THE
ECONOMIC
JOURNAL

The Economic Journal, **130** (*July*), 1173–1199 DOI: 10.1093/ej/ueaa013

THE ECOLOGICAL IMPACT OF TRANSPORTATION INFRASTRUCTURE*

Sam Asher, Teevrat Garg and Paul Novosad

There is a long-standing debate over whether new roads unavoidably lead to environmental damage, especially forest loss, but causal identification has been elusive. Using multiple causal identification strategies, we study the construction of new rural roads to over 100,000 villages and the upgrading of 10,000 kilometers of national highways in India. The new rural roads had precisely zero effect on local deforestation. In contrast, the highway upgrades caused substantial forest loss, which appears to be driven by increased timber demand along the transportation corridors. In terms of forests, last mile connectivity had a negligible environmental cost, while expansion of major corridors had important environmental impacts.

Does human economic progress have an unavoidable environmental cost? This is a central question for policymakers pursuing sustainable development and has been a long-standing debate in both the conservation and economics literatures (Arrow *et al.*, 1995; Grossman and Krueger, 1995; Stern *et al.*, 1996; Andreoni and Levinson, 2001; Foster and Rosenzweig, 2003; Dasgupta, 2007; Alix-Garcia *et al.*, 2013). A key pillar of economic development is large-scale investment in transportation infrastructure that reduces the costs of moving goods and people across space. Concern has been expressed about the potential environmental cost of such investments, and of increased trade more generally (Copeland and Taylor, 1994; Antweiler *et al.*, 2001; Copeland and Taylor, 2004; Frankel and Rose, 2005), but researchers have struggled to identify causal estimates of the impact of transportation infrastructure on local environmental quality.

The most omnipresent of transportation investments are roads. We focus on the impact of road construction and expansion on forest loss as it is among the primary environmental concerns associated with new road construction. Forest cover loss is both globally and locally important, generating global greenhouse emissions (IPCC, 2007; Jayachandran *et al.*, 2017 and local health externalities Bauch *et al.*, 2015; Garg, 2019). Analysis by the IPCC suggests that restoring and protecting forests could yield almost a sixth of the emission mitigation required to prevent runaway climate change by 2030 (IPCC, 2019).

Because of the high cost and high expected return of roads, their placement typically depends on various economic and political factors, making causal identification of their impacts difficult. For example, new roads may be targeted to regions with expanding agricultural land use; these roads

---

* Corresponding author: Paul Novosad, Department of Economics, Dartmouth College, Hanover, NH 03755, USA. Email: paul.novosad@dartmouth.edu

This paper was received on 29 January 2018 and accepted on 31 May 2019. The Editor was Nezih Guner.

The data and codes for this paper are available on the Journal website. They were checked for their ability to replicate the results presented in the paper.

We are grateful for feedback on earlier versions of this paper from Prashant Bharadwaj, Jenn Burney, Jonah Busch, Eric Edmonds, Paul Ferraro, Gordon Hanson, Maulik Jagnani, Erzo Luttmer, Erin Mansur, Gordon McCord, and seminar participants at UC San Diego, UC Irvine, San Diego State University, Claremont McKenna College, George Washington University, the 2018 Pacific Development Economics Conference, the 2018 Annual Meetings of the Agricultural and Applied Economics Association and the 2017 CU Boulder Environmental and Energy Economics Workshop. Ryu Matsuura provided excellent research assistance. Garg acknowledges funding from the Center for Global Transformation at UCSD, ESRC Centre for Climate Change Economics and Policy, the Grantham Foundation for the Protection of the Environment.

may be a response to activities that are already causing forest cover reduction, making it difficult to isolate the direct impact of the roads. While many earlier studies have documented changes in forest cover following the construction of new roads, none have addressed the endogeneity of road placement beyond the inclusion of control variables and in a few cases, location fixed effects. Further, most of these studies have focused on large highways built into the Amazon rainforest (Pfaff, 1999; Pfaff *et al.*, 2007; Weinhold and Reis, 2008); while these highways are important in terms of potential deforestation, their impacts are of uncertain relevance for the set of potential rural roads and highways that policymakers in developing countries are considering today. The majority of road projects in the decades ahead are likely to be last-mile roads to people not currently connected to the road network and upgrades of existing transportation corridors into modern highways.

In this article, we take advantage of a validated satellite-based measure of forest cover (Vegetation Continuous Fields [VCF]), which makes it possible to study the impacts of two large-scale transportation projects in India. The first of these was an initiative to upgrade two major transportation corridors: the 6,000 km 'Golden Quadrilateral' network (GQ) connecting the country's four largest cities, and the comparably-sized 'North–South and East–West' network (NS–EW) connecting the country's four cardinal endpoints in a cross. Both corridors were already used for cross-city transportation before 2000, but over the following 15 years they were upgraded into world class divided highways. The second project was a rural road construction programme, under which over 100,000 new paved rural feeder roads were built, five kilometers in length on average, providing new connections to over 100 million rural residents. Each project has exceeded $10 billion in cost to date and has caused a significant reallocation of local economic activity (Ghani *et al.*, 2016; Asher and Novosad, 2020).

Theoretically, the effect of road investments on local forest cover can be positive or negative. New roads can increase forest cover loss by: (*i*) providing external markets for forest resources, especially timber and firewood, (*ii*) providing external markets for agricultural products, motivating extensification of agriculture into forested land, and (*iii*) increasing the value of land for settlement and industry, resulting in forest clearing. On the other hand, paved roads could also reduce forest cover loss by (*i*) improving local household and industry access to substitutes for local forest resources, especially firewood, and (*ii*) providing access to external output and labour markets, lowering the relative returns to clearing forests for agricultural land as well as to harvesting other forest products such as firewood. Given the substantially different nature of rural feeder roads and national highways, we can also expect the importance of any of these channels to vary by the type of road.

To evaluate the impact of rural roads, we first use a regression discontinuity (RD) approach, exploiting an implementation rule that discontinuously raised the probability of road construction in villages with a population above an arbitrary threshold. Second, we use a difference-in-differences specification that exploits the exact timing of road construction. Both approaches show zero effects of new roads on forest cover. The estimates are precise; we can rule out gains larger than 0.6% and losses greater than 0.2% in forest cover up to five years after roads are completed. Further, we find zero effects for sample subgroups where we might expect losses to be greater, such as villages with greater baseline forest cover or with very poor or forest-dependent residents. We also find zero change in household firewood use in treated villages. We do identify marginal (0.5%) reductions in forest cover during the road construction period; these reductions are reversed soon after roads are completed, but there is no evidence that forest cover continues to rise. We show that ignoring these construction period effects could lead to biased impact estimates. These roads have no effect on forest cover in spite of significantly altering economic opportunities for people in villages (Asher and Novosad, 2020; Adukia *et al.*, 2020).

Causal identification for impacts of highways is much more difficult than for rural roads, because in almost all cases, new highways are small in number and are built along existing transportation corridors. We take the approach of comparing changes in forest cover in areas that are near and that are far from the new highways. While we do not have data covering the period before the construction of the GQ, the NS–EW highway route provides a plausible counterfactual, in that it is a highway of comparable size and importance that was announced simultaneously and on a similar construction schedule to the GQ, but its construction was pushed back by approximately eight years due to bureaucratic delays. Ghani *et al.* (2016) take a similar approach in comparing these two networks to study the impacts of the GQ on manufacturing activity.[1]

In sharp contrast to rural roads, we find that the highway upgrades have had substantial negative effects on forest cover. Following construction of the GQ, we find a 20% decline in forest cover in a 100 kilometer band around the highway, an effect that persists for at least eight years. We find no change in forest cover along the NS–EW corridor until construction accelerates in 2008, at which point we also observe local forest cover loss. The timing of relative forest loss around the construction of each corridor supports a causal interpretation of these estimates. Because forest cover in India is rising on average during the sample period, these are net effects on forest cover, combining increases in deforestation and reductions in afforestation.

These highways appear to have depleted forest cover by increasing timber demand in their vicinity, which has wide ranging effects into the hinterlands of the transport corridors. Following the construction of the GQ, we find a substantial upward trend break in employment in proximate firms that use timber and wood as primary inputs, as well as employment in logging firms. Additional tests reject the competing mechanisms; there are no increases in agricultural land use or changes in local firewood consumption along the highway corridor.

This article makes three central contributions. First, we generate the first causal estimates of the impact of large scale transportation infrastructure investments on natural resource depletion.[2] In so doing, we contribute to a long literature on the trade-offs and synergies between economic development and environmental conservation.[3]

Second, this is the first article to show that the impact of roads on deforestation is a function of which markets are being connected by those roads. Last-mile rural roads provide connectivity to small local markets, facilitating exits from agriculture but without significantly changing indus-

---

[1] On the impacts of the Golden Quadrilateral on firms in India, see also Datta (2012) and Khanna (2016).

[2] Many studies describe cross-sectional relationships between roads and forest cover or forest loss (Chomitz and Gray, 1996; Angelsen and Kaimowitz, 1999; Pfaff, 1999; Cropper *et al.*, 2001; Geist and Lambin, 2002; Deng *et al.*, 2011; Barber *et al.*, 2014; Li *et al.*, 2015; Dasgupta and Wheeler, 2016). A small number of studies examine forest loss in areas with new roads but do not address the endogeneity of road placement (Pfaff *et al.*, 2007; Weinhold and Reis, 2008). The closest study to ours is ongoing work by Kaczan (2020), who uses a difference-in-differences design similar to our first strategy (but does not look at highways), finding that India's new rural roads marginally increased forest cover. The differences may arise because Kaczan (2020) does not distinguish between construction and post-construction periods, and includes villages that never receive roads as part of the control group. We show in Section 3 that both of these choices may lead to biased treatment effects.

[3] On the general relationship between economic development and the environment, see Den Butter and Verbruggen (1994), Arrow *et al.* (1995), Grossman and Krueger (1995), Stern *et al.* (1996), Andreoni and Levinson (2001), Dasgupta *et al.* (2002), Foster and Rosenzweig (2003) and Stern (2004). On deforestation specifically, see Koop and Tole (1999), Burgess *et al.* (2012), Alix-Garcia *et al.* (2013), and Jayachandran *et al.* (2017). Assunção *et al.* (2017) provide causal evidence that rural *electrification* mitigated forest loss in Brazil. For an exhaustive review on drivers of deforestation, see Busch and Ferretti-Gallon (2017). For a literature review on impacts of highways and rural roads on outcomes other than the environment, see Asher and Novosad (2020).

try's access to forest products (Asher and Novosad, 2020). In contrast, highways dramatically change the geographic distribution of industry (Ghani *et al.*, 2016); in India at least, this appears to have substantial environmental consequences.

Our estimates are particularly relevant as the infrastructure agenda in sub-Saharan Africa and South and Southeast Asia is likely to prioritise exactly the kinds of infrastructure investments that we study here—new feeder roads and expansion of existing corridors—as opposed to the large highways through virgin rainforest that have been the subject of much of the earlier work on roads and deforestation. China's Belt and Road Initiative is the signature example, which aims to promote the construction of large scale highway corridors across Southeast and Central Asia, most of which are expansions of existing roadways (Reed and Trubetskoy, 2019). In sub-Saharan Africa, where only 30% of rural people live within two kilometers of a road, last mile access is a major policy priority (Roberts *et al.*, 2006).

Finally, we raise an important methodological issue in the literature on estimating impacts of infrastructure. Large-scale infrastructure often takes many years to build and involves significant land clearing and economic activity during the construction process. In both our examination of highways and of rural roads, we find that forest loss begins during the construction period; in either case, estimates based strictly on the timing of infrastructure completion would underestimate the environmental impact of roads.

The rest of the article is organised as follows. The next section describes India's rural road and highway construction programmes. In Section 2, we describe the data on forest cover and roads, as well as other secondary data sets used in our analysis. Section 3 presents empirical strategy and results describing the impact of rural roads on deforestation. Section 4 presents the empirical strategy and impacts of highway expansions, and Section 5 concludes.

1. Background: Road Construction Programmes in India

In 1999 and 2000, the Government of India launched two major road construction programmes—one aimed at upgrading several national highway corridors and the other at connecting the remainder of India's population to the road network. Together, these programmes marked the largest expansion of road infrastructure in Indian history and came at a joint cost exceeding $50 billion. This section provides background information on both road construction programme.

1.1. *Rural Roads*

In 2000, the Indian government launched the Pradhan Mantri Gram Sadak Yojana (PMGSY), or the Prime Minister's Village Roads Scheme. The primary objective of the programme was to provide new paved roads to previously unconnected villages, although in practice this also involved upgrading low quality roads in already connected villages. By 2015, over 400,000 kilometers of new roads were built, providing new access to the national road network to over 100 million rural people in over 100,000 villages. Over 70% of new rural roads were routes that terminated in villages.

Rural road construction began toward the end of 2001 and was continuing steadily through the end of the sample period in 2014 (see Online Appendix Figure A1). Villages were selected for roads based on a set of guidelines issued by a national government body, the National Rural Roads Development Authority. Notably, the programme prioritised construction of roads to larger villages; district-level implementation plans were to first target all villages with populations greater than 1,000, followed by villages with a population greater than 500, and finally those with a population greater than 250.[4]

The rules were applied on a state-by-state basis, allowing states to move from one threshold to another on their own timelines. In practice, there were several other prioritisation guidelines and political patronage undoubtedly played a role, so that a village's population relative to the threshold significantly influenced its likelihood of receiving a road but was not definitive. For instance, smaller villages could be connected if they were along the least-cost path between larger prioritised villages, and proximate villages could combine their populations to attain the eligibility thresholds. For more details, see Asher and Novosad (2020) and National Rural Roads Development Agency (2005).

1.2. *National Highways*

In 1999, the Indian government announced a plan to modernise its major highways, the National Highways Development Project. The first component of the project was the upgrading and widening of the GQ highway corridor, so named because it connected the four major cities in India: New Delhi, Mumbai, Chennai and Kolkata. The second component was a similar upgrading of the NS–EW corridor, which would connect the furthest corners of the country from Srinagar in the north to Kanyakumari in the south, and from Porbandar in the west to Silchar in the east. Panel A of Figure 1 shows both highway corridors along with the major cities that were connected by them.

While the GQ and NS–EW projects were commissioned around the same time, the government prioritised the implementation of the GQ and construction of the NS–EW was substantially delayed. Construction on the GQ began in 2001; 80% was completed by 2004 and 95% by 2006. In contrast, by 2006 only 10% of the NS-EW corridor was completed, almost half of which was a set of highways which were shared with the GQ (Ghani *et al.*, 2016). By 2010, 72% of the NS–EW was completed, and 90% was completed by 2015. The delay in the construction of the NS–EW allows us to use the NS–EW corridor as a counterfactual for changes in forest cover in the GQ corridor during and immediately following substantial completion of the GQ.

Before these highways were widened and upgraded, the GQ and NS–EW routes were already significant transportation corridors, but their road quality and congestion were highly variable. The upgrading of these networks dramatically improved their quality and reliability; these were the first major long-distance divided highway networks to be developed in India. The construction of the GQ changed national supply networks and led to a substantial reallocation of manufacturing firms into the GQ corridor (Datta, 2012; Ghani *et al.*, 2016; Khanna, 2016). The economic impact of the NS–EW corridor has so far been little studied due to its delayed completion date.

2. Data

To estimate the effects of new roads on forest cover, we combine five different national data sources. We use a validated high resolution satellite-based measure of forest cover. Data on rural roads come from the administrative implementation data generated by the rural road construction

---

[4] Strictly speaking, the allocation was based on habitation population rather than village population. A habitation is a smaller unit of aggregation than the village; there are between one and three habitations in each village. In practice, habitation populations were pooled to the village level in many cases (see below). We aggregate to the village level because neither additional data nor maps are available at the habitation level.

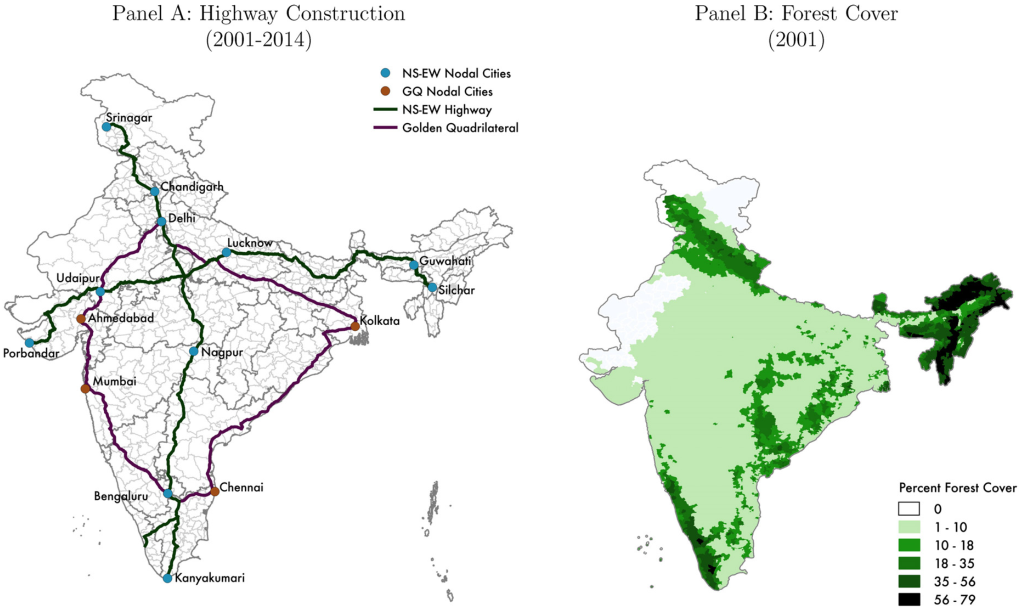


Fig. 1. *Panel A Shows a Map of the GQ and NS–EW Corridor Highways. Panel B Shows a Heat Map of Forest Cover in 2001. Areas Are Shaded According to Average Share of Each Pixel That is Covered by Forest.*

programme, and geographic data on new major highway networks come from national highway maps. While these data sets form the basis of our core specifications, we also use data from the 1991, 2001 and 2011 Population Censuses and third through sixth rounds of the Economic Census to control for location characteristics and explore mechanisms of treatment effects. All of these are census data sets that describe the entire population of India and are geocoded to the village, town and subdistrict levels. This section describes the details of how we prepare and combine all of these data sets. Table 1 shows summary statistics for all variables used.

2.1. *Forest Cover*

Detailed and reliable administrative records on forest cover and deforestation rarely exist, especially in developing countries. Instead, we obtain high resolution time series estimates of forest cover using a standardised publicly-available satellite-based data set. VCF is available at 250 m resolution and provides annual tree cover from 2000–14 in the form of the percentage of each pixel under forest cover (Townshend *et al.*, 2011). For our primary specification, we define forest cover as the total log pixel value plus one in a given geographic area.[5] Results are robust to using the average percentage of forest cover in each village.

The VCF measure is a prediction of the percentage of a pixel that is covered by forest, generated from a machine learning model based on a combination of images from MODIS and samples from higher resolution satellites. The measure employs not only the visible bandwidth but also other bandwidths. For example, VCF uses thermal signatures because forested areas tend to be

---

[5] Results are robust to using the inverse hyperbolic sine transformation instead of log plus one.

Table 1. *Summary Statistics.*

	Mean	SD	Observations
Village-level statistics			
New road before 2011	0.17	0.38	256,885
Road completion year	2007	2	45,338
Population share with no assets (2002)	0.69	0.31	171,249
Population share Scheduled Tribes (2001)	0.22	0.39	256,885
Agricultural share of village land (2001)	0.64	0.28	372,246
Share energy from firewood (2001)	0.67	0.26	409,298
Share energy from imports (2001)	0.07	0.09	409,298
Share energy from local non-wood (2001)	0.26	0.26	409,298
Subdistrict-level statistics			
Average forest cover (2000)	12.76	14.66	4,019
Average forest cover (2014)	14.69	14.49	4,019
Distance to Golden Quadrilateral	218.50	212.33	4,019
Distance to North–South East–West	191.48	155.68	4,019
Employment in wood-using firms	141.41	299.77	4,019
Employment in logging firms	9.78	92.66	4,019

Notes: The table shows summary statistics for the samples used for village- and subdistrict-level analyses. Road completion year is shown only for villages that received new roads between 2001 and 2011. The sample for the first four village-level variables consists of the set of villages that did not have a road at baseline. The sample for agricultural land and energy shares consists of all villages with non-zero forest cover at baseline.

cooler than non-forested plantation areas, allowing VCF to (partially) distinguish between forest cover and plantations. To the extent that thermal signatures and other correlates can distinguish forests from non-forest plantations, VCF substantially improves upon the Normalized Differenced Vegetation Index (NDVI) that has been widely used in understanding the causes of deforestation (for example, Foster and Rosenzweig, 2003). For all analyses, we restrict the sample of villages to those that had non-zero forest cover in 2000, a year predating the construction of all roads considered in this research. This is also the earliest year that these forest cover data sets are available.[6]

Some earlier studies have used the Global Forest Cover (GFC) data set, which describes baseline forest cover in the year 2000, and a binary indicator for the year of deforestation for each 30m × 30m pixel. In the GFC data, a pixel is considered deforested if over 90% of 2000 forest was lost by a given year, or reforested if a pixel goes from zero forest in 2000 to positive forest cover by 2012 (Hansen *et al.*, 2013). While GFC and VCF are both based on satellite imagery, GFC is less useful for the study of forest cover in India, because forest change in India is not well summarised by a binary deforestation indicator. The VCF measures suggest that forest cover rose 15% over the sample period, an estimate consistent with official and international sources. Because most of these gains are in areas that had some pre-existing forest, they are not recorded by GFC. GFC also does not describe partial forest loss, while VCF does. We can replicate GFC estimates by restricting the VCF data to forest losses, but they miss a significant share of forest change in the sample period. Because 92% of villages are larger than the VCF cell size, the resolution advantage of GFC would be minimal. In the cross-section data from the year 2000, VCF and GFC have a correlation coefficient of 0.92 with each other, as compared to respective correlation coefficients of 0.71 and 0.67 with an NDVI measure based on the choices of Foster and Rosenzweig (2003). Online Appendix Figure A2 presents heat maps of forest cover in 2000

---

[6] Fewer than 10% of villages have zero forest cover in 2000; 95% of these villages have less than 1% forest cover in 2014; the mean of forest cover for pixels with non-zero forest is 12.76% in 2000.

according to these three data sets, which convey clearly the similarity between VCF and GFC, and the difference of both of these from NDVI.

We matched forest cover data to the 2011 Population Census village, town and subdistrict boundaries using geographic boundary data purchased from ML InfoMap. In remote parts of India, we received only settlement centroids rather than village boundaries. We generated Thiessen polygons for these villages; all results are robust to excluding this set of villages. Panel B of Figure 1 shows a heat map of baseline forest cover in India. While contiguous areas of very dense forest are geographically concentrated, areas with 20–40% of their land covered by forest are found throughout the country.

2.2. *Rural Roads*

We scraped village-level administrative data describing the construction of rural roads from the programme's online management portal.[7] For each road, the data provide the names of connected villages, the date when the contract for road construction was awarded, and the date of road completion. While data were reported at the sub-village (habitation) level, we aggregated the data to the village level to match our other data sources. We define a village as treated if any habitation in the village was provided with a new road. The data construction and scraping approach is described in detail in Asher and Novosad (2020). The data set describes over 100,000 new roads built between 2001 and 2014; we limit our sample to areas with non-zero forest cover and no paved road in the baseline year, leaving approximately 65,000 new roads in the analysis sample.[8]

2.3. *Highways*

Construction dates and geocoordinates for the GQ and NS–EW corridors were generously shared with us by Ghani *et al.* (2016). We linked these to the village, town and subdistrict polygons described above by calculating straight line distances from polygon centroids to the nearest point on each highway.

2.4. *Population and Economic Censuses*

We matched all villages and towns from the 1991, 2001 and 2011 population censuses using a combination of incomplete keys provided by the Registrar General and a set of fuzzy matching algorithms based on village and town names. The population censuses describe village and town public goods, village amenities (such as schools and medical centres) and household characteristics, including the primary source of cooking fuel. Fuel use is reported as the share of households in a location using firewood (68% of households at baseline), imported fuels (chiefly propane, 8%) or local non-wood fuels (crop residue and dung, 22%) as a primary source of energy. Fuel use is reported at the subdistrict level in 2001 and at the village level in 2011.

The Economic Censuses are complete enumerations of all non-farm establishments undertaken in 1990, 1998, 2005 and 2013, including informal and non-manufacturing firms. We matched

---

[7] The data is publicly available at http://omms.nic.in.

[8] Results are robust to including upgrades and/or villages with no forest cover at baseline. These would be expected to attenuate non-zero treatment effects, thus their exclusion if anything biases us against finding zero effects.

these on village names to the three population censuses using a fuzzy matching algorithm.[9] The Economic Census reports total employment and industry for all firms. We create variables describing total employment in (*i*) firms engaged in logging, and (*ii*) firms whose primary input is raw lumber, which include sawmilling and planing of wood, manufacture of wooden products such as furniture and wooden containers, manufacture of cork, and manufacture of pulp and paper products. The industry categorisation for the 2005 Economic Census places logging firms in the same industry category as firms engaged in the conservation of forest plantations, management of forest tree nurseries and other afforestation categories. We therefore exclude 2005 from analysis of employment in logging firms.

3. Impacts of Rural Feeder Roads on Forest Cover

This section describes the impact of new feeder roads on local deforestation. The main challenge to causal identification of the impacts of rural roads is endogeneity. Because roads are costly to build, their placement is typically correlated with other factors that could also be predictors of deforestation. For example, roads could be targeted to places that are expected to grow or to places that are lagging economically. Road placement may also depend on geographic (e.g., slope, terrain, soil quality) or political factors. Any of these scenarios would bias OLS estimates of the effect of new roads on deforestation.[10] Causal identification of the impact of new roads therefore relies on some kind of variation in road placement or timing that is plausibly exogenous. To study the impact of rural roads, we rely on (*i*) an implementation rule that led to a discontinuity in the probability of a village getting a new road based on arbitrary population cutoffs, and (*ii*) variation in the specific year that a targeted village was treated. We focus our analysis on forest cover in the vicinity of connected villages. Because newly connected rural villages are mostly small and isolated, and because most of the new roads terminate in villages rather than providing new long distance corridors, these roads are unlikely to have had important general equilibrium effects on more distant areas.

3.1. *Rural Roads: Regression Discontinuity Specification*

We begin by exploiting the eligibility rule that prioritised villages for new roads based on arbitrary population thresholds. Given the imperfect compliance with these eligibility rules (described in Section 1), we employ a fuzzy RD design. We limit the RD analysis to states in which administrators adhered closely to population threshold rules.[11]

We use an optimal bandwidth local linear RD specification (Imbens and Lemieux, 2008; Imbens and Kalyanaraman, 2012; Gelman and Imbens, 2019) to identify the change in forest cover caused by a new road at the treatment threshold. We use the following two stage least squares specification:

---

[9] All of the keys matching economic and population census can be downloaded at www.devdatalab.org/shrug (Asher *et al*., 2019).

[10] Online Appendix Table A1 shows estimates from cross-sectional OLS regressions of village-level log forest cover in 2001 on an indicator variable that takes the value one if a village has a paved road in 2001. While the bivariate relationship is strongly negative and highly statistically significant, the estimate gets progressively closer to zero as we add village-level controls and fixed effects, implying substantial selection on observables in the presence of roads. Selection on unobservables is plausibly also important, making the OLS estimates unreliable for causal inference.

[11] We identified these states with the help of officials at NRRDA. They include Chhattisgarh, Gujarat, Madhya Pradesh, Maharashtra, Orissa and Rajasthan. The difference-in-differences analysis below uses all states that built any roads in the sample period.

$$
\begin{aligned}
Treatment_{vds} = \gamma_0 + \gamma_1 \cdot (pop_{vds} \geq T_s) + \gamma_2(pop_{vds} - T_s) \\
+ \gamma_3(pop_{vds} - T_s) \cdot (pop_{vds} \geq T_s) + \nu_d + \boldsymbol{\theta} \boldsymbol{X}_{vds} + \epsilon_{vds}, \qquad (1)
\end{aligned}
$$

$$
\begin{aligned}
Forest_{vds} = \beta_0 + \beta_1 \cdot Treatment_{vds} + \beta_2(pop_{vds} - T_s) \\
+ \beta_3(pop_{vds} - T_s) \cdot (pop_{vds} \geq T_s) + \mu_d + \boldsymbol{\kappa} \boldsymbol{X}_{vds} + \eta_{vds}. \qquad (2)
\end{aligned}
$$

$Forest_{vds}$ is forest cover in village v, district d and state s, and $Treatment_{vds}$ is an indicator equal to one if a new road was built in village v. pop_{vds} is the population of village v and T_s is the treatment threshold used in state s.[12] μ_d and νd are district fixed effects; we find virtually identical results with fixed effects at higher or lower geographic scales. $\boldsymbol{X}_{vds}$ is a control for baseline forest cover; like the fixed effects, the control is unnecessary for identification but improves precision. This is a cross-sectional regression where β_1 identifies the effect of new roads on forest cover in a given year. Outcomes are measured in the final year in the sample data, which is 2013.[13]

Online Appendix Figure A3 shows RD balance tests for a set of variables measured in the baseline period; Online Appendix Table A2 presents the regression estimates on these tests using equation (1). None of the RD estimates are significantly different from zero at baseline. Online Appendix Figure A4 shows that the density of the running variable is continuous around the treatment threshold (McCrary, 2008).

3.2. *Rural Roads: Regression Discontinuity Results*

Figure 2 shows a graphical representation of the RD estimates of the impact of rural roads on forest cover. Panel A shows the first stage; the Y axis shows the share of sample villages that received new roads by 2013 under PMGSY as a function of their population relative to the treatment threshold. Villages above the threshold are about 16 percentage points more likely to receive new roads and the discontinuity is evident. Panel B shows the first stage estimate separately for each outcome year; each point in the figure represents the γ_1 coefficient from equation (1), where the dependent variable takes the value one if a village received a new road by the year indicated on the *X* axis. We can see that roads built before 2007 were not prioritised according to the population threshold rule; the first stage of the RD becomes noticeable after 2008 and continues to rise until 2014.

Panel C of Figure 2 plots village-level log forest cover in 2013 against the population relative to the treatment threshold, in population bins. If roads significantly affected local forest cover, we would expect to see a discontinuity at the treatment threshold analogous to that in Panel A; no such treatment effect is evident. Panel D shows the reduced form treatment effect of above-threshold population on forest cover (β_1) in each year separately; as in Panel B, each point is

---

[12] The treatment threshold varies with state because some states used a threshold of 500 and others were using a threshold of 1,000. States used the lower treatment threshold when they had few villages with population over 1,000 that did not already have roads. Officials at the National Rural Roads Development Agency provided us with information on which states were using which cutoffs, which we then verified in the data. Madhya Pradesh used both the 500 and 1,000 treatment thresholds for roads built in the same period; we include separate fixed effects for the set of villages in the neighbourhood of each threshold. Because the optimal RD bandwidth is close to 100, there is no overlapping between these two groups. Few villages around the lowest population threshold of 250 received roads so we do not use this threshold for analysis.

[13] We find similar results if we pool outcome years from 2010 through 2013 and cluster standard errors at the village level (not shown). Standard errors are slightly smaller with this alternative approach, at the cost of putting more weight on roads which have been built for shorter periods of time.

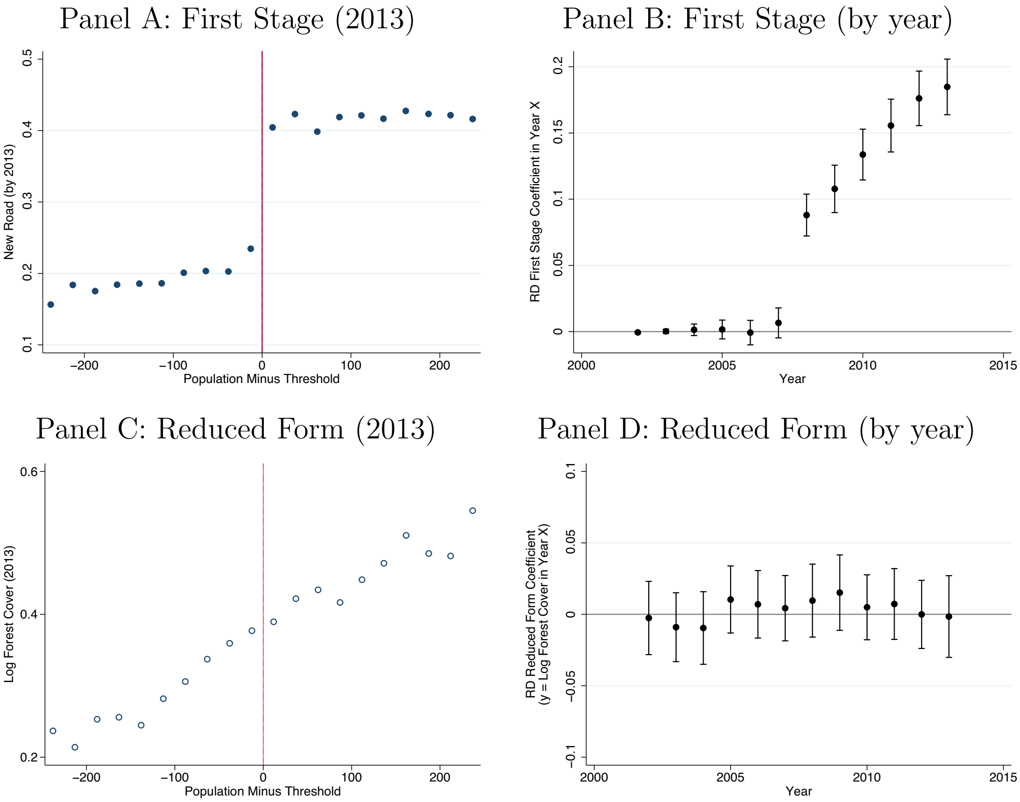


Fig. 2. *RD Estimates of Impact of Rural Roads on Forest Cover.*

Notes: The figure shows RD estimates of the impact of new rural roads on local deforestation. Panel A shows the first stage probability of a village receiving a new road before 2013 as a function of its population relative to the population threshold. Each point shows the mean of the Y variable in a given population bin. Panel B shows the first stage RD estimate of a village receiving a new road by the year indicated on the *X* axis. Each point is an estimate from an RD first stage regression. Panel C is analogous to Panel B; the dependent variable is the log of forest cover in 2013. The points show the mean of this variable in each population bin; population is shown relative to the population treatment threshold. Panel D shows reduced form RD estimates of the impact of being above the population threshold on forest cover in each year on the *X* axis. All estimates in Panels B and D use the same specification as Table 2, and include district-population threshold fixed effects and a control for baseline forest cover.

an estimate from a separate regression, where the dependent variable is the log of forest cover for the year on the *X* axis. If the new rural roads significantly affected forest cover, we would expect to see a change in the coefficient following 2008 when administrators began to adhere to the population implementation rule. Instead, the effect is very close to zero both before and after 2008, indicating that new rural roads had negligible effects on forest cover.

Table 2 shows analogous regression estimates, where the dependent variable is forest cover as measured in 2013. Column 1 shows the first stage estimate of a 16 percentage point increase in the probability of road treatment for villages just above the eligibility threshold. Columns 2 and 3 confirm there is no reduced form effect on either log or average forest cover. Columns 4 through 6 test for treatment effects in villages that might be expected to respond more to new roads. These

Table 2. *RD Estimates of Impact of Rural Roads on Forest Cover.*

	First stage	Reduced form					IV	
	Any road	Log forest	Avg forest	High baseline	High ST	Low assets	Log forest	Avg forest
Above population threshold	0.185*** (0.011)	−0.002 (0.015)	−0.007 (0.106)	−0.011 (0.018)	−0.008 (0.020)	0.007 (0.023)		
New road							−0.008 (0.079)	−0.373 (0.703)
N	22,365	22,365	22,365	11,214	11,174	8,875	22,368	22,368
R^2	0.25	0.81	0.62	0.72	0.83	0.80	0.81	0.42

Notes: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. The table shows RD treatment estimates of the effect of new village roads on local forest cover, estimated with equation (1). In Column 1, the dependent variable is an indicator that takes the value one if a village received a new road in the sample period. Above population threshold is an indicator for a village population being above the treatment threshold. Columns 2 through 6 show reduced form estimates of the effect of being above the treatment population threshold. The dependent variables in Columns 2 and 3 respectively are log village forest cover and average covered share of each village pixel; the data source is VCF. Columns 4 through 6 run the log forest cover specification on subgroups defined respectively by (*i*) above-median forest cover villages, (*ii*) above median share of Scheduled Tribes in a village, and (*iii*) below median baseline village assets. Columns 7 and 8 show IV estimates of the treatment effects of new roads, using respectively log and average forest cover as dependent variables. The outcome variable in Columns 2 through 8 is measured in 2013. All estimates include district-population threshold fixed effects and a control for baseline forest cover.

are: villages with above-median baseline forest cover (Column 4); villages with above-median population shares of constitutionally described ‘backward’ communities (Scheduled Tribes) who often derive livelihoods from forests (Column 5); and villages with below median assets, who might depend more on forests for fuelwood (Column 6). There is no evidence of impacts of roads in any of these groups.[14] Columns 7 and 8 show IV estimates on log and average forest cover. The IV estimates respectively rule out a 0.14 gain and a 0.11 loss in log forest cover with 95% confidence, or approximately a one percentage point change in average forest cover. The average treated village in the sample received a new road in 2008, so these estimates reflect cumulative forest change five years after a village is connected. Results are robust to different controls or fixed effects and different bandwidth choices.[15] Online Appendix Table A5 uses the RD specification to show further that there are no changes in household fuel use following completion of a new road.

3.3. *Rural Roads: Difference-in-Differences Specification*

The RD design estimates causal impacts of roads under minimal assumptions, but is limited to estimating a LATE in the neighbourhood of the treatment threshold in states that closely followed implementation rules on population thresholds. We can make greater use of our data and obtain tighter treatment estimates using a difference-in-differences specification that exploits the differential timing of road treatment in each village. For this empirical test, we limit the sample of villages to those that received a road at *some* point during the road construction programme, and use outcomes in later-treated villages as a control group for villages that were treated earlier.

---

[14] Online Appendix Table A3 shows further that roads do not significantly affect forest cover in villages defined by high or low town distance, market access, nor in villages in subdistricts with above median employment in the logging sector or in industries that are heavy consumers of wood.

[15] Results at many different bandwidths are shown in Online Appendix Table A4.

We specifically estimate the following equation:

$$Forest_{vdt} = \beta_1 \cdot Award_{vdt} + \beta_2 \cdot Complete_{vdt} + \boldsymbol{\alpha}_v + \boldsymbol{\gamma}_{dt} + \boldsymbol{X}_v \cdot \boldsymbol{\nu}_t + \eta_{vdt}. \tag{3}$$

$Forest_{vdt}$ is a measure of forest cover in village v and district d in year t. $Award_{vdt}$ is an indicator that takes the value one for the years where a contract has been awarded for the construction of a road to village v but the road construction is not yet complete. $Complete_{vdt}$ is an indicator that takes the value one for all years following the completion of a new road to village v. We separate these two periods because the road construction process may have effects on forest cover (such as clearing of forested area to make room for the physical placement of roads) that are theoretically distinct from the economic effects of a village having a new road. Village fixed effects ($\boldsymbol{\alpha}_v$) control for all village-level time-invariant unobservables, while district-year fixed effects ($\boldsymbol{\gamma}_{dt}$) control for any pattern of regional shocks.[16] We also interact a vector of baseline village controls $\boldsymbol{X}_v$ (baseline forest cover, village population and distance from the village to the nearest towns) with year fixed effects. These control for any differential time path of forest cover that is correlated with baseline village characteristics. These controls are particularly important because larger villages are more likely to be treated earlier due to programme implementation rules. Standard errors are clustered at the village level to account for serial correlation.

We can interpret β_1 and β_2 as the effects of road construction activities and the effects of new roads, respectively; both coefficients describe outcomes relative to the period before any construction began. We restrict our sample from the universe of villages in India to those that had no road in 2000 and had a road completed during the study period. We do this so as not to compare villages that received new roads with those that did not; the endogeneity problem in such a comparison is severe.[17] Identification rests on the assumption that, among the set of villages that received roads in the sample period, there are no other systematic changes specific to villages in the years that roads were awarded and completed that are not caused by the roads themselves.

3.4. *Rural Roads: Difference-in-Differences Results*

The difference-in-difference estimates of the impact of rural roads on village-level forest cover are summarised by Figure 3. These graphs show the residual of log forest cover—after taking out fixed effects and controls described above—as a function of the number of years elapsed since a road was completed in a given village. Panel A shows all previously-unconnected villages that received new roads between 2001 and 2014. Panel B restricts the set of villages to those with above median forest cover in 2000. We show only four years before and after road construction because wider windows have more variable sample composition across estimates; this occurs because we observe different length of pre- and post-periods for different villages depending on their date of treatment.[18] Two patterns are evident in the figure. First, there is a statistically significant reduction in forest cover approximately two years before road construction is complete. Second, forest cover marginally increases in the four years after road completion recovering some or all of the pre-treatment drop.

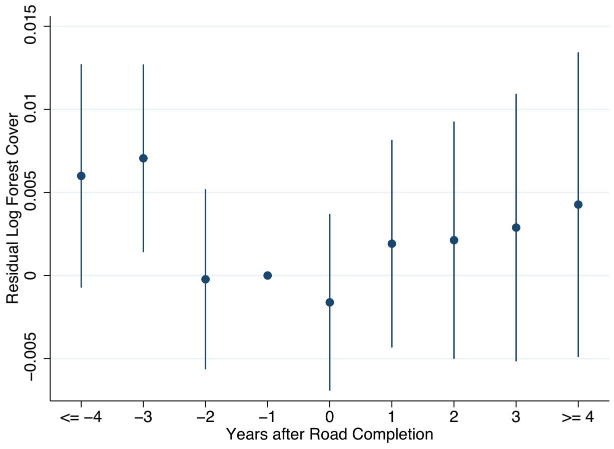
---

[16] Results are unchanged by replacing these with state-year or subdistrict-year fixed effects.

[17] As we show above, a minority of roads were allocated strictly due to the village population thresholds. There are enough of these to estimate an RD test on local compliers, but not enough to assume that all treated villages are selected as good as randomly.

[18] Online Appendix Figure A5 shows a wider time window around treatment; the pattern is the same.

Panel A: Full Sample



Panel B: High Baseline Forest

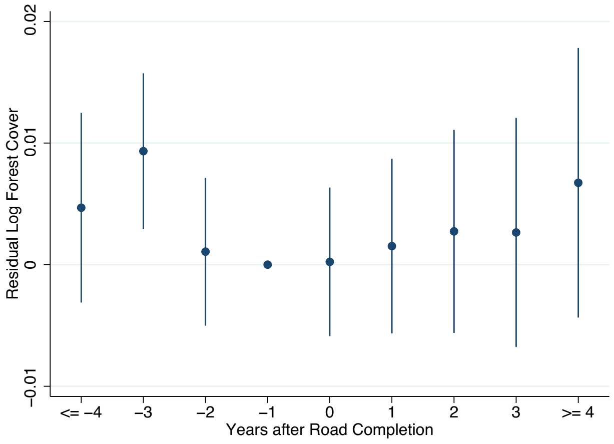


Fig. 3. *Difference-in-Differences Estimates of Impact of Rural Roads on Forest Cover.*
Notes: The figure shows year-by-year estimates of log forest cover in villages that received new roads between 2001 and 2013. Villages are grouped on the X axis according to the year relative to road completion. Each point thus shows the average value of log forest cover in villages in a given year relative to the treatment year, controlling for village fixed effects, district × year fixed effects, baseline population × year and baseline log forest cover × year interactions. Standard errors are clustered at the village level. The year before road completion is omitted ($t = -1$); forest cover is thus shown relative to this period.

Given that these rural roads took one to two years to build, this pattern is consistent with a small degree of forest loss (approximately 0.5%) during the road construction period, with partial or complete recovery afterward. We test this directly in Table 3, which shows estimates from equation (3). Our main estimate in Column 1 shows that villages lose 0.5% of their forest cover during the period between the awarding of a road construction contract and the completion of a road. However, that forest loss is fully restored in the period after the road has been completed;

Table 3. *Difference-in-Differences Estimates of Impact of Rural Roads on Forest Cover.*

	Log forest		Average forest	
	(1)	(2)	(3)	(4)
Award period	−0.005***		−0.033***	
	(0.002)		(0.013)	
Completion period	0.002	0.005***	0.009	0.013
	(0.002)	(0.002)	(0.015)	(0.012)
District-year F.E.	Yes	Yes	Yes	Yes
Village F.E.	Yes	Yes	Yes	Yes
N	688,275	688,275	688,275	688,275
R^2	0.94	0.94	0.92	0.92

Notes: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. The table shows difference-in-differences estimates of the impact of new village roads on local forest cover. We define forest cover as log village forest cover (Columns 1 and 2) and average covered share of each village pixel (Columns 3 and 4); the sample consists strictly of villages that received new roads between 2001 and 2013, and were not accessible by paved road in 2001. Award period is an indicator variable that takes the value one for years after a road contract was awarded and before the road was completed. Completion period is an indicator variable that marks the years after a village's new road was built. All regressions include district × year fixed effects, village fixed effects, baseline population × year fixed effects, and baseline forest × year fixed effects. Standard errors are clustered at the village level to correct for serial correlation.

the estimate of 0.002 log points on the completion indicator can be interpreted as the difference in forest cover between the post-road and the pre-award periods. Relative to the pre-award period, we can rule out gains larger than 0.6% and declines larger than 0.2% in forest cover. In Column 2, we show that failing to account for the award period would lead to the estimation of a marginal forest cover gain of 0.5% because it would incorrectly attribute the construction period loss to the pretrend. This result highlights the importance of accounting for the construction period when studying the environmental impacts of new infrastructure. Columns 3 and 4 present estimates where forest cover is measured as the average share of each pixel that is covered by forest; results are similar. These estimates are based on different lengths of post-construction periods in different villages, but on average they show effects for four years after treatment.[19]

Table 4 shows these estimates along the same dimensions of heterogeneity described above. Effects are broadly similar whether we cut the sample on baseline forest cover, population share of Scheduled Tribes, or asset poverty. There is thus no evidence that our zero results are hiding differential positive and negative effects in different places.[20] It is also unlikely that outmigration of individuals following road construction is significantly biasing our findings; these rural roads are not associated with significant population change at the village level (Asher and Novosad, 2020). Rural-to-urban migration has also been much slower in India over the sample period than in other countries at comparable levels of income.

---

[19] Online Appendix Table A6 shows that these estimates are robust to a range of specifications including the use of village time trends, subdistrict-year fixed effects (instead of district-year fixed effects) and using a limited sample of roads for which we have at least four (or five) years of both pre-treatment and post-treatment data. Online Appendix Table A7 shows additional specifications. Column 1 adds villages that did not receive roads in the sample period, the specification used in Kaczan (2020). Like Kaczan (2020), we find a positive treatment coefficient; however, Column 2 shows that this is not robust to the inclusion of village-specific time trends, indicating that never-treated villages are on different forest cover trends from treated villages. Columns 3 and 4 show that our main estimate is robust to village-specific time trends. Column 5 and 6 define the treated area as a circle around the village with a radius of 5 km and 50 km, respectively; as in the main specification, we find no treatment effects at these radii.

[20] Equally, we find no effects of roads on forest cover when splitting the sample on distance to the nearest town or on market access, nor in subdistricts with above median employment in logging or in industries with high consumption of wood (Online Appendix Table A8).

Table 4. *Rural Roads and Deforestation: Heterogeneity of Difference-in-Differences Estimates.*

	Baseline forest		ST share		Asset poverty	
	High	Low	High	Low	Poor	Not poor
Award period	-0.005^{*}	-0.005^{***}	-0.003	-0.006^{***}	-0.004	-0.006^{***}
	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
Completion period	-0.002	0.002	0.000	0.003	0.001	0.001
	(0.004)	(0.002)	(0.003)	(0.003)	(0.004)	(0.003)
N	341,280	346,455	344,010	343,860	265,470	422,430
R^2	0.86	0.92	0.93	0.95	0.94	0.95

Notes: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. The table shows difference-in-differences estimates of the impact of new village roads on local forest cover, along three dimensions of heterogeneity. Forest cover is defined as log village forest cover; the data source is VCF. Columns 1 and 2 respectively show estimates for villages with above and below median baseline forest cover. Columns 3 and 4 respectively show estimates for villages and above and below median population share of members of Scheduled Tribes. Columns 5 and 6 respectively show estimates for below- and above-median shares of households who report no assets in the 2002 Below Poverty Line. The sample consists strictly of villages that received new roads between 2001 and 2013, and were not accessible by paved road in 2001. Award period is an indicator variable that takes the value one for years after a road contract was awarded and before the road was completed. Completion period is an indicator variable that marks the years after a village's new road was built. All regressions include district × year fixed effects, village fixed effects, baseline population × year fixed effects, and baseline forest × year fixed effects. Standard errors are clustered at the village level to correct for serial correlation.

The panel estimates confirm the finding in the RD analysis, using a different set of villages with a different local average treatment effect; the evidence is clear that new rural roads have had a negligible effect on local forest cover.

4. Impacts of Major Highways on Forest Cover

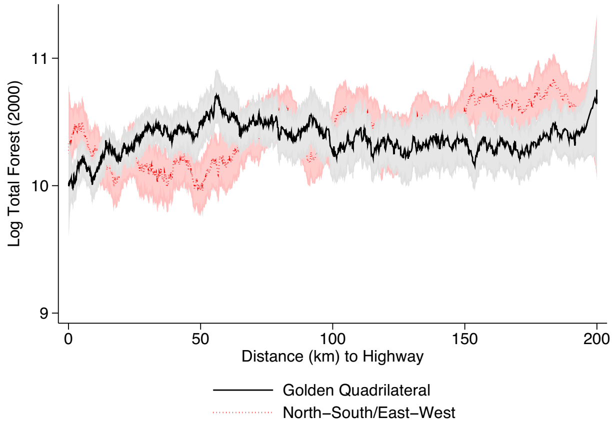
In this section, we aim to identify the causal impact of highways on local forest cover. The identification challenge is that highways are typically built to connect cities with current or anticipated economic growth; if economic growth is correlated with forest cover changes for any reason other than the direct effect of highways, then we cannot interpret the correlation between highways and forests as a causal effect.

We therefore focus on a set of places that happen to be in between the targeted endpoints of India's new highways, as in Ghani *et al.* (2016). Both the GQ and the NS–EW corridors were upgraded with the objective of improving connections between India's major cities and regions; the connection of secondary cities and intermediate places on the route was a secondary priority. Because these intermediate regions were targeted incidentally rather than directly, the placement of the highways is less likely to be driven by existing or anticipated economic growth.

We can further generate a plausible counterfactual that describes how forest cover would have changed in the absence of the highway upgrades. Like the GQ, the NS–EW route was an important transportation corridor in 2000 and was to be upgraded before 2005 as part of NHDP, but the project did not begin in earnest until several years after the GQ was completed. Our main estimates examine forest changes along the GQ corridor during and after the construction years, as compared to regions further from the GQ. We then test for effects along the NS–EW route using a similar specification, showing there are no effects along the second corridor until after 2008, as would be expected given the construction delay.

As a starting point, Figure 4 plots kernel-smoothed local regression estimates of mean forest cover and forest cover change as a function of distance from each highway. Initial forest cover (Panel A) is broadly similar across the two highways. Panel B shows forest cover change from

Panel A: Pre-Expansion Forest Cover (2000)



Panel B: Forest Cover Change (2000-2008)

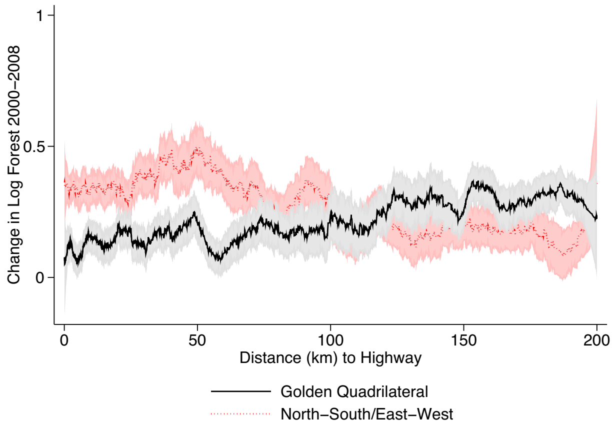


Fig. 4. *Forest Cover and Forest Cover Change Along Highway Corridors (2000–8).*
Notes: Panel A shows a kernel-smoothed regression of log subdistrict forest cover in 2000 on distance to the corridors where the GQ and NS–EW highways will be expanded. Panel B plots kernel-smoothed regression estimates of change in log subdistrict forest cover from 2000 to 2008 against distance to each highway network. By 2008, there was very little construction on the NS–EW corridor, so we treat it here as a control group. The plots display means that are unadjusted for any fixed effects or controls. The shaded areas display 95% confidence intervals.

2000–2008, also by distance to each highway. Relative to the NS–EW (dashed line), forest cover within 100 km of the GQ (solid line) falls substantially between 2000 and 2008. At further distances the effects are similar across the two highways, though there may be smaller relative gains for the GQ. We present this as suggestive evidence of relative forest loss along the GQ corridor during and after its construction. The rest of this section generates formal tests for change, controlling for fixed effects and other factors that may have simultaneously influenced forest change.

4.1. *Highways: Empirical Specification*

The simplest form of the difference-in-differences specification is described by the following equation:

$$Forest_{ist} = \beta_0 + \beta_1 CLOSE_{is} + \beta_2 POST_t + \beta_3 CLOSE_{is} \times POST_t + \epsilon_{ist}. \tag{4}$$

In this specification, i indexes a subdistrict in state s and time t, $CLOSE_{is}$ is an indicator for subdistricts close to the highway, and *POST* indicates years following the completion of the highway. $Forest_{ist}$ is a measure of forest cover in subdistrict i and state s at time t, usually log total forest cover. β_3 describes the differential change in forest between locations that are near and far from the highway network after the highway is built, controlling for the same geographic difference before the highway was built. If new highways cause deforestation, we expect β_3 to be less than zero. We conduct our analysis at the subdistrict level, because subdistricts are contiguous regions that cover the whole of India for which we can calculate a range of demographic and socioeconomic controls. We weight results by subdistrict area.[21] There are approximately 4,000 subdistricts in India.

We extend this simple specification in three ways. First, because we do not have strong priors on which distances are near and which are far, we use a flexible set of distance indicators to non-parametrically identify highway effects at a range of distances. Estimates can still be interpreted as the difference from a given band to the omitted (most remote) distance band. This ensures that our result is not dependent upon a particular definition of closeness. Second, because the construction of India's national highways were multiyear projects, we separate the $POST_t$ indicator into multiple periods to capture construction and post-construction effects. Third, we add a wide set of fixed effects and controls to improve precision and reduce bias from omitted variables. The most flexible estimating equation is:

$$Forest_{ist} = \sum_{d=1}^{D} \sum_{t=2001}^{2014} \beta_{d,t}(DIST_i \in (d^-, d^+), YEAR = t) + \gamma_{st} + \boldsymbol{X}_i \cdot \boldsymbol{\nu}_t + \psi_d + \eta_{ist}. \tag{5}$$

The distance to the highway is divided into D bands, the boundaries of which are indexed by d. We include a distance band fixed effect ψ_d, state-year fixed effect γ_{st} and a vector of subdistrict controls ($\boldsymbol{X}_i$) interacted with year fixed effects ($\boldsymbol{\nu}_t$). The latter control for any differential time path of forest cover that is correlated with baseline subdistrict characteristics. Controls are the same as in equation (3). We include locations up to a distance $D + E$ from the GQ; the outer boundary (D, E) is the omitted distance category against which the other estimates can be compared. Unless otherwise specified, we define (D, E) as the 200–300 km distance band.[22] $\beta_{d,t}$ identifies the change in forest cover from 2000 to year t, at distance range d from the highway, relative to the omitted distance range (D, E). The $\beta_{d,t}$ coefficients can thus be directly interpreted as the effect of highway construction on forest cover after t years. If new highways cause proximate forest cover loss, we expect $\beta_{d,t}$ to take on negative values for low values of d in the periods t after highway construction has begun. For graphs, we include a set of indicator variables $\beta_{d,2000}$ which

---

[21] Results from a town- and village-level analysis with subdistrict clusters deliver nearly identical results. We could in principle conduct analysis at the grid cell level, but this would require imputation for control variables not available at the grid cell level.

[22] Alternate choices of the range of the omitted group, including using the remainder of the country does not appreciably affect our estimates.

describe baseline forest cover as a function of distance from the highway.[23] Standard errors are clustered at the subdistrict level to account for serial correlation. Because the regression above may have hundreds of coefficients, we pool years or distances in different specifications to improve interpretability.

We exclude areas within 200 km of the nodal towns on the highway routes, as we wish to identify effects on intermediate regions rather than at the highway end points, as in Goswami Ghani and Kerr Ghani *et al.* (2016). Estimates of NS–EW treatment effects omit areas that are within 200 km of the GQ as they are plausibly being treated by the other highway network. We do not omit NS–EW regions from the GQ regressions because NS–EW construction had barely begun during the periods of interest for the GQ analysis; however, regression results are not changed by omitting places within 200 km of NS–EW.

4.2. *Highways: Estimates on Forest Cover*

Panel A of Figure 5 plots coefficient estimates from a single estimation of equation (5), with distances from the GQ highway divided into 10 km bands, and years divided into a single pre-construction year (2000), the construction period (2001–4), and two post construction periods (2005–8 and 2009–12). All estimates describe the difference between a given 10 km distance band from the GQ and the omitted category of 250–300 km.[24] The solid black line describes baseline forest cover as a function of distance from the GQ corridor. The remaining lines show that forest cover within 100 km of the GQ declines rapidly during the GQ construction period and then continues to fall in the years following construction. Effects are slightly smaller in the 100–150 km bandwidth, and statistically indistinguishable from zero at a distances greater than 150 km from the highway.

To alleviate the concern that these forest losses are explained by existing trends in forest loss along existing highway corridors, we run the same estimation for subdistricts along the NS–EW corridor and show results in Panel B. As predicted, there are no differential changes in forest cover close to the NS–EW route before 2008. Net forest loss along NS–EW begins in the 2009–12 period, and the distance effects then look similar to those of the GQ.[25] This distance pattern of forest cover loss is similar to that found in the Amazon (Pfaff *et al.*, 2007). Effects along the NS–EW corridor may be slightly smaller than along the GQ both because construction took place slowly and was still incomplete at the end of the sample period in 2014; the network structure of highways mean that the value of any particular segment depends on the completion status

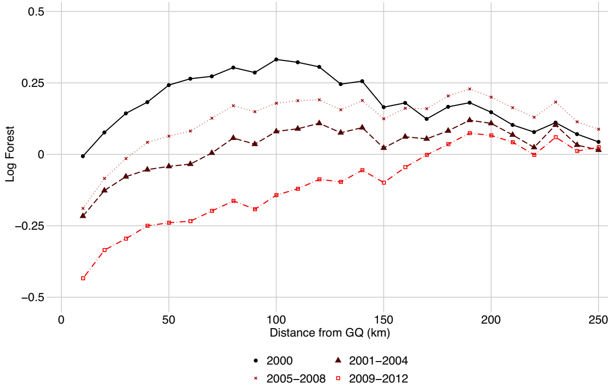
---

[23] We do not include subdistrict fixed effects because we want to generate coefficients on the distance band indicators for the omitted year 2000—these coefficients describe the baseline differences in forest cover between places that were near and far from the highway. However, inclusion of subdistrict fixed effects does not meaningfully change the results. We use state-year fixed effects rather than district-year fixed effects because we wish to test for meaningful effects of distance from highways that may extend beyond the radius of districts. District-year fixed effects would absorb true effects of the GQ that span distances larger than districts. The analysis of Goswami Ghani and Kerr (Ghani *et al.*, 2016) is entirely district level, giving us reason to expect meaningful cross-district effects. As expected, the inclusion of district-year fixed effects attenuates our results slightly but does not change the direction of effects nor eliminate statistical significance.

[24] We include coefficients for the 200–250 km bands in order to plot treatment effects at these ranges. Effects in closer bands are very similar if we restrict the distance indicators to 200 km and use 200–300 km as the omitted group, because there are few differences across years in the 200–250 km range.

[25] Standard errors are omitted in the figure for visual clarity. For the GQ, differences between the 2000 estimates and the 2001–4 estimates are statistically significant at the 1% level for all estimates up to 150 km. For the NS–EW, differences between the 2000 estimates and the 2009–12 treatment estimates are statistically distinguishable at the 1% level until the 180 km estimate.

Panel A: Golden Quadrilateral



Panel B: North-South/East-West

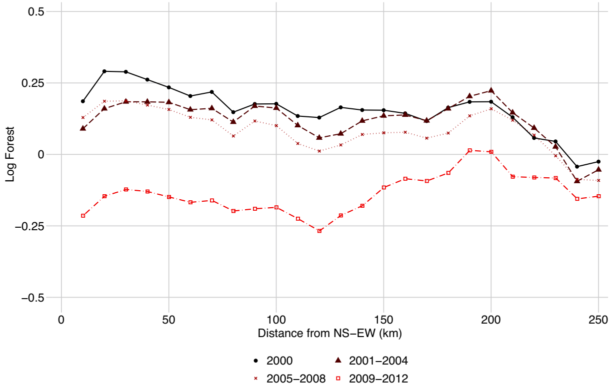


Fig. 5. *Difference-in-Differences Estimates of Impact of Highways on Forest Cover, by Distance Bands.*
Notes: The figure shows point estimates from equation (5), with distance from the GQ highway network (Panel A) and distance from the NS–EW highway network (Panel B) divided into 10 km bands. Each point on the graph shows, for a given set of years (shown in the legend), the average value of log forest cover at a given distance band from the given highway network, relative to the omitted distance band of 290 to 300 km from the highway. All estimates control for state × year fixed effects, baseline population × year and baseline log forest cover × year interactions.

of other segments. Concentration of industry along the GQ corridor may have also reduced the importance of NS–EW as an intercity transportation corridor by the time the NS–EW upgrades took place.

Table 5 presents regression estimates from equation (5), with distances in 50 km bands for legibility. Each estimate describes the difference in forest cover between a given distance band and the omitted category of 200–300 km. Columns (1) and (2) describe the impact of the GQ on forest cover. The top four rows of the table show estimates of construction period impacts on forest cover at various distance bands. Places within 50 km of the new highway network lose 27 log points of forest cover (Column 1) or 1.3 percentage points of forest cover (Column 2, on a base of 7.5%), and the effects shrink at greater distances. The next four rows show similar effects (still relative to year 2000) in the post-construction period of 2005–8. The final four rows

Table 5. *Difference-in-Differences Estimates of Impact of Highways on Forest Cover.*

	GQ (treatment)		NSEW (placebo)	
	Log forest	Avg forest	Log forest	Average forest
GQ construction period × (0–50 km)	-0.265^{***}	-1.306^{***}	-0.038	0.183
	(0.055)	(0.248)	(0.072)	(0.305)
GQ construction period × (50–100 km)	-0.278^{***}	-1.215^{***}	-0.001	0.077
	(0.056)	(0.247)	(0.065)	(0.286)
GQ construction period × (100–150 km)	-0.221^{***}	-1.085^{***}	-0.001	-0.176
	(0.051)	(0.235)	(0.060)	(0.267)
GQ construction period × (150–200 km)	-0.102^{**}	-0.444^{**}	0.023	0.004
	(0.044)	(0.198)	(0.057)	(0.264)
GQ post period × (0–50 km)	-0.210^{***}	-1.161^{***}	-0.013	0.413
	(0.061)	(0.253)	(0.062)	(0.319)
GQ post period × (50–100 km)	-0.185^{***}	-1.023^{***}	0.022	0.106
	(0.060)	(0.249)	(0.061)	(0.308)
GQ post period × (100–150 km)	-0.131^{**}	-0.855^{***}	0.019	-0.200
	(0.058)	(0.226)	(0.060)	(0.294)
GQ post period × (150–200 km)	-0.008	-0.198	0.028	-0.012
	(0.051)	(0.197)	(0.060)	(0.301)
Distance 0–50 km	0.022	0.145	-0.014	0.265
	(0.021)	(0.100)	(0.020)	(0.722)
Distance 50–100 km	0.021	0.151	-0.021	0.394
	(0.020)	(0.099)	(0.018)	(0.717)
Distance 100–150 km	0.026	0.133	-0.006	-0.582
	(0.019)	(0.083)	(0.021)	(0.784)
Distance 150–200 km	0.015	0.075	-0.000	-0.560
	(0.014)	(0.060)	(0.016)	(0.697)
N	26,766	26,766	19,062	19,062
R^2	0.89	0.91	0.92	0.85

Notes: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. The table shows treatment estimates for the impact of the construction of the GQ highway network on forest cover in the proximity of the highway, according to equation (5). We define forest cover as log subdistrict forest cover (Columns 1 and 3) and average covered share of each subdistrict pixel (Columns 2 and 4); the data source is VCF. The distance variables are indicators that identify places within a given distance band from the GQ (Columns 1 and 2) or the NS–EW highway network (Columns 3 and 4). The omitted category is the band of places at a distance of 200–300 km from the highway network. These distance band indicators are then interacted with time period indicators. The construction period (rows 1 through 4) is 2001 to 2004. The post period (rows 5 through 8) is 2005 to 2008. Columns 3 and 4 estimate a placebo specification with distances to the NS–EW highway network, where construction had barely begun by 2008. The sample includes data from 2000 to 2008; 2000 is the omitted period. We omit years after 2008 as the placebo group is treated in those years. In Columns 3 and 4, we exclude places within 150 km of the GQ network to prevent sample contamination. All estimates include state-year fixed effects and standard errors are clustered at the subdistrict level to account for serial correlation.

show estimates of baseline differences between the GQ and the regions further away; the level differences are small relative to the treatment effects.

Columns 3 and 4 of Table 5 show comparable estimates along the NS–EW corridor for the same time periods. There are no detectable changes in forest cover close to the NS–EW in the time period when significant forest loss took place near the GQ. Along with Figure 5, this should alleviate any concern that the GQ treatment effects are driven by generalised deforestation along existing highway corridors from 2001–8. These results are robust to instrumenting for highway location using straight line instruments connecting the nodal cities of the highway network, as employed by Ghani *et al.* (2016); we present analogous reduced form estimates to the above in Online Appendix Table A9. These estimates alleviate the concern that the particular routing of the GQ (but not the NS–EW) was specifically targeted to places that may have already been losing forest cover.

An alternate estimation approach would be to exploit the timing of construction of each segment of the highway and to study forest cover in the years before and after the nearest segment to a given location was built. This would be analogous to the difference-in-differences estimation used to study the impact of rural roads in Section 3. The results, which are consistent with the findings above, are presented in Online Appendix Table A10.[26] One useful feature of this analysis is that it is directly analogous to our difference-in-differences estimates of the impacts of rural roads (equation (3) and Table 3), making clear that the differential effects of highways and rural roads are not an artifact of the different empirical strategies used to evaluate them.

A final concern with these estimates of forest change is that they could be describing displacement of forest loss from the hinterlands to the highway corridors, or even net afforestation in the hinterlands. This concern arises frequently in studies of transportation projects with national scale, and is typically only resolved by assumption through a structural modelling approach, which is beyond the scope of this article. This said, large displacement effects are made less plausible by the low quality of the broader road network, and by the high transportation costs during the sample period, weakening market connections with the hinterlands of these highways. While we cannot entirely rule out that there may be some displacement effects, Panel B of Figure 4 suggests that effects are driven by the highway corridor regions rather than the hinterlands; forest cover diverges between the places close to the GQ and NS–EW, but not in those far away.

4.2.1. *Mechanisms for highway effects*

We consider four possible mechanisms for the forest cover loss caused by India's major highway networks: (*i*) increased demand for timber products by firms due to local growth, (*ii*) increased demand for firewood due to shifts in household fuel consumption, (*iii*) expansion of agriculture into previously forested lands, and (*iv*) clearing of trees for settlements and industry. This section presents suggestive evidence that the deforestation along India's major highways is predominantly caused by increased logging driven by local timber demand.

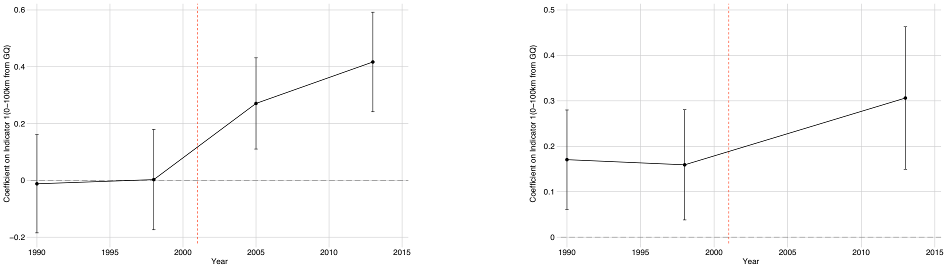
To identify potential mechanisms, we use the regression specification used to identify effects of highways on forest cover equation (5), with data from the economic and population censuses which were undertaken in various periods between 1990 and 2013. Because the results above suggest forest cover changes within 100 km of the highway network, we use 100 km distance bands and define 200–300 km as the omitted distance band.[27] The years in the sample are determined by census availability.

We first look at changes in employment in the list of industries that are directly downstream from timber harvesting (described in Section 2). Panel A of Figure 6 shows results from a regression of log employment in major wood-consuming sectors on the usual set of year-distance-band fixed effects. We graph the point estimates on the 0–100 km coefficient in each year that the Economic Census is available (i.e., the coefficients $\beta_{0-100\text{km},1990}$, $\beta_{0-100\text{km},1998}$, $\beta_{0-100\text{km},2005}$, and $\beta_{0-100\text{km},2013}$ from equation (5)). These estimates can be interpreted as the difference in residual log employment between the 0–100 km distance band and the 200–300 km distance band, after controlling for state-year fixed effects and the controls described above. In 1990 and

---

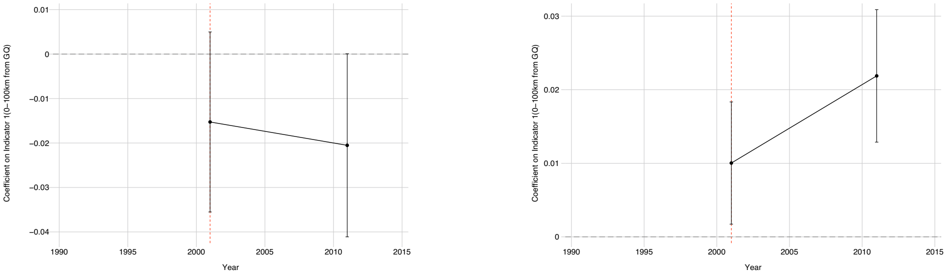
[26] We did not use this as a primary specification because there are substantial network effects in highway construction. A region without an upgraded highway segment may experience greater transportation access if the regions around it experience upgrades. Equally, a nearby upgrade may not be of much value if it is an isolated upgrade to a not-yet-upgraded part of the corridor. We therefore would expect estimates from a specification exploiting timing to underestimate the full impacts of the highway construction, which is what we find in Online Appendix Table A10.

[27] We find virtually identical results when we use 50 km distance bins.



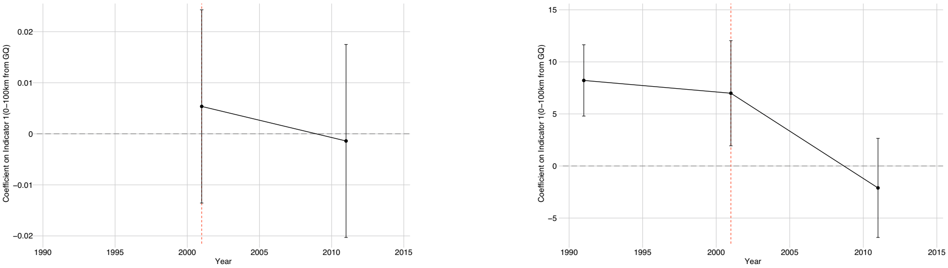
Panel A: Employment in Wood-Using Firms

Panel B: Employment in Logging Firms



Panel C: Share of Energy from Firewood

Panel D: Share of Energy from Imported Fuels



Panel E: Share of Energy from Local Non-Wood Sources

Panel F: Agricultural Share of Village Land

Fig. 6. *Mechanism Tests for Impact of Highways on Deforestation.*
Notes: The figure shows point estimates from equation (5), with distances from the GQ highway network specified in 100 km bands. Each figure shows the point estimate on the 0–100 km distance indicator, interacted with the year shown in the *X* axis. The omitted category is the set of places that are 200–300 kilometers from the GQ network. The dependent variables in Panels A and B are log employment in respectively wood-consuming firms and in logging firms. Logging was not specified in the 2005 Economic Census so this point is omitted. In Panels C through E the dependent variable is the share of households' cooking fuel that takes the form of (C) firewood; (D) imported fuels, primarily propane; and (E) crop residue and animal waste. In Panel F, the dependent variable is the share of village land dedicated to agriculture. All estimates are from regressions with state-year fixed effects and standard errors are clustered at the subdistrict level. Online Appendix Table A12 shows the full set of estimates from the regressions that produced these graphs.

1998 (before the GQ was begun), there is no significant difference between areas close to the highway corridor and areas that are far from it, nor is there a significant trend. By 2005, we see a 5% increase in employment in wood-consuming firms in the GQ corridor relative to the hinterland, which continues to rise through 2013. Panel B shows similar results for employment in logging firms; we omit 2005 because logging firms were not distinguished from firms engaged in afforestation in the 2005 Economic Census. These two graphs suggest that demand for wood from downstream firms is a plausible explanation for local deforestation after construction of the GQ.[28] Logging firms were more common along the transport corridor even before the GQ was built, but there is no suggestion of a pretrend that could explain what we see after highway construction. Note that employment in other sectors of the economy, which also consume wood, exhibit similar treatment effects close to the new highways (Ghani *et al.*, 2016).

Panels C through E of Figure 6 show the effects of the GQ upgrades on household fuel consumption, for which data are available in 2001 and 2011. We observe marginal increases in firewood and imported wood use, and comparable reductions in use of local non-wood fuels. These effects are not statistically significantly different from zero, nor are they large enough to explain a 20% reduction in forest cover in the neighbourhood of the GQ corridor. Panel F shows that land use shifts slightly *away* from agricultural uses following the construction of the GQ, breaking a previous upward trend, making agricultural extensification an unlikely explanation for the treatment effects.

It is difficult to directly test the last hypothesis that net forest loss has come from the expansion of land dedicated to settlement and industry, because data on land dedicated to settlement and industry only becomes available in the 2011 Population Census. However, it is implausible that settlement and industrial expansion could explain a 20% reduction in forest cover in a distance band as wide as 100 km around a 6,000 km long highway corridor. In 2011, only 6.7% of of rural land was used for settlement and industry.

The large relative forest losses along the GQ corridor are unlikely to be direct effects of the road construction process. Although we found evidence of these effects in the construction of rural roads in Section 3, these were very local, temporary, and an order of magnitude lower in size. It is also implausible that direct construction effects would extend more than a few kilometers from the roadway.

To summarize, we find evidence that expansion of industry demand for timber can explain forest loss in the GQ corridor, and we can rule out agricultural expansion, changes in household fuel consumption and settlement expansion as mechanisms.

5. Conclusion

The development, maintenance and expansion of transportation infrastructure is an important driver and correlate of economic development around the world. In this article, we provide causal estimates of the ecological impact of two transportation investments with global significance: India's massive expansion of rural roads and its upgrading of national highways. Using identification strategies established in the literature, we find that (*i*) the new rural roads had negligible effects on forest cover, and (*ii*) the highway expansions had a large negative effect on forest cover, which may have been driven by the expansion of wood-using industries. Methodologically, we demonstrate the critical importance of accounting for endogeneity and separately estimating the effects of the construction period from the post-completion period.

Because of the different implementation and spatial structure of rural roads and highways, it was necessary to use different empirical tests to evaluate the environmental impacts of each. However, our differential results are unlikely to be the result of having used these different

---

[28] Online Appendix Tables A11 and A12 show complete regression results for all mechanism tests. As would be expected, employment in logging firms is more geographically diffuse, as those firms reach further into the GQ's hinterland.

strategies. Our analysis of rural roads ruled out even small effects on forest cover extending six years after the onset of construction and four years after road completion—a time horizon in which the highway expansions already had substantial impacts on local forest cover. We also find no effects of rural roads on forests in places with high local demand for wood from either individuals or firms—which are the most likely drivers of forest change along highway corridors. Finally, while most of our analysis of rural roads focuses on local effects, we also found no effects at the wider distance horizons at which forest loss responds to highway expansion.

Globally, road expansion is expected to dramatically increase through the course of the twenty-first century. Some additional 25 million kilometers of road infrastructure is projected to be built by 2050, a 60% increase over 2010 levels. Nine out of ten of these roads will be built in developing countries (Laurance *et al.*, 2014). At the same time, tropical forests in developing countries are increasingly under threat. These forests not only provide global carbon benefits but also provide important local ecosystems which support biodiversity as well as the generally poor populations that rely on them (Barrett *et al.*, 2016). Against the background of this tension between economic development and environmental conservation, understanding the relationship between roads and forests is fundamental to a successful strategy for sustainable development.

Crucially, we show that the impact of road construction depends on what those roads connect. The fiscal costs of the two large scale transportation investments that we study were similar, but they had vastly different environmental consequences. Expansion of existing highway corridors caused changes in the spatial distribution of industry, which had dramatic effects on forest use in India. In contrast, building roads to connect smallholder farmers to new markets had virtually no impact on local forests, even for those farmers most likely to draw some part of their livelihoods from those forests.

Johns Hopkins University SAIS (Asher was at the World Bank Research Group when this paper was accepted)
UC San Diego
Dartmouth College

Additional Supporting Information may be found in the online version of this article:

Online Appendix
Replication Package

References

Adukia, A., Asher, S. and Novosad, P. (2020). 'Educational investment responses to economic opportunity: Evidence from Indian road construction', *American Economic Journal: Applied Economics*, vol. 12(1), pp. 348–76.

Alix-Garcia, J., McIntosh, C., Sims, K.R. and Welch, J.R. (2013). 'The ecological footprint of poverty alleviation: Evidence from Mexico's oportunidades program', *Review of Economics and Statistics*, vol. 95(2), pp. 417–35.

Andreoni, J. and Levinson, A. (2001). 'The simple analytics of the environmental Kuznets curve', *Journal of Public Economics*, vol. 80(2), pp. 269–86.

Angelsen, A. and Kaimowitz, D. (1999). 'Rethinking the causes of deforestation: Lessons from economic models', *The World Bank Research Observer*, vol. 14(1), pp. 73–98.

Antweiler, W., Copeland, B.R. and Taylor, S.M. (2001). 'Is free trade good for the environment?', *American Economic Review*, vol. 91(4), pp. 877–908.

Arrow, K., Bolin, B., Costanza, R., Dasgupta, P., Folke, C., Holling, C.S., Jansson, B.O., Levin, S., Mäler, K.G. and Perrings, C. (1995). 'Economic growth, carrying capacity, and the environment', *Ecological Economics*, vol. 15(2), pp. 91–5.

Asher, S., Lunt, T., Matsuura, R. and Novosad, P. (2019). 'The socioeconomic high-resolution rural-urban geographic dataset on India (SHRUG)', Working Paper.

Asher, S. and Novosad, P. (2020). 'Rural roads and local economic development', *American Economic Review*, vol. 110(3), pp. 797–823.

Assunção, J., Lipscomb, M., Mobarak, A.M. and Szerman, D. (2017). 'Agricultural productivity and deforestation in Brazil', Mimeo.

Barber, C.P., Cochrane, M.A., Souza, C.M. and Laurance, W.F. (2014). 'Roads, deforestation, and the mitigating effect of protected areas in the Amazon', *Biological Conservation*, vol. 177, pp. 203–9.

Barrett, C.B., Garg, T. and McBride, L. (2016). 'Well-being dynamics and poverty traps', *Annual Review of Resource Economics*, vol. 8, pp. 303–27.

Bauch, S.C., Birkenbach, A.M., Pattanayak, S.K. and Sills, E.O. (2015). 'Public health impacts of ecosystem change in the Brazilian Amazon', *Proceedings of the National Academy of Sciences*, vol. 112(24), pp. 7414–9.

Burgess, R., Hansen, M., Olken, B.A., Potapov, P. and Sieber, S. (2012). 'The political economy of deforestation in the tropics', *The Quarterly Journal of Economics*, vol. 127(4), pp. 1707–54.

Busch, J. and Ferretti-Gallon, K. (2017). 'What drives deforestation and what stops it? A meta-analysis', *Review of Environmental Economics and Policy*, vol. 11(1), pp. 3–23.

Chomitz, K.M. and Gray, D.A. (1996). 'Roads, land use, and deforestation: a spatial model applied to Belize', *The World Bank Economic Review*, vol. 10(3), pp. 487–512.

Copeland, B.R. and Taylor, M.S. (1994). 'North–South trade and the environment', *The Quarterly Journal of Economics*, vol. 109(3), pp. 755–87.

Copeland, B.R. and Taylor, S.M. (2004). 'Trade, growth, and the environment', *Journal of Economic Literature*, vol. 42(1), pp. 7–71.

Cropper, M., Puri, J., Griffiths, C., Barbier, E.B. and Burgess, J.C. (2001). 'Predicting the location of deforestation: The role of roads and protected areas in north Thailand', *Land Economics*, vol. 77(2), pp. 172–86.

Dasgupta, P. (2007). 'The idea of sustainable development', *Sustainability Science*, vol. 2(1), pp. 5–11.

Dasgupta, S., Laplante, B., Wang, H. and Wheeler, D. (2002). 'Confronting the environmental Kuznets curve', *Journal of Economic Perspectives*, vol. 16(1), pp. 147–68.

Dasgupta, S. and Wheeler, D. (2016). 'Minimizing ecological damage from road improvement in tropical forests', World Bank Policy Research Working Paper No. 7826.N

Datta, S. (2012). 'The impact of improved highways on Indian firms', *Journal of Development Economics*, vol. 99(1), pp. 46–57.

Den Butter, F. and Verbruggen, H. (1994). 'Measuring the trade-off between economic growth and a clean environment', *Environmental and Resource Economics*, vol. 4(2), pp. 187–208.

Deng, X., Huang, J., Uchida, E., Rozelle, S. and Gibson, J. (2011). 'Pressure cookers or pressure valves: Do roads lead to deforestation in China?', *Journal of Environmental Economics and Management*, vol. 61(1), pp. 79–94.

Foster, A.D. and Rosenzweig, M.R. (2003). 'Economic growth and the rise of forests', *The Quarterly Journal of Economics*, vol. 118(2), pp. 601–37.

Frankel, J.A. and Rose, A.K. (2005). 'Is trade good or bad for the environment? Sorting out the causality', *Review of Economics and Statistics*, vol. 87(1), pp. 85–91.

Garg, T. (2019). 'Ecosystems and human health: The local benefits of forest cover in Indonesia', *Journal of Environmental Economics and Management*, vol. 98, p. 102271.

Geist, H.J. and Lambin, E.F. (2002). 'Proximate causes and underlying driving forces of tropical deforestation: Tropical forests are disappearing as the result of many pressures, both local and regional, acting in various combinations in different geographical locations', *BioScience*, vol. 52(2), pp. 143–50.

Gelman, A. and Imbens, G. (2019). 'Why high-order polynomials should not be used in regression discontinuity designs', *Journal of Business & Economic Statistics*, vol. 37(3).

Ghani, E., Goswami, A.G. and Kerr, W.R. (2016). 'Highway to success: The impact of the Golden Quadrilateral project for the location and performance of Indian manufacturing', ECONOMIC JOURNAL, vol. 126(591), pp. 317–57.

Grossman, G.M. and Krueger, A.B. (1995). 'Economic growth and the environment', *The Quarterly Journal of Economics*, vol. 110(2), pp. 353–77, 10.2307/2118443.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O. and Townshend, J.R.G. (2013). 'High-resolution global maps of 21st-century forest cover change', *Science*, vol. 342(6160), pp. 850–3, ISSN 0036-8075, 10.1126/science.1244693.

Imbens, G. and Kalyanaraman, K. (2012). 'Optimal bandwidth choice for the regression discontinuity estimator', *Review of Economic Studies*, vol. 79(3), pp. 933–59.

Imbens, G.W. and Lemieux, T. (2008). 'Regression discontinuity designs: A guide to practice', *Journal of Econometrics*, vol. 142(2), pp. 615–35.

IPCC. (2007). 'The fourth assessment report of the intergovernmental panel on climate change', Geneva, Switzerland.

IPCC. (2019). 'Global Warming of 1.5 Degree C: Special Report on the Impacts of Global Warming', Geneva, Switzerland.

Jayachandran, S., de Laat, J., Lambin, E.F., Stanton, C.Y., Audy, R. and Thomas, N.E. (2017). 'Cash for carbon: A randomized trial of payments for ecosystem services to reduce deforestation', *Science*, vol. 357(6348), pp. 267–73.

Kaczan, D.J. (2020). 'Can roads contribute to forest transitions?', *World Development*, vol. 129, p. 104898.

Khanna, G. (2016). ‘Road oft taken: The route to spatial development’, Mimeo.

Koop, G. and Tole, L. (1999). ‘Is there an environmental Kuznets curve for deforestation?’, *Journal of Development Economics*, vol. 58(1), pp. 231–44.

Laurance, W.F., Clements, G.R., Sloan, S., O’connell, C.S., Mueller, N.D., Goosem, M., Venter, O., Edwards, D.P., Phalan, B. and Balmford, A. (2014). ‘A global strategy for road building’, *Nature*, vol. 513(7517), p. 229.

Li, M., De Pinto, A., Ulimwengu, J.M., You, L. and Robertson, R.D. (2015). ‘Impacts of road expansion on deforestation and biological carbon loss in the Democratic Republic of Congo’, *Environmental and Resource Economics*, vol. 60(3), pp. 433–69.

McCrary, J. (2008). ‘Manipulation of the running variable in the regression discontinuity design: A density test’, *Journal of Econometrics*, vol. 142(2), pp. 698–714, ISSN 03044076, 10.1016/j.jeconom.2007.05.005.

National Rural Roads Development Agency.(2005). ‘Pradhan Mantri Gram Sadak Yojana—operations manual’, Ministry of Rural Development, Government of India.

Pfaff, A., Robalino, J., Walker, R., Aldrich, S., Caldas, M., Reis, E., Perz, S., Bohrer, C., Arima, E. and Laurance, W. (2007). ‘Road investments, spatial spillovers, and deforestation in the Brazilian Amazon’, *Journal of Regional Science*, vol. 47(1), pp. 109–23.

Pfaff, A.S. (1999). ‘What drives deforestation in the Brazilian Amazon?: Evidence from satellite and socioeconomic data’, *Journal of Environmental Economics and Management*, vol. 37(1), pp. 26–43.

Reed, T. and Trubetskoy, A. (2019). ‘The Belt and Road Initiative and the Value of Urban Land’, Mimeo.

Roberts, P., Shyam, K. and Rastogi, C. (2006). ‘Rural access index: a key development indicator’, The World Bank.

Stern, D.I. (2004). ‘The rise and fall of the environmental Kuznets curve’, *World Development*, vol. 32(8), pp. 1419–39.

Stern, D.I., Common, M.S. and Barbier, E.B. (1996). ‘Economic growth and environmental degradation: the environmental Kuznets curve and sustainable development’, *World Development*, vol. 24(7), pp. 1151–60.

Townshend, J., Hansen, M., Carroll, M., DiMiceli, C., Sohlberg, R. and Huang, C. (2011). ‘User guide for the modis vegetation continuous fields product collection 5 version 1’.

Weinhold, D. and Reis, E. (2008). ‘Transportation costs and the spatial distribution of land use in the Brazilian Amazon’, *Global Environmental Change*, vol. 18(1), pp. 54–68.