Songlines*

Sotiris Kampanelis
Cardiff University†

Aldo Elizalde
Queen's University Belfast‡

Yannis M. Ioannides
Tufts University§

ASSA Version: December 26, 2024

Abstract

This study examines the hypothesis that Europeans' adoption of Indigenous knowledge during colonisation was pivotal in shaping economic development in the long-run. We use the case of Australia, where Aboriginal knowledge of the landscape was critical to colonial exploration and settlement. To quantify the effects of this knowledge, we construct a newly digitised and georeferenced dataset of trade routes created by Aboriginal people based on oral traditions, known as *Songlines*. Our results indicate that *Songlines* are strongly associated with current economic activity as measured by satellite light density at night and, alternatively, population density. As a counterfactual, we construct *Natural Routes*—environmentally optimal travel paths—and show that *Songlines* dominate them in predicting colonial exploration routes, early settlements, and modern economic activity. We attribute this association to path dependence and agglomeration effects that emanate from the transport infrastructure built by Europeans roughly along the *Songlines*, which have induced agglomeration of economic activity.

Keywords: Aboriginal trade routes; Songlines; colonialism; agglomeration; Australia

JEL Codes: N77; O10; R12; Z10; Z13

---

*We thank Luis Angeles, Panle Barwick, Jeanet Bentzen, Yuxian Chen, Arcangelo Dimico, Evangelos V. Dioikitopoulos, Andrew Foster, Oded Galor, Luís Guimarães, Edward Glaeser, Pauline Grosjean, Eduardo Hidalgo, Erik Hornung, Stelios Michalopoulos, Federico Masera, Frederic Robert-Nicoud, Jordan Rosenthal-Kay, Nathan Nunn, Sascha O. Becker, Ömer Özak, Elias Papaioannou, Adam Storeygard, Alex Trew, Serena Trucchi, and David Weil for helpful comments and discussions on our work. We would also like to thank the audiences at the seminars, conferences, and workshops in CRETE 2022, ASREC-London, ASREC-Australasia, IESG-AERIP, Queen's University Belfast, RES 2023, RIDGE-TSW Bogota Forum, RIDGE-TSW Chile Forum, NBER-Urban Economics 2023, Uppsala Workshop in Economic History, Brown University, University of Toronto, Urban Economics Association in Washington DC, and University of California Irvine. We also thank Maria Bableki and Charlie Joyez for excellent research assistance. We are grateful for financial support from the British Academy/Leverhulme Small Research Grants. The paper has been presented under an earlier title: "How Local Knowledge Fuelled Colonial Settlement and Long-Run Economic Activity: Evidence from Australia."

†Cardiff Business School, Cardiff University; kampaneliss@cardiff.ac.uk
‡Department of Economics, Queen's University Belfast; a.elizalde@qub.ac.uk
§Department of Economics, Tufts University; Yannis.Ioannides@tufts.edu

1. Introduction

The economic legacy of European colonisation has been a subject of considerable scholarly interest (Sokoloff and Engerman, 2000; Acemoglu et al., 2001; Acemoglu et al., 2002; Glaeser et al., 2004; and Dell, 2010). A central focus of this research is to understand the drivers behind the patterns of European settlement and their long-term economic consequences (Easterly and Levine, 2016). To explain these patterns, scholars have traditionally relied on the hypothesis of first-nature geography, according to which the natural features of the landscape played a decisive role (Sokoloff and Engerman, 2000; Acemoglu et al., 2001). However, a less explored dimension in the literature is the role of second-nature geography, that is in our case the role of activity associated with the local knowledge possessed by the Indigenous population, in shaping European colonisation and its aftermath (Caruana-Galizia et al., 2021).

There are numerous historical episodes that illustrate how the adoption of Indigenous knowledge shaped European colonial expansion. A notable episode is Plymouth Colony, the second English colony in the United States, founded by Pilgrims in 1620. Historians have documented that the use of native knowledge of corn cultivation enabled settlers to survive the first years of colonisation, ultimately ensuring the continuity of the colony and contributing to the expansion of European colonisation (Mann, 2005).[1] Despite the fact that anthropologists and historians have widely recognised this important channel, the economic literature has yet not fully investigated its potential economic impacts. Moreover, the underlying mechanisms that explain the path-dependence in these economic effects remain unexplored. We aim to fill this gap by focusing on Australia where the adoption of Aboriginal knowledge of the landscape was an integral part of colonial exploration and settlement. As Reynolds (1980) puts it:

"*The natives were the parties who first guided the White Man through the intricacies of their forests, led them to their Rivers, their springs, and rich pastures, assisted them in keeping their stock, watched their working oxen, tracked their stray Horses, and rendered other essential assistance . . . The knowledge of their Country was thus acquired, was turned to account.*"

By the time of contact with Europeans in the early seventeenth century, Aboriginal people had developed an extensive trading network based on oral traditions (Flood, 1983; Veth and O'Connor, 2013; Kerwin, 2010; McBryde, 2000). This oral tradition is known as *Songlines*, which contains information that connects specific landmarks that al-

[1] For a recent overview of why colonisation is considered an important factor explaining the origins of modern economic growth, see Koyama and Rubin (2022), Ch. 6.

lowed Aboriginal people to navigate long distances based on songs and stories (Chatwin, 1987; Norris and Harney, 2014). Anthropologists refer to this "oral network" as "multitudinous", extending across Australia in all directions (Mountford, 1976; Kerwin, 2010, p. 63). Recent anthropological evidence indicates that this trading network provided Europeans with vital knowledge to explore and settle mainland Australia during colonisation. For example, Kerwin (2010) noted that "*European exploration of Australia...was made possible by Aboriginal trading paths...Surveyors quickly expanded European occupation and territories, and they coloured the map of Australia in their image*". We argue that Europeans used this knowledge in building transport networks to connect the main colonies, which led to improved connectivity and an increase in economic activity roughly along these routes. This had a profound and long-lasting impact on the evolution of modern economic activity and urbanisation patterns in Australia.

To quantitatively capture this knowledge, we create a new georeferenced dataset of *Songlines*. We do so by digitising and utilising anthropological data on Aboriginal trading sites published in a series of research articles by Frederick David Mc̀Carthy (1939). These publications provided one of the most comprehensive descriptions of all known (as of the time of his writing) points of origin and destination associated with Aboriginal people prior to colonisation. We acquire extensive information about these sites through careful reading and compilation. This yields a catalogue of 1,026 origins and destinations that span mainland Australia.

Armed with this unique data, we develop a novel georeferenced map of *Songlines* in Australia. To achieve this, we employ a least cost path algorithm (LCP) to identify optimal routes linking the 1,026 origins and destinations. Our approach is based on the hypothesis that Aboriginal people selected paths based on their physiological constraints and environmental conditions, as indicated by anthropological research (Kerwin, 2010). The algorithm uses the Human Mobility Index (HMI) developed by Özak (2018). The HMI estimates the time it would take people to move across a square kilometre of land in pre-industrial times, taking into account temperature, relative humidity, cloud cover, slope, type of terrain and risk of heat exhaustion. We characterise 513 distinct *Songlines* that comprise the entire network.[2]

To construct our unit of analysis, we divide Australia into 10 km x 10 km grid cells. For each unit, we measure contemporary economic

[2] We also rely on the non-georeferenced maps of Mc̀Carthy (1939) to create the *Songlines* and thus construct an alternative variable of interest. Besides the detailed anthropological description of the different origin and destination points that Mc̀Carthy (1939) documented in his publications, he also included a series of maps illustrating some of the *Songlines*. We digitise and georeference these maps.

activity via satellite nighttime light density data for 2015.[3] We also consider population density as an alternative proxy , which we observe at the same grid cell level.[4] Our empirical examination begins by controlling for local government fixed effects to account for unobserved, government-specific characteristics. We find a positive and statistically significant association between *Songlines* and current economic activity. Specifically, we find that the presence of a *Songline* increases the probability of observing economic activity in Australia by 3% today.

It is particularly interesting to determine whether the spatial features of Australia's modern economy have been shaped by first-nature geography, that is by the landscape, or by the second-nature influence of *Songlines*, an outcome of Aboriginal economic and cultural activity. Our estimation models control for a range of geographic and climatic features so as to mitigate potential confounding factors. In addition, we examine historical mining activities as a possible explanation for rapid urbanisation during the colonial gold rush in Australia. We find that, even after controlling for these factors, our results remain unchallenged.

To provide causal evidence, we perform a contiguous-pair analysis similar to methodologies employed by Dube et al. (2010) and Michalopoulos and Papaioannou (2013). This approach leverages grid cell pairs, where one grid cell hosts a *Songline* and its contiguous neighbour does not, so as to isolate the effect of *Songlines* on economic activity. By including fixed effects for each grid cell pair, we control for all shared unobserved characteristics. Hence, any variation in potential outcomes is likely to be associated with the presence of a *Songline*. Indeed, our results show that grid cells containing *Songlines* yield significantly higher levels of current economic activity compared to their adjacent neighbours that do not have them.

We then proceed to validate the plausibility of our identification assumption. Specifically, we seek to establish whether our results are due to adoption of Aboriginal knowledge by Europeans while they explored and settled Australia during the colonial period, and not to the features of Australian geography. Simply put, European explorers may have deviated from Aboriginal routes due to places with either favourable environmental features, or less Aboriginal hostility against them, or just because they evaluated the suitability of the landscape differently. This would mean that Australia's spatial development would have followed different geographical patterns. We address this issue in a number of different ways.

[3] In the Appendix E, we also perform several association tests that confirm the validity of light density as an economic activity measure for Australia, even for arid deserts up to (and more) 500km from the coast.

[4] We thank Andrew Foster and Edward Glaeser for insisting that we also use population density as an outcome.

First, by drawing on the diaries of European explorers, we present historical evidence highlighting the crucial role of Aboriginal companions in facilitating European exploration efforts. Second, we exploit exogenous variation in European exploration and settlement by constructing a measure of *Natural Routes*, inspired by the methodology of Barjamovic et al. (2019). These *Routes* capture the most geographically favorable paths for travel from coastal grid cells to the interior, representing the routes Europeans might have plausibly taken when exploring inland Australia unguided and on their own. Our results are surprising: we demonstrate that *Songlines* are more predictive than *Natural Routes* of both the presence of European exploration routes during the colonial period and current levels of economic activity. Third, we investigate whether the positive effects of *Songlines* persist in areas with local "frictions" that might have hindered the adoption of Aboriginal knowledge during colonisation. That is, using data on Aboriginal massacre sites (Ryan et al., 2017), we find that the effect of *Songlines* on economic activity weakens near these sites but strengthens as the distance from them increases.

We next examine the dynamic effects of *Songlines* on the emergence of new settlements in Australia using two strategies. First, we find that *Songlines* strongly predict the establishment of cities before 1900, with early cities typically established closer to a *Songline*. However, given that most major cities were coastal, their locations may reflect other factors unrelated to Aboriginal knowledge. To address this, we use data on the year each of 740 non-coastal towns were established (Elder, 2024), finding consistent results.

We extend our analysis by conducting robustness checks to validate our baseline results. These include homogenising the sample by retaining grid cells intersected by the *Songlines* and their adjacent cells. To address potential measurement error, we reconstruct the *Songlines* by incorporating additional parameters in our Least Cost Path (LCP) algorithm. We also include dummy variables for all origin and destination points to account for their potential influence. Furthermore, we consider the network centrality of *Songlines*, address spatial autocorrelation, and mitigate omitted variable bias. Lastly, we control for Aboriginal language fixed effects to account for pre-colonial factors and incorporate early colonial economic activity as an additional control.

We address the mechanisms that can explain our main results. We hypothesise that the influence of *Songlines* can be partially attributed to path dependence and associated agglomeration effects resulting from the early European transport infrastructure. Specifically, when Europeans settled in new areas, they built transport networks connecting these settlements to major colonies (Kerwin, 2010). This likely caused these routes to become increasingly associated with economic activity as a result of increased connectivity over time. To quantitatively investigate this mechanism, we digitise and georeference a series of maps showing the 19th and 20th century Australian railway and road networks. We find a strong positive relationship between *Songlines* and early transport infrastructure.

We address the counterfactual that early transport infrastructure might have been designed with very different objectives. We do so by focusing on larger areas made up of smaller 10km x 10km grid cells and identifying the top 75% and 90% of places with the highest density of *Natural Routes* (as a proxy for cost-effective construction conditions). Then, we examine whether early infrastructure aligns more closely with these *Routes* or with the *Songlines*. Our findings reveal that early infrastructure follows the *Songlines* rather than the *Natural Routes*.

This paper aims at several contributions to the literature. First, it provides the first quantitative evidence that European colonisers leveraged Aboriginal knowledge by means of "oral routes." This orally transmitted knowledge shaped their exploration and settlement patterns and thus affected economic development in the long run. Second, it advances research on urban development and spatial dynamics (Glaeser et al., 1992; Duranton and Puga, 2014; Barsanetti, 2021) by demonstrating that these Aboriginal routes influenced early European settlements and have continued impacting contemporary economic activity. Third, it enriches the literature on pre-colonial factors affecting modern outcomes (Michalopoulos and Papaioannou, 2013; Angeles and Elizalde (2017)) by showing their persistent effects even in developed countries like Australia. Additionally, the study expands research on historical trade routes (Flückiger et al., 2021; Barjamovic et al., 2019) by focusing on societies without centralised states and physical infrastructure and highlighting how Aboriginal routes facilitated urban development. Lastly, it establishes the long-term impacts of historical events (Nunn, 2008; Valencia Caicedo, 2019), demonstrating that the adoption of Aboriginal knowledge by colonisers had lasting effects: it shaped European settlement patterns and influenced significantly modern urbanisation and economic geography in Australia.

The remainder of the paper proceeds as follows. Section 2 outlines historical evidence on how Europeans adopted Aboriginal knowledge during colonisation in Australia. Section 3 describes the data and presents the main empirical results. Section 4 validates the identification assumption. Section 5 explores the dynamic effects of *Songlines* on urbanisation. Section 6 presents robustness checks. Section 7 focuses on the mechanisms. Section 8 concludes. An extensive set of additional details backing up our analysis is given in the Appendix.

2. Adopting Knowledge from Aboriginal Guides

Prior to colonisation in Australia, Aboriginal people had developed an extensive system of oral traditions via which they passed on to subsequent generations their knowledge of the landscape via songs and stories. This oral tradition is known as *Songlines*. Wositsky and Harney (1999) define *Songlines* as "*epic creation songs passed to present generations by a line of singers...[which] provide maps for the country,"* and some of these *"describe a path crossing the entire Australian continent."*[5]

European explorers noted the utility of this Aboriginal knowledge during the early stages of colonisation (Reynolds, 1980). Unfamiliar with the Australian landscape, they depended on Aboriginal guides to navigate, locate water and food, and communicate with Indigenous communities. Aboriginal expertise was vital to the success of many expeditions, with explorers employing two main strategies to benefit from this knowledge: recruiting Aboriginal guides from the starting point of explorations or engaging with local Aboriginal people during the journey (Clarke, 2008; McLaren, 1996).

We provide evidence about the importance of Aboriginal guides during European colonial exploration in a comprehensive dataset, which may be accessed via an online repository.[6] This dataset documents 76 inland explorers from the 19th century who were accompanied by Aboriginal guides. More specifically, for each explorer, we find at least one reliable source suggesting that they were accompanied by one or more Aboriginal guide(s) during their explorations. Based on this data, we document an extensive narrative, including examples from explorers' diaries, which offers substantial evidence of the crucial role that Aboriginal people played in European explorations. This systematic compilation is based on several sources like the *explorers' diaries, the Australian Dictionary of Biography, local council websites*, and documents from the *National Library of Australia*. Although it is highly likely that some expeditions were not fully documented by the explorers themselves or later writers, arguably our data covers the most important explorers and their Aboriginal guides.

Starting from the first colonies in Southeastern Australia, like Sydney and Melbourne, Aboriginal guides were vital to European exploration efforts. For instance, Hamilton Hume frequently relied on Aboriginal expertise throughout his expeditions since his first journey in 1814 at the age of 17. Hume's respect for Aboriginal expertise per-

[5] This is reminiscent of the prevailing theory of how the Homeric epics were passed on over many years by the "Singers of Tales" (who performed from memory) until the epics were written down in the sixth century BCE (the so-called Pisistratean Recension) (Lord, 1960). The *Songlines* were popularised by Chatwin (1987).

[6]

sisted throughout his career, especially during his 1828 expedition with Charles Sturt. His ability to communicate and negotiate with Aboriginal communities was key to securing safe passage and gaining valuable local knowledge, greatly contributing to the success of his explorations (Hume, 1966). Similarly, Charles Throsby depended on an Aboriginal guide named Broughton (also known as Toodwick or Toodwit) during his expeditions from 1817 to 1821 (Campbell, 2005). George Evans also benefited from Aboriginal guidance, particularly during his 1812 expedition from Jervis Bay to the Shoalhaven River. He was led by Bundle, a Gundungurra man, and followed a traditional trade route, *a Songline*, used by the Dharug and Gandangara peoples to reach Wiradjuri territory (Bathurst Local Aboriginal Consultative Committee, 2011).

In Central Australia, the harsh landscapes posed significant challenges for explorers like Ernest Giles and John McDouall Stuart. Giles' survival often depended on his Aboriginal guide Tommy Oldham, whose expertise in finding water sources, such as Queen Victoria Spring, saved the expedition from disaster (Giles, 1889). Similarly, Stuart's reliance on Aboriginal trackers during his expeditions through Central Australia was essential, though his struggles after the departure of an Aboriginal guide in 1858 underscore the difficulty of surviving without Indigenous knowledge (Stuart, 1865).

Western Australia offers further examples of successful collaboration between explorers and Aboriginal guides. The Forrest brothers, John and Alexander, were heavily reliant on their guide Windich, whose knowledge of water sources and forage for horses was crucial in their exploration of Lake Moore and Lake Barlee in the 1860s and 1870s (Forrest, 1875). Likewise, Peter Egerton Warburton benefited from the guidance of Charley, whose tracking skills and knowledge of water sources were indispensable to Warburton's expeditions (Warburton, 1875).

In Northern Australia, explorers like Ludwig Leichhardt and Edmund Kennedy also depended on Aboriginal guides. Leichhardt's expeditions, particularly his first in 1844, were marked by the essential contributions of an Aboriginal guide named Harry Brown, who corrected Leichhardt's course and ensured the party's survival (Leichhardt, 1847). Similarly, Kennedy's guide named Jackey Jackey demonstrated loyalty and bravery, remaining with Kennedy until his death and ultimately reaching the supply ship alone.[7]

The role of Aboriginal guides in Australian exploration proved crucial in opening up the Australian interior to European settlement (Reynolds, 1990; McLaren, 1996; Clarke, 2008). Soon after the first British colony at Sydney Cove was established in 1788, exploration of inland Australia was pursued in search for more land for agricultural production. In the early 1810s, expeditions were organised to cross the Blue Mountains,

[7] Please refer to section D in Appendix for further details.

a large mountainous region bordering Sydney. This expedition led to the discovery of Bathurst by the explorer George Evans, which became the first inland settlement in Australia.

Importantly, following the European expeditions that opened up new areas for permanent settlements, construction of transport infrastructures emerged as a necessary tool for connecting these regions. For example, shortly after the discovery of the Bathurst Plain by explorer George Evans, a road was built to connect it with the main colonies and ports in the broader area of Sydney (Herald, 1912 [Online]). More examples of railways and roads that ran along *Songlines* can be found in the anthropology literature (Kerwin, 2010).

Indeed, while it is plausible that the adoption of Aboriginal knowledge aided Europeans in locating suitable settlements, this became less necessary once these were established, mapped, and connected to the main colonies. Yet, the influence of *Songlines* may have survived even after Europeans no longer needed them to travel between settlements. Specifically, the construction of transport infrastructure by Europeans roughly along *Songlines* created path dependence and agglomeration effects that fostered the concentration of economic activity in these areas as a result of increased connectivity over time. We argue that this has profoundly impacted the evolution of the geographical patterns of urbanisation and economic activity in Australia.

3. Data and Empirical Results

Our data construction and selection of variables is aimed at measuring key dimensions of Australian urbanisation. We place particular emphasis on the *Songlines* as well as counterfactual proxy routes that are defined so as to let us examine both first- as well as second-nature geography, given the topography of Australia, and perform extensive robustness checks.

3.1. Construction of Variables

The novel variables of interest in this study consist of all *Songlines* along the Australian land mass. Each *Songline* is defined by an origin and a destination, obtained from the trilogy of research articles on Aboriginal trade sites reported by Mćarthy (1939).[8]

Mćarthy (1939) used techniques that have been emulated more recently by other anthropologists in constructing datasets that record pre-colonial characteristics of ethnic groups, such as Murdock (1967). Specifically, he utilised early written accounts from the time of contact

[8]Frederick David McCarthy was an anthropologist and archaeologist.

of settlers with natives to construct the dataset.[9] We manually obtained this information from the accounts by identifying the origin and destination of each *Songline* in Australia. Two such examples, associated with procuring red ochre and pituri, are discussed next.[10]

Example 1: *"Jessop...says that at "these three places, Noarlunga, Augusta and Aroona, situated at distances of 150 miles in a direct line from south to north, where they interchanged their respective earths or clays, the natives drove also a good trade in skins with those who lined further inland""*.

Example 2: *"Johnston and Cleland... state that " it is probable that the supply of pituri of the Ooldea area is secured from the Musgrave and Everard Ranges... along the recognized and known trade routes between these areas... there is a native trade route passing more or less north-east from Alice Springs via Aritunga, Ambalindum, the Plenty, MacDonald Downs, The Sandover, thence via the native wells to the Georgina at Lake Nash""*.

We geocoded all recorded trading sites. This generated a set of 1,026 origins and destinations, encompassing all Australian states except Tasmania. They label 513 *Songlines* and are the primary variables of interest. The set of 513 routes is a small subset of the universe of all possible connections $(513 \times 512/2)$, suggesting that Aboriginal people did optimise with respect to all connection possibilities.

Drawing the routes between all origins and destinations requires taking into account climatic, environmental, and physiological constraints on people. Aboriginal people, having traveled across all of Australia for thousands of years, must have been able to select via trial and error travel routes whose environmental and climatic conditions suited them best, considering their biological endurance and food and water conservation knowledge and technologies.

Therefore, we employ the Human Mobility Index (HMI) developed by Özak (2018) to impute the most cost-effective route between each documented pair of origin and destination points via a least-cost path

[9] For example, Peter Beveridge, a British colonist, spent twenty three years (1845 to 1968) documenting Aboriginal practises, including trade, among various tribes on the Lower Murray River, now Sawn Hill.

[10] They demonstrate the trade of red ochre, a pigment used by Aboriginal people for ceremonies, in composing rock art, and to decorate various objects. and pituri, respectively, between South Australia and the Northern Territory.

(LCP).[11,12] Other than the information provided by the maps and sketches in MćCarthy (1939), we do not know those exactly. Importantly, this index gives the travel time on each square kilometer of land in pre-industrial times, taking into account temperature, relative humidity, cloud cover, slope, type of terrain (e.g. natural trails or loose sand), and risk of heat exhaustion, which refers to the effects of heat on an individual's metabolic rate and speed to avoid exhaustion. Thus, the HMI is based on an estimate of the time it takes a person to traverse each square kilometre of terrain.

[11] The LCP is a distance analysis function that finds optimal pathways between two locations, considering additional parameters such as obstacles. For instance, terrain altitude and path slope can be considered as costs that are typically avoided by travelers. In the present paper, the main variables of interest are the time taken to traverse a location, as greater travel time is generally seen as less favourable for walking. To illustrate, if a mountain range separates two locations, a person may opt to to bypass it by means of a more level but possibly longer route. Therefore, the objective is travel time and the LCP is employed to identify the most advantageous paths.

[12] Since the HMI raster file lacks data on the initial layer of coastal grid cells in Australia, we were unable to employ the least-cost path algorithm for establishing *Songlines* from coastal sites. As a result, we have excluded these particular grid cells from our analysis. But even if we had created trade routes along including coastal cells, it would have introduced bias into our findings, favouring our primary hypothesis. This potential bias may have arisen from the increased likelihood of *Songlines* predicting economic activity, given the concentrated development along the Australian coast. Therefore, by excluding these coastal grid cells, we effectively mitigate these biases in our study.

Figure 1. Aboriginal Trading Sites and *Songlines* in Australia

Notes: This map shows our reconstruction of the approximate location of *Songlines* in Australia, based on the work of Mc̍Carthy (1939). White lines indicate *Songlines*, constructed via a least cost path algorithm using HMI. Blue dots and triangles (1,026) denote trading sites (origins and destinations). Blue triangles (275) signify exclusions as part of a robustness check for analysis (see section 6.4). *Songlines* were optimally constructed by authors using Özak (2018) with origins and destinations from Mc̍Carthy (1939).

Figure 1 illustrates the entire network of *Songlines* in Australia as reconstructed by the authors based on the work of Mc̍Carthy (1939). The map reveals a vast web of routes encompassing the entire interior of Australia, in addition to the coastal routes.

3.2. Construction of Unit of Analysis and Outcomes

We use a grid cell of 10 km x 10 km, which covers an area of 100 km^2, as our unit of analysis. To construct an indicator of economic activity for each grid cell, we utilise data from the Visible Infrared Imaging Radiometer Suite (VIIRS) sensor, which captures detailed light imagery of human-built facilities for the year 2015. This remote sensing data is filtered to exclude background noise, solar and lunar pollution, and features unrelated to electric lighting such as sunlight, glare, moonlight, aurora, fires and volcanoes, as described in Elvidge et al. (2017).

The VIIRS data differs from the widely used DMSP (Defence Meteorological Satellite Programme) nightlight data in terms of the pixel footprint, which is 45 times smaller. This improved resolution allows for more precise light data, resulting in a higher degree of accuracy when performing local analysis.[13,14]

In addition to luminosity data, we also construct an alternative outcome measuring population density at the grid cell level for the year 2015 by leveraging data from the Global Human Settlement Layer (GHSL) (European Commission, 2023). More specifically, the spatial raster dataset depicts the distribution of the residential population, expressed as the number of people at 30 arc sec resolution. We then compute the average population for each observation. Since our observations (cells) are equal areas (10 km x 10 km), the average population can also be considered as population density. Utilising GHSL urban population data offers the advantage of leveraging high-resolution satellite imagery to reassign individuals from larger, sparsely populated administrative units to areas with visible buildings, thus providing a comprehensive representation of urban populations, even in the smallest settlements.

3.3. Covariates and Methodological Issues

We focus on whether *Songlines* had an influence on spatial features of the Australian economy. We posit that Aboriginal people had an intimate knowledge of their environment, which enabled them to choose "efficient" so as to be able to exchange goods and engage in cultural activities and customs over long distances. Consequently, a key question is to identify the forces which helped determine the spatial features of Australian urbanisation: Were they determined by first-nature geography of the Australian landscape, or were they influenced, in a second-nature fashion, by *Songlines*?

In order to address the above question, we assemble a large dataset of geographical and climatic variables at the grid level to compare grid cells independently of their inherent natural features. This allows us to compare areas with otherwise similar characteristics that vary with respect to proximity to Aboriginal trade routes. The majority of the exogenous variables are commonly used in the economic development, history, and geography literatures. These variables include the coordinates of each grid cell, agricultural suitability, elevation, ruggedness (standard deviation of elevation), rainfall and standard deviation of rainfall, temperature and standard deviation of temperature, distance to the sea, and a binary variable that identifies whether the cell

[13] We use the "vcm-orm-ntl" version where the background pixels (without light) are set equal to zero.

[14] This measure can also address the phenomenon of blooming: the spillover of light onto cells tangential to the "actual" light. In particular, to minimise the potential for spurious results due to the spread of nocturnal light, the measure increases the accuracy of our empirical analysis, reducing the problem of brightness.

abuts the coast.[15]

In order to further control for potential confounding factors, we introduce distance from Sydney and state capitals as additional controls. These variables, while not fully exogenous to each cell, are likely to have a significant impact even at a distance. Moreover, we use the distance of each cell to the nearest historical mine as an additional control variable.[16] We have excluded active mines, because their number and location could be correlated with contemporary economic activity in the respective state, which would make them endogenous and thus "bad" controls. Still, distance to a historical mine is also relevant, as it takes into account the gold rush period in the 19th century, which had a great impact on the economy of Australia at the local level.

We also control for water presence in each cell as a proportion of total area, providing an exogenous measure of water availability. This variable accounts for suitability to such activities as shipping, agriculture, and fishing, let alone human survival.[17]

One concern is that Aboriginal people may have followed specific environmental conditions while travelling. We address this concern by correlating geoclimatic variables one by one with *Songlines* in Table 1. Results show a positive but insignificant association between *Songlines* and distance from the sea, contrary to expectation since coastal areas are particularly suitable for human habitation. The water percentage variable exhibits a positive correlation, implying that they likely sought places with more inland water. Elevation appears to be relevant for Aboriginal travellers, not surprisingly because it would make walking more cumbersome and so would ruggedness. Interestingly, Aboriginal people did not necessarily orientate their routes towards fertile plains and rain-prone areas, as the corresponding coefficients (for suitability for agriculture and precipitation) are not significant.

Table 1 does not support the hypothesis that the Aboriginal people necessarily sought access to specific environmental conditions, since their traditional routes could also lead through completely dry and non-plain regions.

---

[15]While the first layer of coastal grid cells has been excluded due to data limitations as explained in Section 3.1, we have created a coastal dummy using the second layer of coastal cells as these cells still offer valuable insight into the proximity of places to the coast.

[16]The historical mines were selected because the less advanced technology of the early 20th century suggests distances from those mines are irrelevant for the modern economy, given today's transport technology and infrastructure and therefore likely to be exogenous.

[17]We have also constructed two additional control variables to account for coastal trade patterns: historical shipwrecks and connectedness. These variables aim to capture the level of coastal trade activity and pre-industrial trading opportunities, respectively. More information about these variables, along with their sources, can be found in the Appendix, Table B.1.

Table 1. Analysis of balance: Environment and *Songlines*

	Dependent Variable: Binary variable for *Songlines*						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Agriculture Suitability	0.020						
	(0.015)						
Elevation		-0.000^{***}					
		(0.000)					
Ruggedness			-0.001^{***}				
			(0.000)				
Precipitation				-0.000			
				(0.000)			
Temperature					0.000		
					(0.000)		
Distance to the Sea						0.001	
						(0.001)	
Water Percentage							0.260^{***}
							(0.099)
N	79731	79731	79731	79731	79731	79731	79731
R^2	0.002	0.035	0.003	0.000	0.000	0.003	0.005

Notes: The unit of observation is a grid cell of 10km X 10km. The dependent variable is a dummy variable that takes the value of 1 if cell *i* has at least one *Songline*, and 0 otherwise. This variable was constructed using anthropological data from Mćcarthy (1939), which describes the trade sites created by Aboriginal people based on oral traditions prior to colonisation. The Human Mobility Index (HMI) from Özak (2018) was used to identify optimal routes between origins and destinations, with a least-cost algorithm. Each column includes a different geographical and climatic variable as follows: Column (1): agricultural suitability; column (2): Elevation; column (3): ruggedness; column (4): precipitation; column (5): temperature; column (6): distance to the sea; column (7): water percentage. The descriptions of the geographical and climatic variables can be found in the Appendix. Robust standard errors clustered at the local government district level were used and the constant term was omitted for space. *, ** and *** mean that the coefficient is statistically significant at 10%, 5% and 1% respectively.

3.4. Empirical Strategy

To examine the impact of *Songlines* on contemporary economic activity, we estimate the following regression, with standard errors clustered at the local government level:

$$Nightlight_i = \alpha_i + \beta Songlines_i + Z_i^{\prime}\rho + \varepsilon_i, \tag{1}$$

where i indexes grid cells for Australia. Our dependent variable $Nightlight_i$ is a dummy variable that takes the value 1, if grid cell i has strictly positive nightlight (and thus economic activity), and 0 otherwise. $Songlines_i$, our main variable of interest, takes the value 1, if grid cell i hosts at least one *Songline*, and 0 otherwise. Z'_i is a vector of climatic, topographic and geographic variables. In addition, our model includes fixed effects for local government districts, α_i, in order to capture any unobserved local characteristics such as market institutions. The coefficient of interest, β, reflects the impact of *Songlines* on current economic activity.

3.4.1. Main Results

Column (1) of Table 2 illustrates the relationship between *Songlines* and contemporary economic activity. Results are positive and significant at the 1% level. We use latitude and longitude in column (2), and add geographical and climatic control variables in columns (3) and (4). All variables enter the model with the expected sign.[18] The most conservative coefficient in column (4) suggests that the probability of having economic activity in a grid cell is 3% greater when a *Songline* is present.

Column (5) of Table 2 reports results with population density as an alternative measure of economic activity at the local level. Remarkably, the coefficient on *Songlines* remains positive and highly significant. Notably, the R-squared in column (5) is larger than in column (4). This difference can be attributed to the fact that our Night Lights measure is discrete and the alternative proxy, Population Density, is continuous. Nevertheless, the fact that the key explanatory variable, *Songlines*, performs equally well confirms the robustness of our approach. Moreover, a larger R-squared in column (5) suggests that *Songlines* did not only establish the foundations of local economic activity, encompassing even non-urban enclaves such as remote manufacturing and mining outposts, but also presence of enduring "magnet" effects on current settlements long after the *Songlines* were composed.[19]

[18]While we have excluded the first layer of coastal grid cells from the sample due to lack of data, Appendix, Section B, reports results that we also consider coastal trade dynamics.

[19]In Table A.2 in the Appendix, we use an alternative variable for our *Songlines*. This variable

Table 2. Baseline results

	Night Lights				Pop. Density
	(1)	(2)	(3)	(4)	(5)
Songlines	0.035***	0.036***	0.034***	0.030***	0.975***
	(0.007)	(0.007)	(0.006)	(0.006)	(0.148)
Latitude		−0.000**	−0.002***	−0.003***	−0.044*
		(0.000)	(0.000)	(0.001)	(0.026)
Longitude		−0.000	0.000	0.003***	0.071**
		(0.000)	(0.000)	(0.001)	(0.033)
Agriculture Suitability			−0.002	−0.003**	−0.151***
			(0.002)	(0.002)	(0.043)
Elevation			0.000	0.000	−0.001
			(0.000)	(0.000)	(0.002)
Ruggedness			0.000**	0.000	0.009
			(0.000)	(0.000)	(0.006)
Precipitation			0.000***	0.000	0.001
			(0.000)	(0.000)	(0.001)
StDev Precipitation			0.000	0.000	0.017
			(0.000)	(0.000)	(0.012)
Temperature			0.002***	0.002**	−0.006
			(0.001)	(0.001)	(0.030)
StDev Temperature			−0.024***	−0.024***	−0.356***
			(0.005)	(0.005)	(0.121)
Coastal Dummy				0.065***	1.059***
				(0.014)	(0.289)
Distance to the Sea				−0.000	−0.015
				(0.000)	(0.010)
Distance to Sydney				0.003***	0.064*
				(0.001)	(0.035)
Distance to State Capital				0.000	−0.020***
				(0.000)	(0.007)
Distance to Historical Mine				−0.002***	−0.056***
				(0.000)	(0.010)
Water Percentage				0.007	−0.681
				(0.020)	(0.602)
Local Gov FE	✓	✓	✓	✓	✓
N	79731	79731	79731	79731	79731
R^2	0.222	0.222	0.228	0.233	0.600

Notes: The unit of observation is a grid cell of 10km X 10km. In columns (1)-(4), the dependent variable is a dummy variable that takes the value of 1 if cell i has nightlight, and 0 otherwise. This variable uses data from Visible Infrared Imaging Radiometer Suite (VIIRS) sensor. In column (5), the dependent variable is the logarithm of human population density, which is the number of individuals per grid cell. The data is derived from national censuses and population registers for the year 2015, with an additional minor adjustment of 1e-8 included. All columns control for local government fixed effects. Columns (1)-(5) include a measure of *Songlines*. This measure is a dummy variable that takes the value of 1 if cell i has at least one *Songline*, and 0 otherwise. This variable was constructed using anthropological data from McCarthy (1939), which describes the trade sites created by Aboriginal people based on oral traditions prior to colonisation. The Human Mobility Index (HMI) from Özak (2018) was used to identify optimal routes between origins and destinations, with a least-cost algorithm. From columns (2) to (5), geographical and climatic control variables were added gradually and incrementally. The descriptions of the geographical and climatic variables can be found in the Appendix. Robust standard errors clustered at the local government district level were used and the constant term was omitted for space. *, ** and *** mean that the coefficient is statistically significant at 10%, 5% and 1% respectively.

3.4.2. Causal results: Contiguous Pairs Analysis

We now investigate whether the presence of *Songlines* causally influences current economic activity in Australia. To this end, we utilise an approach similar to Dube et al. (2010) and Michalopoulos and Papaioannou (2013) and perform a contiguous pair analysis. Specifically, we develop a estimation model that includes *all* pairs consisting of $10km \times 10km$ grid cells hosting *Songlines and* their neighbouring grid cells without them. For example, if grid cell A has a *Songline* and is adjacent to or shares a corner with grid cells B and C without one, both pairs will be included in our sample. Our specification for adjacent grid cells includes a fixed effect for each pair in order to attribute any difference in contemporary economic activity between the two grid cells to *Songlines*, i.e. accounting for all unobserved factors among neighbouring grid cells, except for the effects of *Songlines*. The specification for contiguous pair grid cells is as follows:

$$Nightlight_{i,(j)} = \alpha_{i,(j)} + \beta Songlines_{i,(j)} + Z'_{i,(j)}\rho + \varepsilon_{i,(j)}. \tag{2}$$

In equation (2), our dependent variable is a dummy variable that takes the value 1 when there is nightlight in grid cell i, which contains *Songlines*, and is also adjacent to grid cell j, which does not contain one. The same structure applies to the $Songlines_{i,(j)}$ and all control variables $Z'_{i,(j)}$. Finally, $\alpha_{i,(j)}$ are the pairs grid cell-fixed effects.

Table 3 presents the results. The coefficient for $Songlines$ is positive and statistically significant at the 99% confidence level. This finding indicates that grid cells with *Songlines* show higher levels of economic activity compared to their neighbouring grid cells without them, even after controlling for all shared unobserved characteristics.

4. Adoption of Aboriginal Knowledge, Geography, or Path Dependence?

In this section, we turn our attention to the plausibility of our identification assumption. As evident in section 2, we contend that it is the adoption of Aboriginal knowledge of the landscape that influenced European exploration and settlement. The construction of transportation infrastructure that followed locked in path dependence, which shaped the spatial distribution of contemporary economic activity in Australia in profound ways.

is constructed using the digitised Mc̒Carthy (1939)'s maps, which illustrate some of the routes described in their work. Figure A.1 shows the *Songlines* based on Mc̒Carthy (1939)'s maps. While we recognise the limitations of early 20th-century cartographic tools, we utilise this data to validate the results obtained from our estimated routes. The alternative variable is a binary variable that assigns a value of 1 if a *Songline* intersects a grid cell, and 0 otherwise. Reassuringly, the results corroborate those from Table 2.

Table 3. Causal evidence

	Dependent Variable: Binary variable for night lights
	(1)
Songlines	0.014***
	(0.002)
Local Gov. FE	✓
Geo & Hist Controls	✓
N	115024
R^2	0.465

Notes: The unit of observation is a grid cell of 10km X 10km. The dependent variable is a dummy variable that takes the value of 1 if cell *i* has nightlight, and 0 otherwise. This variable uses data from Visible Infrared Imaging Radiometer Suite (VIIRS) sensor. All columns control for local government fixed effects, as well as for a set of geographical, climatic and historical variables, which include: the coordinates of each grid cell, agricultural suitability, elevation, ruggedness (standard deviation of elevation), rainfall and standard deviation of rainfall, temperature and standard deviation of temperature, distance to the sea, distance to Sydney, distance the state capital, distance to historical mines, coastal presence, and water percentage. The descriptions of the geographical, climatic, and historical variables can be found in the Appendix. The regression in the table is based on an analysis of contiguous pairs, which include pair-fixed effects, following equation equation 2 from section 3.4.2. Robust standard errors clustered at the local government district level were used and the constant term was omitted for space. *, ** and *** mean that the coefficient is statistically significant at 10%, 5% and 1% respectively.

However, an important concern arising from our empirical findings relates to the distinction between the adoption of Aboriginal knowledge, on the one hand, and the direct effects of inherent characteristics of the Australian topography and the indirect effects of how the Aborigines gained that knowledge, on the other. Specifically, it remains essential to assess whether our results genuinely reflect such adoption or they can be attributed to other favorable environmental attributes that may have guided European exploration of Australia. In such a scenario, Europeans did not have to adopt the routes used by Aboriginal people. For example, while Europeans established early cities in arid regions like Alice Springs, whose development was due to the Overland Telegraph Line in the 19th century, rather than to the Gold Rush. Other areas, such as the Daintree Rainforest and Cape York Peninsula, remained largely unsettled or sparsely populated well into the 21st century. Therefore, we seek to demonstrate empirically that Australia's economic trajectory would have followed different geographical patterns if Europeans had not adopted Aboriginal knowledge during colonisation. A contrary finding would suggest that *Songlines* had limited or insignificant influence on how Europeans chose sites for exploration and settlement.

To address these concerns and in view of the detailed historical evidence discussed in section 2, we present two approaches in the subsequent subsections.

4.1. Aboriginal Knowledge or Just Geography?

The first approach leverages exogenous variation in European exploration and settlement across Australia. To this end, we follow Barjamovic et al. (2019) and construct *Natural Routes* in Australia. We define these *Natural Routes* as optimal travel routes derived from environmental conditions that characterize inland Australia from *each to every other* coastal grid cell. To achieve this, we train our least-coast path algorithm to identify them as those that are feasible for humans to traverse between *all* pairs of coastal grid cells. This produced more than 2 million *Natural Routes* in Australia. The number of intersections of these optimal travel routes within each grid cell indicates the density of *Natural Routes*, its degree of connectedness. Our measure of *Natural Routes* is computed as the logarithm of the number of such optimal routes that intersect with each grid cell.

Figure 2 shows the heat map with the values for *Natural Routes*, where high density means higher connectedness. Since this measure uses the HMI index as a ”travel cost” parameter, it can be considered exogenous to any human action. Thus, *our Natural Routes* variable exogenously represents optimal locations for travel and settlement in Australia.

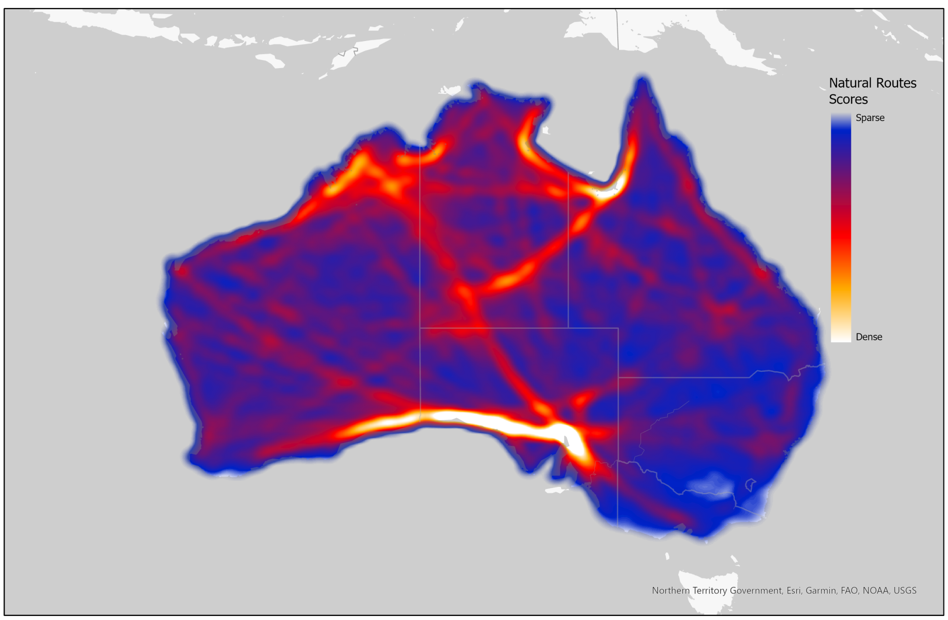


Figure 2. Natural Routes Scores

Notes: This heat map shows the ratings of *Natural Routes* in Australia. The *Routes* were created using a least cost path algorithm to identify optimal travel routes from each coastal pixel to the interior. The *Routes* were optimally constructed by the authors using the Human Mobility Index of Özak (2018).

We begin by testing the extent to which Europeans relied on *Natural Routes* or *Songlines* when exploring Australia. To do so, we digitise the map published by Robinson (1927) showing all early European inland explorations routes from 1813 to 1901 (see Figure A.2 in the Appendix).[20] We then construct a dummy variable, *Exploration Routes,* that takes the value 1 if there is at least one European exploration route in a grid cell and 0 otherwise, and use it as a dependent variable.

Column (1) of Table 4 presents the results of a"horse race" analysis, comparing the predictive performance of *Natural Routes* and *Songlines* in explaining *Exploration Routes,* The reported coefficients in Table 4 are standardised, which allows us to obtain a direct comparison of the relative importance of these two variables. Therefore, these coefficients reflect the effect on the outcome variables associated with a one-standard deviation increase in the predictor variables.

Table 4. *Natural Routes versus Songlines*

	Exploration Routes	Night Lights	Pop. Density
	(1)	(2)	(3)
Natural Routes	0.006*	0.004***	0.099***
	(0.003)	(0.001)	(0.032)
Songlines	0.017***	0.009***	0.282***
	(0.003)	(0.002)	(0.043)
Local Gov FE	✓	✓	✓
Geo & Hist Controls	✓	✓	✓
N	79731	79731	79731
R^2	0.069	0.233	0.600

Notes: The unit of observation is a grid cell of 10km X 10km. OLS estimates are shown with robust standard errors clustered at the local government district level. All columns show standardised coefficients. *Natural Routes* is the logarithm of the number of optimal travel routes that intersects within each pixel. *Songlines* is a dummy variable that takes the value of 1 if cell *i* has at least one *Songline*, and 0 otherwise. In column (1), the dependent variable is an indicator that takes the value of 1 if there exists at least one European exploration route in a grid cell, and 0, otherwise. In column (2), the dependent variable is an indicator that takes the value of 1 if grid cell *i* has nightlight, and 0 otherwise. In column (3), the dependent variable is the logarithm of human population density, which is the number of individuals per grid cell. The data is derived from national censuses and population registers for the year 2015, with an additional minor adjustment of 1e-8 included. All columns control for local government fixed effects, as well as for a set of geographical, climatic and historical variables, which include: the coordinates of each grid cell, agricultural suitability, elevation, ruggedness (standard deviation of elevation), rainfall and standard deviation of rainfall, temperature and standard deviation of temperature, distance to the sea, distance to Sydney, distance the state capital, distance to historical mines, coastal presence, and water percentage. The descriptions of the geographical, climatic, and historical variables can be found in the Appendix. The constant term was omitted for space. *, ** and *** mean that the coefficient is statistically significant at 10%, 5% and 1% respectively.

The results are notable. We observe that a one-standard deviation increase in *Natural Routes* corresponds to a significantly smaller effect (approximately one-third) on Exploration Routes compared to our main measure of *Songlines*. Moreover, it is noteworthy that the coefficient on *Natural Routes* shows statistical significance only at the margins, while the estimate for *Songlines* remains significant at the 1% level.

Table 4, columns (2) and (3), report the effects of these variables when they are jointly present as regressors on contemporary economic activity. The dependent variable in column (2) is the indicator for night lights and in column (3) the outcome is the measure of population density. We find that while both coefficients show statistically significant

[20] Although this is a detailed map showing both nodes and sections of the early explorations, we should also note the limited technology of geographical tools in the early 20th century.

effects, the coefficient for *Natural Routes* is having approximately half as large as the effect attributable to *Songlines*.[21,22]

These results provide compelling support for the hypothesis that the adoption of Aboriginal knowledge by Europeans during colonisation significantly contributed to the spatial patterns of modern economic activity in Australia. However, it is crucial to acknowledge that since both *Natural Routes* and *Songlines* are statistically significant, other factors also contribute to our understanding of how Australia's economy unfolded over time.

4.2. Aboriginal Knowledge or Just Path Dependence?

Next, we discuss a second approach to validate our identification strategy. Specifically, we explore the potential impact of local "frictions" that may have hindered the ability of Europeans to adopt Aboriginal knowledge during colonisation. We do so by accounting for the locations where massacres have been documented to have occurred. Those are places where Europeans encountered hostility or resistance from Aboriginal people during exploration. If we were to find a positive and significant effect of Aboriginal paths on economic activity near massacre sites, that would plausibly contradict our claim that the level of contact itself (and the presumed transmission of knowledge) between Europeans and Aboriginal people influenced economic activity in the long run. In other words, we expect that in areas near sites of massacres, Europeans would have found it difficult to acquire local knowledge due to hostility with local tribes, even if they had initially recruited Aboriginal people at the start of their expeditions. In this case, the presence of a *Songline* should not have a positive impact on present-day economic activity. Therefore, focusing on outcomes in the segments of Aboriginal paths near massacre sites may serve as plausible counterfactuals to the routes we assume Europeans were aware of.

We test the above idea empirically by utilising data from Ryan et al. (2017). These researchers construct a database on massacres and their location from 1788 to 1930.[23] The dataset provides the places of

[21] We also examine (but do not report) the effects of *Songlines* on night lights, excluding *Natural Routes*, and vice versa. The *Songlines* yield a coefficient of 0.0166, significant at the 99% level, while the *Natural Routes* show a coefficient of 0.0099, significant at the 95% level. This corroborates the stronger influence of *Songlines* compared to *Natural routes*.

[22] In the Appendix, we also employ a different strategy to compare the effects of *Natural Routes* and *Songlines*. We construct a dummy variable which takes the number 1 only for those grid cells which are intersected by a high number (above the 90th percentile) of travel optimal paths constructed as explained above. We then replicate the "horse-race" analysis. Table A.15 supports the findings of Table 4, as the coefficients of *Songlines* are consistently higher and more statistically significant than those of *Natural routes*.

[23] The authors distinguish between "reprisal" and "opportunity" massacres. The first category refers to a massacre that was carried out in response to the alleged killing or kidnapping of an

426 massacres occuring between 1788-1930, of which almost all, i.e. 412, were against Aboriginal people. Keeping all reprisal massacres that proxy for extreme hostility between Aboriginals and Europeans, we carry out our main analysis by keeping progressively observations, linked to grid cells, that are within 5km, 10km, 15 km, and 20 km, away from each massacre site.

Table 5. Adoption “Frictions”

	Dependent Variable: Dummy variable for night light			
	(1) grid cells within 5km	(2) grid cells within 10km	(3) grid cells within 15km	(4) grid cells within 20km
Songlines	0.258	0.051	0.066*	0.055**
	(0.172)	(0.064)	(0.037)	(0.028)
N	141	528	1173	1995
R^2	0.078	0.231	0.280	0.303

Notes: The unit of observation is a grid cell of 10km X 10km. The dependent variable is a dummy variable that takes the value of 1 if cell *i* has nightlight,and 0 otherwise. This variable uses data from the Visible Infrared Imaging Radiometer Suite (VIIRS) sensor. The analysis only includes pixels that are 5 km (column (1)), 10km (column (2), 15km (column (3), and 20km (column (4)) closer to a massacre. All columns control for local government fixed effects, as well as for a set of geographical, climatic and historical variables, which include: the coordinates of each grid cell, agricultural suitability, elevation, ruggedness (standard deviation of elevation), rainfall and standard deviation of rainfall, temperature and standard deviation of temperature, distance to the sea, distance to Sydney, distance the state capital, distance to historical mines, coastal presence, and water percentage. The descriptions of the geographical, climatic, and historical variables can be found in the Appendix. The constant term was omitted for space. *, ** and *** mean that the coefficient is statistically significant at 10%, 5% and 1% respectively.

Interestingly, Table 5 shows that the effect of *Songlines* on Night Lights is non-significant if we keep observations up to 10km from the hostility sites in columns (1) and (2). However, if we increase the radius to 15km in column (3), the coefficient of interest becomes slightly significant at the 90% confidence level, while when we increase the radius up to 20km in column (4), their effect becomes larger and significant at the 95% confidence level. Consequently, we may argue that there is a strong connection between ease of acquisition of knowledge of Aboriginal paths by Europeans and the location of contemporary economic activity.[24]

5. Urban Evolution of Australia

In this section, we examine the dynamic effects of *Songlines* on new settlements in Australia. This analysis allows us to investigate the possible shift in informational significance of *Songlines* in the face of other historical shocks, such as the gold rush between the 1850s and 1910s, World Wars I and II, and immigration in the 20th century. This is important to consider because while early European settlements such as

Aboriginal person (among other hostilities) while the second refers to a massacre in response to less violent events such as Preventing Aboriginal people from accessing a waterhole or ceremonial ground.

[24]As an additional test, in the Appendix Table A.3, we exclude the grid cells close to massacres sites up to 20km. Our coefficient on *Songlines* remains positive and highly significant.

Table 6. Establishment of Cities

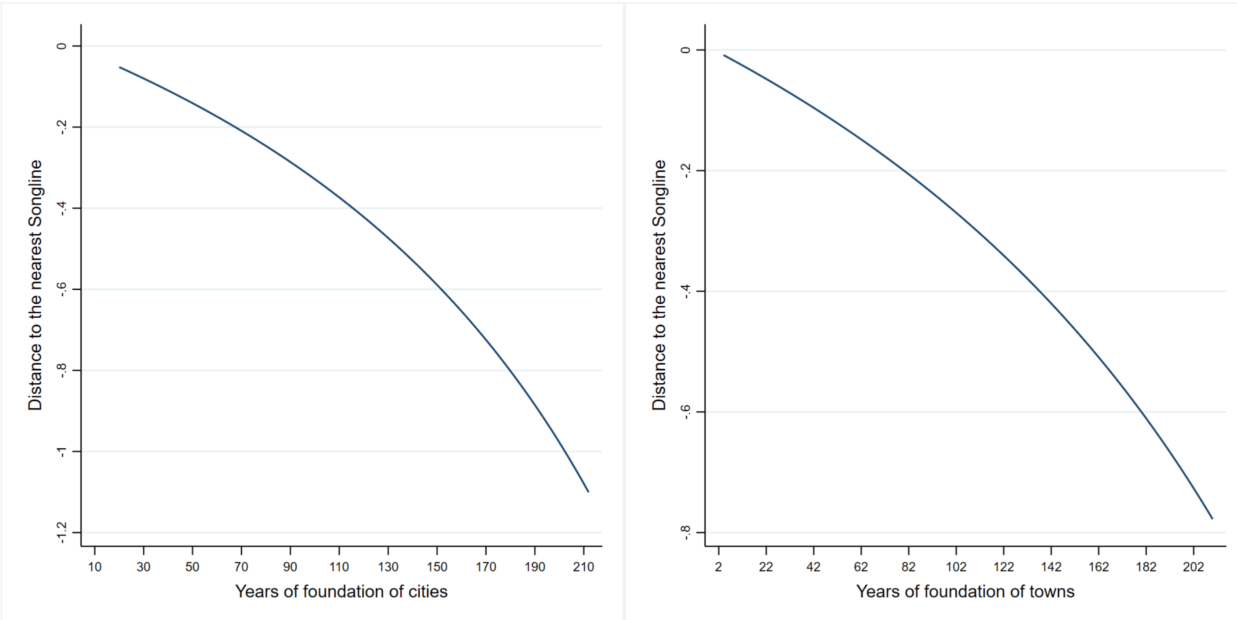
	All Cities	Cities est. before 1850	Cities est. in 1850-1900	Cities est. in 1900-1950	Cities est. in 1950-2000
	(1)	(2)	(3)	(4)	(5)
Songlines	0.004***	0.001*	0.002***	0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
Local Gov FE	✓	✓	✓	✓	✓
Geo & Hist Controls	✓	✓	✓	✓	✓
N	79731	79731	79731	79731	79731
R^2	0.054	0.089	0.029	0.018	0.010

Notes: The unit of observation is a grid cell of $10km \times 10km$. OLS estimates are shown with robust standard errors clustered at the local government district level. In column (1), the dependent variable is a dummy variable that takes the value of 1 if grid cell i has a major city that was established by Europeans between 1788 and 2000 in Australia. In column (2), the dependent variable is a dummy variable that takes the value of 1 if grid cell i has a major city that was established by Europeans between 1788 and 1850. In column (3), the dependent variable is a dummy variable that takes the value of 1 if grid cell i has a major city that was established by Europeans between 1850 and 1900. In column (4), the dependent variable is a dummy variable that takes the value of 1 if grid cell i has a major city that was established by Europeans between 1900 and 1950. In column (5), the dependent variable is a dummy variable that takes the value of 1 if grid cell i has a major city that was established by Europeans between 1950 and 2000. Dummy variables from all columns were constructed using data from Kampanelis (2019) on the year of establishment of 249 major cities in Australia. Songlines are defined in the same way as in Table 2. All columns control for local government fixed effects, as well as for a set of geographical, climatic and historical variables, which include: the coordinates of each grid cell, agricultural suitability, elevation, ruggedness (standard deviation of elevation), rainfall and standard deviation of rainfall, temperature and standard deviation of temperature, distance to the sea, distance to Sydney, distance the state capital, distance to historical mines, coastal presence, and water percentage. The descriptions of the geographical, climatic, and historical variables can be found in the Appendix. The constant term was omitted for space. *, ** and *** mean that the coefficient is statistically significant at 10%, 5% and 1% respectively.

Sydney or Melbourne may have had their roots in Aboriginal settlements, this is less likely to be the case for later settlements.

To examine the above argument, we use data from Kampanelis (2019) on the year of establishment of 249 major cities in Australia, which comprise a large portion of the total population. To analyse this data, we partition the cities into four groups according to their year of establishment: cities established between 1788 and 1850, 1850 and 1900, 1900 and 1950, and 1950 and 2000. We then create four dummy variables for each grid cell by assigning the value 1 if grid cell i is associated with one of these groups of cities and use them as dependent variables.

Table 6 reports results of regressing the dummy variables for the year when a city was founded against the presence of *Songlines*. Results in column (1), which includes all cities, are positive and significant at the 99% confidence level. Columns 2 and 3 show that *Songlines* have a positive and significant effect on cities established before 1900. However, this effect does not extend to later years; cities established after 1900 do not show a significant effect from them. This finding accords with the notion that the likely effects of technological progress, population growth, oil discoveries and other factors, which have enabled the establishment of cities in locations that were previously inhospitable for human settlement, have rendered obsolete the informational advantage of Aboriginal knowledge.



(a) Emergence of cities and proximity to *Songlines*

(b) Emergence of towns and proximity to *Songlines*

Figure 3. Urban evolution and proximity to *Songlines*

Notes: These figures show the marginal association between the distance to the nearest *Songline* to major cities (Figure 3a) and towns (Figure 3b) and their corresponding year of establishment in Australia. Data on the foundation years of cities comes from Kampanelis (2019), and the data on the foundation years of towns comes from Elder (2024).

To further investigate the spatial expansion of the Australian urban system in relation to *Songlines*, we first calculate the distance of each city from the nearest *Songline*. We then examine the relationship between these distances and the founding year of each city by modelling how the establishment of cities has evolved over time relative to their proximity to *Songlines*. The results in Figure 3a demonstrate a concave decreasing pattern. This finding implies that as European colonisation evolved, newer settlements are located further away from *Songlines*, with the marginal increase in the distance decreasing as time since foundation decreases. We interpret this result as evidence of *accelerating spatial agglomeration* around *Songlines* as the urban system thickens.

Certainly, given that Australia's largest cities are predominantly coastal, they possibly benefited more from early interactions between explorers and Aboriginal communities. To address this issue, we extend our analysis to the evolution of economic activity into inland and typically arid regions, which may have developed under different patterns, not strictly linked to the *Songlines*.

To do so, we draw data from Elder (2024), who provides detailed information on the years of establishment of 1,143 towns. From this dataset, we select 866 inland towns, each situated over 10 kilometers from the nearest coastline. Additionally, we exclude 127 towns established close to cities already examined in the preceding analysis. The results from Figure 3b align with our earlier findings (t-statistic=-2.64), showing that towns closer to the *Songlines* emerged earlier. This supports the idea that the *Songlines* acted as focal points for the development of economic activity for more than two centuries. Even though inland areas are not as thickly settled as coastal ones, and the marginal effect of year of settlement weakens over time, clustering near *Songlines* is still present.

6. Robustness Checks

6.1. Homogenising the Sample

We now focus on the robustness of our baseline results. We begin by addressing the inhomogeneity that characterises the Australian landscape. Our goal is to confirm that our estimates remain robust and not affected by the heterogeneity of regions, including those with arid desert terrain. To achieve this, we excluded the top and bottom 25% of grid cells with the highest temperature and the lowest levels of precipitation, respectively.

Columns (1) and (2) of Table A.4 in the Appendix demonstrate that our primary variable of interest, *Songlines*, remains statistically significant even after homogenising the sample by eliminating the grid cells below the 25-percentile of precipitation. Furthermore, its magnitude is comparable to the coefficient of the baseline regression (4) in Table 2, which suggests that there is no notable difference between coastal and inland locations, that is more and less hospitable, areas in Australia, in the effects of *Songlines*.

6.2. Neighbouring Analysis

6.2.1. Do Routes through Neighbouring Cells Matter?

We proceed by employing a more conservative grid cell selection, That is, we include *only* grid cells intersected by *Songlines and* their respective *neighbouring* grid cells that may or may not contain one. This approach allows us to limit the influence of cells that are far away from *Songlines*, such as uninhabited deserts, which could act as a "noisy" control group in our primary sample.

Table A.6 provides evidence for our hypothesis using our restricted sample. Coefficients of the main variables of interest are positive and statistically significant. Fixed effects for local governments are included for all grid cells along Aboriginal paths and their neighbourhoods, facilitating comparison of environmentally similar grid cells with the only difference being the presence of *Songlines*.

6.3. Measurement Error and Alternative Proxies

6.3.1. Does Grid Cell Size Matter?

Our reliance on $10km \times 10km$ grid cells is another source of concern, possibly due to measurement errors resulting from the fact that our main binary variable is constructed obased on the Human Mobility Index (HMI). We address this by working with $50km \times 50km$ grid cells so as to account for potential miscoding in the construction of *Songlines*. This provides us with areas of 2500 km^2 as the units of analysis. Table A.5 shows positive and significant impacts of *Songlines* on current economic activity in all columns, with column (4) representing our most conservative specification, where only the grid cell size differs, but all covariates, which are defined as in regressions reported in Table 2, remaining the same. The effect is larger than the corresponding value, reported in Table 2, suggesting that our original estimates are closer to the lower bound.

6.3.2. Alternative Sources of Measurement Error

A potentially greater source of measurement error is that our baseline results rely on just one source of data, namely Mćarthy (1939), for the location of *Songlines*. This suggests a possible limitation, as the HMI may not fully align with the travel patterns of Aboriginal people. To address this issue, we employ a number of strategies. First, we gather coordinate data from McDonald and Clayton (2016) on the locations of Aboriginal rock arts to proxy for the existence of human activity in Australia prior to European arrival.[25,26] We use the Aboriginal rock art data and the HMI as the main parameters in our weighted least-cost path algorithm and construct a dummy variable for the Aboriginal paths. In Table A.7, column (1), the new dummy variable has a statistically significant positive coefficient, indicating that our results are consistent with alternative data sources on the location of *Songlines*.

In columns (2) and (3) of Table A.7, we use as weights either the Mćarthy (1939)'s digitised maps and the rock art data, or just the Mćarthy (1939)'s digitised maps, respectively. The results in both columns indicate that *Songlines* impact current economic activity.

Finally, we conduct an analysis to assess to what extent our main *Songlines* dataset based on the HMI overlaps with the three other datasets of the *Songlines* based on, one, the HMI & Rock Art (RA); two, McCarthy's historical maps (DIG) & RA; and three, DIG. For example,

[25] For areas with multiple rock art sites, the centroid of the area is calculated to serve as a gravity point for our algorithm.

[26] The rock art data represent a synthesis of heritage places with rock art(s) included in the World Heritage List (WHL), Commonwealth Heritage List (CHL), National Heritage List (NHL), and Register of the National Estate (RNE).

using our main dataset, which comprises only data from the HMI, and a dataset containing information from both the HMI and the RA, we keep grid cells whose locations appear in both datasets. Columns (4)-(6) of Table A.7 indicate that, even when excluding *Songlines* with potential measurement errors, our findings remain consistent with the baseline results. All coefficients for *Songlines* are positive and statistically significant at the 1% level.

6.4. Trading Sites

Although we have sought to address measurement error in reconstructing the *Songlines* in the previous section in several ways, it might still be argued that measurement error may still be a concern. Specifically, it may be argued that the routes calculated by the least cost path algorithm using the HMI might not plausibly be the actual trails that Aboriginal people used to travel on. To test the credibility of our variable of interest, we focus on all starting and destination sites documented by MćCarthy (1939). Specifically, we investigate whether contemporary economic activity is also associated with locations where Aboriginal people traded goods in the past.

Moreover, we believe that the above analysis allows for a sharper identification of the effects of *Songlines* on contemporary economic activity. Since our data on trading sites (i.e., starting or destination points of *Songlines*) covers the entire Australian territory, we can exclude all coastal areas, thereby focusing only on sites in the interior, which was the most unfamiliar and inhospitable to traverse terrain for European colonisers. Any association between early trading sites and contemporary economic activity would therefore strongly indicate that Europeans selected their settlement locations based on Aboriginal knowledge of the landscape.

Therefore, we conduct an alternative estimation of equation (1) by replacing our main explanatory variable of interest, the *Songlines*, with a dummy variable that takes the value of 1, if at least one Aboriginal trading site is present in a grid cell, and 0 otherwise. Anticipating our detailed analysis, we think that by using recorded Aboriginal trading sites we avoid relying on constructed routes and thus rely on direct evidence, in the form of a smaller but directly observed information set less prone to measurement error, in support of our claim. In Panel A of Table 7, column (1) shows a positive and strong relationship between Trading Sites and Night Lights when all grid cells in the sample are used.

We extend the analysis in columns (2)-(6) of Table 7 in Panel A by gradually excluding all grid cells that are 50 km, 100 km, 200 km, 300 km and 500 km from the coastline, respectively, to remove the potential influence of proximity to the coast. Column (2) shows the results

Table 7. Trading Sites

Panel A:	Dependent Variable: Binary dummy for night light					
	(1) all cells	(2) excl. 50km	(3) excl. 100km	(4) excl. 200km	(5) excl. 300km	(6) excl. 500km
Trading Sites	0.197***	0.168***	0.168***	0.169***	0.159***	0.070***
	(0.023)	(0.026)	(0.028)	(0.029)	(0.033)	(0.020)
Local Gov FE	✓	✓	✓	✓	✓	✓
Geo & Hist Controls	✓	✓	✓	✓	✓	✓
N	79731	72474	65382	52803	41838	23744
R^2	0.234	0.167	0.151	0.113	0.063	0.032

Panel B:	Dependent Variable: Population Density					
	(1) all cells	(2) excl. 50km	(3) excl. 100km	(4) excl. 200km	(5) excl. 300km	(6) excl. 500km
Trading Sites	4.087***	3.863***	3.814***	3.909***	3.692***	3.051***
	(0.358)	(0.429)	(0.469)	(0.569)	(0.607)	(0.439)
Local Gov FE	✓	✓	✓	✓	✓	✓
Geo & Hist Controls	✓	✓	✓	✓	✓	✓
N	79731	72474	65382	52803	41838	23744
R^2	0.600	0.582	0.570	0.529	0.477	0.281

Notes: The unit of observation is a grid cell of 10km X 10km. In Panel A, the dependent variable is a dummy variable that takes the value of 1 if cell i has nightlight,and 0 otherwise. This variable uses data from the Visible Infrared Imaging Radiometer Suite (VIIRS) sensor. In Panel B, the dependent variable is the logarithm of human population density, which is the number of individuals per grid cell. The data is derived from national censuses and population registers for the year 2015, with an additional minor adjustment of 1e-8 included. Trading Sites is a dummy variable that takes the value of 1 if at least one Aboriginal trading point is present in each cell and 0 otherwise. In column (1), all grid cells in the sample are used. From columns (2)-(6), cells are progressively excluded that are 50km (col. 2), 100km (col 3), 200km (col 4), 300km (col 5) and 500km (col 6) close to the shoreline. All columns control for local government fixed effects, as well as for a set of geographical, climatic and historical variables, which include: the coordinates of each grid cell, agricultural suitability, elevation, ruggedness (standard deviation of elevation), rainfall and standard deviation of rainfall, temperature and standard deviation of temperature, distance to the sea, distance to Sydney, distance the state capital, distance to historical mines, coastal presence, and water percentage. The descriptions of the geographical, climatic, and historical variables can be found in the Appendix. Robust standard errors clustered at the local government district level were used and the constant term was omitted for space. *, ** and *** mean that the coefficient is statistically significant at 10%, 5% and 1% respectively.

with the least restrictive sample, which excludes only grid cells within 50 km of the coastline. As a result, 275 trading sites are excluded. These sites are marked by blue triangles in Figure 1. The respective coefficient is positive and significant, indicating that our results are not predominantly influenced by proximity to the coast. Notably, this is a remarkable result given that major Australian cities such as Sydney and Melbourne, which might coexist with coastal trading sites due to their natural resource abundance, have already been dropped. The next four columns report results with successively more restricted samples, leaving up to only 25% of the interior of Australia. However, the respective coefficients remain positive and significant. Therefore, we argue that we have dispelled concerns that economic activity by European colonisers and pre-colonial Aboriginal trading activity coincided randomly.

In Panel B of Table 7, we repeat the same analysis from Panel A with population density as the dependent variable instead of Night Lights. The coefficient on Trading Sites once more demonstrates statistical significance at the 1% threshold. This analysis effectively allays several important concerns. First, it dispels doubts that our main outcome—Night Lights—is not accurately reflecting economic activity within inland locations across Australia. The observed brightness seems to stem not from places like mines, factories, or power plants, but predominantly from urban population centres. Second, we mitigate concerns that our results could be influenced by biases related to Australia's coastal development — a region where a significant portion of the Australian population has concentrated, and most notably along the Eastern Seaboard, ever since European settlers arrived. Specifically, 85% of Australian population lives within 50km of the coast, which we have actually excluded in column (2). Third, comparing across from column (1) to column (6), we see that the estimated coefficient of trading sites remains highly statistically significant and falls numerically by only 25%. This dispels the notion that the informational impact of trading sites on contemporary urban density is confounded by proximity to the coast.[27]

6.5. Songlines as a Network

We examine next the robustness of our *Songlines* variable by considering as alternative additional controls, that is network centrality measures that highlight network aspects of the system of Aboriginal routes. As we documented earlier, the *Songlines* were used for numer-

[27] We provide additional validation for these concerns by conducting an additional check. Table A.8 reports results with the same set of alternative samples using, as earlier, *Songlines* as the main variable of interest instead of Trading Sites. The coefficients reported in Table A.8 further strengthen the main results of the paper.

ous cultural activities in addition to trade. If we were to assume that the Aboriginal communities, associated with origins and destinations, used them for the purpose of exchange of commodities that were differentiated by origin, we may study the resulting trading system by the tools of international trade. That is, following Allen and Arkolakis (2014) and Chen et al. (2022), we may obtain a network description of the *Songlines*. This allows us to define centrality measures, based on the matrix with elements the inverse traveling costs, $\mathbf{T} = [1/\tau_{ij}]$. as computed earlier.

We examine the following alternative measures of centrality: degree, betweenness, Katz-Bonacich, and eigenvector centrality (Bloch et al., 2023). Degree is the most local of them, and in our setting is equivalent to the average inverse distance from each node to every other node. Eigenvector is the most global in the sense that it aggregates the centrality of every other node. Betweenness and Katz-Bonacich fall in between. The latter assigns a basic centrality to all nodes and accounts for the "lengths" of paths by discounting them in order obtain a finite measure. Its extreme values coincide with eigenvector centrality. An advantage of these controls is that they bring measures of the entire topography of the Songlines to bear on the estimation, instead of treating them as mere characteristics of grid cells, when the latter lack any geographical attributes. The network measures employed here may also be interpreted as *market access.*

Table A.9 in Appendix reports estimation results with those centrality measures as alternative controls. Interestingly, their inclusion always yields highly significant coefficients, and in most cases slightly strengthens the coefficient of *Songlines*. We argue that this reaffirms its robustness when spatial aspects of the travel network are controlled for.

6.6. Additional Robustness Checks

6.6.1. Trading Routes or Trading Points?

To isolate the effect of our *Songlines* per se, we control for or exclude all origin and destination points (cells) of them from our sample. Mćarthy (1939) refers to these places as trading points, and they might have developed into modern centers of economic activity (cities or towns). Thus, by considering their role, we distinguish between the effects of pre-colonial trading points and trading routes. Therefore, In Table A.10, in column (1) we control for the places having a trading point, while in column (2), we exclude these places. For column (3), we create buffers with a diameter of 50 km (beam 25 km) around the trading sites and exclude all grid cells within these buffers. The results show that the coefficient of *Songlines* remains economically and statistically significant after considering grid cells associated to origins and destinations. We underscore that the presence of trading sites as an additional regressor does not eliminate the significance of *Songlines*, but trading sites themselves are statistically very significant and numerically more important.

6.6.2. Excluding States

Australia, a vast continent, is microgeopraphically very diverse as compared to Europe. To underscore this point, we note that Western Australia is mainly composed of sandy and sparsely populated deserts; New South Wales is heavily urbanised and hosts large cities such as Sydney. In order to test whether there is any particular state that could influence our results, we exclude sequentially each of the states individually in Table A.11. The coefficients associated with our main variable demonstrate that despite microgeographic and environmental heterogeneity across Australian states, *Songlines* maintain its positive and significant effect.

6.6.3. Spatial Autocorrelation

We assess the potential impact of spatial autocorrelation on our results by using the large-cluster approach, as proposed by Bester et al. (2011) and implemented by Kelly (2020). The latter conducts a Monte Carlo simulation and finds that the large-cluster approach with three clusters provides rather conservative and robust estimates for different patterns of spatial correlation. As Table A.12 shows, the spatially corrected coefficient suggests that the results of our study are unlikely to be driven by spatial autocorrelation.

6.6.4. Ommited Variables

In order to address omitted variable bias concerns, we apply the technique proposed by Oster (2019) to address the issue of omitted variable bias, which occurs when some relevant factors (unobserved controls) are left out of the model, potentially distorting the results. Oster's approach improves our estimates by increasing the R^2 value—an indicator of how much of the variation in the data is explained by the model—by 30% compared to our most conservative estimate in column 4 Table 1 ($R_{\max} = 1.3 \times R^2$). Increasing the R^2, Oster's algorithm assumes that we employ all (observable and unobservable) variables as controls, allowing for the calculation of a new β coefficient and a parameter called δ. For the model to be free of omitted variable bias, the interval between our most conservative β (from column 4 Table 1) and the new β coefficient should safely exclude zero, and the δ parameter must be greater than one. The parameter δ reflects how important the unobserved variables would need to be, compared to the observed ones, in order to significantly affect the model. If δ is greater than one, it suggests that the model is unlikely to be influenced by omitted variable bias. As shown in Table A.13, both these conditions are met: the new β coefficient is very close to the baseline estimate, and δ is greater than 6, which indicates that the model is highly unlikely to suffer from omitted variable bias.

Furthermore, we use the spatial first differences (SFD) design proposed by Druckenmiller and Hsiang (2018). This method accounts for the fact that adjacent units of analysis that are spatially closer to each other offer better counterfactuals, as unobservable confounders are not systematically correlated between them. To this end, we compare cell c to cell $c-1$ and assume that the expected potential outcome for each would be the same if they received the same treatment. We then distinguish each pair of adjacent cells in both west-east and north-south directions and regress the spatial first differences as follows:

$$\Delta Luminosity_i = \beta_{fsd} TradeRoutes_i + \varepsilon_i. \tag{3}$$

In equation (3), β_{fsd} captures the effect on luminosity when a random grid cell c is treated while grid cell $c-1$ is not (or vice versa). Omitted by default is a hidden factor Y_{fsd} Cov, a vector of covariates for each observation (grid cell). Column (1) of Table A.14 calculates the differences in the west-east direction and column (2) calculates the differences in the north-south direction. We observe a positive and statistically significant effect of *Songlines* on current economic activity. This result lends confidence that spatially correlated unobserved heterogeneity and omitted variables are not biasing our findings.[28]

6.6.5. Pre-colonial Institutions

We now address concerns related to the impact of pre-colonial Aboriginal institutions along the lines of Michalopoulos and Papaioannou (2013) and Angeles and Elizalde (2017) on long-term economic development. To address this, and given the limitations of our data regarding the societal and political organisation of Aboriginal tribes prior to European contact, we utilise the work of O'Grady et al. (1966), who compiled extensive data on Aboriginal linguistics. Specifically, they developed a map illustrating the boundaries of approximately 450 Aboriginal spoken languages in Australia before European colonisation. While acknowledging the lack of precision of the delineated boundaries, this map likely represents the only available source containing rich data on pre-colonial societal structures in Australia. After matching each grid

[28] Our results remain highly significant when bootstrapped standard errors are used (Roodman et al., 2019)

cell (its centroid) with the corresponding language group, we construct language fixed effects that aim to control for unobserved societal heterogeneity across different Aboriginal groups before colonisation. Column (1) in Table 8 presents the results of our baseline OLS model, which also includes pre-colonial language fixed effects. Interestingly, the coefficient of our main variable of interest remains positive and highly significant as our baseline findings.

6.6.6. Early Colonial Economic Activity

Last, to investigate the long-term effects of early economic activities introduced by colonisers, we follow Bruhn and Gallego, 2012 and utilise the dataset compiled by Elder (2024). This dataset, in addition to detailing the foundation year of each town as described in Section 5, also provides a thorough record of their local histories. With this information, we then identify the principal early colonial economic activity for each inland town. For example, if a town's early development was driven by both gold mining and cattle ranching, we classify gold mining as the most significant early economic activity.

We compiled a dataset of 867 towns, each classified by its principal economic activity during its early years. These activities are grouped into eight categories: (1) mining (e.g., gold, copper), (2) primary sector (e.g., agriculture, cattle, fishing), (3) secondary sector (e.g., sawmills, butter and cheese factories), (4) services (e.g., telegraph stations, banks), (5) transport/transfer hubs (e.g., stopovers for travellers), (6) leisure/tourism (e.g., resorts, tourist attractions), (7) military (e.g., army bases, training camps). For all other places we assume zero early colonial economic activity. Subsequently, each grid cell is matched with its corresponding town and its early colonial economic activity, thus permitting us constructing fixed effects.

Due to the geographic constraints and lack of necessary infrastructure and resources, the initial settlements in Australia were sparse and concentrated, leaving much of the inland regions largely uninhabited and devoid of significant economic activity (Reynolds et al., 2007). To address potential bias arising from the high diversification of early economic activity in major cities (as described in Section 5), we exclude all grid cells that include them. Subsequently, our hand-collected dataset may efficiently control for early colonial activities. The coefficient of our main variable of interest in column (2) in Table 8 remains positive and significant.

Table 8. Early and Pre-Colonial Activity

Dependent Variable: Binary dummy for night light		
	Pre-Colonial Institutions	Early Colonial Activity
	(1)	(2)
Songlines	0.0252***	0.0219***
	(0.0046)	(0.0045)
Local Gov FE	✓	✓
Pre-colonial Language FE	✓	
Early Ec Activity FE		✓
Geo & Hist Controls	✓	✓
N	79731	79599
R^2	0.269	0.233

Notes: The unit of observation is a grid cell of 10km x 10km. OLS estimates are shown with robust standard errors clustered at the local government district level. Songlines is a dummy variable that takes the value of 1 if a grid cell has at least one Songline, and 0 otherwise. In both columns (1) and (2), the dependent variable is a dummy variable that takes the value of 1 if cell i has nightlight, and 0 otherwise. Both columns include controls for local government fixed effects, as well as geographical and historical variables, including coordinates, elevation, and climatic factors. Column (1) incorporates pre-colonial language fixed effects and column (2) incorporates early economic activity fixed effects. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

7. Mechanisms and Path Dependence

Why do *Songlines* matter for contemporary economic activity in Australia? We argue that this association can be partly attributed to path dependence and agglomeration effects resulting from the impact of the *Songlines* on the development of early European transportation infrastructure during colonisation. By having closely followed those routes they operated as a lock-in effect. This fact is well supported by recent anthropological evidence (Kerwin, 2010). Importantly, and in line with the historical findings presented in section 2, when Europeans established settlements in new regions, they constructed transportation networks to connect these areas to larger colonies. This development likely induced the concentration of economic activities by enhancing connectivity over time.

The mechanism described above is consistent with recent studies examining the impact of historical trade routes on contemporary outcomes (Dalgaard et al., 2018; Flückiger et al., 2021; Ahmad and Chicoine, 2021). These studies argue that regions that experienced a higher degree of exposure to these trade networks developed better connectivity patterns due to reduced transport costs, leading to increased levels of goods exchange and subsequent infrastructure investments. As a result, these regions presently exhibit higher economic outcomes than those with a lower legacy of historical trade routes.

To assess th misechanism, we digitised and georeferenced maps of

Table 9. Mechanisms

	Early Railways	Early Highways
	(1)	(2)
Songlines	0.013**	0.021***
	(0.006)	(0.007)
Local Gov FE	✓	✓
Geo & Hist Controls	✓	✓
N	79731	79731
R^2	0.200	0.137

Notes: The unit of observation is a grid cell of 10km X 10km. OLS estimates are shown with robust standard errors clustered at the local government district level. In columns (1), the dependent variable is a dummy variable that takes the value of 1 if cell *i* has at least one railway built between 1880 and 1920, and 0 otherwise. In column (2), the dependent variable is dummy variable that takes the value of 1 if cell *i* has at least one highway built in Australia until the early 1950s, and 0 otherwise. The measure uses only major interstate and state highways. Both measures were constructed by digitising and georeferencing a series of historical maps, with sources provided in the description of variable section in the Appendix. All columns control for local government fixed effects, as well as for a set of geographical, climatic and historical variables, which include: the coordinates of each grid cell, agricultural suitability, elevation, ruggedness (standard deviation of elevation), rainfall and standard deviation of rainfall, temperature and standard deviation of temperature, distance to the sea, distance to Sydney, distance the state capital, distance to historical mines, coastal presence, and water percentage. The descriptions of the geographical, climatic, and historical variables can be found in the Appendix. The constant term was omitted for space. *, ** and *** mean that the coefficient is statistically significant at 10%, 5% and 1% respectively.

early transport infrastructure in Australia. Figure A.3 depicts all early railways built between 1880 and 1920, which was created using a series of historical maps. Similarly, Figure A.4 displays all early highways constructed by Europeans until the early 1950s, with only the major interstate and state highways in Australia being digitised. We construct dummy variables for each map to measure the presence of early transport networks in Australia. Specifically, we measure early railway infrastructure by assigning a dummy variable of 1 to grid cells with at least one railway and 0 otherwise. This methodology is applied in like manner to the other map on early highways.

Our analysis evolves around two strategies. First, we regress each of the dummies just defined on our main variable of interest, with all of our baseline control variables and local government fixed effects being included. Table 9, Columns (1)-(2), show that both early railways and highways were significantly correlated with *Songlines*. The historical evidence we provide is consistent with these results, suggesting that current economic activity is associated with *Songlines*.

Our second strategy seeks to determine whether the effects of early transport networks on contemporary economic activity is due to the influence of adoption of Aboriginal knowledge during colonisation via the *Songlines*, or they were directly affected by geographic or economic factors. Put differently, it could be argued that mid-20th-century infrastructure (as measured in our digitised maps of railways and highways circa 1950) was deliberately designed to facilitate economic growth and associated agglomeration forces, such as the concentration of farmers in fertile, easily traversable areas or the development of infrastructure tied to coal production, rather than being guided by *Songlines*. Thus, early railways or highways could have been developed to link key settlements based on overall optimization rather than such traditional pathways as the *Songlines*. This would imply that connections, whether by rail or road, likely followed routes deemed "economically efficient"—in other words, routes that minimised construction costs by passing through the most traversable areas.

To explore this further, we perform the following analysis. We split up the study area into 50x50 km grid cells, each containing 25 smaller 10x10 km grid cells (our primary units of analysis). Within each 50x50 km grid cell, we identify the 10x10 km grid cells that fall at or above the 75th and 90th percentiles in *Natural Routes* density, as described in Section 4. This approach allows us to highlight the most navigable grid cells at the 50x50 km grid cell level, rather than using the universe of our sample. Therefore, for each 50x50 km area, we assess which smaller grid cells would have been the most likely crossing points if infrastructure planning in the 20th century had been driven by other forces forces, rather than *Songlines*.

For each 10x10 km grid cell, we construct two binary variables to indicate the presence of early railways and early highways, as detailed above. We also create a binary variable to record whether the grid cell falls within the high-density *Natural Route* category, based on the 75th or 90th percentiles. Next, we conduct separate regressions for the binary variables representing railways and highways, each performed twice: one using the high-density *Natural Route* indicator at the 75th percentile and the other using the indicator at the 90th percentile, along with the *Songlines* variable and baseline control variables. The results are presented in Table 10. They reveal a positive and statistically significant effect of *Songlines* on the presence of railways (columns 1 and 2), while the high-density *Natural Route* variables at both the 75th and 90th percentiles do not show significant coefficients. For highways (columns 3 and 4), both *Songlines* and high-density *Natural Routes* are significant predictors. However, the impact of *Songlines* is nearly three times larger than that of high-density *Natural Routes*.

These findings highlight the considerable impact of *Songlines* on the development of early infrastructure during the 20th century, demonstrating their stronger gravitational pull compared to *Natural Routes*.

Table 10. Impact of *Songlines* and *Natural Routes* on Early Infrastructure

	Early Railways		Early Highways	
	(1)	(2)	(3)	(4)
Songlines	0.0040**	0.0040**	0.0066***	0.0065***
	(0.0017)	(0.0017)	(0.0020)	(0.0020)
High-Density Natural Routes (90th)	−0.0004		0.0027***	
	(0.0008)		(0.0010)	
High-Density Natural Routes (75th)		−0.0007		0.0024**
		(0.0010)		(0.0011)
Local Gov FE	✓	✓	✓	✓
Geo & Hist Controls	✓	✓	✓	✓
R^2	0.205	0.205	0.142	0.142
N	79731	79731	79731	79731

Notes: The unit of observation is a grid cell of 10km x 10km. OLS estimates are shown with robust standard errors clustered at the local government district level. The dependent variable in columns (1) and (2) indicates the presence of early railways, while in columns (3) and (4), it indicates early highways. Songlines is a dummy variable equal to 1 if a grid cell contains at least one Songline. High-density Natural Routes are binary variables taking the value 1 if a grid cell falls within the top 75th percentile (columns 2 and 4) or 90th percentile (columns 1 and 3) of Natural Routes within 50x50 km² areas containing 25 grid cells of 10km x 10km. All columns include local government fixed effects and control for geographical, climatic, and historical variables, including coordinates, agricultural suitability, elevation, ruggedness, rainfall, temperature, proximity to the sea, Sydney, the state capital, historical mines, and water percentage. Detailed variable descriptions are in the Appendix. The constant term was omitted for space. *, **, and *** indicate significance at the 10

8. Conclusion

This paper examines the long-term economic impact of the adoption of Indigenous knowledge by Europeans during colonisation. We focus on a unique setting, a "natural" experiment in colonial Australia. We argue that Europeans relied on Aboriginal knowledge of the landscape for their exploration and settlement by means of oral maps, known as *Songlines*. We construct a novel georeferenced dataset of *Songlines* and demonstrate that areas located along these traditional routes show higher levels of economic activity today.

To address potential confounding factors and establish a causal relationship, we employed several strategies. We compared the predictive power of *Songlines* against environmentally optimal travel (*Natural Routes*) and found that *Songlines* were more strongly associated with both historical European exploration paths and current economic activity. We also utilise a number of additional diagnostics, such as contiguous pair analyses and many others, and control for a range of geographic and historical variables to ensure that our results are not driven by unobserved regional characteristics.

The long-term effects of *Songlines* appear to stem from path dependence and agglomeration effects facilitated by transport infrastructure built by Europeans along these routes during the colonial period. Early settlements tended to emerge closer to *Songlines* and economic activity became increasingly concentrated in these areas over time.

These results contribute to a deeper understanding of how Indigenous knowledge systems can impact long-term economic outcomes. They highlight the importance of considering second-nature geography—human activity and cultural practices—in addition to natural geographic features when analysing economic development over the long-run. By demonstrating the lasting economic significance of *Songlines*, our study highlights the role of Indigenous peoples not just as passive participants but as active contributors to the shaping of modern economies.

References

Acemoglu, Daron, Simon Johnson, and James A Robinson (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review* 91(5), pp. 1369–1401.

— (2002). Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *The Quarterly Journal of Economics* 117(4), pp. 1231–1294.

Ahmad, Zofia and Luke Chicoine (2021). Silk Roads to Riches: Persistence Along an Ancient Trade Network. *Available at SSRN 3760490*.

Allen, Treb and Costas Arkolakis (2014). Trade and the Topography of the Spatial Economy. *The Quarterly Journal of Economics* 129(3), pp. 1085–1140.

Angeles, Luis and Aldo Elizalde (2017). Pre-colonial institutions and socioeconomic development: The case of Latin America. *Journal of Development Economics* 124, pp. 22–40.

Bakker, Jan David, Stephan Maurer, Jörn-Steffen Pischke, and Ferdinand Rauch (2021). Of mice and merchants: connectedness and the location of economic activity in the Iron Age. *Review of Economics and Statistics* 103(4), pp. 652–665.

Barjamovic, Gojko, Thomas Chaney, Kerem Coşar, and Ali Hortaçsu (2019). Trade, merchants, and the lost cities of the bronze age. *The Quarterly Journal of Economics* 134(3), pp. 1455–1503.

Barsanetti, Bruno (2021). Cities on pre-Columbian paths. *Journal of Urban Economics* 122, p. 103317.

Bathurst Local Aboriginal Consultative Committee (2011). The Wiradjuri people. Bathurst, N.S.W.: Bathurst Local Aboriginal Consultative Committee.

Bester, C Alan, Timothy G Conley, and Christian B Hansen (2011). Inference with dependent data using cluster covariance estimators. *Journal of Econometrics* 165(2), pp. 137–151.

Bloch, Francis, Matthew O Jackson, and Pietro Tebaldi (2023). Centrality measures in networks. *Social Choice and Welfare* 61, pp. 413–453.

Blyton, Greg (2015). Harry Brown (c. 1819-1854): Contribution of an Aboriginal guide in Australian exploration. *Aboriginal History* 39, pp. 63–82.

Bruhn, Miriam and Francisco A Gallego (2012). Good, bad, and ugly colonial activities: do they matter for economic development? *Review of Economics and Statistics* 94(2), pp. 433–461.

Campbell, Keith (2005). Broughton (c.1798–c.1850). *Australian Dictionary of Biography*. URL: https://adb.anu.edu.au/biography/broughton-12820.

Caruana-Galizia, P, T Hashino, and MS Schulze (2021). Underlying sources of growth: first and second nature geography. *Cambridge Economic History of the Modern World* 1, pp. 339–68.

Chatwin, Bruce (1987). The Songlines. Franklin Press.

Chen, Yuxian, Yannis M Ioannides, and Ferdinand Rauch (2022). Asymmetric Trading Costs and Ancient Greek Cities. *CEPR Discussion Paper No. DP17204*.

Clarke, Philip A (2008). Aboriginal plant collectors. Rosenberg Pub.

Cleary, Tania (1993). Poignant Regalia: 19th Century Aboriginal Breastplates & Images: A Catalogue of Aboriginal Breastplates Held in Public, Regional and Private Collections in New South Wales, Victoria, Tasmania, Queensland, South Australia, Western Australia and the Australian Capital Territory. Exhibition venues: Greenway Gallery, Hyde Park Barracks, 26 May-4 July 1993. Glebe, NSW: Historic Houses Trust of New South Wales.

Dalgaard, Carl-Johan, Nicolai Kaarsen, Ola Olsson, and Pablo Selaya (2018). Roman roads to prosperity: Persistence and non-persistence of public goods provision.

Dell, Melissa (2010). The persistent effects of Peru's mining mita. *Econometrica* 78(6), pp. 1863–1903.

Druckenmiller, Hannah and Solomon Hsiang (2018). Accounting for unobservable heterogeneity in cross section using spatial first differences. *techreport*. National Bureau of Economic Research.

Dube, Arindrajit, T William Lester, and Michael Reich (2010). Minimum wage effects across state borders: Estimates using contiguous counties. *The Review of Economics and Statistics* 92(4), pp. 945–964.

Duranton, Gilles and Diego Puga (2014). The growth of cities. *Handbook of Economic Growth* 2, pp. 781–853.

Easterly, William and Ross Levine (2016). The European origins of economic development. *Journal of Economic Growth* 21, pp. 225–257.

Elder, Bruce (2024). *Aussie Towns: A Definitive Travel Guide to Australian Towns and Destinations*. Accessed: 2024-09-10. URL: https://www.aussietowns.com.au/.

Elvidge, Christopher D, Kimberly Baugh, Mikhail Zhizhin, Feng Chi Hsu, and Tilottama Ghosh (2017). VIIRS night-time lights. *International Journal of Remote Sensing* 38(21), pp. 5860–5879.

European Commission (2023). *GHSL Data Package 2023*. URL: https://ghsl.jrc.ec.europa.eu/documents/GHSL_Data_Package_2023.pdf?t=1683540422 (visited on 09/06/2023).

Fischer, Guenther, Freddy Nachtergaele, Sylvia Prieler, HT Van Velthuizen, Luc Verelst, and David Wiberg (2008). Global agro-ecological zones assessment for agriculture (GAEZ 2008). *IIASA, Laxenburg, Austria and FAO, Rome, Italy* 10.

Flood, Josephine (1983). Archaeology of the Dreamtime: The Story of Prehistoric Australia and its People. Collins Publishers Australia.

Flückiger, Matthias, Erik Hornung, Mario Larch, Markus Ludwig, and Allard Mees (2021). Roman transport network connectivity and economic integration. *The Review of Economic Studies*.

Forrest, John (1875). Explorations in Australia. London: Sampson Low, Marston, Low, & Searle.

Geospatial Data (n.d.). Australian Government Department of Climate Change and Energy Efficiency. URL: https://fed.dcceew.gov.au/pages/terms-of-use (visited on 08/28/2023).

Giles, Ernest (1889). Australia Twice Traversed: The Romance of Exploration, Being a Narrative Compiled from the Journals of Five Exploring Expeditions into and through Central South Australia and Western Australia, from 1872 to 1876. Vol. 1. London: Sampson Low, Marston, Searle, & Rivington. URL: https://books.google.gr/books/about/Australia_Twice_Traversed.html?id=gws0AAAAYAAJ&redir_esc=y.

Glaeser, Edward L, Hedi D Kallal, Jose A Scheinkman, and Andrei Shleifer (1992). Growth in cities. *Journal of Political Economy* 100(6), pp. 1126–1152.

Glaeser, Edward L, Rafael La Porta, Florencio Lopez-de Silanes, and Andrei Shleifer (2004). Do institutions cause growth? *Journal of Economic Growth* 9(3), pp. 271–303.

Herald, The Sydney Morning (1912 [Online]). *Early history. The discovery of Bathurst*. URL: https://trove.nla.gov.au/newspaper/article/15374708.

Hijmans, Robert J, Susan E Cameron, Juan L Parra, Peter G Jones, and Andy Jarvis (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology: A Journal of the Royal Meteorological Society* 25(15), pp. 1965–1978.

Hume, S. H. (1966). Hamilton Hume (1797–1873). *Australian Dictionary of Biography*. URL: https://adb.anu.edu.au/biography/hume-hamilton-2211.

Kampanelis, Sotiris (2019). It's time for Westernization: the advantages of the early start for long-term economic development at the local level. *Oxford Economic Papers* 71(4), pp. 996–1025.

Kelly, Morgan (2020). Direct standard errors for regressions with spatially autocorrelated residuals.

Kerwin, Dale (2010). Aboriginal Dreaming Paths and Trading Routes: The colonisation of the Australian economic landscape. Sussex Academic Press.

Koyama, Mark and Jared Rubin (2022). How the world became rich: The historical origins of economic growth. John Wiley & Sons.

Leichhardt, Ludwig (1847). Journal of an Overland Expedition in Australia, From Moreton Bay to Port Essington, A Distance of Upwards of 3000 Miles, During the Years 1844-1845. London: T. & W. Boone.

Lord, Albert B (1960). The Singer of Tales. Cambridge (Mass).

Mann, Charles C (2005). 1491 The Americas before Columbus. Granta Books, London.

McBryde, Isabel (2000). Travellers in storied landscapes: a case study in exchanges and heritage. *Aboriginal History* 24, pp. 152–174.

MćCarthy, Frederick David (1939). "Trade in Aboriginal Australia, and "Trade" Relationships with Torres Strait, New Guinea and Malaya". *Oceania* 9(4), pp. 405–438.

McDonald, Jo and Lucia Clayton (2016). Rock art thematic study. *Report to the Department of the Environment and the Australian Heritage Council*.

McLaren, Glen (1996). Beyond Leichhardt: Bushcraft and the exploration of Australia. *(No Title)*.

Michalopoulos, Stelios and Elias Papaioannou (2013). Pre-colonial ethnic institutions and contemporary African development. *Econometrica* 81(1), pp. 113–152.

Mountford, Charles Pearcy (1976). Nomads of the Australian desert. Adelaide: Rigby.
Murdock, George Peter (1967). Ethnographic atlas: a summary. *Ethnology* 6(2), pp. 109–236.
Norris, Ray P and Bill Yidumduma Harney (2014). Songlines and navigation in Wardaman and other Australian Aboriginal cultures. *Journal of Astronomical History and Heritage* 17.
Nunn, Nathan (2008). The long-term effects of Africa's slave trades. *The Quarterly Journal of Economics* 123(1), pp. 139–176.
O'Grady, Geoffrey N, Carl F Voegelin, and Florence M Voegelin (1966). Languages of the world: Indo-Pacific fascicle six. *Anthropological Linguistics* 8(2), pp. 1–197.
Oster, Emily (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics* 37(2), pp. 187–204.
Özak, Ömer (2018). Distance to the pre-industrial technological frontier and economic development. *Journal of Economic Growth* 23(2), pp. 175–221.
Reynolds, Henry (1980). The land, the explorers and the Aborigines. *Australian Historical Studies* 19(75), pp. 213–226.
— (1990). *With the white people: the crucial role of Aborigines in the exploration and development of Australia.*
Reynolds, James F, D Mark Stafford Smith, Eric F Lambin, BL Turner, Michael Mortimore, Simon PJ Batterbury, Thomas E Downing, Hadi Dowlatabadi, Roberto J Fernández, Jeffrey E Herrick, et al. (2007). Global desertification: building a science for dryland development. *science* 316(5826), pp. 847–851.
Robinson, H.E.C. (1927). *The discovery and exploration of Australia, 1519 to 1901 by sea and land.*
Roodman, David, Morten Ørregaard Nielsen, James G MacKinnon, and Matthew D Webb (2019). Fast and wild: Bootstrap inference in Stata using boottest. *The Stata Journal* 19(1), pp. 4–60.
Ryan, Lyndall et al. (2017). *Colonial Frontier Massacres in Australia 1788-1930*. Newcastle. URL: http://hdl.handle.net/1959.13/1340762.
Sokoloff, Kenneth L and Stanley L Engerman (2000). Institutions, factor endowments, and paths of development in the new world. *Journal of Economic Perspectives* 14(3), pp. 217–232.
Stuart, John McDouall (1865). Explorations in Australia: The Journals of John McDouall Stuart during the Years 1858, 1859, 1860, 1861, 1862, When He Fixed the Centre of the Continent and Successfully Crossed It from Sea to Sea. Ed. by William Hardman. 2nd. London: Saunders, Otley, and Co.
Troy, Jakelin (1993). King Plates: A History of Aboriginal Gorgets. Aboriginal Studies Press.

Valencia Caicedo, Felipe (2019). The mission: Human capital transmission, economic persistence, and culture in South America. *The Quarterly Journal of Economics* 134(1), pp. 507–556.

Veth, Peter and Sue O'Connor (2013). The past 50,000 years: an archaeological view. In: *The Cambridge History of Australia Volume 1: Indigenous and Colonial Australia*. Cambridge University Press, pp. 17–42.

Warburton, Peter E. (1875). Journey Across the Western Interior of Australia: With an Introduction and Additions. London: Low.

Wositsky, J and Bill Y Harney (1999). Born under the paperbark tree. Marlston, SA, JB Books.

Appendices

A. Tables and Figures

Table A.1. Summary Statistics

	N	Mean	St.Dev.	Min	Max
Songlines	79731	0.09	0.29	0	1
Night Light	79731	0.06	0.23	0	1
Population Density	79731	-12.95	7.03	-17.75	7.57
Exploration Routes	79731	0.15	0.36	0	1
Early Railways	79731	0.05	0.23	0	1
Early Highways	79731	0.08	0.27	0	1
All Cities	79731	0.00	0.04	0	1
Cities est. before 1850	79731	0.00	0.03	0	1
Cities est. in 1850-1900	79731	0.00	0.03	0	1
Cities est. in 1900-1950	79731	0.00	0.01	0	1
Cities est. in 1950-2000	79731	0.00	0.01	0	1
Latitude	79731	708.86	67.08	568.01	877.01
Longitude	79731	-87.57	106.24	-305.45	113.55
Agriculture Suitability	79731	1.06	0.57	0	7
Elevation	79731	281.17	183.93	-15.36	1814.49
Ruggedness	79731	16.41	23.30	0.15	395.88
Precipitation	79731	448.12	308.74	128.97	3751.36
StDev Precipitation	79731	6.52	15.53	0	646.34
Temperature	79731	217.48	37.94	49.59	293.02
StDev Temperature	79731	1.09	1.46	0	23.45
Coastal Dummy	79731	0.02	0.14	0	1
Distance to the Sea	79731	35.58	24.33	0.50	93.89
Distance to Sydney	79731	208.39	93.69	1.82	386.52
Distance to State Capital	79731	94.97	51.00	0.36	249.82
Distance to Historical Mine	79731	17.44	12.31	0.03	65.83
Water Percentage	79731	0.02	0.08	0	1

Table A.2. *Songlines* from McCarthy's digitised maps

	Dependent Variable: Binary variable for night lights
	(1)
Songlines	0.0139***
	(0.0034)
Local Gov. FE	✓
Geo & Hist Controls	✓
N	79731
R^2	0.236

Notes: The unit of observation is a grid cell of 10km X 10km. The dependent variable is a dummy variable that takes the value of 1 if cell *i* has nightlight, and 0 otherwise. This variable uses data from Visible Infrared Imaging Radiometer Suite (VIIRS) sensor. All columns control for local government fixed effects, as well as for a set of geographical, climatic and historical variables, which include: the coordinates of each grid cell, agricultural suitability, elevation, ruggedness (standard deviation of elevation), rainfall and standard deviation of rainfall, temperature and standard deviation of temperature, distance to the sea, distance to Sydney, distance the state capital, distance to historical mines, coastal presence, and water percentage. The descriptions of the geographical, climatic, and historical variables can be found in the Appendix. Robust standard errors clustered at the local government district level were used and the constant term was omitted for space. *, ** and *** mean that the coefficient is statistically significant at 10%, 5% and 1% respectively.

Table A.3. Adoption "Frictions": Sites of Massacres Excluded

	Dependent Variable: Dummy variable for night light			
	(1) excl. grid cells within 5km of a massacre & with Songlines	(2) excl. grid cells within 10km of a massacre & with Songlines	(3) excl. grid cells within 15km of a massacre & with Songlines	(4) excl. grid cells within 20km of a massacre & with Songlines
Songlines	0.029***	0.028***	0.026***	0.025***
	(0.005)	(0.005)	(0.005)	(0.005)
N	79703	79621	79506	79365
R^2	0.232	0.231	0.230	0.228

*Notes:*The unit of observation is a grid cell of 10km X 10km. The dependent variable is a dummy variable that takes the value of 1 if cell *i* has nightlight,and 0 otherwise. This variable uses data from the Visible Infrared Imaging Radiometer Suite (VIIRS) sensor. The analysis excludes grid cells that are 5 km (column (1)), 10km (column (2), 15km (column (3), and 20km (column (4)) closer to a massacre and have a Songline. All columns control for local government fixed effects, as well as for a set of geographical, climatic and historical variables, which include: the coordinates of each grid cell, agricultural suitability, elevation, ruggedness (standard deviation of elevation), rainfall and standard deviation of rainfall, temperature and standard deviation of temperature, distance to the sea, distance to Sydney, distance the state capital, distance to historical mines, coastal presence, and water percentage. The descriptions of the geographical, climatic, and historical variables can be found in the Appendix. The constant term was omitted for space. *, ** and *** mean that the coefficient is statistically significant at 10%, 5% and 1% respectively.

Table A.4. Homogeneous sample

	Dependent Variable: Binary variable for night light	
	(1)	(2)
	≥25% Precipitation	≤25% Temperature
Songlines	0.041***	0.025***
	(0.006)	(0.006)
Local Gov FE	✓	✓
Geo & Hist Controls	✓	✓
N	59457	59678
R^2	0.238	0.253

Notes: The unit of observation is a grid cell of 10km X 10km. The dependent variable is a dummy variable that takes the value of 1 if cell *i* has nightlight, and 0 otherwise. This variable uses data from Visible Infrared Imaging Radiometer Suite (VIIRS) sensor. In column (1), cells with the bottom 25% rainfall and temperature were excluded. In column (2), cells with the top 25% rainfall and temperature were excluded. All columns control for local government fixed effects, as well as for a set of geographical, climatic and historical variables, which include: the coordinates of each grid cell, agricultural suitability, elevation, ruggedness (standard deviation of elevation), rainfall and standard deviation of rainfall, temperature and standard deviation of temperature, distance to the sea, distance to Sydney, distance the state capital, distance to historical mines, coastal presence, and water percentage. The descriptions of the geographical, climatic, and historical variables can be found in the Appendix. Robust standard errors clustered at the local government district level were used and the constant term was omitted for space. *, ** and *** mean that the coefficient is statistically significant at 10%, 5% and 1% respectively.

Figure A.1. *Songlines* in Australia: McCarthy's maps

Notes: This map shows *Songlines* in Australia, based on the work of McCarthy (1939). Red lines indicate trade routes. Authors' own digitalisation using McCarthy (1939)

Table A.5. 50 X 50 km cell analysis

	Dependent Variable: Binary variable for night light			
	(1)	(2)	(3)	(4)
Songlines	0.051**	0.056**	0.055**	0.048**
	(0.023)	(0.023)	(0.024)	(0.023)
Latitude		−0.022***	−0.024	−0.051
		(0.008)	(0.027)	(0.046)
Longitude		−0.011	−0.011	0.089**
		(0.015)	(0.015)	(0.045)
Agriculture Suitability			−0.010	−0.074*
			(0.026)	(0.044)
Elevation			0.000	−0.000
			(0.000)	(0.000)
Ruggedness			−0.000	−0.001
			(0.001)	(0.001)
Precipitation			0.000	−0.000*
			(0.000)	(0.000)
StDev Precipitation			0.001**	0.001***
			(0.000)	(0.000)
Temperature			−0.000	−0.005
			(0.004)	(0.003)
StDev Temperature			0.014	0.019
			(0.022)	(0.022)
Coastal Dummy				0.117**
				(0.046)
Distance to the Sea				0.000
				(0.001)
Distance to Sydney				0.010*
				(0.005)
Distance to State Capital				0.002**
				(0.001)
Distance to Historical Mine				−0.012***
				(0.001)
Water Percentage				0.005
				(0.004)
Local Gov FE	✓	✓	✓	✓
N	3444	3444	3444	3444
R^2	0.286	0.289	0.294	0.334

Notes: The unit of observation is a grid cell of 50km X 50km. The dependent variable is a dummy variable that takes the value of 1 if cell *i* has nightlight, and 0 otherwise. This variable uses data from Visible Infrared Imaging Radiometer Suite (VIIRS) sensor. All columns control for local government fixed effects. Columns (1)-(4) include a measure of *Songlines*. This measure is is a dummy variable that takes the value of 1 if cell *i* has at least one *Songline*, and 0 otherwise. This variable was constructed using anthropological data from Mc̀Carthy (1939), which describes the trade routes created by Aboriginal people based on oral traditions prior to colonisation. The Human Mobility Index (HMI) from Özak (2018) was used to identify optimal routes between origins and destinations, with a least-cost algorithm. From columns (2) to (4), geographical and climatic control variables were added gradually and incrementally. The descriptions of the geographical and climatic variables can be found in the Appendix. Robust standard errors clustered at the local government district level were used and the constant term was omitted for space. *, ** and *** mean that the coefficient is statistically significant at 10%, 5% and 1% respectively.

Table A.6. Neighbouring analysis

	Dependent Variable: Binary variable for night lights
	(1)
Songlines	0.030***
	(0.006)
Local Gov. FE	✓
Geo & Hist Controls	✓
N	18188
R^2	0.233

Notes: The unit of observation is a grid cell of 10km X 10km. The dependent variable is a dummy variable that takes the value of 1 if cell i has nightlight, and 0 otherwise. This variable uses data from Visible Infrared Imaging Radiometer Suite (VIIRS) sensor. All columns control for local government fixed effects, as well as for a set of geographical, climatic and historical variables, which include: the coordinates of each grid cell, agricultural suitability, elevation, ruggedness (standard deviation of elevation), rainfall and standard deviation of rainfall, temperature and standard deviation of temperature, distance to the sea, distance to Sydney, distance the state capital, distance to historical mines, coastal presence, and water percentage. The descriptions of the geographical, climatic, and historical variables can be found in the Appendix. Regression excludes from the sample cells with Aboriginal trade routes and all their tangents that may or may not host a *Songline*. Robust standard errors clustered at the local government district level were used and the constant term was omitted for space. *, ** and *** mean that the coefficient is statistically significant at 10%, 5% and 1% respectively.

Figure A.2. European exploration routes by Robinson (1927)

Notes: This map shows the European explorations routes as drawn by Robinson (1927).

Table A.7. Alternative sources

	Dependent Variable: Binary variable for night light					
	(1) HMI & RA	(2) DIG & RA	(3) DIG	(4) HMI=HMI & RA	(5) HMI=DIG & RA	(6) HMI=DIG
Songlines	0.012^{***}	0.012^{***}	0.011^{***}	0.049^{***}	0.049^{***}	0.051^{***}
	(0.002)	(0.003)	(0.003)	(0.008)	(0.008)	(0.008)
Local Gov FE	✓	✓	✓	✓	✓	✓
Geo & Hist Controls	✓	✓	✓	✓	✓	✓
N	79731	79731	79731	59266	62922	62323
R^2	0.232	0.232	0.232	0.223	0.218	0.217

Notes: The unit of observation is a grid cell of 10km X 10km. The dependent variable is a dummy variable that takes the value of 1 if cell *i* has nightlight,and 0 otherwise. This variable uses data from the Visible Infrared Imaging Radiometer Suite (VIIRS) sensor. Columns (1)-(3), Trade Routes is a measure of *Songlines*. In columns (1)-(3), the Aboriginal routes were created based on a variety of sources. In column (1), the variable uses a weighted least-cost algorithm based on the Human Mobility Index (HMI) from Özak (2018) and data on the location of Aboriginal rock art. In column (2), a weighted least-cost algorithm based on historical maps from McCarthy (1939) and the location of Aboriginal petroglyphs is used. In column (3), the variable uses a weighted least-cost algorithm based only on the historical maps of McCarthy (1939). Columns (4)-(6) use only cells for the OLS regressions where the measurement of *Songlines* based on HMI alone 'agrees' with the existence of a route. Column (4) uses a dummy variable that takes the value 1 if cell *i* hosts a *Songline* based on the HMI and the location of Aboriginal rock art. Column (5) uses a dummy variable that takes the value 1 if cell *i* hosts a *Songline* based on the historical maps of McCarthy (1939) and the location of Aboriginal rock art. Column (6) uses a dummy variable that takes the value 1 if cell *i* hosts a *Songline* based on the historical maps of McCarthy (1939). All columns control for local government fixed effects, as well as for a set of geographical, climatic and historical variables, which include: the coordinates of each grid cell, agricultural suitability, elevation, ruggedness (standard deviation of elevation), rainfall and standard deviation of rainfall, temperature and standard deviation of temperature, distance to the sea, distance to Sydney, distance the state capital, distance to historical mines, coastal presence, and water percentage. The descriptions of the geographical, climatic, and historical variables can be found in the Appendix. Robust standard errors clustered at the local government district level were used and the constant term was omitted for space. *, ** and *** mean that the coefficient is statistically significant at 10%, 5% and 1% respectively.

Table A.8. *Songlines*, Night Lights & Population Density

Panel A:	Dependent Variable: Binary dummy for night light					
	(1) all cells	(2) excl. 50km	(3) excl. 100km	(4) excl. 200km	(5) excl. 300km	(6) excl. 500km
Songlines	0.030***	0.024***	0.021***	0.018***	0.015***	0.007**
	(0.006)	(0.005)	(0.005)	(0.005)	(0.004)	(0.003)
Local Gov FE	✓	✓	✓	✓	✓	✓
Geo & Hist Controls	✓	✓	✓	✓	✓	✓
N	79731	72474	65382	52803	41838	23744
R^2	0.233	0.165	0.149	0.111	0.060	0.031
Panel B:	**Dependent Variable: Population Density**					
	(1) all cells	(2) excl. 50km	(3) excl. 100km	(4) excl. 200km	(5) excl. 300km	(6) excl. 500km
Trade Routes	0.975***	0.851***	0.758***	0.684***	0.582***	0.662***
	(0.148)	(0.147)	(0.129)	(0.133)	(0.108)	(0.133)
Local Gov FE	✓	✓	✓	✓	✓	✓
Geo & Hist Controls	✓	✓	✓	✓	✓	✓
N	79731	72474	65382	52803	41838	23744
R^2	0.600	0.582	0.570	0.528	0.476	0.281

Notes: The unit of observation is a grid cell of 10km X 10km. In Panel A, the dependent variable is a dummy variable that takes the value of 1 if cell *i* has nightlight,and 0 otherwise. This variable uses data from the Visible Infrared Imaging Radiometer Suite (VIIRS) sensor. In Panel B, the dependent variable is the logarithm of human population density, which is the number of individuals per grid cell. The data is derived from national censuses and population registers for the year 2015, with an additional minor adjustment of 1e-8 included. *Songlines* is a dummy variable that takes the value of 1 if cell *i* has at least one Aboriginal trade route, and 0 otherwise. In column (1), all grid cells in the sample are used. From columns (2)-(6), cells are progressively excluded that are 50km (col. 2), 100km (col 3), 200km (col 4), 300km (col 5) and 500km (col 6) close to the shoreline. All columns control for local government fixed effects, as well as for a set of geographical, climatic and historical variables, which include: the coordinates of each grid cell, agricultural suitability, elevation, ruggedness (standard deviation of elevation), rainfall and standard deviation of rainfall, temperature and standard deviation of temperature, distance to the sea, distance to Sydney, distance the state capital, distance to historical mines, coastal presence, and water percentage. The descriptions of the geographical, climatic, and historical variables can be found in the Appendix. Robust standard errors clustered at the local government district level were used and the constant term was omitted for space. *, ** and *** mean that the coefficient is statistically significant at 10%, 5% and 1% respectively.

Table A.9. Network analysis

	Dependent Variable: Binary variable for night light					
	(1)	(2)	(3)	(4)	(5)	(6)
Songlines	0.027***	0.029***	0.030***	0.030***	0.030***	0.023***
	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)
Centrality: Mean Degree	1.023***					
	(0.204)					
Centrality: Mean Betweenness		1.044***				
		(0.197)				
Centrality: Mean Eigenvector			1.423***			
			(0.270)			
Centrality: K-B(b=1)				1.423***		
				(0.270)		
Centrality: K-B(b=0)					1.423***	
					(0.270)	
Centrality: K-B(b=0.5)						0.180***
						(0.022)
Local Gov FE	✓	✓	✓	✓	✓	✓
Geo & Hist Controls	✓	✓	✓	✓	✓	✓
N	79731	79731	79731	79731	79731	79731
R^2	0.234	0.234	0.233	0.233	0.233	0.235

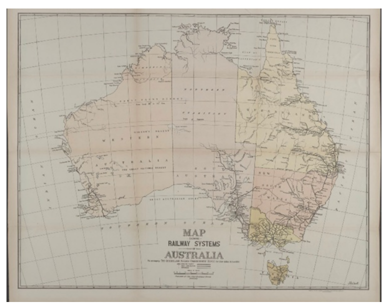
Notes: The unit of observation is a grid cell of 10km X 10km. The dependent variable is a dummy variable that takes the value of 1 if cell *i* has nightlight,and 0 otherwise. This variable uses data from the Visible Infrared Imaging Radiometer Suite (VIIRS) sensor. All columns control for local government fixed effects, as well as for a set of geographical, climatic and historical variables, which include: the coordinates of each grid cell, agricultural suitability, elevation, ruggedness (standard deviation of elevation), rainfall and standard deviation of rainfall, temperature and standard deviation of temperature, distance to the sea, distance to Sydney, distance the state capital, distance to historical mines, coastal presence, and water percentage. Degree, betweenness, eigenvector and Katz-Bonacich (K-B) centrality measures are straightforward applications to weighted networks of the respective concepts (Bloch *et al.* (2023)]. The descriptions of the geographical, climatic, and historical variables can be found in the Appendix. Robust standard errors clustered at the local government district level were used and the constant term was omitted for space. *, ** and *** mean that the coefficient is statistically significant at 10%, 5% and 1% respectively.

Table A.10. Considering starting & destination points

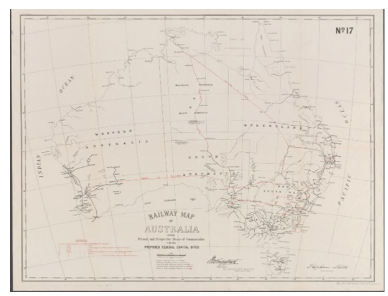
	Dependent Variable: Binary variable for night light		
	(1)	(2)	(3)
Songlines	0.0067***	0.0236***	0.0160***
	(0.0014)	(0.0050)	(0.0053)
Trading Sites	0.0111***	—	—
	(0.0013)	—	—
Local Gov FE	✓	✓	✓
Geo & Hist Controls	✓	✓	✓
N	79731	79425	72495
R^2	0.239	0.234	0.208

Notes: The unit of observation is a grid cell of 10km X 10km. The dependent variable is a dummy variable that takes the value of 1 if cell *i* has nightlight, and 0 otherwise. This variable uses data from the Visible Infrared Imaging Radiometer Suite (VIIRS) sensor. In column (1), we control for the places having either a starting or a destination point. In column (2), cells associated with starting and destination points were excluded. In column (3), buffers of 50km around starting and destination points were created, and all cells within these buffers were excluded. All columns control for local government fixed effects, as well as for a set of geographical, climatic, and historical variables, which include: the coordinates of each grid cell, agricultural suitability, elevation, ruggedness (standard deviation of elevation), rainfall and standard deviation of rainfall, temperature and standard deviation of temperature, distance to the sea, distance to Sydney, distance to the state capital, distance to historical mines, coastal presence, and water percentage. Descriptions of these variables can be found in the Appendix. Robust standard errors clustered at the local government district level were used, and the constant term was omitted for space. *, **, and *** mean that the coefficient is statistically significant at 10%, 5%, and 1%, respectively.

(a) 1880



(b) 1890



(c) 1900

(d) 1920

Figure A.3. Maps of the railway network in Australia between 1880-1920.

Notes: This figure shows the development of railway infrastructure in Australia between 1880 and 1920.

Table A.11. Excluding Australian States

	Dependent Variable: Binary variable for night light					
	(1) Western Australia	(2) Northern Territory	(3) South Australia	(4) Queensland	(5) Victoria	(6) New South Wales
Songlines	0.030***	0.031***	0.031***	0.030***	0.030***	0.030***
	(0.006)	(0.006)	(0.005)	(0.007)	(0.005)	(0.006)
Local Gov FE	✓	✓	✓	✓	✓	✓
Geo & Hist Controls	✓	✓	✓	✓	✓	✓
N	50776	65959	69826	62758	77547	71789
R^2	0.260	0.243	0.229	0.241	0.202	0.222

Notes: The unit of observation is a grid cell of 10km X 10km. The dependent variable is a dummy variable that takes the value of 1 if cell *i* has nightlight,and 0 otherwise. This variable uses data from the Visible Infrared Imaging Radiometer Suite (VIIRS) sensor. In column (1), cells associated with the state of Western Australia were excluded. In column (2), cells associated with the state of Northern Territory were omitted. In column (3), cells linked to the state of South Australia were dropped. In column (4), cells linked to the state of Queensland were excluded. In column (5), cells corresponding to the state of Victoria were dropped. In column (6), cells corresponding to the state of New South Wales were omitted. All columns control for local government fixed effects, as well as for a set of geographical, climatic and historical variables, which include: the coordinates of each grid cell, agricultural suitability, elevation, ruggedness (standard deviation of elevation), rainfall and standard deviation of rainfall, temperature and standard deviation of temperature, distance to the sea, distance to Sydney, distance the state capital, distance to historical mines, coastal presence, and water percentage. The descriptions of the geographical, climatic, and historical variables can be found in the Appendix. Robust standard errors clustered at the local government district level were used and the constant term was omitted for space. *, ** and *** mean that the coefficient is statistically significant at 10%, 5% and 1% respectively.

Table A.12. Spatial Autocorrelation

	Dependent Variable: Binary variable for night light
	(1)
Songlines	0.030**
	(0.004)
Local Gov FE	✓
Geo & Hist Controls	✓
N	79731
R^2	0.233

Notes: The unit of observation is a grid cell of 10km X 10km. The dependent variable is a dummy variable that takes the value of 1 if cell i has nightlight,and 0 otherwise. This variable uses data from the Visible Infrared Imaging Radiometer Suite (VIIRS) sensor. Column (1) controls for local government fixed effects, as well as for a set of geographical, climatic and historical variables, which include: the coordinates of each grid cell, agricultural suitability, elevation, ruggedness (standard deviation of elevation), rainfall and standard deviation of rainfall, temperature and standard deviation of temperature, distance to the sea, distance to Sydney, distance the state capital, distance to historical mines, coastal presence, and water percentage. The descriptions of the geographical, climatic, and historical variables can be found in the Appendix. Robust standard errors clustered at the local government district level were used and the constant term was omitted for space. *, ** and *** mean that the coefficient is statistically significant at 10%, 5% and 1% respectively.

Figure A.4. Highway network in Australia by 1950

Notes: This map shows the Highway network in Australia by 1950

Table A.13. Omitted Variable Bias (Oster's test)

	Dependent Variable: Binary variable for night light			
	(1) Baseline specification coefficient ḃ	(2) Identified Set (ḃ(Rmax,Delta=1),b)	(3) Exclude Zero	(4) Absolute Delta (δ)
Songlines	0.030	$[0.028, 0.030]$	✓	$6.6 > 1$
Local Gov FE	✓	✓		
Geo & Hist Controls	✓	✓		
N	79731	79731		

Notes: The unit of observation is a grid cell of 10km X 10km. The dependent variable is a dummy variable that takes the value of 1 if cell i has nightlight,and 0 otherwise. This variable uses data from the Visible Infrared Imaging Radiometer Suite (VIIRS) sensor. In column (1), the conservative coefficient is given. Column (2) shows the range of coefficients between the conservative estimate (b) and the new b coefficient (ḃ(Rmax,Delta=1)) obtained on the basis of the technique proposed by Oster (2019). In column (3), ✓indicates that (ḃ(Rmax,Delta=1),b) safely excludes zero, indicating that omitted variable bias is ruled out. In column (4), 10.3 ¿ 1 means that the parameter δ exceeds one, indicating that the model from equation 1 is most likely not affected by omitted variable bias. The regression controls for local government fixed effects, as well as for a set of geographical, climatic and historical variables, which include: the coordinates of each grid cell, agricultural suitability, elevation, ruggedness (standard deviation of elevation), rainfall and standard deviation of rainfall, temperature and standard deviation of temperature, distance to the sea, distance to Sydney, distance the state capital, distance to historical mines, coastal presence, and water percentage. The descriptions of the geographical, climatic, and historical variables can be found in the Appendix. Robust standard errors clustered at the local government district level were used and the constant term was omitted for space. *, ** and *** mean that the coefficient is statistically significant at 10%, 5% and 1% respectively.

Table A.14. Omitted Variable Bias (Spatial First Differences)

	Dependent Variable: Binary variable for night light	
	(1) West-East direction	(2) North-South direction
Songlines (Differences)	0.021***	0.02***
	(0.003)	(0.003)
Bootstraped t-statistic	[3.99]	[3.96]
N	79730	79730

Notes: The unit of observation is a grid cell of 10km X 10km. The dependent variable is a dummy variable that takes the value of 1 if cell i has nightlight,and 0 otherwise. This variable uses data from the Visible Infrared Imaging Radiometer Suite (VIIRS) sensor. In column (1), the estimate calculates spatial first differences (SFD) in the west-east direction as proposed by Druckenmiller and Hsiang, 2018, while in column (2) this is calculated for the north-south directions. Robust standard errors clustered at the local government district level were used and the constant term was omitted for space. *, ** and *** mean that the coefficient is statistically significant at 10%, 5% and 1% respectively.

B. Coastal Trade Dynamics

While the main analysis has excluded the first layer of coastal grid cells due to data limitations, coastal trade may still have played an important role in shaping modern economic activity and urbanisation patterns in Australia. Therefore, we address this issue in three ways.

First, we construct a measure based on recorded shipwrecks along the Australian coastline, which serves as a proxy for coastal trade. More specifically, we work with the Australian National Shipwrecks database (*Geospatial Data* n.d.) and count the number of shipwrecks from 1622 to 1900 within a 50-kilometer radius encircling each coastal cell. Then, within our dataset comprising non-coastal cells, we measure the distance of each cell to its nearest coastal counterpart, and assign to it its inverse so that cells that are closer to the coast are assigned a higher weight. We then associate with each observation a magnitude which is the product of the inverted distance and the number of shipwrecks. Column 1 of Table A.16 reports a coefficient on the number of historical shipwrecks, thus adjusted and referred to as Coastal Trade, which is positive and significant. Notably, the coefficient of our main variable of interest, Trade Routes, remains positive and significant.

Second, we utilise data from Bakker et al. (2021) who define a connectedness variable – a coastal indicator for pre-industrial trading opportunities based on the shape of the shoreline. We calculate the average connectedness for each coastal cell within a circular buffer of a 50km radius, and we repeat the same process with the inverse distances as we describe above. That is, as we move further into the interior of Australia, the cells are assigned smaller weights. Then, each

Table A.15. *Natural Routes* above 90th Centile *versus Songlines*

	European exploration	All Cities	Early Railways	Early Highways	Night Lights	Pop. Density
	(1)	(2)	(3)	(4)	(5)	(6)
Natural Routes above the 90th centile	0.0085**	0.0002	0.0012	0.0080**	0.0029***	0.1093***
	(0.0037)	(0.0002)	(0.0014)	(0.0037)	(0.0011)	(0.0325)
Songlines	0.0166***	0.0011***	0.0038**	0.0062***	0.0088***	0.2813***
	(0.0028)	(0.0003)	(0.0016)	(0.0019)	(0.0016)	(0.0428)
Local Gov FE	✓	✓	✓	✓	✓	✓
Geo & Hist Controls	✓	✓	✓	✓	✓	✓
N	79731	79731	79731	79731	79731	79731
R^2	0.075	0.060	0.205	0.143	0.237	0.602

Notes: The unit of observation is a grid cell of 10km X 10km. OLS estimates are shown with robust standard errors clustered at the local government district level. All columns show standardised coefficients. Natural Routes is a dummy variable that intersects within each pixel that takes the value of 1 if cell i is intersected by a number of optimal routes which is above the 90th centile of the numbers of optimal routes that intersect all cells, and 0 otherwise. Aboriginal Trade Routes is a dummy variable that takes the value of 1 if cell i has at least one Aboriginal trade route, and 0 otherwise. In column (1), the dependent variable is an indicator that takes the value of 1 if there exists at least one European exploration route in a grid cell, and 0, otherwise. In column (2), the dependent variable is a indicator that takes value of 1 if grid cell i is associated to the establishment of a city in Australia between 1788 and 2000, and 0 otherwise. In column (3), the dependent variable is an indicator that takes value 1 if at least one railway was built in grid cell i between 1880 and 1920, and zero otherwise. In column (4), the dependent variable is an indicator that takes value 1 if at least one highway was built in grid cell i until the early 1950s, and zero otherwise. In column (5), the dependent variable is an indicator that takes the value of 1 if grid cell i has nightlight, and 0 otherwise. In column (6), the dependent variable is the logarithm of human population density, which is the number of individuals per grid cell. The data is derived from national censuses and population registers for the year 2015, with an additional minor adjustment of 1e-8 included. All columns control for local government fixed effects, as well as for a set of geographical, climatic and historical variables, which include: the coordinates of each grid cell, agricultural suitability, elevation, ruggedness (standard deviation of elevation), rainfall and standard deviation of rainfall, temperature and standard deviation of temperature, distance to the sea, distance to Sydney, distance the state capital, distance to historical mines, coastal presence, and water percentage. The descriptions of the geographical, climatic, and historical variables can be found in the Appendix. The constant term was omitted for space. *, ** and *** mean that the coefficient is statistically significant at 10%, 5% and 1% respectively.

Table A.16. Coastal Trade Dynamics

	Dependent Variable: Binary variable for night light			
	(1)	(2)	(3)	(4)
Songlines	0.031***	0.029***	0.026***	0.026***
	(0.005)	(0.005)	(0.005)	(0.005)
Coastal Trade	0.028***			0.021***
	(0.006)			(0.005)
Connectedness		−0.007**		0.001
		(0.003)		(0.002)
Local Gov FE	✓	✓	✓	✓
Longitude FE			✓	✓
Longitude FE X Dist. to Sea			✓	✓
N	79731	78238	79731	78238
R^2	0.234	0.231	0.251	0.253

Notes: The unit of observation is a grid cell of 10km X 10km. The dependent variable is a dummy variable that takes the value of 1 if cell *i* has nightlight, and 0 otherwise. This variable uses data from Visible Infrared Imaging Radiometer Suite (VIIRS) sensor. Trade Routes is a dummy variable that takes the value of 1 if cell *i* has at least one Aboriginal trade route, and 0 otherwise. All columns control for local government fixed effects. In column (1) Coastal Trade is variable measuring coastal trade intensity. In column (2), Connectedness measures the mean connectedness of the closest coastal cell within 50km. In column (3) and (4), it is added longitude fixed effects and the interaction between longitude fixed effects and proximity to coast. All regressions controls for local government fixed effects, as well as for a set of geographical, climatic and historical variables, which include: the coordinates of each grid cell, agricultural suitability, elevation, ruggedness (standard deviation of elevation), rainfall and standard deviation of rainfall, temperature and standard deviation of temperature, distance to the sea, distance to Sydney, distance the state capital, distance to historical mines, coastal presence, and water percentage. The descriptions of the geographical and climatic variables can be found in the Appendix. Robust standard errors clustered at the local government district level were used and the constant term was omitted for space. *, ** and *** mean that the coefficient is statistically significant at 10%, 5% and 1% respectively.

cell gets a value which is the multiplication of its weight and the average connectedness of its closest coastal cell.[29] Column 2 of Table A.16 shows that the coefficient on connectedness is negative and significant. This may suggest that contemporary economic development in Australia has not been influenced significantly by the coastal trade activities, perhaps because of the landscape. Indeed, the northern part of the country, which benefits from greater geographical connectivity to Polynesia and Asia via the sea, has not necessarily experienced markedly greater development compared to the southern part. The latter hosts prominent urban centers, such as Sydney and Melbourne.

Third, in column 3 of Table A.16, we control for longitude fixed effects coupled with their interaction with proximity to the sea. These controls aim at reducing the risk of confounding effects stemming from location-specific variables that are constant over time within each longitude and for potential influence of coastal areas on trade patterns and economic opportunities as one moves closer or farther from the coast. The coefficient on Aboriginal trade routes remains unchallenged. Lastly, in column 4 of Table A.16, we add together all controls of coastal trade along with longitude fixed effects and their interaction with proximity to the sea. Nevertheless, the coefficient of Aboriginal trade routes re-

[29] We lose 1,493 observations by creating this control due to the fact that the closest coastal cells have missing values of connectedness within a circular buffer of a 50km radius.

mains unaffected while the effect of connectedness not only is reversed but also becomes insignificant. This underscores the robustness of our findings and reinforces the notion that Aboriginal trade routes continue to exert a significant impact on contemporary economic activity in Australia, irrespective of coastal trade dynamics.

C. Variable definitions

Table B.1. Variable definitions and sources

Variable	Description	Source
Main variables		
Songlines	Cost-effective routes between origins and destinations using (weighted) least-cost path analysis.	Authors' elaboration from McCarthy (1939)'s historical descriptions and the Human Mobility Index (HMI) as a cost weight drawn from Özak (2018).
Night light	Dummy variable showing if there is light within a cell.	Authors' elaboration using VIIRS's lights in 2015 from Elvidge et al. (2017) and NOOA/National Centers for Environmental information: here
Latitude/Longitude	Geographic coordinates of the cell.	Authors' elaboration using ArcGIS.
Agriculture Suitability	Average value of seven key soil dimensions important for crop production: nutrient availability, nutrient retention capacity, rooting conditions, oxygen availability to roots, excess salts, toxicities, and work-ability. It corresponds to the average value of the surface area of the cell.	Authors' elaboration using data from Fischer et al. (2008): here
Elevation	Average altitude in meters above or below the sea level. It corresponds to the average value of the surface area of the cell.	Author's elaboration using data from DIVA-GIS.
Ruggedness	Standard deviation of the altitude (in meters above or below the sea level) of the territory corresponding to the cell.	Author's elaboration using data from DIVA-GIS.
Precipitation and (StDev Precipitation)	Mean and (standard deviation) of annual precipitation, in millimeters. They correspond to the average value of the surface area of the cell.	Author's elaboration using data from WorldClim by Hijmans et al. (2005).
Temperature and (StDev Temperature)	Annual average and (standard deviation) of temperature, in degrees of Celsius. It corresponds to the average value of the surface area of the cell.	Author's elaboration using data from WorldClim by Hijmans et al. (2005).
Distance to the Sea	The geodesic distance from the centroid of each cell to the nearest coastline, in kilometres.	Authors' elaboration using ArcGIS.
Coastal Dummy	Dummy showing whether cells face a coast using the second layer of coastal cells.	Authors' elaboration using ArcGIS.
Distance to Sydney	The geodesic distance from the centroid of each cell to Sydney, in kilometres.	Authors' elaboration using ArcGIS.
Distance to State Capital	The geodesic distance from the centroids of the cells of each State to their corresponding State capital, in kilometres.	Authors' elaboration using ArcGIS.
Distance to Historical Mine	The geodesic distance from the centroids of the cells to the closest historical mine, in kilometres.	Author's elaboration using data from the Australian mining history association: here
Water Percentage	Percentage of water due to rivers, canals, and lakes of the cell.	Author's elaboration using DIVA-GIS.
Other variables		
Early Railways	Dummy variable showing whether there is an early railway between 1880-1920 in the grid cell.	Author's elaboration digitising historical maps of railways. See Figure A.3 in the Appendix.
Early Highways	Dummy variable showing whether there is an early highway by 1950 in the grid cell.	Author's elaboration digitising a historical map of the highway network in 1950. See Figure A.4 in the Appendix.
Exploration Routes	Dummy variable showing whether there is a European exploration route during colonisation in the grid cell.	Author's elaboration digitising a historical map by Robinson (1927). See Figure A.2 in the Appendix.
Cities 1788-2000	The foundation year of the most populated cities (that cover almost 90% of the population) in Australia.	Data from Kampanelis (2019).
Towns	The foundation year of non-coastal towns in Australia.	Data from Elder (2024).
Population density	Logarithm of human population density (number of persons per grid cell) based on counts consistent with national censuses and population registers for 2015 plus a tiny number (1e-8).	Author's elaboration using population density data from Global Human Settlement Layer (GHSL): here. European Commission (2023).
Centralities	Authors' computations based on historical descriptions and the Human Mobility Index (HMI) as a cost weight drawn from Özak (2018).	See definitions for degree, betweenness, eigenvector and K-B centralities in 6.5
Trading Sites	A dummy variable that takes the value of 1, if at least one Aboriginal trading point (origins and destinations) is present in each cell, and 0 otherwise.	Authors' elaboration from McCarthy (1939)'s historical descriptions.
Alternative sources	Cost-effective routes constructed based on origins and destinations drawn from McCarthy (1939), rock art locations, and historical maps of McCarthy (1939). The routes then rely on a (weighted) least-cost path analysis.	Authors' elaboration from McCarthy (1939)'s historical descriptions, Human Mobility Index (HMI) by Özak (2018), rock art locations drawn from the Centre for Rock Art Research and Management here, and McCarthy (1939)'s maps.
Natural Routes	The variable "Natural Routes" measures the traversability of locations in Australia based on geographically intrinsic routes. It quantifies the number of optimal travel paths intersecting each pixel on a 10 km x 10 km map, with higher scores indicating better Natural Routes.	Authors' computations based on the Human Mobility Index (HMI) as a cost weight drawn from Özak (2018), using ArcGIS.
Coastal Trade	The multiplication of the number of shipwrecks between 1622 and 1900 and the inverted distance related to the closest coastal cell. Further explanation in Section B of Appendix.	Authors' elaboration using data from the Australian National Shipwrecks database: here
Connectedness	The mean connectedness of the closest coastal cell within 50km.	Authors' elaboration using data from Bakker et al. (2021)

D. Aboriginal Guides

The exploration of Australia during the 19th century was heavily dependent on the knowledge and skills of Aboriginal guides. These guides were instrumental in helping European explorers navigate the challenging and unfamiliar landscapes of Australia. Their expertise was vital in locating water, finding food, and facilitating communication with Indigenous communities. Explorers who employed Aboriginal guides had two main strategies: they could either recruit Aboriginal men from settled areas to join the expedition from the start or engage with local Aboriginal people when they encountered them during their journey to gain valuable knowledge directly from them (Clarke, 2008; McLaren, 1996). In this Appendix, we present comprehensive evidence on the Aboriginal companions of European explorers. **Table X** includes an extensive list of **76** documented inland explorers, primarily from the 19th century, along with compelling evidence demonstrating that they were indeed accompanied by Aboriginal guides. For each explorer, we also cite the source(s) that substantiate this evidence. These source(s), among others, include published explorers' diaries, the Australian Dictionary of Biography, websites of local councils, as well as documents from the National Library of Australia. While we acknowledge that it is impossible to document every explorer who ventured across Australia—particularly since many expeditions may have gone unrecorded—to the best of our knowledge, this is the first systematic compilation of consistent data on inland explorers and their Aboriginal guides. Based on this evidence, in the following subsections, we present key examples highlighting the importance of Aboriginal companions for a selected group of explorers from among all those presented in the Appendix. They are organized by specific regions across Australia.

D.1. Southeastern Australia: The Vital Role of Indigenous Knowledge

In Southeastern Australia, where two of the most significant urban centers, Sydney and Melbourne, emerged, Aboriginal guides played a crucial role in supporting European explorations. Hamilton Hume, a renowned Australian explorer, often relied on the invaluable guidance of Aboriginal people during his numerous expeditions. From his earliest explorations, Hume recognized the importance of Aboriginal knowledge and skills in navigating the challenging Australian terrain. On his first journey at the age of 17 in 1814, he was accompanied by an Aboriginal man who helped him traverse the Berrima-Bong Bong district, part of the Southern Highlands of New South Wales (Hume, 1966) . Hume's respect for Aboriginal expertise continued throughout his career, most notably during his collaboration with Charles Sturt on the 1828 expedition to the interior. Hume's ability to communicate and negotiate with Aboriginal communities was crucial in securing safe passage and gaining local insights, which significantly contributed to the success of his explorations. In addition to Hume, other explorers such as Charles

Throsby who explored the southeastern Australia from 1817 to 1821, similarly relied on Aboriginal guidance. Broughton, whose Aboriginal name was recorded as Toodwick, Toodood, or Toodwit, became a trusted figure, serving as a guide and translator on several of Throsby's expeditions (Campbell, 2005). In the same vein, George William Evans was accompanied by Aboriginal guides during his explorations from 1812 up to 1818. One notable instance is his 1812 expedition from Jervis Bay to the Shoalhaven River, where he was guided by an Aboriginal man named Bundle (also known as Bon-del or Bundel) from the Gundungurra area. In November 1813, he led a team of five men to survey a route through the mountains, aiming to reach the fertile plains beyond for agricultural expansion to support the Port Jackson colony. They traced the course of the Wambool (Macquarie River) westward into Wiradjuri territory, extending their journey to Killongbutta, approximately 40 kilometers from Bathurst. The path they took was along a well-established route (i.e., a songline) historically used by the Dharug and Gandangara peoples for trade with the Wiradjuri (Bathurst Local Aboriginal Consultative Committee, 2011).

D.2. Central Australia: The Harsh Realities and Aboriginal Guidance

The stark and unforgiving landscapes of central Australia presented formidable challenges to early explorers. Ernest Giles, whose expeditions in the 1870s ventured into the arid interior, often relied on Aboriginal guides to survive. The presence of guides like Tommy Oldham, a Wirangu man, was crucial in navigating the desolate deserts, finding scarce water sources, and avoiding natural dangers. On one significant occasion, Oldham discovered Queen Victoria Spring, a vital water source that saved Giles' expedition from dying of thirst. In his personal diary, Australia Twice Traversed (Giles, 1889), Giles notes:

"It was Mr. Tietkens's turn to steer, and he started on foot ahead of the string of camels for that purpose. He gave Tommy his little riding-bull, the best leading camel we have, and told him to go on top of a white sandhill to our left, a little south of us, and try if he could find any fresh blacks' tracks, or other indications of water. I did not know that Tommy had gone, nor could I see that Tietkens was walking—it was an extraordinary event when the whole string of camels could be seen at once in a line in this country—and we had been traveling some two miles and a half when Alec Ross and Peter Nicholls declared that they heard Tommy calling out "water!" I never will believe these things until they are proved, so I kept the party still going on. However, even I, soon ceased to doubt, for Tommy came rushing through the scrubs full gallop, and, between a scream and a howl, yelled out quite loud enough now even for me to hear, "Water! water! plenty water here! come on! come on! this

way! this way! come on, Mr. Giles! mine been find 'em plenty water!" I checked his excitement a moment and asked whether it was a native well he had found, and should we have to work at it with the shovel? Tommy said, "No fear shovel, that fellow water sit down meself (i.e. itself) along a ground, camel he drink 'em meself.Of course we turned the long string after him."[30]

Giles' explorations clearly demonstrate the critical impact of Aboriginal expertise; when skilled guides were involved, the survival and success of his party increased significantly. In the same vein, another significant explorer, John McDouall Stuart, also found that Aboriginal guidance was essential during his expeditions through central Australia. Stuart's journeys often involved traversing extremely harsh environments, and the Aboriginal guides' knowledge of the land proved to be the difference between success and failure. Stuart embarked on an expedition on 14 May 1858 with an assistant and an Aboriginal tracker, carrying provisions for four weeks, to explore beyond Lake Torrens and Lake Gairdner in search of grazing land. The group reached as far as Coober Pedy before turning South and then West. However, on 3 August, the Aboriginal tracker departed, leaving Stuart and his team to face significant challenges. With their supplies and water nearly depleted and their horses lame, they struggled to reach T. M. Gibson's outstation at Streaky Bay by the 22nd of August.[31] After resting for ten days, Stuart returned to Adelaide (Stuart, 1865). The departure of the Aboriginal tracker on the 3rd of August clearly marked a turning point in Stuart's expedition. Without the essential guidance and knowledge provided by the Aboriginal guide, the European party encountered severe difficulties, including the depletion of their supplies and water, as well as the deteriorating condition of their horses.

D.3. Western Australia: Success Through Collaboration

In Western Australia, the expeditions of John and Alexander Forrest are prime examples of successful collaboration between European explorers and Aboriginal guides. Windich, an Aboriginal guide who accompanied the Forrest brothers, played a pivotal role in their exploration of Lake Moore and Lake Barlee in the late 1860s and early 1870s. His deep understanding of the region's geography, including the intri-

[30] Mr. William Henry Tietkens (1844–1933) served as second-in-command on several of Giles's expeditions.

[31] T. M. Gibson's outstation refers to a remote or secondary station or homestead that was part of a larger pastoral property owned or managed by someone named T. M. Gibson. In the context of John McDouall Stuart's exploration, it would have been a place where the explorers could find shelter, supplies, and possibly assistance during their journey. Outstations were common in the Australian outback during the 19th century and served as critical support points for explorers, stockmen, and travellers, especially in isolated regions far from main settlements.

cate web of waterholes and their seasonal variations, was instrumental in the success of these expeditions. Windich was on particularly good terms with the Forrest family, and both John and Alexander Forrest heavily relied on him in their daily search for drinking water and horse feed. Windich usually acted as the scout, adept at finding native wells or waterholes in the rocky outcrops. While the Forrest brothers, being skilled surveyors, were capable of navigating using astronomical observations and were not at risk of getting lost in the inland deserts, they depended greatly on their Aboriginal tracker for finding drinking water and forage for their horses. The mutual respect and reliance between the Forrest brothers and Windich highlight how critical Indigenous knowledge was in unlocking the mysteries of Western Australia's remote landscapes (Forrest, 1875).[32]

Similarly, the explorations of Peter Egerton Warburton also exemplify the importance of Aboriginal guides in Western Australia. His mid-1850s expeditions benefited immensely from the guidance of an Aboriginal companion named Charley, whose expertise in navigating the vast, often hostile terrains was indispensable. Warburton credited Charley with saving the expedition multiple times by seeking out water and being sent ahead to scout paths and resources. In his book, called Journey Across the Western Interior of Australia (Warburton, 1875), which is a published account of his diaries from his explorations, he notes:

"At 1 p.m. Charley returned with news that they had tracked the natives round from north-west to south-west, and found their well. Lewis remained there to clear it out, and I shall move on towards it this afternoon, in time to water the camels."

In another section of the same book he mentions:

"Sent Lewis and Charley at daylight to follow up the native tracks; Charley returned at 10 a.m. with news of a good well about five miles to the north-west. Lewis stayed at the well to save the camel."[33]

Warburton's diary notes suggest that his journeys were deeply intertwined with Charley's knowledge, underscoring his essential role as an Aboriginal guide. Charley's contribution to Warburton's expedition was twofold. Firstly, his ability to expertly track Aboriginal paths was crucial in guiding the entire Warburton party through the challenging terrain. Secondly, his deep knowledge of the natural environment and local geography was vital to the survival of the expedition, ensuring that the party could find essential resources such as water in the harsh and unforgiving landscape.

[32] Other Aboriginal tracker(s) such as Tommy Pierre were also Aboriginal guides with Forrest's party.

[33] William Lewis was a European member of Warburton's expedition across the western interior of Australia.

D.4. Northern Australia: The Mixed Fortunes of Exploration

Northern Australia, with its remote and often hostile environment, posed unique challenges to explorers like Ludwig Leichhardt. His expeditions from the Darling Downs to Port Essington in the 1840s were marked by a reliance on Aboriginal guides, particularly Harry Brown and Charley Fisher, whose contributions were crucial for the survival and success of the exploration party. Brown, known for his exceptional navigation skills, frequently scouted ahead to find the best routes and locate vital water sources, such as the freshwater lagoon that became known as "Brown's Lagoons." His role was so essential that even when Leichhardt realized they had insufficient supplies during his first major expedition in 1844, he chose to send two European members (Caleb and Hodgson) back while retaining both Brown and Fisher (Leichhardt, 1847). Brown's importance was further highlighted when he corrected Leichhardt's course during a critical moment, ensuring the party remained on the right path. More specifically, during the second expedition in 1846, Leichhardt mistakenly led the party in the wrong direction. Brown, with his deep knowledge of the terrain, realized the error and informed Mann, another member of the expedition, that they were off course.[34] Despite Leichhardt's initial confidence in his navigational skills, Brown's insight proved correct, and he was called upon to lead the party through the difficult terrain, ultimately ensuring their survival and continued progress (Blyton, 2015). Among the exceptions to expeditions accompanied by Aboriginal guides that ultimately were not successful is Ludwig Leichhardt's third and final journey in 1848. Despite the presence of two Aboriginal guides, Wommai and Billy, the expedition ended in tragedy, with the entire party disappearing without a trace. This outcome contrasts sharply with Leichhardt's earlier expeditions, where the presence of skilled guides like Harry Brown had been critical to their success. Similarly, Edmund Kennedy and his guide Jackey-Jackey formed a partnership that was crucial in their exploration of northern Queensland. Despite the many dangers they faced, Jackey-Jackey's intimate knowledge of the land and his ability to navigate through dense rainforests and swamps were invaluable to Kennedy's expedition. Jackey Jackey's loyalty and bravery were particularly evident towards the end of the expedition when he remained with Kennedy until his death and ultimately reached the supply ship alone.

[34] John F. Mann was one of the European members of Ludwig Leichhardt's second expedition in 1846.

D.5. Conclusion on the Role of Aboriginal Guides

The role of Aboriginal guides in the exploration of Australia extends far beyond the immediate survival of the explorers they accompanied. These guides were often recruited from survivors of the early waves of British colonization and were present at pivotal moments in Australian exploration history. The partnerships formed between explorers and Aboriginal companions—such as those between Mathew Flinders and Boongaree, George Grey and Kaiber, John Mitchell and Barney, Throsby and Toodwick, Edward Eyre and Wylie, Ludwig Leichhardt and Charley Fisher, Peter Warburton and Charley, Ernest Giles and Jimmy, and John Forrest and Windich—were instrumental in opening up the Australian interior to European settlement (Reynolds, 1990; McLaren, 1996; Clarke, 2008). Apart from the most well-known explorers, lesser-known figures such as James Blackman was also accompanied by an Aboriginal servant named Aaron during his exploration from Bathurst to the Cudgegong River in 1821, while Nathan Buchanan relied on a Warumungu man named Jack during his explorations in the Tennant Creek area in 1896. Similarly, William Hann, a less known explorer, was guided by an Aboriginal man named Jerry during his 1872 Northern Expedition of Cape York Peninsula. Many of these Aboriginal guides were later recognized for their contributions through the awarding of brass or copper breastplates, acknowledging their indispensable role in these exploratory missions (Cleary, 1993; Troy, 1993). Their deep understanding of the land, encoded in their ecological knowledge and cultural practices, provided European explorers with the tools they needed to survive and thrive in an unfamiliar and often unforgiving environment. The tragic fate of explorers like Burke and Wills, who failed to fully integrate Aboriginal guides into their team, highlights the dangers of neglecting this vital resource. Their experience stands as a stark reminder of the critical importance of Aboriginal expertise in the exploration of Australia (Clarke, 2008; Cleary, 1993).

E. Luminosity as a Measure Local Income

In our analysis, we utilize luminosity as an indicator of economic activity. While previous studies have demonstrated the efficacy of using light as a proxy for regional development in a few developed countries (Ebener et al., 2005; Ghosh et al., 2010; Sutton et al., 2007), the applicability of this approach to Australia warrants elaboration. Despite being considered a developed economy, Australia's unique historical trajectory, particularly stemming from European colonization, alongside its varied geomorphology, has resulted in a predominantly coastal urban distribution with vast, sparsely inhabited interior regions. Although our primary analysis excludes the coastal areas hosting major

Australian cities, questions may be raised regarding the ability of nightlights to accurately capture economic activity in both the interior and the coastal part.

Hence, we explore the correlation between our measure of light density (the continuous measure in logs) and two conventional indicators of local economic growth: mean income and median income. Specifically, we draw upon data from the Australian Bureau of Statistics, which furnishes mean and median employee income values at the Local Government Area (LGA) level spanning 2013 to 2018. Our preference for 2015 data aligns with the temporal consistency of our light and population density metrics, both also sourced from 2015.

Each pixel in our analysis is assigned to its corresponding LGA. In instances where a pixel is intersected by more than one LGA, we join it with the LGA that covers its larger part. Subsequently, these pixels inherit the mean and median income values of their respective LGAs. Leveraging both mean and median income offers a comprehensive perspective: while mean income provides a central tendency measure, the median is more robust to outliers, rendering it a robust measure in the presence of skewed distributions or income inequality.

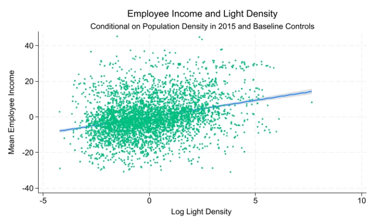
Given that our income data pertains to larger geographic areas than our pixel analysis, we restrict our correlation analysis to 4379 pixels with non-zero values for both population density and nightlights. This is approximately 5,5% of our total sample. This approach excludes pixels that lack income, light, or population data, thus enhancing the robustness of our correlation assessments.

Figure B.1 demonstrates the robust positive correlation between our luminosity index and both income metrics, signalling significant associations. This positive relationship persists across both measures, affirming the strength of the correlation.

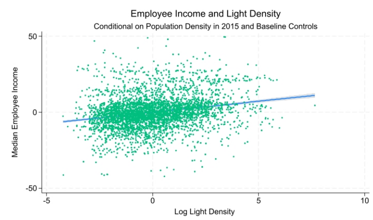
Figure B.1 also delves deeper into these associations by excluding coastal regions within the 100km to 500km range. Remarkably, even with this exclusion, the correlations remain high and the slopes of the relationships become steeper. This finding is particularly pronounced when focusing on the remote interior regions of Australia, traditionally characterized by lower development levels.

Conversely, we conduct a complementary analysis by retaining the coastal areas and excluding the interior regions extending beyond 100km to 500km from the coast. Despite this spatial shift, the positive correlations persist, reaffirming the reliability of our luminosity index as an indicator of economic activity at the local level across diverse geographic contexts.

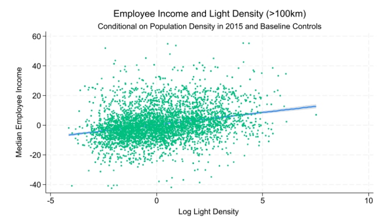
In sum, our comprehensive analysis underscores the efficacy of luminosity as a reliable proxy for economic activity in Australia, even at the granular local level, thereby bolstering its utility for our empirical research.



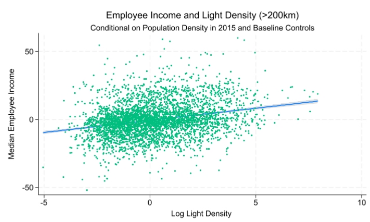
(a) Mean Income



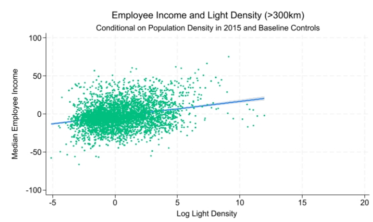
(b) Median Income



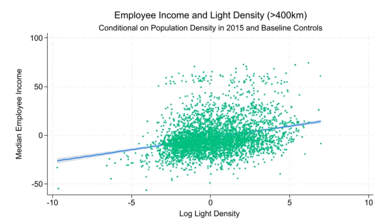
(c) > 100 km



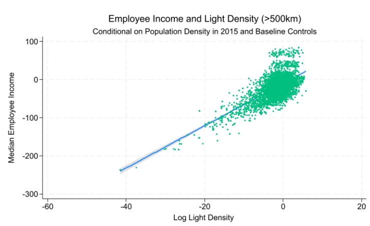
(d) > 200 km



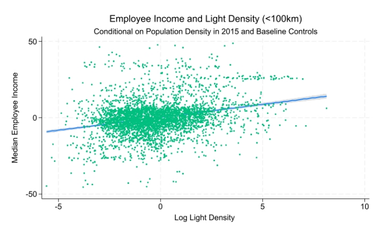
(e) > 300 km



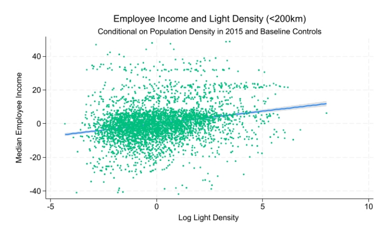
(f) > 400 km



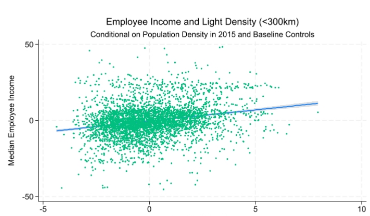
(g) > 500 km



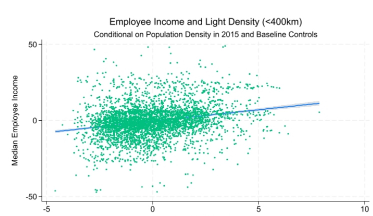
(h) < 100 km



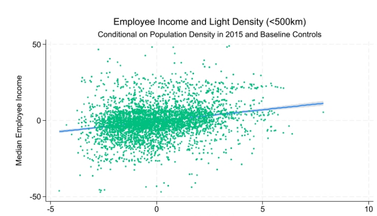
(i) < 200 km



(j) < 300 km



(k) < 400 km



(l) < 500 km

Figure B.1. Relationship between luminosity and income conditional on population density and baseline controls.

F. Appendix References

References

[1] Ebener, S., Murray, C., Tandon, A., and Elvidge, C.C. (2005). *From Wealth to Health: Modelling the Distribution of Income per Capita at the Sub-National Level Using Night-Time Light Imagery*. International Journal of Health Geographics, 4(1), 5.

[2] Ghosh, T., Powell, R.L., Elvidge, C.D., Baugh, K.E., Sutton, P.C., and Anderson, S. (2010). *Shedding Light on the Global Distribution of Economic Activity*. The Open Geography Journal, 3(1), 148–161.

[3] Sutton, P. C., Elvidge, C. D., and Ghosh, T. (2007). *Estimation of Gross Domestic Product at Sub-National Scales Using Nighttime Satellite Imagery*. International Journal of Ecological Economics & Statistics, 8(S07), 5–21.