WPS8350

Policy Research Working Paper 8350

The WEB of Transport Corridors in South Asia

Background Paper

Wider Economic Benefits of Investments in Transport Corridors and the Role of Complementary Policies

Martin Melecky
Siddharth Sharma
Hari Subhash

WORLD BANK GROUP
South Asia Region, Office of the Chief Economist
&
Finance, Competitiveness and Innovation Global Practice
February 2018

Policy Research Working Paper 8350

Abstract

This paper estimates the impact of the Golden Quadrilateral and North-South-East-West Highways in India on welfare, social inclusion, and environmental quality. The analysis uses district-level data for 1994–2011 and the difference-in-difference method. The results suggest that the highways shifted employment from the farm to the nonfarm sector, and that this shift was accompanied by an increase in output per capita. However, there is no evidence of an impact on household expenditure per capita, the poverty rate, or the incidence of regular wage employment. The results suggest that the highways caused an increase in air pollution. The effects of the highways are heterogeneous, depending on conditions in local factor and product markets.

This paper is a product of the Office of the Chief Economist, South Asia Region and the Finance, Competitiveness and Innovation Global Practice. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at mmelecky@worldbank.org and ssharma@worldbank.org.

Wider Economic Benefits of Investments in Transport Corridors and the Role of Complementary Policies*

Martin Melecky, Siddharth Sharma,# and Hari Subhash

World Bank

Keywords: Large-scale Infrastructure Investments; Estimated Impacts; Wider Economic Benefits; Income and Consumption; Poverty; Jobs and Gender Equality; Air Pollution; India

JEL classification: R40, H54, F15, R12

* The authors thank Ruifan Shi for excellent research support. They thank Martin Rama, Marianne Fay, Bill Maloney, Uwe Deichmann, Somik Lall, Yasuyuki Sawada, Arjun Goswami, Jay Menon, Akio Okamura, Takayuki Urade, and Duncan Overfield for suggestions and comments on earlier drafts of the paper.
# Corresponding author: ssharma1@worldbank.org.

1. Introduction

Policy makers are increasingly interested in appraising large transport infrastructure proposals using rigorous methods beyond a simple cost-benefit analysis. They want these appraisals to be comprehensive and cover the wider economic benefits (WEBs) of these investments—that is, benefits that go beyond savings of travel time and vehicle operating costs. They would like to ensure that such transport infrastructure projects benefit both the big firms involved in foreign trade and the small enterprises that could exploit the increased connectivity and market access to pursue new economic opportunities. They are also interested in mitigating the possible negative impact of corridors—be it on environmental degradation, income inequality, or social exclusion of women—as well as in identifying and compensating any potential losers due to corridor investments. They are especially interested in ascertaining the net wider economic benefits of transport corridor investment—that is, benefits incurred by households net of cost (ADB et al. forthcoming), so they can design better investments and complementary policies to increase net WEBs. Yet, as pointed out in Roberts et al. (forthcoming), there is a lack of research on complementary reforms and institutions needed to inform policy on the design of transport corridor programs to achieve wider economic benefits.

To help fill this gap, this paper evaluates the impact of two major highway systems in India built largely through major upgrades of old highways— the Golden Quadrilateral (GQ) and North-South-East-West (NSEW) highway systems. It uses the difference-in-difference (DiD) method, which is increasingly used in the academic literature to estimate the impacts of infrastructure investments after they have been built (ex post) (Melo et al., 2013; Redding and Turner 2014; Berg et al. 2015; Roberts et al. forthcoming). Applying it to district-level data from 1994 to 2011, the paper first estimates the impacts the GQ and NSEW highway systems had on a set of development outcomes (WEBs). Second, the paper examines the dependence of these district-level impacts on initial conditions in factor (capital, labor, land) markets, product markets, and institutions (governance) by using interaction terms in the DiD estimation.

Appraisals of infrastructure projects often rely on ex post impact estimates of similar, completed projects (see, for example, Laird and Venables 2017). Those impact estimates are commonly based on a before-and-after comparison of outcomes in locations affected by the existing project. This method is technically suspect. A before-and-after comparison does not account for the confounding impact of other contemporaneous factors that could affect the relevant outcomes, such as important policy reforms and macroeconomic shocks. Another challenge for the appraisal is that the impacts of projects might depend on initial market conditions. If so, impact estimates based an older projects should be adjusted to take account of the initial conditions of the project being appraised.

In recent years, many studies have addressed the technical issues in impact evaluation by using more granular spatial data and adopting the DiD method. This method compares the change in the outcomes of interest across locations affected by the project (the "treatment group") and those not affected by it (the "control" group). This double differencing identifies the impact of the project, provided that the confounding factors or preexisting trends are on average similar across the control and treatment group (see, for example, Redding and Turner 2014, Melecky 2017, and

Roberts et al. 2018, who review papers that applied this method to transport infrastructure investment).

In the case of India, the DiD method has been applied to the measure the impact of the GQ highway in studies such as Datta (2012) and Ghani, Goswami, and Kerr (2016). Our approach is like that used by Ghani, Goswami, and Kerr (2016), in which the "treatment" (that is, the change in the connectivity of a location) is measured by the distance from the GQ highway. We build on their study by examining the impacts of the GQ as well as the NSEW (a more recent highway network) on a wider set of development outcomes. The WEBs include measures of welfare (district-level GDP and household consumption, district-level poverty); labor market inclusion (total and female employment in regular-wage jobs); and environmental quality (thickness of aerosol particles, carbon dioxide emissions, and nitrogen oxide emissions).[1]

We also build on existing studies by examining whether the impacts of the GQ and NSEW highways depended on a set of initial market conditions. For instance, Bosker, Deichmann, and Roberts (2015) analyze the impacts of the National Express Network (NEN) in China on real income, at both the aggregate and district levels. They simultaneously consider how these impacts have been affected by restrictions on migration associated with the country's permanent household registration (Hukou) system. Similarly, some studies have considered the complementary role of protected area status in mitigating the impacts of road construction on deforestation (Cropper et al. 2001; Damania and Wheeler 2015; Dasgupta and Wheeler 2016).

The motivation for this analysis is to better understand the complementarities between connective infrastructure and initial conditions in product and factor markets that could be relevant for policy design. Estimates of these complementarities are also useful for project/program appraisals to refine impact estimates from previous projects and better reflect the initial conditions for the project at hand.

In this paper, the capital market conditions are measured by indicators of household and firm-level access to banking, while labor market conditions are measured by rates of literacy and secondary schooling. Conditions in land markets are measured by indicators of land use constraints, and those in product markets are measured by indicators of industrial composition and private ownership. The initial quality of institutions (governance) is measured by a state-level index that looks at the following dimensions: non-transport infrastructure; social services; fiscal performance; justice, law, and order; and the quality of the legislature.

Applying the DiD method with ordinary least squares (OLS) regressions, we find that the improved connectivity due to the GQ and NSEW highway networks contributed significantly to structural transformation of the treated districts, with significant shifts from farm to non-farm employment for both highway networks. The results also suggest that the highway networks had a positive impact on the district-level per capita GDP. However, this impact was not widely shared. We do not find similar positive effects on household consumption, poverty, and regular wage

[1] We initially included deforestation in the set of outcomes but had to drop it from the estimation after closer inspection of the data, which revealed an inadequate level of variation in the extent of forest cover across districts and time.

employment across the districts, particularly for females. Moreover, we find evidence of significant trade-offs in the GQ and NSEW impacts on WEBs, estimating that while GDP increased significantly, air quality decreased significantly.

There is a concern that the route of the highways was correlated with unobserved factors that vary across locales that could have affected the outcomes of interest. Previous studies such as Ghani, Goswami, and Kerr (2016), have employed an instrumental variables (IV) approach to examine robustness to the concern of endogenous highway placement. Following them, we instrument the actual distance of a district from the highway networks (the treatment) with its distance from the straight-line path connecting the relevant nodal cities (major metropolitan areas) of those networks. The IV results are more ambiguous than the OLS results: While the estimated impact on air pollution is statistically significant, those on other outcomes, including district-level per capita GDP and the composition of employment, are not. As we discuss later in this paper, this could be because this type of IV strategy has shortcomings when trying to estimate the impacts of two large highway networks simultaneously.[2]

Given that estimated impacts on the average treated district are ambiguous for most outcomes, we focus next on examining a possible heterogeneity in such impacts. We do so by introducing interaction terms between the treatment variable (nearness to highway) and a set of variables measuring the initial market conditions and quantity of institutions in the study districts. In effect, these are triple-differencing estimates.

The results suggest that initial local market conditions are important in determining the impact of highway upgrades and the ways in which their wider economic benefits are shared across geographic units. The explanation for this heterogeneity could relate to the dynamics of the most likely mechanism through which highways affect development outcomes: accelerating trade across connected locations. The increased trade is facilitated by a reallocation of resources to more productive firms and to sectors with comparative advantage. The effectiveness of such reallocation could in turn depend on initial conditions in labor, land, capital, and product markets.[3] For example, in districts with comparative advantage in manufacturing, limited availability of land for non-farm uses could reduce the gains from a highway by constraining the reallocation of land toward manufacturing. In a district with comparative advantage in agricultural products, limited availability of farm land could constrain the potential to gain from a highway.

Examining this hypothesis with the GQ and NSEW, our results first suggest that access to credit by non-farm enterprises can amplify the positive effect of the NSEW on women's employment in regular-wage jobs, while almost equally reducing farm employment of women in districts connected to the NSEW. Thus, firms' access to credit seems to be complementary to highways. In contrast, access to bank payment and saving services could reduce any positive impact of the GQ on poverty and household consumption. Our conjecture is that this conditional impact could still be desirable if the access to formal payment and savings methods incentivizes people to save and

[2] Ghani, Goswami, and Kerr 2016 focus on a period during which the NSEW was largely unbuilt. Thus, unlike this study, they can focus on the impact of the GQ and ignore the impacts of the NSEW.

[3] For a recent literature review of how connective infrastructure affects development and the mechanisms behind it, see Roberts et al. forthcoming.

accumulate assets. However, the lack of household asset data prevents us from testing this hypothesis.

Second, labor market conditions measured by the literacy level could have amplified the positive impact of the NSEW on the per capita GDP of connected districts. Higher levels of literacy and secondary education also seem to be associated with a lower negative environmental impact of the highways. Similarly, a higher level of secondary education is associated with a more positive impact of the GQ on regular-wage employment, particularly for women, and a more negative impact on farm employment. Overall, the estimated interactions suggest that the low average level of schooling in India may have prevented a wider sharing of the benefits from connectivity.

Third, certain land market conditions—particularly a greater area of crop land—can constrain local firms and households from pursuing the higher-value non-agricultural (non-farm) opportunities opened by greater market access due to the GQ. This constraining effect also applies to non-farm employment for women. In contrast, NSEW districts experienced greater GDP increases in conditions of greater crop land availability. One explanation is that some areas located near the NSEW have rich potential in tradable farm products and agro-processing. Their gains from market access thus depend favorably on the availability of cropland.

Fourth, product market conditions such as the initial industrial composition could be important in ensuring that transport corridors aid poverty reduction. We find that the impacts of the GQ on poverty reduction were significantly enhanced by the higher share of agro-processing in manufacturing. Hence, a strong base in agro-processing could have helped unskilled rural workers move off the farm to find a job. Districts with a larger agro-processing are also likely to have better quality of "soft infrastructure" (like warehouses and cold chains) in rural areas. The result thus indicates a complementarity between hard and soft connective infrastructure.

Finally, better state-level governance is associated with a more positive impact of the GQ on the incidence of regular-wage jobs. One explanation is that better governed areas were better at foreseeing and implementing policies that enhanced the highways' impact on job creation. Note, however, that our governance measure is at the state-level and could reflect other unobserved differences across states.

The rest of the paper is organized as follows. Section 2 describes India's highways. Section 3 describes the data used in the paper. Section 4 explains the methodological approach. Section 5 then discusses the main results of the empirical analysis, and Section 6 concludes.

2. India's Highways and the Golden Quadrilateral (GQ) and North-South East-West (NSEW) Highway Systems

The Golden Quadrilateral is a large-scale highway construction and improvement project connecting India's four top metropolitan cities—Delhi, Mumbai, Chennai, and Kolkata—thereby, forming a quadrilateral. The overall length of the quadrilateral is 5,846 kilometers (km), consisting of four/six-lane express highways. The Golden Quadrilateral project was launched in 2001 as the first phrase of National Highways Development Project (NHDP), was two-thirds complete by

2005, and was mostly finished in 2007 (Ghani, Goswami, and Kerr 2016). The entire length of the quadrilateral was operational by January 2012.[4]

As the second phase of NHDP, the North-South-East-West Corridor (NS–EW), is the largest ongoing highway project in India. It consists of building 7,142 kilometers of four/six-lane expressways connecting Srinagar in the north to Kanyakumari in the south, and Silchar in the east to Porbandar in the west. Upgrades equivalent to 13 percent of the NS–EW network were initially planned to begin in Phase One alongside the GQ upgrades, with the remainder scheduled to be completed by 2007. However, work on the NS–EW corridor was pushed into Phase II and later, because of issues with land acquisition, zoning permits, and the like. In total, 2 percent of the work was completed by the end of 2002, 4 percent by the end of 2004, and 10 percent by the end of 2006. These figures include the overlapping portions with the GQ network that represent about 40 percent of the NS–EW progress by 2006. As of May 31, 2017, 6,568 of 7,142 kilometers in the project had been completed.[5]

In combination with the GQ, the NS-EW Corridor forms a key part of the Indian highway network, connecting many of India's important manufacturing, agricultural, and commercial centers. A study (Datta 2012) finds that, even in the short period of three years, firms located in districts that were not major metropolitan areas (non-nodal districts) along the GQ network witnessed a larger decline in the average input inventory (measured in terms of the number of days of production for which the inventory held was sufficient) relative to those located on other highways. The study also finds that firms in districts closer to the GQ network were more likely to switch their primary input suppliers than firms farther away. These results suggest improved efficiency and sourcing for establishments on the GQ network after its upgrade.

Another recent study (Ghani, Goswami, and Ker 2016) finds that the GQ upgrades led to a substantial growth in the activity of manufacturers in the formal sector. The growth included higher entry rates, expansion in productivity by incumbents, adjustments in the spatial sorting of industries (location and concentration of industries across districts) and improved allocative efficiency in the manufacturing industries initially located along the GQ network. The study also estimates a 49 percent overall output increase (which is equivalent to $3.8 billion) from initial values for the average district located near the GQ network.

3. Data

We use a district-level panel data set to estimate the impact of the highways. Districts are the primary administrative unit of India below the state level. In the last year in our data set, 2010–11, there were about 640 districts in India.

Our main source for the district-level data is the South Asia Spatial Database, a database being developed by the World Bank's Office of the Chief Economist for South Asia.[6] It is intended to

[4] http://www.roadtraffic-technology.com/projects/golden-quadrilateral-highway-network/.

[5] http://www.nhai.org/WHATITIS.asp.

[6] The database was in testing mode and still under development when this paper was being written, and not publicly available. The authors were able to download data upon special request. The World Bank intends to make the database public.

bring together data on India and other South Asian economies from a range of sources such as official censuses, administrative records, surveys, satellite imagery, and official maps into a single spatial data platform. It covers two points in time (2000–01 and 2010–11) and has four administrative levels (ranging in spatial detail from state or province, district, town, or village to even hundreds of thousands of gridded cells, or "tiles").

We select those district-level variables from this database that measure specific economic outcomes of interest and initial market conditions that could have interacted with the highways to generate heterogeneity in impacts. We supplement the spatial database variables with additional district-level measures of welfare and labor market conditions that we derived from various rounds of India's National Sample Survey (NSS) of Employment and Unemployment. The NSS labor force data are available for 1994–95, 1999–2000, 2004–05, and 2010–11. Thus, our panel data set covers four years if the outcome variable is derived from the NSS data, and two years otherwise. Sources of data are listed in table A.1 in the appendix.

Table 1 lists the main outcome variables used in our study and their primary data sources. The outcome variables correspond to the following categories: economic welfare, economic (labor market) inclusion, and environmental quality. The measures of economic welfare are district-level per capita GDP, mean household per capita expenditure, and the percentage of households above the poverty line. Our preferred measure of economic (labor market) inclusion is the percentage of regular-wage jobs in employment, disaggregated by gender. This choice reflects the observation that, in India, regular-wage employment has been associated with better economic outcomes than other forms of employment, such as self-employment and "casual" wage labor (Chatterjee et al. 2016). The measures of environmental quality include the percentage of particulate matter in the air, as well as measures of nitrogen oxide (NO) air pollution. A measure of deforestation was discarded because of limited variation in the data. In addition to these final outcome variables, we also estimate impacts on the breakdown of total employment by farm and non-farm jobs. The latter is an intermediate outcome of interest that indicates possible structural transformation in the local economy.

Table 2 lists the main variables reflecting initial market conditions, and their sources. They are categorized by type of market: labor, land, capital, and product. The main labor market variables are measures of human capital as of 2001: literacy rate and the percentage of those with a secondary school or higher educational qualification. The land market variables capture the nature of land endowments in districts. They measure the extent of land that is suitable/available for agriculture, as well as the mineral production capacity of the district. The capital market variables measure household access to bank services (that is, to formal bank accounts), and firms' access to bank loans, as of 2001. The product market variables include a measure of product diversification, and the share of private firms in industrial establishments; both are intended to proxy for product market competition. A third product market variable is the share of agro-processing in manufacturing. This variable proxies for the initial level of opportunity for factory work available to low-skilled workers—the majority of the workforce—particularly in rural areas. We presume that a large agro-processing sectors also signals better supply chain infrastructure (such as warehouses, cold chains, and other logistical facilities) in rural areas.

Our state-level measure of governance is drawn from Mundle, Chowdhury, and Sikdar (2016). This study scores the 19 largest Indian states along five dimensions: infrastructure; social services; fiscal performance; justice, law and order; and the quality of the legislature. The scores are based largely on "output" measures such as proportion of trials completed in less than three years, one of the indicators for justice, law and order dimension, and development expenditure as a share of total expenditure, one of the indicators for the fiscal performance dimension. The overall Governance Performance Index (GPI) combines the score on these five dimensions. We use the 2001 value of GPI as an initial interaction condition in our estimations.[7] This is shown in table 3.

A major challenge in putting together the final district database was the matching and harmonization of districts across time. Many new districts have been created, leading to changes in district boundaries and names. Currently, there are a total of 707 districts in India, compared to 640 in the 2011 Census of India and 593 recorded in the 2001 Census of India. We have addressed this issue by mapping newly created districts back to their unique parent district in 1999.[8] For instance, if district X in 1999 was split into districts Y and Z by 2010, we combine the 2010 data for Y and Z to recreate the parent district X in 2010. In addition to aggregating new districts to their 1999 parent district, we have dropped districts from the remote states of Jammu and Kashmir and northeastern India from our analysis. This is standard practice in district-level studies on India, including previous studies on the impact of the GQ highway. After these steps, our data set consists of around 425 districts per year.

Measures of Distance from the GQ and NSEW Highways

We have merged geo-coded data on the location of the GQ and NSEW networks into the district database.[9] This information is used to calculate the distances of district centroids (the geographic center of the district area) from the nearest points on the GQ and NSEW networks. Figure 1 plots the highways networks (excluding parts of NSEW that were not built by 2010) and the distance of every district centroid from its nearest point on each highway. We also categorize districts into four distance bands from each highway: nodal district (major metropolitan area at which the

[7] We could not use alternative, better-known state-level measures of governance dimensions, such as the World Bank's state-level Ease of Doing Business indicator, because to our knowledge none of the other measures are available for the initial period of our study.

[8] We chose 1999 as the reference year because there was a wave of new state creation between 1999 and 2001, leading to a sharp rise in the number of districts and significant renaming of districts, all in the space of a few years. Mapping districts to their 1999 definitions is therefore the more conservative approach to district harmonization.

[9] We are grateful to Ejaz Ghani and coauthors Arti Grover Goswani and William Kerr (Ghani, Goswami, and Kerr 2016) for sharing these data, which they compiled using official highway maps. The merger of the GQ data was relatively straightforward because the network was largely completed by 2005, and we could simply use the final GQ network map to calculate distances of district centroids to the highway. The NSEW distance calculation was more complicated because some sections of the network, particularly in its eastern leg, were not completed by 2010. We ignored those incomplete sections of the NSEW when calculating the district distances to NSEW.

highways start and end),[10] 0–40 km from the highway, 40–100 km from the highway, and more than 100 km from the highway.[11] Thus, there are eight distance bands in total.

Table 4 lists the joint distribution of districts across the GQ and NSEW across these distance bands. As explained later in the methodology section of this paper, we use these distance bands to assign districts to the "treatment" and "control" groups in our difference-in-difference estimation. Specifically, the 0–40 distance band from the GQ (or the NSEW) identifies the GQ (or NSEW) treatment districts, while the control districts are those more than 100 km away from *both* highways. Table 4 therefore shows that our sample contains about 70 GQ treatment districts and about 40 NSEW treatment districts. There are about 200 districts in the common control group.

One concern with the analysis is that districts close to the GQ could also be close to the NSEW, which would make it hard to distinguish their impacts from each other. However, table 4 shows that there is little overlap across districts in terms of nearness to these highways. Most of the districts that are close to the GQ (0–40 km from the GQ network) are more than 100 km from the NSEW network, and vice versa. This increases our confidence in being able to distinguish between the impacts of the GQ and NSEW highways.

We use a district's distance from the straight lines connecting important nodes of a highway networks as an instrument for its distance from the actual highway. Figure 2 depicts these straight-line counterfactuals of the GQ and NSEW highways, and their distances from the district centroids. The nodal cities used to construct the straight-line version of the GQ are Delhi, Kolkata, Chennai, Bangalore, and Mumbai.[12] The nodes used to construct the NSEW straight-line versions are cities at its northern, western, and southern extremities (Jalandhar, Porbandar, and Kanniakumari, respectively), and Jhansi in central India (where the East-West and North-South arms of the NSEW cross).[13] We ignore the arm of NSEW going east from Jhansi because it was largely unbuilt in 2010.

4. Estimation Methodology

We use the difference-in-difference methodology to estimate the impact of the highways on district-level outcomes of interest. This method compares the change in the outcome of interest after the highway was built in districts located close to the new highways (the "treatment districts") to those located far from them (the "control districts"). The first differencing—that is, looking at the change in the outcome after highway construction—controls for the confounding effect of unobserved factors that do not change over time. For instance, districts that are near the highways

[10] As in Ghani, Goswami, and Kerr (2016), we assign nodal districts to a separate category, and do not consider them as treated districts. Nodal districts correspond to major metropolitan areas and their peripheries, and as such, are distinct from the average Indian district. Table A.2 in the appendix lists the nodal districts.

[11] For comparability, we chose these distance bands to correspond to those used in the Ghani et al. (2016) study of GQ. The distance cutoffs are not the same because they calculate distance to the highway from the district's nearest edge, while we do so from the districts centroid.

[12] There is an additional kink at the center of the coastal Prakasam district (in Andhra Pradesh). Given the shape of India's eastern coastline between Kolkata and Chennai, this additional node was added to ensure that the straight-line segment joining those cities passes through land.

[13] There is an additional straight-line branch from Salem (in Tamil Nadu) to Kochi, to reflect a short sub-branch in the actual NSEW highway.

could have been relatively productive even before the highways were built. The second differencing—that is, comparing the change across treatment and control districts—controls for the confounding effect of unobserved factors common to control and treatment districts that do vary over time. For instance, reforms undertaken around the same time as the highway construction could have led to a general rise in productivity across India. Thus, the identification assumption behind this approach is that unobserved time-varying factors had the same impact across control and treatment districts.

4.1. Estimating the Average Impacts of the Highways

Formally, the underlying regression specification can be described as follows:

$$Y_{i,t} = \beta \times Highway_i \times Post_t^{Highway} + \emptyset_i + \varphi_t + \varepsilon_{i,t} \ . \quad (1)$$

This regression is estimated on district-level panel data. Here, $Y_{i,\,t}$ is an outcome of interest in district *i* and year *t*. The dummy variable $Post_t^{Highway}$ is equal to one in years after the highway completion, and zero in years prior to that. The dummy variable $Highway_i$ is equal to one in districts close to the new highways (the treatment districts) and zero otherwise. $\emptyset_i$ is a set of district fixed effects that controls for time-invariant district-level factors, and φ_t is a set of year dummies that control for unobserved time-varying factors common to all districts. The impact of the highways is estimated by β, the coefficient on the treatment term (the interaction $Highway_i \times Post_t^{Highway}$), which measures how the change in the outcome after the highway was built differed across control and treatment districts.

We adjust this basic specification to account for the fact that we are simultaneously estimating the impacts of *two* highway networks, the GQ and the NSEW. There are two factors to consider in this regard. First, in estimating the impact of either highway network, it is important to control for the presence of the other one. Second, the two networks could have had different impacts. In other words, there were two sets of treatment districts: those proximate to GQ, and those proximate to NSEW.

Following Ghani, Goswani, and Kerr (2016), we assign districts into distance bands based on proximity of district centroid to the GQ. The bands are: more than 100 km from the nearest GQ point, 40–00 km from the GQ, 0–40 km from the GQ, and nodal districts. We then interact indicators for each district distance band with a variable indicating the years after the GQ was built. We then repeat this process for the NSEW. Thus, the specification we estimate is as follows:

$$Y_{i,t} = \beta^{GQ} \times GQ_i \times Post_t^{GQ} + \beta^{NSEW} \times NSEW_i \times Post_t^{NSEW} + \emptyset_i + \varphi_t + \varepsilon_{i,t}. \quad (2)$$

Here, GQ_i (respectively, $NSEW_i$) is a vector of dummies indicating the distance band from the GQ (respectively, NSEW) to which district *i* belongs, while $Post_t^{GQ}$ (respectively, $Post_t^{NSEW}$) is a dummy equal to one in the years after GQ (respectively, NSEW) completion. The omitted distance band dummy corresponds to districts more than 100 km from the highway (GQ or NSEW). $\emptyset_i$ is a set of district fixed effects, and φ_t is a set of year dummies (or state-year dummies).

As explained in the data description section, depending on the outcome, the panel data set covers either four years (1994–95, 2000–01, 2004–05, and 2010–11) or two (2000–01 and 2010–11). Because the GQ network was largely complete by 2005, the indicator $Post_t^{GQ}$ is set equal to one in the years 2004–05 and 2010–11, and zero otherwise. Work on NSEW started after 2005, and therefore $Post_t^{NSEW}$ is set equal to one only in 2010–11. Only segments for the NSEW that were complete in 2010 are considered when assigning districts to distance bands around NSEW.

The impact of the GQ is measured by the β^{GQ}, corresponding to the 0–40 km distance band from GQ, to be denoted by $\beta^{GQ,0-40}$ hereafter. Because we are controlling for $NSEW_i \times Post_t^{NSEW}$, $\beta^{GQ,0-40}$ in effect measures how the post-GQ change in the outcome differed between districts 0–40 km from GQ (the GQ treatment group) and districts more than 100 km from both highways (the control group). Similarly, the impact of NSEW is measured by the β^{NSEW}, corresponding to the 0–40 distance band from the NSEW, denoted by $\beta^{NSEW,0-40}$. Our main results tables thus report these two βs.

In our preferred specification, we replace the year fixed effects with more flexible state-year fixed effects. This controls for unobserved state-level differences in growth patterns, which is important given the documented divergence in economic growth across Indian states (Government of India, Ministry of Finance 2017).

4.2. Examining Robustness to Endogenous Highway Routes: Instrumental Variables Estimation

OLS estimates of the highway impacts could be biased if the placement of the highways was correlated with unobserved factors affecting local developmental outcomes. For example, it could be that the placement of the highways was deliberately tilted toward locations with poor growth prospects, with the objective of achieving more spatially "balanced" growth. Datta (2012) argues that such endogenous placement is not a major concern with the GQ and NSEW projects because they were largely upgrade projects, and as such, their routes were pre-determined by the existing highway networks connecting their major nodal cities. While we agree with this reasoning, a concern remains that the highway planners still had some choices to make between alternative existing highway segments.

Ghani, Goswami, and Kerr (2016) approach this issue by using an instrumental variables (IV) estimation strategy as a robustness check. The idea is to instrument for the treatment (proximity to the highway) with a variable that is correlated with the treatment but arguably uncorrelated with unobserved factors affecting the outcomes being considered. Their IV strategy relies on the fact that because the highways were meant to improve the connectivity between certain pre-specified nodal cities, the straight-line path connecting those nodes would predict the path of the highway while remaining uncorrelated with unobserved determinants of local developmental outcomes. Following their approach, we create a dummy variable indicating districts whose centroids are within 40 kilometers of the straight lines connecting the GQ nodes, and a similar dummy variable for the NSEW. These variables are used to instrument for the GQ and NSEW treatment variables in a two-stage least squares (2SLS) estimation of equation (2).

4.3. *Estimating Conditional Impacts: Could Highway Impacts Have Depended on Market Conditions?*

We also test the hypotheses that the impact of the highways depended on conditions in factor and product markets. Roberts et al. (forthcoming) use a simple policy model to argue that gaining a greater understanding of the conditional impacts of transport corridor investments—including highways—is needed to better inform policy decisions on the design of corridor investment programs. However, they also document that attempts to estimate conditional impacts of connectivity are scarce in the literature. Melecky (2017) discusses that because nonlinear structural general equilibrium (GE) models are commonly log-linearized using only the first-order Taylor series expansion (not a higher-order one), interaction terms typically disappear from GE econometrics. If the second (or higher) order Taylor series expansion is used the conditional effects of transport corridors could be directly motivated and derived from existing structural GE models. The resulting econometrics would then involve working with interactive terms between the connectivity to a transport corridor and market conditions as well as institutions (depending on the theory).

Our regression estimation thus exploits the information on varying initial market conditions across districts and adopts a difference-in-difference-in-difference approach by interacting the treatment ($Highway_i \times Post_t^{Highway}$) with variables capturing initial conditions in districts. The exact specification for the conditional impacts is as follows:

$$Y_{i,t} = \beta^{GQ} \times GQ_i \times Post_t^{GQ} + \beta^{NSEW} \times NSEW_i \times Post_t^{NSEW} + \delta^{GQ} \times GQ_i \times Post_t^{GQ} \times Z_i + \delta^{NSEW} \times NSEW_i \times Post_t^{NSEW} \times Z_i + \gamma_1 Post_t^{GQ} \times Z_i + \gamma_2 Post_t^{NSEW} \times Z_i + \emptyset_i + \varphi_t + \varepsilon_{i,t} \,. \quad (3)$$

Here, Z_i is a vector of initial conditions of interest in district *i*. The effect of initial conditions on the impact of the highways is estimated by the δs, the coefficients on the triple interaction term between $Highway_i \times Post_t^{Highway}$ and the Z_i s.

For illustration, suppose that the Z_i in question is a variable measuring the efficiency of land markets, with higher values indicating more efficiency. The corresponding δ^{GQ} (respectively, δ^{NSEW}) coefficient measures how the impact of the GQ (respectively, the NSEW) depends on the efficiency of land markets. A positive estimate of this δ^{GQ} would imply that the impact of the GQ on the outcomes of interest was more positive in districts with more efficient land markets.

Many potential Z_i variables are available for each factor and product market. The estimation of interaction effects is also complicated by the fact that potential Z_i variables could be correlated, which makes it important to check if an estimated interaction term is robust to controlling for interactions with other potential Z_i variables. Hence, we face a practical issue of choosing a parsimonious regression specification without omitting important Z_i variables. We address this issue by employing a simple iterative algorithm that starts with the full set of interaction variables and progressively drops interaction terms with low *p*-values, until we are left with a small set of interaction terms. Because of this procedure, the set of Z_i variables in the final specification varies across outcome variables. The procedure ensures that the results shown are robust to controlling for other interaction terms.

5. Results

5.1. Average Impacts of the Highways: OLS Estimates

Tables 5–7 show the results of estimating equation (2) by OLS. This is our baseline difference-in-difference specification measuring the average impact of the highways. The tables report estimates from regressions including a full set of state-year fixed effects, our preferred specification, as it controls for unobserved state-level variables which could have been changing significantly during the study period. The impact of the GQ is given by $\beta^{GQ,0-40}$, the coefficient on the GQ treatment term $GQ_i^{0-40} \times Post_t^{GQ}$, which measures how the change in outcomes varied across treatment districts (those 0–40 km from the GQ network) and control districts (those more than 100 km away). Likewise, the impact of NSEW is given by $\beta^{NSEW,0-40}$, the coefficient on the NSEW treatment term $NSEW_i^{0-40} \times Post_t^{NSEW}$.

The first interesting result is the impact of the highways on structural transformation—that is, on the movement of labor from farm to non-farm work. According to the point estimate of $\beta^{GQ,0-40}$ in table 5, column 5, which is statistically significant at the 1 percent level, the GQ highway increased the share of non-farm employment by 1.6 percentage points. This is a big impact because the baseline (that is, control district) increase in the share of non-farm employment during this period was 2.5 percentage points. The GQ also increased the share of non-farm employment among females by 2.4 percentage points (column 6). The estimate of $\beta^{NSEW,0-40}$ is also positive and statistically significant. Specifically, the NSEW highway appears to have raised the share of non-farm employment by about 2.5 percentage points in the general population, as well as among females.

The positive impacts of the highway construction on the farm to non-farm transition suggests that limited access to markets is a reason for the slow structural transformation observed in India. The lack of market access has reduced employment opportunities beyond the farm, while keeping workers trapped in low-productivity agriculture by shielding it from competition. It is important to note, however, that this hypothesis does not imply that highways should lead to a movement away from farm work in *every* district. In districts with a strong inherent comparative advantage in agriculture, it could theoretically lead to a movement into the farm. The results only suggest that the movement from farm employment is the case for the *average* treatment district.

As for the final development outcomes of interest, the results do not indicate a statistically significant impact of the highways on the share of regular-wage employment, our preferred measure of better labor market inclusion (table 5, columns 1 and 2). We also do not find statistically significant impacts of either highway on measures of welfare such as mean household per capita consumption expenditure and the percentage of households above the poverty line (table

6). However, the regressions suggest that the GQ highway had a statistically significant positive impact on district output per capita. Looking at the results for the GDP per capita (in logs) variable, the point estimate of $\beta^{GQ,0-40}$ implies that the highway increased growth in GDP per capita over 2001–11 by 4 percentage points. This increase is of significant magnitude even with the baseline increase of 27 percent. The corresponding estimated impact of NSEW is statistically insignificant.[14]

The generally weak results on wage employment and welfare indicators are at odds with the seemingly positive impact on GDP per capita, and the prior finding that the GQ had a significant impact on output, productivity, and wages in formal manufacturing (Ghani, Goswami, and Kerr 2016). The formal manufacturing sector still comprises a small share of output and employment—particularly jobs—in most Indian districts. It is possible that increased market access from the highways mainly benefitted formal manufacturing and the relatively small number of skilled workers employed in that sector. This was enough to cause a significant increase in total (and per capita) district GDP. But because the growth in formal manufacturing started from such a small base, it did not lead to a detectable increase in the total number of regular-wage jobs, or in the incomes of low-skilled individuals. The results thus suggest that the economic benefits from the highways were not shared widely.

Interestingly, our results also suggest that the structural change caused by the highways was not in itself enough to lead to a significant increase in "good" (that is, regular-wage) jobs. Nor was it enough to cause significant poverty reduction. The opening up of market access thus drove workers off the farm, but not necessarily into better paying jobs. This could have happened because most farm workers are uneducated and could not access well-paying jobs in other sectors. Other labor market frictions, such as barriers to rural-urban migration or restrictive labor regulation of formal jobs, could also have played a role.

Another potential explanation for these results is the possibility of migration among districts in response to the highways. That is, the highways did increase the demand for non-farm labor significantly, but immigration into the treated districts was sizable enough to wipe out any equilibrium wage increase. This explanation is at odds with the well-established facts that India's rates of domestic migration are low by international standards and that most domestic migration in India is within-district (Kone et al. forthcoming). But it bears further investigation, possibly with district-level migration data.

Finally, we look at the highways' impacts on environmental measures (table 7). While we do not detect a significant impact on CO_2 and NO levels in the air, it appears that the GQ highway led to an increase in air pollution related to particulate matter. Specifically, the GQ is estimated to have increased "aerosol optical thickness," a measure of particulate pollution, by 0.02 point, relative to a baseline increase of 0.06 point. This result indicates a significant trade-off between economic benefits and pollution.

[14] The OLS results (available upon request) also suggest that night lights intensity, a commonly used proxy for economic activity, was not affected significantly.

5.2. IV Estimates

Next, we estimate equation (2) by 2SLS, using the straight-line path connecting nodal cities as instruments for the actual path of the highways. Specifically, the instrument for the GQ (respectively, NSEW) treatment variable—the dummy indicating whether a district centroid is within 40 kilometers of the highway—is a dummy variable indicating whether a district centroid is within 40 kilometers of the straight lines connecting the GQ (respectively, NSEW). We interact these with the respective *Post* dummies to generate the counterparts of the treatment variable in the panel.

Unlike the original specification (equation 2), we no longer include $GQ_i^{40-100} \times Post_t^{GQ}$ and $NSEW_i^{40-100} \times Post_t^{GQ}$ in the set of explanatory variables, and also drop nodal districts. The reason for this is that we have just one instrument per highway, and therefore cannot instrument for multiple distance bands. In this sparser version of equation (2), the treatment districts stay the same, but the control districts—the omitted distance bands—are those more than 40 kilometers away from both highways, excluding nodal districts. Tables A.3–A.5 in the appendix show the OLS estimates of this modified version of equation (2), which are qualitatively similar to the OLS results discussed in section 4.1. Thus, the OLS results are robust to using this sparser specification.

Table 8 presents the first-stage results. This is estimated on the cross-section of districts in a single year. As expected, the straight-line distance dummies are significantly correlated with the actual distance dummies. The instruments pass the F-test of joint statistical significance in predicting the treatment variables.[15] That said, the joint distribution of the straight-line and actual distance dummies (tables 9 and 10) gives us some reservations about the instruments. Almost half the districts within 40 kilometers of the GQ are *outside* the 40-kilometer band around the straight-line counterpart of GQ. In a sense, the instrument's correlation with the instrumented variable is driven by the subsample of districts that are far away from the highway.

The 2SLS results are presented in tables 11–13. While we have reservations about the instruments, the results do add to the general impression that the highways did not have transformational impacts on welfare and labor market outcomes in the average treatment district. Indeed, unlike the OLS estimates, the 2SLS estimates of the impact of the GQ on GDP per capita and the share of non-farm employment are statistically not significant. The only OLS result that is clearly robust to IV is the positive impact on air pollution, as measured by aerosol optical thickness (table 13, column 2). The 2SLS impact on non-farm employment among females is also marginally significant, and the sign consistent with the OLS results.

[15] It is also notable that the F-test values are generally not as large as those reported in Ghani, Goswami, and Kerr (2016). This seems puzzling because of the similarities in our data sets, as well as the treatment and instrumental variables. However, our sample of districts is significantly larger than that in Ghani, Goswami, and Kerr (2016) because they drop districts that were not sampled (or had an insignificantly small sample of firms) in the Annual Survey of Industries, a firm-level data set that is the source of their main outcome variables.

5.3 How the Impacts Could Have Depended on Product and Factor Market Conditions: Estimates of Interaction Effects

Overall, the results discussed so far suggest that while the GQ may have increased industrial output, employment, and wages (Ghani, Goswami, and Kerr 2016), it did not have significant impacts on a range of other measures of economic outcomes (wider economic benefits) in the average district. It appears to have, on average, increased district output per capita, but this result is not robust to the instrumental variables estimation. The impacts of the NSEW are even more elusive.

Because the average district impacts on economic outcomes are low (apart from particulate pollution), we examine the heterogeneity in the impacts of the highways across districts. This section presents OLS estimates of the triple-difference specification spelled out in equation 3.[16] Recall that we are interested in estimating the coefficients on the interactions of the treatment ($Highway_i \times Post_t^{Highway}$) with variables Z_i capturing initial conditions in a district's product and factor markets, and state-level governance. The coefficients on these triple interactions, denoted by the vectors δ^{GQ} and δ^{NSEW}, measure how the impact of the highways depended on those initial conditions. The main results, which look at interactions with product and factor market conditions, are shown in tables 14–16. The tables are arranged by outcome categories: table 14, welfare outcomes; table 15, labor market outcomes; and table 16, environmental outcomes.

As noted, we used a simple iterative procedure to reduce the number of extraneous interaction variables; thus, the set of interaction variables in the final specification varies across outcomes. These regressions include state-year dummies to control for state-level shocks.

Table 17 shows the interaction results when we also include a measure of state-level governance performance among the interaction terms. These regressions use year dummies instead of state-year dummies because the governance variable does not vary within states.

Our broad hypothesis is that gaining the full benefits of market access could have depended on certain factor endowments like skills (which are immobile in the short to medium term), and on product and factor market efficiency. If this is correct, then low average levels of factor endowments or market efficiency could explain why the *average* district did not experience widespread benefits from the highway construction. Identifying such complementary factors can reveal how the highway construction could be combined with complementary public interventions to maximize the wider economic benefits generated by these highways.

5.3.1. Capital Markets (people's access to bank accounts and firms' access to credit)

The results for capital markets are intriguing, and hint at a complex relationship between market access and access to formal financial services.

[16] Given our reservations about the instrumental variables available, we no longer present 2SLS counterparts of these regressions.

When looking at the impact of the GQ on mean household consumption expenditure per capita, the estimated coefficient on the interaction between the GQ treatment and a measure of households' access to bank accounts is negative and significant (table 14).[17] This implies that the impact of the GQ on household expenditure was less positive in districts where households had better access to formal savings accounts. Given that the main use of these formal accounts is to save, one potential explanation for this is that even though the highways increased household earnings, a larger fraction of that additional income was saved in locations where households had better access to formal channels of saving. This could also reflect an unmet demand for channels of asset diversification, and the possible intention of households to build resilience to shocks.

Another potential explanation is that highways help attract capital to unbanked locations. This would explain why they have less impact on some outcomes in areas where access to finance is better initially. There is some support for this conjecture in recent research on the impact of connective infrastructure in China (Banerjee, Duflo, and Qian 2012). However, in our case, this hypothesis is not consistent with the estimated interaction effect of *firms'* access to finance and the NSEW. When looking at the share of non-farm jobs among females, we find that the estimated coefficient on the interaction between the NSEW treatment and a measure of firms' access to formal financing is positive and significant (table 15).Consistent with this finding, the estimated interaction effect of firms' access to formal financing and the NSEW on female farm employment share is negative (table 15).[18] This result suggests that firms' access to formal loans could have enhanced the impact of the highways on structural transformation and increased the movement of women off the farm.

5.3.2. Labor Markets (Skills)

The results indicate a complementarity between highways and skills. Looking at the impact on the share of non-farm employment, the coefficient on the triple interaction between a measure of secondary schooling (the initial percentage of those 15 years old or older who have a secondary school education) and $GQ_i \times Post_t^{GQ}$ is positive and statistically significant at the 1 percent level (table 15). This result implies that the GQ had a significantly more positive impact on non-farm employment in districts with a greater share of secondary education. The point estimate of the δ^{GQ} implies that relative to a district at the 10th percentile of the secondary schooling measure, a district at the 90th percentile experienced an impact of the GQ on the share of non-farm employment that

[17] Also, the estimated coefficient on the interaction between the GQ treatment and the banking access measure is negative and significant when the outcome being examined is the reduction in the poverty headcount. Because the headcount measure is derived from the consumption expenditure measure, these interaction results are consistent with each other.

[18] When looking at the impact on the share of regular-wage jobs in female employment, the interaction between firms' access to finance and NSEW is negative. This is in apparent contrast to the positive interaction impact of firms' access to finance and NSEW on non-farm jobs. However, it reflects a positive interaction impact of firms' access to finance and NSEW on total female employment, which is the denominator for estimating the share of regular-wage jobs in female employment. Overall, the interaction between firms' access to finance and NSEW increased female non-farm employment and decreased female farm employment, and the former was the dominant effect. There was no significant interaction impact on the absolute number of regular-wage jobs among females, but because the interaction impact on total female employment was positive, the interaction impact on the ratio of regular-wage jobs to total jobs was negative.

was 7 percentage points larger. Similarly, we observe a positive and statistically significant interaction of the secondary schooling variable with $NSEW_i \times Post_t^{NSEW}$.[19] The same pattern is also observed concerning the impact on the share of non-farm employment among females.

Education also seems to have enhanced the impact of the NSEW on regular-wage jobs, although in this case it is basic literacy that seems to have mattered. The estimated coefficient on the interaction between the literacy rate and $NSEW_i \times Post_t^{NSEW}$ is positive and significant for the share of regular-wage jobs in total employment (table 15). It implies that moving from the 10th percentile of the literacy measure to its 90th percentile can increase the NSEW on the share of regular wage jobs by 12 percentage points.

Results for GDP per capita as an outcome also highlight a strong complementarity between connectivity and skills (table 14). While the results for the GQ are not statistically significant, in the case of the NSEW, the coefficient on the triple interaction between basic education (the initial literacy rate among those 7 years old or older) and $NSEW_i \times Post_t^{NSEW}$ is positive and statistically significant at the 5 percent level. The implied complementarity is large in magnitude: that is, moving from the 10th percentile of the literacy measure to its 90th percentile can increase the NSEW impact on per capita GDP by 17 percentage points. Recall that baseline growth in GDP per capita during 2001–11 was about 27 percent.

As shown in table 16, the negative impact of highways (the GQ) on particulate matter air pollution (aerosol optical thickness) is mitigated by having a better educated population (as measured by the share of the population with completed secondary schooling). This result could arise because more educated individuals buy higher quality, environmentally cleaner vehicles as their income rises. Another possible explanation for this interaction is that the way in which economic activity changes after the highway is completed is more environmentally friendly in more educated districts. Notably, districts with higher levels of secondary schooling experience a greater impact of the GQ on the shift from farm to non-farm jobs. This shift would have reduced pollution from the burning of straw in farms, a major contributor to particulate pollution in India.

5.3.3. Land Markets

Our results suggest that the impact of the highways also depended significantly on the share of cropland in a district's total land area. This interaction variable can be seen as a proxy for a district's comparative advantage in agriculture. A higher share of cropland also signals a bigger constraint on the availability of land for industrial purposes.

Specifically, when looking at the impact of the GQ on the share of non-farm jobs among females, and on the share of regular-wage jobs among both males and females, the coefficient on the interaction of the GQ treatment with the cropland measure is negative and statistically

[19] The interaction term on $GQ_i \times Post_t^{GQ}$ and a measure of literacy, however, is negative and statistically significant. This is puzzling but, given that the regression is also controlling for the interaction between secondary schooling and $GQ_i \times Post_t^{GQ}$, this could be an artifact of the correlation between these two educational measures. We do not observe this puzzling pattern for NSEW, and in general, the coefficients on interaction with educational measures go in the positive direction.

significant (table 15). For example, the estimates of the interaction term imply that moving from the 10th to the 90th percentile of the share of cropland would reduce the impact of the GQ on non-farm employment among females by 5 percentage points, and that on the share of regular- wage jobs by 0.1 percentage points. This finding is consistent with the hypothesis that the extent to which improved market access shifted people out of farm jobs was negatively related to comparative advantage in agriculture (or it depended positively on the availability of industrial land). Consistent with this hypothesis—when looking at impacts on household consumption per capita—the coefficient on the interaction of the GQ treatment with the cropland measure is *negative* and statistical significant (table 14). The GQ districts gained less from market access if they had less land available for non-farm use.

The results also suggest that the interaction with land conditions differed across the GQ and the NSEW. First, a greater share of cropland is not significantly associated with a lower positive effect of the NSEW on non-farm jobs. This finding suggests that unlike for the GQ, a larger share of crop land did not impede the structural change induced by the NSEW (table 15). Second, looking at GDP per capita, the coefficient on the interaction of the NSEW treatment with the cropland measure is *positive* and statistical significant: that is, having more cropland seems to have enhanced the impact of the NSEW highway on per capita output (table 14) in connected districts. A potential explanation for these results is that some districts in the areas connected by the NSEW began to specialize in high-value tradable farm products and agro-processing. Among these districts, having more cropland increased the gains from market access, while not necessarily impeding the creation of non-farm jobs in agro-related industries.

5.3.4. Product Markets

We also test the hypothesis that the gains from highways depend on efficiency and competition in product markets. For this part of the analysis, we rely on proxies for product market competition at the district level, such as the share of the private sector in formal manufacturing. We do not observe a statistically significant and consistent interaction between these measures and the impact of the highways; perhaps it reflects the crude quality of the available measures.

There is, however, a positive and statistically significant interaction between the share of agro-processing in local industry and the impact of the GQ on alleviating rural poverty (table 15). There could be two reasons for this complementarity. First, as discussed, for the structural change brought by the GQ to have translated into widespread benefits, the availability of suitable jobs for the vast reserves of low-skilled workers leaving the farm could have been critical. Agro-processing could be an important source of suitable jobs for rural workers. Second, the size of the agro-processing sector could be acting as a proxy measure for the quality of "soft infrastructure" (such as warehouses and cold chains) in rural areas. This result is indicative of a complementarity between hard and soft connective infrastructure.

5.3.5 Governance

In addition to specific conditions in factor and product markets, the impact of connectivity could also depend on cross-cutting institutions such as the quality of governance. Better governed areas, for instance, might have been better at enacting complementary policies in locations affected by the new highways. We test this idea by including a measure of state-level governance performance in 2001 as an additional interaction term in the regressions. As noted, we cannot include state-year fixed effects if we want to estimate this interaction. Hence, the results with governance interactions should be interpreted with the caveat that they may not be robust to controlling for unobserved state-level shocks. Accordingly, we present these as additional results, and do not include the governance interaction in our main results which are shown in tables 14–16.

We first note that the statistical significance levels and signs of the product and labor market interaction effects shown earlier survive the replacement of state-year fixed effects with year fixed effects, and the addition of a governance interaction term. Thus, the estimated interaction effects of product and labor market conditions do not change in this alternative specification. Second, the newly included governance and highways interaction terms are not statistically significant for most outcomes, with the important exception of labor market inclusion measures. These results are shown in table 17.[20] Looking at the share of regular-wage jobs in total employment, our preferred measure of "good" job creation, the interaction of governance with the GQ is positive and significant. The same effect is observed for the share of regular-wage jobs among females. These findings suggest that better governance enabled a wider sharing of the jobs impacts of highways.

6. Conclusion

This paper suggests that improved connectivity through highways led to some structural change in districts on or near the highways and increased air pollution. It also finds some evidence of a beneficial impact on the output of connected districts, but this result is less robust. On average, the paper finds no evidence of widely shared economic benefits as measured by household expenditure and poverty rates. It provides evidence that a wider sharing of benefits from large transport infrastructure investments across districts and people could be significantly dependent on complementary policies and institutions.

For the GQ, our estimations suggest that the highway did cause a significant reduction in poverty, but only in districts with a preexisting larger agro-processing base—not in other districts. Locations that did not have land restricted to agricultural use experienced an increase in wage employment among females. For the NSEW, higher levels of education could have enhanced the impact on wage employment.

The results are largely consistent with the idea that transport corridors affect economic outcomes by increasing market access and trade across connected locations. This process requires a reallocation of land, labor, and capital to economic sectors with a comparative advantage. The GQ, for example, led to a significant reallocation of workers from the farm to the non-farm sector.

[20] Table 12 corresponds to the same set of outcomes as table 10, replacing state-year dummies with year dummies and adding the governance interactions. Notably, the significance and sign of the other interaction terms is similar in these tables.

The benefits from transport corridors—and how widely they are shared—therefore depend on how effectively land, labor, and capital can move to new sectors of promising activity. This effectiveness—or ineffectiveness, for that matter—could be the key reason why the impacts of highways like the GQ depend on initial market conditions.

Our results also illustrate the potential trade-off impacts of corridors involving environmental quality. The estimates suggest that the GQ increased particulate air pollution. While not surprising given the likely increase in traffic, this potential negative impact is largely absent from policy decisions. Economic benefits could thus come at environmental costs that could be mitigated.

Our key policy message for the design of corridors is that the efficiency of markets and the effectiveness of local policies and institutions could complement corridor investments. These complementary factors play an important role in determining whether the overall spatial impacts of large transport investments on development outcomes will indeed be transformative and produce the expected wider economic benefits. Investing in human capital and better governance is important, as are policies that could mitigate the potential negative effects on environmental quality.

The design of corridor investments and their complementary policies should be based on a better understanding of the underling mechanisms through which the corridor in question could affect development outcomes. We are thus not suggesting a simple extrapolation of the patterns observed in our case study. For example, in the case of the GQ, the availability of land for non-farm uses and the level of secondary education seem to have been the main constraints for a wider sharing of its socioeconomic benefits. But this does not imply that agricultural-intensive locations or areas with low education and skills are not suitable for corridor placement. Instead, the lesson to draw is that policy makers should focus on identifying the underlying constraints in the local context of prospective corridors—be it policies and institutions behind constrained use of land or frictions behind low availability of skills in the local labor market.

Table 1. Summary of Outcome Variables

Outcome variable	Source code	Outcome type	No. of obs.	25th percentile	Median value	75th percentile	Mean value	Standard deviation
Aerosol radius	B	Environment	854	8.12	15.55	25.60	18.96	15.88
Aerosol optical thickness	B	Environment	854	0.28	0.36	0.52	0.39	0.15
Nitrogen dioxide level	B	Environment	854	137.25	167	207	175.63	58.9
Percentage households above the poverty line, total	C	Welfare	1661	0.43	0.59	0.74	0.59	0.2
Percentage households above the poverty line, rural	C	Welfare	1661	0.41	0.6	0.78	0.59	0.23
Percentage households above the poverty line, urban	C	Welfare	1626	0.43	0.58	0.75	0.59	0.2
Regular wage employed	C	Inclusion	1626	0.07	0.12	0.19	0.14	0.1
Regular wage employed, female	C	Inclusion	1626	0.03	0.06	0.13	0.1	0.11
Farm employed, female	D	Structure	854	53.11	73.85	84.1	66.17	22.98
Farm employed, total	D	Structure	854	49.00	64.55	73.79	59.13	19.85
Non-farm employed, female	D	Structure	854	15.90	26.15	46.84	33.7	22.83
Non-farm employed, total	D	Structure	854	26.21	35.42	50.9	40.79	19.84
Log GDP per capita	E	Welfare	854	5.43	5.84	6.21	5.79	0.61
Log mean household consumption	I	Welfare	1661	6.20	6.36	6.52	6.37	0.24

Note: obs. = observations.

Table 2. Summary of Market Condition Variables

Market variable	Source code	Market type	No. of obs.	25th percen-tile	Median value	75th percen-tile	Mean value	Standard deviation
Households' access to banking services	G	Capital market	1708	24.26	32.6	44.05	34.86	14.09
Access to financial services, private non-farm enterprises	H	Capital market	1708	1.50	2.75	4.50	3.61	3.21
Literacy rate, 7+ years (percent of population group)	D	Labor market	1708	55.70	63.6	72.30	63.50	12.66
Secondary education completion rate, 15+ years (percent of population group)	I	Labor market	1708	17.00	22.8	28.95	23.20	8.46
Cropland (percent of area)	A	Land market	1708	29.05	60.1	88.30	56.79	32.25
Food/beverage/tobacco / manufacturing (percent of establishments)	J	Product market	1708	13.20	20.8	32.25	25.31	17.34
Diversification index of non-farm enterprises, ISIC 3.1 2-digit (index)	H	Product market	1708	3.60	4.7	5.87	4.83	1.50
Non-agricultural enterprises by ownership - privately owned (percent of establishments)	H	Product market	1708	90.70	93.6	95.80	92.50	4.79

Note: ISIC = International Standard Industrial Classification; obs. = observations.

Table 3. Governance Measure

State	GPI score 2001	State	GPI score 2001
Gujarat	0.66	Assam	0.43
Tamil Nadu	0.60	Madhya Pradesh	0.38
Punjab	0.60	Uttarakhand	0.36
Kerala	0.57	Odisha	0.35
Haryana	0.55	Rajasthan	0.34
Andhra Pradesh	0.53	Jharkhand	0.27
Karnataka	0.51	Uttar Pradesh	0.19
Himachal Pradesh	0.50	Bihar	0.16
Chattisgarh	0.48		
West Bengal	0.44		

Source: Mundle, Chowdhury, and Sikdar 2016.
Note: GPI = Governance Performance Index.

Table 4. District Distribution by Distance to the Highway (in kilometers)

a. By distance bands

		NSEW			
		Nodal	**0–40**	**40–100**	**> 100**
GQ	**Nodal**	4			4
	0–40	1	2	10	60
	40–100	[Complete]	6	10	60
	> 100	1	32	43	194

Note: GQ = Golden Quadrilateral; NSEW = North-South-East-West Highway.

b. By distance dummy
(nodal districts excluded) (in kilometers)

		NSEW		
		0–40	**>40**	**Total**
GQ	**0–40**	2	70	72
	>40	38	307	345
	Total	40	377	417

Note: GQ = Golden Quadrilateral; NSEW = North-South-East-West Highway.

Table 5. OLS Estimation of Average Impacts of Highways on Welfare Outcomes

Variables	(1) Mean household consumption (in logs)	(2) GDP per capita (current USD, in logs)	(3) Percentage households above the poverty line, total	(4) Percentage households above the poverty line, rural	(5) Percentage households above the poverty line, urban
PostGQ*GQ (0–40)	0.00041	0.0402***	-0.0027	-0.0048	0.0151
	(0.019)	(0.015)	(0.016)	(0.018)	(0.020)
PostNSEW*NSEW (0–40)	-0.0127	0.00418	0.00635	0.00878	0.0366
	(0.028)	(0.020)	(0.024)	(0.027)	(0.030)
PostGQ*GQ (40–100)	-0.019	0.0137	-0.0055	-0.0044	0.00264
	(0.019)	(0.015)	(0.016)	(0.018)	(0.021)
PostNSEW*NSEW (40–100)	-0.0096	-0.0054	0.0229	0.0173	0.0422*
	(0.023)	(0.016)	(0.020)	(0.022)	(0.025)
PostGQ*GQ (nodal)	0.0936	-0.0156	-0.0208	-0.0647	-0.0072
	(0.074)	(0.047)	(0.064)	(0.072)	(0.079)
PostNSEW*NSEW (nodal)	0.0112	0.0395	0.0824	0.109	0.099
	(0.085)	(0.055)	(0.074)	(0.083)	(0.092)
State year fixed effects	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	1,661	854	1,661	1,661	1,626
R-squared	0.808	0.994	0.799	0.807	0.688

Note: The ordinary least squares (OLS) estimation is based on four distance bands on the proximity of district centroid to highways (0–40km, 40–100km, more than 100 km, and nodal districts) interacted with post-treatments. All the regressions control for state-year and district fixed effects. GQ = Golden Quadrilateral; NSEW = North-South-East-West Highway: USD = U.S. dollars.
*** p<0.01, * p<0.1

Table 6. OLS Estimation of the Average Impact of Highways on Inclusion and Labor Market Structure Outcomes

Variables	(1) Regular wage employed, total	(2) Regular wage employed, female	(3) Farm employed, total	(4) Farm employed, female	(5) Non-farm employed, total	(6) Non-farm employed, female
PostGQ*GQ (0–40)	-0.0017	-0.0075	-1.432**	-2.454***	1.578**	2.400***
	(0.008)	(0.011)	(0.617)	(0.825)	(0.612)	(0.821)
PostNSEW*NSEW (0–40)	0.0128	-0.0063	-2.414***	-2.426**	2.448***	2.542**
	(0.012)	(0.016)	(0.810)	(1.082)	(0.803)	(1.077)
PostGQ*GQ (40–100)	-0.0131	-0.0017	-0.725	-1.136	0.767	1.143
	(0.008)	(0.011)	(0.609)	(0.815)	(0.605)	(0.811)
PostNSEW*NSEW (40–100)	-0.0005	-0.0157	-0.314	-0.663	0.184	0.615
	(0.010)	(0.013)	(0.646)	(0.864)	(0.641)	(0.859)
PostGQ*GQ (nodal)	0.0906***	0.0237	-0.906	-2.347	0.73	1.922
	(0.031)	(0.042)	(1.921)	(2.567)	(1.905)	(2.555)
PostNSEW*NSEW (nodal)	0.0449	0.103**	0.108	-0.529	-0.0159	0.758
	(0.036)	(0.049)	(2.265)	(3.027)	(2.246)	(3.012)
State year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,626	1,626	854	854	854	854
R-squared	0.791	0.689	0.99	0.986	0.99	0.986

Note: The ordinary least squares (OLS) estimation is based on four distance bands on the proximity of district centroid to highways (0–40km, 40–100km, more than 100 km, and nodal districts) interacted with post-treatments. All the regressions control for state-year and district fixed effects. GQ = Golden Quadrilateral; NSEW = North-South-East-West Highway.
*** p<0.01, ** p<0.05

Table 7. OLS Estimation of Average Impact of Highways on Environmental Outcomes

Variables	(1) Aerosol particle radius (in logs)	(2) Aerosol optical thickness (thickness scale 0–1)	(3) Nitrogen dioxide levels (billion molecules/mm^2) in logs
PostGQ*GQ (0–40)	-0.724	0.0259***	1.006
	(0.843)	(0.006)	(1.852)
PostNSEW*NSEW (0–40)	-0.405	0.00953	-1.54
	(1.106)	(0.008)	(2.429)
PostGQ*GQ (40–100)	0.77	0.0230***	1.495
	(0.832)	(0.006)	(1.828)
PostNSEW*NSEW (40–100)	0.547	0.00046	2.87
	(0.882)	(0.006)	(1.938)
PostGQ*GQ (nodal)	-0.166	0.0229	9.886*
	(2.623)	(0.018)	(5.763)
PostNSEW*NSEW (nodal)	-2.943	0.0157	-1.771
	(3.092)	(0.021)	(6.794)
State year fixed effects	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes
Observations	854	854	854
R-squared	0.97	0.984	0.989

Note: The ordinary least squares (OLS) estimation is based on four distance bands on the proximity of district centroid to highways (0–40km, 40–100km, more than 100 km, and nodal districts) interacted with post-treatments. All the regressions control for state-year and district fixed effects. GQ = Golden Quadrilateral; NSEW = North-South-East-West Highway.
*** p<0.01, * p<0.1

Table 8. IV Estimation - First Stage (cross-section of districts)

Variables	(1) GQ 0–40	(2) NSEW 0–40
GQStraightDum (0–40)	0.377***	-0.0311
	(0.0495)	(0.0348)
NSEWStraightDum (0–40)	0.0176	0.450***
	(0.0599)	(0.0421)
State fixed effects	Y	Y
Observations	417	417
R-squared	0.233	0.375
F statistic	4.745	9.393

Note: GQ = Golden Quadrilateral; NSEW = North-South-East-West Highway.
*** $p<0.01$

Table 9. District Distribution by Actual Distance versus Straight-Line Distance to GQ Highway (nodal districts excluded)

		GQ distance		
		0–40 km	> 40 km	Total
GQ straight-line distance	0–40 km	35	30	65
	> 40 km	37	315	352
	Total	72	345	417

Note: GQ = Golden Quadrilateral.

Table 10. District Distribution by Actual Distance versus Straight-Line Distance to NSEW Highway (nodal districts excluded)

		NSEW distance		
		0–40 km	> 40 km	Total
NSEW straight-line distance	0–40 km	25	19	44
	> 40 km	15	358	373
	Total	40	377	417

Note: NSEW = North-South-East-West Highway.

Table 11. 2SLS Estimation of the Average Impact of Highways on Welfare Outcomes

Variables	(1) Mean household consump-tion (in logs)	(2) GDP per capita (current USD, in logs)	(3) Percentage households above the poverty line, total	(4) Percentage households above the poverty line, rural	(5) Percentage households above the poverty line, urban
PostGQ*GQ (0–40)	0.0387	0.027	0.0337	0.0198	0.0321
	(0.041)	(0.027)	(0.036)	(0.040)	(0.044)
PostNSEW*NSEW (0–40)	0.0139	-0.0246	0.0443	0.0299	0.0587
	(0.048)	(0.027)	(0.041)	(0.046)	(0.051)
State year fixed effects	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	1,637	834	1,637	1,637	1,602
R-squared	0.801	0.993	0.795	0.806	0.682
F statistic	9.479	132.8	9.127	9.769	4.905

Note: 2SLS = two-stage least squares; GQ = Golden Quadrilateral; NSEW = North-South-East-West Highway; USD = U.S. dollars.

Table 12. 2SLS Estimation of the Average Impact of Highways on Inclusion and Labor Market Structure Outcomes

Variables	(1) Regular wage employed as a shared of total employed, total	(2) Regular wage employed as a shared of total employed, female	(3) Farm employed, total	(4) Farm employed, female	(5) Non-farm employed, total	(6) Non-farm employed, female
PostGQ*GQ (0–40)	0.0093	0.0161	-0.747	-1.555	1.21	1.732
	(0.017)	(0.023)	(1.121)	(1.486)	(1.111)	(1.476)
PostNSEW*NSEW (0–40)	0.011	-0.0135	-1.256	-2.351	1.315	2.618*
	(0.020)	(0.027)	(1.120)	(1.485)	(1.110)	(1.475)
State year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,602	1,602	834	834	834	834
R-squared	0.771	0.669	0.988	0.985	0.988	0.985
F statistic	7.672	4.618	72.68	56.57	73.92	56.5

Note: 2SLS = two-stage least squares; GQ = Golden Quadrilateral; NSEW = North- South-East-West Highway.
* p<0.1

Table 13. 2SLS Estimation of the Average Impact of Highways on Environmental Outcomes

Variables	(1) Aerosol particle radius (in logs)	(2) Aerosol optical thickness (thickness scale 0–1)	(3) Nitrogen dioxide levels (billion molecules/mm^2) in logs
PostGQ*GQ (0–40)	1.384	0.0378***	3.714
	(1.547)	(0.011)	(3.279)
PostNSEW*NSEW (0–40)	-1.619	0.0171	-5.034
	(1.546)	(0.011)	(3.275)
State year fixed effects	Yes	Yes	Yes
District fixed effect	Yes	Yes	Yes
Observations	834	834	834
R-squared	0.969	0.983	0.988
F statistic	27.37	50.47	76

Note: GQ = Golden Quadrilateral; NSEW = North-South-East-West Highway.
*** p<0.01

Table 14. OLS Estimation of Impacts Conditional on Market Variables for Welfare Outcomes

Market variable	Interacted treatment	Mean household consumption (in logs)	GDP per capita (current USD, in logs)	Percentage households above the poverty line, total	Percentage households above the poverty line, rural	Percentage households above the poverty line, urban
Households' access to banking services	GQ	-0.0031*		-0.0032*	-0.0041**	
Access to financial services, private non-agricultural enterprises	GQ			0.0096	0.0116	
Literacy rate, 7+ years (percent of population group)	GQ					
Secondary education completion rate,15+ years, total (percent of population group)	GQ					
Cropland (percent of area)	GQ	-0.0017**		-0.0007		-0.001
Food/beverage/tobacco/ manufacturing (percent of establishments)	GQ	0.0015	-0.0012	0.0024**	0.0026**	0.0018
Diversification index of non-agricultural enterprises, ISIC 3.1 2-digit (index)	GQ					
Non-agricultural enterprises by ownership, privately owned	GQ	0.0128**	-0.0051	0.007	0.0072	
Households' access to banking services	NSEW		-0.0028			
Access to financial services, private non-agricultural enterprises	NSEW	-0.0275*		-0.0195	-0.0126	-0.0225
Literacy rate, 7+ years (percent of population group)	NSEW	0.0032	0.0054**			0.0037
Secondary education completion rate,15+ years, total (percent of population group)	NSEW					
Cropland (percent of area)	NSEW	-0.0015	0.0022**			
Food/beverage/tobacco/ manufacturing (percent of establishments)	NSEW					0.002
Diversification index of non-agricultural enterprises, ISIC 3.1 2-digit (index)	NSEW					
Non-farm enterprises by ownership, privately owned	NSEW					
Number of observations		1661	854	1661	1661	1626
Adjusted R-squared		0.3	0.8	0.3	0.4	0

Note: GQ = Golden Quadrilateral; ISIC = International Standard Industrial Classification; NSEW = North-South-East-West Highway; USD = U.S dollars.
** p<0.05, * p<0.1

Table 15. OLS Estimation of Impacts Conditional on Market Variables for Labor Market Inclusion and Structural Change Outcomes

Market variable	Interacted treatment	Regular wage employed, total	Regular wage employed, female	Farm employed, total	Farm employed, female	Non-farm employed, total	Non-farm employed, female
Households' access to banking services	GQ		-0.0019*				
Access to financial services, private non-agricultural enterprises	GQ				0.436		-0.427
Literacy rate, 7+ years (percent of population group)	GQ	-0.0011		0.1379**		-0.1712**	
Secondary education completion rate,15+ years, total (percent of population group)	GQ			-0.3385***	-0.3819***	0.3463***	0.3781***
Cropland (percent of area)	GQ	-0.0008***			0.058*		-0.0602**
Food/beverage/tobacco manufacturing (percent of establishments)	GQ	-0.0005			-0.0748		0.0567
Diversification index of non-agricultural enterprises, ISIC 3.1 2-digit (index)	GQ		-0.0083				
Non-agricultural enterprises by ownership, privately owned	GQ		-0.005	0.6972***	1.2225***	-0.6264***	-1.2085***
Households' access to banking services	NSEW	-0.0018					
Access to financial services, private non-agricultural enterprises	NSEW		-0.0197**		-1.3426**		1.2659**
Literacy rate, 7+ years (percent of population group)	NSEW	0.0037***	0.0042**				
Secondary education completion rate,15+ years, total (percent of population group)	NSEW				-0.2408*		0.245*
Cropland (percent of area)	NSEW						
Food/Beverage/Tobacco Manufacturing (percent of establishments)	NSEW	-0.0009	-0.001	0.0588			
Diversification index of Non-agricultural enterprises, ISIC 3.1 2-digit (index)	NSEW	-0.0124				0.8233	
Non-agricultural enterprises by	NSEW			-0.32	-0.5972	0.56*	0.7291*

ownership, privately owned						
Number of observations	1626	1626	854	854	854	854
Adjusted R-squared	0	-0.2	0	0.3	0	0.3

Note: GQ = Golden Quadrilateral; ISIC = International Standard Industrial Classification; NSEW = North-South-East-West Highway.

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

Table 16. OLS Estimation of Impacts Conditional on Market Variables for Environmental Outcomes

Market variable	Interacted treatment	Aerosol particle radius (percent of small particles) in logs	Aerosol optical thickness (thickness scale 0–1)	Nitrogen dioxide levels (billion molecules/mm^2) in logs
Households' access to banking services	GQ	-0.1558**		
Access to financial services, private non-agricultural enterprises	GQ		0.0034	
Literacy rate, 7+ years (percent of population group)	GQ	0.2869***		
Secondary education completion rate,15+ years, total (percent of population group)	GQ		-0.0015**	
Cropland (percent of area)	GQ		0.0003	
Food/beverage/tobacco manufacturing (percent of establishments)	GQ			
Diversification index of non-agricultural enterprises, ISIC 3.1 2-digit (index)	GQ	1.086**	-0.0076*	1.9606
Non-agricultural enterprises by ownership, privately owned	GQ			
Households' access to banking services	NSEW		-0.0012	
Access to financial services, private non-agricultural enterprises	NSEW		0.0066	1.363
Literacy rate, 7+ years (percent of population group)	NSEW		0.001	-0.5962**
Secondary education completion rate,15+ years, total (percent of population group)	NSEW			
Cropland (percent of area)	NSEW	-0.1641***	0.0008**	
Food/beverage/tobacco manufacturing (percent of establishments)	NSEW	-0.1569***		
Diversification index of non-agricultural enterprises, ISIC 3.1 2-digit (index)	NSEW			
Non-agricultural enterprises by ownership, privately owned	NSEW		-0.0048	1.2378
Number of observations		854	854	854
Adjusted R-squared		0.6	0.6	0.6

Note: GQ = Golden Quadrilateral; ISIC = International Standard Industrial Classification; NSEW = North-South-East-West Highway.
*** p<0.01, ** p<0.05, * p<0.1

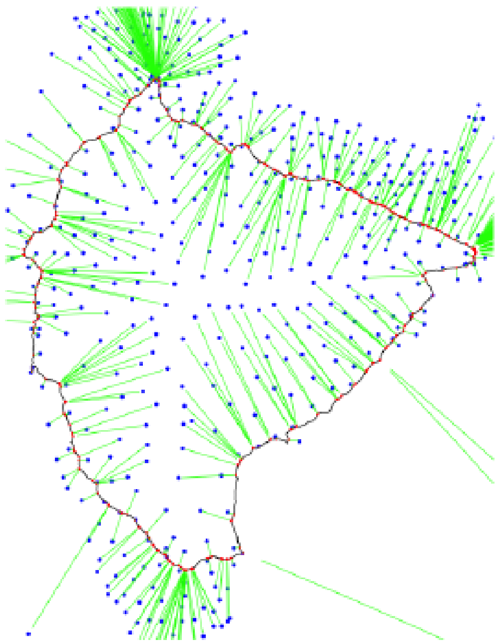
Table 17. Conditional Impacts on Labor Market Inclusion and Structural Change Outcomes with Governance Interactions Included

Market variable	Interacted treatment	Regular wage employed, total	Regular wage employed, female	Farm employed, total	Farm employed, female	Non-farm employed, total	Non-farm employed, female
Households' access to banking services	GQ		-0.0027***				
Access to financial services, private non-agricultural enterprises	GQ				0.6323**		-0.6441**
Literacy rate, 7+ years (percent of population group)	GQ	-0.0014		0.1707**		-0.1952**	
Secondary education completion rate,15+ years, total (percent of population group)	GQ			-0.3131***	-0.4516***	0.3254***	0.4507***
Cropland (percent of area)	GQ	-0.0004			0.0292		- 0.0301
Food/beverage/tobacco/ manufacturing (percent of establishments)	GQ		-0.0244***				
Diversification index of Non-agricultural enterprises, ISIC 3.1–3.2 digit (index)	GQ		-0.0062	0.5971***	1.2805***	-0.5135**	- 1.2697***
Non-agricultural enterprises by ownership, privately owned	GQ	-0.0002			-0.0484		0.0288
Governance Index	GQ	0.1351**	0.2282***	-1.3668	6.311	0.3746	-6.6951
Households' access to banking services	NSEW	-0.0007					
Access to financial services, private non-agricultural enterprises	NSEW		-0.0169**		-1.1709**		1.0624*
Literacy rate, 7+ years (percent of population group)	NSEW	0.0033**	0.0042**				
Secondary education completion rate,15+ years, total (percent of population group)	NSEW				-0.157		0.1408
Cropland (percent of area)	NSEW						
Food/beverage/tobacco manufacturing (percent of establishments)	NSEW	-0.007.2				1.0729*	
Diversification index of non-agricultural enterprises, ISIC 3.1 2-digit (index)	NSEW			-0.383	-0.5704	0.7046**	0.7212*
Non-agricultural enterprises by ownership, privately owned	NSEW	-0.0005	-0.0006	0.0504			
Governance index	NSEW	-0.1582	-0.0341	-6.037	0.1802	2.9084	1.5415
Number of observations		1626	1626	854	854	854	854
Adjusted R-squared		0	-0.2	0	0.3	0	0.3

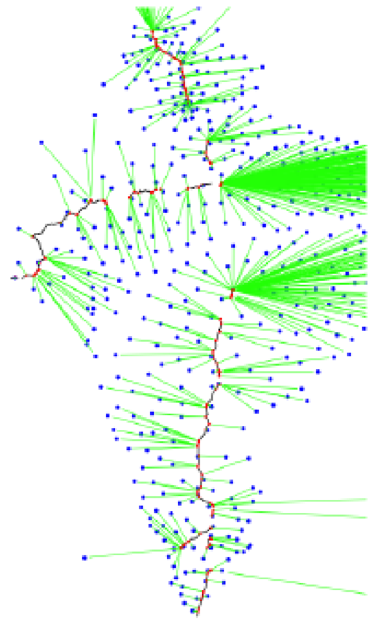
Note: This table reports regressions estimate with year fixed effects instead of state-year fixed effects. GQ = Golden Quadrilateral; International Standard Industrial Classification; NSEW = North-South-East-West Highway.
*** p<0.01, ** p<0.05, * p<0.1]

Figure 1. GQ and NSEW Highway Networks, with Distances from District Centroids

a. Golden Quadrilateral (GQ)



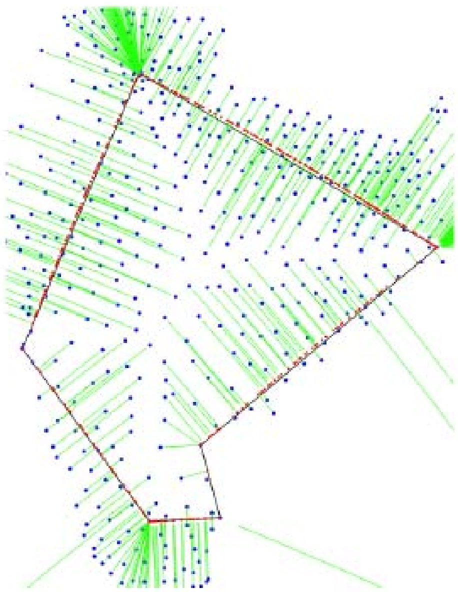
b. North-South-East-West Highway



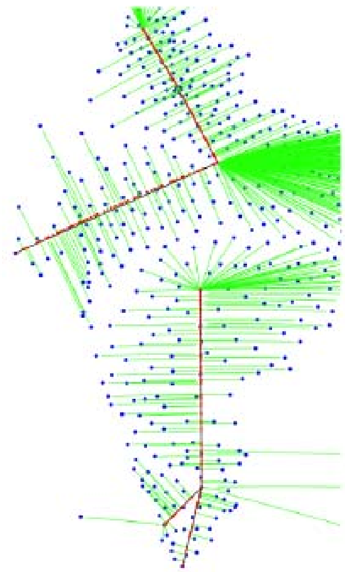
Note: GQ = Golden Quadrilateral; NSEW = North-South-East-West Highway. The green lines depict the shortest distance from district centroids (blue dots) to the actual highway placement (in red).

Figure 2. Straight-Line Versions of GQ and NSEW Highways Networks, with Distances from District Centroids

a. Golden Quadrilateral (GQ)



b. North-South-East-West Highway



Note: The green lines depict the shortest distance from district centroids (blue dots) to the straight line connecting the nodal cities (in red)—south west arm for the GQ adjusted with a kink not to cross the ocean.

Appendix

Table A.1. List of Data Source Codes

Source code	Name
A	MODIS Land Cover Type I product (MODIS). Information and images obtained from National Aeronautics and Space Administration (NASA) Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, https://lpdaac.usgs.gov/data_access.
B	National Aeronautics and Space Administration (NASA) Earth Observations (NEO-ND), http://neo.sci.gsfc.nasa.gov/view.php?datasetId=AURA_NO2_M.
C	National Sample Survey (NSS) Labour Force Survey
D	Census of India–Primary Census Abstract (PHC–PCA), Office of the Registrar General and Census Commissioner, Government of India, http://censusindia.gov.in.
E	State-wise Gross District Domestic Product (DDP), Directorate of Economics and Statistics, Planning Commission, Government of India, http://planningcommission.nic.in/plans/stateplan/index.php?state=ssphdbody.htm.
F	DSMP-OLS Radiance Calibrated Nighttime Lights (RCNTL). National Oceanic and Atmospheric Administration's National Geophysical Data Center (NGDC), http://ngdc.noaa.gov/eog/dmsp/download_radcal.html.
G	Census of India–House Listing and Housing Census (PHC–HH), Office of the Registrar General and Census Commissioner, Government of India, http://censusindia.gov.in.
H	Economic Census (EC) of 2005, Central Statistical Office, Ministry of Statistics & Programme Implementation of India, http://164.100.34.62/index.php/catalog/21/.
I	The Household Consumption Expenditure Survey of National Sample Survey (NSS–HCE), National Sample Survey Office (NSSO), the Ministry of Statistics and Programme Implementation, Government of India, http://mospi.nic.in/Mospi_New/site/inner.aspx?status=2&menu_id=71.
J	Annual Survey of Industries (ASI), Central Statistical Office (Industrial Statistics Wing), the Ministry of Statistics and Programme Implementation, Government of India, http://mospi.nic.in/Mospi_New/site/inner.aspx?status=2&menu_id=92.

Table A.2. Nodal Districts

State	District	GQ distance	NSEW distance	GQ straight-line distance	NSEW straight-line distance
Gujarat	Porbandar	317.5313	1.886075	405.7717	0
Haryana	Faridabad	0.010304	31.76897	18.71244	1.205224
Haryana	Gurgaon	0.829699	41.30815	13.04913	35.2665
Karnataka	Bangalore	0.220969	0.220968	0	67.73664
Karnataka	Bangalore rural	21.7138	8.152857	16.66858	69.38587
Kerala	Ernakulam	322.5242	88.827	338.6767	0
Maharashtra	Mumbai	3.37185	386.4158	0	405.7835
Maharashtra	Mumbai suburban	2.350186	374.9798	2.743553	391.7226
NCT of Delhi	Central	1.125572	1.739703	3.797182	2.927443
NCT of Delhi	East	4.644253	4.744387	6.357544	8.826364
NCT of Delhi	New Delhi	0.88681	4.325667	0	0
NCT of Delhi	North	6.318545	0.884115	11.2142	5.506637
NCT of Delhi	North East	6.524806	5.371046	12.4347	10.39161
NCT of Delhi	North West	18.12909	6.215507	21.45096	2.096619
NCT of Delhi	South	8.611861	13.71957	6.114379	3.935371
NCT of Delhi	South West	11.10865	23.30152	15.79743	18.6445
NCT of Delhi	West	13.10197	13.69143	15.16745	10.22288
Punjab	Jalandhar	326.8173	10.6747	330.6815	0
Punjab	Kapurthala	347.822	24.04326	351.6391	21.75046
Tamil Nadu	Chennai	0.964845	228.1311	0	221.3636
Tamil Nadu	Kanniyakumari	478.2926	24.44338	517.4671	0
Tamil Nadu	Salem	93.34125	10.74619	144.6043	0
Tamil Nadu	Thiruvallur	21.39638	203.1678	20.25113	188.3025
Uttar Pradesh	Gautam Buddha Nagar	19.5113	37.03685	3.616564	18.94192
Uttar Pradesh	Jhansi	114.0565	39.72296	217.3647	0

West Bengal	Haora	5.959926	940.6368	16.69946	984.182
West Bengal	Jalpaiguri	389.6221	1059.672	421.8346	1020.105
West Bengal	Kolkata	5.347588	970.4988	0	1012.148

Note: GQ = Golden Quadrilateral; NCT = National Capital Territory. NSEW = North-South-East-West Highway.

Table A.3. OLS Estimation of Average Impact of Highways on Welfare Outcomes

Variables	(1) Mean household consumption (in logs)	(2) GDP per capita (current USD, in logs)	(3) Percentage households above the poverty line, total	(4) Percentage households above the poverty line, rural	(5) Percentage households above the poverty line, urban
PostGQ*GQ (0–40)	0.00282	0.0329**	-0.0024	-0.0038	0.0112
	(0.018)	(0.014)	(0.015)	(0.017)	(0.019)
PostNSEW*NSEW (0–40)	-0.0092	0.00202	-0.0015	0.00304	0.0215
	(0.027)	(0.019)	(0.024)	(0.026)	(0.029)
State year fixed effects	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	1,637	834	1,637	1,637	1,602
R-squared	0.802	0.993	0.796	0.806	0.683

Note: The ordinary least squares (OLS) estimation is based on a distance dummy on the proximity of district centroid to highways [GQ (0–40) = 1 when district centroid is within 40 km of highway] interacted with post-treatments. All the regressions control for state-year and district fixed effects. GQ = Golden Quadrilateral; NSEW = North-South-East-West Highway; USD = U.S. dollars.

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

Table A.4. Average Impact of Highways on Inclusion and Labor Market Structure Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	Regular wage employed, total	Regular wage employed, female	Farm employed, total	Farm employed, female	Non-farm employed, total	Non-farm employed, female
PostGQ*GQ (0–40)	0.00197	-0.0083	-1.200**	-2.060***	1.336**	2.011***
	(0.007)	(0.010)	(0.587)	(0.779)	(0.582)	(0.774)
PostNSEW*NSEW (0–40)	0.0137	-0.0002	-2.217***	-2.116**	2.291***	2.236**
	(0.011)	(0.015)	(0.784)	(1.041)	(0.777)	(1.035)
State year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,602	1,602	834	834	834	834
R-squared	0.771	0.671	0.988	0.985	0.988	0.985

Note: The OLS estimation is based on a distance dummy on the proximity of district centroid to highways (GQ (0–40) =1 when district centroid is within 40 km of highway) interacted with post-treatments. All the regressions control for state-year and district fixed effects. GQ = Golden Quadrilateral; NSEW = North-South-East-West Highway.

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

Table A.5. Average Impact of Highways on Environmental Outcomes

	(1)	(2)	(3)
Variables	Aerosol particle radius (in logs)	Aerosol optical thickness (thickness scale 0–1)	Nitrogen dioxide levels (billion molecules/mm^2) in logs
PostGQ*GQ (0–40)	-1.125	0.0187***	0.574
	(0.800)	(0.006)	(1.708)
PostNSEW*NSEW (0–40)	-0.742	0.00928	-2.386
	(1.069)	(0.007)	(2.283)
State year fixed effects	Yes	Yes	Yes
District fixed effect	Yes	Yes	Yes
Observations	834	834	834
R-squared	0.969	0.983	0.989

Note: The ordinary least squares (OLS) estimation is based on a distance dummy on the proximity of district centroid to highways [GQ (0–40) =1 when district centroid is within 40 km of highway] interacted with post-treatments. All the regressions control for state-year and district fixed effects. GQ = Golden Quadrilateral; NSEW = North-South-East-West Highway.
*** p<0.01

References

ADB, DFID, JICA, and WBG (Asian Development Bank, UK Department for International Development, Japan International Cooperation Agency, and World Bank Group). Forthcoming. *The WEB of Transport Corridors in South Asia*. Washington, DC: World Bank.

Banerjee, A., E. Duflo, and N. Qian. 2012. "On the Road: Transportation and Infrastructure Growth in China." NBER Working Paper 17897, National Bureau of Economic Research, Cambridge, MA.

Berg, C. N., U. Deichmann, Y. Liu, and H. Selod. 2017. "Transport Policies and Development." *The Journal of Development Studies* 53 (4): 465–80.

Bosker, E. M., U. Deichmann, and M. Roberts. 2015. "Hukou and Highways: The Impact of China's Spatial Development Policies on Urbanization and Regional Inequality." Policy Research Working Paper 7350, World Bank, Washington, DC.

Chatterjee, U., R. Murgai, A. Narayan, and M. Rama. 2016. *Pathways to Reducing Poverty and Sharing Prosperity in India: Lessons from the Last Two Decades*. New Delhi: World Bank.

Cropper, M., J. Puri, C. Griffiths, E. B. Barbier, and J. C. Burgess. 2001. "Predicting the Location of Deforestation: The Role of Roads and Protected Areas in North Thailand." *Land Economics* 77 (2): 172–86.

Damania, R., and D. Wheeler. 2015. "Road improvement and Deforestation in the Congo Basin Countries." Policy Research Working Paper 7274, World Bank, Washington, DC.

Dasgupta, S., and D. Wheeler. 2016. "Minimizing Ecological Damage from Road Improvement in Tropical Forests." Policy Research Working Paper 7826, World Bank, Washington, DC.

Datta, S. 2012. "The Impact of Improved Highways on Indian Firms." *Journal of Development Economics* 99 (1): 6–57.

Ghani, E., A. G. Goswami, and W. R. Kerr. 2016. "Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing," *Economic Journal*, Royal Economic Society, 126 (591): 317–57.

India, Ministry of Finance. 2017. Economic Survey 2016–17.

Kone, Z. L., M. Liu, A. Mattoo, C. Ozden, and S. Sharma. Forthcoming. "Internal Borders and Migration in India." *Journal of Economic Geography*.

Laird, J. J., and A. J. Venables. 2017. "Transport Investment and Economic Performance: A Framework for Project Appraisal." *Transport Policy* 56 (2017): 1–11.

Melecky, M. 2017. "Appraisal Econometrics for Proposed Transport Corridors: Optimal Placement, Intervention Design, and Wider Economic Benefits." Policy Research Working Paper 8269, World Bank, Washington, DC.

Melo, P. C., D. J. Graham, and R. Brage-Ardao. 2013. "The Productivity of Transport Infrastructure Investment: A Meta-analysis of Empirical Evidence." *Regional Science and Urban Economics* 43 (5): 695−706.

Mundle, S., S. Chowdhury, and S. Sikdar, 2016. "Governance Performance of Indian States: 2001–02 and 2011–12," Working Paper Number 164, National Institute of Public Finance and Policy (NIPFP), Delhi.

Redding, S. J., and M. A. Turner. 2014. "Transportation Costs and the Spatial Organization of Economic Activity." NBER Working Paper 20235, National Bureau of Economic Research, Cambridge, MA.

Roberts, M., M. Melecky, T. Bougna, and Y. Xu. 2018. "The Estimated Wider Economic Benefits of Transport Corridors: A Critical Review of the Literature." World Bank Policy Research Paper 8302, World Bank, Washington, DC.