Can roads contribute to forest transitions?[*]

David J. Kaczan[†]

Job Market Paper

April 17, 2017 (link to most recent version)

Abstract

New road construction is widely seen to aid economic development. New roads have also been shown to increase deforestation, suggesting development-environment tradeoffs for the locations where studies have focused to date. Yet in other settings, multiple mechanisms could support the opposite road-forest relationship, alleviating this tradeoff. New roads could lower deforestation or promote reforestation by: (1) raising the relative productivity of labor in non-agricultural sectors, thereby reducing agricultural activity; (2) facilitating price convergence across regional forest-product markets, raising profits from forest management or plantations; and (3) encouraging substitution from locally collected fuelwood to other energy sources. India's Rural Roads Program offers an opportunity to explore these hypotheses in a large country with little previous investigation of the road-forest relationship. Program rules prioritized construction by village population ranges, which differentiated construction timing across villages. I exploit this with a generalized difference-in-differences estimation strategy, having combined satellite, survey and census data to create a village-level, countrywide panel for 2000-2014. As a robustness check, I instrument for road construction in a cross section regression using the discrete thresholds between population ranges. For India as a whole, I find that road construction contributed to tree cover expansion, in great contrast to the existing empirical road-forest literature. Further, I demonstrate considerable variation in road impacts across economic settings within India: frontier settings saw reductions in tree cover due to new roads, while less isolated settings with more agricultural development saw increases in tree cover. These results inform the spatial targeting of roads, while broadening the set of mechanisms used to explain forest transitions (the reversal of forest cover loss).

---

[*] The support of committee members Alex Pfaff, Brian Murray, Subhrendu Pattanayak, Elizabeth Shapiro (Duke University) and Ashwini Chhatre (Indian School of Business, ISB) is gratefully acknowledged. Shilpa Aggarwal (ISB) provided data and advice. Nikita Lakhotia, Varun Goel, Arnab Dutta, Harry Fisher (Researching Rainfed Agriculture Network) provided data. Sanjog Sahu (University of Arizona) and Satya Prasanna (ISB) provided helpful field assistance. E. Somanathan, Ruth DeFries, Jennifer Orgill, Ben McCartney, and Faraz Usmani provided useful comments. The input of participants at the University of Colorado Environmental and Resource Economics Workshop, the 2016 FLARE Conference, and ISB and Duke University seminars is acknowledged. This project received financial support from the Nicholas Institute for Environmental Policy Solutions and the Duke University Department of Economics. Any errors are my own.

[†] Economist (Young Professionals' Program, starting Sept. 2017), World Bank Group. This research was undertaken while the author was a PhD Candidate (Environmental and Resource Economics) at the Sanford School of Public Policy and Nicholas School of the Environment, Box 90328, Duke University, Durham, NC, 27708, USA. Email: david.kaczan@duke.edu

1. Introduction

Transport costs affect socio-economic development. Expansion of road networks in rural areas lowers transport costs, facilitating economic benefits through specialization, trade, and improved public service delivery. These benefits have been observed to contribute to agricultural productivity improvements, economic growth, and poverty alleviation in a range of countries (Aggarwal, 2014; Asher and Novosad, 2016; Bell and Van Dillen, 2012; Gollin and Rogerson, 2014; Khandker et al., 2009; van de Walle, 2002; Warr, 2010).

Expansion of road networks also affects forest cover. The empirical literature describes roads as one of the strongest and most consistent determinants of deforestation, particularly in tropical frontier forests (Angelsen and Kaimowitz, 1999; Ferretti-gallon and Busch, 2014; Geist and Lambin, 2002; Pfaff et al., 2013; Rudel et al., 2009). Roads decrease input costs, and in some cases, increase prices received for agricultural products. The area in which agriculture is profitable consequently expands, causing deforestation (in the absence of countervailing institutional constraints), a process described by the enduring von Thünen (1826) model. While not the sole cause of deforestation, roads and subsequent agricultural expansion have contributed to the high rates of forest loss observed globally over the past several decades. Between 2000 and 2012, 2.3 million km^2 of forest land was converted to other uses (Hansen and et al., 2013) (an area larger than Mexico). This implies the loss of a variety of ecosystem services, including carbon sequestration, water provision, and biodiversity (Foley et al., 2005). Concerns have been raised that investments in roads globally – up to 25 million new kilometers of road by 2050 – will exacerbate these trends (Caro et al., 2014; Laurance et al., 2015, 2014).

Yet there are reasons to believe that roads may not always cause deforestation and could, in particular settings, reduce deforestation or even encourage reforestation. First, roads could change the relative productivity of labor in non-agricultural and agricultural sectors in ways that reduce agricultural activity at low-productivity forest margins, and thereby reduce deforestation or increase reforestation at those margins. Second, roads could facilitate price convergence between rural and urban forest products markets, increasing the profitability of improved forest management or forest plantations. Third, roads could encourage substitution from local fuelwood to imported energy sources such as kerosene or compressed natural gas, reducing the pressure on forests. Whether one or more of these pathways is of sufficient magnitude to deliver net forest increase – given countervailing impacts of roads and other drivers of land cover change – is likely to be a function of local economic and geographic settings. Testing for the existence of such settings, and exploring their characteristics, is this paper's objective.

A majority of previous quantitative studies of deforestation have focused on frontier forests (i.e. forests with limited prior human settlement, limited or no property rights, and situated at or close to the interface between forest and cleared land), particularly in Central and South America (Ferretti-gallon and Busch, 2014; Rudel et al., 2009). This geographic focus is understandable: Forest frontiers feature the most dramatic land cover shifts, and South America has experienced a large absolute loss of forest cover (FAO, 2010) and a very fast rate of tropical rainforest loss (Hansen and et al., 2013). Other contexts may offer substantially different insights. India, the focus of this study, offers settings characterized by extensive histories of settlement and agricultural development, high rural population densities and (relatively) well-defined property rights. This implies a predominance of mixed-use landscapes, 'mosaics' rather than tracts of wilderness (although it should be noted that these forests may still provide a variety of important ecosystem services, see Agrawal et al., 2014). Further, forest cover in India is growing overall (World Bank, 2015). Importantly, India also offers the opportunity to causally test this paper's central hypothesis – that roads may increase tree cover above underlying trends in particular settings – via a recent large-scale rural roads building program. The program has natural experiment characteristics, allowing for the use of evaluation techniques that avoid many of the endogeneity problems that otherwise challenge causal identification of road impacts (van de Walle, 2009).

The program, *Pradhan Mantri Gram Sadak Yojana* (PMGSY) ('Prime Minister's Rural Roads Program') constructs one-lane all-weather sealed roads that provide access to previously unconnected villages across India. Its first stage, the focus of this study, ran from 2001 to 2013 and provided road access to over 110 million people. Construction was prioritized by village population ranges, directing road construction first toward larger villages in each district – thus allowing for evaluation of road impacts using a generalized difference-in-differences strategy.

I apply this strategy to a unique, village-level panel dataset comprised of roads data, remotely-sensed forest change data (Vegetation Continuous Fields (VCF), Townshend et al., 2011; Global Forest Change (GFC), Hansen et al., 2013), and village and district level socio-economic data taken from Indian censuses and National Sample Surveys. My panel includes every Indian village for which matches can be made (over 435,000). The value of this panel is that it documents road construction and forest change within villages across time, allowing for the control of all sources of time-invariant selection bias at a highly localized (village) scale. This is relatively unusual in the roads evaluation literature, particularly in the context of a natural experiment that simultaneously provides plausible control against time-varying selection bias.

The program rules imply that the year in which a village received a road was a function of its own population and of the population distribution in the surrounding district. Populations were determined by a census that took place just before the program commenced. The impact of the village's population (and associated characteristics) at the time of program commencement is time-invariant and thus controlled for by village fixed effects. Meanwhile, the population distribution (i.e. how many villages were higher in priority for roads in that district) is plausibly exogenous to time-varying characteristics of a particular village after the program commenced. Studies that lack these characteristics (observations of roads over time, and rule-based placement), risk bias due to the strong selection pressures that are typically present in rural road placement decisions (van de Walle, 2009; Warr, 2010). I test and support the assumptions underlying this strategy with an event-study analysis and checks on pre-treatment trends. As a further robustness check, I use the thresholds between population categories as instruments for new road construction in a cross section regression.

My empirical analysis shows that the causal impact of new roads on tree cover is positive for India overall, on the order of 0.6 to 1.2 percent relative to baseline tree cover, above underlying trends across the study period (2000-14). While previous studies provided indirect support for the possibility of positive road impacts (Deng et al., 2011; Foster and Rosenzweig, 2003), this is the first direct empirical evidence of such from any country, to my knowledge. The magnitude of this impact is small in absolute terms, yet is substantial given the short 15-year study period: land use decisions are not instantaneous, and trees grow slowly. Moreover, any such positive finding is noteworthy in the context of the deforestation literature that until now overwhelming concludes that new roads are a negative influence of forest cover (reviewed in section 2).

This positive average treatment effect suggests that new rural roads are contributing to India's forest transition. A forest transition is a reversal of forest loss on a national scale (Mather, 1992)[1]. This phenomenon is often explained in the literature with reference to human migration (urbanization), forest product price changes, agricultural intensification, technological change (such as the transition from wood to coal), and/or institutional change[2] (Barbier et al., 2010; Perz, 2007). Until now, roads have been largely

---

[1] A forest transition leads to an approximate 'U-shaped' (or more accurately, reverse 'J-shaped') national forest cover function over time. Early studies documented this relationship for European countries (France, Denmark, Switzerland, and others) which passed through transition points in the 19th century (Mather et al., 2000, 1999). More recent forest transitions have been documented in Asian countries, notably China, India, and Vietnam (FAO, 2010; Foster and Rosenzweig, 2003; Mather, 2007). In India, the proportion of forested land increased from around 10 percent in 1971 to over 23 percent in 2012 (World Bank, 2015).

[2] Policy changes may complement, precipitate or diminish the economic forces underlying forest transitions. In France, for example, management and land tenure changes encouraged private forest investment (Mather, 1992). In India,

neglected as possible contributors to forest transitions. My positive road impact finding is consistent with forest transitions theory given that linkages in labor and commodities markets (linkages that may be provided by roads) are implicit in the mechanisms proposed for forest transitions. I return to this point in the discussion (section 6).

A second key finding is considerable spatial heterogeneity in road impacts. Just as India differs from Central and South America, settings within India differ from each other. Within my study area, I hypothesize and then show that road impacts vary in predictable ways. Settings more characteristic of frontiers – those with more initial tree cover, and located at greater distances from urban centers – show negative road impacts (relative tree cover loss) from new road access, while more agriculturally developed and proximate settings drive the finding of positive impact from new roads (relative tree cover gain) observed for India as a whole.

This finding of significantly heterogeneous impacts, indeed differences in sign, is based on tree cover levels (VCF data) which reflect both deforestation and reforestation. The conclusion also finds support, however, when I consider deforestation rates alone (as captured by GFC, a separate data source). Deforestation rates slowed very slightly due to new road connections within the more agriculturally established settings, and increased slightly due to road connections in frontier type settings. Given that existing deforestation studies have focused on forest frontiers, the results from the more 'frontier-like' context are concordant with that literature's negative road-impacts finding. I reconcile these opposing road impacts across settings using a modified von Thünen model, proposing that inter-sectoral transfers in labor may be responsible.

In reporting these findings for India, I do not suggest that the concerns raised about roads' forest impacts globally (Caro et al., 2014; Laurance et al., 2015, 2014) are erroneous or unjustified. As I show, and as is clear from the existing literature, roads are often destructive. My findings simply suggest that in some settings (those characterized by mosaic landscapes, high rural population densities, and extensive agricultural development), roads give rise to benefits that may previously have been overlooked. On a practical level, my results offer the possibility of targeting new roads to avoid environmental costs and/or to generate environmental benefits. In combination with existing studies (e.g. Aggarwal, 2014; Asher and

---

joint forest management (JFM) was introduced in 1991 to devolve forest management to more local levels and thus improve conservation; and a widespread tree planting program of 'social forestry' increased tree cover in the 1970s and 80s (before the period considered in this study) (Mather, 2007).

Novosad, 2016) and future research on socio-economic outcomes, these results can help determine road placement criteria likely to lead to joint social-environmental benefits from new roads.

In the next section I briefly review the road-deforestation literature and describe possible mechanisms for positive road impacts. I also present a modification of the von Thünen model that can account for opposing road impacts across settings. I then describe the multiple datasets used to construct this study's panel in section 3 before expanding on the empirical strategy in section 4. Section 5 describes the results, and section 6 discusses these in the context of the forest transitions literature.

2. The Relationship between Roads and Forest Cover

2.1. Relevant Literature

Early empirical literature on drivers of deforestation used national level indicators, and highlighted population density, agricultural expansion and wood production as key causal factors (e.g. Allen and Barnes, 1985; Rudel, 1989). Subsequent availability of satellite-based land cover data allowed for detailed econometric analysis of deforestation and its spatial correlates, including infrastructure. Early examples using these methods were the first to highlight transport costs, usually in the form of road proximity, as important deforestation factors (Chomitz and Gray, 1996; Cropper et al., 2001; Deininger and Minten, 1998; Nelson and Hellerstein, 1997; Pfaff, 1999). These studies explained their findings with reference to the von Thünen (1826) model, in which land rents are determined by production costs and commodity prices, which are themselves functions of transport costs. Four meta-analyses summarized these and numerous subsequent empirical studies, concluding that roads are among the strongest and most consistent determinants of deforestation (Angelsen and Kaimowitz, 1999; Ferretti-gallon and Busch, 2014; Geist and Lambin, 2002; Rudel et al., 2009).

However, debate exists over the magnitude and direction of road impacts in settings where considerable prior clearing has occurred. Far fewer studies have considered such settings. The deforestation literature gravitates towards places experiencing extensive deforestation, typically frontier forests with large tracts of undisturbed forest and minimal property rights. There is also disproportionate geographic representation from South and Central America (for example, more than 60 percent of studies considered by Ferretti-gallon and Busch, 2014, and more than 50 percent of studies considered by Rudel et al., 2009). A number of authors have highlighted the need to consider road impacts in other settings, including on relatively local scales i.e. within countries (e.g. Chomitz, 2007; Pfaff et al., 2009).

Yet previous studies that compare settings were inconclusive on the question of whether positive road impacts occur. Andersen et al. (2002) analyzed road impacts in the Brazilian Amazon with county-level data, and concluded that sufficiently high prior clearing would cause a reversal of road-induced deforestation. However, Pfaff et al. (2007) who undertook pixel-level analysis of the same region, did not reach this conclusion although acknowledged its possibility in other settings. They reported a reduction in roads' impact due to prior development but never a negative impact. In a later study, Pfaff et al. (2016) refined this finding by demonstrating non-monotonic road impacts for the Brazilian Amazon: small and negative in undeveloped and highly developed settings, and large and negative in moderately developed settings. Deng et al. (2011) analyzed prior road impacts on forest cover in China's Jiangxi Province, but found no robust road impacts either positive or negative. Indirect support for positive road impacts was found by Foster and Rosenzweig (2003) who, in a household sample within 250 Indian villages, showed that rising incomes drive forest product demand and consequent expansion in forest cover[3]. They did not explicitly test the impact of roads, yet roads as drivers of economic development may contribute to the increased demand for forest products they observed.

2.2. Potential Pathways

Foster and Rosenzweig (2003) thus highlighted increased demand for forest products (which incentivizes plantations and improved forest management) as a potential road impact mechanism[4]. In addition, I propose three other mechanisms: (1) agricultural intensification, (2) inter-sectoral labor shifts, and (3) energy substitution (Figure 1). All rest on the assumption that roads change prices due to price convergence between newly linked markets. I describe these mechanisms conceptually, before presenting an extension of the von Thünen model to explain how opposite impacts could manifest across space based on labor shifts between economic sectors. I also explore labor allocation in a simple model of household decision-making in the Supplementary Section (section 11).

The agricultural intensification pathway entails a relative shift in agricultural production from within forested areas or near forest margins, to non-forest areas. Such a shift could manifest in a number of ways: vegetable growing, cropping, or stall-fed animal production (e.g. small dairies), for example, could increase in prevalence relative to more land-expansive (but less labor productive) grazing on forestlands or along forest margins. These activity shifts could promote natural regrowth, and/or lower the opportunity cost of

---

[3] They rule out agricultural productivity increases as drivers of forest expansion (the 'Borlaug hypothesis').

[4] This hypothesis (not necessarily considering roads) is known as the 'forest scarcity path' in discussions of forest transitions (Barbier et al., 2010; Rudel et al., 2005).

replanting[5]. Conditions promoting the shift towards intensification include improved agricultural product prices or reduced agricultural input prices, both of which could result from lower transport costs. Of course, an increase in land-intensive agriculture in itself does not preclude simultaneous growth in land-expansive agriculture or grazing (Angelsen and Kaimowitz, 2001). However, limited local supply of labor or capital, and limitations in the mobility of these factors, may cause a trade-off between agricultural types (explored in more detail in section 2.3 and formally in the Supplementary Section). Such a tradeoff, particularly in labor (i.e. localized rural labor markets), is likely in the context of this study given the relatively low levels of migration between rural areas in India (Foster and Rosenzweig, 2003, 2004)[6].

The second hypothesized pathway simply extends this idea of sectoral shifts in labor to non-agricultural activities. Reduced input prices or increased output prices of non-agricultural goods and services (due to lower transport costs) could lead to greater employment in small commerce or manufacturing enterprises within newly connected villages. Similarly, roads could increase opportunities for labor to move to urban work opportunities nearby but outside the newly connected village (i.e. temporary migration or commuting). In both, roads could facilitate a labor shift from low-productivity agricultural or grazing sectors to the non-agricultural sector, reducing pressure on forests[7].

Third, increased trade due to roads increases the diversity of the consumption bundle available to households (Aggarwal, 2014). In the case of energy consumption, households may substitute firewood for non-forest fuels, such as kerosene and liquid natural gas (Baland et al., 2006; Veld et al., 2006). Commercial non-forest fuels are widely used in urban India, but less so in rural areas, where firewood remains the primary fuel for approximately 70 percent of households[8]. Firewood is generally collected at zero monetary

---

[5] Support for this pathway comes from Burton (2011), who quantified agricultural production and land use in 17 villages in Himachal Pradesh. In response to increased vegetable prices, labor for grazing on forested or marginal land shifted toward more lucrative vegetable cultivation, with a concordant decrease in total land use. Across villages, reduced grazing was correlated with increased forest density.

[6] Foster and Rosenzweig (2004) reported that less than 11 percent of their representative rural sample (male, aged 20-37) had moved from their village of birth in 1999.

[7] Support for this pathway comes from Asher and Novosad (2016) in their assessment of the employment impacts of the PMGSY program. They report a ten percentage point reduction in the proportion of workers engaged in agriculture, and an equivalent increase in non-agricultural wage employment, following a new road connection. This reallocation is strongest in villages close to large cities, suggesting that access to urban opportunities is primarily responsible.

[8] Liquid natural gas accounts for approximately 15 percent of rural India’s energy consumption (and is growing).

cost, but at considerable time cost, from village forests, common lands, roadsides, and private fields (Heltberg et al., 2000). A shift from firewood collection to other labor uses could occur if substitute energy sources become available and affordable, reducing pressure on forests (Burton, 2011).

While I describe these mechanisms in terms of a potential positive impact on tree cover, each could have negative effects under different conditions. Commodity price or wage changes due to increased market connectivity could be negative or positive depending on initial prices and local scarcity or abundance. Property rights and their enforcement may further mediate road impacts. The extent to which agricultural expansion is possible depends on whether boundaries between agricultural land (owned by individuals or communities) and forest land (owned by communities or the state) are delineated and respected. Similarly, it is conceivable that higher demand for forest products could increase rather than decrease deforestation (or forest degradation) particularly in cases where property rights are weak (Agrawal and Yadama, 1997)[9]. And as discussed in section 2, road impacts will also depend on prior development (including existing land use) (Pfaff et al., 2016)[10]. This large number of possible outcomes motivates firstly my use of reduced form analysis for estimating forest area response to roads, and secondly, my focus on settings. Causal analysis of the impact of mechanisms in isolation is desirable but beyond the scope of this study.

2.3. A Model of Heterogeneous Road Impacts

The von Thünen model (Figure 2A) provides a useful theoretical starting point given its prevalence in the deforestation literature. Land is allocated to its use of highest rent. Formally, rent, r, is a function of output levels, y, output prices, p, transport costs, v, distance to market, d, wage costs, w, capital costs, k, and land rights enforcement costs, c. In a model of two land uses, agriculture (A), and forest (unused land):

$$rent_A = p_A y_A - w_A y_A - k - c - v_A d$$

---

[9] Open-access forests, and poorly managed commonly held forests will face increased pressure from higher forest products prices. In contrast, well-defined and enforced property rights will allow landowners to capitalize on increasing prices, encouraging improved management of existing forests or new plantations (Robinson et al., 2013).

[10] How exactly prior land use would impact marginal clearing is not clear from theory alone. One possibility is that localities that have undergone significant deforestation already are unlikely to offer high returns to further clearing (as clearing would likely occur on the most productive land first). Another possibility is that initial deforestation encourages further deforestation due to economies of scale, conglomeration effects, or by simply providing access to more remote locations.

The agricultural frontier is defined as the distance from the market at which agricultural cultivation is no longer profitable, i.e. $rent_A(d) = 0$. Hence:

$$d^{frontier} = \frac{p_A y_A - w_A y_A - k - c}{v}$$

Among other predictions, a decrease in transport costs, v, extends the distance to the frontier, implying deforestation. Modification of this model is necessary to predict the reforestation possibilities I raise in section 2.2. Imagine the addition of a plantation rent curve. One possibility is that increased rents to plantations could drive reforestation either on the frontier (if forestry and agriculture curves are closely aligned), or closer to markets if plantation rent starts higher but falls more rapidly than agricultural rents as a function of distance.[11]

An alternative explanation arises from consideration of other factors of production (Figure 2B). Assuming that profits accrue partially to land and partially to labor (rather than only to land), I allow wages per unit of agricultural output w_A, to be a function of per unit profits, $p_A - v_A d$. Hence:

$$rent_A + y_A w_A = p_A y_A - k - c - v_A d y_A$$

And thus $w_A = w_A(p_A - v_A d)$, where $w_A'(.) > 0$. This function gives the wage per unit of output. The wage per unit of labor, $\widetilde{w_A}$ (the basis for an individual labor-provider's decision) includes the agent's productivity (output, y, per time unit), γ. Hence: $\widetilde{w_A} = \gamma * w_A(p_A - v_A d)$. This is the 'realized' value marginal product of labor (i.e. that accruing to the laborer). As in the traditional von Thünen model (although omitted above for brevity), I include an additional economic sector that produces non-agricultural agricultural output, y_N.[12] The same assumptions apply with two exceptions: (1) demand for land (per unit of labor utilized) is less than that of the agricultural sector, and transport costs, v_N are greater. The latter simply represents the notion that proximity to urban areas is more important for profitability for the non-agricultural sector than for the agricultural sector due to the need for market access. Labor productivity in

---

[11] It should be noted that the original von Thünen model does not predict mixed land use mosaics. For a given distance, the land use with the highest rent dominates. In reality, mixed uses are likely due to local heterogeneity.

[12] In this generalized framework, I call this second sector the 'non-agricultural' sector. In reality, it may be any sector, agricultural or otherwise, which is distinct from the first simply based on land use and the productivity impact of transport costs.

this sector is normalized relative to that in the first sector. Of course, no wages will be paid in either sector if the return to labor is zero:

$$\widetilde{w_A} = \begin{cases} \gamma * w_A(p_A - v_A d), & p_A - v_A d > 0 \\ 0, & p_A - v_A d < 0 \end{cases}$$

$$\widetilde{w_N} = \begin{cases} w_N(p_N - v_N d), & p_N - v_N d > 0 \\ 0, & p_N - v_N d < 0 \end{cases}$$

Graphically this is represented by the wage curves' intercept with the x-axis, and is analogous to the von Thünen zero rent condition. Given a uniform distribution of productivity across agents in the local labor market, agents choose to work in one sector or the other depending on their realized value marginal product of labor. The pool of available labor in each location is assumed to be fixed (as discussed in section 2.2 and footnote 6). Labor may select into either sector, or remain in surplus: $l_A + l_N \leq L$. Labor remaining in surplus is that of low productivity and/or high reservation wages. This setup is consistent with the ratios of agricultural to non-agricultural labor seen across India. There is a weak but statistically significant positive correlation at the district level ($r = 0.20, p < 0.01$) between the proportion of agricultural to total labor, and the average distance from villages in that district to the closest urban area. Districts with a majority of non-agricultural employment have low average distance (i.e. greatest access) to urban areas, almost all within 50 kilometers.[13]

At the frontier, deforestation occurs if the agricultural labor force (and thus agricultural area) expands in response to reduced transport costs. Given the differential transport costs between sectors, there is no viable non-agricultural sector at the frontier (i.e. the local labor force is employed in agriculture or not employed at all) and hence marginal changes in v_N do not lead to positive wages in the non-agricultural sector, $\widetilde{w_N}$. There is no possibility of inter-sectoral labor substitution, and growth in profits (implying both increased land rent and increased value marginal product of labor) can only lead to clearing on the frontier.

In intermediate positions between the frontier and the market, where there is a viable non-agricultural sector, sectors compete for labor. Marginal reductions in v_N and v_A increase demand for labor in both sectors, although disproportionately so in the non-agricultural sector because $v_N > v_A$. If the net transfer of labor from agriculture to non-agriculture is greater than the uptake of underutilized labor by the agricultural

---

[13] Data for this statistic and other analyses are described in section 3.

sector, agriculture will shrink, reducing pressure on forests. The reduced pressure on forests is because agriculture (and thus agricultural labor) has greater forest impacts than non-agricultural activity.

While reliant on a number of assumptions (importantly, differential labor productivity changes in response to transport cost changes between sectors, and labor market constraints), this analysis shows how opposing road impacts may be reconciled with von Thünen's monotonic predictions, and is useful given the observed heterogeneous impacts presented subsequently in this paper. It should also be noted that this mechanism is complementary with the previous mechanism described – a rise in plantation forestry due to increased plantation land rent – if plantations have lower labor demand than agriculture (as is likely).

3. Data

I construct a unique, village-level, countrywide panel of satellite-derived land cover data, socio-economic and demographic survey and census data, and PMGSY roads data, for 2000-14.

3.1. Land Cover and Slope

I use two publically available satellite-based products, Vegetation Continuous Fields (VCF) MOD44B (v. 51) (Townshend et al., 2011) and Global Forest Change (GFC) (Hansen and et al., 2013). VCF provides annual 2000-14 tree cover at a resolution of 250 meters, specifically, the percentage of each pixel covered by woody vegetation. The advantage of this product over other similar products is that it provides a continuous measure of tree cover without arbitrary thresholds between cover classes. GFC provides annual 2000-13 forest loss with a resolution of 30 meters. Unlike VCF, data are binary, i.e., a change in pixel value indicates the near-complete removal of tree canopy from that pixel. This gives the deforestation rate, but not cumulative forest cover or forest condition. Like other satellite-based data, these products do not distinguish between plantations and natural forests. I source slope data from the Esri online global terrain layer (Esri, 2013). This product uses NASA Shuttle Radar Topography Mission data with resolution of approximately 30 meters. I create village-level variables by averaging pixel values from the remotely sensed data products in each year over a uniform circular land area (radii of 5 and 10 kilometers) around each Indian village.[14] The location of villages is determined by plotting geographic coordinates extracted from the online portal *India Place Finder* (Sagara, 2011).

---

[14] This leads to some overlap of areas when villages are close together. As a robustness check, I remove up to 90 percent of the sample (randomly), which greatly reduces overlap. Results remain consistent.

3.2. Road Data

Data on PMGSY road construction at habitation level, covering the period 2000-12, were sourced from the Indian Government's Online Management and Monitoring System (OMMS)[15]. Habitations are sub-units of villages (which are standardized administrative units for census purposes). PMGSY data includes habitation name, year of connection approval, habitation population, baseline connectivity, and whether a new road or an upgrade was constructed. I create road connectivity variables by aggregating habitation data to the village level. These include a binary variable indicating village connectivity if at least one habitation within the village is connected; a continuous variable indicating the proportion of a village's population connected (based on habitation-level population data), and similarly, a continuous variable indicating the sum of the village's population connected.

3.3. Socio-Economic and Price Data

The 2001 Census (Ministry of Home Affairs, 2016) gives village-level demographic, socio-economic, land use, and infrastructure information, for the baseline year. The census does not use the village coding system used by the roads data, necessitating fuzzy merging based on district, sub-district, and village names (which contain considerable spelling differences). I successfully match 83 percent of villages. Summary statistics of census 2001 data shows the relative socio-economic disadvantage faced by villages without a road connection (Table 1).

4. Empirical Approach

4.1. The PMGSY Roads Program

The first phase of PMGSY operated between 2001 and 2013 and aimed to increase agricultural incomes via farm-to-market access, provide employment opportunities, and reduce rural poverty (Ministry of Rural Development, 2012a)[16]. Construction took place in 32 out of 36 states/territories, in 663 out of 676 districts.

---

[15] Processed records from OMMS were kindly provided by Shilpa Aggarwal (Indian School of Business) for this project.

[16] Two recent papers suggest these goals were at least partially met. Aggarwal (2014) found that PMGSY roads led to increased use of chemical fertilizer and hybrid seeds, and changes in school attendance (increases for children 5-14 years of age, and decreases for older teenagers (14-20 years age) likely due to both increased access to schools and increased access to work opportunities). Aggarwal (2014) also showed a change in the mix of labor market activity with increased animal raising, textiles and retail, and an increase in consumption variety – specifically a switch from non-perishable foods (cereals and pules, for example) to increased variety in perishable dairy and vegetables, as well

Approximately 45 percent of habitations in the participating districts were unconnected to (i.e. beyond 500 meters of) the national road network in 2000. At the end of the first phase, 175,674 habitations (11.5 percent of the total eligible) had been connected via approximately 360,000 kilometers of new road, giving approximately 110 million people all-weather road access for the first time. Total expenditure was around USD 26 billion (including road upgrades and maintenance). Funding was provided by the federal government, sourced from a fuel tax. A second phase of the program (not part of the present study) is ongoing.

Importantly for causal analysis, road construction was prioritized by habitation population (determined by the 2001 census). National guidelines stipulated that within each district, habitations with a population of 1000 or above were connected first, those with a population of 500-999 second, and those smaller only thereafter (some states included a category of 250-499 in addition) (Ministry of Rural Development, 2012a). The population distribution within each district determined the order of construction within that district, and thus the likelihood that a particular habitation received a new road connection in a particular year. Resulting probabilities of road construction as a function of population categories provides evidence that the rule was in general adhered to (Table 2 and Table 3). In addition, a map of eligible road connections was made in advance of the program (the "core network") based on the minimum number of connections necessary to connect eligible habitations. Only connections specified in the core network could be subsequently constructed under the program. Only rural roads were eligible, and roads constructed had to provide a link to a formerly unconnected habitation. Both phases of the program upgraded existing roads as a second priority.

4.2. Identification Strategy

4.2.1. Exploiting Differences in Timing

I use a generalized difference-in-differences strategy with village fixed effects to identify the reduced-form impact of new road connections on tree cover. Treatment is receiving a road. Villages that do not receive roads act as controls; tree cover change in these villages is captured by the estimated underlying time trend. Villages that receive roads at some point in the study period are control villages prior to receiving the road and treatment villages after receiving the road. This approach compares the change in tree cover within

---

as greater non-food consumption. These changes are consistent with predictions based on reduced transportation costs. Asher and Novosad (2016) presented evidence of socio-economic transitions using a regression discontinuity approach and a highly detailed (individual level) panel. Labor switched toward non-agricultural activities, facilitated by local rural-urban migration rather than within-village sectoral shifts.

villages that received roads with the change in tree cover within villages that did not receive a road (or have not yet).

The primary specification is:

$$F_{v,t} = \beta_1 R_{vt} + \boldsymbol{\beta} R_{vt} * \boldsymbol{X}_v + \theta_{st} + \theta_v + \varepsilon_{vt}$$

Where $F_{v,t}$ is average tree cover in time t around village v; R_{tv} is an indicator of road connectivity in time t in village v, and $\boldsymbol{X}_v$ are key time-invariant covariates[17]. I use three specifications of the road variable: a binary variable indicating village connectivity if at least one habitation within the village is connected, $I(R_{th} = 1)$; a continuous variable indicating the proportion of a village's population, N_v connected (where habitation population is given by N_h), $\sum_{h \in v} I(R_{th} = 1) * \frac{N_h}{N_v}$; and similarly, a continuous variable indicating the sum of the village's population connected, $\sum_{h \in v} I(R_{th} = 1) * N_h$.

Considerable changes in Indian tree cover can be expected irrespective of road impacts. To control for underlying trends I use state-year or district-year fixed effects, θ_{st}. Village fixed effects, θ_v, control for all time-invariant factors (i.e. remove time-invariant sources of selection bias). The state-year or district-year fixed effects flexibly and locally (i.e. at the state or district level) control for the underlying trend, allowing for estimation of the treatment effect above the underlying trend (the difference-in-differences estimator), represented by β_1. The treatment effect is an estimate of the average difference between the pre-treatment and post-treatment levels of tree cover, above the underlying trend. Identification is possible using a variety of subsamples: (1) all villages (i.e. those with and those without a road at baseline), (2) only initially unconnected villages, or (3) only initially unconnected but treated villages (i.e. villages in their pre-treatment state serve as control villages). Standard errors are clustered at the district level.

This strategy avoids some of the endogeneity problems often faced by studies of road impacts, specifically selection bias. Selection bias occurs because road placement is not random. Roads may be built in places already experiencing economic growth (correlated with many outcomes of interest, including land use change) or they may be built to places of low growth in an attempt to boost development (van de Walle, 2009). In my approach, selection bias driven by time-invariant factors, such as unobserved baseline village characteristics, is controlled by the village fixed effects. Selection bias driven by time-varying factors is more problematic: roads may be built to service forest plantations installed during the program period, which are unobserved across time and thus cannot be captured by controls or fixed-effects. I rely on the

---

[17] All panel regressions are linear OLS, estimated on Stata with user-created program '*reghdfe'* (Correia, 2014).

natural experiment characteristics of the PMGSY program to reduce the possibility of this bias, and then test for its presence. The prioritization rule means that the time in which a road is built to a village is a function of (1) the 2001 population of the village (a time-invariant property), and (2) the number of villages with higher priority for a road connection in the district (a characteristic external to the village in question). This reduces the likelihood that the timing of road connections are able to respond to time-varying factors within villages, although does not completely mitigate the possibility. Given this, I test for selection bias in an examination of pre-treatment trends, as well as present estimates disaggregated by time period.

4.2.2. Testing for Selection Bias

I regress treatment (whether village v received a road) on the change in forest cover over the first few years of the program:

$$R_v = \beta_1 \Delta F_v^{2000-0x} + X_v + \theta_s + \varepsilon_v$$

Where R_v indicates road treatment (at any point in time during the program), $\Delta F_v^{2000-0x}$ indicates the change in tree cover around the village between years 2000 and either 2003, 2004 or 2005. This variable is the coefficient of a simple linear regression of tree cover on year and thus incorporates information from intermediate as well as start and end-points in each village's time series. In some specifications a suite of controls or state fixed effects are included. Few roads were constructed in these early years of the program: 2.6, 5.0 and 7.5 percent of villages unconnected at baseline had been connected by 2003, 2004, and 2005 respectively (Table 3). While the small number of connections that are constructed within these periods could affect the observed trends, this possibility would bias results towards a relationship between tree change and treatment and thus make my pre-trends test more conservative.

4.2.3. Exploring Heterogeneity

A holistic description of baseline settings (and thus heterogeneity of settings) can be achieved by grouping like villages according to multiple baseline characteristics. I use k-means clustering, a procedure for partitioning observations into k classes in which each observation is attributed to a class based on its baseline characteristics. Each characteristic is standardized so receives equal weight. The characteristics used are (1) baseline agricultural development indicators (irrigated area, cropped area); (2) baseline infrastructure and services indicators (presence of a grid electricity connection, credit society, primary school, and primary health center); (3) baseline geographic descriptors (forest cover, distance to nearest urban area, and average slope); and (4) baseline scheduled tribe population (see Table 1 for summary statistics).[18] Attribution of villages to classes aims to minimize the within-class sum of squares (WSS), i.e. to minimize the squared Euclidian distance between the values of all characteristics $\boldsymbol{x_v}$ for a particular village *v*, and the means (across villages) of those characteristics, $\overline{\boldsymbol{x}}_k$, within a particular class, *k*:

$$\arg\min \sum_k \sum_v ||\boldsymbol{x}_v^{(k)} - \overline{\boldsymbol{x}}_k||^2$$

Attribution proceeds iteratively using a two-step procedure (the k-means algorithm). K initial means are randomly generated, and classes are created by associating each observation with the nearest randomly generated mean. Class means are then updated before observations are reattributed based on the updated means. This process iterates until convergence is reached (i.e. a stable WSS) and no observations change class between iterations. Given the potential for the initial random means to determine outcomes I repeat the above procedure 200 times with different starting points and select the partition most frequently observed.

4.2.4. Robustness: Exploiting Threshold Discontinuities

I provide supporting evidence for the observed treatment effect using instrumental variables regression. My instruments are the population thresholds, specified at habitation level, which guide construction prioritization (Figure 3 and Figure 4). The identifying assumption is that movement across the population threshold (a discrete change) increases the probability of receiving treatment, yet does not affect unobserved covariates that may affect tree cover, once population itself is controlled. I use flexible population covariates to achieve the latter. Thresholds apply to habitation populations in 2001 (at the start of the program). Given that habitations are sub-units of villages, and that tree cover data and census data is at the village level, I restrict the sample to villages with only one habitation (approximately 79 percent of the matched sample). Thresholds, T, are time-invariant instruments. I convert the panel to a cross section of villages by taking the difference in tree cover between endline and baseline observations, $F_v^{2000-2014}$. I specify road treatment, R_v, as construction of a new road in any year during the program. The IV predicts road treatment, rendering its predicted values exogenous to the extent that the identifying assumption holds. I specify the first stage as:

---

[18] Inclusion of different or additional attributes will change class outcomes; however, given that most variables available in my dataset are correlated to some extent with those in this set, the change is relatively minor.

$$R_v = \beta_0 + \beta_1 \boldsymbol{I}(pop_v > T) + \boldsymbol{\beta} \sum_{q=1}^{4} pop_v^q + \boldsymbol{\beta} X_v + \theta_d + \varepsilon_v$$

I then regress the change in tree cover against predicted treatment in the second stage:

$$\Delta F_v^{2000-2014} = \beta_0 + \beta_1 \widehat{R_v} + \boldsymbol{\beta} \sum_{q=1}^{4} pop_v^q + \boldsymbol{\beta} X_v + \theta_d + \varepsilon_v$$

Controls, X_v, population controls, $\sum_{q=1}^{4} pop_v^q$, and district level fixed effects, θ_d, are included in both stages. I use a window of data (T ± 250) around thresholds. This strategy has one advantage relative to the difference-in-differences strategy, the potential to avoid bias from time varying sources of endogeneity. However, the use of a time invariant-instrument prevents use of the rich panel data, and furthermore estimates treatment effects only for the window of villages around the threshold (thus giving local average treatment effects). Unbiased estimates also rely on accurate population data reporting at the start of the program. Consequently, I use this strategy as a qualitative check of the difference-in-differences estimates rather than the basis for my key reported results.

5. Results

5.1. Main treatment effects

I first consider the observed change between baseline and endline tree cover around villages, distinguished by road treatment. While both groups indicate an increase in tree cover (as expected given underlying trends in India, see World Bank, 2015) (Figure 6), there is a slightly greater gain (0.08 percentage points) among villages that received a road ($p < 0.001$, under both parametric and non-parametric distribution assumptions). Such a raw test of means does not control for differences between treatment and control groups, or for annual anomalies that could disproportionately affect one group. Regression results that do, however, support this finding. My primary specification suggests that roads increased tree cover by 0.05 to 0.122 percentage points ($p < 0.05$) (Table 4). These coefficients represent the change in level of tree cover in the zone around a village (defined at radii of both 5 and 10 km from the village center point), relative to the underlying, flexibly specified trend in tree cover. Coefficient values are at the low end of this range under more localized time-trends, i.e. with district-year rather than state-year fixed effects[19]. While this

---

[19] District-year fixed-effects provide a more localized (and thus more accurate) estimation of the underlying trend relative to state-year fixed effects, however, may represent excessive control. This is because the decision-making unit for treatment is a district-year, i.e. in every year, a district decides on (or at least submits for approval to higher levels

increase in tree cover is small in absolute terms, so is the baseline: average tree cover in zones around unconnected villages is 11.45 percent (and 9.94 percent around all villages, connected and unconnected). In relative terms, these coefficients represent a 0.6 to 1.2 percent increase above underlying trends. Findings are relatively consistent in magnitude across different versions of the treatment variable once average village size is accounted for (807 persons in initially unconnected villages). Magnitudes and significance patterns are similar under 5 km radius zones. This finding is a considerable departure from the strong negative road impacts reported elsewhere in the literature (reviewed in section 2 and discussed further in the following section).

I apply this specification to three samples: initially unconnected villages (Table 4), all villages (Table 5), and only treated villages (Table 6). My specification, based on village fixed effects, identifies a treatment effect based on change over time within villages. In this context, the role of the control villages is to estimate the underlying trend, represented by the district-year or state-year fixed effects, which is then subtracted from the change observed among treated villages. I find positive, significant average treatment effects under all samples, although coefficients based on the all-villages sample are smaller (0.04-0.07 percentage points). My preferred sample is all unconnected villages at baseline, which provides for strong comparability between treatment and control groups.

5.2. Testing for Selection Bias

Threats to causal identification from time-varying sources of selection bias are not addressed by my village fixed effects strategy. To explore the potential for this bias I look for statistical evidence of differences in pre-treatment trends between treatment and control groups. A significant relationship between the change in tree cover prior to treatment and the treatment itself would suggest that roads are constructed in response to economic activities or local policy decisions that are themselves the drivers of land use change. Significance in this relationship can be seen in the simplest regression specification across all three pre-treatment periods (first column of each period in Table 7). However, the non-significance that results from when state fixed effects are included, and/or when a small suite of (time-invariant) controls are included, indicates that pre-trends may be explained simply by time-invariant rather than time-varying factors. Village-level fixed effects control for these factors – and do so much more comprehensively than state fixed effects or controls – giving grounds for confidence in the treatment effect estimates presented above.

---

of government) the road projects to undertake (Ministry of Rural Development, 2012b). Consequently, district-year fixed effects control for factors that determine the treatment decision as well as the underlying tree cover trend.

I further explore endogeneity possibilities by estimating the causal impact within periods relative to treatment for each village: 1-3 years before and after, and greater than 4 years before and after (excluding the construction year). These estimations serve as a placebo test by estimating the relationship between treatment and tree cover prior to treatment taking place (i.e. shifting the construction indicator back in time). Results show no significance in pre-treatment years and significance following treatment (Table 8 (which also contains the same check but on the subsamples described in next section) and Figure 12). Evidence of reverse causality would show up as significance in pre-treatment years.

5.3. Exploring Heterogeneous Settings

I consider treatment effects specific to classes of villages identified using k-means cluster analysis. The optimal number of classes, k, may be determined through inspection of the change in the within-class sum of squares (WSS) statistic as k increases. I consider $k \in \{1, \dots 12\}$. WSS decreases approximately linearly until a change in slope around k = 6 (Figure 8), suggesting an optimum at this point (Makles, 2012). However, an arguably overriding criterion for choosing k is the ability to make useful interpretations of the resulting partition. More than three classes tends to result in small classes splitting off due to extremes in a single characteristic (e.g. presence of a medical center). I present two and three class solutions, which have relatively large classes and differ on a number of baseline characteristics.

In the two-partition analysis (Table 10, columns 2-3), approximately three quarters of the sample (class 1) is characterized by low initial tree cover, relatively flat topography, and greater areas of irrigation and crop cover. This class has above average infrastructure and services (health, education, credit society and electricity) and is located relatively close to urban centers. It has a below average proportion of scheduled tribe population. The remaining one quarter of the sample (class 2) has opposite characteristics. Geographically, class 1 is centered in the densely populated Gangetic plain of Uttar Pradesh, Bihar and West Bengal, as well as parts of the central highlands of Rajasthan and Madhya Pradesh. Class 2 is found in the Himalayas (north and northeast), and scattered throughout the eastern peninsula plateau in Chhattisgarh and Odisha (Figure 9).[20] The three-partition analysis (Table 10, columns 4-6) returns one class of 60 percent of villages that looks very similar to class 1 in the two-partition analysis (hence is also called class 1). The remaining 40 percent are relatively distant from urban areas and have less irrigation and crop

---

[20] The density of villages included in this analysis is greater in the northern half of India. This is due to my selection of only unconnected villages at baseline (for comparability). Infrastructure and service provision is generally much better in the south of India, and so there were far fewer southern Indian villages eligible for participation in the PMGSY program.

cover. One class within this group of villages has low initial tree cover and relatively high scheduled tribe population (class 2A, approximately 25 percent of villages) while the other has high initial tree cover and a smaller proportion of scheduled tribe population (class 2B, approximately 12 percent of villages). This class is also found on much hillier land concentrated in the Himalayas, while class 2A is more common on rolling hills in the eastern peninsula plateau areas and western plains of Rajasthan (Figure 10).

I estimate the impact of roads on tree cover within classes by interacting class indicators with the treatment variable (Table 11). Immediately apparent is a strong and statistically significant bifurcation in road impacts between classes. The more agriculturally developed, closer, better serviced villages (classes 1) show positive road impacts, specifically a 0.10 to 0.25 percentage point increase ($p < 0.01$) between the before and after periods relative to underlying trends. The hillier, more distant, more treed areas (class 2B) show negative road impacts, specifically a 0.24 to 0.48 percentage point decrease ($p < 0.01$) relative to underlying trends in the 3 class partition. The raw data depicts this result visually, with opposite skews in tree cover change for treated and untreated villages (Figure 11), as do event study plots of coefficients for disaggregated time periods Figure 12). In relative terms (relative to both underlying trends and initial tree cover), these coefficients represent a 1.2 to 3.2 percent increase in class 1 villages' tree cover, and a 0.7 to 1.1 percent decrease in class 2B villages' tree cover. I discuss the magnitude of these results in section 6 below. It should be noted that the underlying trend in tree cover in all classes is positive (with and without roads), with a faster rate of increase among class 1, intermediate in class 2B, and the slowest in class 2A.

Results for class 2 in the two class partition are negative also but not significant under district-level differential time trends. The weaker robustness of this result is not surprising given that class 2 is more broadly defined than class 2B. The villages that differ between these classes primarily form class 2A, which shows countervailing positive road impacts (although with mixed significance).

5.4. Corroborating Evidence

I first repeat the difference-in-differences analysis using my alternative outcome measure, the Landsat-derived annual deforestation rate data (GFC) (previous estimates use the MODIS-derived VCF data). Deforestation is not, in aggregate, a serious problem in India.[21] The high-resolution satellite data I use shows a small amount of deforestation (which may include the cutting of managed or planted forests), amounting to 0.024 percent of India's total land area per year (Table 12), concentrated mainly in the northeast and east

---

[21] The FAO Forest Resources Assessment, for instance, reports zero change in 'primary or old growth' forest cover (a measure of deforestation) over the last 15 years (FAO, 2015).

of the country. A number of villages (34.3 percent) report no deforestation over the 13-year time series. I find no statistically significant change in deforestation rate due to roads overall (Table 12). However, I find some evidence of road-induced changes after disaggregating by subsample. There is a 0.002 percentage point decrease ($p < 0.01$) in the annual deforestation rate in class 1, and a 0.006 percentage point increase ($p < 0.1$) in class 2B under state-level time trends, qualitatively consistent with the previously presented findings. Estimates are not significant with district-level time trends. These estimates are small in absolute terms, unsurprising given that underlying deforestation rates are low. In relative terms, these coefficients represent a 10 percent increase in the deforestation rate in Class 2B, and a 47 percent decrease in Class 1, calculated at mean values of baseline deforestation in each class.

I use the IV strategy as a robustness check on the main difference-in-differences results. Estimates are consistent in sign when significant, although as expected are weaker (due to the more limited use of data, see section 4.2.4). A positive average treatment effect is observed overall from use of the T=1000 threshold. The magnitude of this effect is larger than that observed under difference-in-differences. This is also to be expected given that in the cross section setup, this estimate describes the total change in tree cover between 2000 and 2014, rather than the more limited before-treatment after-treatment change in the case of difference-in-differences (which takes an average over these periods, reducing the possible observable effect size). Breakdown by class shows that the positive effect is driven by class 1, consistent with the previous heterogeneity analysis. No significant overall effect is observed when using the T=500 threshold. However, a negative and significant effect is seen in class 2B, as expected. Non-significant coefficient estimates are also negative, suggesting that the local average treatment effect around the 500-population threshold is lower than that around the 1000 population threshold. This could imply differential application of thresholds across regions, with more frontier-like regions (within districts and within classes) that see negative road-forest impacts having a higher proportion of smaller villages.

6. Discussion

The average difference-in-differences treatment effect, which is small in absolute terms, should be interpreted with a number of considerations in mind. First, the period of observation is relatively short, 2000-14. The median year of construction for treated villages is 2007 and thus the post-construction period for the typical village is 5-6 years once construction time is taking into account (results are robust to assumptions of 0-3 years of construction time)[22]. Given that tree regrowth is slow, and that households and firms take some time to respond to price and wage changes, estimates of tree cover change that were much greater than those observed would be very surprising. Treatment effects should also be considered in the context of low baseline tree cover in India: estimates represent a relative change of 0.6 to 1.2 percent of initial tree cover, above underlying trends. It should also be noted that VCF is a relatively conservative measure. For comparison, the VCF data has baseline average tree cover of 11.4 percent nationwide compared to 23.8 percent reported by the Forest Resources Assessment (FAO, 2015). The latter is based on a combination of field survey data from the Indian government and Landsat satellite imagery. This does not pose a threat to the validity of findings given that the satellite imagery and processing is consistent across space and time, but it does imply that these estimates of tree cover change are also conservative.

Given the economic and geographic diversity of India, the results in different settings are arguably at least as important as the national average treatment effect. The key difference observed – simultaneous positive and negative road impacts in different settings – is consistent with the conceptual model I propose based on differential changes in the value product of labor across economic sectors (agricultural versus non-agricultural, for instance). While Indian forests are not, in general, frontier forests in any way like those of South America (where much of the prior research on the drivers of deforestation has focused) some Indian settings arguably have more frontier-like characteristics than others. Class 2B, in particular, is more distant from urban centers, is less agriculturally developed, and has more initial forest cover. Negative road impacts are evident here. Opportunities for forest conversion are likely relatively high, and opportunities for local economic transitions away from forest-impacting activities potentially less. By contrast, class 1 is closer to urban centers, has better infrastructure and services (although it should be noted that these villages are not necessarily wealthier), and has less initial tree cover. Economic opportunities from further clearing are likely limited, and opportunities from forest gain or from activities associated with reduced forest pressure are likely greater.

The positive average treatment effect I find suggests that road construction is contributing to India's forest transition – the expansion of a country's forest cover following an extended period of deforestation (Mather, 1992). Forest transitions literature generally describes the forest transition phenomenon as proceeding in stages: (1) a period of initial high forest cover and low deforestation (i.e. a period in which there are large tracts of undeveloped land); (2) a period of high deforestation as agriculture expands in

---

[22] Program rules specified that construction was to be completed within one year or within two years if difficult engineering works were required, following a road's approval date. The rules also specified financial penalties for contractors who did not meet these construction time requirements (Ministry of Rural Development, 2012a).

response to increased demand and population growth; (3) a period of forest cover stabilization as agricultural development reaches the limits of land suitability and intensification reduces land demand via productivity increases; and finally, (4) a reforestation period as demand increases for forest products and urbanization further shrinks the agricultural labor force (Angelsen, 2007; Barbier et al., 2010; Rudel et al., 2005) (Figure 13). My average treatment effect suggests roads contribute to the processes driving the upswing in the fourth stage.

Additionally, my setting-specific finding, of simultaneous forest loss and forest gain in different locations, is consistent with a spatially differentiated conception of forest transitions in which economic and agricultural development proceed at different paces in different settings. I suggest that roads may contribute to both parts of the forest transitions curve: deforestation (the initial downward trajectory in countries' forest cover) and subsequent reforestation (the post-transition upswing in forest cover). While the existing literature describes countries' forest cover change as a temporal rather than a spatial phenomenon (as noted by Angelsen, 2007), differences in development pace implies that the stages of the forest transition are likely occurring simultaneously in different settings. Hence, while forest transitions unfold over decades or centuries, i.e., over periods longer than the 15 years of data available for this study, differences in settings allow me to see road impacts at what are arguably distinct stages of the forest transition. Arguably, more frontier-like conditions, such as those seen in class 2B, approximate stages 1-3 above, while the more agriculturally developed conditions approximate stage (4)[23]. To summarize, my results suggest that roads, as both a cause and a consequence of the economic development underlying forest transitions, contribute to forest transitions in the upswing (stage 4). To the extent the spatial variation proxies for temporal change, these results also demonstrate roads' contribution to the loss of tree cover, i.e. stages 1-2 of the forest transitions model. Where they do the latter, their impact is consistent with conclusions in the existing deforestation literature, which documents strong negative road impacts in frontier settings (Pfaff, 1999; Rudel et al., 2009).

Consideration of institutional heterogeneity is not currently part of this analysis, although this omission should not threaten identification. Over the last 30 years, India's forest policies have featured devolution of state power through joint forest management, state-driven reforestation programs, and restitution of forest rights, among other large-scale initiatives (Agrawal and Ostrom, 2016; Kumar and Kerr, 2012; Mather, 2007). Road impacts are almost certainly modified by the legacy of differential implementation of these

---

[23] In the case of India these distinctions are not enormous; I observe negative trends in tree cover in only the top decile of baseline tree cover, indicating that a large majority of India is either not changing or in the upswing of the forest transitions curve. Average tree cover change in class 2B is positive although lower than that in class 1.

programs across regions. In addition, new roads likely facilitate policy implementation. Roads may allow for better enforcement of state property rights, or improved state support for local enforcement of community rights, through improved access for forest officers (Ghate et al., 2009). Alternatively, increased access could lead to conflict over previously isolated resources. The institutional setting, such as the tenure regime under which different forests are managed, is also likely to modify the road-forest relationship. Open-access forests face increasing pressure under higher forest products prices. Commonly-held forests face similar pressures but not necessarily the same outcome: improved collective action that conserves forest may be incentivized under these conditions (Agrawal and Yadama, 1997). Private property rights, meanwhile, should allow landowners to capitalize on increasing forest products prices, encouraging the protection of existing forests or the establishment of new plantations (Robinson et al., 2013). Some of these institutional considerations will be the focus of future analysis.

7. Conclusion

These empirics first suggest that at least part of the rise in India's tree cover, and thus India's forest transition, can be attributed to the construction of new roads in rural settings. I estimate a positive causal impact for India overall, on the order of 0.6 to 1.2 percent relative to baseline tree cover, above underlying trends across the study period (2000-14). While previous studies provided indirect or theoretical support for the possibility of positive road impacts (Deng et al., 2011; Foster and Rosenzweig, 2003), this is the first direct evidence of such to my knowledge. While small in absolute terms, these estimates of positive tree cover outcomes from new roads apply to a relatively short period. Should trends continue the impact of the program would compound over time. I note, however, that countervailing general equilibrium effects could mitigate the long-term impact[24].

Regardless of the specific magnitude, this positive average treatment effect runs contrary to the conventional wisdom that new roads drive deforestation. This is not to say that the previous understanding, based on substantial empirical evidence and well-developed theory, is erroneous, or that concern regarding the environmental impact of road construction (e.g. Caro et al., 2014; Laurance et al., 2014b) is unfounded.

---

[24] For example, there may be a localized price rise for plantation timber following connection of a village to the road network. If many villages experience the same rise, and respond in concert by establishing new plantations, eventually the price of plantation goods will fall again, mitigating further incentives for new plantations. However, the equilibrium local price, and thus the equilibrium level of plantations, should still be higher than the initial price (how long that takes and whether such an equilibrium is meaningful, given many other contemporaneous economic changes, is another matter).

Nor does it mean that the net provision of environmental services must be positive from expanded tree cover, even in this study location[25]. Instead, the results simply demonstrate that particular settings, which have not been the subject of much prior scholarship, can give rise to positive road impacts, raising the possibility that improved spatial targeting of roads can not only avoid environmental costs but also deliver environmental benefits. I present evidence regarding the nature of settings in which this occurs in my heterogeneity analysis. The fact that I observe negative road impacts in frontier-like settings helps reconcile my results with the broader deforestation literature, which has focused on frontier-like settings (primarily in central and South America). However, in India, where forest cover is rising overall, negative road impacts on frontier settings are outweighed by positive road impacts in other settings.

I present models of potential mechanisms, and reduced-form results consistent with these models' predictions, but do not present evidence for the mechanisms themselves. It is possible that the relative impact of the potential mechanisms (described in section 0) can be causally explored with additional data, but doing so is beyond the scope of this paper. Additionally, neither deforestation data nor forest cover data distinguishes between natural forests and plantations, and past studies suggest that plantations are particularly important in India's land cover change (Foster and Rosenzweig, 2003; Mather, 2007; Rudel et al., 2005). Insights on mechanisms will be the focus of future extensions to this work. A second line of extension will be to quantify two further types of heterogeneity with important policy implications: ecological quality, and socio-economic impacts. The environmental benefits and costs of new road would be better revealed by forest quality metrics such as biodiversity, forest structure, and carbon sequestration, in addition to simple tree cover (see footnote 25). Secondly, socio-economic impacts of roads vary as a function of setting also. Quantifying over space both types of variation – environmental quality and socio-

---

[25] Environmental benefits of tree cover increase are not quantified in this study. Tree cover increases may occur due to new plantations, reforestation on previously cleared land, or densification of existing forest. While these secondary forests many contribute to hydrological services and carbon sequestration, their biodiversity benefits are typically much lower than those of primary forests (Barlow et al., 2007). In the context of my finding of opposing impacts in different settings, it is possible that higher value forests are being lost while (more expansive) low value forests are being gained. Thus despite the net positive impact in total tree cover, environmental services could be diminished due to differences in forest quality. Even within a particular positively-impacted locality, secondary forest gains could come partially at the expense of higher value primary forests. Similarly, secondary forests may have adverse impacts on water tables (Jackson et al., 2005). Data limitations mean that ecological and hydrological impacts are beyond the scope of this study.

economic impacts – reveals settings that lead to environmental-social trade-offs, and synergies, from rural road construction.

8. References

Aggarwal, S., 2014. Do rural roads create pathways out of poverty? Evidence from India, Job Market Paper.

Agrawal, A., Ostrom, E., 2016. Collective action, property rights, and decentralization in resource use in India and Nepal. Polit. Soc. 29, 485–514.

Agrawal, A., Wollenberg, E., Persha, L., 2014. Governing agriculture-forest landscapes to achieve climate change mitigation. Glob. Environ. Chang. 29, 270–280. doi:10.1016/j.gloenvcha.2014.10.001

Agrawal, A., Yadama, G., 1997. How do Local Institutions Mediate Market and Population Pressures on Resources? Forest Panchayats in Kumaon, India. Dev. Change 28, 435–465. doi:10.1111/1467-7660.00050

Allen, J.C., Barnes, D.F., 1985. The causes of deforestation in developing countries. Ann. Assoc. Am. Geogr. 75, 163–184.

Andersen, L.., Granger, C.W.J., Reis, E.J., Weinhold;, D., Wunder, S., 2002. The Dynamics of Deforestation and Economic Growth in the Brazilian Amazon. Cambridge University Press, Cambridge, UK.

Angelsen, A., 2007. Forest Cover Change in Space and Time: Combining the von Thünen and Forest Transition Theories. World Bank Policy Res. Work. Pap. 1–43. doi:doi:10.1596/1813-9450-4117

Angelsen, A., Kaimowitz, D., 2001. Introduction: the Role of Agricultural Technologies in Tropical Deforestation. Agric. Technol. Deforestation 1–17.

Angelsen, A., Kaimowitz, D., 1999. Rethinking the causes of deforestation: Lessons from economic models. World Bank Res. Obs. 14, 73–98.

Asher, S., Novosad, P., 2016. Market Access and Structural Transformation: Evidence from Rural Roads in India, Job Market Paper.

Baland, J.-M., Bardhan, P., Das, S., Mookherjee, D., Sarkar, R., 2006. Managing the environmental consequence of growth: Forest Degradation in the Indian Mid-Himalayas. India Policy Forum.

Barbier, E.B., Burgess, J.C., Grainger, A., 2010. The forest transition: Towards a more comprehensive theoretical framework. Land use policy 27, 98–107. doi:10.1016/j.landusepol.2009.02.001

Barlow, J., Gardner, T.A., Araujo, I.S., Ávila-Pires, T.C., Bonaldo, A.B., Costa, J.E., Esposito, M.C., Ferreira, L. V., Hawes, J., Hernandez, M.I.M., Hoogmoed, M.S., Leite, R.N., Lo-Man-Hung, N.F., Malcolm, J.R., Martins, M.B., Mestre, L. a. M., Miranda-Santos, R., Nunes-Gutjahr, A.L., Overal, W.L., Parry, L., Peters, S.L., Ribeiro-Junior, M.A., Silva, M.N.F. da, Motta, C. da S., Peres, C.A., 2007. Quantifying the biodiversity value of tropical primary, secondary, and plantation forests. Proc. Natl. Acad. Sci. 104, 18555–18560. doi:10.1073/pnas.0703333104

Bell, C., Van Dillen, S., 2012. How Does India's Rural Roads Program Affect the Grassroots? Findings from a Survey in Orissa. Land Econ. 90, 372–394.

Burton, G.P., 2011. Increasing forest cover amidst expanding commercial agriculture: An agent-based simulation of markets and rural livelihoods in the Indian Himalayas. University of Illinois at Urbana-Champaign.

Caro, T., Dobson, A., Marshall, A.J., Peres, C.A., 2014. Compromise solutions between conservation and road building in the tropics. Curr. Biol. 24, R722–R725. doi:10.1016/j.cub.2014.07.007

Chomitz, K.M., 2007. At Loggerheads ? Agricultural expansion, poverty reduction, and environment in the tropical forests, The World Bank.

Chomitz, K.M., Gray, D. a, 1996. Roads, Land Use, and Deforestation: A Spatial Model Applied to Belize. World Bank Econ. Rev. 10, 487–512. doi:10.1093/wber/10.3.487

Correia, S., 2014. reghdfe: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects.

Cropper, M., Puri, J., Griffiths, C., 2001. Predicting the Location of Deforestation : The Role of Roads and Protected Areas in North Thailand. Land Econ. 77, 172–186. doi:10.3368/le.77.2.172

Deininger, K., Minten, B., 1998. Determinants of Deforestation and the Economics of Protection: An Application to Mexico. Am. J. Agric. Econ. 84, 943–960.

Deng, X., Huang, J., Uchida, E., Rozelle, S., Gibson, J., 2011. Pressure cookers or pressure valves: Do roads lead to deforestation in China? J. Environ. Econ. Manage. 61, 79–94. doi:10.1016/j.jeem.2010.04.005

Esri, 2013. Layer Description: Terrain, ArcGIS online [WWW Document]. URL www.arcgis.com/home/item.html?id=58a541efc59545e6b7137f961d7de883

FAO, 2015. Global Forest Resources Assessment 2015. Desk reference. Food and Agriculture

Organization of the United Nations, Rome. doi:10.1002/2014GB005021

FAO, 2010. Global Forest Resources Assessment 2010.

Ferretti-gallon, K., Busch, J., 2014. What Drives Deforestation and What Stops It ? A Meta-Analysis of Spatially Explicit Econometric Studies (No. 361), Center for Global Development Working Paper.

Foley, J. a, Defries, R., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, G.C., Gibbs, H.K., Helkowski, J.H., Holloway, T., Howard, E. a, Kucharik, C.J., Monfreda, C., Patz, J. a, Prentice, I.C., Ramankutty, N., Snyder, P.K., 2005. Global consequences of land use. Science 309, 570–4. doi:10.1126/science.1111772

Foster, A., Rosenzweig, M., 2003. Economic growth and the rise of forests. Q. J. Econ. 118, 601–637.

Foster, A.D., Rosenzweig, M.R., 2004. Agricultural Productivity Growth, Rural Economic Diversity, and Economic Reforms: India, 1970–2000. Econ. Dev 52, 509–542.

Geist, H.J., Lambin, E.F., 2002. Proximate Causes and Underlying Driving Forces of Tropical Deforestation. Bioscience 52, 143–150.

Ghate, R., Mehra, D., Nagendra, H., 2009. Local institutions as mediators of the impact of markets on non-timber forest product extraction in central India. Environ. Conserv. 36, 51. doi:10.1017/S0376892909005311

Gollin, D., Rogerson, R., 2014. Productivity, transport costs and subsistence agriculture. J. Dev. Econ. 107, 38–48. doi:10.1016/j.jdeveco.2013.10.007

Hansen, M.C., et al., 2013. High-resolution global maps of 21st-century forest cover change. Science 342, 850–3. doi:10.1126/science.1244693

Heltberg, R., Arndt, T.C., Sekhar, N.U., 2000. Fuelwood consumption and Forest Degradation : Fuelwood Consumption A Household Model for Domestic Energy in Rural India Substitution. Land Econ. 76, 213–232. doi:10.2307/3147225

Jackson, R.B., Jobbágy, E.G., Avissar, R., Roy, S.B., Barrett, D.J., Cook, C.W., Farley, K.A., le Maitre, D.C., McCarl, B.A., Murray, B.C., 2005. Trading Water for Carbon with Biological Carbon Sequestration. Science (80-.). 310, 1944–1947. doi:10.1126/science.1119282

Khandker, S., Bakht, Z., Koolwal, G., 2009. The Poverty Impact of Rural Roads: Evidence from Bangladesh. Econ. Dev. Cult. Change 57, 685–722. doi:10.1086/598765

Kumar, K., Kerr, J.M., 2012. Democratic Assertions: The Making of India’s Recognition of Forest Rights

Act. Dev. Change 43, 751–771. doi:10.1111/j.1467-7660.2012.01777.x

Laurance, W.F., Clements, G.R., Sloan, S., O/'Connell, C.S., Mueller, N.D., Goosem, M., Venter, O., Edwards, D.P., Phalan, B., Balmford, A., Van Der Ree, R., Arrea, I.B., 2014. A global strategy for road building. Nature 513, 229–232. doi:10.1038/nature13717

Laurance, W.F., Peletier-Jellema, A., Geenen, B., Koster, H., Verweij, P., Van Dijck, P., Lovejoy, T.E., Schleicher, J., Van Kuijk, M., 2015. Reducing the global environmental impacts of rapid infrastructure expansion. Curr. Biol. 25, R259–R262. doi:10.1016/j.cub.2015.02.050

Makles, A., 2012. Stata tip 110: How to get the optimal k-means cluster solution. Stata J. 12, 347–351.

Mather, A., 2007. Recent Asian forest transitions in relation to forest transition theory. Int. For. Rev. 9, 491–502. doi:10.1505/ifor.9.1.491

Mather, A.S., 1992. The Forest Transition. Area 24, 367–379.

Mather, A.S., Fairbairn, J., Needle, C.L., 1999. TIhe Course and Drivers of the Forest Transition: the case of France. J. Rural Stud. 15, 65–90.

Mather, A.S., Fairbairn, J., Road, E., 2000. From Floods to Reforestation: The Forest Transition in Switzerland. Environ. Hist. Camb. 6, 399–421.

Ministry of Home Affairs, 2016. Indian Census. Office of the Registrar General and Census Commissioner [WWW Document]. URL http://censusindia.gov.in/ (accessed 7.8.16).

Ministry of Rural Development, 2012a. Pradhan Mantri Gram Sadak Yojana – Programme Guidelines. New Delhi.

Ministry of Rural Development, 2012b. Pradhan Mantri Gram Sadak Yojana - Program Guidlines. New Delhi.

Nelson, G.C., Hellerstein, D., 1997. Do Roads Cause Deforestation ? Using Satellite Images in Econometric Analysis of Land Use. Am. J. Agric. Econ. 79, 80–88.

Perz, S.G., 2007. Grand Theory and Context-Specificity in the Study of Forest Dynamics: Forest Transition Theory and Other Directions. Prof. Geogr. 59, 105–114. doi:10.1111/j.1467-9272.2007.00594.x

Pfaff, A., 1999. What drives deforestation in the Brazilian Amazon?: evidence from satellite and socioeconomic data. J. Environ. Econ. Manage. 37, 26–43.

Pfaff, A., Amacher, G.S., Sills, E.O., Coren, M.J., Streck, C., Lawlor, K., 2013. Deforestation and Forest Degradation: Concerns, Causes, Policies, and Their Impacts, Encyclopedia of Energy, Natural Resource, and Environmental Economics. doi:10.1016/B978-0-12-375067-9.00052-8

Pfaff, A., Barbieri, A., Ludewigs, T., Merry, F., Perz, S., Reis, E., 2009. Road Impacts in Brazilian Amazonia. Geophys. Monogr. Ser.

Pfaff, A., Robalino, J., Reis, E.J., Walker, R., Perz, S., Laurance, W.F., Bohrer, C., Aldrich, S., Arima, E., Caldas, M., Kirby, K., 2016. Can strategically locating roads "clean" development? Roads' heterogeneous impacts on the Brazilian Amazon forest frontier.

Pfaff, A., Robalino, J., Walker, R., 2007. Road Investments, Spatial Spillovers, and Deforestation in the Brazilian Amazon. J. Reg. Sci. 47, 109–123.

Robinson, B.E., Holland, M.B., Naughton-Treves, L., 2013. Does secure land tenure save forests? A meta-analysis of the relationship between land tenure and tropical deforestation. Glob. Environ. Chang. doi:10.1016/j.gloenvcha.2013.05.012

Rudel, T.K., 1989. Population, development, and tropical deforestation: A cross-national study. Rural Sociol. 54, 327–338.

Rudel, T.K., Coomes, O.T., Moran, E., Achard, F., Angelsen, A., Xu, J., Lambin, E., 2005. Forest transitions: towards a global understanding of land use change. Glob. Environ. Chang. 15, 23–31. doi:10.1016/j.gloenvcha.2004.11.001

Rudel, T.K., Defries, R., Asner, G.P., Laurance, W.F., 2009. Changing drivers of deforestation and new opportunities for conservation. Conserv. Biol. 23, 1396–1405. doi:10.1111/j.1523-1739.2009.01332.x

Sagara, T., 2011. India Place Finder [WWW Document]. URL http://india.csis.u-tokyo.ac.jp/

Townshend, J., Hansen, M., Carroll, M., DiMiceli, C., Sohlberg, R., Huang, C., 2011. User Guide for the MODIS Vegetation Continuous Fields product Collection 5 version 1.

van de Walle, D., 2009. Impact evaluation of rural road projects. J. Dev. Eff. 1, 15–36. doi:10.1080/19439340902727701

van de Walle, D., 2002. Choosing Rural Road Investments to Help Reduce Poverty. World Dev. 30, 575–589.

Veld, K. van't, Narain, U., Gupta, S., Chopra, N., Singh, S., 2006. India's Firewood Crisis Re-examined, Resources for the Future Working Paper.

von Thunen, J.H., 1826. The Isolated State (1996 trans., eds. Hall, P. and Wartenberg, C. M.). Pergamon Press, Oxford, UK.

Warr, P., 2010. Roads and poverty in rural laos: An econometric analysis. Pacific Econ. Rev. 15, 152–169. doi:10.1111/j.1468-0106.2009.00494.x

World Bank, 2015. Forest area (percent of land area) [WWW Document]. URL http://data.worldbank.org/indicator/AG.LND.FRST.ZS (accessed 10.5.16).

9. Figures

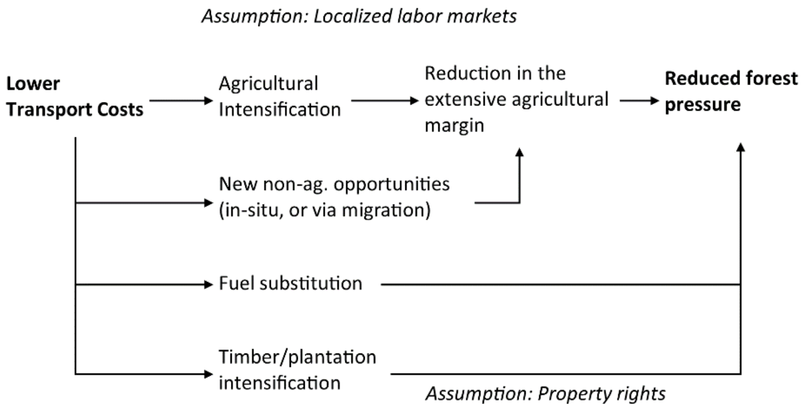


Figure 1: Hypothesized relationships between roads (lower transport costs) and land cover change.

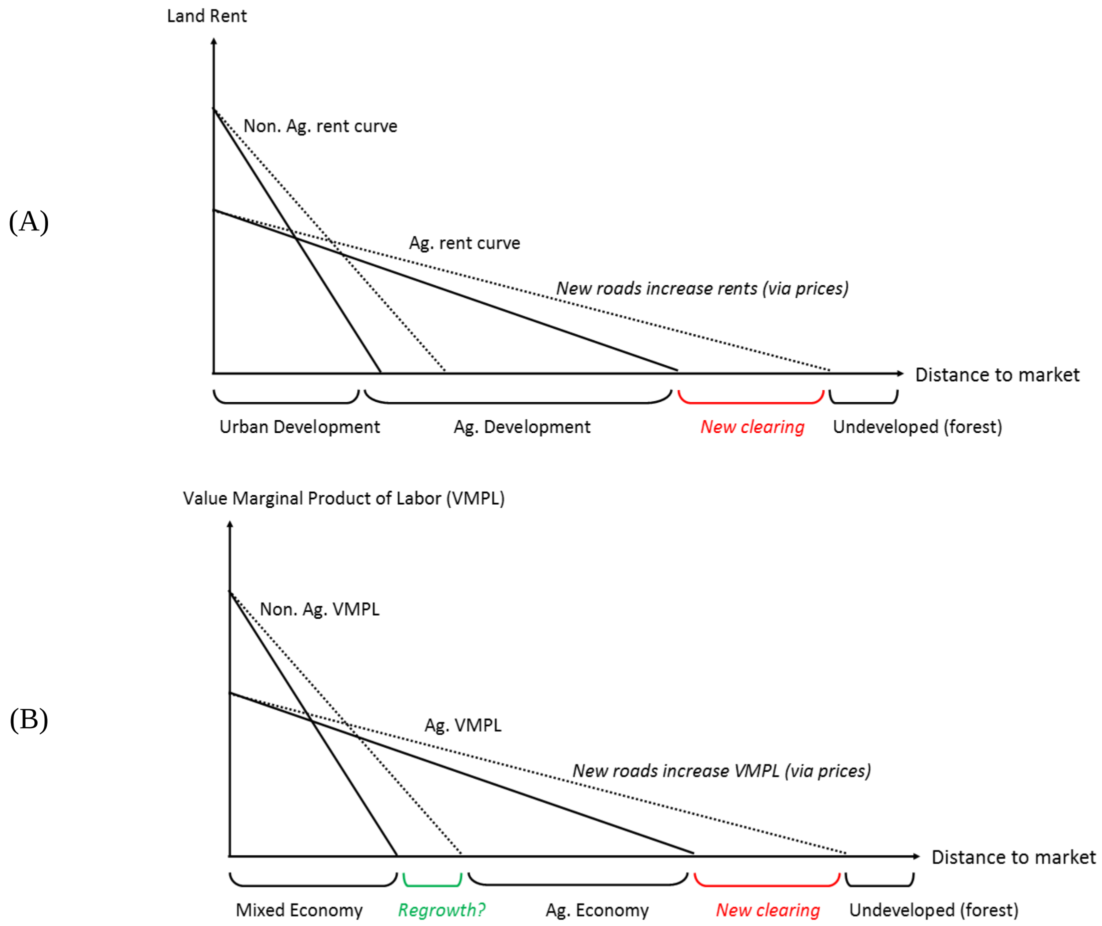


Figure 2: Models of land use as a function of distance to market. (A) Traditional von Thünen model. New roads increase land rents, increasing clearing on the frontier where agriculture becomes profitable. (B) Altered von Thünen model: new roads increase the value marginal product of labor, encouraging a move of labor towards the agricultural sector in frontier areas, and a move towards the non-agricultural sector in non-frontier areas (due to differential change in profitability across sectors).

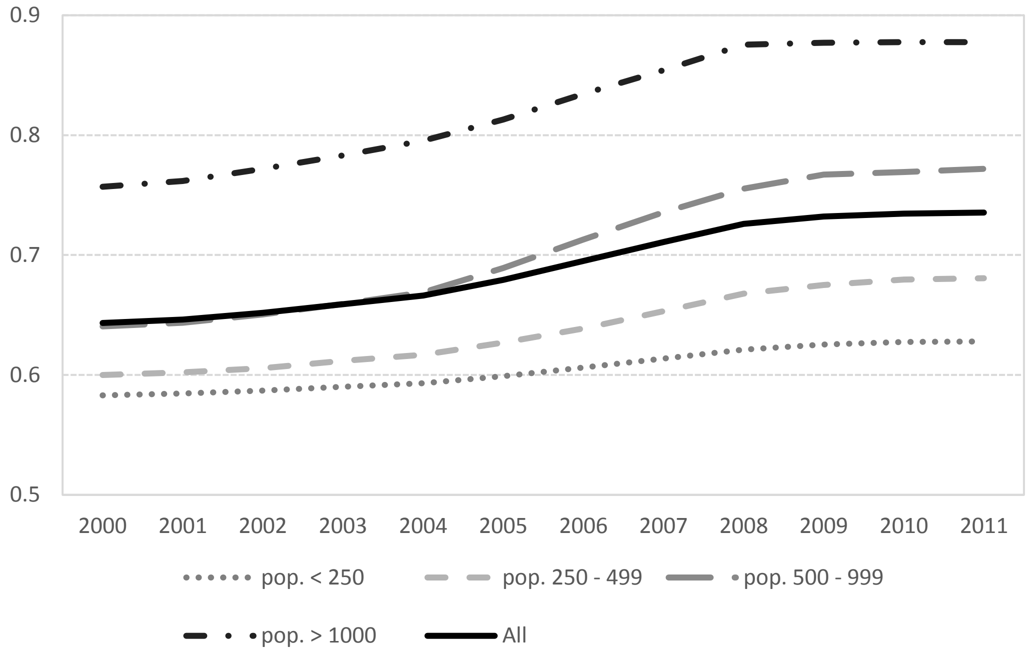


Figure 3: Cumulative probability of a habitation receiving a new road by population category over time (all habitations included, those connected and unconnected in baseline).

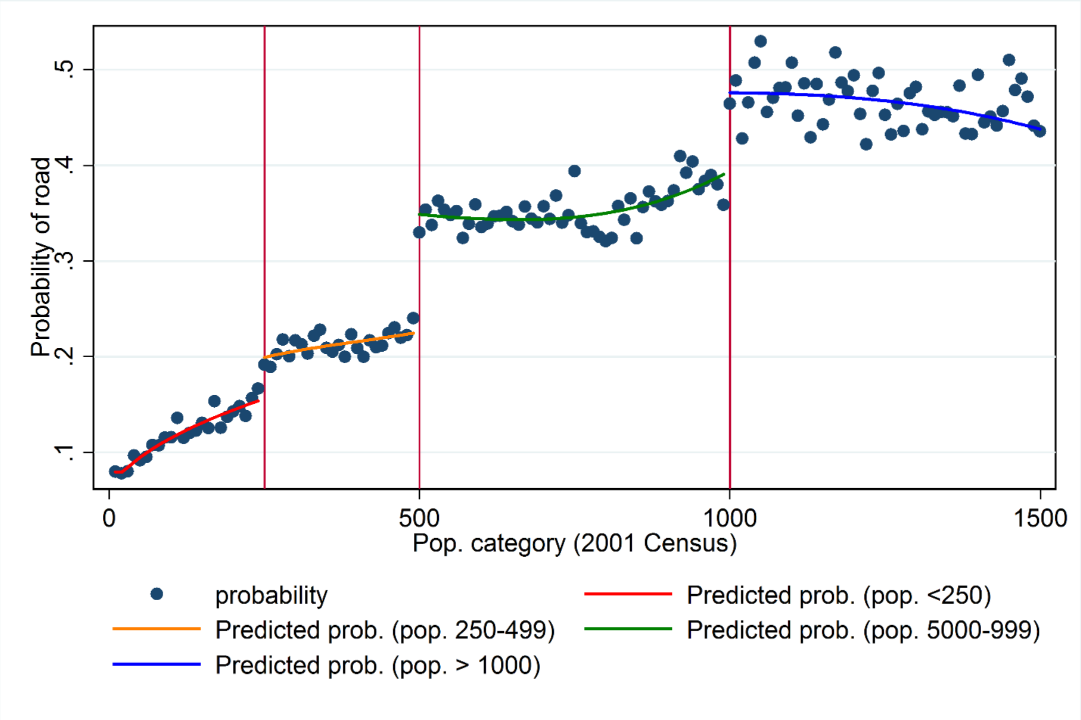


Figure 4: Probability of receiving a PMGSY road project by habitation populations using Census 2001 population data (note: zero population habitations removed, 99th percentile and above removed, only villages with one habitation can be considered, villages unconnected in baseline only).

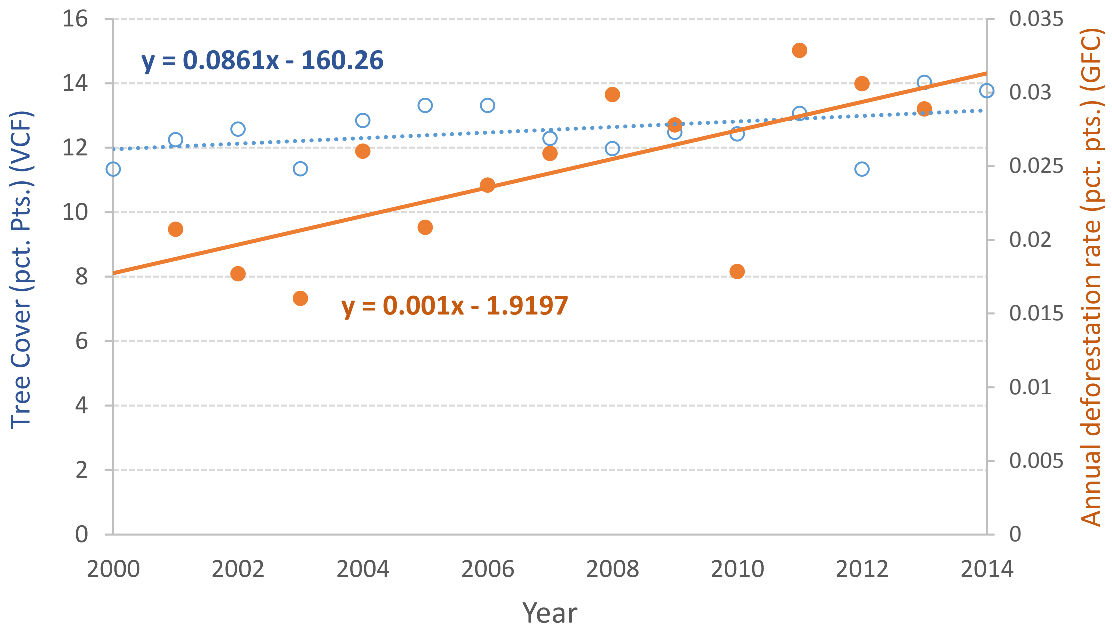


Figure 5: Forest cover level (MOD44B v.51, Vegetation Continuous Fields, Townshend, et al. 2011) (dash line, hollow data points) and annual deforestation rate (Global Forest Change, Hansen, et al. 2013) (solid line, colored data points) for India, 2000-14. Linear trends with equations included on graph. Data points calculated using pixel level data, giving percentages as a function of total India land area.

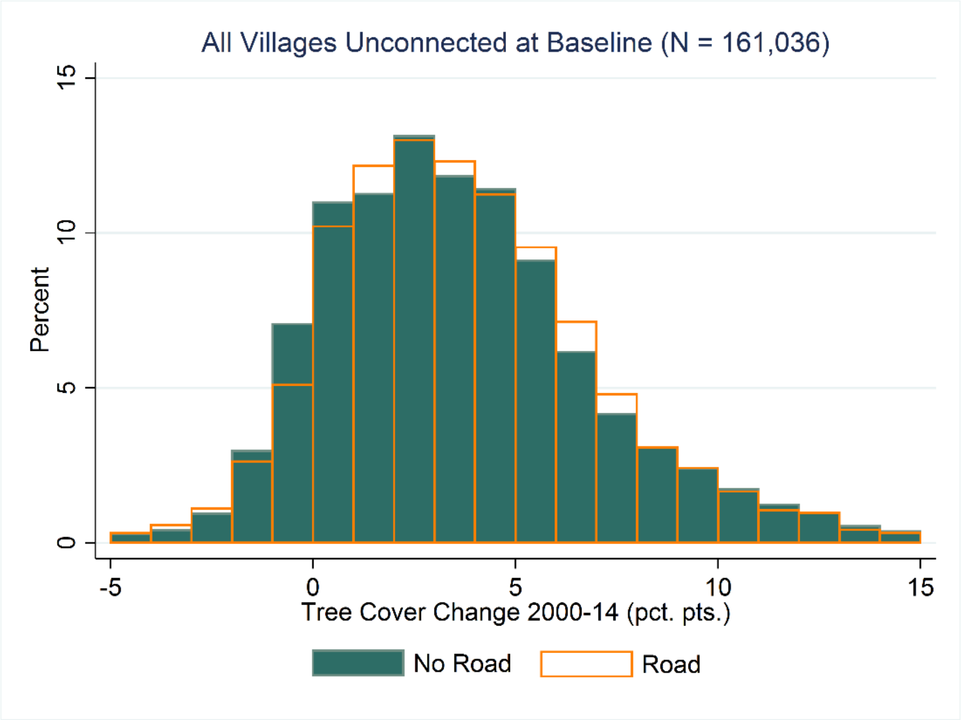


Figure 6: Overlaid histograms of tree cover change between baseline and endline. It should be noted that the MOD44B Vegetation Continuous Fields data contains annual fluctuations, and so the absolute changes represented here are a partially a function of the anomalies in the two years used to calculate the change (2000 and 2014).

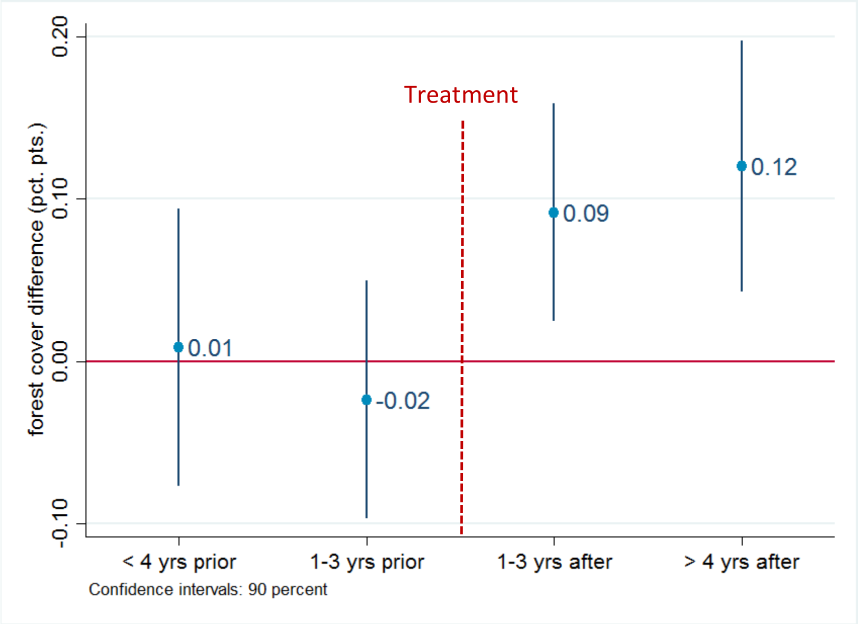


Figure 7: Event study plot, average treatment effects for periods before and after treatment for all unconnected villages at baseline (N = 161,036).

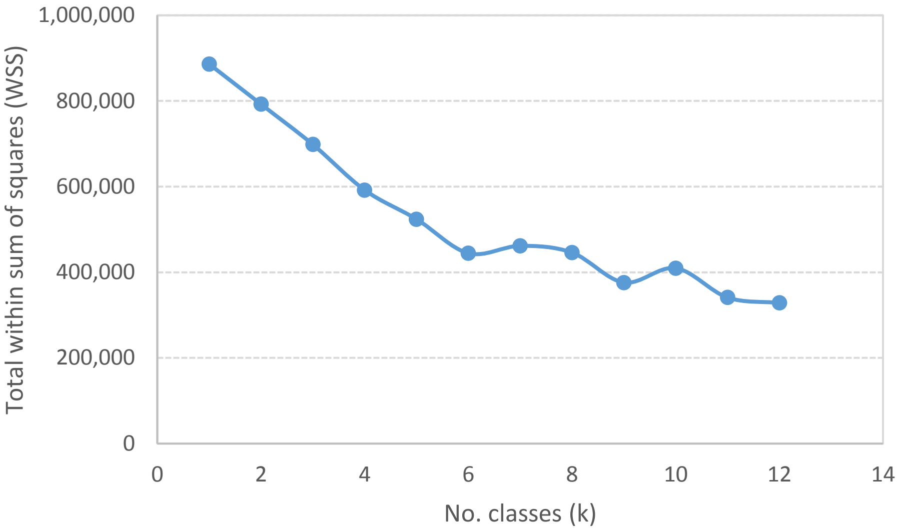


Figure 8: Total within sum of squares as a function of the number of classes.

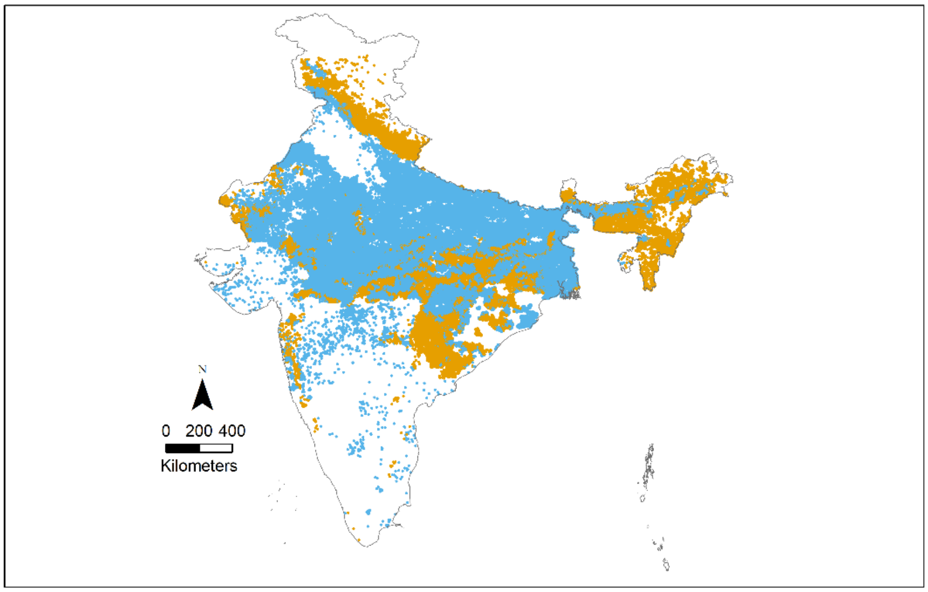


Figure 9: Location of villages in each class (k=2 partitions): Class 1 = blue, Class 2 = orange.

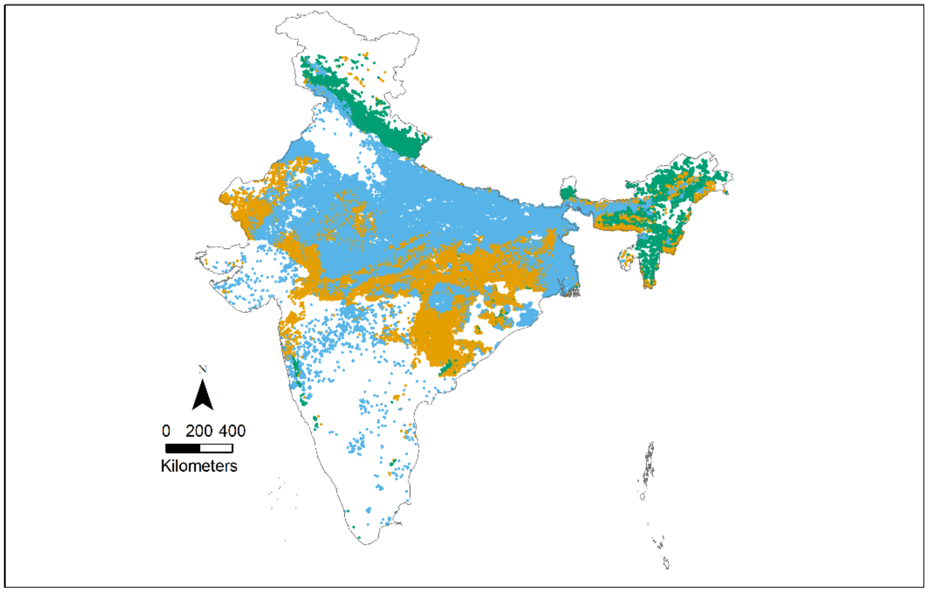
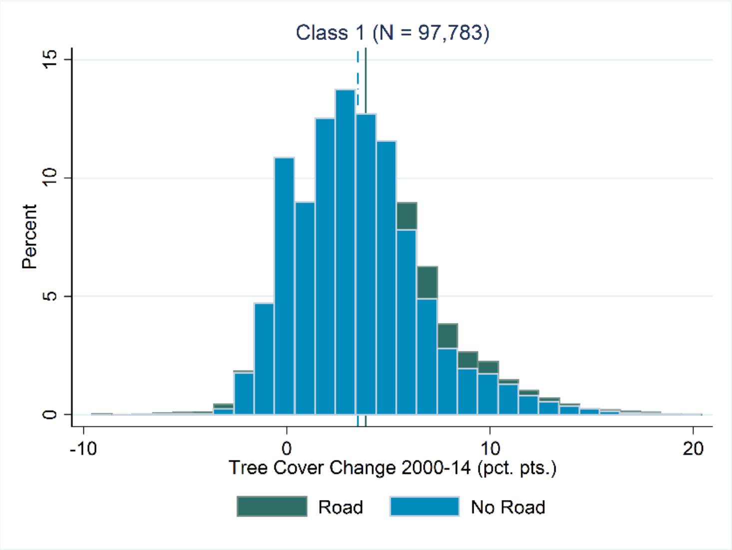


Figure 10: Location of villages in each class (k=3 partitions): Class 1 = blue, Class 2A = orange, Class 2B = green.

Class 1



Class 2B

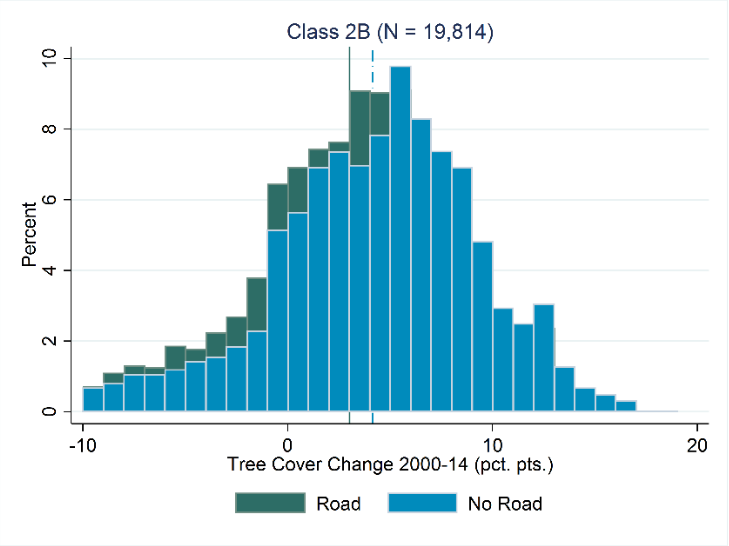


Figure 11: Overlaid histograms of tree cover change between baseline and endline for class 1 and class 2B villages of the 3 class partition (the two most opposite classes in terms of road impacts). It should be noted that the MOD44B Vegetation Continuous Fields data contains annual fluctuations, and so the absolute changes represented here are a partially a function of the anomalies in the two years used to calculate the change (2000 and 2014).

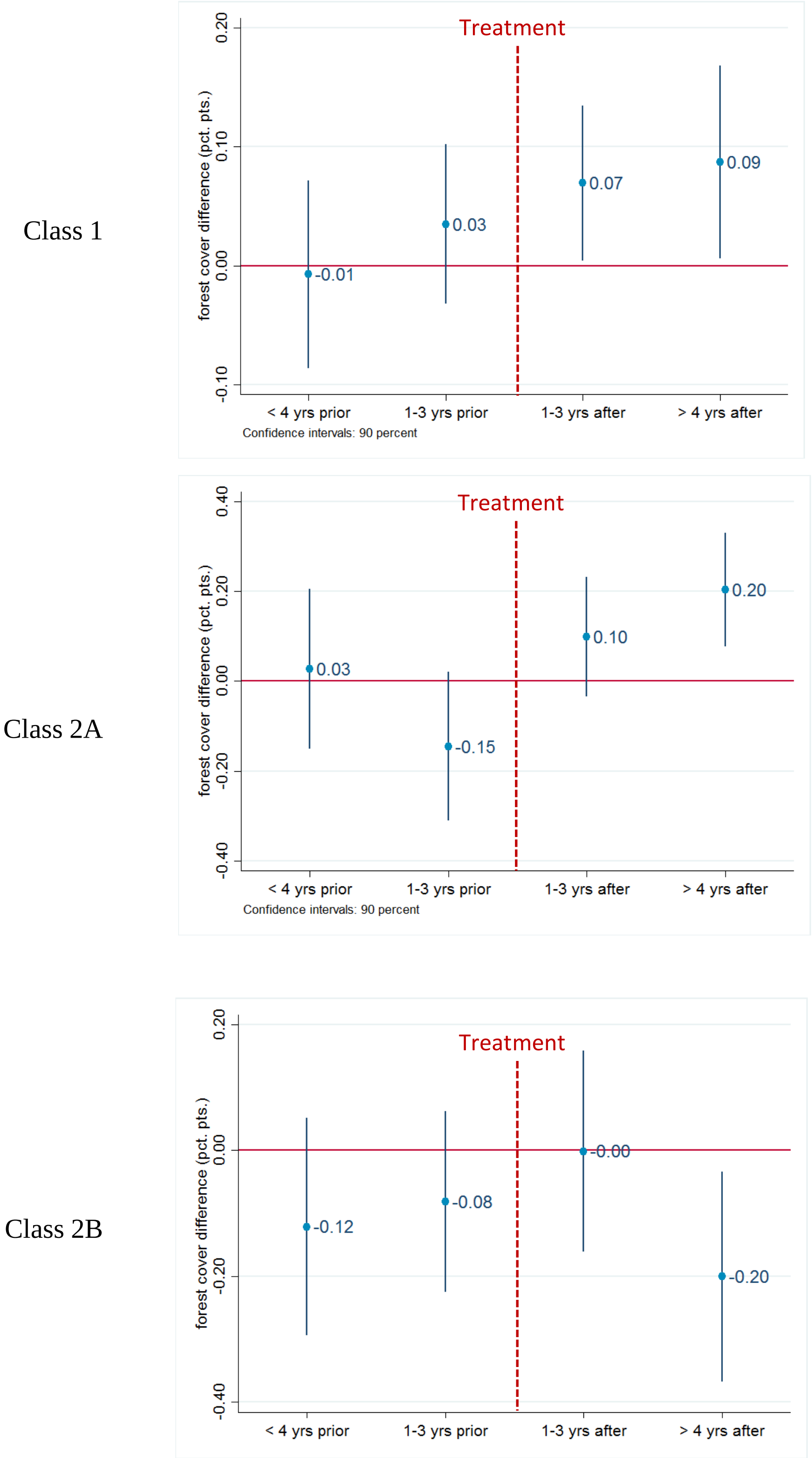


Figure 12: Event study plot. Average treatment effects for periods before and after treatment for classes in the 3-class partition. All villages unconnected at baseline.

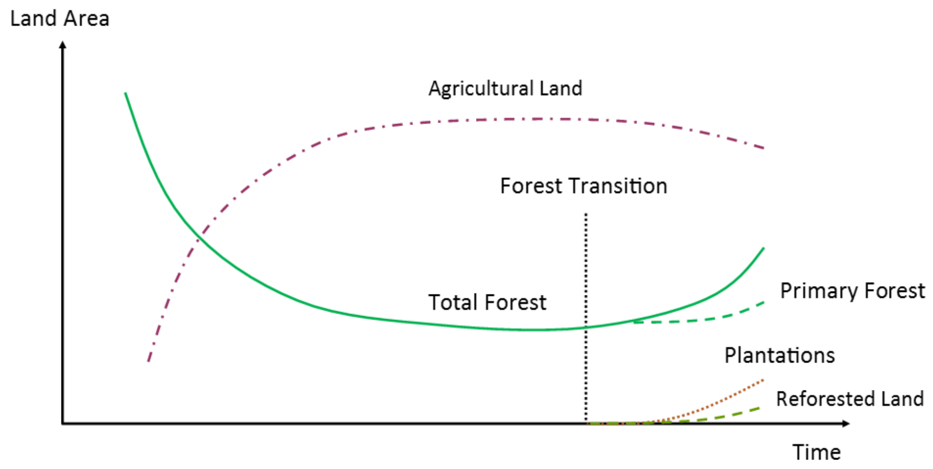


Figure 13: The forest transitions hypothesis. Forest cover falls and then rises over time as economic and agricultural development proceeds (adapted from Barbier, *et al.* 2010).

10. Tables

Table 1: Summary statistics (mean, standard deviation in parentheses): All villages (2001 census) by baseline road access status (connected or unconnected with a paved road).

	Unconnected (2001)	**Connected (2001)**	**All**
Proportion of villages	0.34	0.66	1.00
Population	817	1762	1439
	(1013)	*(2298)*	*(2006)*
Population growth (2001-11) (percent)	0.25	0.19	0.21
	(2.77)	*(1.39)*	*(1.99)*
Scheduled tribe population (avg. proportion)	0.19	0.13	0.15
	(0.35)	*(0.27)*	*(0.30)*
Electricity connection (proportion of villages)	0.57	0.90	0.79
	(0.49)	*(0.30)*	*(0.41)*
Credit Society (proportion of villages)	0.07	0.27	0.20
	(0.25)	*(0.45)*	*(0.40)*
Primary School (proportion of villages)	0.69	0.89	0.82
	(0.46)	*(0.32)*	*(0.39)*
Primary health center (proportion of villages)	0.06	0.23	0.17
	(0.24)	*(0.42)*	*(0.38)*
Distance to nearest town (kilometers)	26.75	20.38	22.56
	(26.18)	*(20.19)*	*(22.63)*

Table 2: Road connections at habitation and village level.

Statistic	**Unit**	
Number:	Habitations	754 813
	Villages	458 560
	Sub-districts	5 310
	Districts`	564
Average population:	Habitation[2]	824
	Village[2]	2 992
Median population:	Habitation[2]	474
	Village[2]	1 335
Average Number:	Habitations per village[2]	1.65
	Villages per sub-district	86.36
	Sub-districts per district	9.41
Median Number:	Habitations per village	2
	Villages per sub-district	56
	Sub-districts per district	7
Connected at baseline[1] (proportion):	Habitations	0.64
	Villages – partial or complete [3]	0.66
Connected at baseline[1] (proportion):	Villages - complete (all habitations within village) [2]	0.57
	Villages - partial (some habitations within village) [2]	0.06
	Villages - none (no habitations within village) [2]	0.37

[1]: Baseline = 2000

[2]: PMGSY data (habitation level)

[3]: Census 2001 data (village level)

Table 3: Road connections over time under the PMGSY program.

Year	**Proportion of villages with road connection**	
Sample:	All	Unconnected at baseline
2000	63.2%	0.0%
2001	63.5%	0.8%
2002	64.2%	2.6%
2003	64.2%	2.6%
2004	65.1%	5.0%
2005	66.0%	7.5%
2006	67.6%	11.8%
2007	69.6%	17.3%
2008	71.4%	22.4%
2009	73.4%	27.6%
2010	74.2%	29.8%
2011	74.4%	30.4%
2012	74.5%	30.6%
Mean (across years)	69.3%	16.7%
No. villages in sample	437,820	161,036

Table 4: OLS panel regression with village-level fixed effects. Sample comprises all villages unconnected at baseline.

Dep. Var: tree cover (MOD44B) (10km) (pct. pts)	**All villages unconnected at baseline**					
Road connection (binary)	0.122***	0.060***				
	(0.034)	*(0.019)*				
Road connection (vil. pop. proportion)			0.128***	0.061***		
			(0.036)	*(0.020)*		
Road connection (vil. pop. sum) (‘000)					0.067***	0.016*
					(0.021)	*(0.009)*
R-square	0.97	0.98	0.97	0.98	0.97	0.98

Dep. Var: tree cover (MOD44B) (5km) (pct. pts)	**All villages unconnected at baseline**					
Road connection (binary)	0.113***	0.053**				
	(0.035)	*(0.021)*				
Road connection (vil. pop. proportion)			0.117***	0.051**		
			(0.036)	*(0.021)*		
Road connection (vil. pop. sum) (‘000)					0.062***	0.014
					(0.022)	*(0.010)*
R-square	0.97	0.98	0.97	0.98	0.97	0.98
State*Year FEs	Yes	No	Yes	No	Yes	No
District*Year FEs	No	Yes	No	Yes	No	Yes
No. obs.	2,093,468	2,093,234	2,093,468	2,093,234	2,093,468	2,093,234
No. Villages	161,036	161,018	161,036	161,018	161,036	161,018

* p<0.10, ** p<0.05, *** p<0.01

Heteroskedasticity-robust standard errors clustered on district

Two year lead time is allowed for road construction

Standard errors in parentheses

Table 5: OLS panel regression with village-level fixed effects. Sample comprises all villages.

Dep. Var: tree cover (MOD44B) (10km) (pct. pts)	**All villages (connected and unconnected at baseline)**					
Road connection (binary)	0.087***	0.039**				
	(0.032)	*(0.016)*				
Road connection (vil. pop. proportion)			0.094***	0.041**		
			(0.033)	*(0.017)*		
Road connection (vil. pop. sum) (‘000)					0.048***	0.005
					(0.018)	*(0.007)*
R-square	0.96	0.98	0.96	0.98	0.96	0.98

Dep. Var: tree cover (MOD44B) (5km) (pct. pts)	**All villages (connected and unconnected at baseline)**					
Road connection (binary)	0.079**	0.031*				
	(0.031)	*(0.018)*				
Road connection (vil. pop. proportion)			0.085***	0.031*		
			(0.033)	*(0.019)*		
Road connection (vil. pop. sum) (‘000)					0.044**	0.004
					(0.018)	*(0.008)*
R-square	0.95	0.97	0.95	0.97	0.95	0.97
State*Year FEs	Yes	No	Yes	No	Yes	No
District*Year FEs	No	Yes	No	Yes	No	Yes
No. obs.	6,567,300	6,567,285	6,567,300	6,567,285	6,567,300	6,567,285
No. Villages	437,820	437,819	437,820	437,819	437,820	437,819

* p<0.10, ** p<0.05, *** p<0.01

Heteroskedasticity-robust standard errors clustered on district

Two year lead time is allowed for road construction

Standard errors in parentheses

Table 6: OLS panel regression with village-level fixed effects. Sample comprises all villages unconnected at baseline and treated during program.

Dep. Var: tree cover (MOD44B) (10km) (pct. pts)	**Treated villages only**					
Road connection (binary)	0.118***	0.054**				
	(0.041)	*(0.024)*				
Road connection (vil. pop. proportion)			0.127***	0.060**		
			(0.041)	*(0.025)*		
Road connection (vil. pop. sum) (‘000)					0.034*	-0.003
					(0.018)	*(0.009)*
R-square	0.97	0.98	0.97	0.98	0.97	0.98
Dep. Var: tree cover (MOD44B) (5km) (pct. pts)	**Treated villages only**					
Road connection (binary)	0.103**	0.044*				
	(0.040)	*(0.026)*				
Road connection (vil. pop. proportion)			0.108***	0.045*		
			(0.040)	*(0.026)*		
Road connection (vil. pop. sum) (‘000)					0.03	-0.006
					(0.018)	*(0.010)*
R-square	0.96	0.98	0.96	0.98	0.96	0.98
State*Year FEs	Yes	No	Yes	No	Yes	No
District*Year FEs	No	Yes	No	Yes	No	Yes
No. obs.	740,025	739,815	740,025	739,815	740,025	739,815
No. Villages	49,335	49,321	49,335	49,321	49,335	49,321

* p<0.10, ** p<0.05, *** p<0.01

Heteroskedasticity-robust standard errors clustered on district

Two year lead time is allowed for road construction

Standard errors in parentheses

Table 7: OLS regression of tree cover trend during early years of program on treatment (receiving a road at any time during the program).

Period:	**2000-05**				**2000-04**				**2000-03**			
Tree Cover Trend	0.058***	0.035**	0.000	0.001	0.043***	0.016	-0.010	-0.007	0.036***	0.015	-0.005	-0.007
	(0.013)	*(0.014)*	*(0.012)*	*(0.012)*	*(0.012)*	*(0.012)*	*(0.010)*	*(0.01)*	*(0.009)*	*(0.009)*	*(0.007)*	*(0.007)*
Constant	0.279***	0.146***	0.224***	0.159**	0.289***	0.151***	0.229***	0.160***	0.300***	0.154***	0.225***	0.162**
	(0.015)	*(0.023)*	*(0.066)*	*(0.063)*	*(0.014)*	*(0.023)*	*(0.064)*	*(0.062)*	*(0.013)*	*(0.022)*	*(0.066)*	*(0.062)*
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
State FEs	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
District FEs	No	No	No	No	No	No	No	No	No	No	No	No
Adjusted R-square	0.01	0.06	0.08	0.11	0	0.06	0.08	0.11	0.01	0.06	0.08	0.11
No. Villages	161,036	137,818	161,036	137,818	161,036	137,818	161,036	137,818	161,036	137,818	161,036	137,818

* p<0.10, ** p<0.05, *** p<0.01

Heteroskedasticity-robust standard errors clustered on district

Dep. Var: Change in MOD44B v.51 over period specified above

Controls: village population, scheduled tribe population proportion, primary school, primary health center, area of irrigated land, distance to nearest urban center, power supply, average slope, initial forest cover.

Table 8: Event study analysis. OLS panel regression with village-level fixed effects. Independent variables are periods before and after treatment. Sample comprises all villages unconnected at baseline.

Dep. Var: tree cover (MOD44B) (10km) (pct. pts)	**All villages**	**Class 1**	**Class 2A**	**Class 2B**
Road connection (binary): ≥ 5 years prior	0.008	-0.007	0.028	-0.121
	(0.052)	*(0.051)*	*(0.108)*	*(0.104)*
Road connection (binary): 1-4 years prior	-0.024	0.035	-0.145	-0.082
	(0.044)	*(0.043)*	*(0.101)*	*(0.087)*
Road connection (binary): 1-4 years after	0.092**	0.069*	0.099	-0.001
	(0.041)	*(0.042)*	*(0.081)*	*(0.096)*
Road connection (binary): ≥ 5 years after	0.120**	0.087*	0.203***	-0.201**
	(0.047)	*(0.052)*	*(0.077)*	*(0.101)*
Adjusted R-square	0.97	0.91	0.93	0.97
State*Year FEs	Yes	Yes	Yes	Yes
No. obs.	2,415,540	1,466,730	605,190	297,195
No. Villages	161,036	97,782	40,346	19,813

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Heteroskedacity-robust standard errors clustered on district

Standard errors in parentheses

Table 9: Instrumental variable regressions (2SLS) on village cross-section. First stage dependent variable is binary road connection (whether village received a road at any point during study period). Sample comprises all villages unconnected at baseline.

Threshold (IV):	**T = 1000**		**T = 500**	
Pred. road connection	1.209*		-0.441	
	(0.701)		*(0.395)*	
Pred. road connection * Class 1		1.249*		-0.191
		(0.646)		*(0.545)*
Pred. road connection * Class 2A		0.599		-0.493
		(1.356)		*(0.346)*
Pred. road connection * Class 2B		3.223		-1.999**
		(2.270)		*(0.800)*
Controls	Yes	Yes	Yes	Yes
Population controls	Yes	Yes	Yes	Yes
Class FEs	Yes	Yes	Yes	Yes
District FEs	Yes	Yes	Yes	Yes
Adjusted R-square	0.6	0.59	0.6	0.59
No. villages	21,581	21,581	52,042	52,042

* p<0.10, ** p<0.05, *** p<0.01

Heteroskedacity-robust standard errors clustered on district

Dep. Var: Change in MOD44B.v51 (2000-14)

Bandwidth: T +/- 250

Table 10: Mean standardized values of baseline descriptive variables for village classes (2 and 3 class partitions). Colors highlight signs and magnitudes (blue are above average, orange are below average).

	K = 2 class partition		**K = 3 class partition**		
	Class 1	Class 2	Class 1	Class 2A	Class 2B
Total Irrigated Area	0.10	-0.25	0.14	-0.19	-0.26
Forest Cover (2000)	-0.36	0.94	-0.33	-0.16	1.91
Crop Cover (2001)	0.44	-1.20	0.56	-0.75	-1.24
Electricity connection	0.06	-0.15	0.10	-0.30	0.13
Credit Society	0.06	-0.17	0.11	-0.20	-0.13
Primary School	0.08	-0.14	0.07	0.06	-0.29
Primary health center	0.03	-0.07	0.05	-0.08	-0.04
Distance to nearest town	-0.26	0.76	-0.31	0.57	0.50
Slope	-0.40	1.11	-0.42	-0.04	2.21
Scheduled tribe population	-0.26	0.72	-0.52	1.30	-0.07
Villages in class (total)	115,398	42,545	97,783	40,346	19,814
Villages in class (percent)	**73.06**	**26.94**	**61.91**	**25.54**	**12.55**

Table 11: OLS panel regression with village-level fixed effects and interaction terms on village class. Coefficients have been adjusted to give treatment effects relative to zero (i.e. no reference group required).

Dep. Var: tree cover (MOD44B) (10km) (pct. pts)	**All villages unconnected at baseline**		**Treated villages only**	
	3 Class Partition			
Road connection * Class 1	0.054	0.109**	0.066	0.074
	(0.062)	*(0.048)*	*(0.066)*	*(0.047)*
Road connection * Class 2A	-0.370***	-0.242***	-0.484***	-0.292***
	(0.121)	*(0.063)*	*(0.163)*	*(0.098)*
Road connection * Class 2B	0.247***	0.095***	0.286***	0.125***
	(0.046)	*(0.021)*	*(0.053)*	*(0.031)*
R-square	0.97	0.98	0.97	0.98
	2 Class Partition			
Road connection * Class 1	0.229***	0.104***	0.259***	0.124***
	(0.044)	*(0.021)*	*(0.052)*	*(0.031)*
Road connection * Class 2	-0.157**	-0.062	-0.172**	-0.093
	(0.074)	*(0.050)*	*(0.085)*	*(0.057)*
R-square	0.97	0.98	0.97	0.98
State*Year FEs	Yes	No	Yes	No
District*Year FEs	No	Yes	No	Yes
No. obs.	2,053,259	2,053,038	630,396	630,227
No. Villages	157,943	157,926	48,492	48,479

* p<0.10, ** p<0.05, *** p<0.01

Heteroskedacity-robust standard errors clustered on district

Two year lead time is assumed for road construction

Standard errors in parentheses

Table 12: OLS panel regression with village-level fixed effects and interaction terms on village class. Coefficients have been adjusted to give treatment effects relative to zero (i.e. no reference group required).

DV: annual deforestation (Hansen, et al. 2013) (pct. pts)	**All villages unconnected at baseline**				**All villages unconnected at baseline (with some deforestation)**			
Road connection	-0.0004		-0.00047		-0.00053		-0.00066	
	(0.00068)		*(0.00055)*		*(0.00098)*		*(0.00077)*	
Road connection * Class 1		-0.00178***		-0.0005		-0.00273***		-0.00073
		(0.00067)		*(0.00033)*		*(0.00104)*		*(0.00052)*
Road connection * Cluster 2A		-0.00018		-0.00157		-0.00043		-0.00216
		(0.00157)		*(0.00171)*		*(0.00212)*		*(0.00220)*
Road connection * Class 2B		0.00594*		0.00222		0.00586*		0.00221
		(0.00326)		*(0.00186)*		*(0.00326)*		*(0.00187)*
Adjusted R-square	0.67	0.67	0.73	0.73	0.66	0.66	0.72	0.72
State*Year FEs	Yes	Yes	No	No	Yes	Yes	No	No
District*Year FEs	No	No	Yes	Yes	No	No	Yes	Yes
No. obs.	1,932,180	1,895,316	1,931,964	1,895,112	1,279,308	1,254,024	1,279,020	1,253,748
No. Villages	161,015	157,943	160,997	157,926	106,609	104,502	106,585	104,479

* p<0.10, ** p<0.05, *** p<0.01

Heteroskedacitity-robust standard errors clustered on district

Two year lead time is assumed for road construction

Standard errors in parentheses

11. Supplementary Section: A Static Model of Household Labor Allocation

I consider decisions at the household level. Households have a fixed time budget,[26] L, which they may allocate to intensive agricultural activities (those not at the forest margin) l_{A1}, extensive agricultural activities (those at the forest margin or within forest areas) l_{A2}, non-agricultural production l_N, and firewood collection, l_F,. Growth in plantations is not considered in this model. To simplify the analysis I treat leisure time as fixed. Hence:

$$L = l_{A1} + l_{A2} + l_N + l_F \tag{1}$$

Firewood is not typically purchased in rural India (Veld, *et al.* 2006), and is consumed within the home. Firewood production, q_F, is a function of forest labor time, and is strictly increasing and concave: $q_F = l_F^{\beta}$, where $\beta < 1$. Non-forest energy (liquefied natural gas or kerosene), q_E, must be purchased at the fixed market price, p_E. Total energy consumption is represented by:

$$E = {l_F}^{\beta} + q_E \tag{2}$$

Agricultural labor produces agricultural goods via increasing and concave production functions: $q_{A1}(l_{A1}) = {l_{A1}}^{\alpha}$, and $q_{A2}(l_{A2}) = {l_{A2}}^{\beta}$, where $\alpha < 1$, $\beta < 1$, and $\alpha > \beta$ represent the higher labor productivity of intensive agricultural production. Equivalent assumptions apply to non-agricultural labor, $q_N(l_N) = {l_N}^{\alpha}$, which captures any other wage-earning activity performed by the household either in the village or outside (based on commuting or temporary migration). All outputs except firewood are sold at market prices p_{A1}, p_{A2}, and p_N. Prices are a function of transport cost only and are thus exogenous at the household level. The modelled effect of road construction is to increase or decrease prices, depending on whether the resulting market linkages provide primarily export opportunities or import opportunities relative to the surrounding region. A household's cash income may thus be represented as:

$$Y = {l_{A1}}^{\alpha} p_{A1} + {l_{A2}}^{\beta} p_{A2} + {l_N}^{\alpha} p_N \tag{3}$$

Household income is spent on the consumption of market goods, q_X, and non-forest energy, q_E:

$$Y = q_E p_E + q_X p_X \tag{4}$$

Households' utility is a Cobb-Douglas function of market goods consumption, q_X, and energy consumption (of either market or forest origin), such that $U(q_X, E) = {q_X}^{\delta} E^{\gamma}$. The household chooses

---

[26] This assumes labor markets are incomplete and households do not hire labor.

the utility-maximizing combination of variables $q_S, q_X, q_E, l_{U,} l_F, l_{A1,} l_{A2}$ and l_N, subject to constraints (1) through (3).

$Max_{q_X, q_E, l_F, l_{A1}\, l_{A2}, l_N} : {q_X}^{\delta} E^{\gamma}$

Subject to:

$L = l_{A1} + l_{A2} + l_F + l_N$

$Y = {l_{A1}}^{\alpha} p_{A1} + {l_{A2}}^{\beta} p_{A2} + {l_N}^{\alpha} p_N$

$Y = q_E p_E + q_X p_X$

$$E = {l_F}^{\beta} + q_E \qquad (5)$$

Substituting constraint (2) into (3), and attaching Lagrangian multipliers gives:

$$\mathcal{L} = {q_X}^{\delta} E^{\gamma} + \lambda_1 \left[\left({l_{A1}}^{\alpha} + {l_{A2}}^{\beta} \right) p_A + {l_N}^{\alpha} p_N - q_E p_E - q_X p_X \right] + \lambda_2 \left[{l_F}^{\beta} + q_E - \mathrm{E} \right] + \lambda_3 \left[l_{A1} + l_{A2} + l_F + l_N - L \right] \qquad (6)$$

Which gives first order conditions at the optimum:

$$\frac{\partial \mathcal{L}}{\partial l_{A1}} = \lambda_3 + \lambda_2 \alpha {l_{A1}}^{\alpha-1} p_A = 0 \text{ (if } l_{A1} > 0) \qquad (7)$$

$$\frac{\partial \mathcal{L}}{\partial l_{A2}} = \lambda_3 + \lambda_2 \beta {l_{A2}}^{\beta-1} p_A = 0 \text{ (if } l_{A2} > 0) \qquad (8)$$

$$\frac{\partial \mathcal{L}}{\partial l_N} = \lambda_3 + \lambda_1 \beta \alpha {l_N}^{\alpha-1} p_N = 0 \text{ (if } l_N > 0) \qquad (9)$$

$$\frac{\partial \mathcal{L}}{\partial l_F} = \lambda_3 + \lambda_2 \beta {l_F}^{\beta-1} = 0 \text{ (if } l_F > 0) \qquad (10)$$

$$\frac{\partial \mathcal{L}}{\partial E} = \gamma {q_X}^{\delta} E^{\gamma-1} - \lambda_2 = 0 \text{ (if } E > 0) \qquad (11)$$

$$\frac{\partial \mathcal{L}}{\partial q_X} = \delta {q_X}^{\delta-1} E^{\gamma} - \lambda_1 p_X = 0 \text{ (if } q_X > 0) \qquad (12)$$

$$\frac{\partial \mathcal{L}}{\partial q_E} = -\lambda_1 p_E + \lambda_2 = 0 \text{ (if } q_E > 0) \qquad (13)$$

Combining equations (7), (8) and (9) gives the marginal rates of transformation between agricultural and non-agricultural activities:

$$\frac{{l_{A2}}^{\beta-1}}{{l_{A1}}^{\alpha-1}} = \frac{\alpha p_{A1}}{\beta p_{A2}} \qquad (14)$$

$$\frac{{l_{A1}}^{\alpha-1}}{{l_N}^{\alpha-1}} = \frac{p_N}{p_{A1}} \qquad (15)$$

$$\frac{{l_N}^{\alpha-1}}{{l_{A2}}^{\beta-1}} = \frac{\beta p_{A2}}{\alpha p_N} \tag{16}$$

Combining these expressions with (10) gives expressions relating forest labor to other forms of labor:

$$\frac{\lambda_1}{\lambda_2} = \frac{1}{p_N}\frac{\beta}{\alpha}\frac{{l_F}^{\beta-1}}{{l_N}^{\alpha-1}} = \frac{1}{p_{A2}}\frac{{l_F}^{\beta-1}}{{l_{A2}}^{\beta-1}} = \frac{1}{p_{A1}}\frac{\beta}{\alpha}\frac{{l_F}^{\beta-1}}{{l_{A1}}^{\alpha-1}} \tag{17}$$

Rearranging (11) and (12) shows that these expressions, at the optimum, are equivalent to the marginal rate of substitution between consumed goods, and also to the inverse of the price of energy (from 13):

$$\frac{\lambda_1}{\lambda_2} = \frac{1}{p_X}\frac{\delta}{\gamma}\frac{E}{q_X} = \frac{1}{p_E} \tag{18}$$

These expressions, (17) and (18), allow me to construct simplified relationships between forms of labor:

$$l_F = l_{A2}\left(\frac{p_{A2}}{p_E}\right)^{\frac{1}{\beta-1}} \tag{19}$$

$$l_N = \left(\frac{\beta}{\alpha}\frac{p_{A2}}{p_N}\right)^{\frac{1}{\alpha-1}} {l_{A2}}^{\frac{\beta-1}{\alpha-1}} \tag{20}$$

$$l_{A1} = \left(\frac{\beta}{\alpha}\frac{p_{A2}}{p_{A1}}\right)^{\frac{1}{\alpha-1}} {l_{A2}}^{\frac{\beta-1}{\alpha-1}} \tag{21}$$

Substituted into the household time constraint, $L = l_{A1} + l_{A2} + l_N + l_F$ (1), these give an implicit function relating extensive agricultural labor and prices:

$$l_{A2} + l_{A2}\left(\frac{p_{A2}}{p_E}\right)^{\frac{1}{\beta-1}} + \left(\frac{\beta}{\alpha}\frac{p_{A2}}{p_N}\right)^{\frac{1}{\alpha-1}} * {l_{A2}}^{\frac{\beta-1}{\alpha-1}} + \left(\frac{\beta}{\alpha}\frac{p_{A2}}{p_{A1}}\right)^{\frac{1}{\alpha-1}} * {l_{A2}}^{\frac{\beta-1}{\alpha-1}} = L \tag{22}$$

And after rearranging (A19), an implicit function relating forest labor and prices:

$$l_F + l_F * \left(\frac{p_E}{p_{A2}}\right)^{\frac{1}{\beta-1}} + {l_F}^{\frac{\beta-1}{\alpha-1}} * \left(\frac{\beta}{\alpha}\frac{p_E}{p_N}\right)^{\frac{1}{\alpha-1}} + {l_F}^{\frac{\beta-1}{\alpha-1}} * \left(\frac{\beta}{\alpha}\frac{p_E}{p_{A1}}\right)^{\frac{1}{\alpha-1}} = L \tag{23}$$

I assume that new roads decrease the market price of firewood substitutes, p_E, and increase the sales-price for the products of non-agricultural labor, p_N. I consider both increases and decreases in the sales-prices of agricultural products (p_{A1} and p_{A2}) to be possible due to new roads. Note that change in prices p_N or p_{A1} could equally represent a reduction in input prices or an increase in the sales-price received.

The key outcome of interest is the impact of household decisions on forest cover. I assume that damage to forests occurs both due to the harvesting of firewood, l_F, and labor spent in agriculture (most likely grazing) near or within forests, l_{A2}. Price changes that encourage both activities are more likely to lead to observed forest degradation and forest degradation. Price changes that encourage one forest damaging activity at the expense of the other will have an ambiguous impact on forests, while those that discourage both activities are more likely to lead to observed forest regrowth, or a slowing of degradation and deforestation trends.

From (22), partial derivatives of extensive agricultural labor have signs: $\frac{\partial l_{A2}}{\partial p_N} < 0, \frac{\partial l_{A2}}{\partial p_{A1}} < 0, \frac{\partial l_{A2}}{\partial p_{A2}} > 0$, and $\frac{\partial l_{A2}}{\partial p_E} < 0$. As expected, higher returns to extensive agriculture cause households to devote more time to extensive agricultural labor. Higher returns in non-agricultural labor, l_N, causes a monotonic shift away from both types of agriculture and away from forest use, again as expected. This represents an inter-sectoral movement from agriculture to small village enterprises such as shops, mills, and other local or commuter-based urban employment. A similar effect is seen from an increase in returns to intensive agriculture. An increase in the market price of energy reduces extensive agriculture as labor is diverted to relatively lower-cost forest firewood collection. From (23), partial derivatives of firewood collection have signs: $\frac{\partial l_F}{\partial p_N} < 0, \frac{\partial l_F}{\partial p_{A1}} < 0$, $\frac{\partial l_F}{\partial p_{A2}} < 0$, and $\frac{\partial l_F}{\partial p_E} > 0$. Higher returns to intensive agriculture reduce firewood collection by reducing time available and providing income for the purchased substitute. Higher energy prices have the opposite effect, as expected due to the relatively lower implicit cost of collecting firewood. In aggregate, these results show:

- Increased returns to intensive agriculture, and increased returns to non-agricultural activities, all else equal, have an unambiguously positive predicted effect on forest cover (i.e. both l_F and l_{A2} are reduced). This is despite an income effect that encourages greater energy consumption overall.

- Increased returns to extensive agriculture (grazing) has an ambiguous predicted effect on forest cover: firewood collection decreases, but extensive agriculture itself increases. The precise net effect is a function of the relative impact (and starting magnitude) of both activities, variables that are highly dependent on local settings.

- Equal, simultaneous, increases in returns in both agricultural sector (p_{A1} and p_{A2}) do not lead to qualitative differences in these results. The labor allocation to both agricultural sectors increase, with disproportionate growth in the more productive sector. The labor allocation to firewood collection decreases. Similarly, the absence of a non-agricultural sector entirely (perhaps representing particularly remote regions) does not change signs. All comparative statics are robust to changes in the utility parameters δ and γ, and robust to changes in the productivity parameters β and α (for β, α < 1).