield et al., 2006](#)). Most cities grew substantially at the extensive margin over the period from 1990 to 2015. The average city expanded its area by almost 50% and the area of capital cities grew faster than that of other cities in the same initial region.<sup>23</sup> Unfortunately, we do not observe a city’s urban extent in every year so that we cannot calculate detailed measures of urban expansion. Instead, our baseline results focus on the universe of cities detected in 1990 and treat their urban extent as

---

<sup>21</sup>Only one pre-treatment coefficient (out of 15) in the long event window is significant using interval estimation. We do not compute bootstrap based confidence bands due to computational speed but note that these would be wider than the point-intervals.

<sup>22</sup>The effect is also larger than the effect of funneling public funds to specific regions documented in the literature on political favoritism (although the level of analysis is not the same). [Hodler and Raschky \(2014\)](#) estimate that being the birth-region of a national leader increases nighttime light intensity by about 3.9%, while [De Luca et al. \(2018\)](#) estimate an increase of 7% to 10% in the ethnic homeland of a leader who is currently in office.

<sup>23</sup>[Table E-4](#) in [Online Appendix E](#) shows that capital cities expanded their average footprint by about 9.9% to 14.2% more than non-capital cities over the period from 1990 to 2015.



fixed (to represent ‘the core’). This avoids potential endogeneity issues in the selection of cities and allows us to focus on increases in density but comes at the cost of neglecting initially less densely developed areas of cities. In [Table II](#) we loosen this assumption by accounting for cities that ultimately merge into a single larger agglomeration and include areas which were initially in the periphery but became part of a single agglomeration by 2015. This allows us to study changes in the light intensity of the overall agglomeration and changes outside of the 1990s core of each city.

TABLE II  
Larger agglomerations and city fringe: Difference-in-differences

	<i>Dependent Variable: ln LIGHTS<sub>cit</sub></i>					
	All Cities			Reformed Regions		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A. Growth of the larger agglomeration</i>						
CAPITAL	0.0818 (0.0285)	0.0738 (0.0279)	0.0821 (0.0282)	0.1093 (0.0300)	0.0922 (0.0292)	0.0898 (0.0286)
<i>Panel B. Growth in the periphery of the city</i>						
CAPITAL	0.0787 (0.0353)	0.0624 (0.0341)	0.0862 (0.0342)	0.1035 (0.0365)	0.0797 (0.0362)	0.0935 (0.0348)
$N$	13373	13373	13373	4466	4466	4466
$N \times \bar{T}$	274408	274408	274408	86455	86455	86455
Fundamentals	–	✓	✓	–	✓	✓
Agglomeration FE	✓	✓	✓	✓	✓	✓
Country-Year FE	✓	✓	–	✓	✓	–
Ini. Region-Year FE	–	–	✓	–	–	✓

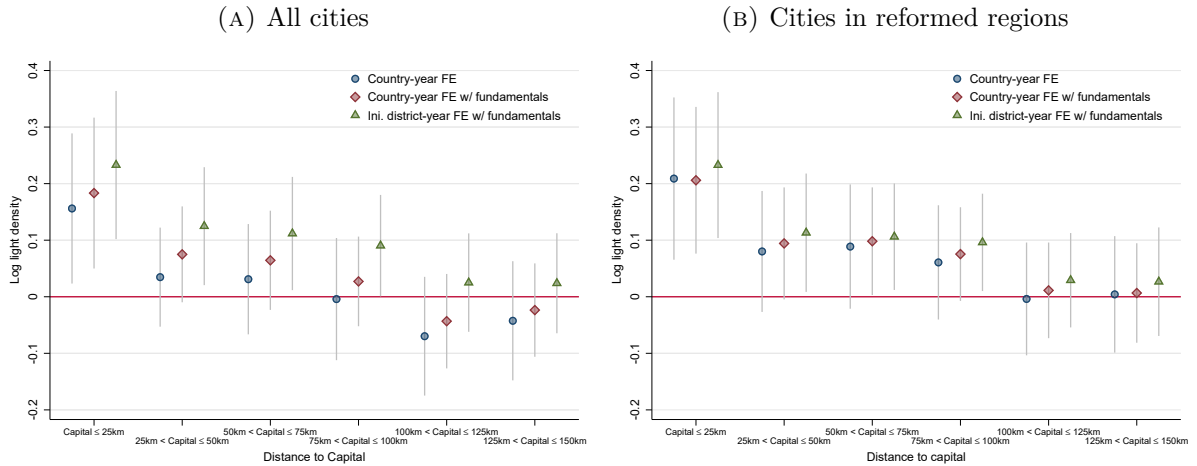
*Notes:* The table reports results from fixed effects regressions of the log of light intensity per square kilometer on capital city status. Panel A reports results based on the larger agglomeration (the envelope over 1990 and 2015 of the urban clusters detected in 1990). Panel B reports the results for the fringe (areas the urban clusters detected in 1990 that meet the detection threshold by 2015). Standard errors clustered on initial regions are provided in parentheses.

Capital cities experience faster overall growth than non-capital cities and their peripheries grow faster as well. Panel A of [Table II](#) reports our six specifications at the level of agglomerations, that is, cities detected in 1990 including the parts that only pass our density threshold of 1,500 people or 50% built-up per sq. km by 2015. The estimates tend to be smaller than our difference-in-differences estimates by about 2–4 percentage points but are otherwise similar. Panel B reports the same set of specifications but focuses on light growth in the periphery, that is, only the area of each agglomeration that is initially less dense and subsequently passes the population threshold. We find that new developments around capital cities are growing at a pace comparable to the larger agglomeration but somewhat slower than the core. The results are statistically significant at conventional levels in all columns, apart from column 2 where the effect is

less precisely estimated but within a standard error of other estimates. Taken together, this strongly suggests that both increasing density in the center and urban expansion are associated with gaining the status as a capital city.

**Spillovers to nearby cities and SUTVA violations:** An important question in our context is whether new subnational cities draw economic activity from their immediate surroundings or whether creating capitals benefits more cities in a region. Moreover, the presence of any such negative or positive spillovers would violate the stable unit treatment value assumption (SUTVA). This assumption requires that the treatment status of any one unit (capital cities) does not affect the treatment status of other units (non-capital cities). Our preferred specification is vulnerable to this problem as it compares the status of cities that are—by virtue of being located in the the same initial region—relatively close-by. We can view this as omitted variables problem. If there are positive spillovers to nearby cities, then our baseline results are attenuated and *vice versa*. Provided that spatial spillovers have a monotonic pattern in distance and some cities are unaffected, it suffices to include dummies capturing the proximity to treated cities and their change in treatment status (see e.g., [Asher et al., 2019](#), for a similar approach).

FIGURE VI  
Spillovers to nearby agglomerations: Difference-in-differences



*Notes:* The figure reports results from fixed effects regressions of the log of light intensity per square kilometer on capital city status as well as several spillover dummies for different distance intervals (panel A for all agglomerations, panel B for agglomerations within reformed areas). Circles represent point estimates from a regression with city and country-year fixed effects, diamonds represent specifications with additional controls for locational fundamental, and triangles represent specifications with initial-region-by-year fixed effects in addition. Panel A shows estimates for all cities. Panel B shows estimates for cities in reformed regions. All regressions include city fixed effects. 95% confidence intervals based on standard clustered on initial regions are provided by the gray error bars.

Figure VI explores this possibility by adding indicators for the (time-varying) distance

to the nearest capital city in the country, where each indicator captures agglomerations in a 25 km ring around the treated agglomeration, starting from 0 km and going up to 150 km.<sup>24</sup> A considerable advantage of this specification over our baseline results is that it allows us to account for capital cities which we did not match to an urban cluster in 1990. Even if the urban extent of some capitals is not observed, the distance of all other urban clusters to these “unobserved” capitals with known coordinates is straightforward to compute, so that we indirectly capture the entire universe of capital cities, including *all* changes in status of nearby cities. The spillovers are identified by cities switching between rings as they move closer to (or further away from) capital cities.<sup>25</sup>

The country-year fixed effects specification (circles in panel A) in [Figure VI](#) provides the least evidence of spillovers. It uses all non-capital agglomerations in the same country as a control, many of which are located in regions far apart from treated agglomerations. The evidence in favor of spillovers becomes stronger once we introduce initial-region-by-year fixed effects (triangles) or limited the sample to cities in reformed regions, as in panel B. Depending on the specification, we find evidence of positive spillovers affecting agglomerations which are up to 75 to 100 km away from a new capital. All specifications suggest a declining pattern of positive treatment effects, where satellite towns close to the new capital grow substantially faster but this effect disappears after a distance of 75 to 100 km. Accounting for these indirect effects increases the estimate for the capital itself, particularly in the sample of cities in reformed regions. We now estimate a treatment effect between 23.2% and 26.2%. This spatial pattern has an important policy implication. Rather just drawing activity and population from its immediate surroundings, creating new capital cities appears to benefit more cities in the reformed region.

**Alternative control groups:** A potential concern is that our baseline results include a variety of cities in the control group, many of which are unlikely to ever become subnational capitals. In fact, future capitals are usually among the biggest and brightest cities in the pre-reform region. Out of the 221 cities which became capitals during the period from 1992 to 2013, the median city was ranked second in terms of its 1990 population in the initial region, while the city at 90th percentile was ranked 10th.<sup>26</sup> Static selection is not a concern in difference-in-differences designs. However, in spite of finding no evidence in favor of pre-trends, our baseline results could still include cities in the control group that are on fundamentally different growth paths than cities which later become subnational capitals.

We use a simple form of nearest-neighbor matching to assess whether the definition of the control group influences our results. We rank all cities in the initial region according

---

<sup>24</sup>We use 150 km as an upper bound for spillovers, but the results are not sensitive to this choice.

<sup>25</sup>Spillovers are identified through proximity to the closest administrative capital. These distances change both when nearby cities become capitals and when they lose their elevated status.

<sup>26</sup>See the discussion of correlates of capital locations in [Online Appendix C](#).

TABLE III  
Different control groups: Matched difference-in-differences

	<i>Dependent Variable: ln LIGHTS<sub>cit</sub></i>					
	Light intensity in 1992			Population in 1990		
	Distance from treated city					
	Any (1)	> 50 km (2)	> 75 km (3)	Any (4)	> 50 km (5)	> 75 km (6)
<i>Panel A. Control cities within ± 2 ranks of treated cities in initial region</i>						
CAPITAL	0.1154 (0.0276)	0.0790 (0.0296)	0.0911 (0.0330)	0.1033 (0.0284)	0.0947 (0.0318)	0.0913 (0.0351)
F-test pre-trends ( <i>p</i> -val.)	0.186	0.447	0.470	0.364	0.190	0.113
<i>N</i>	625	413	354	618	433	370
<i>N</i> × $\bar{T}$	13578	8963	7694	13437	9403	8038
<i>Panel B. Control cities within ± 3 ranks of treated cities in initial region</i>						
CAPITAL	0.1241 (0.0277)	0.0906 (0.0286)	0.0970 (0.0319)	0.1055 (0.0287)	0.1050 (0.0320)	0.0931 (0.0362)
F-test pre-trends ( <i>p</i> -val.)	0.270	0.457	0.250	0.468	0.476	0.384
<i>N</i>	784	497	411	767	516	431
<i>N</i> × $\bar{T}$	17032	10781	8923	16670	11205	9366
<i>Panel C. Control cities within ± 4 ranks of treated cities in initial region</i>						
CAPITAL	0.1243 (0.0276)	0.0888 (0.0277)	0.0945 (0.0318)	0.1048 (0.0276)	0.1061 (0.0316)	0.0966 (0.0357)
F-test pre-trends ( <i>p</i> -val.)	0.194	0.458	0.408	0.406	0.460	0.340
<i>N</i>	913	564	457	887	581	469
<i>N</i> × $\bar{T}$	19831	12229	9924	19270	12609	10190

*Notes:* The table reports results from fixed effects regressions of the log of light intensity per square kilometer on capital city status. Panels A to C match treated cities to a varying number of control cities on the basis of their rank in terms of light intensity or population within the initial region. All regressions include city fixed effects, initial-region by-year fixed effects, and time-varying coefficients on the fundamentals. We report an F-test for pre-trends tests for the null hypothesis that all leading terms in the equivalent event-study specification are jointly zero. Standard errors clustered on initial regions are provided in parentheses.

to their initial light intensity or population in 1990 and designate all cities that are within  $k$ -ranks of the treated city as potential controls, where  $k$  ranges from 2 to 4 positions.<sup>27</sup> This creates a trade-off. While selecting among a subset of comparable cities in the initial region makes it more likely that these are good controls, positive spillovers imply that nearby cities are affected by the change in status of the capital city and therefore, as we just showed, represent a treatment group of their own.

Table III reports a range of results addressing these issues, all of which are based on the the most restrictive specification with initial-region-by-year fixed effects. By definition, we are now only using variation from cities in reformed regions. Reassuringly, every single estimate indicates a positive and significant effect of capital city status on city growth. We find effects in columns 1 and 4 that are close to our difference-in-differences results no matter if we use initial light intensity or estimates of city population in 1990 to define the control group, or if we consider only two, three or four similarly ranked cities. We also conduct simple omnibus tests for pre-trends using the equivalent event-study specification for each of these samples. In every case, we fail to reject the null hypothesis that the coefficients on all leading terms are jointly zero by a wide margin. The remaining columns remove observations whose minimum distance to a capital city is smaller than 50 or 75 km to mitigate the concerns about potential SUTVA violations.<sup>28</sup> There is some indication that the effects could be smaller once cities affected by spillovers are excluded. However, all estimates are well within two standard errors of one another and based on a different specification with substantially less variation in distance to treated cities than the more comprehensive spillover analysis presented above.<sup>29</sup>

**Other robustness checks:** We conducted a range of other checks verifying our analysis. We only briefly summarize their results here and report the corresponding tables in Online Appendix E. Our baseline estimates are robust to accounting for spatial autocorrelation (see Figure E-4) or using different versions of the light data, provided that there is some adjustment for bottom-coding (see Table E-2). In fact, the bottom correction and the non-filtered series from NOAA produce almost identical results. Correcting for top-coding then has a similar effect in terms of increasing the estimated magnitudes by another 2–3 percentage points. The estimated effects are robust to considering only cities that have a substantially larger initial population in 1990 and

---

<sup>27</sup>This approach is similar to Becker et al. (2021), who construct controls for Bonn—the temporary capital of the Federal Republic of Germany from 1949 until 1990—using cities ranked 20 places below and above Bonn in terms of their 1939 population.

<sup>28</sup>We disregard the time variation in distances to capital cities in this table to construct a conservative test which excludes all cities which were *ever* located within 50 or 75 km of a capital city. Results using time-varying distances are similar.

<sup>29</sup>In Table E-5 in Online Appendix E, we repeat this matching exercise using all similarly ranked cities in the country as controls in a specification with country-year fixed effects (to not limit the comparison to the same initial region). The results are qualitatively similar in these samples as well.

rise somewhat with initial city size (see [Table E-3](#)). Finally, none of these perturbations results in significant pre-treatment trends.

## 5. Treatment heterogeneity

We now examine two types of treatment heterogeneity to understand better which combination of factors affects economic activity in capital cities. The first are economic fundamentals and the second are factors limiting the capital city premium, such as a lack of economies of scale in the administrative region and settled urban networks. Of course, heterogeneity in the treatment effect implies that selection into the treatment could also be heterogeneous. We test for pre-trends by interacting these static variables with the entire event-time path.

**Locational fundamentals:** It is an open question in the literature whether political factors, such as designating subnational capitals, can substitute for the lack of good economic and geographic fundamentals or whether they, at best, complement these fundamentals. The allocation of subnational capitals and public investments may have large effects in the hinterland, i.e., in contexts with fewer local advantages or locations primarily suited for agricultural production, and induce little change in areas where (trade-related) fundamentals are strong to begin with, or vice versa. Recent evidence suggests that the growth potential of secondary cities in the agricultural hinterland might be limited. [Gollin et al. \(2016\)](#), for example, highlight that recent urbanization in resource-dependent economies has been concentrated in low productivity “consumption cities.” Urbanization in low productivity cities may be exacerbated by elevating cities to capitals in less favorable locations.<sup>30</sup>

[Table IV](#) tackles the question of fundamentals in our setting. We present two sets of results. Panel A shows results from regressions where we group our large set of potentially relevant fundamentals into aggregate indexes and reduce the underlying dimensionality by extracting the first principal component from three groups of fundamentals (internal trade, external trade, agriculture). Each column takes a group of fundamentals and interacts it with the treatment status. Panel B repeats this analysis using a representative fundamental from each group, as composite indexes are difficult to interpret. All variables are standardized to have mean zero and unit variance to facilitate comparisons across different specifications. We initially omit the time-varying coefficients on the control fundamentals in this specification, but this omission hardly affects the qualitative results.

---

<sup>30</sup>Other results in the literature can be viewed through this lens. For example, [Becker et al. \(2021\)](#) document that Bonn’s temporary status as the national capital of (West) Germany created little development apart from direct public employment. The city narrowly won its status over Frankfurt, which had considerably stronger fundamentals in the 1940s, and was always considered a temporary component of the division of Germany.



TABLE IV  
Heterogeneity in fundamentals: Difference-in-differences

	<i>Dependent Variable: <math>\ln \text{LIGHTS}_{cit}</math></i>				
	(1)	(2)	(3)	(4)	(5)
<i>Panel A. Principal components for each group of fundamentals</i>					
CAPITAL	0.1924 (0.0417)	0.1426 (0.0316)	0.1433 (0.0296)	0.1882 (0.0390)	0.1393 (0.0424)
CAPITAL $\times$ INT. TRADE	0.0957 (0.0390)			0.0860 (0.0381)	0.0481 (0.0366)
CAPITAL $\times$ EXT. TRADE		0.0027 (0.0170)		-0.0059 (0.0169)	-0.0047 (0.0156)
CAPITAL $\times$ AGRICULTURE			-0.0650 (0.0172)	-0.0600 (0.0162)	-0.0571 (0.0163)
<i>Panel B. Selected variables for each group of fundamentals</i>					
CAPITAL	0.2688 (0.0490)	0.1427 (0.0316)	0.1381 (0.0298)	0.2697 (0.0454)	0.2026 (0.0498)
CAPITAL $\times$ MARKET ACCESS	0.1293 (0.0342)			0.1348 (0.0316)	0.0957 (0.0297)
CAPITAL $\times$ DIST. TO COAST		-0.0012 (0.0264)		0.0234 (0.0238)	0.0101 (0.0227)
CAPITAL $\times$ WHEAT SUITABILITY			-0.0657 (0.0212)	-0.0638 (0.0194)	-0.0626 (0.0195)
Fundamentals	—	—	—	—	✓
$N$	8338	8338	8338	8338	8338
$N \times \bar{T}$	182529	182529	182529	182529	182529

*Notes:* The table reports results from fixed effects regressions of the log of light intensity per square kilometer on capital city status and interactions of the status with a particular fundamental. The interactions of the capital city status with some other variable  $\tilde{z}$  are standardized such that  $\tilde{z} \equiv (z - \bar{z})/\sigma_z$ . All first principal components are scaled to represent better suitability. All regressions include city fixed effects and initial-region-by-year fixed effects. Standard errors clustered on initial regions are provided in parentheses.

All heterogeneity analyses are based on the full difference-in-differences specification for cities in reformed regions (without accounting for spillovers).

Columns 1 to 3 in panel A show individual regressions where the capital city status is interacted with an index of how easy it is to trade internally, trade externally, or produce agricultural goods around the location of the city. Column 4 and 5 include all variables at the same time and add the time-varying coefficients on the fundamentals. In nearly all specifications, we find strong evidence of complementarity between gaining the political advantage of a capital city and economic fundamentals. A two standard deviation decrease in the index of internal trade offsets the positive capital city effect in column 4, although this effect is no longer significant in column 5. External trade integration appears to matter little for the relative growth rates of subnational capitals,

whereas cities which become capitals in agricultural locations attract considerably less activity than those located in other locations. While only suggestive, these results are in line with a reduced importance of agriculture for city locations or productivity and a greater importance of connectivity-related fundamentals today ([Henderson et al., 2018](#)).

Panel B unpacks these three groups. Column 1 interacts the treatment status with a city’s internal market access (to other cities in the country) in 1990. Here we observe a strong interaction effect. A city which becomes a capital in a location with a level of market access that is a standard deviation above the mean experiences an additional increase in light intensity of 13.8%. Given that most of our reforms occur in developing countries, this finding echos [Brühlhart et al. \(2020\)](#), who show that market access remains a strong determinant of regional productivity in developing countries, even as its importance declines in developed economies. Column 2 uses distance to coast as a measure of external market access. The coefficient points in the expected direction, but the estimated effect is small and insignificant. Column 3 uses wheat suitability as a proxy for locations in the agricultural hinterland. Mirroring the results from above, it shows that greater suitability for agriculture is negatively correlated with the growth of capital cities.

[Figure E-6 in Online Appendix E](#) shows that the event-study estimates of internal trade and market access rise in line with the baseline capital city effect, while the negative effect of agricultural fundamentals only starts to appear in the medium run and is barely significant. Moreover, the estimates of the pre-treatment coefficients and their sub-*t* confidence intervals strongly suggest that heterogeneous selection in terms of these fundamentals is unlikely.

**Limits to the capital city premium:** Our results thus far suggest that creating more subnational capitals and, hence, more first-order units benefits these cities and regions irrespective of their size. It is unlikely that subnational capital cities which rule over ever smaller territories would stand to benefit in the same way, as those who are the capitals of more populated regions. The literature on the optimal size of local jurisdictions highlights a number of relevant trade-offs, ranging from scale economies, over externalities in the provision of public goods, to preference homogeneity ([Oates, 1972](#); [Alesina et al., 2004](#); [Coate and Knight, 2007](#)). While our empirical framework is not suited to directly address most of these questions, we can examine if the positive effects of gaining capital city status documented here depend on scale. This captures the intensive margin of the “political importance” treatment, since we expect the premium to increase in the size of the region the capital administers and diminish when a region contains few other cities.<sup>31</sup>

---

<sup>31</sup>In a related study, [Grossman and Lewis \(2014\)](#) argue that Uganda has become so heavily decentralized that the intergovernmental bargaining power of a single first-order unit has been substantially weakened as a result.

TABLE V  
Economies of scale: Difference-in-differences

	<i>Dependent Variable: <math>\ln \text{LIGHTS}_{cit}</math></i>					
	Pop (region) (1)	(2)	Urban pop (region) (3)	(4)	No. cities (region) (5)	(6)
CAPITAL	0.1739 (0.0489)	0.1228 (0.0475)	0.2730 (0.0719)	0.1801 (0.0738)	0.2093 (0.0427)	0.1927 (0.0453)
CAPITAL $\times$ SCALE	0.0741 (0.0339)	0.0582 (0.0311)	0.1432 (0.0431)	0.0995 (0.0401)	0.1025 (0.0301)	0.1108 (0.0299)
Fundamentals	—	✓	—	✓	—	✓
City FE	✓	✓	✓	✓	✓	✓
Ini. Region-Year FE	✓	✓	✓	✓	✓	✓
$N$	8438	8438	8438	8438	8438	8438
$N \times \bar{T}$	184687	184687	184687	184687	184687	184687

*Notes:* The table reports results from fixed effects regressions of the log of light intensity per square kilometer on capital city status and interactions of the status with a the log of regional population, log of urban population in regions, and the log number of cities within the region. All regressions include the base term of scale variable. The interactions of the capital city status with some other variable measuring the scale of the new region ( $\tilde{z}$ ) are standardized such that  $\tilde{z} \equiv (z - \bar{z})/\sigma_z$ . Standard errors clustered on initial regions are provided in parentheses.

Table V reports a series of regressions in which we interact the capital city status with a variable measuring the scale of the region. Columns 1 and 2 use the regional population in 1990. Columns 3 and 4 use the urban population of the region in 1990, which we take as a proxy for the size of the non-agricultural economy, while 5 and 6 use counts of the number of cities in the resulting region. The measures of scale are standardized and time-varying (they change whenever a region is reformed). No matter in which way we specify this interaction with scale, we find evidence suggesting a trade-off: new capital cities benefit only when they rule over larger regions. Since the reforms in our baseline sample are splits of larger regions, this implies that creating new capitals of increasingly smaller regions weakens the effect of this reform on economic activity. A two standard deviation decrease in scale all but wipes out the capital city effect in all specifications. This could occur through a variety of channels, such as limited public investments in smaller regions or a small migratory response when the population of the new region is smaller.

Figure E-7 in Online Appendix E shows the corresponding event studies for each column of the table. We continue to observe an upward trend in the estimates for the capital city effect, at average levels of the scaling variable, and also observe an increase after treatment in the estimated interaction effects. Pre-trends appear to be flat, with the exception of the estimate two periods before treatment when the scaling variable is the urban population in the initial region.

In Online Appendix F, we examine whether “mother” capitals—i.e., cities that rule

over a smaller area after a regional split but maintain their political status—experience a change in economic activity. We find no evidence suggesting their activity responds in either direction (but the relatively wide confidence intervals include a range of potential effects). We also briefly examine the role of preference heterogeneity in [Online Appendix E](#). There, we use ethnic diversity in the initial region as a proxy for preference heterogeneity ([Table E-6](#)).<sup>32</sup> We find no evidence in favor of the hypothesis that capital cities grow at different speeds when they are located in more ethnically homogeneous (or diverse) regions. While this appears to contradict recent research suggesting that initial diversity inhibits agglomeration ([Eberle et al., 2020](#)), it only shows that such an effect is unlikely to run through changes in the density of capital cities relative to other cities.<sup>33</sup>

Next, we examine whether redesigning the territorial structure has different effects on the spatial equilibrium if a country agglomerated early or is still urbanizing today. Designating new capital cities could have little to no impact on migration in well-established urban networks with limited population pressures. In contrast, similar interventions in developing countries with growing populations and ongoing structural transformation could lead to a lasting shift in the location of economic activity. To test this conjecture and avoid constructing potentially endogenous sample splits, we rely on the country-level classification into early and late developing countries provided by [Henderson et al. \(2018\)](#).<sup>34</sup>

The results in [Table VI](#) show a clear pattern. No matter if we interact the capital city status with an indicator for late development according to education, urbanization or GDP in 1950, we always find that the effects are driven by late developers. The results for early developers are small and insignificant at conventional levels in nearly all samples apart from column 1, whereas the effect for late developers is remarkably stable across the different sample splits. The effect for late developing countries ranges from about 12.8% to 19.6%, depending on whether we account for time varying fundamentals or not, but varies little across the different grouping variables. [Figure E-8](#) in [Online Appendix E](#) shows the corresponding event studies and sup- $t$  confidence bands for each column of the table. No matter how the sample is split, we observe flat pre-trends and a steep rise after the reform in the sample of late developers. The results for early developers follow no systematic, or statistically significant, pattern.

---

<sup>32</sup>Similar to [Eberle et al. \(2020\)](#), we measure the ethno-linguistic fractionalization of initial regions using GHSL population data for 1975 and an algorithm developed by [Desmet et al. \(2020\)](#) which distributes data from the World Language Mapping System (WLMS)—the vector version of the Ethnologue project—on a  $5 \times 5$  km grid.

<sup>33</sup>[Eberle et al. \(2020\)](#) suggest that this is a result of diverse groups spreading over smaller cities. Given that we typically include country-year or initial-region-by-year fixed effects, we would not pick up an effect if all cities are similarly affected by initial diversity.

<sup>34</sup>[Henderson et al. \(2018\)](#) use a simple algorithm to let the data decide at which point the unexplained variance over the ‘late’ and ‘early’ samples is minimized. Their dependent variable is a contemporary cross-section of light intensity in a grid cell, while they define ‘late’ or ‘early’ according to urbanization, schooling and GDP per capita in 1950.

TABLE VI  
Late versus early developing countries: Difference-in-differences

	<i>Dependent Variable: <math>\ln \text{LIGHTS}_{cit}</math></i>					
	Late developer according to					
	Education 1950 (1)	(2)	Urbanization 1950 (3)	(4)	GDP per capita 1950 (5)	(6)
CAPITAL $\times$ EARLY	0.0381 (0.0203)	0.0214 (0.0224)	0.0769 (0.0562)	0.0653 (0.0567)	0.0239 (0.0284)	0.0140 (0.0329)
CAPITAL $\times$ LATE	0.1788 (0.0478)	0.1494 (0.0478)	0.1529 (0.0352)	0.1206 (0.0362)	0.1584 (0.0359)	0.1258 (0.0374)
Fundamentals	–	✓	–	✓	–	✓
City FE	✓	✓	✓	✓	✓	✓
Ini. Region-Year FE	✓	✓	✓	✓	✓	✓
$N$	7178	7178	8438	8438	8170	8170
$N \times \bar{T}$	157023	157023	184687	184687	178982	178982

*Notes:* The table reports results from fixed effects regressions of the log of light intensity per square kilometer on capital city status in early and late developing countries defined according to [Henderson et al. \(2018\)](#). All regressions include city fixed effects and initial-region-by-year fixed effects. Standard errors clustered on initial provinces are provided in parentheses.

[Online Appendix E](#) shows that these late developer results can not be fully explained by differences in political systems. Expanding on [Ades and Glaeser \(1995\)](#), we find that capitals in autocracies grow faster compared to their democratic counterparts. Yet, the size of the effect is much smaller than the early vs. late distinction and only present in the full sample of cities (see [Table E-7](#)). A potentially important distinction is that first-order administrative regions, including their capitals, have more political power in federal countries (such as states in Germany or India) than in more centralized countries (such as departments in France or provinces in China). [Table E-8](#) shows that the capital city premium exists in both federal and centralized countries but tends to be about twice as large in the former.

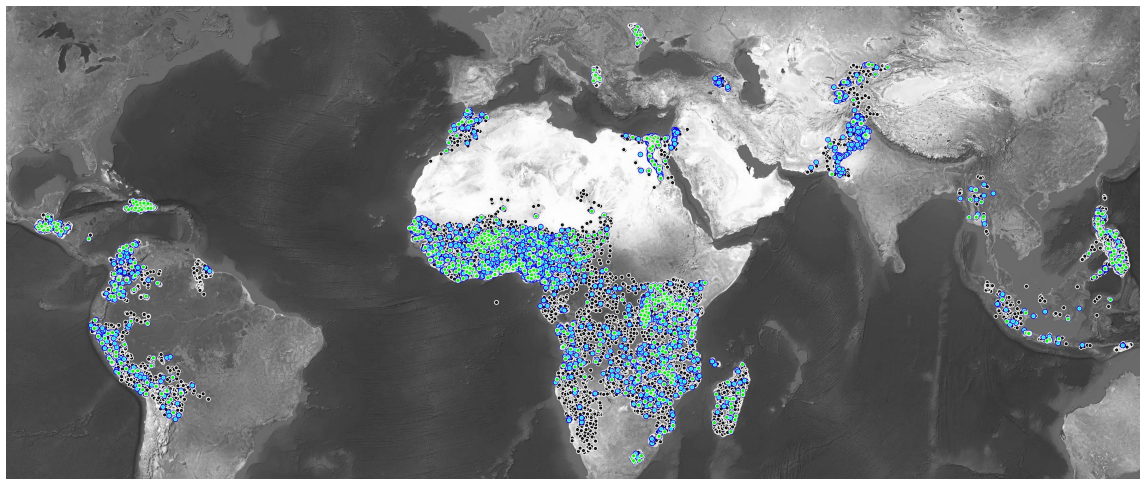
Taken together, our estimates imply that there is strong heterogeneity in the medium-run effect of gaining capital city status. We interpret this as evidence that “territorial politics” can have a substantial effect on the location of economic activity and urbanization in developing countries, but that this effect varies by locational fundamentals, the territorial design of the administrative structure, federalism, and the age of the urban network.

## 6. Individual-level evidence

We now examine individual-level evidence and compare outcomes among residents in capital cities to residents of non-capital cities in the same initial region. This additional

layer of analysis serves two purposes. First, it allows us to validate the results obtained with nighttime lights using indicators that directly measure development outcomes (and a different estimation strategy). Second, the survey data enable us to examine specific dimensions in which residents of capital cities might be doing better and frame the mechanisms we examine in the next section. For example, we can test if households in capital cities have better access to electricity, more schooling or better health outcomes.<sup>35</sup>

FIGURE VII  
DHS coverage



*Notes:* The figure illustrates the spatial distribution of DHS clusters (dark blue with gray borders), DHS clusters matched to our cities (light blue with dark blue borders) and DHS cluster matched to capitals (green with light gray borders).

We compile a global sample of geocoded Demographic and Health Surveys (DHS) to test for differences in individual and household-level outcomes between residents of capitals and non-capital cities.<sup>36</sup> Figure VII plots our global coverage of DHS clusters in gray (about 80,000 clusters containing roughly 1.36 million households), the ones we are able to match to any of our cities in blue (about 25,000 clusters), and those that are located in capital cities at the time of the survey in green (56% of the matched clusters). Note that capital cities are heavily over-represented among urban DHS clusters (56% of matched clusters are capitals).

Our strategy is to pool all of this data and run regressions comparing household-level or individual-level outcomes in non-capital cities in the same initial administrative region to outcomes in capital cities. We control for the fundamentals of each city, including initial population density, and household-level characteristics so that we estimate the capital city premium (or penalty) in otherwise comparable locations and households. We study two different samples—all city dwellers and those that were born in the city—as

<sup>35</sup>We adjust the approach of [Henderson et al. \(2020\)](#) to our setting, where we exclusively focus on differences between different types of cities, rather urban-rural differences.

<sup>36</sup>The data are described in detail in [Online Appendix A](#).



TABLE VII  
Amenities in capitals: Within-region regressions

	<i>Dependent Variable:</i>				
	DHS WEALTH INDEX	ELEC- TRICITY	YEARS OF EDU $\geq$ 8	LN YEARS SCHOOLING	INFANT MOR- TALITY
	(1)	(2)	(3)	(4)	(5)
<i>Panel A. All city dwellers</i>					
CAPITAL	0.3413 (0.0391)	0.0901 (0.0135)	0.0073 (0.0021)	0.0096 (0.0055)	-7.2745 (2.2455)
Outcome mean	4.4603	0.6848	0.0334	1.318	65.4767
N	144960	144960	236500	238579	651132
<i>Panel B. City natives / born in city</i>					
CAPITAL	0.1951 (0.0495)	0.0324 (0.0142)	0.0054 (0.0038)	-0.0001 (0.0082)	-8.2015 (3.7040)
Outcome mean	4.3129	0.7500	0.0302	1.2903	67.6722
N	39023	39023	65096	66036	183576
DHS controls	✓	✓	✓	✓	✓
Fundamentals	✓	✓	✓	✓	✓
Ini. Region-Year FE	✓	✓	✓	✓	✓

*Notes:* The table reports results from regressions of various DHS measures on the current capital status of a city (i.e., the capital status in the corresponding survey-year). Panel A reports results for all respondents/ households currently residing in a city. Panel B reports results for those that were born in the city of current residence. The dependent variables are the DHS wealth index, a dummy for the presence of electricity, a dummy for 8 years of schooling (or more) for respondents 16 years or older, log years of schooling for respondents 16 years or older, and infant mortality. All columns include initial-region-by-year fixed effects, an indicator for the national capital, and DHS controls (log household size, female household head, log age of the household head, and three indicators for completed primary, secondary and higher education of the household head). Columns 1 and 2 use data at the level of households. Columns 3 and 4 use individual level data and add age, age squared, and a female indicator of the respondent to the household level controls. All columns include locational fundamentals, as well as an indicator if the respondent lives in the national capital city. We allow all those additional controls to have different effect in different survey years. In column 5 we use respondent-child-level data with the following controls: a gender dummy, an indicator variable for multiple children, and a set of period of birth dummies (each period corresponds to a decade, e.g., 1990s). Note that due to degree of freedom issues we allow the fundamentals in column 5 to only vary across 5 year periods. Standard errors clustered on the agglomeration provided in parentheses.

a first test to see if migrants moving to capital cities differ from those who have always lived there.

Table VII presents the results of specifications regressing household-level or individual-level outcomes on a capital dummy at the time of the survey. Column 1 highlights that households in capitals accumulate, on average, more assets than households in non-capital cities. The effect is larger for all city dwellers (Panel A) than for city natives (Panel B), but makes up close to 7.5% of the sample average in both cases. Column 2 shows that

residents of capital cities are about 9 percentage points more likely to have electricity in their household compared to residents of other cities within the same initial region, but this falls to 3 percentage points when we restrict the sample to city natives. Turning to schooling, we find that residents of capital cities that are 16 years or older are 7.3 percentage points more likely to have completed eight years of schooling (column 3 in Panel A). While this effect appears small, only around 3.3% of all respondents over the age of 16 have completed eight years of education, so that this about 22% of the sample average. We find weaker results in terms of years of schooling (about 1% more) for all residents of capital cities. Remarkably, the schooling results disappear when we only look at city natives (Panel B), suggesting that the migrant population of capital cities is somewhat more educated than the native population. Finally, we also find a sizable reduction in infant mortality of around 7.3 to 8.2 children per 1000 births. This is up to 12.1% of mean infant mortality.

Taken together, these results imply that residents of capital cities in developing countries enjoy a number of benefits compared to those in non-capital cities. The increase in household wealth suggests that our main finding on density also runs through increased productivity and higher wages. The proximity to government appears to manifest itself in substantially better access to electricity and health outcomes. However, the pattern of larger effects for all city dwellers compared to city natives suggests that some of these results may be driven by selective migration into capital cities. Such migration flows could be a response to a city becoming a subnational capital, which is why we return to this issue when studying potential mechanisms.

## 7. Mechanisms

There are a range of mechanisms which could explain why capital cities attract more economic activity. We first examine changes in urban built-up and population at the city level to investigate whether new capital cities are, in fact, becoming denser than non-capital cities. Second, we use the DHS data from above to examine whether capital cities attract more highly skilled migrants. Third, we utilize city-level data on international financial flows to examine whether capital cities mainly receive public funds in the form of development projects or also an influx of private investment from abroad.

**Urban built-up and population growth:** Ideally, we would like to have city-level measures for the housing stock, housing prices, and annual data on city-level population. As these are not available for our global sample of cities, we construct proxies based on remotely-sensed and census-derived data.

Table VIII examines changes in urban land cover within a city.<sup>37</sup> We derive three

---

<sup>37</sup>Figure E-10 in Online Appendix E reports the corresponding event studies for the preferred

TABLE VIII  
Built-up area within city: Difference-in-differences

	<i>Dependent Variable: Built-up area within city</i>					
	NDBI		UI	NDVI		
	(1)	(2)		(4)	(5)	(6)
CAPITAL	0.8342 (0.2201)	0.6659 (0.2292)	1.2889 (0.2968)	0.9594 (0.3157)	-0.9809 (0.2463)	-0.6233 (0.2670)
Fundamentals	–	✓	–	✓	–	✓
City FE	✓	✓	✓	✓	✓	✓
Ini. Region-Year FE	✓	✓	✓	✓	✓	✓
$N$	8438	8438	8438	8438	8438	8438
$N \times \bar{T}$	178866	178866	178866	178866	178866	178866

*Notes:* The table reports results from fixed effects regressions of the Normalized Difference Built-Up Index (NDBI), Urban Index (UI), and the Normalized Difference Vegetation Index (NDVI) on capital city status. All coefficients are scaled by 100 for exposition. Standard errors clustered on initial regions are provided in parentheses.

frequently used spectral indices of land cover by creating annual composites from daily images taken by the Landsat 5 and Landsat 7 satellites over the entire period from 1987 to 2018. The Normalized Difference Built-up Index (NDBI), the Urban Index (UI) and the Normalized Difference Vegetation Index (NDVI) are well-established spectral indices of urban and non-urban land cover, which are typically used as inputs for more advanced land cover classifications (e.g. the USGS data used by [Burchfield et al., 2006](#)). Since we already know the urban extent of each city in 1990, taking the average value of each index within each city and year allows us to track annual changes in urban land cover within the city core, just as in our main specification. Columns 1 to 6 show that each of these measures either indicates a significant increase in urban land cover (due to new housing, businesses or infrastructure) or a corresponding decrease in vegetation in case of the NDVI when a city gains capital status. Depending on the sample, the estimated effect sizes are about 10–15% of the typical within-city standard deviation for these indices. [Online Appendix E](#) shows that we also observe increases in built-up in the larger agglomerations, including areas outside of the 1990s core ([Table E-9](#)). Hence, part of the increase in economic activity is directly reflected in housing and other physical infrastructure.

The increase in urban built-up is mirrored by a sizable increase in population. [Table E-10](#) in [Online Appendix E](#) uses long differences from the census-derived GHSL population data in 1990, 2000 and 2015 to show that cities which were capitals for a longer period experience stronger population growth. There is some variation across time-periods and specifications but, on average, subnational capitals appear to grow slightly more than a

specification with time-varying effects of the static fundamentals.

percentage point per year faster than other cities. We interpret this as indirect evidence that short and medium term increases in light intensity at the city level partly reflect an influx of population from other locations.<sup>38</sup>

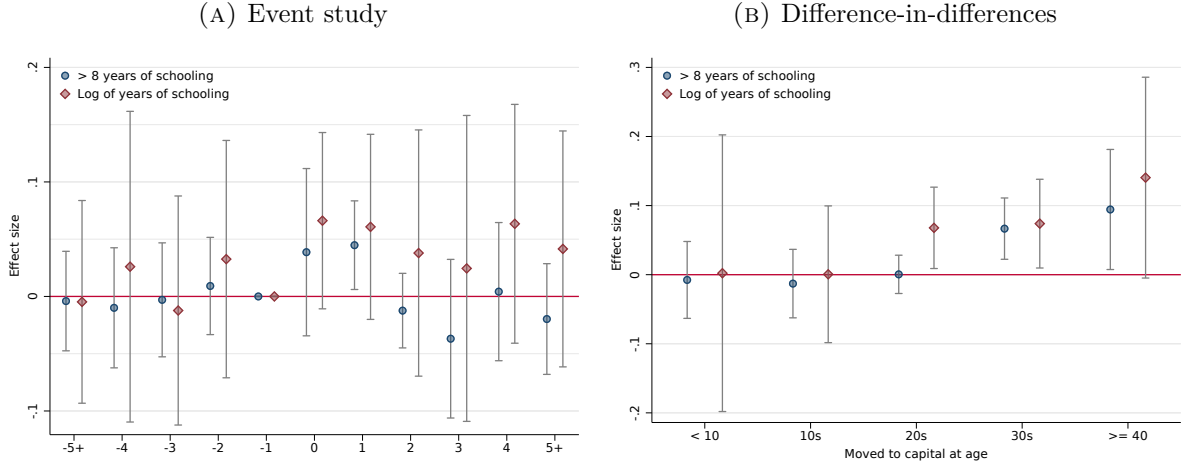
**Selective migration:** We now return to the question of whether migrants to capital cities are different from other city-migrants. The DHS surveys allow to go beyond studying cross-sectional differences among city dwellers. For a subset of 7,539 migrants living in cities that have become capitals in our sample, we know how many years they have resided at their current location. We use this information to create a synthetic panel of recent migrants living in cities that eventually become capitals. For each of these individuals, we observe whether they moved to a (new) capital before or after it has gained its political status, plus some initial characteristics, such as their age when they moved. We use this cohort-level panel in two ways. First, we set up an event study mimicking our main results, where we examine the schooling attainment of migrants that have arrived shortly before or after a city became a capital. Second, we collapse these data to a difference-in-differences specification where we allow the treatment effect to vary by cohort. All specifications use only variation within the subset of cities that become capitals, individuals within the same city-year, and individuals within the same cohort during the time of the move. We focus on the educational attainment of migrants only, since human capital is embodied and not tied to the place of current residence.

Figure VIII illustrates the corresponding results. Panel A shows that we find little evidence that migration is selective in terms of educational attainment prior to a city becoming a capital city. Migrants arriving two years or more before a city becomes a capital seem to be no more or less likely to have completed eight years of schooling than those in the year prior to the reform. In the year after the status change, we observe an uptick in the probability of migrants having at least eight years of schooling (around 4.5 percentage points). However, the uncertainty around most of these event-study coefficients is relatively large, especially when it comes to log years of schooling, and we do not know if these migrants have acquired this additional human capital before or after they moved. Panel B addresses both points by reducing the specification to before-and-after comparisons for different cohorts. We find that migrants arriving after a city became a capital that were in at least in their 30s (and perhaps already in their 20s)

---

<sup>38</sup>Given the pattern of spillovers documented earlier, this increase is likely to come from rural areas and cities far away from the capital (although we have no direct evidence documenting such a migratory response). However, we do find some evidence that the larger agglomeration appears to grow faster than the core. Table E-11 in Online Appendix E shows analogous regressions. Due to the limited time period and measurement errors inherent in the spatial and temporal interpolation of census data, we can no longer specify regressions that mirror our main identification strategy. Instead, we ask how much faster the population density of capital cities grows relative to that of other cities in the same initial region (and with similar fundamentals).

FIGURE VIII  
Selective migration: Within city evidence



*Notes:* The figure plots estimates of educational attainment from the DHS surveys on different treatment dummies. Panel A reports event-study results from fixed effects regressions of a dummy for more than 8 years of schooling for migrants of all ages (blue circles) and log years of schooling for migrants of all ages (red triangles) on a binned sequence of treatment change dummies defined as the year of the last move of a migrant minus the year a city changed its status. Panel B reports the results of a difference-in-differences specification with cohort-level heterogeneity. All specification include city-year and cohort-at-move fixed effects (defined four bins from ‘0 to 10’ to ‘older than 40’ at the time of the last move), a gender dummy, age and age squared. 95% confidence intervals based on standard clustered on the city level are provided by the gray error bars.

at the time of the move appear to have higher human capital.<sup>39</sup> Given their age at the time of the move, this strongly suggests that these migrants have acquired their human capital before they moved to a new capital city. We take this as direct evidence that capital cities attract a (moderately) more skilled migrant population than other cities.

**Public and private investments by sector:** The previous finding raises the question of whether capital cities only attract public investments (and public employment) or whether we also observe a private response. [Bairoch \(1991\)](#), for example, describes places where educated bureaucrats and property owners rule in the absence of industry as “parasite cities.” Lacking census data on the employment structure of cities, we use two proxies for public and private investments: *i*) geocoded data on World Bank projects from 1995 to 2014 and similar data on Chinese-financed development projects from 2001 to 2014, and *ii*) the fDi Markets database on private foreign direct investments, which is available over the period from 2003 to 2018.<sup>40</sup>

World Bank projects are usually carried out in close cooperation with the national government in the recipient country and even substitute for some of its basic functions in

<sup>39</sup>Those that moved in their 30s or later have a 6.6–9.4 pp higher probability to have eight years of schooling and 7.5–15% more years of schooling than those in the same age group who moved before the city became a capital.

<sup>40</sup>fDi Market is proprietary and available via subscription at [www.fdimarkets.com](http://www.fdimarkets.com).

particularly poor countries. The data includes all projects approved in the World Bank IBRD/IDA lending lines over this period ([AidData, 2017](#)), including many infrastructure investments. It contains more than US\$630 billion in commitments (in 2011 dollars) which were spent on 5,684 projects in 61,243 locations. We supplement this data with geocoded project level data on China’s global footprint of official financial flows over the period from 2000 to 2014 ([Bluhm et al., 2020](#)). The data include 3,485 projects (worth US\$273.6 billion in 2014 dollars) in 6,184 locations across the globe. China invests heavily in economic infrastructure and services, ranging from roads over seaports to power grids.

Both the World Bank and China project locations were geocoded ex post and include precision codes indicating if an exact building, city or administrative region were identified.<sup>41</sup> We use a subset of these data which were coded to be either exact or near to the exact location. We then match these projects to our universe of capital and non-capital cities if they fall within 10 km of the city centroid. Aiddata codes the sector of each project following a variant of the OECD Creditor Reporting System (CRS). This allows us to distinguish investment in, say, government from investments in water and sanitation. We examine two outcomes for overall and aid per sector in a series of cross-sectional regressions: whether a city had any project committed over the entire available period (e.g. 1995 to 2014 for the World Bank) and the total amount committed. On the right hand side we have the share of years in which a city was a capital. All regressions include initial-region fixed effects, control for the full set of fundamentals (including initial population), and include a dummy for national capitals. Similar to the long difference approach taken with population, this help us to understand whether capital cities attract more or less projects funded from abroad than comparable cities, but this evidence does not show that new capital cities attract funding immediately after gaining status.<sup>42</sup>

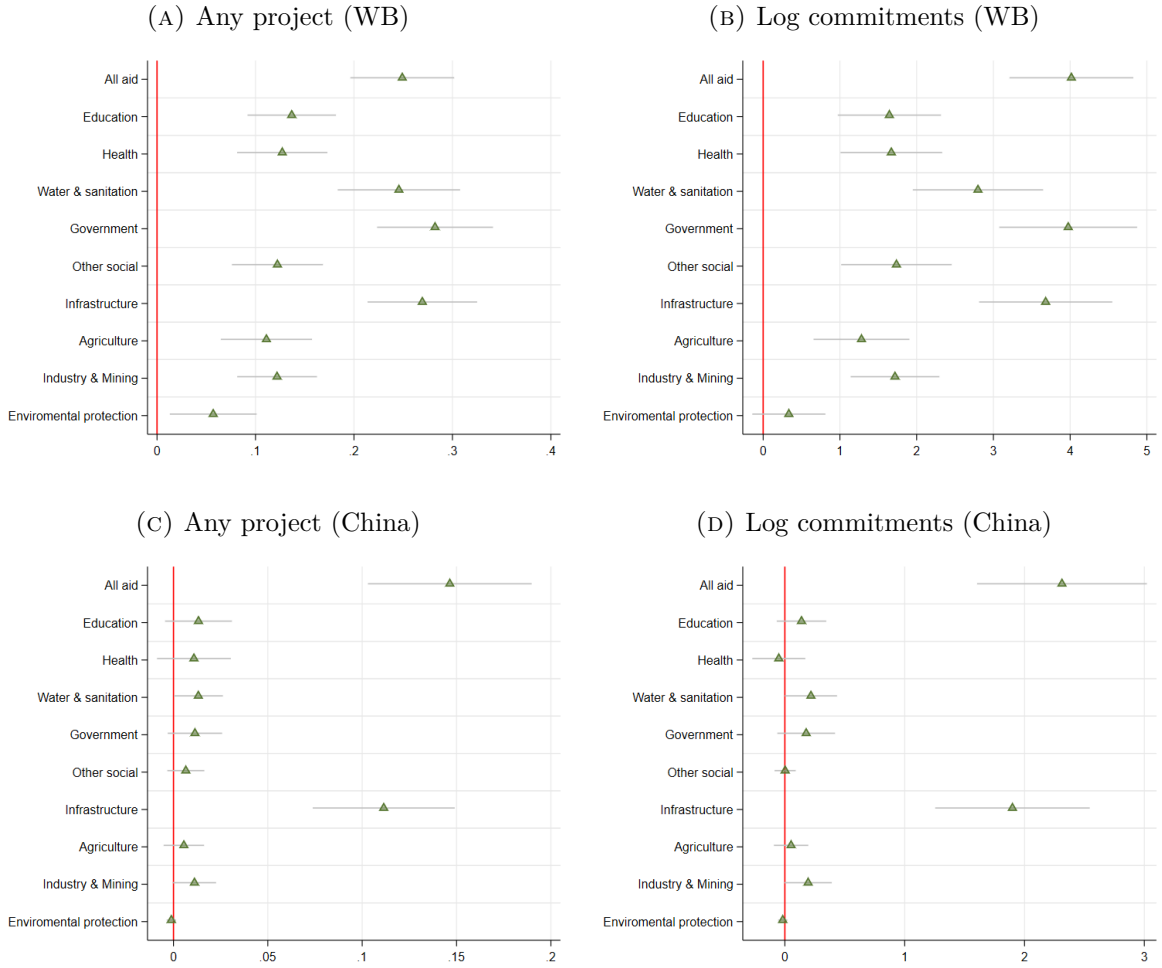
[Figure IX](#) reports the regression results per sector and shows several interesting patterns. First, capital cities are considerably more likely to receive any development project than other cities in the same region. For example, a city that has been a capital throughout the entire period has a 24.9 percentage point higher probability of receiving any World Bank project and was about 14.6 percentage points more likely to receive a Chinese-funded project than a non-capital city. Second, the sectoral composition of projects suggest that capital cities primarily attract funds for water and sanitation, government and civil society, and infrastructure when it comes to funding by the World Bank. China, on the other hand, appears to invest relatively more in the physical infrastructure of subnational capitals compared to other cities or sectors. In fact, most of the other sectors, apart from water and sanitation or industry and mining show no

---

<sup>41</sup>Restricting the sample to the two highest precision codes makes sure that we do not mechanically find more projects in subnational capitals, as projects for which less precise geographic information is available are often geocoded to regional capitals.

<sup>42</sup>Running event studies with these data matched to cities leads to results which have large uncertainties and could encompass a large range of positive or negative effect sizes.

FIGURE IX  
Capitals and aid by sector



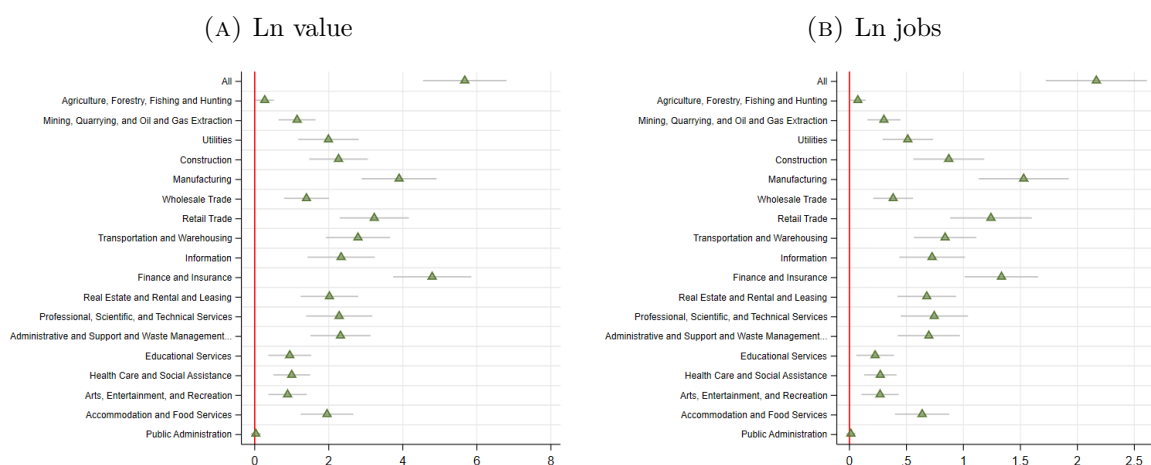
*Notes:* The figure plots estimates from regressions of development projects in a particular sector on the fraction of years a city was a capital. Panels A and C show the results from regressions with binary variables on the left hand side indicating whether a city has received any project in a particular sector by the World Bank and China, respectively. Panels B and D show the results from regressions with the log of 1 + commitments in USD on the left hand side indicating the amount of commitments a city has received in a particular sector by the World Bank and China, respectively. The definition of sectors follows the OECD's Common Reporting Standard (see [Online Appendix A](#) for details). 95% confidence intervals based on standard errors clustered on initial regions are provided as error bars.

differences across capital and non-capital cities. Subnational capitals thus attract more public funds from international donors or creditors and these funds primarily go to sectors that improve living standards, government operations, and connectivity. This pattern of investments lines up well with the DHS evidence presented above (and the related evidence in [Gollin et al., 2016](#); [Henderson and Turner, 2020](#); [Henderson et al., 2020](#)). Of course, we cannot rule out that this allocation is the product of regional favoritism or corruption ([Hodler and Raschky, 2014](#); [De Luca et al., 2018](#)) and cannot test within our framework whether these projects are effective.



Having documented that capital cities attract more public funds begs the question whether the increase in activity and density is primarily the product of more public employment (and the related investments) or whether private employment increases as well. This is far from a theoretical conjecture. A significant part of the increase in the urban extent of Mamuju after 1990 documented in [Figure IV](#) occurs in the area of the city where the new provincial headquarters were built. Recent findings on public employment multipliers appear to depend on the context and usually come from developed countries.<sup>43</sup> Public investment multipliers could be substantially larger in developing countries where the initial capital stock is low.

FIGURE X  
Capitals and FDI by industry



*Notes:* The figure plots estimates from regressions of FDI projects in a particular sector on the fraction of years a city was a capital. Panel A shows the results from regressions with the log of  $1 +$  value of FDI projects in USD on the left hand side indicating the amount of FDI a city has received in a particular sector. Panel B shows the results from regressions with the log of  $1 +$  estimated number of jobs of FDI projects in the city on the left hand side indicating the increase in private employment a city has experienced in a particular sector. The definition of sectors follows the NAICS 2-digit sector classification (see [Online Appendix A](#) for details). 95% confidence intervals based on standard errors clustered on initial regions are provided as error bars.

The fDi Markets tracks global FDI investments and joint ventures by sector, provided that they lead to a new physical operation in the host country. Similar to the China data, the data are not primarily based on official statistics but collected from media, industry organisations and investment agencies. The fDi Markets data reports the host city of the project, the value of the investment and an estimate of the jobs created

<sup>43</sup>[Faggio and Overman \(2014\)](#), for example, find no evidence of increases in private employment in response to substantial increases in public employment in more rural areas of the UK. [Becker et al. \(2021\)](#) study the relocation of the German government to Bonn in 1949 and [Faggio et al. \(2019\)](#) study the move back to Berlin in 1999. Public jobs crowded out private employment in Bonn almost one-to-one, while Berlin's service sector benefited (with 0.55 private jobs for every public job). [Jofre-Monseny et al. \(2020\)](#) use the capital city status of Spanish cities to identify the effects of public employment and find evidence of significant crowd-in of private sector employment (1 to 1.3 jobs).

that can be connected to the investment. We use a subset of the global data for cities in reformed regions. We geocode each host city and match it to an urban cluster or administrative center when it is within 10 kilometers of an FDI project. We run the same set of regressions used for aid projects, only this time the cumulative value of FDI projects in a city or the cumulative number of projected jobs a company plans to create are on the left hand side, and the fraction of years a city was a capital from 2003 to 2018 plus fundamentals and initial-region fixed effects are on the right hand side. The FDI data uses the North American Industry Classification System (NAICS). We aggregate their highly detailed classification to the 2-digit NAICS industry level, which delivers a level of aggregation similar to the CRS codes used for aid projects.

Figure X suggests that subnational capitals attract considerably more FDI than non-capital cities. The value of FDI projects in capitals is about 5.7 log points larger and these project come with an 8-fold increase in the number of jobs compared to projects in non-capital cities in the same initial region. The sectoral composition of FDI also shows an interesting pattern. Capital cities attract more high-value projects with more manufacturing, finance and insurance, and retail jobs than other cities. However, differences to other cities in terms of private investments in public administration and agriculture are negligible. This finding suggests that core industries with international linkages preferably locate themselves in subnational capitals.

Although the evidence documented on these mechanisms is only a series of partial correlations, it generally supports the conjecture that capital cities attract substantial public and private investments. Our findings on medium-run population changes also make it unlikely that the fast growth of subnational capitals can only be attributed to increases in public employment.<sup>44</sup>

## 8. Conclusion

Our results provide the first evidence that the recent proliferation of administrative units and the corresponding change in status of some cities to first-order administrative capitals affect the location of economic activity in developing countries. Leveraging a new and global panel of administrative reforms from 1987 until 2018, we find that new capital

---

<sup>44</sup>Public employment shares in developing countries are often not particularly high (e.g. provincial level public employment ranges from about 8 to 47 per thousand in Indonesia, see [OECD, 2016](#)), so that the population increases documented above would imply an implausibly large expansion of the public sector. To see this, consider an initial public employment share of 5%. If the population of a capital city grows by 30% from 1990 to 2015 as a consequence of the change in capital status, then public employment would have to rise 6-fold to explain all of this increase. If the increases are more modest, for example a doubling in public employment, then the associated multipliers would have to be much larger than what is suggested by the literature ([Faggio and Overman, 2014](#); [Becker et al., 2021](#); [Jofre-Monseny et al., 2020](#)). This is in line with [Bai and Jia \(2021\)](#), who document similar results for the very long-run development of China’s prefectures when they host provincial capitals.

cities attract significantly more economic activity in the short and medium run. These benefits spill over to nearby cities but are heterogeneous across key dimensions. Capitals in inferior locations, as defined by a lack of internal market access or high agricultural suitability, experience considerably weaker growth than those in superior locations. The capital premium varies along the intensive margin of political importance, as measured by the size of the resulting administrative unit, and depends inversely on the age of urban network.

We interpret these findings and our analysis of likely channels as evidence that subnational capitals are focal points for migrants and business within regions. If these capitals coincide with productive locations, then accelerating agglomeration in these cities is likely to impact aggregate welfare positively. More broadly, these findings are relevant to policy-makers who decide to decentralize based on various political considerations. Territorial politics and public investments can be a tool for steering agglomeration in rapidly urbanizing developing countries. However, we also illustrate how ineffective such policies are if they target unfavorable locations in the hinterland or when their implementation no longer delivers sufficient economies of scale.

The global data we provide in this paper opens the door to studying various questions about cities and their role in the administrative hierarchy. So far, we know little about the politics behind the observed locations of subnational capital cities apart from a few historical cases. Our paper only exploits the varied nature of the underlying motivations and their unpredictability but makes no contribution to untangling them. We leave such questions for future research and hope that our global data of administrative cities will prove useful in exploring them.

## References

- Ades, A. F. and E. L. Glaeser (1995). Trade and circuses: Explaining urban giants. *Quarterly Journal of Economics* 110(1), 195–227.
- Agarwal, V. (2017). India and its new states: An analysis of performance of divided states - pre and post bifurcation. *Strides* 2(1), 31–46.
- AidData (2017). `WorldBank_GeocodedResearchRelease_Level1_v1.4.2` geocoded dataset. Aid Data Williamsburg, VA and Washington, DC. AidData. Accessed on 02/09/2020, <http://aiddata.org/research-datasets>.
- Alesina, A., R. Baqir, and C. Hoxby (2004). Political jurisdictions in heterogenous communities. *Journal of Political Economy* 112(2), 348–396.
- Allen, T. and D. Donaldson (2020, November). Persistence and path dependence in the spatial economy. Working Paper 28059, National Bureau of Economic Research.
- Asher, S., J. P. Chauvin, and P. Novosad (2019, June). Rural spillovers of urban growth. Discussion Paper IDB-DP-691, Inter-American Development Bank.
- Bai, Y. and R. Jia (2021). The economic consequences of political hierarchy: Evidence from regime changes in China, 1000-2000 C.E. *Review of Economics and Statistics*. forthcoming.
- Bairoch, P. (1991). *Cities and economic development: From the dawn of history to the present*. Chicago, IL: University of Chicago Press.
- Baragwanath, K., R. Goldblatt, G. Hanson, and A. K. Khandelwal (2019). Detecting urban markets with satellite imagery: An application to India. *Journal of Urban Economics*, 103173.
- 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, 103124.
- Bazzi, S. and M. Gudgeon (2021). The political boundaries of ethnic divisions. *American Economic Journal: Applied Economics* 13(1), 235–66.
- Becker, S. O., S. Heblich, and D. M. Sturm (2021). The impact of public employment: Evidence from Bonn. *Journal of Urban Economics* 122, 103291.
- Bleakley, H. and J. Lin (2012). Portage and path dependence. *Quarterly Journal of Economics* 127(2), 587–644.
- Bluhm, R., A. Dreher, A. Fuchs, B. Parks, A. Strange, and M. J. Tierney (2020, May). Connective Financing - Chinese Infrastructure Projects and the Diffusion of Economic Activity in Developing Countries. CEPR Discussion Papers 14818, C.E.P.R. Discussion Papers.
- Bluhm, R., R. Hodler, and P. Schaudt (2021). Ethnofederalism and ethnic voting. CESifo Working Paper Series 9314, CESifo.
- Bluhm, R. and M. Krause (2022). Top lights: Bright cities and their contribution to economic development. *Journal of Development Economics* 157, 102880.
- Borusyak, K., X. Jaravel, and J. Spiess (2021). Revisiting event study designs: Robust and efficient estimation. ArXiv preprint 2108.12419.
- Brühlhart, M., K. Desmet, and G.-P. Klinke (2020). The shrinking advantage of market potential. *Journal of Development Economics*, 102529.
- 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.
- Callaway, B. and P. H. Sant’Anna (2021). Difference-in-differences with multiple time

- periods. *Journal of Econometrics* 225(2), 200–230.
- Campante, F. and D. Yanagizawa-Drott (2017, 12). Long-Range Growth: Economic Development in the Global Network of Air Links\*. *Quarterly Journal of Economics* 133(3), 1395–1458.
- Campante, F. R. and Q.-A. Do (2014). Isolated capital cities, accountability, and corruption: Evidence from us states. *American Economic Review* 104(8), 2456–81.
- Campante, F. R., Q.-A. Do, and B. Guimaraes (2019). Capital cities, conflict, and misgovernance. *American Economic Journal: Applied Economics* 11(3), 298–337.
- Chambru, C., E. Henry, and B. Marx (2021, December). The dynamic consequences of state-building: Evidence from the French Revolution. CEPR Discussion Paper 16815, Centre for Economic Policy Research.
- Chauvin, J. P., E. Glaeser, Y. Ma, and K. Tobio (2017). What is different about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States. *Journal of Urban Economics* 98, 17–49.
- Coate, S. and B. Knight (2007). Socially optimal districting: a theoretical and empirical exploration. *Quarterly Journal of Economics* 122(4), 1409–1471.
- Combes, P.-P., G. Duranton, and L. Gobillon (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics* 63(2), 723–742.
- Combes, P.-P., G. Duranton, L. Gobillon, and S. Roux (2010). Estimating agglomeration economies with history, geology, and worker effects. In *Agglomeration economics*, pp. 15–66. University of Chicago Press.
- Davis, D. R. and D. E. Weinstein (2002). Bones, bombs, and break points: The geography of economic activity. *American Economic Review* 92(5), 1269–1289.
- Davis, J. C. and J. Henderson (2003). Evidence on the political economy of the urbanization process. *Journal of Urban Economics* 53(1), 98–125.
- De Luca, G., R. Hodler, P. A. Raschky, and M. Valsecchi (2018). Ethnic favoritism: An axiom of politics? *Journal of Development Economics* 132(C), 115–129.
- Depetris-Chauvin, E. and D. N. Weil (2018). Malaria and early African development: Evidence from the sickle cell trait. *Economic Journal* 128(610), 1207–1234.
- Desmet, K., J. F. Gomes, and I. Ortuño-Ortín (2020). The geography of linguistic diversity and the provision of public goods. *Journal of Development Economics* 143, 102384.
- Dijkstra, L., A. J. Florczyk, S. Freire, T. Kemper, M. Melchiorri, M. Pesaresi, and M. Schiavina (2021). Applying the degree of urbanisation to the globe: A new harmonised definition reveals a different picture of global urbanisation. *Journal of Urban Economics* 125, 103312.
- Donaldson, D. and R. Hornbeck (2016). Railroads and American economic growth: A “market access” approach. *Quarterly Journal of Economics* 131(2), 799–858.
- Düben, C. and M. Krause (2021). The emperor’s geography - City locations, nature, and institutional optimization. Mimeograph.
- Eberle, U. J., J. V. Henderson, D. Rohner, and K. Schmidheiny (2020). Ethnolinguistic diversity and urban agglomeration. *Proceedings of the National Academy of Sciences* 117(28), 16250–16257.
- Faggio, G. and H. Overman (2014). The effect of public sector employment on local labour markets. *Journal of Urban Economics* 79, 91–107.
- Faggio, G., T. Schluter, and P. vom Berge (2019). Interaction of public and private employment: Evidence from a German government move. Discussion paper.

- Glaeser, E. L. (1999). Learning in cities. *Journal of Urban Economics* 46(2), 254–277.
- Glaeser, E. L. and J. D. Gottlieb (2008). The Economics of Place-Making Policies. *Brookings Papers on Economic Activity* 39(1 (Spring)), 155–253.
- Gollin, D., R. Jedwab, and D. Vollrath (2016). Urbanization with and without industrialization. *Journal of Economic Growth* 21(1), 35–70.
- Gollin, D., M. Kirchberger, and D. Lagakos (2021). Do urban wage premia reflect lower amenities? Evidence from Africa. *Journal of Urban Economics* 121, 103301.
- Grossman, G. and J. I. Lewis (2014). Administrative unit proliferation. *American Political Science Review* 108(1), 196–217.
- Henderson, J. V., V. Liu, C. Peng, and A. Storeygard (2020). Demographic and health outcomes by degree of urbanisation: Perspectives from a new classification of urban areas. Technical report, Brussels: European Commission.
- Henderson, J. V., 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(1), 357–406.
- Henderson, J. V. and M. A. Turner (2020, August). Urbanization in the developing world: Too early or too slow? *Journal of Economic Perspectives* 34(3), 150–73.
- Hodler, R. and P. A. Raschky (2014). Regional favoritism. *Quarterly Journal of Economics* 129(2), 995–1033.
- Jofre-Monseny, J., J. I. Silva, and J. Vázquez-Grenno (2020). Local labor market effects of public employment. *Regional Science and Urban Economics* 82, 103406. Local public policy evaluation.
- Kessler, A. S., N. A. Hansen, and C. Lessmann (2011, 03). Interregional redistribution and mobility in federations: A positive approach. *Review of Economic Studies* 78(4), 1345–1378.
- Kline, P. (2010). Place based policies, heterogeneity, and agglomeration. *American Economic Review* 100(2), 383–87.
- Kramon, E. and D. N. Posner (2011). Kenya’s new constitution. *Journal of Democracy* 22(2), 89–103.
- Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy* 99(3), 483–499.
- Law, G. (2010). *Administrative subdivisions of countries*. Jefferson, NC: McFarland & Company. The official reference for the Statoids.com database.
- Lessmann, C. (2014). Spatial inequality and development – Is there an inverted-U relationship? *Journal of Development Economics* 106, 35–51.
- Michaels, G. and F. Rauch (2018). Resetting the urban network: 117–2012. *Economic Journal* 128(608), 378–412.
- Miguel, E. and G. Roland (2011). The long-run impact of bombing Vietnam. *Journal of Development Economics* 96(1), 1–15.
- Montiel Olea, J. L. and M. Plagborg-Møller (2019). Simultaneous confidence bands: Theory, implementation, and an application to SVARs. *Journal of Applied Econometrics* 34(1), 1–17.
- Neumark, D. and H. Simpson (2015). Place-based policies. 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*, pp. 1197–1287. Elsevier.
- Nunn, N. and D. Puga (2012). Ruggedness: The blessing of bad geography in Africa. *Review of Economics and Statistics* 94(1), 20–36.



- Oates, W. E. (1972). *Fiscal federalism*. New York, NY: Harcourt Brace Janvonvich.
- OECD (2016). *OECD Economic Surveys: Indonesia 2016*.
- Roca, J. D. L. and D. Puga (2016, 07). Learning by Working in Big Cities. *Review of Economic Studies* 84(1), 106–142.
- Rosenthal, S. S. and W. C. Strange (2004). Evidence on the nature and sources of agglomeration economies. In *Handbook of Regional and Urban Economics*, Volume 4, pp. 2119–2171. Elsevier.
- Roth, J. (2021). Pre-test with caution: Event-study estimates after testing for parallel trends. *American Economic Review: Insights*. forthcoming.
- Rozenfeld, H. D., D. Rybski, X. Gabaix, and H. A. Makse (2011). The area and population of cities: New insights from a different perspective on cities. *American Economic Review* 101(5), 2205–25.
- Schmidheiny, K. and S. Siegloch (2019). On event study designs and distributed-lag models: Equivalence, generalization and practical implications. CESifo Working Paper Series 7481, CESifo Group Munich.
- Storeygard, A. (2016). Farther on down the road: Transport costs, trade and urban growth in sub-Saharan Africa. *Review of Economic Studies* 83(3), 1263–1295.
- Sun, L. and S. Abraham (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics* 225(2), 175–199.

# Online appendix

<b>A</b>	<b>Data Appendix</b>	<b>ii</b>
A-1	Remotely-sensed data . . . . .	ii
A-2	DHS data . . . . .	vii
A-3	Investment data . . . . .	xiii
A-4	Summary statistics . . . . .	xv
<b>B</b>	<b>Tracking capital cities and subnational units</b>	<b>xvi</b>
B-1	Administrative units over time . . . . .	xvi
B-2	Capital cities over time . . . . .	xviii
<b>C</b>	<b>Capital locations</b>	<b>xx</b>
<b>D</b>	<b>Selection issues: City detection</b>	<b>xxii</b>
<b>E</b>	<b>Additional results</b>	<b>xxiii</b>
E-1	Additional figures . . . . .	xxiii
E-2	Additional tables . . . . .	xxxi
<b>F</b>	<b>Former capitals and “mother” capitals</b>	<b>xxxvii</b>
F-1	Former capitals . . . . .	xxxvii
F-2	“Mother” capitals . . . . .	xxxix

## A. Data Appendix

### A-1. Remotely-sensed data

**Light density:** We calculate our light density measures by taking the average value per pixel within the year (to mitigate between multiple satellites) and then summing the average pixel values across our city shapes before dividing by the city area. The luminosity data is based on the raw NOAA data of the OLS-DMSP (stable light) product <https://sos.noaa.gov/catalog/datasets/nighttime-lights/>. The bottom correction is implemented following Storeygard (2016) and the top coding correction following Bluhm and Krause (2022).

**Population:** density within our cities is calculating by first taking the sum of population based on the Global Human Settlement Layer Raster (GHSL R2018A/2019A) and then dividing by the area of our cities (<https://ghsl.jrc.ec.europa.eu>).

**Ruggedness:** We calculate average ruggedness within 25km of our cities by taking the average pixel value of the ruggedness raster provided by Diego Puga (available at <https://diegopuga.org/data/rugged/> Nunn and Puga (2012)).

**Malaria suitability:** Malaria Ecology Index from Kiszewski et al. (2004) in raster format for GIS (<https://sites.google.com/site/gordoncmccord/datasets?authuser=0>).

**Market access:** Own calculation based on GHSL raster. See main text for details.

**River within 25km:** We generate a dummy for all cities located within 25km of river, based on our city coordinates and river shapes from Natural Earth 1:10m grid version 4.1.0 (<https://www.naturalearthdata.com/downloads/10m-raster-data/>).

**Lake within 25km:** We generate a dummy for all cities located within 25km of a lake, based on our city coordinates and “lake centerlines” from Natural Earth 1:10m grid version 4.1.0 (<https://www.naturalearthdata.com/downloads/10m-raster-data/>).

**Port within 25km:** We generate a dummy for all cities located within 25km of a port, based on our city coordinates and port locations obtained from the World Port Index 2010 (<https://msi.nga.mil/Publications/WPI>).

**Coast within 25km:** We generate a dummy for all cities located within 25km of the coast, based on our city coordinates and the coastlines from from Natural Earth 1:10m grid version 4.1.0 (<https://www.naturalearthdata.com/downloads/10m-raster-data/>).

**Precipitation:** Average precipitation is calculated within 25km buffers of our city coordinates, we average yearly values from Jan 1990 to Dec 2014 from the monthly totals. The precipitation data is obtained from NOA (Version 4.01 [https://psl.noaa.gov/data/gridded/data.UDel\\_AirT\\_Precip.html](https://psl.noaa.gov/data/gridded/data.UDel_AirT_Precip.html)).

**Elevation:** Average elevation within a 25km buffer of the city is calculated based on the SRTM Version 4.1 raster (Jarvis et al., 2008).

**Temperature:** Average temperature is calculated for 25km buffers around our cities. We use the average temperature from Jan 1990 to Dec 2014 from the monthly totals as inputs, which is obtained from NOA (Version 4.01 [https://psl.noaa.gov/data/gridded/data.UDel\\_AirT\\_Precip.html](https://psl.noaa.gov/data/gridded/data.UDel_AirT_Precip.html)).

**Wheat suitability:** Average wheat suitability is calculated for 25km buffers around our city coordinates. The wheat suitability values are obtained from the FAO GAEZ Agro-climatically attainable yield for intermediate input level rain-fed wheat for the baseline period 1961-1990 at 5 arc minutes.

**Built-up:** We calculate built-up and vegetation measures using the entire archive of Landsat images from 1987 until 2018, available at a resolution of 30m from Landsat 5 and 7 in *Google Earth Engine*. The measures are based on spectral bands, denoted by  $\rho_x$ , and calculated as follows:  $NDBI = (\rho_{SWIR1} - \rho_{NIR}) / (\rho_{SWIR1} + \rho_{NIR})$ ,  $UI = (\rho_{SWIR2} - \rho_{NIR}) / ((\rho_{SWIR2} + \rho_{NIR}))$ , and  $NDVI = (\rho_{NIR} - \rho_{red}) / (NIR + red)$ . Prior to calculation we extract the average values of cloud free images of the Landsat input as is standard in such calculations.

TABLE A-1  
Countries in sample

COUNTRY	LATE (schooling)	LATE (URBANIZATION)	LATE (GDP)
Afghanistan	1	1	1
Albania	1	1	1
Algeria	1	1	1
Angola	.	1	1
Argentina	0	0	0
Australia	0	0	0
Austria	0	0	0
Bangladesh	1	1	1
Belarus	.	1	.
Belgium	0	0	0
Benin	1	1	1
Bolivia	1	1	1

*Continued on next page*

Table A-1 – *Continued from previous page*

COUNTRY	LATE (schooling)	LATE (URBANIZATION)	LATE (GDP)
Brazil	1	1	1
Bulgaria	0	1	1
Burkina Faso	.	1	1
Burundi	1	1	1
Cambodia	1	1	1
Cameroon	1	1	1
Canada	0	0	0
Central African Republic	1	1	1
Chad	.	1	1
Chile	0	0	0
China	1	1	1
Colombia	1	1	1
Congo	1	1	1
Costa Rica	0	1	1
Cuba	0	0	1
Czech Republic	0	0	0
Cote d'Ivoire	1	1	1
Democratic Republic of the Congo	1	1	1
Denmark	0	0	0
Dominican Republic	1	1	1
Ecuador	1	1	1
Egypt	1	1	1
Eritrea	.	1	.
Estonia	0	0	.
Ethiopia	.	1	1
Finland	0	0	0
France	0	0	0
Georgia	.	0	.
Germany	0	0	0
Ghana	1	1	1
Greece	0	0	1
Guatemala	1	1	1
Guinea	.	1	1
Haiti	1	1	1
Honduras	1	1	1
Hungary	0	0	0
India	1	1	1
Indonesia	1	1	1
Iran	1	1	1
Iraq	1	1	1
Ireland	0	0	0
Italy	0	0	0
Japan	0	0	1
Kazakhstan	1	0	.
Kenya	1	1	1
Korea, North	.	1	1
Korea, South	0	1	1

*Continued on next page*

Table A-1 – *Continued from previous page*

COUNTRY	LATE (schooling)	LATE (URBANIZATION)	LATE (GDP)
Kyrgyzstan	0	1	.
Laos	1	1	1
Latvia	0	0	.
Lesotho	1	1	1
Liberia	1	1	1
Madagascar	.	1	1
Malawi	1	1	1
Malaysia	1	1	1
Mali	1	1	1
Mauritania	1	1	1
Mexico	1	0	0
Moldova	0	1	.
Mongolia	1	1	1
Morocco	1	1	1
Mozambique	1	1	1
Myanmar	1	1	1
Nepal	1	1	1
Netherlands	0	0	0
New Zealand	0	0	0
Nicaragua	1	1	1
Niger	1	1	1
Nigeria	.	1	1
Norway	0	0	0
Oman	.	1	1
Pakistan	1	1	1
Panama	0	1	1
Papa New Guinea	1	1	.
Paraguay	1	1	1
Peru	1	0	0
Philippines	1	1	1
Poland	0	0	0
Portugal	1	1	1
Romania	0	1	1
Russian Federation	0	0	.
Rwanda	1	1	1
Saudi Arabia	1	1	1
Senegal	1	1	1
Sierra Leone	1	1	1
Slovakia	0	1	.
Somalia	.	1	1
South Africa	0	0	0
South Sudan	1	1	1
Spain	0	0	1
Sri-Lanka	0	1	1
Sudan	1	1	1
Sweden	0	0	0
Switzerland	0	0	0

*Continued on next page*



Table A-1 – *Continued from previous page*

COUNTRY	LATE (schooling)	LATE (URBANIZATION)	LATE (GDP)
Syrian Arab Republic	1	1	0
Taiwan	0	1	1
Tajikistan	0	1	.
Thailand	1	1	1
Togo	1	1	1
Tunisia	1	1	1
Turkey	1	1	1
Turkmenistan	.	0	.
United Kingdom	0	0	0
Uganda	1	1	1
Ukraine	0	1	.
United Arab Emirates	1	0	0
United Republic of Tanzania	1	1	1
United States of America	0	0	0
Uruguay	0	0	0
Uzbekistan	.	1	.
Venezuela	1	0	0
Vietnam	1	1	1
Yemen	1	1	1
Zambia	1	1	1
Zimbabwe	1	1	1

*Notes:* The table depicts the countries covered in our study. The three LATE dummies refer to the country classification of [Henderson et al. \(2018\)](#).

## A-2. DHS data

**DHS wealth index:** is taken directly from the DHS surveys (v190). In general the DHS describes their wealth index as being: “... a composite measure of a household’s cumulative living standard. The wealth index is calculated using easy-to-collect data on a household’s ownership of selected assets, such as televisions and bicycles; materials used for housing construction; and types of water access and sanitation facilities” (<https://www.dhsprogram.com/topics/wealth-index/wealth-index-construction.cfm>). Note that the specific assets considered are country dependent.

**Electricity indicator:** is an indicator variable for the availability of electricity in the household (V119).

**Save water indicator:** is a indicator variables set to unity if the respondent household has access to either: protected wells or springs, boreholes, packaged water, and rainwater (v113) (see [Henderson et al., 2020](#), for a similar classification).

**Improved sanitation indicator:** is a indicator variable equaling unity if the respondent household has access to either shared or non-shared facilities that flush/pour to piped sewer systems, septic tank, pit latrine; ventilated improved pit latrine, pit latrine with slab and composting toilets, as well as flushing to unknown locations (v116). Again we follow [Henderson et al. \(2020\)](#) who also use the DHS-WHO joint monitoring program definitions.

**At least 8 years of schooling indicator:** Is a dummy variable that is unity if the respondent has completed 8 or more years of schooling (based on V107) and zero otherwise. It is only defined for respondents who are at least 16 years old.

**Infant mortality:** is defined as the probability of dying before the first birthday. The corresponding rate is normalized as a ratio per 1000 live births. The variable is constructed based on the “age at death” responses about the children of female respondents (variables b13-1 to b13-20). As common in the literature, we use the individual-child-level data to compute this measure and multiply the resulting dummy by 1000 to estimate a rate (“per thousand births”).

**Log household size:** is the log of the number of household members (v136).

**Female head of household indicator:** defined according to the reported gender of the household head (v151).

**Log head of household age:** is the log of age (in years) of the household head (v152).

**Household head completed primary education indicator:** is calculated based an the educational achievement variable (v149) if the respondent is the household head (v150). The indicator is unity if the household head has completed primary education or started but not finished secondary education (v149). The indicator is zero otherwise.

**Household head completed secondary education indicator:** is calculated based an the educational achievement variable (v149) if the respondent is the household head (v150). The indicator is unity if the household head has completed secondary education (v149). The indicator is zero otherwise.

**Household head completed higher education indicator:** is calculated based an the educational achievement variable (v149) if the respondent is the household head (v150). The indicator is unity if the household head has engaged in higher education (v149). The indicator is zero otherwise.

**Age** in years of the respondent (v012 and mv012) in the DHS. Also included as a squared term.

**Female** indicator taking unity for all respondents in the IR dataset of the DHS and zero for all respondents in the MR dataset of the DHS.

**Sex** indicator for respondents children, takes unity if the respondent child is female ( $b4_{01}-b4_{20}$ ).

**Multiple birth** indicator is unity if a respondents child was born either as a twin or multiple ( $b0_{01}-b0_{20}$ ).

**Period of birth indicator:** Indicator for the period of birth of the reported children (by decade, i.e, 1990s).

TABLE A-2  
DHS survey sample

ISO	INTERVIEW YEAR	RESPONDENTS	URBAN	FEMALE
AGO	2006	163	91	100
AGO	2007	78	100	100
AGO	2011	3548	93	100
ALB	2008	506	96	71
ARG	2008	171	100	73
BDI	2010	649	72	68
BDI	2011	1654	99	61
BDI	2012	1464	96	100
BDI	2013	32	100	100
BEN	1996	641	100	78
BEN	2001	1997	90	69
BEN	2011	1898	97	77
BEN	2012	2649	96	77
BFA	1992	2021	100	77
BFA	1993	1693	68	78
BFA	1998	222	100	66
BFA	1999	7	100	29
BFA	2003	2440	98	76
BFA	2010	3452	97	68
BOL	2008	7358	99	75
BRA	2008	23	100	74
CAF	1994	313	27	78
CAF	1995	238	46	82
CIV	1994	4771	88	76
CIV	1998	1169	100	79
CIV	1999	940	100	76
CIV	2011	1822	100	67
CIV	2012	2412	100	68
CMR	2004	5224	98	67
CMR	2011	7456	99	68
COD	2007	10028	98	69
COD	2013	5768	99	70
COL	2010	18276	99	100
DOM	2007	15403	96	52
DOM	2013	6292	96	51
EGY	1992	4713	83	100
EGY	1995	6595	74	100
EGY	2000	7228	78	100
EGY	2003	3452	81	100
EGY	2005	7661	75	100
EGY	2008	6975	71	100
GHA	1993	1354	88	80
GHA	1994	98	100	89
GHA	1998	930	96	75
GHA	1999	620	88	76
GHA	2003	2515	94	55

*Continued on next page*

Table A-2 – *Continued from previous page*

ISO	INTERVIEW YEAR	RESPONDENTS	URBAN	FEMALE
GHA	2008	2645	95	53
GHA	2013	35	100	66
GIN	1999	2149	97	75
GIN	2005	2004	95	65
GIN	2012	3275	98	69
HND	2011	2483	99	80
HTI	2000	3378	89	81
HTI	2006	2879	91	73
HTI	2007	109	100	56
HTI	2012	7020	95	63
IDN	2003	2385	98	77
KEN	2003	4002	95	82
KEN	2008	1087	93	72
KEN	2009	557	100	69
KGZ	2012	2210	81	80
LBR	2006	149	71	52
LBR	2007	2822	100	57
LBR	2008	792	100	100
LBR	2009	167	100	100
LBR	2011	778	100	100
LBR	2013	1673	100	71
LSO	2004	443	87	74
LSO	2005	89	100	73
LSO	2009	354	100	73
LSO	2010	206	100	73
MAR	2003	3854	99	100
MDA	2005	2969	100	76
MDG	1997	1216	96	100
MDG	2008	1906	95	70
MDG	2009	137	100	70
MDG	2011	95	34	100
MDG	2013	172	58	100
MLI	1995	1062	100	80
MLI	1996	1121	100	79
MLI	2001	2869	100	78
MLI	2006	3177	100	76
MLI	2012	2758	100	72
MLI	2013	505	87	67
MOZ	2009	3256	99	100
MOZ	2011	4946	99	77
MWI	2000	835	97	80
MWI	2004	270	100	75
MWI	2005	345	100	74
MWI	2010	1211	100	74
MWI	2012	828	100	100
NER	1992	2201	47	81
NER	1998	1195	92	69
NGA	2003	3168	87	75

*Continued on next page*

Table A-2 – *Continued from previous page*

ISO	INTERVIEW YEAR	RESPONDENTS	URBAN	FEMALE
NGA	2008	10456	89	67
NGA	2010	1481	85	100
NGA	2013	13150	90	68
PAK	2006	2462	93	100
PER	2000	10256	99	100
PER	2004	3794	100	100
PER	2009	8710	99	100
PHL	2003	5862	96	76
PHL	2008	3241	97	100
RWA	2005	1128	94	69
RWA	2008	1174	96	47
RWA	2010	197	31	69
RWA	2011	1532	92	68
SEN	1992	817	82	80
SEN	1993	1584	80	81
SEN	1997	3752	94	67
SEN	2005	11160	88	77
SEN	2008	8782	85	100
SEN	2009	1424	89	100
SEN	2010	2831	87	75
SEN	2011	2940	92	75
SEN	2012	998	93	100
SEN	2013	1080	88	100
SLE	2008	2799	99	70
SLE	2013	4732	100	69
TGO	1998	3441	93	70
TGO	2013	2328	98	70
TJK	2012	2096	90	100
TZA	1999	1395	100	54
TZA	2003	331	100	100
TZA	2004	2053	97	94
TZA	2007	952	95	100
TZA	2008	844	96	100
TZA	2009	408	100	82
TZA	2010	1227	98	81
TZA	2011	463	96	100
TZA	2012	1941	95	100
UGA	2000	473	94	77
UGA	2001	414	100	81
UGA	2006	1231	100	79
UGA	2008	31	100	77
UGA	2009	860	94	100
UGA	2011	6756	94	89
ZMB	2007	2416	100	52
ZMB	2013	4613	96	53
ZWE	1999	1431	100	68
ZWE	2005	2729	98	56
ZWE	2006	228	89	53

*Continued on next page*



Table A-2 – *Continued from previous page*

ISO	INTERVIEW YEAR	RESPONDENTS	URBAN	FEMALE
ZWE	2010	1905	97	56
ZWE	2011	1018	100	58

*Notes:* The table depicts the DHS survey included in our sample. The survey years the number of respondents in each survey that we can match to our data as well as the percentage of urban dwellers and female respondents within each DHS survey.

### A-3. Investment data

**Development aid (World Bank):** Development aid provided by the World Bank is obtained [AidData \(2017\)](#). This geocoded dataset includes all projects approved from 1995-2014 in the World Bank IBRD/IDA lending lines. It tracks more than \$630 billion in commitments for 5,684 projects across 61,243 locations. We construct several aid variables following the sectoral classification. The sectoral classification are in order; Education, health, water supply & sanitation, government and civil society, other social infrastructure & services, economic infrastructure and services, agriculture forestry and fishing, industry and mining and construction, and environmental protection. They correspond to the broadest classification of the project types provided by the World Bank. Note that any project can have multiple (up to 5) project classifications. In such cases the same project appears under multiple headings.

**Development aid (China):** Development aid like financial flows for China are obtained from AidData's Geocoded Global Chinese Official Finance Dataset, Version 1.1.1. ([Bluhm et al., 2020](#)). This dataset geolocates Chinese Government-financed projects that were implemented between 2000-2014. It captures 3,485 projects worth \$273.6 billion in total official financing. The dataset includes both Chinese aid and non-concessional official financing. We construct several aid variables following the sectoral classification. The sectoral classification are in order; Education, health, water supply & sanitation, government and civil society, other social infrastructure & services, economic infrastructure and services, agriculture forestry and fishing, industry and mining and construction, and environmental protection. They correspond to the broadest classification of the project types provided by the World Bank. Note that any project can have multiple (up to 5) project classifications. In such cases the same project appears under multiple headings.

**FDI:** The raw data for our FDI outcomes (dummy, log investment value, and log estimated jobs) comes from the fDi Markets database (<https://www.fdimarkets.com>) a service provided by the Financial Times group. The database contains in detail information on FDI projects across the world for the period 2003 until 2018, including information about the investing company the origin country the company is based and much more. Important for us the database has the estimated jobs created the value spend, the host city name and if the project is a greenfield investment. We geocoded the projects using the same OSM algorithm we employed for the location of the capital cities using the host city information. In a next step we match the FDI to our cities if the projects host city (which do not need to meet any population threshold) fall within a 10km buffer of our detected cities. Finally, we summarize the invested dollar value and the estimated jobs by the host city location and take the logs of them. Note that we only gathered data for our

reformed areas, since the terms of use only allow us to use 10% of their sample. The data is then aggregated to the NAICS 2 digit level. The 2 digits NAICS classification we use are in order: Agriculture, Forestry, Fishing and Hunting; Mining, Quarrying, and Oil and Gas Extraction; Utilities ; Construction; Manufacturing; Wholesale Trade; Retail Trade; Transportation and Warehousing; Information; Finance and Insurance; Real Estate and Rental and Leasing; Professional, Scientific, and Technical Services; Administrative and Support and Waste Management and Remediation Services; Educational Services; Health Care and Social Assistance; Arts, Entertainment, and Recreation; Accommodation and Food Services; Public Administration.

## A-4. Summary statistics

TABLE A-3  
Summary statistics: Fundamentals

	Mean	SD	Min	Max	N
<i>Panel A. Cities (all)</i>					
Log light density	2.95	1.29	1.26	7.65	515,934
Log population 1990	10.84	0.88	9.25	17.06	515,934
Ruggedness	14.49	15.43	0.46	120.22	515,934
Malaria suitability	0.01	0.02	0.00	0.17	515,934
Market access (pop 1990 based)	10.32	1.30	3.46	13.55	515,934
River within 25km	0.35	0.48	0.00	1.00	515,934
Lake within 25km	0.02	0.14	0.00	1.00	515,934
Port within 25km	0.05	0.22	0.00	1.00	515,934
Coast within 25km	0.16	0.37	0.00	1.00	515,934
Distance to coast	371.14	362.96	2.57	2,504.02	515,934
Average precipitation	9.29	5.38	0.05	81.39	515,934
Average elevation	458.67	577.53	-26.41	5,023.05	515,934
Average temperature	19.94	6.89	-7.59	32.09	515,934
Wheat suitability	2,296.89	2,074.38	0.00	7,252.34	515,934
<i>Panel B. Cities (within reformed areas)</i>					
Log light density	2.34	1.12	1.26	7.51	182,048
Log population 1990	10.80	0.83	9.90	16.80	182,048
Ruggedness	13.50	15.61	0.53	110.43	182,048
Malaria suitability	0.01	0.03	0.00	0.16	182,048
Market access (pop 1990 based)	10.61	1.31	3.48	13.55	182,048
River within 25km	0.38	0.49	0.00	1.00	182,048
Lake within 25km	0.02	0.13	0.00	1.00	182,048
Port within 25km	0.03	0.16	0.00	1.00	182,048
Coast within 25km	0.11	0.31	0.00	1.00	182,048
Distance to coast	480.58	378.24	2.57	2,188.57	182,048
Average precipitation	9.77	4.53	0.05	75.78	182,048
Average elevation	486.82	600.83	-25.44	5,023.05	182,048
Average temperature	22.24	5.67	-5.49	30.60	182,048
Wheat suitability	1,982.01	1,760.43	0.00	6,886.30	182,048

*Notes:* Panel A of the table reports the summary statistics for our sample of all cities. Panel B reports summary statistics for the sample of cities located within reformed regions.

## B. Tracking capital cities and subnational units

We separately track changes in the geography of subnational units and capitals over time, and cross-reference both results at the end to minimize the scope for error. We start cataloging subnational capitals using the two most comprehensive databases available today (i.e., the Statoids database, [Law, 2010](#) and the City Population database, [Brinkhoff, 2020](#)). We use the Global Administrative Unit Layers (GAUL) vector data as a baseline to track subnational units over time, which only records the spatial extent of administrative units but contains no information on their capitals. The three databases have varying temporal coverage. Statoids often tracks capitals and subnational units back to the founding of a country and is usually accurate (up until 2013/2014), but lacks any spatial information. City Population and GAUL cover short time periods, from 1998 until 2020 and 1990 until 2014, respectively.

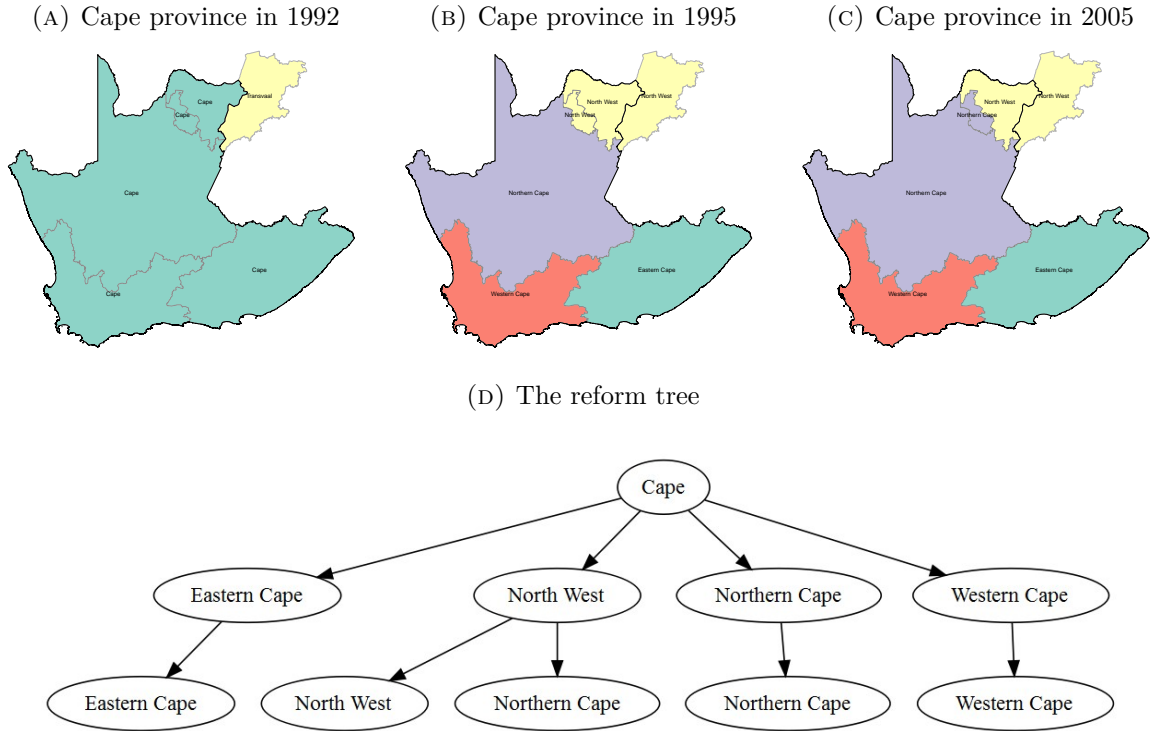
### B-1. Administrative units over time

We begin by backing out a reform tree from the GAUL data using a simple spatial algorithm. For any pair of two years, we create the spatial intersection of the two vector data sets. This creates new areas or new affiliations whenever a border is moved, deleted or created. We then cycle forward by intersecting the result of the previous intersection with the next year of official data and so on. During each iteration, we also record the current region identifier and add it to an identification string which in the last year contains 24 (i.e. 2014 – 1990) identifiers.

We obtain two data sets in this manner. The first is a spatial data set of micro-regions, which in the final year contains the smallest spatial unit whose borders were *not reformed* in any of the preceding years. We call this unit a splinter. The second is a kind of evolutionary tree for each contemporary splinter, summarizing its entire history of regional affiliations and its respective administrative center back until 1990. Note that splinters only result from border reforms that cut across borders from the previous year. If borders are simply abolished, no new splinter will be created but the identity of the region changes. Hence, the combination of the spatial splinter data set and the reform tree identifies all administrative reforms in a general and spatially consistent manner. Moreover, the reform tree allows us to easily compare the results to other non-spatial data sources, such as City Population or Statoids.

[Figure B-1](#) provides an illustration of the two data sets created by this process. It shows the reform history of Cape Province in South Africa from 1992 onward (the green area in panel A). In 1994, the Cape Province was split into four new regions (panel B). Three of the successor provinces are congruent with the former province, while the fourth region (North-West) includes some areas of the former Transvaal (the neighboring province to the north east, marked in yellow in panel A). Furthermore, a part of the

FIGURE B-1  
Reform History of Cape Province, South Africa



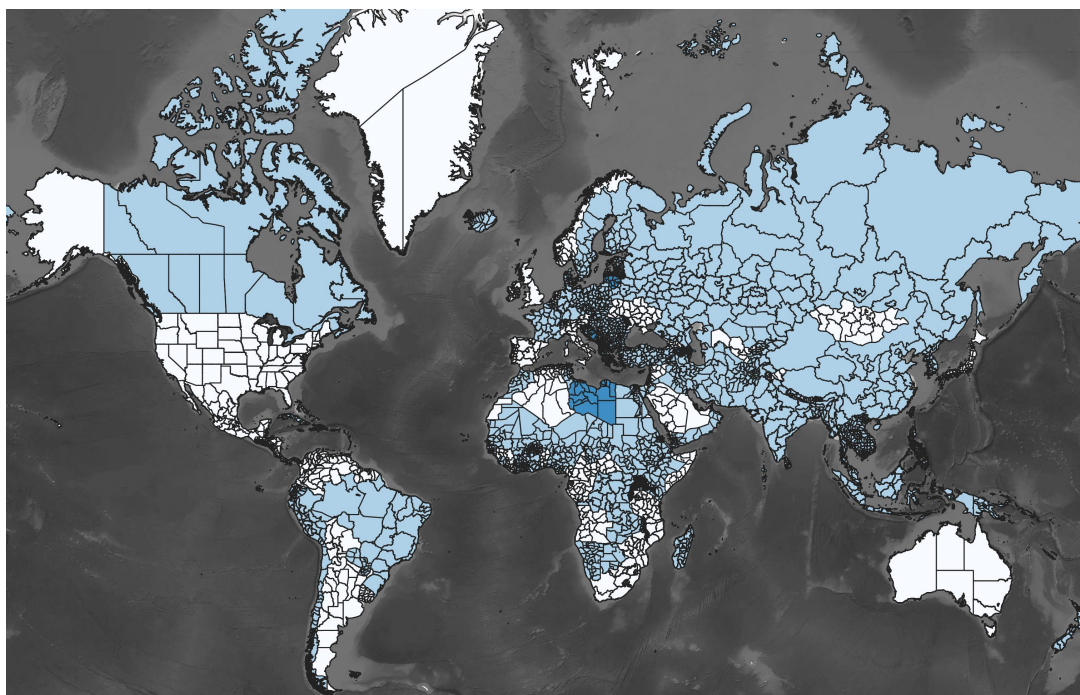
*Notes:* Panel a to c illustrate initial and successor regions of the Cape Province in South Africa. Panel d) illustrates the evolutionary tree for the splinters which were formerly part of Cape Province, South Africa. The last level represents the situation after the 2005 reform.

North-West was assigned to the Northern Cape in 2005 (see the yellow area in panel B, which turns purple in Panel C). As a result, all splinters of Cape Province are affiliated with at least two different administrative regions over this period (panel D).

Next, we compare the resulting reform tree with Statoids and City Population to document discrepancies (of which there are many). First, the different sources do not always agree on what unit constitutes the first-order administrative level. GAUL sometimes contain macro regions, which have no political function and are easily identified using other data sources. Whenever we detect a case in which GAUL seems to disagree with other sources, or misses a reform entirely, we collect additional spatial data for these regions. From 2000 onward, AidData's GeoBoundaries database and GADM provide a lot of high quality data, although neither of them is without error. Data for the early 1990s is harder to obtain and sometimes requires us to digitize offline maps. In rare cases, we were able to recover the correct shapes merging regions. Uganda, for example, consecutively split its larger regions into smaller units, so that the most recent vector data was sufficient to reconstruct an administrative map for each year. In summary, we found that around 40% of all countries in GAUL had missing or incomplete data during the period from 1990 to 2014 (see [Figure B-2](#) for an illustration).



FIGURE B-2  
Corrections made in GAUL data from 1990-2014



*Notes:* The figure plots the corrected GAUL countries. Countries in white are correct in GAUL, bright blue are those we had to fix, dark blue are those we are unable to fix, because we lack correct maps for one or more years during the sample period.

Finally, we extended the corrected sample to full period from 1987 to 2018. Extending the sample from 2014 onward is straightforward, since many statistical offices upload official vector files and we could use newer version of AidData's GeoBoundaries database and GADM to fill in the gaps. Extending backward from 1989 to 1987 was more cumbersome. We relied mostly the 1980s and early 1990s editions of the Atlas Britannica.

## B-2. Capital cities over time

This workflow starts out with two lists of capital city-years obtained from Statoids and City Population. The lists were provided to two trained coders who independently cross-referenced and checked each entry for inconsistencies. The coders resolved any differences using additional data sources such as the CIA Factbook, Wikipedia, or secondary literature. A third coder compared these two sets of results and resolved differences, if there were any, in a final arbitration process.

Next the two expert coders geocoded the locations of all administrative cities, i.e., the longitude and latitude of the city centroids using the OpenStreetMap's (OSM) Nominatim API and the Google Maps' geocoding API. OSM and Google accurately identified the coordinates of most cities without any problems. Unfortunately, not all cities were coded automatically and some cities were not coded correctly. In those cases, we manually identified the coordinates of the city. In Uganda, for example, we had to manually geocode

around 60 out of 136 administrative centers. The manual coding included another arbitration layer in case of disagreements.

Finally, we merge the remotely-sensed universe of urban clusters in 1990 and 2015 with the coordinates of administrative cities. We consider exact matches all cases where the centroid of a capital city falls within 3 km of an urban cluster. In the few instances where no urban cluster is within this distance of an administrative center, we proceed by matching on names. Any cluster within 50 km of a capital city with almost the same name, defined as a Levenshtein edit distance of less than 3, is considered a match.

## C. Capital locations

We now take a closer look at the political geography determinants of capital locations within regions and provide some descriptives on which cities are likely to become capitals within a new administrative region.

We take our inspiration from [Bai and Jia \(2021\)](#), who propose that central government planners in historical China face a trade-off when determining the location of regional capital cities. Being close to citizens implies that the administrative location can efficiently exercise control (levy taxes and provide services at a low cost). Proximity to the national capital, in turn, makes the local administration more accountable to the national government and minimizes the cost of delivering local taxes to the central government (for similar arguments see [Bardhan and Mookherjee, 2000](#); [Campante and Do, 2014](#)). Note that the argument holds for all locations of national capital, which might be isolated itself to put distance between autocratic rulers and their populations ([Campante et al., 2019](#)). The optimal solution to this problem minimizes a location’s ‘hierarchical distance’: the distance to all citizens within a province and the national capital (with some weight on either objective). Of course, other factors are likely to play a role in these location decisions today, which is why we consider a range of additional variables from proximity to the coast to the size distribution of cities in the initial region.

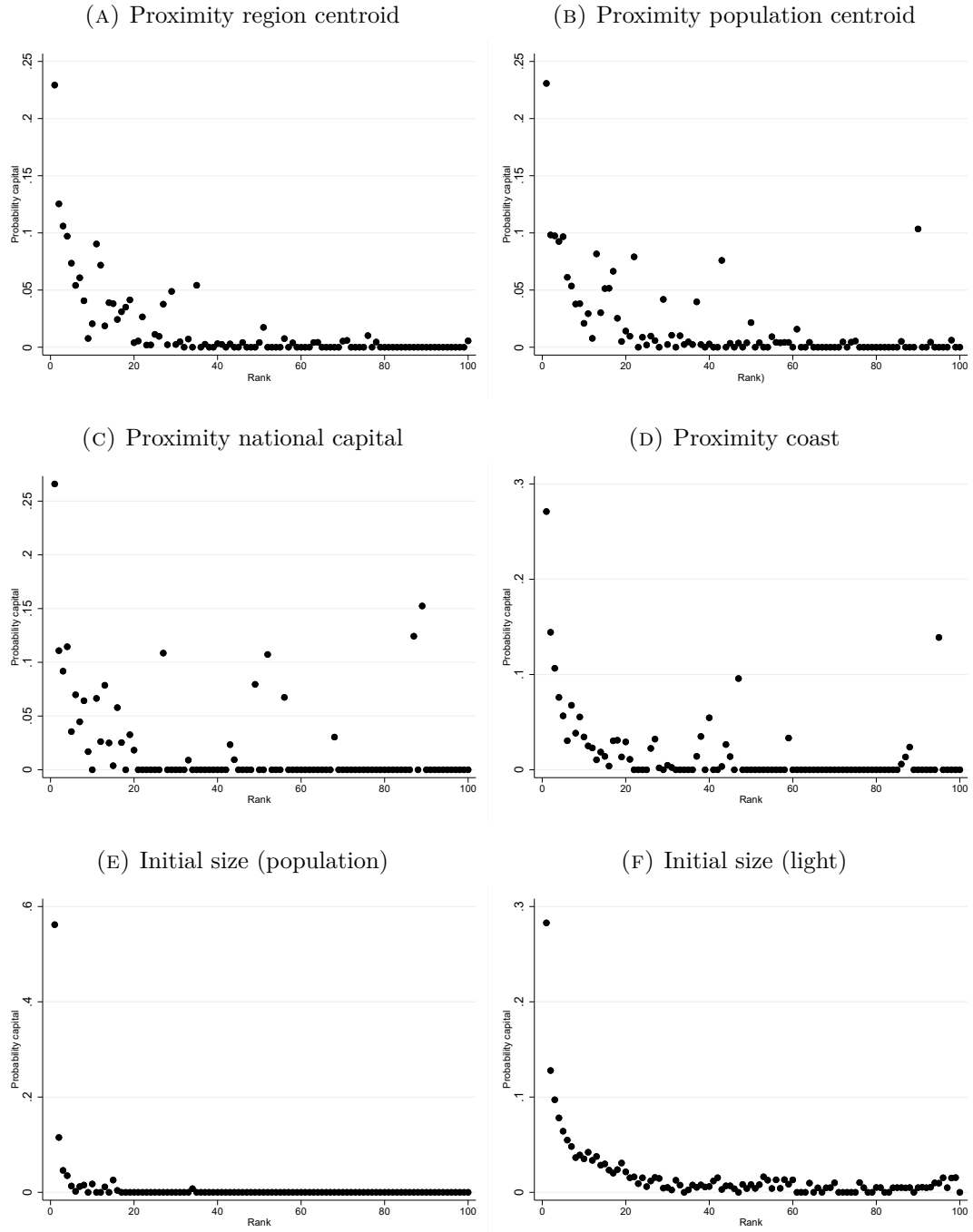
Panels A to C of [Figure C-1](#) provide some evidence in favor of the idea that hierarchical distance also matters in our global sample of contemporary capital city reforms. We rank cities within regions with respect to their distance to the region centroid in panel A, their distance to the population-weighted centroid in panel B, or their distance to the national capital in panel C. In all three cases, cities that occupy lower ranks (are closer) are considerably more likely to become a capital when a region is split. Panel D adds the proximity to the coast as a proxy for the external trade orientation and documents a similar pattern. We find a few outliers where high ranks have a high probability of becoming a capital (due to a few regions in South Asia with relatively “remote” capitals).

Finally, we examine initial size, either based on population or light density, as a predictor of gaining the status as a regional capital.<sup>1</sup> Panel E shows a strong relationship between the initial size of a city and the probability of becoming the region’s capital. The largest city in a region is also the region’s capital in almost 60% of the cases, the second-largest city in around 17% of cases, while the chances of being a capital for the third and fourth-largest cities are in the single digits. Cities that rank five or higher have an average probability below 1%. The relationship wakens if we rank by initial light, where the decline in the probability is smooth, and the largest city becomes the capital in only 30% of cases (panel F).

---

<sup>1</sup>Note that the largest city does not minimize the distance to all citizens by definition, although there is a high correlation 0.64.

FIGURE C-1  
Determinants of capital locations within regions: City ranks



*Notes:* This figure shows scatter plots of the average probability that a city becomes a capital across over the distribution of city characteristics along various dimensions. Panel A ranks cities in terms of proximity to the regional centroid. Panel B ranks cities with respect to proximity to the population-weighted centroid of a region. Panel C uses the proximity to the national capital and panel D the proximity to coast. Panels E and F rank cities based on their initial size, either by population or light density.

## D. Selection issues: City detection

Throughout the main text, we focus on the cities that we were able to detect in 1990. We then analyze changes in the core and in the larger agglomeration, including new developments in these cities from 1990 until 2015. Defining the sample of cities avoids a sample selection issue that we illustrate in more detail in this appendix.

The selection effect arises since the status of a city as a subnational capital also influences the detection likelihood in 2015. Our main result is that cities grow faster once they gain capital city status. Recall that we only observe urban boundaries at two points in time (1990 and 2015). If a small city becomes a subnational capital in the interim and grows faster as a result, it is more likely to cross our detection thresholds and classified as an urban cluster (or city) in 2015. Suppose we track light density (or other outcomes) in these cities over the entire period, even though they are only detected later. In that case, we include this dynamic selection bias and, with that, the possibility of pre-trends.

We design a simple test to illustrate this selection effect. We regress the change in detection status from 1990 to 2015 on the share of years a city is a subnational capital during the same period. The change in status is the first difference of a binary variable indicating whether a city was detected in a particular year in the union of urban clusters found in either in 1990 or 2015. [Table D-1](#) reports the results from several specifications, where we incrementally add country and initial-region fixed effects for our two samples. Columns 1 to 3 show that a city that becomes a capital halfway through the period from 1990 to 2015 has a 7.3 to 11.8 percentage points higher probability of being detected in 2015. The estimated effect sizes are smaller for the sample of cities in reformed regions, but the overall pattern remains the same. Obtaining the status as a first-order capital during the sample significantly increases the likelihood of detection in 2015.

TABLE D-1  
City detection probability

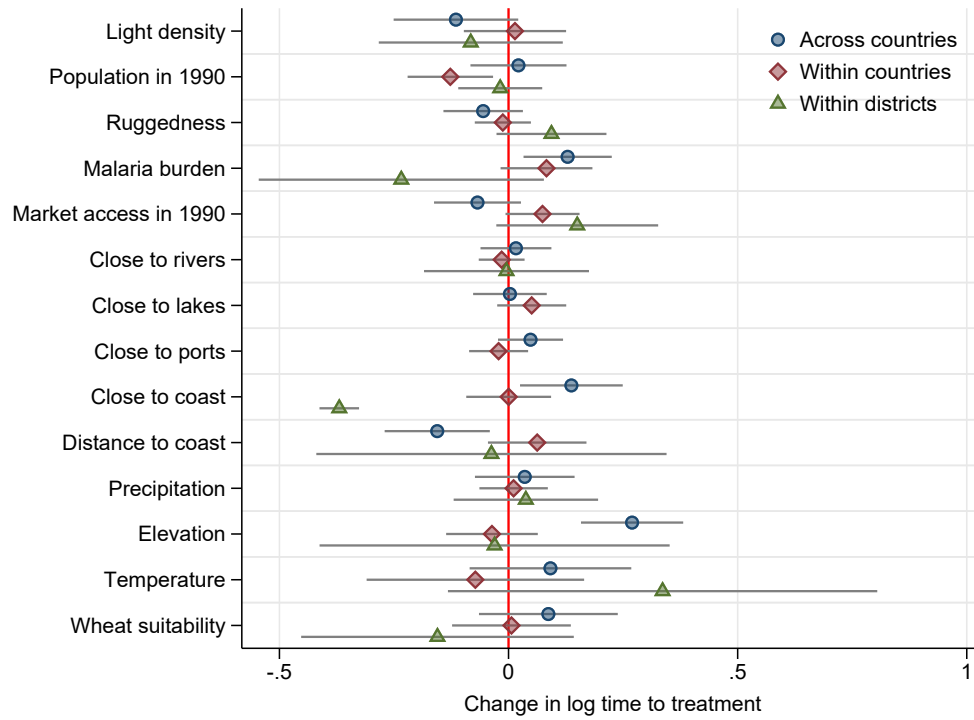
	<i>Dependent Variable: <math>\Delta \text{DETECTED}_{ci}</math></i>					
	All Cities			Reformed Regions		
	(1)	(2)	(3)	(4)	(5)	(6)
CAPITAL	0.1401 (0.0396)	0.2184 (0.0404)	0.2464 (0.0475)	0.1478 (0.0583)	0.2269 (0.0635)	0.2672 (0.0736)
Fundamentals	✓	✓	✓	✓	✓	✓
Country FE	—	✓	✓	—	✓	✓
Initial-Region FE	—	—	✓	—	—	✓
City-unions	27906	27906	27906	10213	10213	10213

*Notes:* The table reports results from a regressions of the change in detection status of a city between 1990 and 2015 on the fraction of years in which a city is a capital. Standard errors clustered on initial regions are provided in parentheses.

## E. Additional results

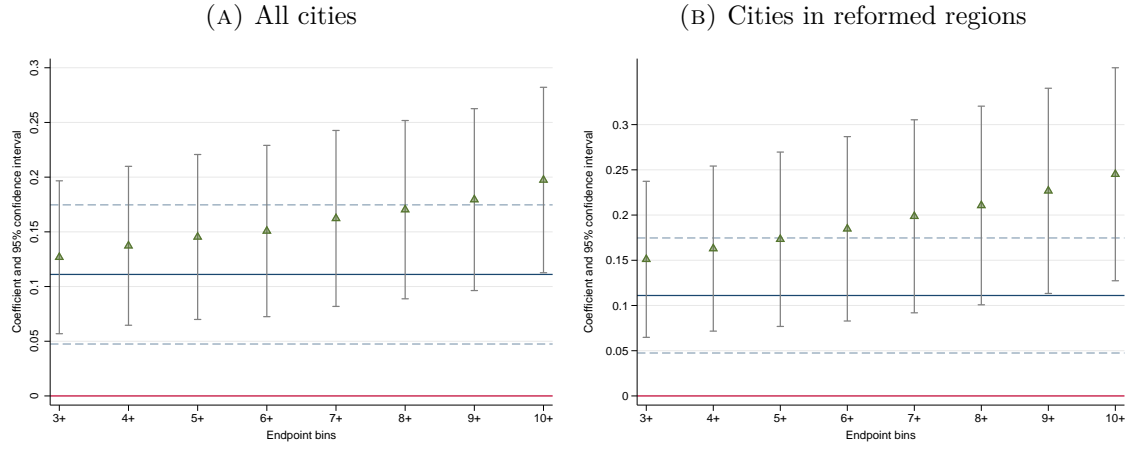
### E-1. Additional figures

FIGURE E-1  
Time to treatment



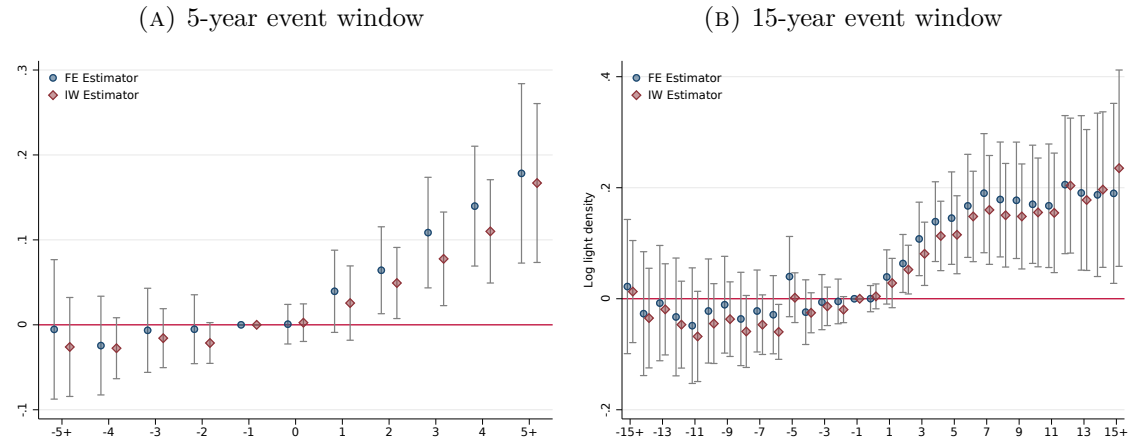
*Notes:* The figure illustrates results from cross-sectional regressions of the time to treatment (in logs plus one) on initial city characteristics. The regressions were run three times, once without fixed effects, with country fixed effects, and with initial region fixed effects. The coefficients are standardized beta coefficients. Some coefficients are omitted in the specification with initial region fixed effect for a lack of within region variation. 95% confidence intervals clustered on initial regions are indicated by the error bars.

FIGURE E-2  
Endpoint binning and medium-run effect size: Event-study estimates



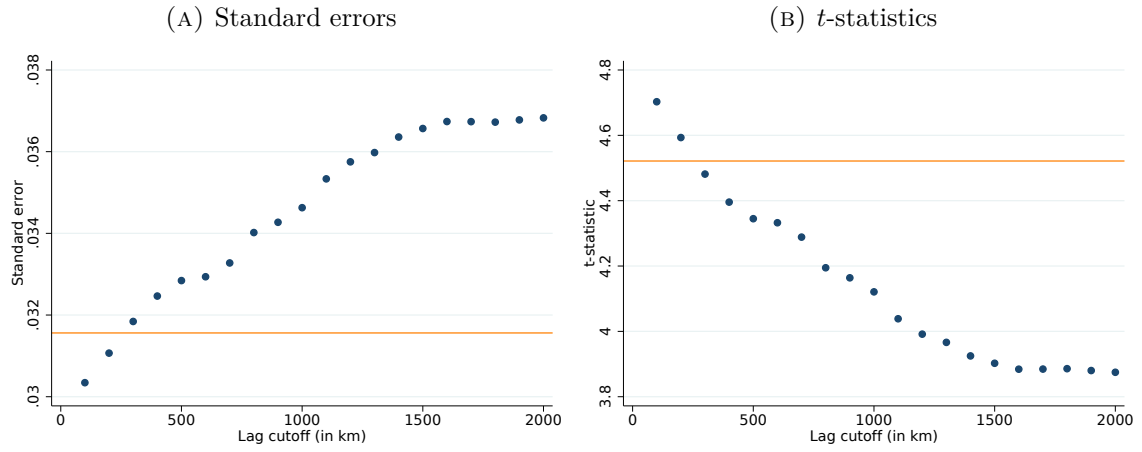
*Notes:* The figure shows point coefficients and 95% confidence intervals of the endpoint bins estimated in several event studies with varying window sizes. The underlying event studies use five pre-treatment periods and extend the event window from 3 (or more) to 10 (or more) periods. The effect in the last pre-period is normalized to zero. Panel A is based on column 3 and panel B is based on column 6 of [Table E-1](#). The blue line indicates the difference-in-differences estimate corresponding to each panel and the dashed blue lines provide the 95% confidence intervals of these estimates.

FIGURE E-3  
TWFE versus IW estimator of dynamic treatment effects



*Notes:* The figure illustrates event-study results from fixed effects regressions of the log of light intensity per square kilometer on a binned sequence of treatment change dummies, city fixed effects, initial-region-by-year fixed effects, time-varying locational fundamentals for a panel that is balanced in calendar time. Circles represent point estimates from two-way fixed effects estimation (TWFE). Diamonds represent point estimates from interaction-weighted (IW) estimation (see [Sun and Abraham, 2021](#)). Panel A shows estimates a five-year event window. Panel A shows estimates a 15-year event window. 95% confidence intervals based on standard clustered on initial regions are provided by the gray error bars.

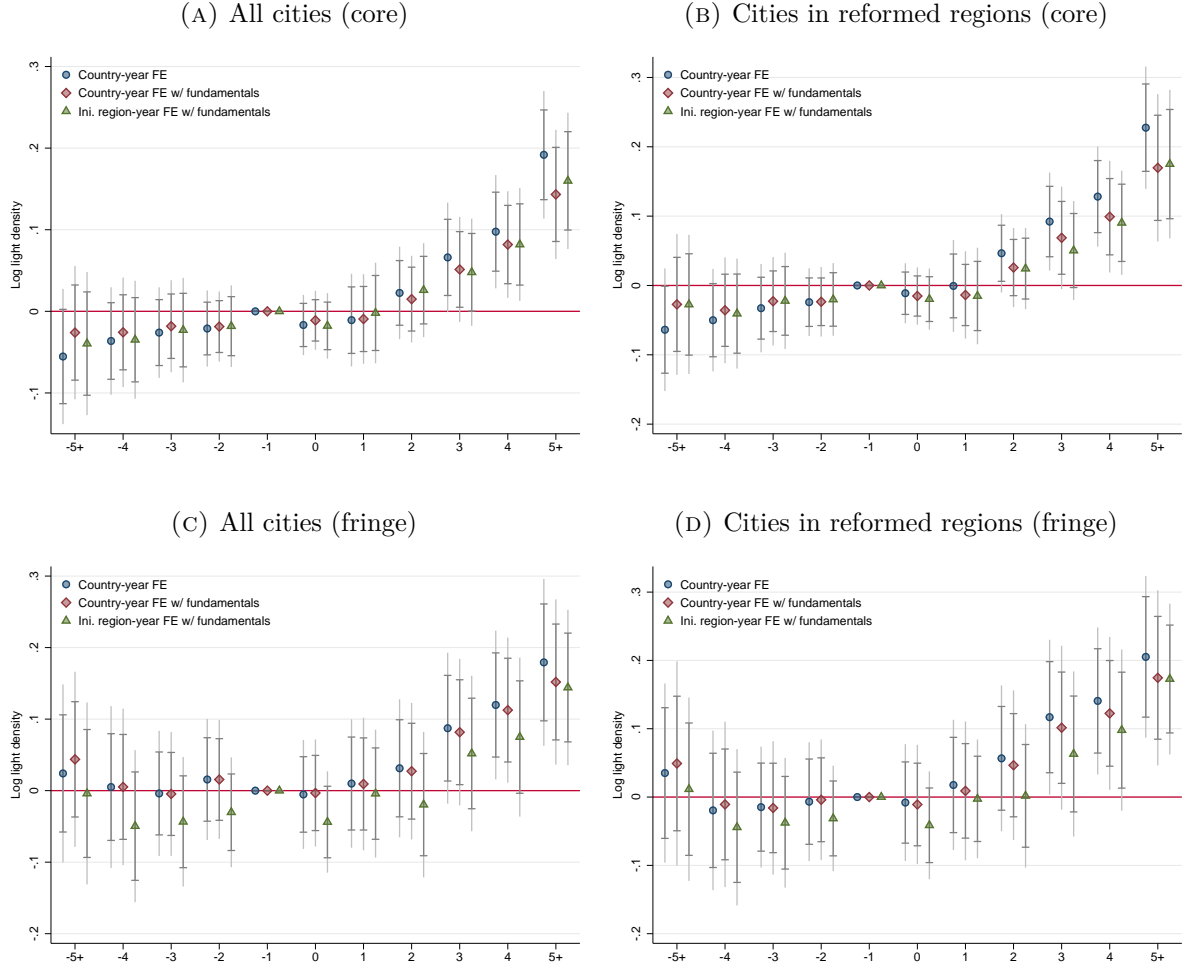
FIGURE E-4  
Accounting for spatial autocorrelation



*Notes:* The figure illustrates results from varying the spatial lag cutoff when estimating standard errors which allow for cross-sectional dependence. All results are based on a variant of column 6 in [Table E-1](#) where we restrict the sample to reformed areas and include city fixed effects, as well as initial-region fixed effects. Here we omit the time-varying effects of the fundamentals for computational reasons (to reduce the size of the regressor matrix small). The estimated effect in this specification is 0.1427 with a standard error of 0.0316. All Conley errors are estimated with a uniform kernel and a time-series HAC with a cutoff of 1,000 years to allow for arbitrary dependence over time. Panel A shows estimates of the resulting standard errors, with the original error clustered on initial regions highlighted in orange. Panel B shows estimates of the resulting  $t$ -statistics, with the original  $t$ -statistic clustered on initial regions highlighted in orange.

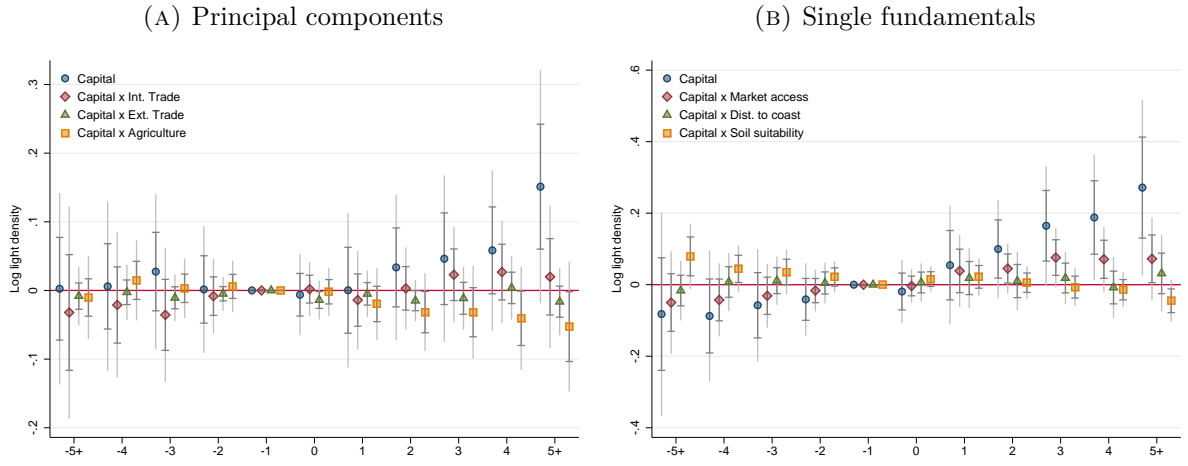


FIGURE E-5  
Agglomerations: Event-study estimates



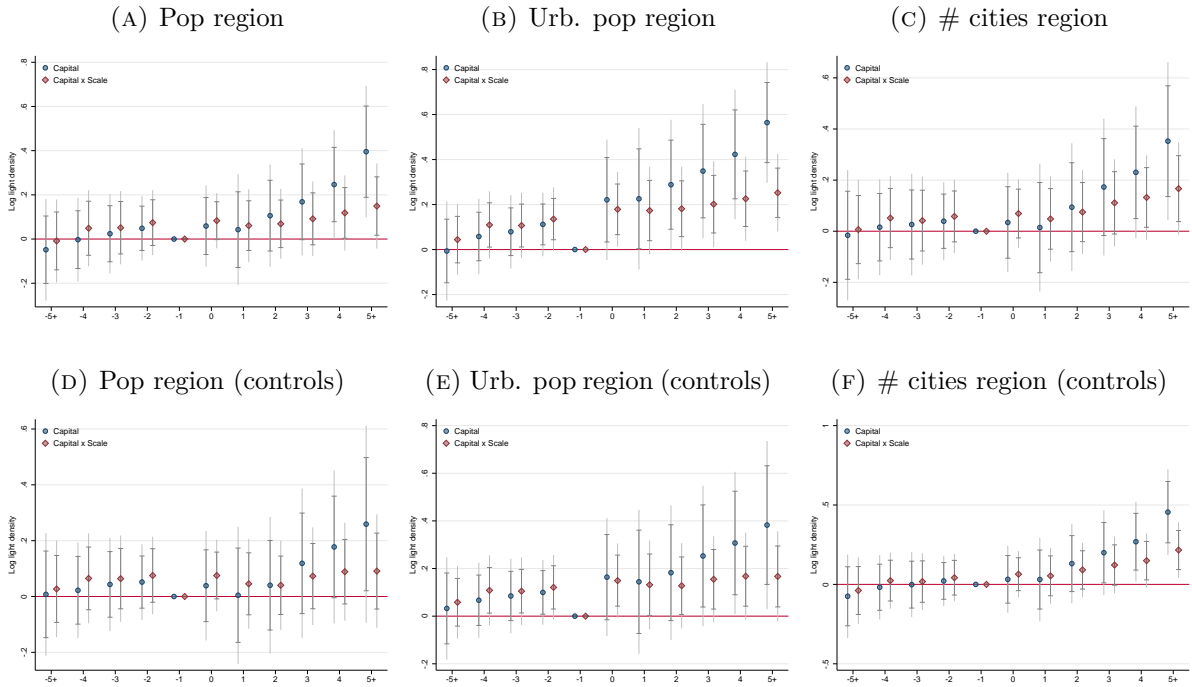
*Notes:* The figure reports event-study estimates corresponding to the difference-in-differences results presented in Table II. The upper panels report results for the core of the agglomeration. The lower panels report results for the fringe (new parts that were added after 1990). Panels A and C show estimates for all cities. Panels B and D show estimates for cities in reformed regions. Circles represent point estimates from a regression with city and country-year fixed effects, diamonds represent specifications with additional controls for locational fundamentals, and triangles represent specifications with initial-region-by-year fixed effects in addition. All regressions include city fixed effects. 95% confidence intervals based on standard clustered on initial regions are provided by the gray error bars. The orange error bars indicate 95% sup- $t$  bootstrap confidence bands with block sampling over initial regions (Montiel Olea and Plagborg-Møller, 2019).

FIGURE E-6  
Fundamentals: Event-study estimates



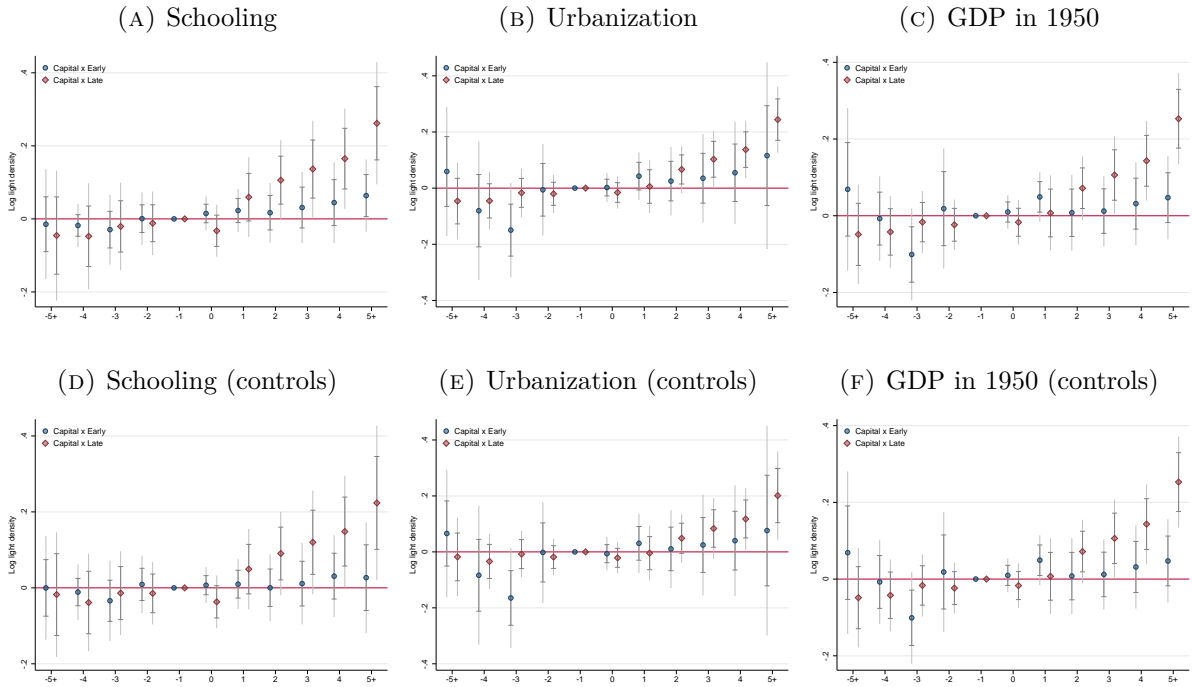
*Notes:* The figure reports event-study estimates corresponding to the difference-in-differences results presented in Table IV. Panel A reports estimates corresponding to column 5 of panel A, whereas panel B reports the estimates corresponding to column 5 of panel B of the table. All regressions include city fixed effects and initial-region-by-year fixed effects. 95% confidence intervals based on standard clustered on initial regions are provided by the gray error bars. The orange error bars indicate 95% sup-*t* bootstrap confidence bands with block sampling over initial regions (Montiel Olea and Plagborg-Møller, 2019).

FIGURE E-7  
Scale: Event-study estimates



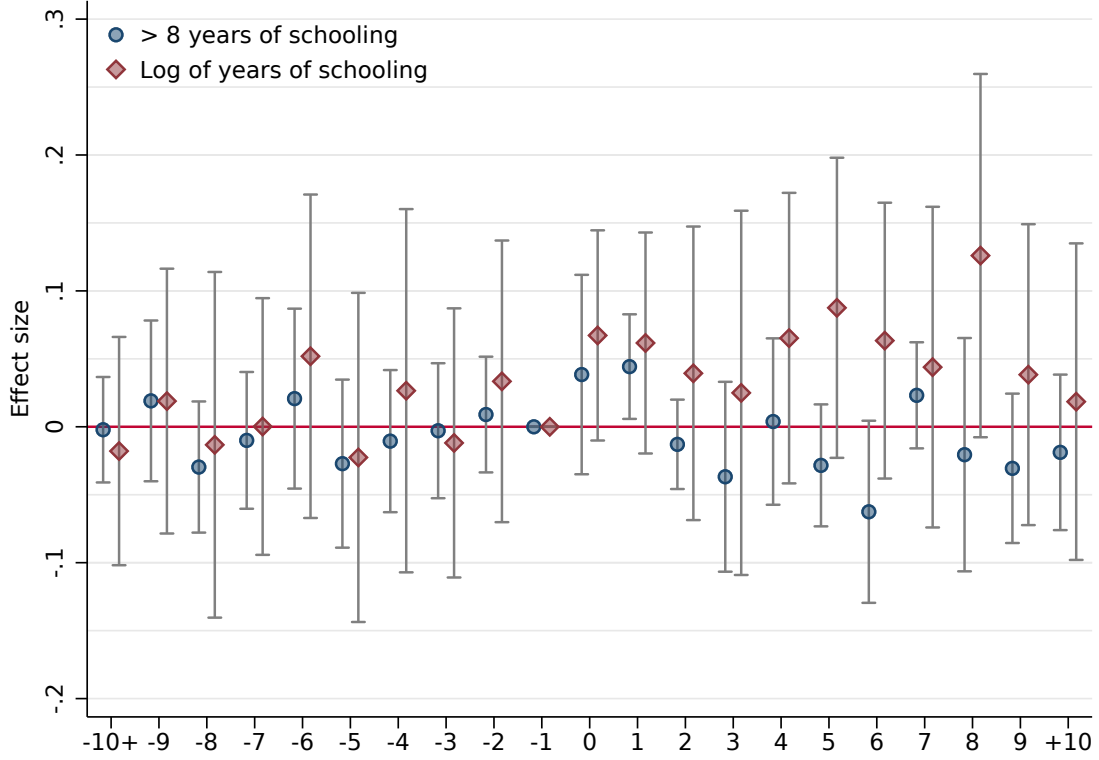
*Notes:* The figure reports event-study estimates corresponding to the difference-in-differences results presented in [Table V](#). Panels A to C report the event studies without controls (corresponding to columns 1, 3 and 5 in the table). Panel D to F report the event studies including controls (corresponding to columns 2, 4 and 6). All regressions include city fixed effects and initial-region-by-year fixed effects. 95% confidence intervals based on standard clustered on initial regions are provided by the gray error bars. The orange error bars indicate 95% sup- $t$  bootstrap confidence bands with block sampling over initial regions ([Montiel Olea and Plagborg-Møller, 2019](#)).

FIGURE E-8  
Early vs. late: Event-study estimates



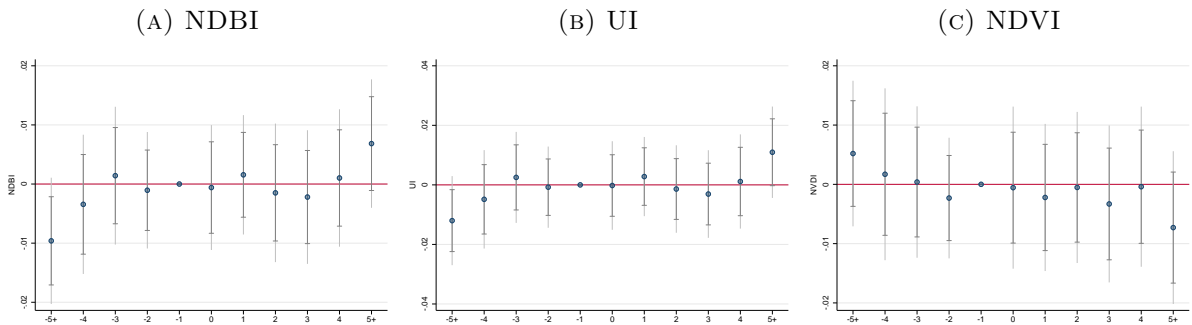
*Notes:* The figure reports event-study estimates corresponding to the difference-in-differences results presented in [Table VI](#). Panels A to C report the event studies without controls (corresponding to columns 1, 3 and 5 of the table). Panels D to F report the event studies including controls (corresponding to columns 2, 4 and 6). All regressions include city fixed effects and initial-region-by-year fixed effects. 95% confidence intervals based on standard clustered on initial regions are provided by the gray error bars. The orange error bars indicate 95% sup- $t$  bootstrap confidence bands with block sampling over initial regions ([Montiel Olea and Plagborg-Møller, 2019](#)).

FIGURE E-9  
Selective migration: Within city evidence (long window)



*Notes:* The figure illustrates event-study results from fixed effects regressions of the more than 8 years of schooling dummy (blue circles) and log years of schooling (red triangles) on the binned sequence of treatment change dummies defined in the text. All specifications include the following individual level controls: A gender dummy, a born in city dummy, age and age squared. All specification include City-year and cohort at move fixed effects as defined in the text. 95% confidence intervals based on standard clustered on the city level are provided by the gray error bars.

FIGURE E-10  
Built-up: Event-study estimates



*Notes:* The figure reports event-study estimates corresponding to the difference-in-differences results presented in [Table VIII](#). Panel A to C cover all six columns of the table. All regressions include city fixed effects and initial-region-by-year fixed effects. 95% confidence intervals based on standard clustered on initial regions are provided by the gray error bars. The orange error bars indicate 95% sup-*t* bootstrap confidence bands with block sampling over initial regions ([Montiel Olea and Plagborg-Møller, 2019](#)).

## E-2. Additional tables

TABLE E-1  
Baseline differences-in-differences

	<i>Dependent Variable: <math>\ln \text{LIGHTS}_{cit}</math></i>					
	All Cities			Reformed Regions		
	(1)	(2)	(3)	(4)	(5)	(6)
CAPITAL	0.1085 (0.0281)	0.0886 (0.0276)	0.1089 (0.0287)	0.1433 (0.0303)	0.1156 (0.0314)	0.1132 (0.0331)
Fundamentals	–	✓	✓	–	✓	✓
City FE	✓	✓	✓	✓	✓	✓
Country-Year FE	✓	✓	–	✓	✓	–
Ini. Region-Year FE	–	–	✓	–	–	✓
$N$	23870	23870	23870	8438	8438	8438
$N \times \bar{T}$	524009	524009	524009	184687	184687	184687

*Notes:* The table reports results from fixed effects regressions of the log of light intensity per square kilometer on capital city status. Standard errors clustered on initial regions are provided in parentheses.

TABLE E-2  
Different light measures

	<i>Dependent Variable: <math>\ln \text{LIGHTS}_{cit}</math></i>				
	Stable lights raw (1)	Stable lights bottom fix (2)	Average lights raw (3)	Bluhm & Krause '18 raw (4)	Bluhm & Krause '18 bottom fix (5)
CAPITAL	0.0395 (0.0311)	0.0899 (0.0344)	0.0876 (0.0306)	0.0899 (0.0344)	0.1132 (0.0331)
Fundamentals	✓	✓	✓	✓	✓
City FE	✓	✓	✓	✓	✓
Ini. Region-Year	✓	✓	✓	✓	✓
$N$	7299	8438	8438	8438	8438
$N \times \bar{T}$	135022	184687	184687	184687	184687

*Notes:* The table reports results from fixed effects regressions of the log of light intensity per square kilometer using different light measures on capital city status. We add one before taking logs of lights per area in km in columns 1 and 4 to keep city-years with no observed light. The raw average lights data record a non-zero light intensity in every city-year. Standard errors clustered on initial regions are provided in parentheses.

TABLE E-3  
Initial city size

	<i>Dependent Variable: <math>\ln \text{LIGHTS}_{cit}</math></i>				
	Initial city size				
	30k	40k	50k	75k	100k
CAPITAL	0.1287 (0.0333)	0.1456 (0.0356)	0.1712 (0.0372)	0.1603 (0.0407)	0.1942 (0.0546)
Fundamentals	✓	✓	✓	✓	✓
City FE	✓	✓	✓	✓	✓
Ini. Region-Year	✓	✓	✓	✓	✓
$N$	5608	4078	3133	1925	1363
$N \times \bar{T}$	122726	89187	68484	42064	29783

*Notes:* The table reports results from fixed effects regressions of the log of light intensity per square kilometer on capital city status. Columns 1 to 5 restrict the estimation samples to cities with an initial population above 30 up to 100k inhabitants. Standard errors clustered on initial regions are provided in parentheses.

TABLE E-4  
City area changes: 1990–2015

	<i>Dependent Variable: <math>\Delta \ln \text{AREA}_{cit}</math></i>			
	All Cities (1)	(2)	Reformed Regions (3)	(4)
CAPITAL	0.1041 (0.0094)	0.0941 (0.0137)	0.1327 (0.0184)	0.0991 (0.0265)
Fundamentals	–	✓	–	✓
Initial-Region FE	✓	✓	✓	✓
Cities	20740	20740	7501	7501

*Notes:* The table reports results from long difference regressions of the change in the log area of a city on the fraction of years in which a city is a capital. The regressions are estimated using the sample of agglomerations, that is, cities which exist in 1990 and have expanded by 2015 or merged into a larger city. Standard errors clustered on initial regions are provided in parentheses.

TABLE E-5  
Different control groups: Countrywide matches

	<i>Dependent Variable: ln LIGHTS<sub>cit</sub></i>					
	Light intensity in 1992			Population in 1990		
	Control city ranks within ... of treated city					
	± 2	± 3	± 4	± 2	± 3	± 4
	(1)	(2)	(3)	(4)	(5)	(6)
CAPITAL	0.0971 (0.0223)	0.0977 (0.0231)	0.0982 (0.0231)	0.0859 (0.0243)	0.0868 (0.0244)	0.0837 (0.0249)
F-test pre-trends ( <i>p</i> -val.)	0.325	0.155	0.146	0.596	0.632	0.763
<i>N</i>	797	1017	1215	765	984	1184
<i>N</i> × <i>T̄</i>	17328	22104	26410	16658	21432	25792

*Notes:* The table reports results from fixed effects regressions of the log of light intensity per square kilometer on capital city status. Panels A to C match treated cities to a varying number of control cities on the basis of their rank in terms of light intensity or population within the entire country. All regressions include city fixed effects, initial-region by-year fixed effects, and time-varying coefficients on the fundamentals. We report an F-test for pre-trends tests for the null hypothesis that all leading terms in the equivalent event-study specification are jointly zero. Standard errors clustered on initial regions are provided in parentheses.

TABLE E-6  
Ethnic diversity

	<i>Dependent Variable: ln LIGHTS<sub>cit</sub></i>					
	All Cities			Reformed Regions		
	(1)	(2)	(3)	(4)	(5)	(6)
CAPITAL	0.1085 (0.0281)	0.0886 (0.0275)	0.1081 (0.0286)	0.1410 (0.0308)	0.1145 (0.0314)	0.1115 (0.0328)
CAPITAL $\times$ ELF	-0.0025 (0.0192)	-0.0029 (0.0184)	0.0067 (0.0196)	0.0132 (0.0205)	0.0082 (0.0207)	0.0124 (0.0214)
Fundamentals	—	✓	✓	—	✓	✓
City FE	✓	✓	✓	✓	✓	✓
Ini. Region-Year FE	✓	✓	✓	✓	✓	✓
$N$	23809	23809	23809	8378	8378	8378
$N \times \bar{T}$	522847	522847	522847	183547	183547	183547

*Notes:* The table reports results from fixed effects regressions of the log of light intensity per square kilometer on capital city status. The interactions of the capital city status with ethnic diversity ( $\tilde{z}$ ) are standardized such that  $\tilde{z} \equiv (z - \bar{z})/\sigma_z$ . Standard errors clustered on initial regions are provided in parentheses.



TABLE E-7  
Democracy

	<i>Dependent Variable: ln LIGHTS<sub>cit</sub></i>					
	All Cities			Reformed Regions		
	(1)	(2)	(3)	(4)	(5)	(6)
CAPITAL	0.1257 (0.0284)	0.1149 (0.0279)	0.1316 (0.0298)	0.1529 (0.0358)	0.1262 (0.0363)	0.1260 (0.0383)
CAPITAL $\times$ DEMOCRACY	-0.0176 (0.0094)	-0.0390 (0.0113)	-0.0415 (0.0119)	-0.0202 (0.0242)	-0.0222 (0.0272)	-0.0264 (0.0278)
Fundamentals	—	✓	✓	—	✓	✓
City FE	✓	✓	✓	✓	✓	✓
Ini. Region-Year FE	—	—	✓	—	—	✓
$N$	23870	23870	23870	8438	8438	8438
$N \times \bar{T}$	522469	522469	522469	184443	184443	184443

*Notes:* The table reports results from fixed effects regressions of the log of light intensity per square kilometer on capital city status interacted with democracy. Standard errors clustered on initial regions are provided in parentheses.

TABLE E-8  
Federalism

	<i>Dependent Variable: ln LIGHTS<sub>cit</sub></i>					
	All Cities			Reformed Regions		
	(1)	(2)	(3)	(4)	(5)	(6)
CAPITAL	0.0934 (0.0308)	0.0779 (0.0302)	0.0913 (0.0319)	0.1231 (0.0337)	0.1010 (0.0337)	0.0981 (0.0358)
CAPITAL $\times$ FEDERAL	0.1262 (0.0600)	0.0899 (0.0571)	0.1257 (0.0554)	0.1469 (0.0578)	0.1063 (0.0587)	0.0987 (0.0601)
Fundamentals	—	✓	✓	—	✓	✓
City FE	✓	✓	✓	✓	✓	✓
Ini. Region-Year FE	—	—	✓	—	—	✓
$N$	23870	23870	23870	8438	8438	8438
$N \times \bar{T}$	522469	522469	522469	184443	184443	184443

*Notes:* The table reports results from fixed effects regressions of the log of light intensity per square kilometer on capital city status interacted with democracy. Federal is a dummy indicating that the country is federally organized (following [Treisman, 2008](#)). Standard errors clustered on initial regions are provided in parentheses.

TABLE E-9  
Larger agglomeration and city fringe: Built-up

	<i>Dependent Variable: Built-up area within city</i>					
	NDBI		UI		NDVI	
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A. Growth of the larger agglomeration</i>						
CAPITAL	0.7235 (0.2200)	0.5693 (0.2251)	1.1174 (0.2939)	0.8241 (0.3071)	-0.9194 (0.2453)	-0.6352 (0.2607)
<i>Panel B. Growth in the periphery of the city</i>						
CAPITAL	0.8584 (0.2811)	0.5581 (0.2680)	0.8233 (0.3503)	0.6599 (0.3445)	-0.3388 (0.3056)	-0.3448 (0.3062)
Fundamentals	–	✓	–	✓	–	✓
Agglomeration FE	✓	✓	✓	✓	✓	✓
Ini. Region-Year FE	✓	✓	✓	✓	✓	✓
$N$	7516	7516	7516	7516	7516	7516
$N \times \bar{T}$	159012	159012	159012	159012	159012	159012

*Notes:* The table reports results from fixed effects regressions of the Normalized Difference Built-Up Index (NDBI), Urban Index (UI), and the Normalized Difference Vegetation Index (NDVI) on capital city status. Panel A reports results based on the larger agglomeration (the envelope over 1990 and 2015 of the urban clusters detected in 1990). Panel B reports the results for the fringe (areas the urban clusters detected in 1990 that meet the detection threshold by 2015). All coefficients are scaled by 100 for exposition. Standard errors clustered on initial regions are provided in parentheses.

TABLE E-10  
Population response: Long differences

	<i>Dependent Variable: <math>\Delta \ln \text{POP}_{ci}</math></i>					
	1990–2000		2000–2015		1990–2015	
	(1)	(2)	(3)	(4)	(5)	(6)
CAPITAL	0.1615 (0.0139)	0.1334 (0.0169)	0.1861 (0.0186)	0.1494 (0.0264)	0.3612 (0.0310)	0.2947 (0.0427)
Fundamentals	–	✓	–	✓	–	✓
Initial-Region FE	✓	✓	✓	✓	✓	✓
$N$	8451	8451	8451	8451	8451	8451

*Notes:* The table reports results from long difference regressions of the change in log population of a city over different epochs on the fraction of years in which a city is a capital. Standard errors clustered on initial regions are provided in parentheses.

TABLE E-11  
Population changes: Agglomeration & city fringe

	<i>Dependent Variable: <math>\Delta \ln \text{POP}_{ci}</math></i>					
	1990–2000		2000–2015		1990–2015	
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A. Growth of the larger agglomeration</i>						
CAPITAL	0.1718 (0.0153)	0.1416 (0.0176)	0.2157 (0.0221)	0.1745 (0.0280)	0.4035 (0.0368)	0.3301 (0.0458)
<i>Panel B. Growth in the periphery of the city</i>						
CAPITAL	0.0741 (0.0743)	0.0764 (0.0752)	0.1484 (0.0345)	0.1609 (0.0345)	0.2340 (0.0845)	0.2459 (0.0850)
Fundamentals	–	✓	–	✓	–	✓
Agglomeration FE	✓	✓	✓	✓	✓	✓
Ini. Region-Year FE	✓	✓	✓	✓	✓	✓
<i>N</i>	7340	7340	7340	7340	7340	7340

*Notes:* The table reports results from long difference regressions of the change in log population of a city on the fraction of years in which a city is a capital. Panel A reports results based on the larger agglomeration (the envelope over 1990 and 2015 of the urban clusters detected in 1990). Panel B reports the results for the fringe (areas the urban clusters detected in 1990 that meet the detection threshold by 2015). Standard errors clustered on initial regions are provided in parentheses.

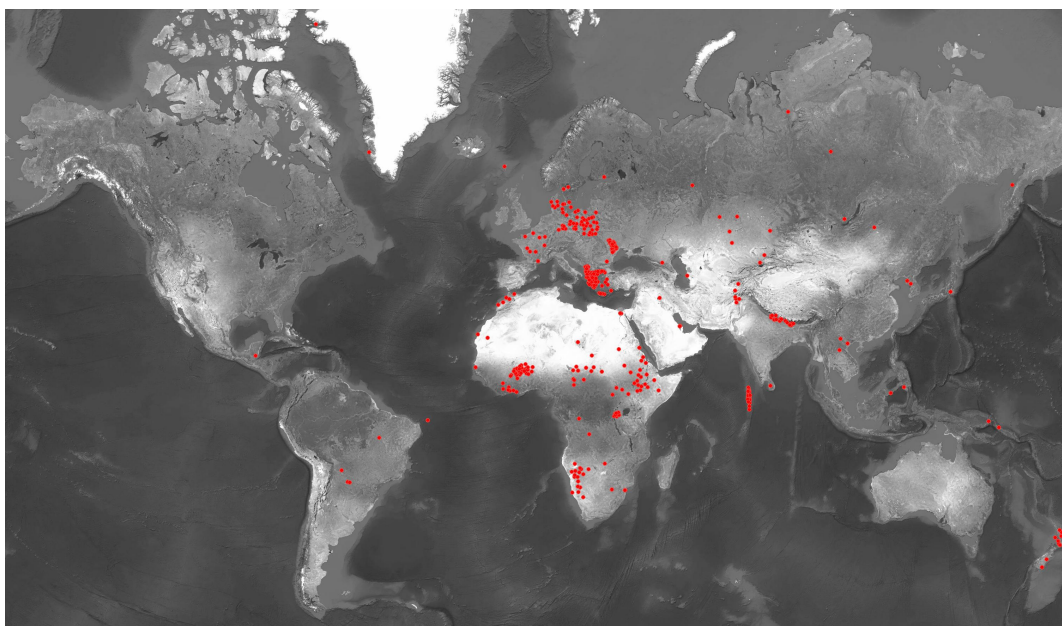
## F. Former capitals and “mother” capitals

This appendix provides descriptive statistics on cities that lose their status as a capital, discusses pre-treatment trends, and the appropriate comparison groups for these cities. We also report evidence on the effects of cities which lose the capital status relative to their peers (cities that remain capitals).

### F-1. Former capitals

Many cities across the globe have lost the status of as a capital during the last three decades (see [Figure F-1](#)). About 94% of the observed 169 status losses in our sample occur during a territorial centralization (mergers of two or more regions). In the other 6% of cases, a different city becomes a capital within the same region.

FIGURE F-1  
Spatial distribution: Capital loss

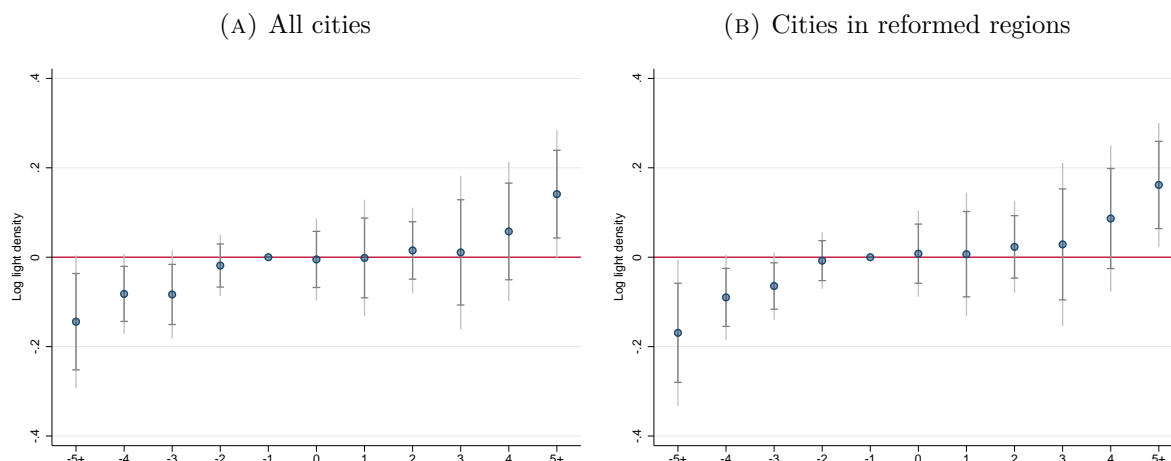


*Notes:* The figure plots all the cities that lose their capital status during the 1987 to 2018 period.

We first turn to our baseline specification which uses other non-capital cities as the control group. [Figure F-2](#) reports event-study estimates using our preferred specification with initial-region-by-year fixed effects and controls for locational fundamentals. There are significant and negative pre-trends. Capital cities that lose their status perform worse relative to non-capitals prior to treatment. Regardless of why this occurs, identification is not feasible in our primary setting.

Of course, it makes more sense economically and statistically to compare capitals that lose their status to cities that remain capitals. Unfortunately, this also implies that we now work with a drastically reduced sample size (of 392 capital cities) and a design that

FIGURE F-2  
Former capitals vs. all cities

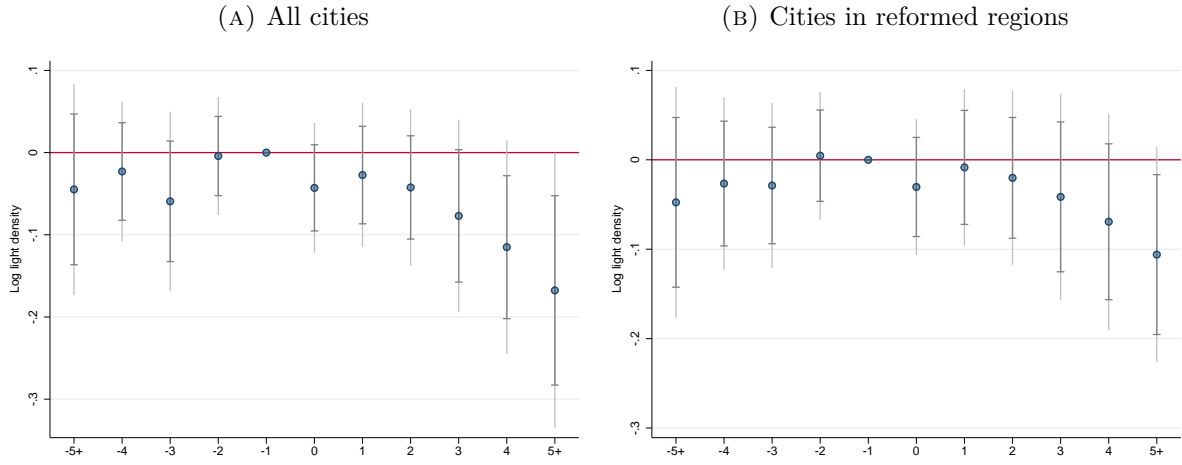


*Notes:* The figure illustrates results from fixed effects regressions of the log of light intensity per square kilometer on the binned sequence of treatment change dummies (capital loss) defined in the text. Panel A shows estimates for all ever capital cities based on a specification with country-year effects. Panel B shows estimates for ever capital cities in reformed regions based on a specification with final-region-by-year fixed effects. All regressions include city fixed effects. 95% confidence intervals based on standard clustered on final regions are provided by the gray error bars.

more closely resembles a staggered event study with a small control group. Moreover, we do not have enough degrees of freedom to allow for time-varying coefficients on the locational fundamentals. In [Figure F-3](#) we run event studies on the set of ever capitals using again binned treatment change indicators for city loss. Note that we exclude cities that become capitals during our sample period. Hence, the comparison groups differ a lot compared to our standard approach. The identifying variation in panel A is based on the difference between cities that are always capitals within the country compared to capitals that lose that status sometime during our sample period. The identifying variation in panel B is restricted to mergers of administrative regions in which one city loses the status and the other city becomes the capital of the whole region. Note that focusing on mergers has also implications for the type fixed effects we can include. Instead of initial-region-by-year fixed effects, we now use final-region-by-year fixed effects. This allows us to compare cities within the at some point merging region and control for unobserved trends in the constituent parts prior to their merger.

The results show a clear pattern. We find no evidence suggesting the presence of pre-trends. Hence, capitals that will subsequently lose their capital status are not declining relative to always capitals prior to treatment. After the capital status is removed, we observe a steady loss of economic activity that takes longer to materialize than our main result but suggests a decline of similar magnitude in the medium-run.

FIGURE F-3  
Former capitals vs. always capitals



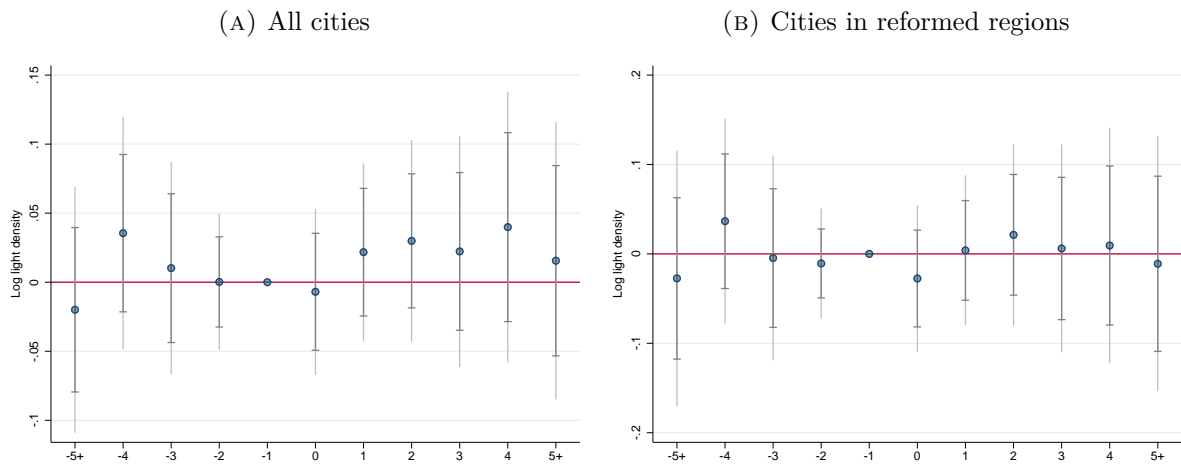
*Notes:* The figure illustrates results from fixed effects regressions of the log of light intensity per square kilometer on the binned sequence of treatment change dummies (capital loss) defined in the text. Panel A shows estimates for all ever capital cities based on a specification with country-year fixed effects. Panel B shows estimates for ever capital cities in reformed regions based on a specification with final-region-by-year fixed effects. All regressions include city fixed effects. 95% confidence intervals based on standard clustered on final regions are provided by the gray error bars.

## F-2. “Mother” capitals

A related issue to the loss of a political premium is territorial decentralization’s effect on existing capitals that lose part of their territory. We refer to these cities as “mother capitals”, i.e., capitals that rule over a smaller jurisdiction after a decentralization reform that creates new additional capitals in the initial region.

We specify the corresponding event for capitals that experience a reduction in their jurisdiction and estimate event studies comparing their performance to the set of always capitals. [Figure F-4](#) presents the results. We find no evidence in favor of pre-treatment trends or any change in activity after a city becomes a “mother capital”. The economic gains of new capital cities appear not to come at the cost of the old ones, at least not in the short to medium run.

FIGURE F-4  
Mother capitals vs. always capitals



*Notes:* The figure illustrates results from fixed effects regressions of the log of light intensity per square kilometer on the binned sequence of treatment change dummies (mother capitals). Panel A shows estimates for all ever capital cities based on a specification with country-year fixed effects. Panel B shows estimates for ever capital cities in reformed regions based on a specification with initial-region-by-year fixed effects. All regressions include city fixed effects. 95% confidence intervals based on standard clustered on initial regions are provided by the gray error bars.

## Additional references

- AidData (2017). `WorldBank_GeocodedResearchRelease_Level1_v1.4.2` geocoded dataset. Aid Data Williamsburg, VA and Washington, DC. AidData. Accessed on 02/09/2020, <http://aiddata.org/research-datasets>.
- Bai, Y. and R. Jia (2021). The economic consequences of political hierarchy: Evidence from regime changes in China, 1000-2000 C.E. *Review of Economics and Statistics*. forthcoming.
- Bardhan, P. K. and D. Mookherjee (2000). Capture and governance at local and national levels. *American Economic Review* 90(2), 135–139.
- Bluhm, R., A. Dreher, A. Fuchs, B. Parks, A. Strange, and M. J. Tierney (2020, May). Connective Financing - Chinese Infrastructure Projects and the Diffusion of Economic Activity in Developing Countries. CEPR Discussion Papers 14818, C.E.P.R. Discussion Papers.
- Bluhm, R. and M. Krause (2022). Top lights: Bright cities and their contribution to economic development. *Journal of Development Economics* 157, 102880.
- Brinkhoff, T. (2020). City population. Technical report.
- Campante, F. R. and Q.-A. Do (2014). Isolated capital cities, accountability, and corruption: Evidence from us states. *American Economic Review* 104(8), 2456–81.
- Campante, F. R., Q.-A. Do, and B. Guimaraes (2019). Capital cities, conflict, and misgovernance. *American Economic Journal: Applied Economics* 11(3), 298–337.
- Henderson, J. V., V. Liu, C. Peng, and A. Storeygard (2020). Demographic and health outcomes by degree of urbanisation: Perspectives from a new classification of urban areas. Technical report, Brussels: European Commission.
- Henderson, J. V., 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(1), 357–406.
- Jarvis, A., E. Guevara, H. Reuter, and A. Nelson (2008). Hole-filled SRTM for the globe: version 4: data grid.
- Kiszewski, A., A. Mellinger, A. Spielman, P. Malaney, S. E. Sachs, and J. Sachs (2004). A global index representing the stability of malaria transmission. *The American Journal of Tropical Medicine and Hygiene* 70(5), 486–498.
- Law, G. (2010). *Administrative subdivisions of countries*. Jefferson, NC: McFarland & Company. The official reference for the Statoids.com database.
- Montiel Olea, J. L. and M. Plagborg-Møller (2019). Simultaneous confidence bands: Theory, implementation, and an application to SVARs. *Journal of Applied Econometrics* 34(1), 1–17.
- Nunn, N. and D. Puga (2012). Ruggedness: The blessing of bad geography in Africa. *Review of Economics and Statistics* 94(1), 20–36.
- Storeygard, A. (2016). Farther on down the road: Transport costs, trade and urban growth in sub-Saharan Africa. *Review of Economic Studies* 83(3), 1263–1295.
- Sun, L. and S. Abraham (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics* 225(2), 175–199.
- Treisman, D. (2008). Dezentralization dataset. Dataset.