RUHR
ECONOMIC PAPERS

Sreeja Jaiswal
Gunther Bensch
Aniket Navalkar
T. Jayaraman

The Socio-Economic and Environmental Impact of a Large Infrastructure Project: The Case of the Konkan Railway in India

rwi

#936

Imprint

Ruhr Economic Papers

Published by

RWI – Leibniz-Institut für Wirtschaftsforschung
Hohenzollernstr. 1-3, 45128 Essen, Germany

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Editors

Prof. Dr. Thomas K. Bauer
RUB, Department of Economics, Empirical Economics
Phone: +49 (0) 234/3 22 83 41, e-mail: thomas.bauer@rub.de

Prof. Dr. Wolfgang Leininger
Technische Universität Dortmund, Department of Economic and Social Sciences
Economics – Microeconomics
Phone: +49 (0) 231/7 55-3297, e-mail: W.Leininger@tu-dortmund.de

Prof. Dr. Volker Clausen
University of Duisburg-Essen, Department of Economics
International Economics
Phone: +49 (0) 201/1 83-3655, e-mail: vclausen@vwl.uni-due.de

Prof. Dr. Ronald Bachmann, Prof. Dr. Manuel Frondel, Prof. Dr. Torsten Schmidt, Prof. Dr. Ansgar Wübker
RWI, Phone: +49 (0) 201/81 49-213, e-mail: presse@rwi-essen.de

Editorial Office

Sabine Weiler
RWI, Phone: +49 (0) 201/81 49-213, e-mail: sabine.weiler@rwi-essen.de

Ruhr Economic Papers #936

Responsible Editor: Manuel Frondel

ISSN 1864-4872 (online) – ISBN 978-3-96973-096-6

Ruhr Economic Papers #936

Sreeja Jaiswal, Gunther Bensch, Aniket Navalkar, and T. Jayaraman

The Socio-Economic and Environmental Impact of a Large Infrastructure Project: The Case of the Konkan Railway in India

Bibliografische Informationen der Deutschen Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at **http://dnb.dnb.de**

RWI is funded by the Federal Government and the federal state of North Rhine-Westphalia.

http://dx.doi.org/10.4419/96973096
ISSN 1864-4872 (online)
ISBN 978-3-96973-096-6

Sreeja Jaiswal, Gunther Bensch, Aniket Navalkar, and T. Jayaraman[1]

The Socio-Economic and Environmental Impact of a Large Infrastructure Project: The Case of the Konkan Railway in India

Abstract

Railways are a key infrastructure that facilitates trade and regional integration with potential consequences on local development and the environment in hitherto backward regions. In this article, we study the medium- to long-term socio-economic and environmental infrastructure impacts for the case of the Konkan Railway, which is one of the biggest railway construction endeavours in independent India. We employ a quasi-experimental mixed-methods design to explore the impact of the Konkan Railway on population, workforce composition and land cover types using census and satellite data. We find that the Konkan Railway led to an increase in the female-to-male sex ratio and a negative effect on the share of male workers among the working population. In combination with qualitative evidence, this suggests that the railway access has reinforced the pre-existing pattern of high levels of male migration. We also find an increase in population and the workforce participation rate without disparate workforce effects across sectors suggesting that the railway had moderate effects across the local economies. In terms of land use, the analysis could not substantiate concerns regarding substantive loss of forest cover induced by the railways. The findings encourage policy makers – in assessing the effects of transport infrastructure – to take into consideration the impact on migration, labour mobility and labour market outcomes in sending and receiving regions.

JEL-Code: N75, O18, O40, R11, R41

Keywords: Infrastructure; railway access; migration; impact evaluation; mixed methods; India

January 2022

1 Sreeja Jaiswal, Centre for Climate Change and Sustainability Studies and MS Swaminathan Research Foundation; Gunther Bensch, RWI; Aniket Navalkar, Centre for Climate Change and Sustainability Studies; T. Jayaraman, Centre for Climate Change and Sustainability Studies and MS Swaminathan Research Foundation. *- This research was made possible by the grant received from the International Initiative for Impact Evaluation (3ie), through their Development Priority Window. We thank 3ie and their anonymous reviewers for their comments and feedback from which this paper has greatly benefited. We are grateful to the Konkan Railway Corporation Limited for providing access to their data and to the Centre for Education and Documentation, Mumbai and Bangalore, for giving us access to their newspaper archives on the Konkan Railway. We thank Anandi Yadav, Sandeep Mahato and Gaurav Gharde for the excellent research assistance provided at various stages of the research. We are further indebted to R Ramachandran for bringing this topic of research to our attention at the TISS Annual Climate Change Conference 2014.* *– All correspondence to: Sreeja Jaiswal, Centre for Climate Change and Sustainability Studies and MS Swaminathan Research Foundation, India, e-mail: sreeja.jaiswal@gmail.com*

1. Introduction

Inadequate transport infrastructure in developing countries is widely seen as an obstacle to economic growth. At the same time, investments in transport infrastructure lock-in resources and development patterns for decades, which is one of the main reasons why large infrastructure projects continue to be contested in academic and political discourse. Projects requiring significant public investment such as dams, highways, railways and airports have been criticized not only for cost overruns and time delays but have also faced public resistance for their anticipated impact on the environment, on displacement, and loss of livelihood of the local population (see for example Cantarelli et al. 2010; Flyvbjerg 2007). In India as well, several large infrastructure projects, especially in regions designated as ecologically sensitive, have been mired in controversy where concerns of environmental preservation have been pitted against development concerns. One such example is the Western Ghats, the region studied in this article, which is a chain of high ranges that runs parallel to the western coast of India and stretches for about 1600 km from north to south. It influences the monsoon weather pattern in India and is globally recognized for its high level of biodiversity, endemism, and unique geomorphic features. Different expert panels have advanced recommendations for its conservation and proposed restrictions on building new transport and energy infrastructure such as highways, railways, or large hydropower projects (Gadgil et al., 2011; Kasturirangan et al., 2013). Most of these recommendations, however, cannot draw on empirical evidence on the long-term socioeconomic and environmental costs and benefits of large infrastructure projects.

This article explores a set of socio-economic and environmental impacts of a controversial large transport project, the Konkan Railway that runs through the Western Ghats. The mainstay of our empirical strategy is the application of a quasi-experimental impact evaluation methodology. We adopt variants of a difference-in-differences approach that rely on census and satellite data spanning the timeframe from 1991 to 2014. This analysis is supported by an assessment of the operational performance of the Konkan Railway, archival research, and key informant interviews aimed mainly at understanding the preference for road versus rail for long-distance travel. Beyond intermediate outcomes such as freight and passenger performance, the main outcomes assessed at the local level are population and workforce composition and land use change. We assume in our quantitative empirical analysis that access to railways affects the socio-economic and environmental character of localities with different treatment intensities, depending on their distance to the railway facility. Accordingly, the distance to railway facilities is used to construct the key treatment variable for which we identify railway impacts. We employ various robustness checks given different hypotheses about the functional form of the relationship between distance and treatment intensity.

We find that the Konkan Railway had a positive effect on population sizes and workforce participation rates in villages in the proximity of railway stations. In the specific context discussed in the qualitative analysis, this did not go along with a stimulus of or disproportionate effect on non-agricultural or industrial economic activities. These results suggest that the Konkan Railway has encouraged moderate economic dynamics across the local economies. Further, the share of male workers to total workers declined and the sex ratio between females and males increased suggesting that the railway access has reinforced the pre-existing pattern of high levels of male outmigration for better economic opportunities. The improved accessibility makes it easier to undertake frequent trips, thereby helping in maintaining social links with the place of origin. From the satellite data analysis, we find a decrease in the agricultural area and an increase in open spaces and built-up area but only insignificant changes in forest cover and water bodies in close proximity to the railway line as compared to units further away. The analysis, hence, could not substantiate the concerns raised during the time of the construction of the Railway regarding loss of forest cover, at least not on a large scale.

Our results on India, a country with the fourth largest railway network in the world by route length, contribute to the existing literature that mainly studies the impact of historic railway networks, often focusing on population growth, trade and industry outcomes at more aggregate levels such as districts, counties or municipalities. Berger and Enflo (2017) estimate that Swedish towns in the 19th century that gained access to railways experienced a substantial relative increase in population. Similarly, in the 19th century, gaining direct access to railways led to a relative increase in population growth in Swiss (Büchel and Kyburz 2020), Spanish (Esteban-Oliver 2020) and German municipalities (Braun and Franke 2019). Braun and Franke (2019) also find that railways increased wages, income and housing values and accelerated the transition away from agriculture. Donaldson (2018) provides evidence that the railroads built in colonial India significantly reduced trade costs and stimulated inter-state trade. Further studies exploit historic data from colonial Africa (Jedwab and Moradi 2016) and the US (Donaldson and Hornbeck 2016; Hornbeck and Rotemberg 2021); Atack et al., 2010) to assess railway access impacts.

Among the studies which exploit more contemporary data to estimate the impact of railways, Wang and Wu (2015) look more specifically at railway infrastructure built between 2002-2005 in two of China's least developed provinces, Qinghai and Tibet. The authors find that the railway significantly increased the GDP per capita of the railway counties by about 33 percent, mainly by having a positive effect on the competitiveness of the manufacturing industry in the national market. On the contrary, the agriculture and service industries were hardly affected by the new railway, which may be explained by the fact they these sectors tend to compete in the regional market. Furthermore, the authors do not find an effect on population and attribute this result to the strictly enforced so-called Hukou system in China, a kind of household registration system that restricts permanent migration and labour mobility (see also Banerjee et al. 2020, Zhou and Logan 2007, and Baum-Snow et al. 2017).. Banerjee, Duflo, and Qian (2020), who do not differentiate between different modes of transportation, find only small positive causal effect of improved access to transportation networks on per capita GDP levels across sectors for China between 1986 and 2003, a result that may be attributed to a lack of factor mobility in the Chinese setting.

To the best of our knowledge, no study so far has examined the relationship between railway built in post-colonial India and its impact on socio-economic or environmental outcomes. More generally, the literature on impacts of large infrastructure in India traditionally focused on the divergent and contradictory values, belief systems, and discursive framings around the ideas of environmental preservation and development channeled by project proponents and the local communities, often looking at immediate, short-term and transitional impacts (see, for example, Drèze, Samson, and Singh 1997; Kothari 1984; Baviskar 1995; Gadgil and Guha 1995; Guha 2000; Rangarajan 1996; Shah 2010; exceptions include D'Souza, Mukhopadhyay, and Kothari 1998; Aiyar and Kaushal 2019; Duflo and Pande 2007; Ghani, Goswami and Kerr 2016).[1]

The present impact evaluation incorporates the perspective that households use migration as a livelihood diversification and income-enhancing strategy and commuting as a strategy for diversification of workplace. In doing so, it encourages development planners and policymakers to consider the impact of improved transport infrastructure on rural-urban mobility, labour market outcomes in the sending and receiving regions and the importance of environmental management practices on land use change outcomes.

The remainder of the paper proceeds as follows: Section 2 provides background on the study context and study object, the Konkan Railway. Section 3 describes the evaluation design, data, and the empirical model. Section 4 discusses the main findings and Section 5 concludes.

2. Konkan Railway: the infrastructure project and its context

The Konkan region is a narrow strip of the western coastline of India. It is bounded by the Arabian Sea on the west and on the east by the Western Ghats. A common feature of the Konkan region is the history of a high level of circular migration. Migration from the region is referred to as circular as the migrants are simultaneously engaged in their place of origin and destination, making frequent journeys back and forth (Hugo 2003). Circularity allows for the maintenance of family, kinship, traditions, and land. Various reasons have been put forward for migration from the region: poverty, lack of employment opportunities, low level of industrialization, exploitative land tenure systems, and the strong pull of Mumbai as an industrial hub (Deshingkar and Grimm 2005; Tumbe 2015). Migration is male-dominated, a process that goes back to the 19th century and even earlier (Sita and Prabhu 1989; Thorat et al. 2011; Tumbe 2012, 2015).

Given the region's historic link to Mumbai, the railway line connecting Mumbai to Mangalore was a critical project to bridge the "missing link" between the two coastal cities and to spur socio-economic and industrial development of the hitherto backward region (Babu 1997; Fernandes 1993; Sardesai and D'Mello 1992). This connection was traditionally established by coastal steamers until the National Highway (NH) 17 (currently renumbered to NH 66) was built in the 1960s and 1970s. NH 17 is a narrow single-lane highway that runs north to south along the hilly terrain of the Western Ghats, making journeys on it accident-prone and time-consuming, particularly during the annual heavy monsoons this region receives (Iversen and Ghorpade 2011; Bhuyan 2016).[2]

Several proposals for a railway line connecting Mumbai to Mangalore were advanced since 1945. However, financial and engineering obstacles continued to frustrate all endeavours. The project again gained political momentum in the late 1980's when political figures from the Konkan region were at the helm of the Ministries of Railway and Finance in the Government of India. The Konkan Railway Corporation Limited (KRCL) was

[1] Jaiswal and Bensch (2018) provide a more extensive review of the evaluation literature on large infrastructure.

[2] In 2011 the National Highway Authority of India (NHAI) proposed the widening and upgrading of the existing NH 17. As of 2021, the expansion and upgradation work is still ongoing at various stretches.

set up in 1990 as a public sector undertaking under the Ministry of Railways. The KRCL was entrusted with the task of raising funds from the market and constructing the Konkan Railway on a Build Operate Transfer (BOT) basis, the first public infrastructure project in India to be executed by a separate entity and on a BOT mode. Another novelty was that shares in the company were not only held by Indian Railways (51 percent), but that also state governments participated in the financing, with the state governments of Maharashtra, Goa, Karnataka, and Kerala covering the remaining 49 percent.

From 1993 onwards, the Konkan Railway was completed in phases from each end. The complete line of 760 km traverses three districts in Maharashtra (Raigarh, Ratnagiri, Sindhudurg), two districts that make up the state of Goa (North and South Goa) and three districts of Karnataka (Uttara Kannada, Dakshina Kannada and Udupi) (Figure 1). With first operations beginning on a stretch of 128 km in early 1996, Konkan Railway became fully operational in January 1998, nearly three years behind its scheduled inauguration date. When completed, the project cost Rs. 33.5 billion (approximately 2.6 billion USD at 2020 prices), equivalent to a cost overrun of 222 percent in comparison to the revised budget proposed by KRCL in 1991. The cost overruns and time delays were mainly due to inflation, technical problems encountered in digging tunnels through soft soil, halting of work for nine months due to litigation and protests, and the additional time and resources required for the work mandated by the commission appointed to arbitrate in the dispute with environmentalists' groups from Goa. These groups argued that the proposed alignment through Goa would cause significant ecological damage and demanded that the alignment be shifted to pass through the mining belt or the hinterland. However, all the appointed commissions approved the route chosen by the KRCL with minor modifications and suggestions for mitigation of environmental damage.

3. Empirical Strategy

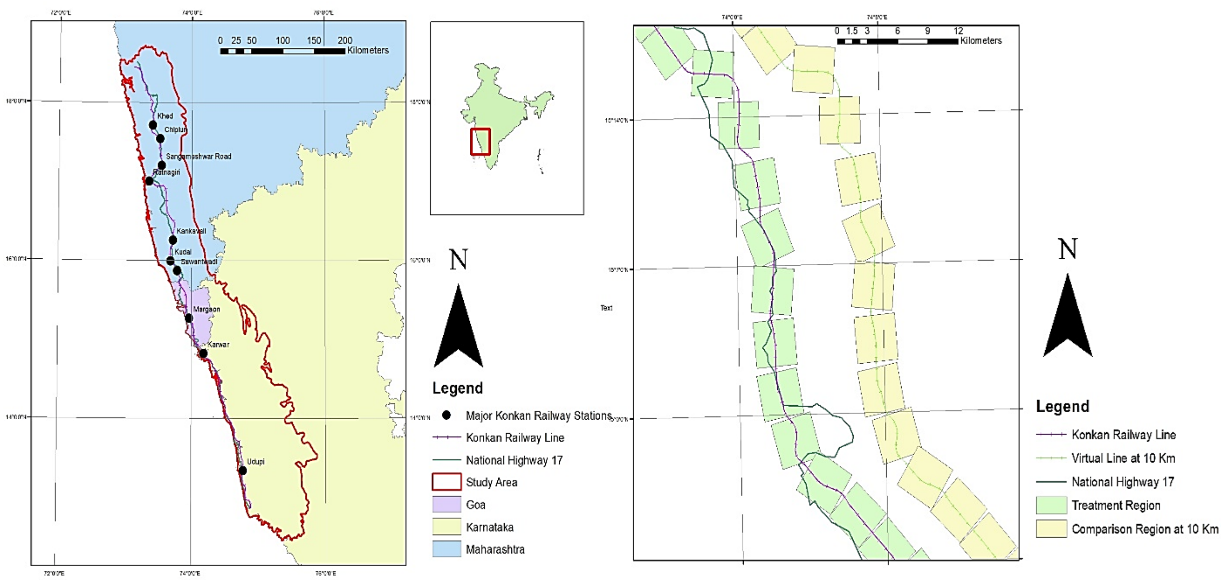
We employ a quasi-experimental design outlined in Section 3.1. The underlying data is presented in Section 3.2, together with mostly qualitative supplementary data used in our mixed-methods analysis.

3.1. Identification strategy

In our main quantitative analysis, we assess the average socioeconomic and environmental impacts of *access to railways*, the *treatment* in evaluation terminology. The design of an identification strategy for an evaluation of a large infrastructure project like the Konkan Railway involves one main methodological challenge: tackling potentially endogenous placement. Initial conditions are likely to determine project placement, as well as to influence the subsequent path of villages and towns affected by the project (Jalan and Ravallion 1998). These initial locality conditions could be physical, geographical, social, or economic. Similar to Wang and Wu (2015), it can be argued that the actual KRCL track due to the challenging geography was largely determined by engineering criteria, which are exogenous to these conditions. KRCL chose an alignment that was as short and flat as possible to optimize load haulage and to be capable of carrying high speed trains, even if this meant building more tunnels.[3] For certain stretches within the limits of the set technical parameters, it tried to avoid heavily built-up locations. This was done to avoid the complications and extra expenditure on compensations related to land acquisitions. Hence, the placement decisions on the railway alignment based on engineering considerations decision helps to mitigate threats of substantive endogeneity in placement.

To minimize remaining bias in the impact estimations, we employ variants of a difference-in-differences (DID) approach. Specifically, we use census and satellite data from different points in time before and after the inauguration of the Konkan Railway to measure the impact of being located close to railways.

Figure 1: Maps of study area and as sample illustration for satellite data impact assessment



(a) Study area (b) Illustration for satellite data impact assessment

Source: The map is generated by processing multi-temporal and multi-spectral satellite data procured from the United States Geological Survey (USGS) and the Indian National Remote Sensing Centre (NRSC)

3.1.1. Satellite imagery analysis using a standard difference-in-differences approach

For the satellite imagery analysis of land use changes, we adopt a standard DID setup:

$$Y_{i,t} = \alpha + \beta_1 P_t + \beta_2 T_i + \beta_3 P_t T_i + \varepsilon_{i,t} \qquad [1]$$

where *i* are 4 km by 4 km squares at time *t*, representing either the baseline or endline (in this case 1991 and 2014, respectively); P_t is the post-treatment time-period dummy, while $\varepsilon_{i,t}$ represents the error term. $Y_{i,t}$ is our land use change outcome variable, β_3 is the DID coefficient on the interaction between P_t and Treatment

[3] The KRCL resurveyed the previous alignment by Southern Railway in 1991 and restricted the curvatures to a radius of 1250 meters or 1.4 degrees and reduced the ruling (i.e. maximum) gradient from 1 in 100 (meaning a steepness of 1/100=1%) to 1 in 150 (steepness1/150 = 0.66%) to optimize speed and haulage. This increased the gross load that a typical Indian locomotive class can haul from 1500 tonnes to 2400 tonnes (Konkan Railway Corporation Limited 1999). The reduced flexibility due to restrictions in curvature also meant that more and longer tunnels were needed. KRCL felt that the additional expense on tunnelling would be balanced by saving in total length and by increased freight capacity.

T_i. We adopt two main alternative definitions of T_i. First, it is defined as the area 2 km on either side of the main KRCL track. Similar to T_i, we define our control group as an area 2 km on either side of a virtual line drawn 20 km in parallel to the east of the main line. Our sample thereby comprises a total of 300 squares measuring 4 km by 4 km along the entire length of the railway track and the virtual line (See Figure 1 (b)). We also try out the analysis with a virtual line drawn 10 km apart.

As an alternative definition of T_i, we define circular buffers around railway stations as treated. In 1998, the Konkan Railway had 52 railway stations excluding the termini stations (for obvious placement endogeneity). The treatment is defined as circular buffers of radius 5 km around the railway station and control buffers are circular buffers of the same radius drawn 20 km away east on the same latitude as the treatment.

3.1.2. Census data analysis using an extended difference-in-differences approach

For the analysis of socio-economic and land use indicators using census data, it seems more appropriate to consider the locality with the new railway *station* as the unit treated with infrastructure access, not the one with the new railway *line*. The exact location of the railway stations is a choice variable, so that the engineering-based exogeneity argument outlined above works less for railway stations than for railway lines. We seek to address the identification challenge through a modified DID approach, inspired by Datta (2012) and Wang and Wu (2015) who use a similar strategy for their work on the impact of improved highways and railways respectively. The estimation equation now looks as follows, using data at the level of locality (village or town) *i*:

$$Y_{i,t} = \alpha + \beta_1 P_t + \beta_3 P_t T_{i,c} + \boldsymbol{\beta_4}'\boldsymbol{X_i}\boldsymbol{P_t} + \beta_5 V_{i,t} + \gamma_i + \varepsilon_{i,t} \qquad [2].$$

This approach accounts for locational factors in three ways. Firstly, we include interactions between P_t and a set of time-invariant locational factors X_i, which include the distance from the nearest highway, latitude, elevation, and distance from the coast.[4] Secondly, we introduce a locality fixed effect γ_i to account for locality-level unobserved factors. Note that we also include a locality dummy $V_{i,t}$ in the model, indicating that the respective locality is a village (and not a town). Unlike the time-constant treatment dummy T_i, this village dummy is not absorbed in the fixed effects since some localities change status over time from village to town or vice versa. Thirdly, to further increase comparability between treatment and comparison units, we restrict our comparison units to the same bio-geographic zone as the treatment units, i.e. to land units with similar altitude, climate, topography and vegetation. Specifically, we include (i) West Coast, (ii) Malabar plain region of Western Ghats and (iii) Mountainous region of Western Ghats up to the termini stations to the north and to the south of the Konkan Railway.

In our main analyses, we show two variants of Equation [2] with different interpretation of the DID coefficient, β_3. In the more straightforward variant, we adopt a binary treatment indicator $T_{i,c}$ with a cutoff distance *c* of localities from the nearest railway station at endline to assign them to a treatment or comparison

[4] This procedure is also inspired by ANCOVA estimations (McKenzie 2012, Hidrobo, Peterman, and Heise 2016 and Chen et al. 2017).

status. T_{10km}, for example, implies that treatment units have a distance of 0 to 10 kms from the nearest railway station, whereas comparison units are more than 10 km (and less than 100 kms) away from their nearest railway station. The second variant involves a modification: since the relationship between distance and treatment intensity is likely to be heterogeneous across distances, we also apply a treatment definition based on multiple distance bands in a similar fashion as done by Ghani, Goswami and Kerr (2016) for highways in India. $\beta_3 P_t T_{i,c}$ in Equation [2] turns into $\sum_{b \in B} \beta_{3,b} P_t T_{i,b}$, where the set *B* contains the four distance bands *b* of localities between 0-5 km, 5-10 km, 10-20 km and 20-30 km. The base category as control are localities with a distance of more than 30 km but less than 100 km from the railway station. The obvious disadvantage of this approach is that it is less powered in statistical terms compared to the binary treatment definition. There are, for example, 720 localities between 0 to 5km from the railway station, 701 villages and 19 towns (see also Table 1 in Section 4.2). The average distance between the stations is 14 km, with a standard deviation of 3.5 km.

Coming to the remaining components of Equation [2], time *t* represents either the baseline or endline in analogy to Equation [2] – now 1991 and 2011, respectively. P_t and $\varepsilon_{i,t}$ are defined as above as the post-treatment time-period dummy and the error term. The set of outcome variables $Y_{i,t}$ in the census data analysis covers a range of demographic and socio-economic indicators, including workforce composition, population composition and land use (see Section 3.2).

3.1.3. Robustness checks

Figure 1 shows how close Konkan Railway was built to the National Highway NH 17, something that is typical to these infrastructures. Equations [2] so far controls for the time-invariant distance to the nearest highway, which is NH 17. Even if the highway was built in the 1960's and 1970's, thus about three decades before the inauguration of the Konkan Railway, it may exert a time-varying effect on our outcomes during the time period observed in our study. We therefore conduct a robustness check where we introduce distance bands *b* into equations [2], now referring to the distance to the nearest highway, $H_{i,b}$:

$$Y_{i,t} = \alpha + \beta_1 P_t + \beta_3 P_t T_{i,c} + \boldsymbol{\beta_4}' \boldsymbol{X_i} \boldsymbol{P_t} + \beta_5 V_{i,t} + \sum_{b \in B} \beta_{6,b} P_t T_{i,c} H_{i,b} + \gamma_i + \varepsilon_{i,t} \quad [3].$$

As can be seen in Equation [3], the distance bands are now interacted with the binary treatment indicator $T_{i,c}$ with a cutoff distance c = 10 km of localities from the nearest railway station at endline. Again, comparison units are more than 10 km and less than 100 kms away.

As a second robustness check, we use the continuous distance $d_{i,1}$, the distance to the nearest railway station after inauguration of the Konkan Railway, directly as treatment variable. The flexible specification of the above Equation [2] using polynomial functional forms looks as follows:

$$Y_{i,t} = \alpha + \beta_1 P_t + \beta_2 P_t d_{i,1} + \beta_3 P_t d_{i,1}^2 + \boldsymbol{\beta_4}' \boldsymbol{X_i} \boldsymbol{P_t} + \beta_5 V_{i,t} + \gamma_i + \varepsilon_{i,t} \quad [4].$$

The validity of the DID identification strategy relies on the assumption that the localities close to the railway station and other localities in our study area have common trends for outcomes of interest, once the fixed effects and controls for baseline locational factors are accounted for. A typical approach to assess the validity of the common trend assumption is to assess pre-treatment trends. Since we lack data for another pre-treatment period to do so, in Section 4.2 we alternatively test baseline levels of the outcomes for our treatment definitions as similar levels at baseline might be suggestive of similar trends as well.

3.2. Data sources and outcomes of interest

For the satellite imagery analysis, the outcome of interest reflects the change in the land use/land cover (LULC). This information was extracted by a so-called post hybrid classification analysis of multi-temporal multi-spectral satellite data. Hybrid classification combines the unsupervised and supervised classification methods and enhances classification accuracy (see Ozesmi and Bauer, 2002; Sader et al., 1995). The satellite dataset comprised of a mix of Landsat and LISS (Linear Integrated Self Scanning Sensor) satellite datasets. Landsat data was used for the base year 1991 while LISS III was used for the years 2003 and 2014. More specifically, the outcome variables represent the shares of the total square area that are covered by the following four LULC classes extracted from the satellite data: open space and built-up area, forest cover, agricultural land as well as water body. The classification follows Anderson's Level 1 classification standardized by the United States Geological Survey (Anderson et al. 1976). Open spaces, for example, include residential, commercial, and industrial built-up area, transportation, mixed urban, dry salt flats, beaches, bare exposed rock, strip mines, quarries, gravel pits, and mixed barren land.

We use three outcome variable categories in the census data analysis: workforce composition, population composition, and land use.[5] The individual indicators for each category are presented in the descriptive statistics section 4.2. The data is retrieved from the Census of India for the years 1991 and 2011, with the former being the baseline data before the Konkan Railway project and the latter the endline data, a decade after its completion. Localities represent the level of observation, of which approximately 7500 villages and 170 towns are in our sample. The census data also captures land use changes within the locality boundaries. The statistics of land use available in the census are compiled from the village land records maintained by the *patwari* or the village accountant. In contrast to the remotely sensed data, the census land use data has several deficiencies with the most obvious being the exclusion of urban land use. Also, the definition of forest in the census includes area legally notified as forest. Area of crops raised in the forest and grazing lands within the forests are included under "forest area" in the census. Further, it does not provide information for instance on social forestry which includes tree planation along railway lines, roadways, rivers and village common land, and built-up land which is important for our study. This calls for exercising caution while comparing, the results from satellite data analysis with census land use. The comparability is further restricted as the census data is from 2010, whereas the satellite data is from 2014. Nevertheless, we also include the results for land use from the census data as it helps us understand the changes within locality boundaries when interpreted in conjunction with the changes in work-force composition.

Further data used in our empirical analysis includes different distance measures, which have been measured

[5] We abstain from using nightlight satellite data as a proxy for economic activity since such data is not available for the pre-treatment period. In addition, the predictive power of nighttime light at smaller geographic scales has recently been disputed in the literature (cf. Goldblatt, Heilmann, and Vaizman 2020).

as Euclidean distances using Geographic Information System (GIS) techniques generated using the digitized village boundaries based on the 2011 (Goa and Maharashtra) and 2001 (Karnataka) village census maps and Digital Elevation Models (DEM).

To complement the quantitative impact analysis, we collected data on intermediate Konkan Railway performance outcomes, namely passenger and freight traffic. Furthermore, 84 semi-structured interviews were carried out with key informants from the following four categories: 60 interviews with elected heads (Sarpanchs) or members of the village self-governing bodies (the Gram Panchayats); seven with KRCL officials and engineers; eleven with representatives from industry and trade and six with activists, environmentalists, and journalists associated with the Konkan Railway realignment controversy during the 1990's.

4. Results and discussion

4.1. Main findings from intermediate outcomes and key informant interviews

To put intermediate outcomes for passenger and freight traffic of the Konkan Railway into perspective, we compare them with data from traffic projection studies commissioned during the planning of the railway. These projections were prepared in 1994 by RITES, an engineering consultancy specializing in transport infrastructure established by the Indian government. The projections were updates of earlier traffic projections given that KRCL anticipated an increased pace of industrialization on the Indian west coast due to a number of mega-industrial projects planned in the region at the time. In particular, the railway was expected to transport mineral, forest and marine resources and to give a stimulus to local petrochemical, metallurgical and food processing industries. However, the quantum of freight traffic fell short of the projections. While the projected freight for the least optimistic scenario for 2005-06 and 2010-11 was 14.32 million tonnes (MT) and 15.43 MT, actual freight carried was 75 percent and 51 percent lower, respectively.

KRCL already noted in its annual report of 1997-1998 that the freight traffic had failed to materialize as traffic flows and economic scenarios had undergone "tremendous changes" during the seven-year construction period (Konkan Railway Corporation 1998). Year-by-year data on freight movement was available from 2003 onwards (Figure 2). Between the year 2003-04 and 2016-17, overall the freight traffic in million tonnes registered a compound annual growth rate (CAGR) of 9 percent. This has been particularly driven by freight traffic terminating on the Konkan Railway stations, which has shown a steady increase with a CAGR of 16 percent. Cross traffic similarly registered a CAGR of 11 percent, but has declined in recent years. Originating freight traffic, however, witnessed a negative CAGR of -4 percent.

Figure 2: Freight movement over time by Type

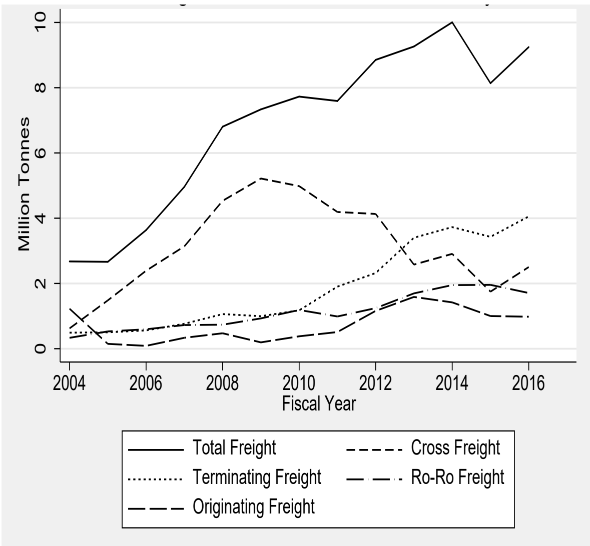
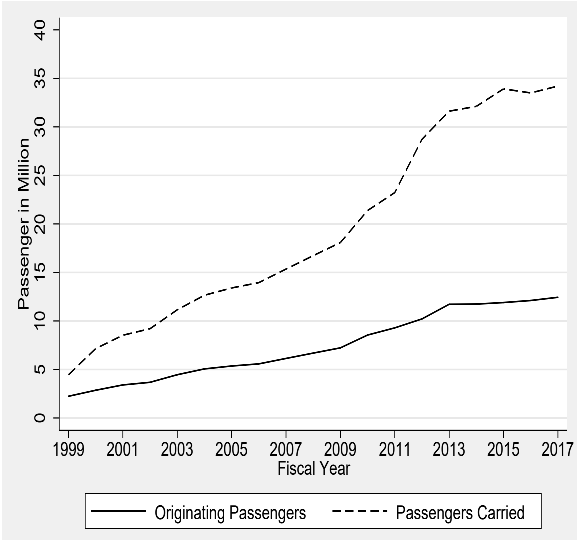


Figure 3: Passenger Traffic over time



Note: Ro-Ro refers to the Roll-on Roll-off service in which loaded goods trucks are carried piggyback on railway flat wagons.
Source: Author's calculation based on KRCL data

We identified several reasons for the shortfall between the projected and actual traffic and the overall bleak freight performance. The RITES study had estimated that industrial traffic from a small number of existing, under construction or planned mega-industrial units in the region would form the bulk of the total freight traffic for the Konkan Railway in the year 2010-11 (8.8 MT or 57 percent). The actual traffic from these identified industrial units was only 2.8 MT or 18 percent of the total traffic in the same year according to our calculations. This can be linked to several factors. First, some of the identified industrial units ended up using alternative modes of transport. For example, food grains were rather shipped through coastal shipping, cement through coastal shipping and road transport and petroleum companies relied more on pipelines (Konkan Railway Corporation 1998). Konkan Railway officials also cited road transport as a primary competitive mode of transport.

Second, some of the industrial units planned in this region were relocated. For example, RITES had estimated that the Konkan Railway would carry copper cathodes from an industrial site that lost its operating license granted by the state government and was therefore relocated in 1994 to another Indian state, Tamil Nadu (Rangarajan 2007; Shrivastava 1994). Growing popular environmental opposition and litigations against iron ore mining activities in western India also contributed to the sharp drop in cross traffic for Konkan Railway in the year 2010-11 discussed above (Arun 2011; PTI 2013, 2014; Shubhashish 2012; Upadhyay 2014). Further reasons for the underperformance of freight transport and the small impact on industrial development in the region according to Konkan Railway officials include the failure of state governments to provide a proper climate for industrial investment (and the overall lack of political will to do so) and the slow development of ports in the west coast to be linked to Konkan Railway. At the same time, Konkan Railway seemed to not have benefitted from connecting the Konkan region via Mumbai to the newly formed coastal industrial cluster south of Mumbai. This region in particular witnessed the post-1991 process of "concentrated decentralization of industries" (Chakravorty 2003, p.135) that combines the concentration of industrial investment in already advanced regions with dispersal of investments within the regions.

Most zones of the Indian Railway earn more revenue from the freight segment (exceptions being North-Eastern Railway and Southern Railway). However, freight earnings of KRCL have surpassed their earnings from the passenger segment only between 2006-2007 and 2010-2011, i.e. in five out of 19 years during the period 1998-1999 to 2016-2017. This also reflects that Konkan Railway has performed well when it comes to the demand for passenger service. The fact that the Konkan Railway has received patronage from the passengers is borne out by data and also by key informant interviews. Between 1998-99 and 2016-17, the CAGR of passengers carried and originating passengers was 11 percent and 9 percent respectively, which corresponds to a 6.7-times and 4.6-fold increase respectively (Figure 2). The number of trains run on the route has also increased substantially. While an average of 30 trains were run on the Konkan Railway route in 2004-05, including 23 passenger and 7 freight trains, the numbers have gone up to 62 including 46 passengers and 16 freight trains in 2014-15. Between 1998-99 and 2015-2016, 40 percent of originating traffic on Konkan Railway has been from stations in Ratnagiri district, which has a long history of male-outmigration to Mumbai, followed by stations in South Goa and Sindhudurg district at 17 per cent each respectively (the share of originating passengers by district is depicted in Appendix C).

The Konkan Railway has routinely increased the frequency of trains and number of halts for popular trains. It has added more coaches to existing trains and run special trains to meet increased demand during religious festivals. The Konkan Railway, being a single-track line, has also gradually added new halt and crossing stations and built new loop lines at a number of stations in order to increase its line capacity. Despite the increasing difficulty to introduce more trains on the network, the Konkan Railway planned to build 21 new stations and 18 new loop lines as of 2016 (Rao 2016; Times News Network 2016).

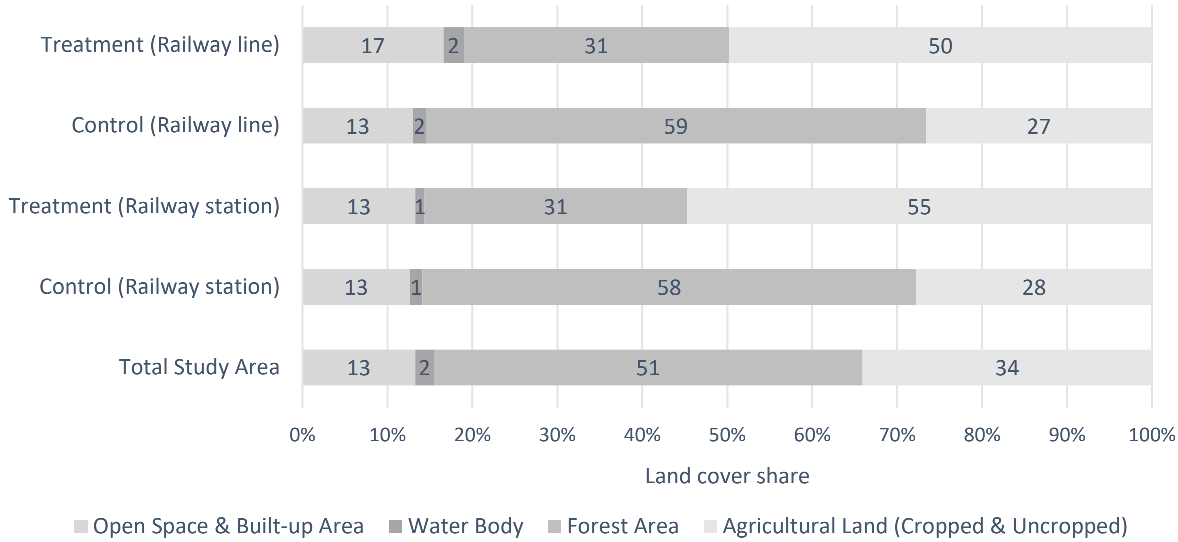
Out of 60 interviews with members of local self-government bodies, 82 percent said that the people of their villages preferred the Konkan Railway to the road for long-distance transport due to cost and time savings, making the travelling more convenient. Train fares in India are in general much lower as compared to private and state-run buses for similar origin-destination combinations due to the heavy subsidies provided by the State, including the policy of Indian Railways to cross-subsidize passenger fares using revenue generated from freight. Public policy, therefore, plays a significant role in shaping people's choice for railways. Over and above the monetary factor, the railways provide users with a tangibly different experience in terms of comfort, on board-facilities (catering and washrooms) and safety, making it, according to the interviewees, a preferred mode of transport especially for families, women and the elderly. Since rail transportation is accessible only through stations, most passenger trips involve a first-mile/ last-mile commute on road, which can change the cost and convenience of the entire trip significantly. From the interviewee responses, we could conclude that buses are preferred to trains in instances where there is no cost-effective way to reach the railway station such as by state transport buses, shared auto-rickshaw or six-seater mini-buses.

We integrate these findings with the results of the econometric data analysis in the next sections. Most importantly, the findings of the freight performance should make us more conservative in our expectations regarding the potential for socioeconomic and environmental impacts of the new Konkan Railway line.

4.2. Descriptive statistics on baseline characteristics

We show descriptive statistics on baseline characteristics for the two main data sources introduced in Section 3.2. To start with the satellite data used to track land use changes, Figure 4 presents land use by the four different land use classes for the baseline year 1991. For our railway line treatment definition, the extent of water bodies to total land cover in the baseline year is 2 percent for both treatment and control areas, open space and built-up space forms 17 percent of land use in treatment area and 13 percent in control area. The major difference is found in the extent of forest cover, which is considerably higher in control than in treatment areas (59 vs. 31 percent), whereas the treatment region has greater extent of agricultural land than the control region (51 vs. 27 percent). For the railway station treatment definition, the extent of open spaces and built-up area is the same for treatment and control regions (13 percent), but the treated area has a greater extent of agricultural land and less forest cover than the control region, with differences similar to the railway line definition. It cannot be excluded that these differences in baseline means, are indicative of structural differences that also affect trends over the observation period. Hence, they need to be accounted for in the interpretation of our results. One likely reason for the differences is the fact that our reservoir of control units is only coming from areas to the east of the railway line as there are partly no land-based control units to the west of the Konkan Railway given the proximity to the coastline of part of the railway line.

Figure 4: Baseline land cover shares by class



Note: Figures based on hybrid classification method for the baseline year 1991; Treatment and Control (Railway line) refer to the 4*4 km squares on the railway line and 20 km away from the railway line respectively; Treatment and Control (Railway station) refer to the circular buffer with 5 km radius around railway stations and around locations 20 km away from the railway station on the same latitude.

Table 1 presents baseline descriptive statistics for the census data, namely the mean and standard deviation for outcome and control variables for localities at different distance bands from the railway station. The last column of the table additionally shows the level of joint significance of the different distance band groups, when controlling for baseline controls. We apply standard significance tests throughout our analysis of census data even though it is based on a census and not a sample, so that inferential statistics and sampling theory are less applicable as we already have data for the whole relevant population (both geographically and temporally) at hand.

Most importantly, we find that the means at baseline are different for all outcomes but the urban population, with varying degrees of absolute differences for the individual outcomes. For the urban population indicator, it has to be noted that the sample is much lower (170 towns compared to 7500 villages underlying all other indicators) and may thus well be underpowered. Mean population outcomes decline as one moves further from the railway stations. The female-to-male sex ratio as another demographic outcome is lower for the distance bands closer to the station than further away, except for the distance band 30-100 km which has the lowest mean sex ratio among all distance bands.

The second set of outcomes relate to the workforce composition. The mean crude workforce participation rate is slightly higher for distance bands further away than those closer to the station. We further differentiate between male and female workers, main and marginal workers, i.e. those who work more or less than six months a year, respectively, and between four employment groups among main workers: cultivators, agricultural labourers, household industry workers, and other workers. Household industry workers refer to people engaged with other household members in activities run at home, such as food production or manufacture of wood products; other workers are neither cultivators, nor agricultural labourers, nor engaged in household industry, and thus comprise government servants, teachers, industrial workers, and business people, among others. There is a difference of 0.8, 3.6 and 2.9 percentage points between distance band 0-5 km and 30-100 km for the share of main workers, female workers and main cultivators, respectively. Main workers clearly dominate the workforce, accounting for 84 to 89 percent among all workers. Henceforth all references to cultivators, agricultural labour, household industry workers and other workers are to those who engage in the respective activity as "main" workers, that is for more than six months a year.

The third set of census outcomes reflect land use and land use changes. The mean forest area increases as one moves further away from the railway station. No such clear trend is visible for cultivated area. The mean cultivated area for the 0-5 km distance band is 211 hectares and for 30-100 km it is 210 hectares. Culturable wasteland in the table refers to land not cultivated in the previous five years. Lastly, the table features the control variables employed in our quantitative analyses. Predictably, elevation, distance from the coast and distance from the highway increases as one moves away from the railway station. All variables are defined in greater detail in the Appendix A.

Overall, the comparison of baseline characteristics indicates that – despite the careful selection of the comparison groups – baseline imbalances exist between the groups compared in the subsequent analysis. As noted earlier, these imbalances do not directly inform about the validity of the difference-in-differences framework. However, they need to be considered as potential confounders to the analysis.

Table 1: Baseline descriptive statistics for outcome and control variables

Outcome Variables	Mean (standard deviation), by distance band 0-5 km	5-10 km	10-20 km	20-30 km	30-100 km	Difference (*p*-value)
Population						
total	2205 (6157)	1474 (2543)	1171 (3206)	915 (1566)	797 (1675)	0.00
if rural	1575 (1896)	1299 (1391)	984 (1103)	813 (1019)	711 (878)	0.00
if urban	25475 (28063)	15087 (13553)	19108 (24459)	10403 (6991)	12326 (12015)	0.55
Females per 1000 males (sex ratio)	1130 (162)	1159 (162)	1201 (207)	1217 (255)	1040 (181)	0.00
Crude workforce participation rate (CWPR), in %	43.2 (10.6)	45.0 (10.3)	47.6 (9.6)	49.3 (10.3)	49.3 (10.8)	0.00
Main workers among workers, in %	87.9 (15.0)	86.3 (15.3)	86.1 (15.6)	84.3 (16.9)	88.6 (15.3)	0.00
Female workers among workers, in %	43.4 (15.0)	46.4 (15.1)	50.1 (14.2)	46.4 (15.1)	39.8 (15.9)	0.00
Shares among main workers, in %						
main cultivators	51.8 (29.5)	58.4 (27.6)	65.6 (26.2)	64.5 (27.6)	54.7 (27.4)	0.00
main agricultural labourers	14.0 (14.1)	14.1 (14.7)	13.1 (15.0)	14.1 (15.8)	26.0 (21.3)	0.00
main household industry workers	1.8 (3.9)	1.6 (4.2)	1.1 (2.4)	0.9 (2.8)	1.1 (2.8)	0.00
main other workers	32.4 (27.0)	25.9 (23.9)	20.3 (21.9)	20.4 (23.3)	18.3 (23.0)	0.00
Forest land, in hectares	66.6 (196.1)	94.3 (395.0)	163.1 (572.4)	285.6 (710.9)	290.9 (628.7)	0.00
Cultivated area, in hectares	210.9 (207.7)	227.9 (253.6)	212.8 (267.3)	197.9 (256.0)	209.6 (289.4)	0.00
Non-agricultural area, in hectares	112.2 (153.6)	125.4 (151.2)	164.3 (286.3)	170.7 (407.3)	85.72 (171.7)	0.00
Culturable wasteland, in hectares	112.3 (156.1)	130.7 (186.4)	208.3 (327.5)	198.7 (344.7)	84.66 (184.7)	0.00
Control Variables						
Elevation at the centroid of the locality, in metres	67.1 (73.9)	83.9 (79.9)	155.3 (156.4)	306.7 (306.8)	630.6 (184.8)	0.00
Latitude at the centroid of the locality, in degrees	16.1 (1.6)	16.4 (1.5)	16.7 (1.3)	16.7 (1.4)	15.3 (1.5)	0.00
Distance from the coast, in kms	15.5 (10.9)	18.6 (11.3)	22.2 (14.4)	24.7 (20.5)	62.5 (19.1)	0.00
Distance from the nearest highway, in kms	3.8 (5.1)	6.2 (5.8)	13.6 (6.7)	23.0 (7.0)	53.7 (20.1)	0.00
Number of observations	720	1184	1360	948	3477	

Note: See Appendix A for a complete definition of outcome variables. Baseline year is 1991. *p*-values refer to *F*-tests on the joint significance of the distance band group coefficients in an ANOVA with the respective baseline outcome on the left-hand side, when controlling for baseline controls. Data on forest land, cultivated area, area not under agricultural uses and culturable wasteland is recorded only for villages. The CWRP represents the total number of workers as a percentage of the total population. "Main" refers to workers who work in the respective sector for more than six months in a year.

Source: Authors' calculation based on Census of India

4.3. Land use change analysis using satellite imagery

This section provides a macro-picture of the environmental impacts of the railway infrastructure intervention. In Table 2, we present summary results for the difference-in-differences analysis on satellite data where the outcome variable is the land use change along railway lines expressed as percentage of the 4*4 km square area. Land use within the village boundaries using census data will be assessed in the following section.

Table 2: Land use change impacts

	Open Space & Built-up Area	Forest Cover	Water Body	Agricultural Land
post	0.059***	0.041	0.003	-0.103***
	(0.013)	(0.025)	(0.005)	(0.019)
treatment	0.038***	-0.273***	0.009*	0.227***
	(0.013)	(0.025)	(0.005)	(0.019)
post x treatment	0.114***	0.027	-0.002	-0.140***
	(0.019)	(0.035)	(0.008)	(0.027)
Control group baseline mean	0.207	0.479	0.020	0.293
R-squared	0.33	0.28	0.01	0.35
Number of observations	600	600	600	600

Note: Results are based on difference-in-differences estimations according to Equation [1] as set out in Section 3.1.1; standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

When comparing the development between 1991 and 2014 of treated regions and comparison regions 20 kms away from the railway line, open space and built-up area have increased by 11.4 percentage points, with a baseline mean of 20.7 percent. This extension was mainly at the expense of agricultural area, which declined by 14 percentage points (baseline comparison group mean of 29.3 percent). Estimates are insignificant for forest cover and, for water bodies as well. The sign, level and significance of the results do not change if we use circular buffers of 5 km radius around the railway stations as treatment and comparison locations 20 kms away (see Appendix B). When we use comparison locations 10 km away from the railway line, we find that the open space category increased by 9.5 percentage points again at the expense of agricultural area that declined by 7.2 percentage points (see again Appendix B). These results may plausibly be attributed to an increase in non-farm activities around the railway stations.

The results for forest cover along the railway line indicate that the Konkan Railway has not led to differential deforestation in the aggregate. Further, between 1991 to 2014, forest cover increased for both the treatment region and the control region 20kms away. For the treatment region it increased from 728 sq kms to 832 sq kms and for the control from 1406 sq km to 1441 sq kms. This result is important, as the establishment and maintenance of physical infrastructure such as a railway line cause immediate loss of vegetation cover where the infrastructure is constructed. One of the major environmental objections against the Konkan Railway was that it would cause large-scale loss of forests cover, a concern raised very often regarding the environmental impact of similar projects in general and for proposed projects in the Western Ghats in particular (Gadgil et al. 2011; Ramachandra, Bharath, and Vinay 2019). The results suggest that the initial loss of forest cover due to the engineering interventions may have been compensated by afforestation activities from KRCL. Over the years, KRCL undertook geo-safety work, plantation of tree saplings and vetveri grass in the land adjoining the track and the railway stations in its effort to stabilize the track and to safeguard the environment. In its Annual Report of 2006-2007, for example, KRCL reported that it had planted around 0.55 million trees along the Konkan Railway alignment.[6]

4.4. Impact analysis using census data

4.4.1. Impacts on population and workforce composition

In our main socio-economic analysis, we look at population and workforce composition using a modified difference-in-differences approach with two definitions of the access-to-infrastructure treatment outlined in Section 3.1.2, the binary and the categorical treatment definition. The results are presented in Figure 5.

The first two graphs (Panel A and B) depict changes in rural and urban population, each corresponding to percentage differences as we defined population in logarithmic terms. Results for total population are basically identical to the results for rural population and therefore not shown. When compared to the respective comparison group, rural population increases between 1991 and 2011 by about 10 percent (significant at 1 percent level) for T_{5km} and T_{10km}, i.e. the binary definitions of treatment referring to localities in a radius of 5 and 10 km, and for the distance band 0-5 km. To the contrary, the distance bands 10-20 km and 20-30 km show (borderline) significantly negative coefficients, which possibly reflects population reallocations in the vicinity of railway stations. Urban population shows a broadly similar picture, where the lack of detection of a significant effect (except for the positive T_{5km} coefficient) may be explained by the low statistical power. Regarding the differences between the binary and categorical treatment definitions, the larger coefficient for the binary treatment, T_{10km}, as compared to the distance band 5-10 km, for example, has to do with the fact that the comparison group in the binary case also includes the distance bands between 10 and 30 km, which have negative coefficients and which are defined as treatment groups in the categorical treatment case. This suggests a non-linear relationship between distance and impact, which will be analysed in more detail in Section 4.4.3 below. For the third population-related indicator, the female-to-male sex ratio, we find a similar pattern, with lower and less significant coefficient values. More specifically, we find that the sex ratio rose only for T_{10km}, by an average of 20 females in the treated localities as compared to the comparison ones, while the baseline mean for the comparison localities is already high at 1107 (the results being significant at the 1 percent level). Mainly due to male-dominated migration, the study area has a historically high female-to-male sex ratio, which comes to a ratio of over 1000 females per 1000 males in 1991 (Table 1), when the average for India was 927 females per 1000 males. The sex ratio remained above 1090 in two districts of the Konkan region for the entire 20th century, for example (Tumbe 2012).[7]

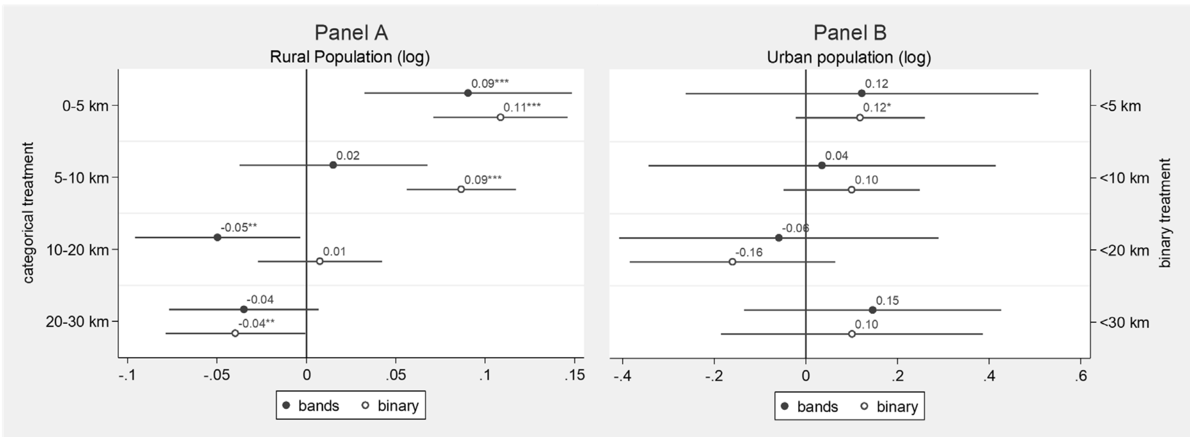
[6] These efforts were also in response to the fact that Konkan Railway falls on the windward side of the Western Ghats which experiences heavy rainfall in the monsoon season averaging around 3000 mm to 3900 mm per annum. Moreover, a major part of the track passes through lateritic terrain. Cuttings through the porous and soft lateritic terrain to make way for the railway line made the Konkan Railway vulnerable to accidents in early years of operation due to rainfall-induced landslides. Accidents due to natural causes have come down drastically in recent years as a result of the engineering solutions and environmental management practices adopted (Garg, Naswa, and Shukla 2013).

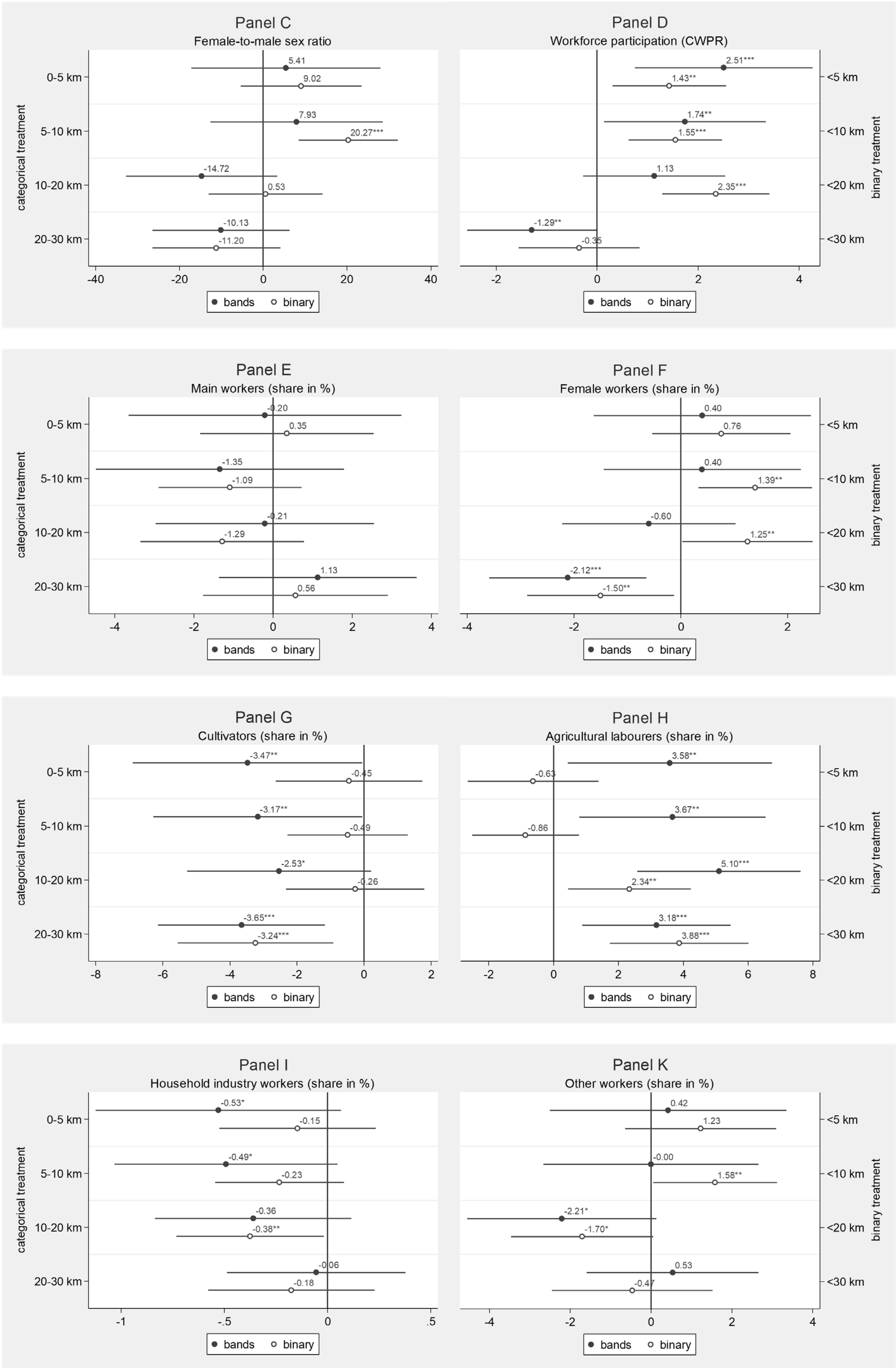
[7] District-level aggregate sex ratios mainly vary because of sex-selective migration, in addition to the prevalent phenomena of sex-selective abortion and sex differentials in mortality, especially during childhood (Kundu and Sahu 1991; Tumbe 2015).

The most consistently significant pattern is found for the increase in the crude workforce participation rate (CWPR, Panel D) closer to the railway stations, namely up to distances of 20 kilometres. The next two indicators inform about compositional changes within the workforce: first, the percentage of main workers (and thus also of marginal workers) to total workers remains largely unaffected. Second, the percentage of female workers to total workers rather increased at the detriment of male workers for the treated localities as compared to comparison ones over the observation period. The share of female workers increased by 1.4 and 1.3 percentage points for T_{10km} and T_{20km}, respectively (significant at 5 percent level). Again, the sign changes once one moves to the most distant binary treatment definition, T_{30km}. Here, also the coefficient for the distance band (20-30 km) is significant.

The share of cultivators among workers (Panel G) decreases significantly by 2.5 to 3.7 percentage points for all distance bands, whereas estimates are close to zero for all binary treatment definitions except for T_{30km}. For agricultural labourers, numbers are mostly the opposite, with positive changes between 3.2 and 5.1 percentage points for all distance bands. The simultaneous decline in cultivators and increase in agriculture labour may occur due to a combination of factors. One factor contributing to the decline in cultivators may be the move away from agriculture to other productive activities with the expansion of rural non-farm activities and access to non-farm employment outside the place of residence. As the decline in cultivators for our treated localities is accompanied by a rise in agricultural labour, another plausible explanation is that the residents and migrants who own farming land in the village choose to informally lease out land to landless agricultural labourers under various forms of tenancy. In the Konkan region, this phenomenon is exacerbated by the fragmentation and sub-division of landholdings and increasing cost of hiring agricultural labour (Tumbe 2012, 2015). In fact, some key informants from rural self-government bodies highlighted that landowners from the village residing in Mumbai either lease out their land to landless agricultural labourers or hand over the supervision to their relatives who in turn lease it out for cultivation on a sharecropping arrangement.

Figure 5: Impacts on population and workforce composition outcomes





Note: The panels depict point estimates and 95% confidence intervals according to the two variants of Equation [2] outlined in Section 3.1.2. *** p<0.01, ** p<0.05, * p<0.1

When it comes to other occupational categories for main workers, we find that household industry workers as percentage of total workers tend to decline, but at levels below 0.5 percentage points; no notable pattern emerges for other workers (Panel I and K).

Bringing the different results together, we find limited evidence that the railway infrastructure provided stimuli towards local employment generation. While the overall workforce participation rate increased, which is indicative of improved economic conditions, non-agricultural jobs did not increase disproportionately. This is consistent with the results on intermediate freight outcomes in Section 4.1. The increase in female-to-male sex ratio and the decrease in the share of male workers in closer proximity to the railway stations further suggests that strengthening the transport system might result in more outmigration, if it is not accompanied by other inputs to create more avenues for employment in the place of residence.

The key informant interviews provide some supporting testimonies that the Konkan Railway has made it easier for migrating male household members to retain ties with their native places and to function more easily from more than one home. Whereas earlier migrants would visit their native villages in Konkan only once or twice a year, the convenience, comfort and lower fares of the Railway has made possible more frequent trips, especially to and from Mumbai.[8] In addition to Mumbai, which continued to exert a strong metropolitan pull on the Konkan region, new destinations such as Surat (in Gujarat), Goa and Raigarh (in Maharashtra) emerged as popular destinations to migrate for work. We further learnt from the interviews that there are a large number of frequent short distance travellers and daily commuters who use the Konkan Railway for more local, inter-district travel for work. For example, a considerable number of people in the southern sub-districts of neighbouring Sindhudurg in Maharashtra commute daily to Goa to work in pharmaceutical factories, industrial estates, as hawkers or street vendors in order to avoid the higher cost of living in Goa. One may therefore conclude that the Konkan Railway has not only amplified the already existing pattern of travel in the region to Mumbai, but also enhanced the available choice of destinations for seeking employment by connecting the region internally and to the rest of the country.

These findings are in line with the larger theoretical and empirical research on how transport connectivity affects commuting and migration decisions. Deshingkar and Grimm (2005) have noted that a "major hindrance in gaining more fully from the many different possibilities for livelihood enhancement through migration is the difficulty in maintaining social and financial links" and that "better infrastructure […] can help to keep social links alive" (p.54). Sorek's (2009) spatial model is one way to explain the impact of transport time and cost on inter-city migration and commute between two regions with different productivity levels. The analysis characterizes an equilibrium pattern under which the ablest workers of the Suburb (low productivity city) migrate to the Metropolis (high productivity city), those with moderate ability commute to work in the Metropolis, and the least able work in the Suburb. Empirically, our results contrast with estimates

[8] In this regard, an illustrative quote from the interviews is the following: "Even before the Konkan Railway, people from the village were settled in Mumbai and would keep coming back to their villages, especially during festivals such as Ganpati and Diwali. After the railway, the people are making more trips to their villages. Earlier they used to come once every six months. Now they come fortnightly and sometimes every weekend" (Sarpanch, Warak Gram Panchayat, District Raigarh).

of infrastructure effects in China referred to earlier where the government has had a long-standing policy of restricting migration from rural to urban areas. For India, Sharma and Chandrasekhar (2014) estimate commuting flows in the year 2009–10 and identify rural unemployment rates, wage differentials, and cost of living as important determinants of commuting. For the two Indian states of Andhra Pradesh and Madhya Pradesh, Deshingkar (2010) finds that belonging to a remote village in a poor region decreases the probability of commuting and increases the probability of migrating, whereas belonging to a well-connected village in a prosperous region increases the likelihood of commuting. The authors attribute the strong increase in commuters in one of the two states, Andhra Pradesh, to the growing road network, improved communications and rapid rate of urbanization, especially the growth of small towns. The other state, Madhya Pradesh, is comparatively backward and has a very poor road network and lower levels of urbanization, which make it more difficult to commute so that people tend to migrate on a short-term basis to large cities. Outside of India, Bah et al. (2003) find that commuting to the regional urban centres of Aba and Port Harcourt in south-eastern Nigeria is encouraged by the efficient and cheap state-subsidized transport system. The deterring effect of costs of living in urban areas on migration from rural to urban areas (instead of commuting) is also found by Baker (1995) in North-West Tanzania.

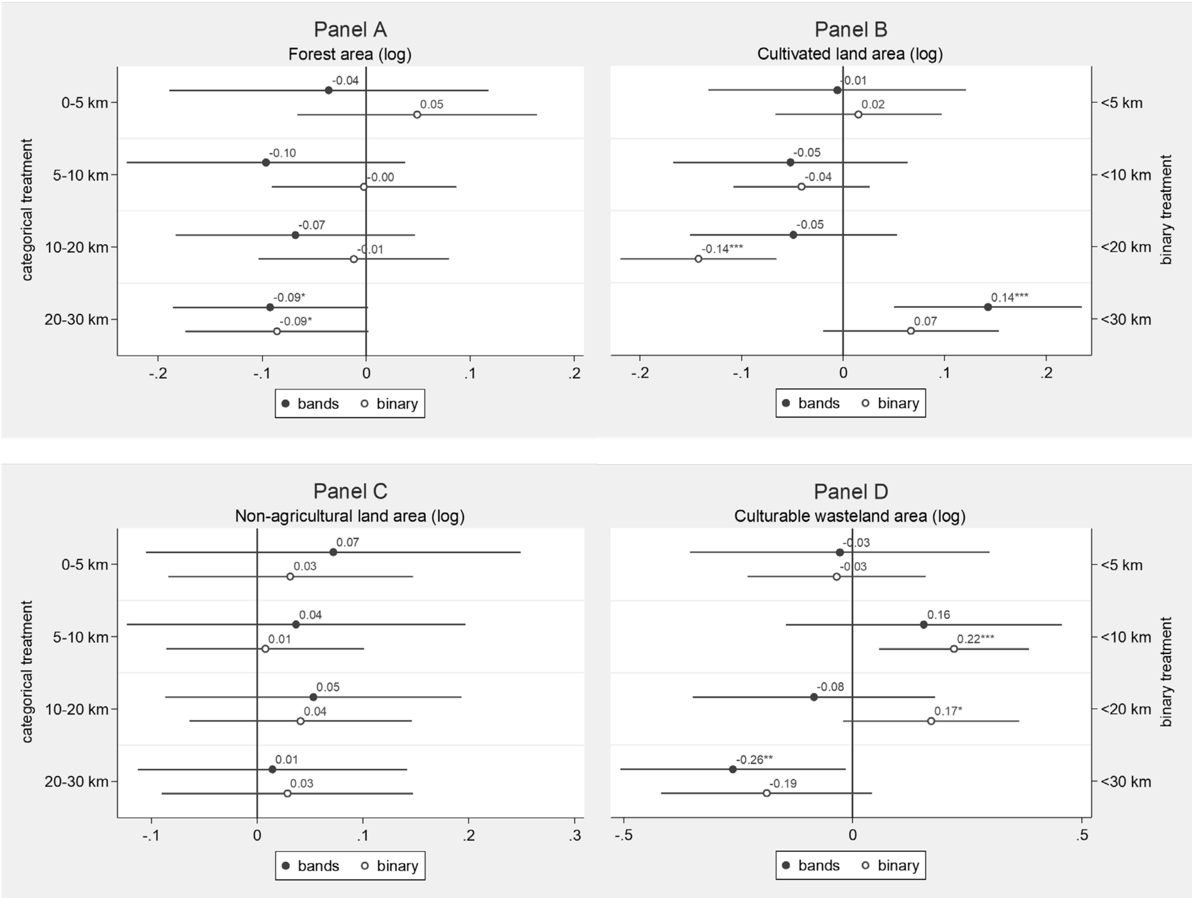
4.4.2. Impacts on land use

The census data allows to distinguish the four land-use categories in Figure 6. Coefficients reflect percentage differences as we defined all area values in logarithmic terms. Between 1998-2010, the impact on forest land seems to be close to zero within locality boundaries close to the railway stations. For villages that are still within a 30 km radius but further away, the impact seems to be rather negative. For T_{30km} and the distance band 20-30 km, the estimate of the forest area decline is significant at the 10 percent level with -9 percent. All other coefficients are insignificant.

For the results of the agricultural land-use change within locality boundaries one finds indications for a decline in the corridor between 5 and 20 km (though only the 10-20 km distance band coefficient is significant), but the estimate for the area adjacent to the railway line is essentially zero and the coefficient for T_{30km} is highly positively significant. The results tend to corroborate the findings from workforce composition where a decline in cultivators was compensated by an increase in agricultural labourers.

Area not under agricultural use in the census dataset for villages, which includes all land occupied by buildings, roads and railways or under water and other land put to uses other than agriculture. Here, all coefficients are positive but imprecisely estimated, so that they are insignificant across all treatment definitions. The picture for culturable wasteland, which is land not cultivated in the previous five years, is very mixed. For one of the four bands, we find near-zero effects, for one we find different signs across the two treatment definitions, and for the other we find opposite signs with partly significant estimates.

Figure 6: Impacts on land use outcomes



Note: The panels depict point estimates and 95% confidence intervals according to the two variants of Equation [2] outlined in Section 3.1.2. *** p<0.01, ** p<0.05, * p<0.1

4.4.3. Robustness checks

In all the earlier difference-in-differences specifications, we controlled for the distance from the highway by interacting it with a post-treatment dummy. In order to test for heterogeneous effects of railway access depending on the simultaneous closeness to the nearby highway, we now interact different highway distance bands with a binary treatment indicator, for which we choose T_{10km}. Estimates therefore have to be compared to those for the T_{10km} treatment definition in the previous sub-sections. Results suggest that impacts are differ across distance bands, but that they are not primarily driven by simultaneous closeness to highways (Appendix B). This analysis, however, has its limitations in that 89 percent of localities in 10 km vicinity to the railway stations are also in 10 km vicinity to the highway. Accordingly, there is relatively little heterogeneity that can be exploited in the estimations.

In a second robustness check, we use distance to railway stations directly in our estimations instead converting it into either a dichotomous or categorical treatment variable. This also allows for potentially non-linear relationships between treatment intensity and outcomes (see Equation [4], section 3.1). The results are again provided in Appendix B. The interpretation of results is less straightforward than for the previous analyses as the coefficient of the simple distance and the quadratic distance have to be jointly considered. Of the fourteen outcome indicators, six show a significant coefficient for the quadratic term. These are rural population, share of cultivators, household industry workers, other workers, forest land and non-agricultural area. A priori, one expects the coefficient of the simple distance to be the opposite of the coefficients for the binary treatment indicators: as treatment is defined by the closeness to Konkan Railway stations, a positive coefficient for a binary treatment indicator implies a negative coefficient for the distance variable. Apart from rural and urban population and share of, cultivated area, this is actually the case for all indicators when considering the binary treatment indicator T_{10km}. Six of these variables show significant coefficients: workforce participation rate, the share of cultivators, other workers, forest land, non-agricultural area and culturable waste land.

5. Conclusion and policy implications

This study used a combination of key informant interviews, passenger and freight performance analysis and quasi-experimental research design to assess the effects of new railway infrastructure in India. It provides evidence of the heterogeneous and non-linear impacts on population composition, sectoral workforce composition and land use change. The impact of the railway line, Konkan Railway, has largely been felt via improved connectivity for passengers and less in terms of freight volume or growth of industries. This, in turn, seems to have reinforced pre-existing patterns of male outmigration from the region and shifts within the agricultural sector. Our study thereby assembled suggestive evidence that supports claims from the existing migration literature highlighting the importance of transport network in enabling workers to respond better to income opportunities generated by migration. We furthermore tentatively interpret an increase in the workforce participation rate and population in localities close to railway stations together as indications for modest indirect and wider economic effects of the presence of railway stations. Deeper analyses of economic survey data may shed more light on supply-side aspects of outcomes regarding the labour market and economic activities.

We had both satellite and census data at hand to assess effects on land use. Both data sources suggest that the impact on forest land is close to zero for localities close to the railway stations. For satellite data this may be explained by afforestation activities undertaken by KRCL along the tracks and stations. These results provide lessons for environmental management for linear infrastructure projects as well as for planning railway investments in ecologically sensitive regions. Several large infrastructure projects are planned or are being built across the world, especially in developing countries to deliver various sustainable development objectives. At the same time, the role of these megaprojects in delivering these objectives is being increasingly contested. An important lesson from our analysis is that large infrastructure projects need to be appraised on a case by case basis, as the large infrastructure that causes disturbances in the ecosystem of a region caused in the short run may be compensated in the long run if adequate measures for environmental management are implemented. Future research may build on the macro-picture provided in this studyto contextualize more localized effects related to forest fragmentation, forest health, on drainage pattern in particular locations, for example.

While several measures have been taken to minimize bias in our impact analysis, our study is subject to two caveats: first, the Konkan region is also crossed by a national highway that has been built three decades before the Konkan Railway. We controlled for the distance to this highway in our analysis and ran supporting robustness checks, but clear isolation of the individual effects may remain elusive, also because of interactions between the two infrastructures. However, in the real-world scenario of infrastructure interventions, it is also helpful to tease out "mediation mechanisms when the project has purposeful selection" (Hansen, Andersen, and White 2011:7). In our case, policymakers will find it useful to focus on devising ways to make the railway and the highway work in more complementary ways, both for passenger and freight transport. Second, interactions may have taken place across locations with different distances to the railway, including spillovers from the treatment to the control groups of our evaluation designs. For example, observed effects on population may partly be driven by migration from areas falling in the comparison group to areas falling in the treatment group. The evidence on displacement effects of railway access is mixed. While Braun and Franke (2019) find little evidence that the effects of the railway on population growth was driven by localised displacement effects, Berger and Enflo (2017) and Büchel and Kyburz (2020) find that the population growth in connected towns and municipalities respectively came at the expense of near-by unconnected localities.

Despite these caveats and qualifications, our analysis contributes to understanding impacts of infrastructure projects in regions prone to outmigration. It raises pertinent questions such as the implications of the improved transport network for circular migration, for daily commuting to work, geographic labour mobility and labour market outcomes for sending and receiving region. The literature on migration studies has gone beyond understanding migration as distress related to negative shocks such as natural calamities to viewing migration also as an employment and income augmenting strategy that is being actively sought out by workers (Bird and Deshingkar 2009; Deshingkar and Anderson 2004; Deshingkar and Grimm 2004). As workers are attracted to higher wages, among other factors to migrate, it is not clear how successful rural development policies which implicitly aim to generate employment in the rural area will be able to reverse this trend (Deshingkar and Anderson 2004). In the interim, it would be advisable to understand ways in which workers can take advantage of improved access to work opportunities outside the place of residence and how transport policy can be formulated to facilitate this.

References

Aiyar, Swaminathan S. Anklesaria, and Neeraj Kaushal. 2019. "Are Resettled Oustees from the Sardar Sarovar Dam Project 'Better Off' Today?" *Economic and Political Weekly* Vol. 54(Issue No. 12).

Anderson, James Richard, Ernest E. Hardy, John T. Roach, and Richard E. Witmer. 1976. *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*. Vol. 964. US Government Printing Office.

Anon. 1994. "Agitation Resurrected." *Down to Earth*, February 28.

Anon. 2013. *Sterlite Industries (I) Ltd. Etc. ... vs Union Of India And Ors. Etc. Etc on 2 April, 2013*.

Arun, S. 2011. "Apex Court Extends Mining Ban to 2 More Karnataka Districts; Sesa Goa, JSW Steel to Feel the Impact." *The Hindu Business Line*, August 26.

Atack, Jeremy, Fred Bateman, Michael Haines, and Robert A. Margo. 2010. "Did Railroads Induce or Follow Economic Growth?" *Social Science History* 34(2):171–97.

Babu, Hemant. 1997. "Dreams Ride on the Konkan Railway." *Indian Express*, October 11.

Bah, Mahmoud, Salmana CissÈ, Bitrina Diyamett, Gouro Diallo, Fred Lerise, David Okali, Enoch Okpara, Janice Olawoye, and Cecilia Tacoli. 2003. "Changing Rural–Urban Linkages in Mali, Nigeria and Tanzania." *Environment and Urbanization* 15(1):13–24. doi: 10.1177/095624780301500104.

Baker, Jonathan. 1995. "Survival and Accumulation Strategies at the Rural-Urban Interface in North-West Tanzania." *Environment and Urbanization* 7(1):117–32.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian. 2020. "On the Road: Access to Transportation Infrastructure and Economic Growth in China." *Journal of Development Economics* 145:102442. doi: 10.1016/j.jdeveco.2020.102442.

Baum-Snow, Nathaniel, Loren Brandt, J. Vernon Henderson, Matthew A. Turner, and Qinghua Zhang. 2017. "Roads, Railroads, and Decentralization of Chinese Cities." *The Review of Economics and Statistics* 99(3):435–48. doi: 10.1162/REST_a_00660.

Berger, Thor, and Kerstin Enflo. 2017. "Locomotives of Local Growth: The Short- and Long-Term Impact of Railroads in Sweden." *Urbanization in Developing Countries: Past and Present* 98:124–38. doi: 10.1016/j.jue.2015.09.001.

Bhuyan, Pranab Jyoti. 2016. "Mumbai-Goa Highway Sees over 1,000 Deaths a Year." *Hindustan Times*, May 16.

Bird, K., and P. Deshingkar. 2009. "Circular Migration in India: Policy Brief No. 4." *Prepared for the World Development Report.*

Braun, Sebastian, and Richard Franke. 2019. *Railways, Growth, and Industrialisation in a Developing German Economy, 1829-1910. MPRA Paper.* University Library of Munich, Germany.

Büchel, Konstantin, and Stephan Kyburz. 2020. "Fast Track to Growth? Railway Access, Population Growth and Local Displacement in 19th Century Switzerland." *Journal of Economic Geography* 20(1):155–95. doi: 10.1093/jeg/lby046.

Cantarelli, Chantal, Bent Flyvbjerg, EJE Molin, and B. van Wee. 2010. "Cost Overruns in Large-Scale Transportation Infrastructure Projects: Explanations and Their Theoretical Embeddedness." *European Journal of Transport Infrastructure Research* 10(1):5–18.

Chakravorty, Sanjoy. 2003. "Industrial Location in Post-Reform India: Patterns of Inter-Regional Divergence and Intra-Regional Convergence." *Journal of Development Studies* 40(2):120–52. doi: 10.1080/00220380412331293797.

Chen, Amy Z., Jeremy Fischer, Andrew Fraker, Neil Buddy Shah, Stuart Shirrell, and Daniel Stein. 2017. "Welfare Impacts of an Entry-Level Solar Home System in Uganda." *Journal of Development Effectiveness* 9(2):277–94. doi: 10.1080/19439342.2017.1307248.

Datta, Saugato. 2012. "The Impact of Improved Highways on Indian Firms." *Journal of Development Economics* 99(1):46–57.

Deshingkar, Priya. 2010. "Migration, Remote Rural Areas and Chronic Poverty in India." *Overseas Development Institute (ODI) Working Paper* 323.

Deshingkar, Priya, and Edward Anderson. 2004. *People on the Move: New Policy Challenges for Increasingly Mobile Populations*. Overseas Development Institute.

Deshingkar, Priya, and Sven Grimm. 2004. "Voluntary Internal Migration: An Update." *London: Overseas Development Institute* 44:4.

Deshingkar, Priya, and Sven Grimm. 2005. *Internal Migration and Development: A Global Perspective*. United Nations Publications.

Donaldson, Dave. 2018. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." *American Economic Review* 108(4–5):899–934. doi: 10.1257/aer.20101199.

Donaldson, Dave, and Richard Hornbeck. 2016. "Railroads and American Economic Growth: A 'Market Access' Approach." *The Quarterly Journal of Economics* 131(2):799–858.

Drèze, Jean, Meera Samson, and Satyajit Singh, eds. 1997. *The Dam and the Nation: Displacement and Resettlement in the Narmada Valley*. New Delhi: Oxford University Press.

D'Souza, Rohan, Pranab Mukhopadhyay, and Ashish Kothari. 1998. "Re-Evaluating Multi-Purpose River Valley Projects: A Case Study of Hirakud, Ukai and IGNP." *Economic and Political Weekly* 297–302.

Duflo, Esther, and Rohini Pande. 2007. "Dams." *The Quarterly Journal of Economics* 122(2):601–46.

Esteban-Oliver, Guillermo. 2020. *The Curse of Geography? Railways and Growth in Spain 1877-1930*. Asociación Española de Historia Económica.

Fernandes, Allywn. 1993. "A West Coast Dream Takes Shape." *Times of India*, July 21, 9.

Flyvbjerg, Bent. 2007. "Policy and Planning for Large-Infrastructure Projects: Problems, Causes, Cures." *Environment and Planning B: Planning and Design* 34(4):578–97. doi: 10.1068/b32111.

Gadgil, Madhav, B. Krishnan, V. Vijayan, R. Borges, R. Sukumar, and G. Subrahmanyam. 2011. *Report of the Western Ghats Ecology Expert Panel: Part 1*. Ministry of environment and forests ,Government of India.

Garg, Amit, Prakriti Naswa, and PR Shukla. 2013. *Impact Assessment and Management Framework for Infrastructure Assets; a Case Study of Konkan Railways*. UNEP Risø Centre.

Ghani, Ejaz, Arti Grover Goswami, and William R. Kerr. 2016. "Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing." *The Economic Journal* 126(591):317–57.

Goldblatt, Ran, Kilian Heilmann, and Yonatan Vaizman. 2020. "Can Medium-Resolution Satellite Imagery Measure Economic Activity at Small Geographies? Evidence from Landsat in Vietnam." *The World Bank Economic Review* 34(3):635–53. doi: 10.1093/wber/lhz001.

Hansen, Henrik, Ole Winckler Andersen, and Howard White. 2011. "Impact Evaluation of Infrastructure Interventions." *Journal of Development Effectiveness* 3(1):1–8. doi: 10.1080/19439342.2011.547659.

Hidrobo, Melissa, Amber Peterman, and Lori Heise. 2016. "The Effect of Cash, Vouchers, and Food Transfers on Intimate Partner Violence: Evidence from a Randomized Experiment in Northern Ecuador." *American Economic Journal: Applied Economics* 8(3):284–303. doi: 10.1257/app.20150048.

Hornbeck, Richard, and Martin Rotemberg. 2021. *Railroads, Market Access, and Aggregate Productivity Growth*. University of Chicago Booth School of Business, mimeo.

Hugo, Graeme. 2003. "Circular Migration: Keeping Development Rolling?" *Https://Www.Migrationpolicy.Org*. Retrieved December 26, 2019 (https://www.migrationpolicy.org/article/circular-migration-keeping-development-rolling).

Iversen, Vegard, and Yashodhan Ghorpade. 2011. "Misfortune, Misfits and What the City Gave and Took: The Stories of South-Indian Child Labour Migrants 1935–2005." *Modern Asian Studies* 45(05):1177–1226. doi: 10.1017/S0026749X10000065.

Jaiswal, Sreeja, and Gunther Bensch. 2018. "A Methodological Framework to Address Gaps in the Evidence on Infrastructure Impacts: The Case of an Indian Railway Project Evaluation." *Journal of Economic Methodology* 1–13. doi: 10.1080/1350178X.2018.1561073.

Jalan, Jyotsna, and Martin Ravallion. 1998. "Are There Dynamic Gains from a Poor-Area Development Program?" *Journal of Public Economics* 67(1):65–85.

Jedwab, Remi, and Alexander Moradi. 2016. "The Permanent Effects of Transportation Revolutions in Poor Countries: Evidence from Africa." *Review of Economics and Statistics* 98(2):268–84.

Kasturirangan, K., C. R. Babu, J. M. Mauskar, Kanchan Chopra, J. Kishwan, D. Shankar, S. Narain, P. S. Roy, A. Tyagi, and I. Chandrasekharan. 2013. *Report of the High Level Working Group on Western Ghats*. Ministry of environment and forests ,Government of India.

Konkan Railway Corporation. 1998. *Annual Report 1997-98*. KRCL.

Konkan Railway Corporation Limited. 1999. *A Treatise on Konkan Railway*. Konkan Railway Corporation Limited.

Kothari, Ashish. 1984. "Narmada Valley Project-Development or Destruction." *Economic and Political Weekly* Vol. 19(Issue No. 22-23).

Kundu, Amitabh, and Mahesh K. Sahu. 1991. "Variation in Sex Ratio-Development Implications." *Economic and Political Weekly* Vol. 26(Issue No. 41).

McKenzie, David. 2012. "Beyond Baseline and Follow-up: The Case for More T in Experiments." *Journal of Development Economics* 99(2):210–21. doi: 10.1016/j.jdeveco.2012.01.002.

Ozesmi, Stacy L., and Marvin E. Bauer. 2002. "Satellite Remote Sensing of Wetlands." *Wetlands Ecology and Management* 10(5):381–402. doi: 10.1023/A:1020908432489.

PTI. 2013. "Mining Ban Hurting Economy, Exports: Anand Sharma." *The Hindu*, December 25.

PTI. 2014. "Economic Survey 2014: Iron Ore Mining Ban Led to Negative Growth in Cargo Traffic - The Economic Times." *The Economic Times*, July 9.

Ramachandra, T. V., Setturu Bharath, and S. Vinay. 2019. "Visualisation of Impacts Due to the Proposed Developmental Projects in the Ecologically Fragile Regions- Kodagu District, Karnataka." *Progress in Disaster Science* 3:100038. doi: 10.1016/j.pdisas.2019.100038.

Rangarajan, Mahesh. 2007. *Environmental Issues in India: A Reader*. Pearson Education India.

Rao, K. Raghavendra. 2016. "Konkan Railway to Start Track-Doubling, Electrification Work after the Monsoon." *The Hindu Business Line*, August 29.

Sader, Steven A., Douglas Ahl, and Wen-Shu Liou. 1995. "Accuracy of Landsat-TM and GIS Rule-Based Methods for Forest Wetland Classification in Maine." *Remote Sensing of Environment* 53(3):133–44. doi: 10.1016/0034-4257(95)00085-F.

Sardesai, Rajdeep, and Ashley D'Mello. 1992. "The Missing Link." *Times of India*, May 3, 11.

Sequeira, S. 2013. *Transport Infrastructure and Firm Performance: Evidence from Southern Africa*. International Growth Centre Working Paper.

Sharma, Ajay, and S. Chandrasekhar. 2014. "Growth of the Urban Shadow, Spatial Distribution of Economic Activities, and Commuting by Workers in Rural and Urban India." *World Development* 61:154–66. doi: 10.1016/j.worlddev.2014.04.003.

Shrivastava, (Rahul. 1994. "Copper Unit Comes a Cropper." *Down to Earth*, May 15.

Shubhashish. 2012. "Mining Ban Hits Sesa Goa Hard, Sales down 86%." *Business Standard India*, October 9.

Sita, K., and K. Seeta Prabhu. 1989. "Levels of Development and Migration: Case of South Konkan." *Economic and Political Weekly* 39–41.

Sorek, Gilad. 2009. "Migration Costs, Commuting Costs and Intercity Population Sorting." *Regional Science and Urban Economics* 39(4):377–85. doi: 10.1016/j.regsciurbeco.2009.03.003.

Thorat, VA, JS Dhekale, HK Patil, and SN Tilekar. 2011. "Determinants of Rural-Urban Migration in Konkan Region of Maharashtra." *Agricultural Economics Research Review* 24:503–9.

Times News Network. 2016. "Konkan Railway Rolls out Plan for Capacity Doubling - Times of India." *The Times of India*. Retrieved November 29, 2017 (http://timesofindia.indiatimes.com/city/mangaluru/Konkan-Railway-rolls-out-plan-for-capacity-doubling/articleshow/56145576.cms).

Tumbe, Chinmay. 2012. "Migration Persistence across Twentieth Century India." *Migration and Development* 1(1):87–112. doi: 10.1080/21632324.2012.716225.

Tumbe, Chinmay. 2015. "Missing Men, Migration and Labour Markets: Evidence from India." *The Indian Journal of Labour Economics* 58(2):245–67. doi: 10.1007/s41027-016-0017-4.

Upadhyay, Anindya. 2014. "Indian Economy Lost Rs 50,000 Crore on Iron Ore Mining Ban - The Economic Times." *The Economic Times*, October 28.

Vartak, Kalyani, Chinmay Tumbe, and Amita Bhide. 2018. "Mass Migration from Rural India: A Restudy of Kunkeri Village in Konkan, Maharashtra, 1961–1987–2017." *Journal of Interdisciplinary Economics* 0260107918776563.

Wang, Yang, and Binzhen Wu. 2015. "Railways and the Local Economy: Evidence from Qingzang Railway." *Economic Development and Cultural Change* 63(3):551–88. doi: 10.1086/680091.

Zhou, Yixing, and John Logan. 2007. "Growth on the Edge: The New Chinese Metropolis." Pp. 140–60 in *Urban China in Transition*, edited by J. Logan. London: Blackwell.

Appendix

Appendix A: Outcome variable definitions

Outcome	Outcome Variable Definition	Availability
Workforce Composition	$Crude\ Workforce\ Participation\ Rate = \frac{Total\ Workers}{Total\ Population} * 100$	Rural & Urban
	$Main\ Workers = \frac{Total\ Main\ Workers}{Total\ Workers} * 100$	Rural & Urban
	$Cultivators = \frac{Main\ Cultivators}{Total\ Main\ Workers} * 100$	Rural & Urban
	$Main\ Agricultural\ Labour = \frac{Main\ Agricultural\ Labour}{Total\ Main\ Workers} * 100$	Rural & Urban
	$Main\ Household\ Industry\ Workers = \frac{Main\ Household\ Industry\ Workers}{Total\ Main\ Workers} * 100$	Rural & Urban
	$Main\ Other\ Workers = \frac{Main\ Other\ Workers}{TotaL\ Main\ Workers} * 100$	Rural & Urban
Population Composition	$Population$	Rural & Urban
	$Rural\ Population$	Rural
	$Urban\ Population$	Urban
	$Sex\ Ratio = \frac{No.\ of\ females}{No.\ of\ Males} * 1000$	Rural & Urban
Land Use	$Forest\ Land = Forest\ Land\ in\ hectares$	Rural
	$Cultivated\ Area = Total\ Irrigated\ Area + Total\ Unirrigated\ Area\ (in\ hectares)$	Rural
	$non - agricultural\ Area, in\ heactres$	Rural
	$Culturable\ Wasteland, in\ hectares$	Rural

Main Workers: Those workers who had worked for the major part of the reference period (i.e. 6 months or more) are termed as Main Workers.

Marginal Workers: Those workers who had not worked for the major part of the reference period (i.e. less than 6 months) are termed as Marginal Workers.

Cultivator: For purposes of the census a person is classified as cultivator if he or she is engaged in cultivation of land owned or held from government or held from private persons or institutions for payment in money, kind or share. Cultivation includes effective supervision or direction in cultivation. A person who has given out her/his land to another person or persons or institution(s) for cultivation for money, kind or share of crop and who does not even supervise or direct cultivation of land, is not treated as cultivator. Similarly, a person working on another person's land for wages in cash or kind or a combination of both (agricultural labourer) is not treated as cultivator.

Agricultural Labourers: A person who works on another person's land for wages in money or kind or share is regarded as an agricultural labourer. She or he has no risk in the cultivation, but merely works on another person's land for wages. An agricultural labourer has no right of lease or contract on land on which she/he works.

Household Industry Workers: Household Industry is defined as an industry conducted by one or more members of the household at home or within the village in rural areas and only within the precincts of the house where the household lives in urban areas. The larger proportion of workers in the household industry consists of members of the household. The industry is not run on the scale of a registered factory which would qualify or has to be registered under the Indian Factories Act. Some of the typical industries that can be conducted on a household industry basis are foodstuffs, such as production of floor, milking or dehusking of paddy, grinding of herbs, production of pickles, preservation of meat etc.; beverages, such as manufacture of country liquor, ice cream, soda water etc.; tobacco products, such as bidi, cigars, textile cotton, jute, wool or silk; manufacture of wood and wood products, paper and paper products etc.

Other Workers: All workers, i.e., those who have been engaged in some economic activity during the last one year, but are not cultivators or agricultural labourers or in Household Industry, are 'Other Workers(OW)'. The type of workers that come under this category of 'OW' include all government servants, municipal employees, teachers, factory workers, plantation workers, those engaged in trade, commerce, business, transport banking, mining, construction, political or social work, priests, entertainment artists, etc.

Non-Workers: A person who did not at all work during the reference period was treated as non-worker. The non-workers broadly constitute Students who did not participate in any economic activity paid or unpaid, dependant such as infants or very elderly people not included in the category of worker, pensioners, beggars, vagrants, prostitutes and persons having unidentified source of income and with unspecified sources of subsistence and not engaged in any economically productive work during the reference period. Others, this category includes all Non-workers who may not come under the above categories such as rentiers, persons living on remittances, agricultural or non-agricultural royalty, convicts in jails or inmates of penal, mental or charitable institutions doing no paid or unpaid work and persons who are seeking/ available for work.

Forest Area: This includes all land classified either as forest under any legal enactment, or administered as forest, whether State owned or private, and whether wooded or maintained as potential forest land. The area of crops raised in the forest and grazing lands or areas open for grazing within the forests remain included under the "forest area".

Non-agricultural Area: This includes all land occupied by buildings, roads and railways or under water, e.g. rivers and canals, and other land put to uses other than agriculture.

Culturable Wasteland: This includes land available for cultivation, whether taken up or not taken up for cultivation once, but not cultivated during the last five years or more in succession including the current year for some reason or the other. Such land may be either fallow or covered with shrubs and jungles, which are not put to any use. They may be accessible or inaccessible and may lie in isolated blocks or within cultivated holdings

Cultivable Area: It is the sum of irrigated area and unirrigated area. The area is assumed to be irrigated for cultivation through such sources as canals (Govt. and Private), tanks, tube -wells, other wells and other sources.

Appendix B: Robustness checks

Difference-in-differences results for land use change in the period 1991 to 2014, with circular buffers of 5 km radius around railway station

	Open Space & Built-up Area	Forest Cover	Water Body	Agricultural Land
post	0.062***	0.059	0.007	-0.126***
	(0.019)	(0.038)	(0.007)	(0.030)
treatment	0.040**	-0.261***	0.006	0.218***
	(0.019)	(0.038)	(0.007)	(0.030)
post x treatment	0.110***	0.033	-0.008	-0.138***
	(0.027)	(0.053)	(0.010)	(0.042)
Control group baseline mean	0.131	0.580	0.017	0.267
R-squared	0.41	0.31	0.01	0.42
Number of observations	208	208	208	208

Note: Results are based on difference-in-differences estimations according to Equation [1] as set out in Section 3.1.1. of the main manuscript; standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Difference-in-differences results for land use change in the period 1991 to 2014, with comparison region 10 km away from railway line

	Open Space & Built-up Area	Forest Cover	Water Body	Agricultural Land
post	0.078***	0.092***	-0.002	-0.170***
	(0.014)	(0.025)	(0.005)	(0.02)
treatment	0.001	-0.172***	0.013**	0.158***
	(0.014)	(0.025)	(0.005)	(0.02)
post x treatment	0.095***	-0.025	0.004	-0.072**
	(0.019)	(0.035)	(0.007)	(0.028)
Control group baseline mean	0.168	0.316	0.024	0.492
R-squared	0.27	0.18	0.03	0.33
Number of observations	600	600	600	600

Note: Results are based on difference-in-differences estimations according to Equation [1] as set out in Section 3.1.1. of the main manuscript; standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Impacts on population and workforce composition outcomes, with highway distance bands interacted with railway station treatment (0-10 km distance)

Panel A

Variables	ln (Population) (1)	ln (Rural Population) (2)	ln (Urban Population) (3)	Sex Ratio (4)	CWPR (5)	Main workers (6)	Female Workers (7)
post	0.66***	0.66***	-0.56*	270.18***	20.93***	-30.24***	28.79***
	(0.06)	(0.06)	(0.32)	(21.62)	(1.69)	(3.29)	(1.95)
post x railway treatment	0.07	0.08	-0.04	47.78	4.65*	7.17	4.29
	(0.09)	(0.09)	(0.14)	(36.09)	(2.82)	(5.50)	(3.25)
post x railway treatment x distance highway 0-5km	-0.00	0.01	0.18	-32.16	-3.83	-6.49	-4.83
	(0.09)	(0.09)	(0.14)	(36.27)	(2.83)	(5.52)	(3.27)
post x railway treatment x distance highway 5-10km	-0.04	-0.05	0.18	-42.97	-4.03	-9.94*	-4.15
	(0.10)	(0.09)	(0.26)	(36.57)	(2.85)	(5.57)	(3.30)
post x railway treatment x distance highway 10-20km	0.02	-0.02		-52.61	-2.86	-7.06	-3.35
	(0.10)	(0.10)		(38.19)	(2.98)	(5.82)	(3.44)
post x railway treatment x distance highway 20-30km	-0.15	-0.15		-3.80	-5.05	-19.51**	-0.01
	(0.15)	(0.15)		(58.33)	(4.55)	(8.88)	(5.26)
R-squared	0.08	0.05	0.52	0.11	0.03	0.11	0.20
Number of localities	7,787	7,700	170	7,787	7,787	7,787	7,787
Number of observations	15,432	15,188	244	15,431	15,432	15,431	15,431

Panel B

Variables	Cultivators (8)	Agricultural labour (9)	Household industry workers (10)	Other workers (11)	Forest land (12)	Cultivated Area (13)	Non-agricultural area (14)	Culturable wasteland (15)
post	1.90	-45.14***	-0.21	43.46***	-0.02	0.83***	-1.20***	-3.52***
	(3.28)	(3.02)	(0.57)	(2.80)	(0.13)	(0.12)	(0.17)	(0.31)
post x railway treatment	9.45*	-8.12	-1.21	-0.13	0.03	-0.13	-0.12	1.28***
	(5.47)	(5.03)	(0.95)	(4.68)	(0.24)	(0.19)	(0.27)	(0.48)
post x railway treatment x distance highway 0-5km	-11.10**	5.50	1.14	4.46	0.00	0.15	0.15	-0.95*
	(5.50)	(5.06)	(0.96)	(4.70)	(0.24)	(0.19)	(0.27)	(0.49)
post x railway treatment x distance highway 5-10km	-12.02**	7.99	1.47	2.56	-0.04	0.07	0.17	-0.92*
	(5.54)	(5.10)	(0.96)	(4.74)	(0.25)	(0.20)	(0.28)	(0.49)
post x railway treatment x distance highway 10-20km	-7.40	5.75	1.03	0.62	-0.23	0.17	0.15	-0.88*
	(5.79)	(5.33)	(1.01)	(4.95)	(0.27)	(0.21)	(0.29)	(0.51)
post x railway treatment x distance highway 20-30km	-10.93	8.32	2.51	0.11	-1.49***	0.16	0.44	0.16
	(8.84)	(8.13)	(1.54)	(7.56)	(0.50)	(0.33)	(0.46)	(0.83)
R-squared	0.30	0.05	0.03	0.30	0.01	0.06	0.06	0.19
Number of localities	7,787	7,787	7,787	7,787	4,970	7,646	7,687	7,029
Number of observations	15,421	15,421	15,421	15,421	8,659	14,589	14,815	11,546

Note: Results are based on difference-in-differences estimations according to Equation [3] as set out in Section 3.1.3, with railway treatment defined as distance to railway station below 10 km (binary treatment definition T_{10km}); standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Impacts on population and workforce composition outcomes, with continuous treatment definition

Panel A

Variables	ln (Population) (1)	ln (Rural Population) (2)	ln (Urban Population) (3)	Sex Ratio (4)	CWPR (5)	Main workers (6)	Female Workers (7)
post	0.61***	0.63***	-0.33	248.80***	20.14***	-28.24***	22.53***
	(0.06)	(0.06)	(0.31)	(24.61)	(1.92)	(3.76)	(2.22)
post x distance from railway station	0.0008	0.0009	0.0033	-0.3463	-0.0902***	0.0341	-0.0126
	(0.0011)	(0.0011)	(0.0080)	(0.4316)	(0.0338)	(0.0659)	(0.0389)
post x (distance from railway station)^2	0.0000**	0.0000**	0.0001	0.0053	0.0003	-0.0001	-0.0001
	(0.0000)	(0.0000)	(0.0001)	(0.0034)	(0.0003)	(0.0005)	(0.0003)
R-squared	0.08	0.05	0.52	0.11	0.03	0.11	0.20
Number of localities	7,858	171	7,775	7,858	7,858	7,858	7,858
Number of observations	15,595	245	15,350	15,573	15,574	15,573	15,573

Panel B

Variables	Cultivators (8)	Agricultural labour (9)	Household industry workers (10)	Other workers (11)	Forest land (12)	Cultivated Area (13)	Non-agricultural area (14)	Culturable wasteland (15)
post	-4.99	-52.08***	0.92	56.16***	-0.16	0.98***	-0.98***	-2.33***
	(3.74)	(3.45)	(0.65)	(3.20)	(0.16)	(0.14)	(0.19)	(0.35)
post x distance from railway station	0.1361**	-0.0179	0.0021	-0.1203**	0.0112***	-0.0003	-0.0132***	-0.0190***
	(0.0655)	(0.0604)	(0.0114)	(0.0561)	(0.0029)	(0.0024)	(0.0034)	(0.0063)
post x (distance from railway station)^2	-0.0011**	0.0005	-0.0002**	0.0008*	-0.0001***	0.0000	0.0001***	0.0000
	(0.0005)	(0.0005)	(0.0001)	(0.0004)	(0.0000)	(0.0000)	(0.0000)	(0.0001)
R-squared	0.29	0.05	0.03	0.30	0.01	0.06	0.06	0.19
Number of localities	7,858	7,858	7,858	7,858	7,775	7,775	7,775	7,775
Number of observations	15,563	15,563	15,563	15,563	15,350	15,350	15,348	15,348

Note: Results are based on difference-in-differences estimations according to Equation [4] as set out in Section 3.1.3.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Appendix C: District-level Konkan Railways passenger shares