Journal of Development Economics 133 (2018) 375–395

Contents lists available at ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Do rural roads create pathways out of poverty? Evidence from India☆

Shilpa Aggarwal

Indian School of Business, Hyderabad, India

ARTICLE INFO

JEL classification:
O18
J24
Q16
R42
F14

Keywords:
Roads
Market integration
Technology adoption
School enrollment
Consumption variety

ABSTRACT

Nearly a third of the world's rural population does not live near a paved road, and it is widely believed that this limits their access to economic opportunities. Using a natural experiment that led to plausibly exogenous variation in the timing and placement of paved roads in Indian villages, this paper provides evidence on the impact of roads on a wide variety of economic outcomes in rural areas. I find evidence consistent with 5 main effects in the village economy. Households in treatment areas report (a) lower prices and (b) increased availability of non-local goods, suggesting greater market integration. Reduced-form evidence suggests that changes in market access caused rural households to (c) increase the use of agricultural technologies, and (d) pull teenaged members out of school to join the labor force. Finally, evidence points to (e) enrollment gains for younger children.

1. Introduction

Poor transportation infrastructure limits access to markets and public services for many residents of developing countries (World Bank, 2007, 2009). In recent years, governments and multilateral organizations, such as the World Bank, have attempted to address this problem by making large investments towards the provision of roads and railroads.[1] However, the impact of these investments is not well-understood as infrastructure placement is usually driven by endogenous economic, political, or social factors. This precludes drawing rigorous conclusions about the primary relationship between transport infrastructure and spatial-integration, as well as its subsequent bearing upon economic and social welfare.

In order to isolate the impact of improved roads on local economic outcomes, this study exploits a rule-based public program from India that led to plausibly exogenous provision of paved roads connecting villages to nearby towns. The program in question - the Prime Minister's rural road scheme (hereafter, PMGSY) - created a federal mandate to bring all villages with a population of at least 500 within reach of the nearest market via an all-weather road. Between the years 2001 and 2010, PMGSY provided paved roads to more than 110 million people,

☆ I am immensely grateful to Jonathan Robinson for an abundance of guidance and support. This paper has benefited substantively from feedback provided by Carlos Dobkin, Jennifer Poole, Nirvikar Singh, and Alan Spearot. I thank Manuel Barron, Pia Basurto, George Bulman, Arun Chandrasekhar, Ritam Chaurey, Jesse Cunha, Sisir Debnath, Rajeev Dehejia, Sarang Deo, Pascaline Dupas, Rob Fairlie, Fred Finan, Johanna Francis, Maitreesh Ghatak, Brian Giera, Tarun Jain, John Leahy, Justin Marion, Paul Novosad, Vishal Singhal, two anonymous referees, and seminar participants at Georgia State University, Mathematica Policy Research, NEUDC 2013, PacDev 2014, SERI Conference 2015, Stanford University, UC Berkeley, UNC Chapel Hill, University of Adelaide, University of San Francisco, Vanderbilt University, and multiple forums at the Indian School of Business and UC Santa Cruz for helpful suggestions. Financial support from the UC Santa Cruz Economics Department is gratefully acknowledged. I thank Pankaj and Sanjay Gupta for access to the agricultural census. Sisir Debnath generously shared village-level mapping for the 2001 and 2011 censuses of India. This project would not have been possible but for the tireless data assistance provided by Bijoy Bhukania and Neha Singhal. All errors are my own. This paper was a part of my doctoral dissertation at UC Santa Cruz.

E-mail address: shilpa_aggarwal@isb.edu.

[1] For instance, the World Bank has spent more than $20 billion on transportation infrastructure projects annually since 2006 (Private Participation in Infrastructure Projects Database, The World Bank).

[2] These numbers are representative of the road connectivity status of rural populations globally. According to the World Bank's Rural Access Index, over 1 billion rural inhabitants (or 31 percent of the world's rural population) live more than 2 km (or 25+ minutes of walking distance) away from the nearest all-weather road. 98 percent of these individuals live in developing countries. See http://www.worldbank.org/transport/transportresults/headline/rural-access.html.

https://doi.org/10.1016/j.jdeveco.2018.01.004
Received 19 February 2017; Received in revised form 30 December 2017; Accepted 12 January 2018
Available online 22 March 2018
0304-3878/

about 14.5 percent of the entire rural population, or 47 percent of the unconnected rural population[2] of India as of the 2001 census.[3] I exploit program roll-out across more than 500 districts over a 10-year period to pin down the causal impact of paved road connectivity, where the counterfactual is an unpaved road.[4] While the ideal identification strategy here would be to consider villages on either side of the population cutoff stipulated by the program, a lack of suitably disaggregated outcome data precludes that. Instead, identification is based on each district's annual exposure to new roads, which is a function of the size distribution of unconnected villages in the district, and is arguably exogenous to economic outcomes.[5]

I analyze the impact of the program on rural households along 5 main outcomes: (i) prices of goods imported from outside the village, (ii) variety in the household consumption basket, (iii) technology adoption in agriculture, (iv) human capital investments in children and adolescents, and (v) occupation choices of adolescents as well as adults. My first main result is that in districts with greater road construction, there was a reduction in prices paid by rural households for goods produced in urban areas. Second, I find that roads led to an increase in variety in the household consumption basket, particularly of goods not produced locally. Both of these responses point to greater market integration as they are likely to have stemmed from more intensive trade between urban and rural areas. Next, I find that in districts which received more paved roads, the use of fertilizer and hybrid seeds increased, and younger kids were more likely to be enrolled in school. Finally, I find evidence that in response to the construction of paved roads, teenagers dropped out of school and started working as labor market opportunities expanded. Similarly, the labor force participation rate of prime-aged women also went up. Many of the occupations adopted by these entrants to the labor force were those where road quality would have a direct bearing on profitability, such as selling perishable goods. While all the findings in this paper is reduced-form in nature, I posit that these changes came about through a decline in transportation costs, and the resultant change in relative prices.

The foremost contribution of this study is in being one of the first in the literature to causally estimate the impact of roads in rural areas. In doing so, this paper adds to an existing body of work on the impact of rural roads, which has largely been reliant on non-random provision of roads (see Dercon et al., 2009; Gibson and Olivia, 2010; Gibson and Rozelle, 2003; Jacoby, 2000; Jacoby and Minten, 2009; Khandker et al., 2009; Khandker and Koolwal, 2011).[6] Outside of this work, much of the existing research on transport infrastructure has focused on railroads and urban highways, and our understanding of the effects of rural roads remains limited. Moreover, these papers study outcomes that are very different than the ones included in the study at hand. For instance, in recent years, a number of papers have looked at the impacts of a large highway upgrade program, also from India ("the golden quadrilateral"), but the focus of these has either been on the urban manufacturing sector, or they are in the macro-development style, with an emphasis on economy-wide income and efficiency gains (see Asturias et al., 2017; Alder, 2017; Datta, 2012; and Ghani et al., 2016). The difference in focus between these papers and mine is a natural consequence of the fact that differences in the placement and reach of transportation infrastructure are likely to generate different qualitative and quantitative impacts. Therefore, even though there has been a proliferation of papers studying the impact of transportation infrastructure recently, we know very little about the contribution of paved roads to the development process of remote areas.

This paper also contributes by providing clean causal impacts of roads. In general, this is an under-researched empirical question as political and bureaucratic capture of public goods is pervasive,[7] and therefore, causal identification is a challenge. Randomized provision of roads is also hard for researchers to implement at scale due to their investment-intensive nature.[8] Given these difficulties, the existing literature on transport infrastructure effects has largely relied on quasi-random variation underpinned by one of two potential sources. Several authors, including Duranton and Turner (2012) and Volpe Martincus et al. (2017), have instrumented for current infrastructure using ancient routes that are no longer in use. In the alternative strategy, Atack et al. (2010), Datta (2012), Faber (2014) Ghani et al. (2016), Jedwab and Moradi (2012), and several others, have utilized variations in the straight line distance between peripheral regions and the (rail)road as a source of quasi-random variation in access. The exogenous rules of the PMGSY program provide me with a new source of quasi-random variation in road access. Moreover, this strategy allows me to measure the impacts of improved accessibility for remote areas, something which the existing empirical studies are unable to capture as these areas are simply too far to be included in the straight line distance method, or to be in the catchment area of ancient routes or current infrastructure.

While the existing literature has largely been concerned with roads only in how they ease the movement of goods, their impact on mobility is potentially all-encompassing, and can facilitate more than just trade. For instance, by making it easier to access labor markets and government services, they can impact a range of economic variables, including (but not limited to) human capital, occupations, and income. While this pervasive nature of potential impacts stemming from road-construction makes roads a compelling development intervention, it also makes the expected impacts hard to model in a tractable fashion. As a result, this paper takes a more program evaluation-type approach to analyzing the impacts of PMGSY, rather than test the predictions from an economic model. The primary channel through which we expect roads to affect economic outcomes is via a reduction in transport costs. As a result, the price of traded goods, inclusive of transportation costs, should go down at import destinations. A recent strand of the trade literature has tried to empirically test this assertion by looking for patterns of price convergence between the origin and destination of traded goods. In his seminal paper on railroad construction in colonial India, Donaldson (forthcoming) finds large reductions in price differences between regions connected by the railroad. Keller and Shiue (2008), in a study set in 19th century Germany, show that the adoption of steam trains led to a decline in grain price-dispersion across 68 markets. In a similar vein, I find a decline in the prices paid by rural households for goods produced in the city. This is an important first step in establishing the efficacy of the program, as paved roads today could have a significantly different impact compared to railroads in the 19th century, when alternative modes of transportation were unavailable and communication infrastructure to enable inter-location information flows was relatively primitive.[9]

[3] The program is still underway as of this writing.

[4] See https://tnrd.gov.in/pmgsy_gallery.html for some before and after pictures of PMGSY roads from the southern Indian state of Tamil Nadu.

[5] I am also able to establish this empirically by showing an absence of pre-trends across districts.

[6] There is also an emerging body of work specifically on the impacts of PMGSY (Adukia et al., 2017; Asher and Novosad, 2016; Shamdasani, 2016), which I discuss below.

[7] This is well-documented in the political economy literature. For instance, Nguyen et al. (2012) and Burgess et al. (2015) provide evidence of mistargeted construction projects in Vietnam and Kenya respectively, on account of nepotism and ethnic favoritism. Rasul and Rogger (2018) highlight the relationship between bureaucratic autonomy and incentives, and the quality and completion rates of public projects in the context of the Nigerian civil service. Khemani (2004) and Rogger (2013) find evidence from India and Nigeria, showing that public goods provision improves when there is a higher degree of political competition.

[8] For instance, in conversations with the author, officials from the Indian ministry of Rural Development have suggested that roads constructed under PMGSY cost between $70,000 and $120,000 per kilometer per lane to build.

[9] A large literature, including Aker (2010); Goyal (2010); Jensen (2007); and Steinwender (2018), has highlighted the role of information frictions in bringing about spatial disparity in prices.

While price-convergence across markets is greatly symptomatic of market integration, it manifests itself only for traded goods. However, the composition of traded goods is endogenous to the presence of a road, and therefore, an alternative test of integration with outside markets is simply the number of goods available locally. I measure this by analyzing the number of varieties of various goods that households report consuming. While there is no study that directly explores the relationship between transportation infrastructure and consumption variety, there is a large literature establishing the variety gains from trade (Feenstra, 1994; Broda and Weinstein, 2010). While not related to variety directly, in a recent paper, Atkin and Donaldson (2015) show that in going from 50 km away from where a good was manufactured to 500 km away, there is an 11 to 22 percent decrease in the likelihood that the good will be available for sale in the local markets. This result underscores the strong influence of remoteness on the availability of goods. In the current context, I find heterogeneous variety impacts by type of good: newly connected households decrease the number of staple foods, and increase the number of non-staple, perishable foods and other non-locally produced goods in their consumption basket. In a framework with CES utility, increase in variety directly enters the utility function in the form of new goods, and is welfare-augmenting by itself. The observed switch from staples (cereals, lentils) to non-staples (dairy, meat, produce) likely has positive welfare implications even in the absence of assumptions on the exact form of the utility function. Specifically, staple foods, while calorie and macro-nutrient rich, are inferior to non-staples in terms of providing micro-nutrients. As such, interventions that cause households to substitute towards micro-nutrient rich foods are important from a policy perspective as the nutrition literature has shown that micronutrient-malnutrition is an important reason behind low productivity in developing countries (Tontisirin et al., 2002; Kennedy et al., 2007). To my knowledge, this is one of the first papers to estimate variety gains from infrastructure provision, the first to show that there may be heterogeneity by good-type in how households adjust their consumption when they move out of relative autarky, and one of the first to use survey data on household consumption to measure variety gains.[10]

There is reason to expect that the expansion in the availability and affordability of goods should not be limited to just consumers, but also be reflected in production and investment choices. Accordingly, I analyze technology adoption decisions, and find that farmers with access to new roads are more likely to increase the use of chemical fertilizer and hybrid seeds on their farms.[11] This can be viewed as a direct test of Suri (2011), who proposes that farmers with high gross returns to inputs such as hybrid seeds may still choose not to adopt them if there are high costs to acquiring these due to poor infrastructure.[12] The hypothesis proposed by Suri is validated by Aggarwal et al. (2017) through a rigorous data-collection exercise in Northern Tanzania showing that transportation costs are a large component of the final price of fertilizer, and consequently, serve to dampen fertilizer adoption by rural farmers. My findings corroborate the adoption effect further by showing that usage goes up when transportation costs decline due to better roads.

It bears mentioning here that there are other potential mechanisms behind why adoption might go up, such as an easing of credit constraints. Roads might alleviate credit constraints by increasing output prices (Khandker et al., 2009), or by increasing the collateral value of land (Gonzalez-Navarro and Quintana-Domeque, 2016; Donaldson and Hornbeck, 2016). Although data limitations preclude me from isolating the exact channels at play, the findings in this paper confirm the association between rural road construction and technology adoption in agriculture.

In terms of human capital, I find that the effect is positive on the school enrollment of 5–14 year olds, and negative on that of 14–20 year-olds. The positive effects could potentially stem from better access to schools. Indeed, there is a rich literature in development that finds large positive effects of school construction on children's school enrollment and attendance (Duflo, 2001; Aaronson and Mazumder, 2013; Burde and Linden, 2013; Kazianga et al., 2013). To the extent that the operative channel in these studies is greater proximity to the school, constructing a road might have similar positive effects by reducing the effective distance (in terms of travel time) and the cost of traveling to school. For instance, Muralidharan and Prakash (2017) analyze precisely the effect of reducing the effective distance to school without constructing any new schools. In order to do so, they use a public program from the Indian state of Bihar that provided bicycles to girls continuing to secondary school, and find a 30 percent gain in enrollment. Alternatively, access gains brought about by roads may improve the labor market returns to education, leading to gains in enrollment (see Adukia et al., 2017 for evidence on middle school children, i.e., 11 to 13 year-olds in the context of the same program).

On the other hand, improved access may open up greater labor market opportunities for children, not just in the future, but also in the present, raising the opportunity cost of schooling, and potentially causing them to drop out. Atkin (2016) provides evidence that new factory openings in Mexico led children to drop out from high school sooner and start working.[13] Similarly, Nelson (2011) finds that improving self-employed households' access to credit causes their kids to drop out of school and start working in the family enterprise. Schady (2004), Kruger (2007), and Shah and Steinberg (2017) find similar effects for even transient labor market shocks (recessions, commodity price busts, and droughts respectively), showing that kids are likelier to be in school when jobs are scarce, and likelier to be working when jobs are abundant. My finding of increased drop-outs among adolescents is in the latter vein.

A discussion of the findings of this paper cannot be complete without placing it in the context of the emerging literature on the impacts of the PMGSY program. As discussed above, the rules underlying the roll-out of this program provide a unique setting to isolate the causal effects of rural road provision, and consequently, since this paper was first written, many others have also studied the impact of PMGSY. Notable among these evaluations are Adukia et al. (2017) on human capital investments, Asher and Novosad (2016) on occupation choice, and Shamdasani (2016) on agricultural investments.

Adukia et al. (2017) show that middle school enrollment (i.e., school enrollment of children of ages 11, 12, and 13) goes up in response to the program, with the gains being larger where nearby markets provide greater returns to education and smaller where the opportunity costs of schooling are higher. This result is a corroboration of my results in two ways. One, it is consistent with enrollment gains for children under the age of 14. More importantly, the heterogeneity in gains by proximity to different kinds of labor markets mirrors the heterogeneity by age that I find. Specifically, opportunity costs of schooling are presumably

[10] Much of the existing trade literature uses countries' import composition to measure variety gains. See, for instance, Arkolakis et al. (2008). Broda and Weinstein (2010) and Handbury and Weinstein (2011) use supermarket scanner data, which provides an alternative measure of household consumption but does not allow the researcher to control for household characteristics. Hillberry and Hummels (2008) analyze this from the firms' perspective and show that trade frictions reduce aggregate trade volumes primarily by reducing the number of goods shipped and the number of establishments shipping.

[11] This increase could be coming about on either the intensive or the extensive margin for a farmer, as the two cannot be disentangled in my data since I only observe the area cultivated using these inputs.

[12] Conventional seeds and manure are usually available locally, while fertilizer and hybrid seeds need to be explicitly procured from agro-input supply shops, which are typically located in larger market centers.

[13] This evidence is far from conclusive as Jensen (2012) finds an increase in school enrollment of young girls in Northern India as job opportunities provided by English language and Information Technology skills expanded.

larger for older adolescents relative to children, and therefore, it stands to reason that they make the smallest gains in school enrollment. Therefore, both the papers tell a story of the impact of rural roads on human capital accumulation that is mediated through the opportunity cost of schooling.

Asher and Novosad (2016) also evaluate PMGSY through the lens of access to urban labor markets, and find evidence of a sectoral reallocation away from agriculture. This is in line with my findings on occupational choice, especially among prime-aged men who move out of agriculture and start working in retail. Note that while it is access to urban markets that is driving the impacts pertaining to both education as well as occupation choice, the driving factor in the latter case is improved mobility.

Shamdasani (2016) studies the impact of PMGSY on agricultural decision-making, and finds a switch from subsistence to market-oriented farming. Consistent with that, there were increases in the use of expensive, productivity-enhancing inputs, such as fertilizer, hybrid seeds, manure, and hired labor. These findings are consistent with those in my paper. While Shamdasani hypothesizes that these effects come about due to increased inter-village mobility, the uptick in input usage suggests that other mechanisms might also be at play, such as prices in the input and output markets. It is not possible however to isolate these mechanisms with currently available data.

It is worth noting here that the 4 papers under discussion are complementary to each other, and collectively, they significantly further our understanding of the largest rural infrastructure upgrade program in the world. The range of outcomes as well as the underlying mechanisms underscore the pervasive nature of impacts generated by transportation infrastructure. A final point worth mentioning here is that all 3 of these studies (other than mine) are based on disaggregated, village-level data. Therefore, the similarities between these results and mine further speak to the validity of the district-based identification strategy. Moreover, this strategy allows the use of the NSS data, and the expansive nature of the NSS affords me the opportunity to study a large number of interesting outcomes not available elsewhere, such as prices, consumption, and labor market participation.

The remainder of this paper proceeds as follows. The next section describes the institutional context in greater detail. Sections 3 describes the data, and against the backdrop of information provided in Sections 2 and 3, Section 4 analyzes program roll-out and compliance. Section 5 lays out the empirical strategy. Section 6 presents the estimation results. Section 7 presents robustness checks, and considers alternative hypotheses. Section 8 discusses the policy implications of the results, and concludes.

2. Context

2.1. Program background

PMGSY is a federally mandated rural road construction program, announced by the Government of India in late 2000.[14] The goal of the program was to provide all-weather road access to the nearest market center (or to a paved road connecting to the market center) within 500 m[15] of all sub-villages (the program refers to these as "habitations") with a population of at least 500 (250 in the case of tribal areas, or areas pre-defined as desert or mountainous). A habitation is a sub-village level entity, and is defined as "a cluster of population, whose location does not change over time".[16] For the purpose of this study, I use the terms sub-village, habitation, and village interchangeably. The population of each village was determined using the 2001 census. The scheme was federally funded,[17] but implemented by individual states.

The mandate of the program was to provide paved roads to only those villages which did not have a paved road within 500 m of the village at the outset of the program, but presumably, all of these villages could access an unpaved dirt or gravel road, i.e., no village was in complete autarky. Those villages that already had a paved road were not eligible to receive a road under this program.[18]

At the outset of the scheme, states were asked to draw up a core network of roads, which was defined as the bare minimum number of roads required to provide access to all eligible villages. Only those roads that were a part of the core network could be constructed under this scheme. Within the core network, construction was to be prioritized using population categories, wherein, villages with a population of 1000 or more were to be connected first, followed by those with a population of 500–1000, ultimately followed by those with a population of 250–500 (if eligible). The rules further stipulated that in each state, villages from lower population categories could start getting connected once all the villages in the immediately larger category were connected. Exceptions were allowed if a smaller (by population category) village lay on the straight path of a road that was being built to a larger village. In this case, the smaller village would get connected sooner. Therefore, the program presents a potentially suitable setting to examine the causal impact of rural roads as the allocation was based on a pre-specified rule.

2.2. Background on administrative units

Before getting into the details of the various data sources used for analysis, it makes sense to understand the various administrative units above the village, the unit of treatment. This is important because data are often either representative only at higher levels, or are made available at higher levels to prevent identification of individuals and households. The village is the basic administrative unit, and there are about 640,000 villages in the country. The next unit of aggregation is the sub-district, variously known as tehsil, taluk, block, or sub-district in different states of the country. There are just shy of 6000 sub-districts nationwide, i.e., an average sub-district has just over 100 villages. Sub-districts roll up to districts, and a district is the unit at which most government programs are implemented. Virtually all government data is also released at this level in its most granular form. There were 593 districts in the country in 2001, which grew to 640 by 2011, therefore

[14] The program website is http://pmgsy.nic.in/pmgsy.asp.

[15] For mountainous areas, this was defined as 1.5 km of path distance. As per an amendment made to the program rules in 2008, in mountainous regions located next to international borders, this distance could be up to 10 km (Ministry of Rural Development, letter no. P-17023/38/2005-RC dated February 29, 2008).

[16] A village will have multiple habitations if it has 2 or more clearly delineated clusters. For instance, there might be two separate clusters of houses on either side of the village well. India has about 640,000 villages comprising of about 950,000 habitations.

[17] This scheme was funded by earmarking 1 Rupee per liter out of the tax on high speed diesel. The funds were disbursed to the states using a pre-determined formula known as "additional central assistance" (ACA). ACA funds are disbursed solely for the implementation of centrally-sponsored schemes, and the center has complete oversight over their utilization. The formula has the following weights: population - 0.6, per capita income (difference between the state and the national average) - 0.25, fiscal performance - 0.075, "special category" status - 0.075. The "special category" designation is reserved for 11 remote and mountainous states, and includes the 7 northeastern states, Assam, Himachal Pradesh, Jammu and Kashmir, and Uttarakhand. Fiscal performance is a composite score assigned based on "tax efforts" (0.025), fiscal management (0.02), and progress on national objectives with respect to population control (0.01), elimination of illiteracy (0.01), timely completion of externally aided projects (0.005), and implementation of land reforms (0.005). I was unable to obtain any further details regarding the various parameters of fiscal performance from primary or secondary sources, but it is likely that there may be some room for discretion in this component of funding allocation.

[18] Refurbishment of pre-existing paved roads ("upgrades") was allowed, but new construction had much greater priority. Specifically, a state could do upgrades only after it had completed all planned new construction, and could spend at most 20 percent of its total PMGSY funding on upgrades. The focus of this paper is only on new roads.

an average district comprises 10 sub-districts and 1000 villages.[19] Districts aggregate up to states, and the state is the biggest sub-national unit. There were 28 states in the country during the period of analysis (a 29th state was created in 2014).

While I will explain this in greater detail while discussing my empirical strategy, my primary unit of analysis is the district. There were 593 districts in the country at the time the program was launched, but only 562 of these are included in the analysis. The remainder were dropped because they were either entirely urban or because they could not be mapped properly between data-sources due to changes in administrative boundaries over time. I now turn to describing the various data sources.

3. Data

3.1. Online management and monitoring system (OMMS)

The Government of India has recently mandated that the ministry in charge of any large public program make all program data publicly available. As a result, village-level road construction data is available online through OMMS. Thus, for the universe of rural habitations, I have data on their baseline level of road-connectivity, population (in order to determine eligibility), whether they got a road under the program, and if so, the year in which the road was approved and built. In all of my analysis, in order to get around issues of implementation and quality, I use the approval date as the date on which the road was built, and use the words “approved” and “built” interchangeably.

3.2. Population census, 2001 and 2011

I use the directory of village amenities included in the 2001 census of India. I match these villages with those from the OMMS, and this match is accomplished over two stages. The 2001 census of India provides a census code for each village. While the OMMS also has village codes, the quality of this data is sub-par. Specifically, there are a lot of missing values for this variable. For cases where this variable was available, I merged the two datasets using the village code. For the remainder, I did a fuzzy match using village names. After both of these matching exercises, I get an 80 percent match rate between the OMMS and the census. I then use these to study differences in baseline characteristics for connected and unconnected villages at the outset of the program. These are presented in Appendix Table A2. Table A2 highlights the fact that at baseline, an average village with a road was significantly different from an average village without one, along all observable parameters. These statistics underscore the setting in which the inhabitants of the average unconnected village lived, and help us contextualize the findings of this paper. Further, they also highlight the stark distinction between the two types of villages, and therefore, caution us against using the connected villages as a control group.

3.3. National sample survey (NSS) data

The main source of data for this study is the NSS, which is a rich, nation-wide, repeated cross-section survey of individuals and households. The only identifier is the district, however. The surveys contain extremely granular household-level information on the quantity and value consumption of more than 350 distinct items, and individual-level information on education and labor-market participation. Even though the unit of observation is the household in the case of consumption data, and the individual in the case of education and employment data, the smallest identifiable unit provided by the Government of India is the district of residence of said individual or household. The NSS data is not a fixed percent sample of the universe it is drawn from, it is however representative at the district level. It is perhaps the most widely used dataset in empirical studies set in India. In this study too, it is the main source of outcome data.

In order to examine the consumption and human capital outcomes, I use data from the rural schedules of rounds 57 (year 2001) to 66 (year 2010) of NSS. However, since some modules are not fielded every year, this translates to consumption data for years 2001–2008 and 2010, and education and employment data for 2004–2006, 2008, and 2010. Since the smallest identifiable unit is the district, this necessitates that my unit of analysis be the district. I discuss this in greater detail in Section 5.

3.4. Agricultural inputs survey

The Ministry of Agriculture conducts a 5-yearly survey on the usage of advanced inputs in agriculture, including the use of fertilizer, hybrid seeds, and pesticides. For this survey, all operational holdings from a randomly selected 7 percent sample of all villages in a sub-district are interviewed about their input use. These responses are then aggregated by crop and plot-size category (these categories are reported as: below 1 hectare (ha), 1–1.99 ha, 2–3.99 ha, 4–9.99 ha, and above 10 ha), and reported at a district level. Therefore, I have a district-crop-size-year panel of operation holdings in rural India, which I aggregate at the district-crop-year level. I use the 2001–02, and the 2006–07 rounds of the survey for this study. To my knowledge, this is the first instance of the use of this survey in the literature.

4. Program roll-out and compliance

Before we proceed with a causal analysis of the impact of PMGSY, we must ensure that guidelines were followed and that there were minimal deviations from the population rule. Accordingly, I analyze compliance with the rule in Fig. 1, where I show the likelihood of road construction for villages based on their population in the 2001 census. The discontinuous jump in the probability distribution of road construction is apparent: as stipulated by the program, larger villages dominated smaller ones in terms of construction priority.[20] However, the prioritization is not completely clean as smaller villages begin to get roads before the larger ones are fully done. This may be explained by two factors. One, the program did allow for out-of-order connectivity if the location of the villages on the path to the market necessitated this. Second, it is possible that there were some deviations from the rule, simply given the scale of the program.

In order to help us understand the source of these deviations better, Table 1 looks at the determinants of road construction under the program over the period 2001–2011. In this table, I report coefficients from regressions where the dependent variable is the likelihood that a village that was unconnected at baseline (i.e., year 2001) had received a road by endline (2011). The independent variables include the population category of the village (i.e., greater than 1000 or 500–1000. Below 500 is the omitted category), the exact population of the village in the 2001 census, the proportion of scheduled castes (i.e., the Hindu lower castes that tend to be economically disadvantaged) in the village population, distance from the nearest town, and indicator variables for a host of public goods. Columns 1 and 2 include state and district fixed effects respectively. If program rules were followed perfectly, we should see positive and significant coefficients on the two population categories, with bigger categories having greater magnitudes; the coefficients on all other variables should be zero. We can see that by endline, villages with a population of 1000 or more were 41 percentage points more likely to have received a road, while those with population 500–1000 were about 25 percentage points more likely to have

[19] New districts are often carved out from old ones as their population increases in order to ease the administrative burden.

[20] Appendix A1 presents cumulative density functions of road connectivity by population category.

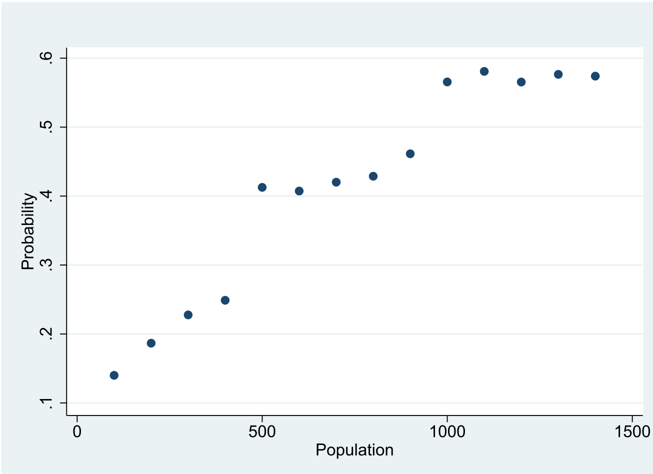


Fig. 1. Road Construction Probability by 2010.

Table 1
Likelihood of paved road construction by 2011.

	1	2	Baseline Mean
500 > p > 1000	0.268***	0.244***	0.22
	(0.033)	(0.030)	
p > 1000	0.410***	0.405***	0.18
	(0.062)	(0.052)	
Population (in '000)	0.007	0.007	0.63
	(0.005)	(0.006)	
SC Population (%)	0.005	−0.001	0.37
	(0.005)	(0.002)	
Distance from Town (in '000 kms)	−0.096	−0.116	0.025
	(0.173)	(0.128)	
Panchayat HQ	0.079***	0.069***	0.08
	(0.019)	(0.015)	
Primary School	0.041**	0.033***	0.78
	(0.016)	(0.011)	
High School	−0.015	−0.011	0.03
	(0.010)	(0.008)	
Adult Literacy Center	0.008	0.002	0.08
	(0.015)	(0.008)	
Primary Health Center	−0.008	−0.007	0.03
	(0.010)	(0.007)	
Commercial Bank	−0.016	−0.014	0.05
	(0.013)	(0.009)	
Post Office	−0.004	0.003	0.24
	(0.007)	(0.004)	
Telephone	−0.006	−0.002	0.26
	(0.007)	(0.005)	
Power Supply	0.001	0.010	0.71
	(0.007)	(0.009)	
Fixed Effects	State	District	
R-squared	0.216	0.288	

Standard errors in parentheses, clustered by state.
***, **, * indicate significance at 1, 5 and 10%.
The likelihood of a village with population less than 500 receiving a paved road by 2011 was 0.13.
Sample of 272,412 villages, all of which were unconnected in 2001 (baseline).

received a road, relative to villages with fewer than 500 inhabitants. Out of the set of 9 public good dummies that I include as controls, two - primary school and *panchayat* headquarters[21] - also turn up statistically significant. While this is unfortunate, it is worth noting that the underlying regressions are extremely high-powered (nearly 300,000 villages), and as such, we should pay attention to not just the p-values but also the magnitude of the coefficients. Note that the coefficients on public goods are completely dwarfed by those on the population categories, in most instances by orders of magnitude. Moreover, even though there is a 7–8 percentage points greater likelihood that a village where the *panchayat* headquarters are located received a road, only 8% of the villages host the headquarters in the first place. Therefore, this is unlikely to bring about an economically meaningful difference (as benchmark, note that 30% of the villages in this sample had received a road by 2011). I try to further alleviate this concern by combining data on village-level amenities from the 2001 and 2011 censuses of India with road construction data from the program to understand if the provision of roads is correlated with the provision of other public goods, such as schools or health centers. Since political influence tends to play an important role in the allocation of all types of public goods, if road provision under the PMGSY was indeed marred by political interference then we should expect to see that PMGSY-beneficiary villages over the 2001–2011 period also receiving other public goods simultaneously. The analysis presented in Appendix Table A3 fails to find such an effect. The treatment villages were no more likely than other villages to have received a school, a health center, a railway station, or a bank branch. There is a 0.2 percent likelihood (significant at the 90% level) that those who received a road also received a post-office, a coefficient I do not believe to be economically meaningful. Finally, there is a 3 percent likelihood that those who received a road also received a bus station, an effect that was likely a consequence of the road.

Finally, while there may have been minor deviations, there are reasons to expect that the government followed the rules. It would have been in the interest of state and district-level politicians to follow the population-based rule of the program as a mechanism to garner votes. For instance, Cole (2009) shows that politicians in India use their influence to get banks to disburse more credit during election years. More generally, even in the absence of "vote buying", the median voter theorem predicts that public goods are allocated in a manner where they benefit the most number of people. As it stands, a far more important corruption concern pertaining to this program would be that the roads were not built at all, and that the funds were appropriated by local politicians and bureaucrats. Two different factors help me mitigate this concern: 1) The government of India was greatly invested in making this scheme transparent to the fullest extent possible.[22] As a result, the program was closely monitored by many different stakeholders and all of the construction details are publicly available,[23] and 2) All of my specifications control for either district or state-level unobservables like corruption. Note however that the results regarding correlation between the provision of various public goods during the intercensal period (Appendix Table A3) provide reassurance that the states did not divert the funds meant for PMGSY towards other alternative uses. Moreover, my analysis is based on roads approved (rather than built), and in case some areas did not get roads as per plan, then the reported estimates represent a lower bound on the causal impact of roads. Finally, since all of my analysis is at the district-level, selection on observables at the village-level should get washed out. Nevertheless, in my data analysis, I deal with this issue by implementing multiple empirical specifications, with and without controlling for observables. My findings stay robust to the inclusion of controls, suggesting that the

[21] "Panchayat" refers to village-level local government in the sub-continent. As part of a decentralization effort, the government of India has devolved some functions, such as maintaining household records, to panchayats. A panchayat typically administers multiple villages in the nearby vicinity, and has its headquarters in one of the villages. The country has about 640,000 villages, of which about 230,000 host the panchayat headquarters.

[22] Indeed, Lehne et al. (2018) document that more than 97 percent of the villages that should have received a paved road under the program based on the data provided by the implementing ministry, were found to have one during the 2011 Census of India.

[23] The program has a three-tier monitoring system at the district, state and federal level. For details, see the program's operation manual, available at http://pmgsy.nic.in/op12.htm.

results are not being driven by selection.

Nevertheless, my empirical analysis consists of a number of robustness checks. I am able to show that there were no pre-trends in outcomes as placebo specifications with roads built during the program period have no predictive power in explaining changes in outcomes over the pre-program period, 1993–1999. I also try to rule out selection into program by controlling for a number of different observable characteristics, and by absorbing unobservables at the district and state level into fixed effects.

5. Empirical strategy

The NSS does not have village-level identifiers, and everything is aggregated to the district. Therefore, I am unable to exploit the program rule of providing roads to villages based on their population category in a regression discontinuity design. Instead, I have to rely on a difference-in-differences strategy to estimate the differences between treatment and control over time. If I had individual-level data on road connectivity status, my estimating equation, for any outcome of interest, *y,* would have been the following:

$$y_{ivdt} = \alpha + \gamma_t + \delta_d + \beta * D_{vdt} + \eta Z_{ivdt} + \varepsilon_{ivdt} \quad (1)$$

where subscript i denotes individuals or households (depending on the outcome of interest), v denotes village, d denotes district, and t denotes survey year. δ is a set of district fixed effects,[24] γ is a set of year fixed effects and Z is a vector of individual/household control variables. D_{idt} is an indicator variable for whether individual i in district d at time t has been exposed to the program, which amounts to an indicator for whether or not a road has been built to his or her village under the program. In this case β captures the average treatment effect of having access to a road for an individual. However, with district-level outcomes, I must aggregate Equation (1) as the following:

$$y_{idt} = \alpha + \gamma_t + \delta_d + \beta * (D_{dt}/N_{dt}) + \eta Z_{idt} + \varepsilon_{idt} \quad (2)$$

where N_{dt} is the population of district d at time t, and D_{dt} is the number of individuals in district d who had received a road under the program by time t. This amounts to using the variations in the percentage of population that received a road in each district in each year. All the other notation in Equation (2) is identical to Equation (1).

It is worth keeping in mind here that the variations in the percentage of population receiving roads in each district are fundamentally a function of variations in the distribution of sizes of unconnected villages in each district. This is because the program rule was applied at the village level, wherein each village's likelihood of receiving a road was an increasing step function of its population, as shown in Fig. 1. When aggregated up to the district, the implication of the rule is that the number of roads built in each district would be some increasing function of the number of villages in each population-size category in that district. It is worth emphasizing that my identification is not based on comparing treated villages with untreated villages, but on comparing districts with varying intensity of treatment.

For some parts of my analysis, I only have access to, or make use of, just 2 rounds of data. In such cases, my estimating equation is given by:

$$y_{idt} = \alpha + \delta_d + \theta * T + \beta * (D_{dt}/N_{dt}) * T + \sigma Z_{idt} + \epsilon_{idt} \quad (3)$$

Here, T is an indicator for the post-treatment period, and θ, the corresponding coefficient. All other notation is identical to Equation (2). Note that all estimates based on this specification only have state fixed effects as there are only 2 rounds of data, and including a district fixed effect would absorb the treatment intensity variable.

Error terms in each specification are clustered at the district level.

In all regressions that are based on the NSS data, household controls include religion, social group (scheduled caste, scheduled tribe, backward caste, or none of these), household type (self-employed or not, agricultural or non-agricultural), size of land owned, and household-size. Where applicable, I also control for an individual's age and gender. As I will discuss in Section 6.4, it turns out that including these controls is not necessary in the sense that their inclusion does not change the magnitude of the treatment effect.[25]

For ease of exposition, I define below the 2 main independent variables that will recur throughout the estimation section.

Roads Built: For any district d at time t, this variable measures that percentage of district d's 2001 population that resides in a village that received a PMGSY road between the years 2001 and t. This is used as the independent variable in nearly all specifications. This is equivalent to the D_{dt}/N_{dt} expression in Equations (2) and (3) above.

Pre-program Roads: For any district d, this variable measures that percentage of district d's 2001 population that resided in a village that already had a road in the year 2001. This variable is used as an independent variable in specifications with state fixed effects. In specifications with district fixed effects, the variable gets absorbed in the fixed effect.

6. Estimation results

6.1. Pre-trends

The fundamental concern with any study in a difference-in-differences setup is that trends might not be parallel, invalidating the results. This concern is especially acute in this case, as districts that had a lot of roads pre-program (and therefore, received fewer roads during the program) might be on a different trajectory relative to the ones that had few roads (and received many more roads during the program). Therefore, before taking the reader through the estimated causal effects of the program, I seek to first establish that the parallel trends assumption holds. In order to do this, I adopt the standard method from the literature, which is to run placebo regressions of roads built during the program on outcomes during a pre-program period. Specifically, I compare outcomes in the year 1999 to those in 1993 as a function of the roads built under the PMGSY program over the period 2001 to 2011.

Of the 5 main outcome variables studied in this paper: prices, consumption variety, technology adoption, school enrollment, and labor force participation, I am able to implement this placebo test for 3 outcomes: consumption variety (Table 2), enrollment, and employment (both in Table 3). In both these tables, the post period is a dummy variable for the year 1999, the baseline year is 1993, and the roads built variable gives the percentage of population that received roads over the entire treatment period up to 2011. If different regions with varying treatment intensities were indeed on parallel trajectories in respect of these outcomes, then we should expect to see the coefficient on the interaction between post and roads built to be zero. Indeed, in virtually all cases (except number of vegetables consumed), the point estimate is statistically indistinguishable from zero. These results bolster our confidence that the results are not picking up spurious effects. Moreover, as we go through the estimation results of the treatment effects, I will continue to marshall further evidence in support of their validity.

[24] All estimating equations were also specified alternately to have state fixed effects, and yield similar results. The results from these specifications, where not presented in the paper, are available on request.

[25] There might be some related concerns about whether this is a sufficient set of controls. While it is hard to establish this, please note that the full set of household-level covariates as well as a district fixed effect are included. There is some likelihood, however, that village-level observables are not adequately controlled. Since my empirical strategy is not suited to doing this, I turn to the literature on some guidance in this regard. Notably, in Muralidharan and Prakash (2017), a paper studying the impact of distribution of free bicycles to on the school enrollment rates of girls in India, Tables 2 and A2 show that the coefficients are identical between specifications that do and do not control for village-level covariates such as public goods (Columns 3 and 4).

Table 2
Placebo test - program roads on 1993–1999 consumption variety.

	Impacts by Item Type											
	Food							Non-Food				
	Cereal	Lentils	Dairy	Meat	Vegetables	Fruit	Processed Food	Contraception	Minor Manufactures	Road Fares	Non-road Fares	Vehicles
Post Dummy	−0.20	0.45***	0.44***	0.50***	0.99***	0.29***	1.08***	0	0.41***	0.13***	0.00	0.01
	(0.24)	(0.13)	(0.07)	(0.08)	(0.24)	(0.08)	(0.07)	(0.00)	(0.04)	(0.03)	(0.01)	(0.01)
Post * Roads Built	1.10	−0.13	−0.25	0.19	3.42***	−0.18	−0.20	−0.01	−0.12	0.13	0.01	0.02
	(0.96)	(0.76)	(0.24)	(0.35)	(0.98)	(0.25)	(0.25)	(0.02)	(0.24)	(0.18)	(0.01)	(0.05)
Observations	1,16,139	1,16,139	1,16,139	1,16,139	1,16,139	1,16,139	1,16,139	1,16,139	1,16,139	1,16,139	1,16,139	1,16,139
R-Squared	0.02	0.06	0.11	0.13	0.11	0.07	0.26	0.00	0.09	0.05	0.01	0.05
Mean of Dep. Var.	2.47	2.12	0.46	0.77	7.00	1.10	0.27	0.00	0.92	0.65	0.03	0.28
SD of Dep. Var.	1.76	1.58	0.63	0.93	3.50	1.20	0.53	0.06	0.77	0.61	0.19	0.46

Notes:
***, **, * indicate significance at 1, 5 and 10%.
All specifications have state fixed effects and household-level controls.
Robust standard errors in parentheses.
The dependent variable is the number of surveyed goods in each category that are consumed by the household.
All specifications are of the type given by Equation 3.

6.2. Prices

Following Donaldson (2013), I argue that if roads indeed led to a reduction in transportation costs, this should be reflected in lower prices of traded goods at import destinations. Consequently, I start by establishing that after getting connected to a town via a road, rural inhabitants paid less for goods manufactured or processed in urban areas than they did before the road was constructed.[26] The data for this analysis comes from the household consumption module of the NSS, which interviews respondent households regarding their consumption of more than 350 distinct items over a 30 day recall period. It should be noted that the survey does not collect information on prices *per se*, but instead asks interviewed households to report their total expenditure on each good they consumed. For a subset of these goods (comprised largely of food items, but also a small number of other items, such as clothing and fuel), the survey collects data on the value as well as the quantity of consumption, enabling me to back out unit values, i.e., value per unit consumed, for these items. By definition, therefore, a unit value can only be calculated when an interviewee reports consuming a positive amount of a particular item. As a result, using unit values to track the behavior of prices poses analytical limitations. Specifically, for any good not consumed in all villages at baseline, we expect treatment households to have a greater extensive margin consumption response as affordability and availability improve (more on this later). This can be problematic as it is reasonable to assume that the average newly-connected village would likely still be paying more than the average already-connected-at-baseline village for imported goods by virtue of being located farther away from the town (see Appendix A2). In pure empirical terms, the unit values corresponding to these households will be missing in the pre-treatment period, but non-missing and greater than the mean in the post-treatment period, due to which a naive inspection of the regression coefficients would suggest that rural prices of urban goods went up after road-building. As a result, I limit this analysis to only those goods that are likely produced in the city (some minimal processing is required) but are consumed nearly universally, even at baseline.

These restrictions leave me with three goods - salt, loose-leaf tea, and matchboxes. The results of this analysis are presented in Table 4, and are based on a regression of the form laid out in Equation (2). The reported coefficients show that for all the 3 selected goods, road construction led to large and significant reductions in the rural prices of processed goods imported from urban areas.[27] It is also notable that the magnitude of the treatment effect is twice as large for salt as it is for tea and matchboxes. This stands to reason: we should expect a road to cause a greater decline in the price of heavier goods since transportation cost is presumably a greater component of the final price of such goods. Moreover, this is also in line with Duranton et al. (2014) who find that a 10 percent increase in highways within a city causes a 5 percent increase in the weight of its exports, and Donaldson (2013) who finds that bilateral trade flows are negatively impacted by the weight per unit value of the good in question, with the treatment effect being 4 percent stronger for every unit increase in the weight per rupee. My results are strikingly similar to Donaldson: at baseline, the average household reported paying about 4 Rupees per kilo of salt and about 130 Rupees per kilo for tea, which amounts to a 32× differential in the weight per unit value ratio. The treatment effect on the price of salt is 122 percent greater than the treatment effect on the price of tea, a 3.8 percent

[26] One might argue that roads connecting villages to cities should also cause urban inhabitants to pay less for goods produced in and imported from rural areas, such as food grains. However, rural roads are much less likely to impact urban prices as urban areas are connected to and served by a large number of villages. In trade/macro parlance, a village is akin to a "small open economy", while the urban area is similar to the "rest of the world".

[27] Note that this is still a lower-bound on the impact of roads as the unit values mask any substitutions towards higher quality products.

Table 3
Placebo test - program roads on 1993–1999 enrollment and employment.

	Enrollment		Employment		
	5–14	14–20	14–20	Adult Men	Adult Women
Post Dummy	0.03***	0.03**	−0.05***	−0.01	−0.02
	(0.01)	(0.01)	(0.01)	(0.00)	(0.02)
Post * Roads Built	−0.04	−0.07	0.06	0	0.08
	(0.05)	(0.06)	(0.07)	(0.02)	(0.07)
Observations	1,45,440	88,325	46,213	74,607	75,373
R-Squared	0.01	0.00	0.01	0.01	0.02
Baseline mean	0.69	0.36	0.39	0.95	0.47

Notes:
***, **, * indicate significance at 1, 5 and 10%.
Robust standard errors in parentheses, clustered at the district level.
All specifications have state fixed effects and household-level controls.
All specifications are of the type given by Equation 3.

Table 4
Impact of Program Intensity on Prices of Universally Consumed Non-local goods.

	Salt	Loose Leaf Tea	Matchbox
Roads Built	−0.21***	−0.09**	−0.11**
	(0.07)	(0.05)	(0.06)
Observations	2,62,663	2,28,490	2,56,543
R-Squared	0.84	0.99	0.99
Mean of Dep. Var.	1.33	−2.03	−0.60
Std Dev of Dep. Var.	0.46	0.34	0.28

Standard errors in parentheses, clustered at the district level.
***, **, * indicate significance at 1, 5 and 10%.
The specification underlying these results is given by Equation 2.
Controls for time and district fixed effects, and individual controls.
Individual controls include household size, religion, social group, and occupation.
Dependent variable is log(price) of each good.
Mean of % Connected by Program: 0.081.

stronger effect per unit weight per rupee.

Finally, the coefficients are not just statistically significant, but also appear economically meaningful in a back-of-the-envelope analysis. Conservatively, if we use the lower bound of 9 percent price reduction for all imported goods (the least of all the three coefficients reported in Table 4), and if we make a conservative estimate that about a third of all consumption in a village is imported,[28] then this translates to a 3 percent reduction in the overall price level. According to the statistics released by the government of India, the monthly per capita expenditure in real terms (at 1987–88 prices) of rural households grew 7.5 percent (from Rupees 179.39 to Rupees 192.93) over the period 2000 to 2010 (Ministry of Statistics and Programme Implementation, 2011), i.e., there was about a 0.75 percent per annum increase in real consumption. This calculation then suggests that the impact of roads through a decrease in prices alone was the equivalent of 4 years of economic growth during this period.

As an aside, it is worth noting here that the standard metric to measure market integration in the trade and infrastructure literature is a reduction in price dispersion. For instance, Donaldson (forthcoming) looks at convergence in the price of salt between origin and destination districts. In his case, the arbitrage argument implies that the price gap is the cost of trading the good between those 2 locations, and therefore, a reduction in the price gap is indicative of a decline in the cost of trading. Things are less clear-cut in the current case as PMGSY did not (directly) connect different villages to each other, but instead, to the nearest market center, which is typically a town. Moreover, the origin and destination of specific goods can also not be identified in my data.

6.3. *Consumption variety*

I now turn to my alternative measure of market integration based on consumption variety. The prediction is that as a village becomes better connected with the rest of the economy, its residents will be able to access goods that are not produced locally. As a result, we should observe that households in newly-connected villages consuming a larger number of goods. For this part of my analysis, I continue using the same NSS data as in Section 6.2. However, I am no longer limited to only those goods for which unit values can be computed (i.e., value and quantity both are available) as I can assign binary consumption indicators based on whether the value consumed of a certain good is zero or non-zero.

My outcome of interest is variety in the consumption basket, which I measure as the number of goods in a particular category (say, fruits or dairy) that are consumed by a household. Note that in this case, consumption of each variable is a binary variable that takes the value 1 for any positive reported amounts, and 0 otherwise, and so is the extensive margin effect.[29] Results are presented in Table 5. The results suggest that among food items, a household that goes from not having a road to having one, consumes 0.4 fewer types of cereals and 0.4 fewer types of lentils. Additionally, there are gains of 0.1, 0.4, and 0.37 in the number consumed of dairy products, fruits, and processed foods respectively. Other food groups also have positive, albeit insignificant coefficients. For non-food items, the estimates are large, positive, and significant.

Many things stand out from this table. One, for food items, we see a marked decrease in the varieties of non-perishable staples (cereal and lentils), and an increase in varieties of perishables and processed food consumed by a household. The increase in processed food varieties is consistent with the transport cost explanation as these foods tend to be produced in urban areas. For locally-produced foods, this outcome is potentially explained by changes in local production patterns, for instance, if access to outside markets allows households to produce goods that are higher up in the value chain (say, milk instead of

[28] This is likely an underestimate: in the year 2001, the split between food and non-food was 50:50, and practically all non-food is imported.

[29] My estimate would be a lower bound on the consumption effect of roads if there are households that completely switch out of consuming a certain good, and substitute it with another, say, if the substituted good is inferior (for instance, a switch from coarse grain to fine grain). The estimated coefficient, in this case, would be 0, since the total number of goods consumed did not change, even though the household potentially moved to a higher indifference curve.

Table 5
Impact of road construction on consumption basket.

	Impacts by Item Type											
	Food							Non-Food				
	Cereals	Lentils	Dairy	Meat	Vegetables	Fruit	Processed Food	Contraceptives	Minor Manufactures	Road Fares	Non-road Fares	Vehicles
Roads Built	−0.36***	−0.35**	0.10**	0.04	−0.22	0.40***	0.37*	0.17***	0.36***	0.29***	0.02	0.18***
	(0.13)	(0.15)	(0.05)	(0.08)	(0.34)	(0.14)	(0.22)	(0.04)	(0.13)	(0.08)	(0.02)	(0.05)
Observations	2,69,572	2,69,572	2,69,572	2,69,572	2,69,572	2,69,572	2,69,572	2,69,572	2,69,572	2,69,572	2,69,572	2,69,572
R-Squared	0.07	0.08	0.04	0.04	0.08	0.07	0.11	0.02	0.05	0.03	0.00	0.05
Mean of Dep. Var.	2.82	2.77	0.89	1.47	10.01	1.66	2.07	0.07	1.30	0.81	0.04	0.36
Std Dev of Dep. Var.	1.33	1.59	0.71	1.29	3.48	1.31	1.55	0.26	0.91	0.64	0.19	0.48
Minimum of Dep. Var.	0	0	0	0	0	0	0	0	0	0	0	0
Median of Dep. Var.	3	3	1	1	10	2	2	0	1	1	0	0
75th Percentile of Dep. Var.	4	4	1	2	12	2	3	0	2	1	0	1
Maximum of Dep. Var.	11	10	5	6	23	10	11	2	9	4	2	2

Notes:
Standard errors in parentheses, clustered at the district level.
***, **, * indicate significance at 1, 5 and 10%.
All specifications have time and district fixed effects, and household-level controls.
Household controls include household size, religion, social group, and occupation.
All regressions based on Equation (2).
Mean of % Connected is 0.081.
The dependent variable is the number of surveyed goods in each category that are consumed by the household.

maize)[30] or if increased competition with other villages allows households to specialize in producing fewer varieties in which they have a comparative advantage.

Two, there is an increase of 0.04 (over a base of 1.47) in the types of "meat" consumed by the average household. Even though the estimated coefficient is insignificant, it should be borne in mind that this has been estimated off a sample with a large number of zeros due to the prevalence of vegetarianism in Indian society.

Three, there are large and significant gains in the many non-food categories of goods consumed by households. For instance, the average household increases its consumption of types of "minor manufactured goods", such as umbrellas and batteries, by 0.36, over a base of 1.3. The growth in types of vehicles owned (increase of 0.18 over baseline mean of 0.36) and the different types of means of surface transport hired (given by the column "road-fares" - increase of 0.29 over 0.81 at baseline) by the household is not only interesting in its own right, it also serve as a robustness check, especially when viewed alongside the absence of effects on non-road means of transportation, comprised of trains, air travel, and travel over water (boat and steamer fares).

Please note that the analysis presented in Table 5 assigns consumption dummies at the household level. An alternative approach would be to assign this dummy at the village level, wherein for any good X, the consumption dummy takes the value 1 if any household in the village reports consuming that good.[31] The village-based strategy has a potential advantage in that the treatment effect is not obscured by individual preferences over goods.[32] The results from a village-level consumption analysis are presented in Appendix Table A4. The qualitative pattern of coefficients is the same as that in Table 5, and a number of point estimates are indeed larger in magnitude. Statistical power suffers however, likely due to a much smaller sample size.

I now turn to the international trade/new economic geography literature in order to benchmark these effects. For most goods considered in the current study, the variety effect is between 10 and 20 percent of the number of varieties consumed by the mean household for that category at baseline. This is comparable to the effects found in the literature as a result of fairly substantial changes in economic conditions. Handbury and Weinstein (2011) find that a doubling of city size (in terms of population) in the US (for example, going from a city the size of Cincinnati to one the size of Atlanta), a marker for large agglomeration and scale economies, is associated with a 20 percent increase in the number of unique varieties available in the supermarkets of that city. Similarly, Broda and Weinstein (2010) use the backdrop of business cycles in the United States (again, economic expansions are strongly correlated with the introduction of new product varieties), and find that over 1999–2003, a period of robust economic growth in the country (average annual growth rate in GDP of about 5 percent per annum), the increase in product variety was about 30 percent. An important difference, however, is the definition of varieties. Whereas these studies are at the barcode level, mine is much more aggregate, making them likely to pick up even those effects that the current study is not designed to pick up. For instance, consider a product, say milk: a barcode-based study will track any differences in milch source, brand, package size, and fat content as different varieties, while the household survey underlying my data just tracks consumption at the broadest category level, "milk". Motivated by this setting, I perform an alternative benchmarking exercise for the rural Indian context by comparing the variety effects of the program to the variety gains brought about by overall growth in the Indian economy over this period. These results are presented in Table A5. The analysis presented in this table is similar to that in Table 5, except that it uses only 2 rounds of data, from 2001 and 2010. Therefore, the post dummy is the time effect, and the interaction between post and roads is the program effect. Notice for all types of non-food items, the coefficient on the interaction between roads and the time dummy is much larger (in some cases, by an order of magnitude) than the coefficient on the time dummy alone. Given that the Indian economy witnessed very rapid growth over this period,[33] these estimates provide remarkable testimony to the effectiveness of infrastructure provision in this regard.

Finally, I turn to Atkin and Donaldson (2015), whose setting is the most comparable to mine, in that they work in a developing country setting and study the impact of remoteness on product availability. They find that CPI enumerators in Ethiopia are about 22 percent less likely to find a good available for sale in a retail location that is 500 miles from where it was manufactured relative to a retail location that is 50 miles from the origin; the corresponding figure for Nigeria is 11 percent.

6.4. *Education & employment*

Having established that road construction did in fact impact market access, I turn to an analysis of human capital accumulation and market participation. I start by looking at the impact of road construction on school enrollment of 5–14 year old children. The results are presented in Panel A of Table 6. In my preferred difference-in-difference specification with district fixed effects (column 4), there is a 5 percentage point increase in enrollment. This finding is of immense importance for public policy. The UN's Millennium Development Goals website notes that as of 2010, primary school enrollment rate stood at 90 percent. These results suggest that rural road construction alone could potentially bridge half of the gap toward achieving universal primary education in India. From an external validity standpoint, it would be useful to isolate the channels through which these gains arise. An obvious channel is improved physical access to the school for children, enabling an easier commute. Moreover, roads might alter the returns to education, increasing the household's incentives to send children to school. Alternatively, this effect could come about due to a number of other channels, such as, increases in family income, relaxing of credit constraints, or greater presence of the teacher in the school due to improved access.[34] However, pinning down the exact mechanisms is not possible with existing data sources.

Panel B presents results from identical analyses for 14-20 year-olds. In this case, the effects are strongly negative, and robust to the inclusion of various covariates and fixed effects. The interpretation is straightforward: going from not having a road to having one, leads to about an 11 percentage point drop in school enrollment, which is an almost 25 percent decline over mean enrollment rates at baseline.[35]

There are a number of important points about Table 6. One, on decomposing by gender, there are no differences in the enrollment gains or losses between girls and boys. This is of great importance in a setting

[30] Shamdasani (2017) also finds evidence of households switching from low-value cereal crops to cash crops in the context of PMGSY. More generally, both Muto and Yamano (2009), and Goyal (2010) find supply responses by farmers to a reduction in search costs due to the introduction of mobile phones. In addition, in Muto and Yamano, this response is limited to perishable foods (bananas), while the non-perishable commodity (maize) stays unaffected.

[31] Note that while villages are not identifiable in the data, it is possible to figure out which households belong to the same village.

[32] On the other hand, there is some likelihood of attenuation due to the presence of individuals in the village who may have procured a good non-locally, such as during a trip to the city.

[33] According to the IMF's World Economic Outlook database, the average annual growth rate of per capita GDP (at constant prices) was 6.3 percent per annum for the period 2001–2010.

[34] This last channel is particularly relevant in the context of India, where the quality of service-provision in the public sector has been found to be extremely poor (Chaudhury et al., 2006).

[35] An alternative interpretation is in terms of network effects: since the program was implemented at the village-level, but these results track changes for the district, it is possible that some of the observed gains and losses from the program arose outside the beneficiary villages. At the district level, the average treatment effect needs to be rescaled by the average treatment size, which in this case is .05. Viewed in this manner, the program led to about a 0.006 percentage point drop in school enrollment for 14-20 year-olds, which translates to a. 01 percent decline over mean.

Table 6
Impact of road construction on school enrollment.

	Overall				Impacts by Gender					
					Girls			Boys		
	1	2	3	4	5	6	7	8	9	10
Panel A: Impact on 5–14 year-olds										
Roads Built	0.04	0.04	0.05*	0.05*	0.03	0.03	0.06*	0.05*	0.05*	0.05*
	(0.03)	(0.03)	(0.03)	(0.03)	(0.031)	(0.031)	(0.031)	(0.027)	(0.027)	(0.029)
Pre-program	0.06**	0.06**			0.05*	0.05*		0.06***	0.06***	
Roads	(0.02)	(0.02)			(0.029)	(0.029)		(0.022)	(0.021)	
Observations	3,22,907	3,22,907	3,22,907	3,22,907	1,51,805	1,51,805	1,51,805	1,71,102	1,71,102	1,71,102
R-Squared	0.039	0.049	0.007	0.017	0.051	0.063	0.021	0.032	0.041	0.013
Baseline mean	0.8	0.8	0.8	0.8	0.77	0.77	0.77	0.83	0.83	0.83
Panel A: Impact on 14–20 year-olds										
Roads Built	−0.09**	−0.08**	−0.12***	−0.11***	−0.09**	−0.09*	−0.09*	−0.08*	−0.08*	−0.11**
	(0.04)	(0.04)	(0.04)	(0.04)	(0.046)	(0.04)	(0.048)	(0.045)	(0.044)	(0.044)
Pre-program	0.00	0.00			0.01	0.00		0	0.00	
Roads	(0.03)	(0.03)			(0.034)	(0.034)		(0.031)	(0.030)	
Observations	2,42,913	2,42,913	2,42,913	2,42,913	1,12,967	1,12,967	1,12,967	1,29,946	1,29,946	1,29,946
R-Squared	0.056	0.09	0.011	0.042	0.077	0.105	0.042	0.051	0.085	0.046
Baseline mean	0.46	0.46	0.46	0.46	0.37	0.37	0.37	0.53	0.53	0.53
Controls	N	Y	N	Y	N	Y	Y	N	Y	Y
District FE	N	N	Y	Y	N	N	Y	N	N	Y
State FE	Y	Y	N	N	Y	Y	N	Y	Y	N

Standard errors in parentheses, clustered at the district level.
***, **, * indicate significance at 1, 5 and 10%.
All specifications have time fixed effects.
Specifications with district FE are of the type given by Equation (2). Those with state FE are also given by Equation (2), but additionally control for pre-program connectivity.
Mean of Roads Built: 0.052.
Where included, household controls are occupation, religion, family size, land size, and social class. Individual controls are age and gender.
The dependent variable is an indicator for whether the respondent reported that the child's primary occupation was going to school.

like India, where investment in girls tends to be disproportionately low due to cultural norms of son preference. My results suggest that even though excludable private resources tend to overwhelmingly be concentrated on male children,[36] the benefits from public goods are potentially enjoyed by both genders equally. Two, in both panels, columns 2 and 4 differ from 1 and 3 in that the former control for household-level observables. Note that the inclusion of these controls does not alter the coefficients. To the extent that household characteristics are correlated with village-level unobservables, this provides additional evidence to rule out selection in road construction. Three, while the first two columns control for fixed effects at the state level, the latter two control for these at the district level. The coefficients on school enrollment remain substantively unaltered across these specifications. Not only does this provide further evidence for the robustness of my estimates, it also enables us to generalize these results to other road construction programs.

While the age-groups of 5–14 and 14–20 were created due to contextual relevance,[37] it may still be informative to analyze the effects of roads on enrollment for each age year separately. Fig. 2 presents the results from this decomposition - the Xs represent the baseline mean of enrollment for each age, and the dots represent the treatment effect. While the biggest changes lie at the tails, the distribution strongly supports the manner in which the ages have been pooled in my regression results.

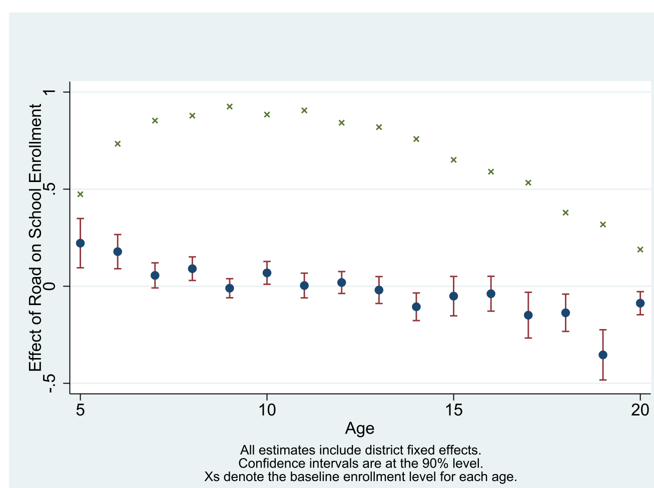


Fig. 2. Effect of Road Construction on School Enrollment by Age.

Table 7 summarizes the next set of results, pertaining to market employment of 14–20 year old children and of adults. Panel A suggests that the school drop-out instance of the 14–20 age group that we witnessed in Table 6, is matched almost one-to-one by increased market employment. As before, these effects do not vary by gender: both girls and boys witness about a 10 percent rise in market employment, which constitutes more than a 40 percent increase over baseline employment levels.[38] Further, this increase in market employment is not limited to

[36] This is also apparent in the great gender disparity in baseline mean enrollment rates, especially for older children.

[37] 14 marks the threshold between primary and secondary education in India. Further, the employment of children below 14 is considered child labor, and is legally punishable.

[38] A breakdown by age, similar to the one for school enrollment, is presented in Fig. 3.

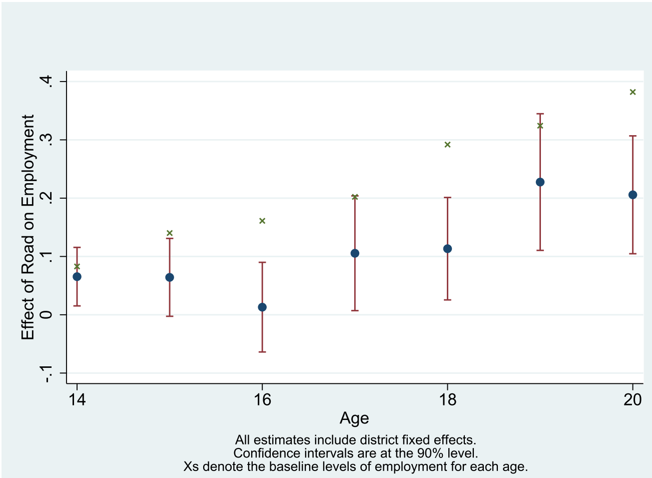


Fig. 3. Effect of road construction on employment by age.

children, as can be evidenced in panel B. On receiving a road, prime-aged women were also 9 percentage points more likely to start working, a 25 percent increase. On the other hand, there is no comparable change for men, which is to be expected, as their employment was nearly universal even at baseline.

I attempt to investigate the mechanisms behind this observed jump in market participation by looking at the occupations that the newly-employed are joining. The results are presented in Table 8. For girls, the most marked increase in employment comes from animal-rearing, followed by textile manufacturing and tailoring. They are less likely than before to be working in forestry, and there is no significant impact on any of the other occupations. For boys, on the other hand, the biggest increase comes from construction,[39] followed by smaller increases in animal-rearing and tailoring. A possible explanation for this is reduced transportation cost as roads might make it possible to transport dairy and meat to the nearest market in a timely fashion. A World Bank report evaluating the performance of PMGSY specifically singles out the growth in dairy farming in program villages, and provides qualitative evidence that roads enable refrigerated milk-collection vans to reach remote villages. This report also talks about increases in house construction activity in rural areas as bulky material like concrete can be transported more easily.[40] The increase in tailoring and making textiles also comes up in the anecdotal evidence provided on the program website as "success stories"[41]: the presence of the road makes it easier for weavers, embroiders, and other similar artisans to sell their crafts in the nearby town. The increases in tailoring may also explain some of the observed increases in school enrollment for younger children. For instance, Heath and Mobarak (2015) show that the advent of garment manufacturing in Bangladesh was associated with enrollment gains for young girl as tailoring jobs require a basic level of numeracy. In looking at occupations for women, I still find the biggest gains in animal rearing. There is also a small increase in textile manufacturing as an occupation. Taken together with the occupational choices of teenaged children, these results suggest that program villages saw the biggest increases in animal-rearing as an occupation, likely due to access to bigger markets. This increase in animal husbandry also constituted a positive supply shock for rural areas themselves, and led to increases in the kinds of dairy and meat products consumed by village inhabitants, which I showed earlier in my analysis of consumption variety in Table 5. Finally, I analyze men's occupation choices, and find that the only significant change came about in the form of large gains in retail as an occupation - prime-aged men are 3 percentage points more likely to work as retailers. This is also in line with the increased market-access hypothesis.

In this context, it is also worth taking note of the positive coefficients on the following outcome variables: the likelihood that a household buys a vehicle (largely bicycles), the number of means of hired road transportation used by the household (both in Table 5), and the likelihood that a treatment village gets a bus stop (Table A3), that have already been discussed in previous sections. When viewed along with the impacts on school enrollment and urban-oriented occupations, the picture that emerges is one of increased local mobility of labor, allowing commuting to nearby urban markets to work.

6.5. Technology adoption

The results thus far provide evidence that road construction led to a reduction in transport costs, and consequently, better access to goods and labor markets. As discussed before, the "reduction in transport costs" channel may also operate in input markets by making it cheaper to either buy the inputs themselves, or by easing credit constraints that hamper technology adoption in agriculture. I test this hypothesis by looking at the area under cultivation using advanced agricultural inputs. Specifically, I look at the adoption of chemical fertilizers and high-yielding seed varieties. Before we analyze the results, it would be useful to understand the underlying data.

The data that I use for this subsection comes from the input survey module of the 2001–02 and 2006–07 rounds of the agricultural census. The data from this survey are reported by the Ministry of Agriculture as district-level aggregates. So, for any district in the country, I have the aggregate acreage, as well as the acreage under modern inputs for all crops grown in that district.

The results are presented in Table 9. From Column 1, for the average crop-district, 22,000 ha of crop area was cultivated using fertilizer at baseline, and would have seen an increase of a little over 10,000 ha in the area under fertilizer use in going from 0 to 100 percent connected. Therefore, the average district, where about 7 percent of the population received new roads, this translates to a 700 ha, or a 3 percent gain in the area under fertilizer per crop. Similarly, for hybrid seeds, there was a 2 percent increase in the area under cultivation per crop. When I break down the analysis by crop type, significant differences emerge: the gains in technology use are entirely concentrated in food crop cultivation, and absent for cash crops. Limiting the analysis to just food crops, and rescaling the coefficients by 0.07 as we just did for the average district, there was an increase of 9 percent in the area under cultivation of food crops using fertilizer, and of 7 percent using hybrid seeds.

A potential explanation for the gains being limited to food crops only might be that cash crops tend to be grown more by bigger (richer) farmers, who are less likely to be constrained by low availability of credit. Alternatively, using the district as the unit of analysis might be masking significant heterogeneity in the pattern of cultivation within the district. Specifically, it is possible that remote regions with low road connectivity do not grow cash crops due to limited market access. In that case, the road construction program is likely to have benefited only those farmers that cultivate food crops.

Boosting fertilizer adoption rates is an important policy goal for many developing countries, and therefore, understanding the mechanisms underlying the gains documented above is important for informing policy. Back of the envelope calculations suggest that if a reduction in the price of fertilizer due to lower transportation cost was the sole mechanism behind increased adoption, then in order to explain the entirety of the observed magnitude of change in demand, the price elasticity of demand would have to be 1.29 (see Appendix A6). A recent

[39] The occupation codes for this category correspond to working as casual labor on private construction sites, and not to working on construction of public works, including roads.

[40] This report can be found at: http://siteresources.worldbank.org/INTSARREGTOPTRANSPORT/1349788-1130967866881/21755701/Rural-Roads-India.pdf.

[41] See http://pmgsy.nic.in/pmgi112.asp#6.

Table 7
Impact of road construction on employment.

Panel A: Impact on 14–20 year-olds	Overall				Impacts by Gender					
					Girls			Boys		
	1	2	3	4	5	6	7	8	9	10
Roads Built	0.10***	0.09***	0.11***	0.10***	0.11***	0.10***	0.09***	0.10**	0.09**	0.12***
	(0.04)	(0.03)	(0.03)	(0.03)	(0.036)	(0.035)	(0.033)	(0.045)	(0.038)	(0.04)
Pre-program Roads	−0.01	0.00			−0.02	−0.02		0.00	−0.02	
	(0.02)	(0.02)			(0.021)	(0.021)		(0.03)	(0.027)	
Indiv Controls	N	Y	N	Y	N	Y	Y	N	Y	Y
District FE	N	N	Y	Y	N	N	Y	N	N	Y
State FE	Y	Y	N	N	Y	Y	N	Y	Y	N
Observations	2,16,366	2,16,366	2,16,366	2,16,366	1,04,066	1,04,066	1,04,066	1,12,300	1,12,300	1,12,300
R-Squared	0.057	0.22	0.006	0.172	0.092	0.16	0.078	0.058	0.244	0.206
Baseline mean	0.24	0.24	0.24	0.24	0.15	0.15	0.15	0.31	0.31	0.31

Panel B: Impact on Prime-Age Individuals	Impacts by Gender							
	Women				Men			
	1	2	3	4	5	6	7	8
Roads Built	0.12**	0.11**	0.09*	0.09	0.01	0.01	0.02	0.01
	(0.054)	(0.054)	(0.06)	(0.057)	(0.01)	(0.009)	(0.01)	(0.011)
Pre-program Roads	−0.12***	−0.12***			−0.02***	−0.02***		
	(0.038)	(0.037)			(0.006)	(0.006)		
Indiv Controls	N	Y	N	Y	N	Y	N	Y
District FE	N	N	Y	Y	N	N	Y	Y
State FE	Y	Y	N	N	Y	Y	N	N
Observations	1,87,735	1,87,735	1,87,735	1,87,735	1,86,528	1,86,528	1,86,528	1,86,528
R-Squared	0.137	0.222	0.004	0.103	0.014	0.049	0.001	0.036
Baseline mean	0.41	0.41	0.41	0.41	0.96	0.96	0.95	0.96

Standard errors in parentheses, clustered at the district level.
***, **, * indicate significance at 1, 5 and 10%.
All specifications have time fixed effects.
Specifications with district FE are of the type given by Equation (2). Those with state FE are also given by Equation (2), but additionally control for pre-program connectivity.
Where included, household controls are occupation, religion, family size, land size, and social class. Individual controls are age and gender.
Mean of Roads Built: 0.052.

Table 8
Impact of road construction on occupation choice.

	Agriculture	Animal Rearing	Forestry	Textile Manufacturing	Tailoring	Manufacturing	Construction	Retail
A: Teenaged Boys								
Roads Built	0.00	0.01*	0.00	0.01	0.01**	0.01	0.05**	0.01
	(0.035)	(0.008)	(0.002)	(0.006)	(0.003)	(0.009)	(0.021)	(0.011)
Observations	1,29,809	1,29,809	1,29,809	1,29,809	1,29,809	1,29,809	1,29,809	1,29,809
R-Squared	0.14	0.01	0.00	0.00	0.00	0.01	0.05	0.02
Mean of Dep. Var.	0.32	0.02	0.00	0.01	0.00	0.02	0.03	0.03
B: Teenaged Girls								
Roads Built	0.01	0.08***	−0.01*	0.01**	0.01***	0.00	0.00	0.00
	(0.034)	(0.020)	(0.004)	(0.007)	(0.004)	(0.006)	(0.006)	(0.004)
Observations	1,12,858	1,12,858	1,12,858	1,12,858	1,12,858	1,12,858	1,12,858	1,12,858
R-Squared	0.09	0.01	0.00	0.01	0.00	0.01	0.01	0.00
Mean of Dep. Var.	0.21	0.04	0.00	0.01	0.01	0.01	0.01	0.00
C: Prime-Aged Women								
Roads Built	−0.03	0.10***	−0.01*	0.01**	0.00	−0.01	0.00	0.00
	(0.048)	(0.031)	(0.003)	(0.007)	(0.004)	(0.012)	(0.019)	(0.007)
Observations	2,18,584	2,18,584	2,18,584	2,18,584	2,18,584	2,18,584	2,18,584	2,18,584
R-Squared	0.07	0.01	0.00	0.00	0.01	0.01	0.01	0.01
Mean of Dep. Var.	0.44	0.08	0.00	0.01	0.01	0.01	0.01	0.01
D: Prime-Aged Men								
Roads Built	−0.02	0.00	0.00	0.01	0.00	−0.01	0.04	0.03**
	(0.032)	(0.006)	(0.003)	(0.005)	(0.005)	(0.015)	(0.041)	(0.016)
Observations	2,16,355	2,16,355	2,16,355	2,16,355	2,16,355	2,16,355	2,16,355	2,16,355
R-Squared	0.07	0.00	0.00	0.00	0.01	0.01	0.03	0.06
Mean of Dep. Var.	0.58	0.01	0.00	0.01	0.01	0.04	0.07	0.06

Standard errors in parentheses, clustered at the district level. ***, **, * indicate significance at 1, 5 and 10%.
All specifications have time and district fixed effects, and household-level controls. Mean of Roads Built: 0.052.
Controls include age, household size, size of land owned, religion, social group, and occupation.
All specifications are of the type given by Equation 2.

Table 9
Impact of road construction on technology adoption in agriculture, 2001–2007.

	All Crops	Cash Crops	Food Crops
Area under Fertilizer			
Post-period Dummy	−43.17	1583.61***	−678.47
	(391.82)	(605.78)	(555.69)
Post * Roads Built	10,266.18***	−2162.47	17,944.83***
	(2524.90)	(2467.75)	(3990.16)
Baseline Mean	22,281.36	7901.07	13,936.63
Baseline Std. Dev	76,771.20	32,036.28	44,764.57
Area under Hybrid Seeds			
Post-period Dummy	692.33	1572.23***	282.12
	(463.56)	(581.45)	(675.55)
Post * Roads Built	6056.85**	−2709.28	13,067.63***
	(2372.22)	(2266.51)	(3740.42)
Baseline Mean	20,187.12	6670.74	12,905.33
Baseline Std. Dev	76,794.03	27,471.54	46,340.40
N	19,087	6666	12,421

Clustered standard errors in parentheses. ***, **, * indicate significance at 1, 5 and 10%.
Includes district fixed effects and district-level covariates (% area irrigated & % smallholders).
Mean of roads built over the analysis period is 0.068.
All specifications are of the type given by Equation 3.

paper by Aggarwal et al. (2018) experimentally varied the price of fertilizer through a discount voucher and estimated the price elasticity of fertilizer demand to be 0.5. An older literature has also estimated this elasticity to be much lower (for instance, elasticity estimates from the US by Binswanger (1974); Ray (1982); Shumway (1983); and McIntosh and Shumway (1994), are well below 1. For the specific case of India, estimates by Subramaniyan and Nirmala (1991) and Dholakia and Majumdar (1995) are similarly low). Using these estimates as a benchmark would suggest that other mechanisms, besides just a reduction in the effective price of fertilizer, are also at play here, such as better output prices, less severe credit and liquidity constraints, or improved access to information and extension services.

7. Robustness and alternative hypotheses

Throughout the discussion of results in this paper, I document a number of findings that support my causal claims. Foremost among these are the results from placebo tests establishing parallel pre-period trends, discussed in Section 6.1. In addition to these tests, I document in Section 6.4 above that the results for human capital outcomes stay similar across a range of different specifications with and without covariates, and with and without fixed effects. This helps me rule out selection on observables in road construction.

Another potential concern is that there was political or bureaucratic capture in the program, calling into question the causality of the observed treatment effects. In order to rule out this possibility, I combine data on village-leve amenities reported by the Census of India in 2001 and 2011, with data on road construction under the program. Specifically, I use as my right hand side variable an indicator for whether or not a village received a road under the PMGSY, and as my left hand side variable, and indicator for whether or not the village received an amenity X between 2001 and 2011. Most of these amenities are of the nature of public goods, although the possibility of the private sector providing such a good cannot be ruled out. My sample is restricted to only those villages which had no road at baseline. These results are presented in Appendix Table A3. It is heartening to note that for most amenities, other than a bus station, there is no correlation between their provision and the provision of a road. There is a small

Table 10
Impact on consumption basket during monsoon.

	Cereal	Lentils	Dairy	Meat	Vegetables	Fruit	Processed Food
Monsoon Dummy	−0.0035***	0.00	−0.0027***	−0.0111***	−0.0050***	−0.0037***	−0.0075***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Roads Built * Monsoon Dummy	0.01	0.01	0.0133*	0.0166*	0.0440***	0.0154*	0.01
	(0.006)	(0.008)	(0.007)	(0.010)	(0.007)	(0.009)	(0.014)
Roads Built	−0.01	−0.0272**	0.0153*	−0.02	0.0213**	0.0292*	0.0801***
	(0.008)	(0.013)	(0.009)	(0.013)	(0.010)	(0.017)	(0.028)
Observations	2,32,772	2,32,772	2,32,772	2,32,772	2,32,772	2,32,772	2,32,772
R-Squared	0.07	0.07	0.04	0.07	0.07	0.10	0.04
Mean of Dep. Var.	0.14	0.22	0.15	0.34	0.10	0.14	0.20

Standard errors in parentheses, clustered at the district level.
***, **, * indicate significance at 1, 5 and 10%.
All specifications have time and district fixed effects, and household-level controls.
Household controls include household size, size of land owned, religion, social group, and occupation.
Estimates are based on a specification that is similar to Equation (2), augmented by the monsoon dummy and its interaction.
Mean of % Connected Post-Program: 0.081.

Table 11
Impact of Road Construction on Employment Location (location dummy = 1 for urban, 0 for rural).

	Overall	Impacts by Group			
		14-20 Boys	14-20 Girls	Prime-Age Men	Prime-Age Women
Post Dummy	0.02***	0.03*	0.00	0.03***	0.02*
	(0.007)	(0.014)	(0.021)	(0.008)	(0.008)
Post * Roads Built	0.13***	0.12	0.36**	0.14***	0.01
	(0.038)	(0.080)	(0.179)	(0.045)	(0.062)
Observations	1,34,860	9787	3106	50,853	13,271
R-Squared	0.01	0.01	0.02	0.01	0.00
Mean of Dep. Var.	0.13	0.16	0.07	0.16	0.06

Standard errors in parentheses, clustered at the district level.
***, **, * indicate significance at 1, 5 and 10%.
All specifications have time and state fixed effects, and household-level controls.
Household controls include household size, size of land owned, religion, social group, and occupation.
All specifications are of the type given by Equation 3.

but statistically significant likelihood of getting a post office (0.2%); however, a post office is unlikely to impact any of the outcomes studied in the paper.

As a final robustness test, I look at consumption effects during the monsoon season. Since the program aimed at providing all-weather roads, its effects were likely to be most keenly felt during the Monsoon when the fair-weather roads to the town are most likely to be flooded or washed out. This is especially true for consumption outcomes, as households are unlikely to make seasonal adjustments to their enrollment or employment decisions. Moreover, any Monsoon-specific effects are unlikely to have come about due to other confounding factors. In order to do this, I combine the information provided by NSS on the date of the survey with consumption information for food, which has a 30 day recall period in the survey. Unfortunately, I am unable to replicate this exercise for non-food items as the survey asks households to report these for a 365-day recall window. Using the Indian Meteorological Department's Monsoon maps as a guide,[42] I create a "monsoon" dummy to indicate whether the household was interviewed during the rainy season, or outside of it. I then interact this dummy with the road construction variable to confirm the robustness of my results, which are presented in Table 10. The specification underlying this table checks for the variety in a household's consumption basket. If the results presented so far are indeed causal, then I should expect to see bigger changes during the monsoon season, and smaller changes outside of it. The pattern of coefficients confirms this hypothesis for perishables and processed food - the categories most likely to have been affected by the roads.

One concern is that what are seemingly program effects might in reality be driven by other factors. One such potential explanation that comes to mind is employment in road construction: if the construction of roads themselves is generating local employment, then the observed outcomes might be short-lived. Further, the results might lose even their short-term generalizability in a setting where construction is managed without tapping the local labor market. I can test this using data on employment location: 2 of the survey rounds (rounds 61 and 66) query all employed individuals regarding the location of their workplace. The responses to this question enable me to ascertain whether an individual's primary place of work is rural or urban. If the mechanism behind the results so far is employment at the local road construction site, then I should not observe individuals commuting to an urban location for work. On the other hand, if the mechanism is increased access to urban areas, I should be able to observe this in individuals' employment location.[43] I present this analysis in Table 11. In program villages, there is an overall 13 percent increase in the number of people reporting their employment location as urban. For teenaged girls and prime-age men, the coefficients are very large (representing an almost 100 percent increase for men, and a 500 percent increase for girls) and significant. Teenaged boys also witnessed a nearly 100 percent increase in the proportion working in urban areas. Further, this increase is borderline significant. The findings for prime-age men suggest that even though we failed to detect any magnitude changes, being connected to the city brought about qualitative shifts in their employment. Additionally, the results from the analysis of occupations in Table 8 also aid in ruling out

[42] Available at http://www.imd.gov.in/.

[43] Any individuals in the survey are those that necessarily live in the rural household, and not emigrants as the survey collects information for only resident individuals.

this explanation. Table 8 shows that none of the gains in market participation are driven by increased employment at public construction sites.[44]

Yet another potential explanation is that the observed outcomes might be driven by selective migration. However, the observed pattern of coefficients is unlikely to fit any sensible hypothesis about selective migration. For instance, for the observed results to conform with greater out-migration, it would have to be true that the families that left were less likely to send their younger children to school, but more likely to send their older children to school. I also try to formally rule out the migration story by analyzing the impact of the program on household size and on the number of prime-aged men, the demographic most likely to emigrate. The results from this analysis are presented in Table A7, and they fail to reject the null of no selective migration.

8. Discussion and conclusion

The results presented in this paper, specifically the ones on market integration, primary education, and technology adoption underscore the great importance of investment in road construction. For instance, the technology adoption results alone have important policy implications as governments in many developing countries provide large subsidies to promote the use of fertilizer and improved seeds. To quote just one example out of several in the developing world, India spent 0.75 percent of its GDP on fertilizer subsidies in 1999–2000 (Gulati and Narayanan, 2003). Similarly, universal primary education is not just a domestic priority in most developing countries, but was also one of the eight millennium development goals.[45] However, despite large budgetary allocations to the education sector,[46] one in every 10 children of primary school age was still out of school in 2012. My results show that it may be possible to close at least part of this gap merely by improving access to the existing schooling infrastructure.

However, the increased probability of older children dropping out of school is both unexpected and unintended. The labor literature documents significant returns to education. In this specific context, a Mincerian regression of wage on education pegs the return to education at 6.9 percent. Therefore, dropping out of school at an earlier age could potentially be reducing the lifetime earnings of these individuals. On the other hand, the expected returns to education in rural India are debatable. Further, even if lifetime earnings were going down, there may not be any welfare losses for individuals with sufficiently high discount rates. Unfortunately, the available data does not allow me to isolate these parameters. Additionally, it must be understood that this paper only analyzes short-run impacts. It is possible that as the income effect begins to dominate the substitution effect, the long-run steady state could correspond to higher enrollment. However, given our normative preference for schooling, policy-makers may still want to design measures to mitigate the short-run effect. Finally, it is worth reiterating the argument made by Banerjee et al. (2012) in the context of their study of railroads in China. They argue that a lack of labor mobility in China resulted in the gains from the railroad to stay localized. India is similarly known for low levels of labor mobility (see, for instance, Munshi and Rosenzweig, 2016),[47] which likely had implications for how rural households responded to PMGSY. For instance, if structural imperfections prevent an individual from migrating to the city to work in a high-skill occupation, better access to good schools is unlikely to increase skill acquisition. In a similar vein, Casaburi et al. (2013) argue that the incidence of the benefits of rural road rehabilitation programs depends on the competitive structure of local markets.

More broadly, my findings highlight the importance of creating a conducive macroeconomic environment for growth that allows agents to make choices that enable them to locate as close to their production-possibility frontier as possible. For instance, it has been notoriously hard to make much progress on fertilizer use due to perceived demand constraints such as lack of credit or limited information. However, I observe a significant uptick in fertilizer usage after road construction. It is noteworthy that this happened without any contemporaneous demand stimulation, and merely because the presence of roads made it more profitable for farmers to adopt this technology.

Apart from the outcomes studied in this paper, roads can potentially impact several other economic variables. Access to credit markets, healthcare, service delivery, and changes to economic geography are some that come to mind. Research is needed on these before we fully understand these partial equilibrium effects of transport infrastructure provision on economic outcomes, or the general equilibrium effects. Finally, an emerging body of work has shown urban infrastructure to create long-term path dependencies (Bleakley and Lin, 2012; Berger and Enflo, 2015; Jedwab et al., 2017), and future research should also explore persistence in the economic effects of rural infrastructure.

[44] The occupation codes included in the category construction pertain to private construction sites. The bulk of this category corresponds to employment as casual labor at private individual homes.

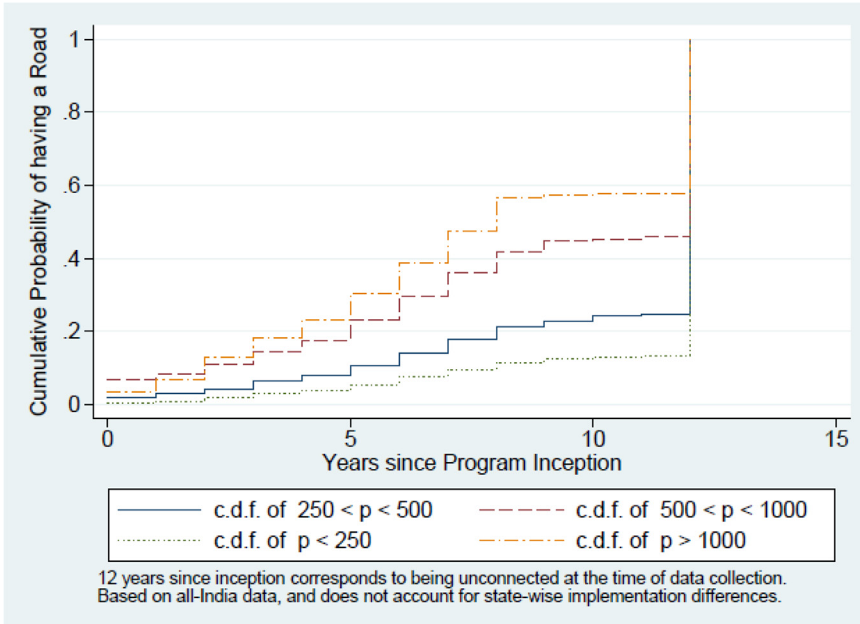
[45] Quality education continues to be a policy priority as part of the new sustainable development goals as well.

[46] According to Dongre et al. (2014), public expenditure on primary education in India amounted to 1.75 percent of GDP in 2011–12.

[47] Colmer (n.d.) reports some statistics: In the 2001 census, 9.5 percent of the population of India had reported migrating in the past decade, an average of 0.9 percent per year. For perspective, 10 percent of the households in the United States migrate internally every year. Of those that migrated over 1991–2001, 82 percent migrated within the same state, and 60 percent within the same district. Moreover, the NSS reports that more than 90 percent of all female migrants do so as a result of getting married, making the rate of labor-market-related migration even lower (NSS, 2010).

Appendix

A.1. CDF of connectivity



A.2. Village-level observables at baseline (Year 2001)

Table A2
Summary statistics of connected & unconnected villages.

Observables	Means		p value
	Connected	Unconnected	Connected = Unconnected
Total Population	929.20 (2345.04)	625.70 (864.47)	<0.0001
SC Population	0.32 (3.69)	0.37 (0.81)	<0.0001
Panchayat HQ	0.19 (0.39)	0.08 (0.27)	<0.0001
Primary School	0.88 (0.32)	0.79 (0.41)	<0.0001
High School	0.07 (0.26)	0.03 (0.17)	<0.0001
Adult Literacy Center	0.16 (0.36)	0.08 (0.28)	<0.0001
Primary Health Center	0.09 (0.29)	0.03 (0.18)	<0.0001
Maternal & Child Welfare Center	0.11 (0.31)	0.06 (0.24)	<0.0001
Commercial Bank	0.13 (0.33)	0.05 (0.22)	<0.0001
Post Office	0.52 (0.71)	0.23 (0.49)	<0.0001
Telegraph	0.05 (0.24)	0.01 (0.11)	<0.0001
Telephone	0.53 (0.50)	0.26 (0.44)	<0.0001
Power Supply	0.90 (0.30)	0.71 (0.46)	<0.0001
Distance from Town	20.78 (21.45)	25.29 (27.31)	<0.0001
Observations	4,77,917	2,80,210	

Standard deviations in parentheses.

A.3. Provision of other public goods

Table A3
Does a PMGSY road predict other public goods?

	Elementary School	High School	Primary Health Center	Post Office	Bus Station	Railway Station	Bank
PMGSY beneficiary Village	0.0027	0.0038	0.0012	0.0016*	0.0285***	−0.0007	−0.0013
	(0.005)	(0.003)	(0.001)	(0.001)	(0.005)	(0.0004)	(0.001)
Observations	1,75,708	1,75,708	1,75,708	1,75,708	1,75,708	1,75,708	1,75,708
R-Squared	0.013	0.005	0.001	0.001	0.005	0	0.001

Standard errors in parentheses, clustered at the state level.
***, **, * indicate significance at 1, 5 and 10%.
All specifications have district fixed effects.
All specifications control for village population category at baseline, i.e., 2001 census.

A.4. Impact of road construction on village-level consumption variety

Table A4
Impact of road construction on consumption variety in the village.

	Impacts by Item Type											
	Food							Non-Food				
	Cereals	Lentils	Dairy	Meat	Vegetables	Fruit	Processed Food	Contraceptives	Minor Manufactures	Road Fares	Non-road Fares	Vehicles
Roads Built	−0.11	0.35	0.27***	0.23*	0.59	0.28	0.48	0.52***	1.06***	0.32***	−0.01	0.09
	(0.23)	(0.24)	(0.10)	(0.12)	(0.41)	(0.25)	(0.30)	(0.10)	(0.32)	(0.11)	(0.03)	(0.08)
Observations	52,029	52,029	52,029	52,029	52,029	52,029	52,029	52,029	52,029	52,029	52,029	52,029
R-Squared	0.17	0.15	0.05	0.12	0.14	0.11	0.15	0.05	0.13	0.07	0.01	0.06
Mean of Dep. Var.	5	4	1	3	14	3	4	0	2	1	0	1
Std Dev of Dep. Var.	2	2	1	1	3	2	2	0	1	1	0	0

Notes:
Standard errors in parentheses, clustered at the district level.
***, **, * indicate significance at 1, 5 and 10%.
All specifications have time and district fixed effects.
Mean of % Connected is 0.081.
The dependent variable is the number of surveyed goods in each category that are consumed by at least 1 household in the village.

A.5. Consumption variety (baseline and endline only)

Table A5
Impact of road construction on consumption basket (2 periods only).

	Impacts by Item Type											
	Food							Non-Food				
	Cereals	Lentils	Dairy	Meat	Vegetables	Fruit	Processed Food	Contraceptives	Minor Manufactures	Road Fares	Non-road Fares	Vehicles
Post Dummy	0.60***	0.52***	0.05***	0.06***	0.75***	0.06*	0.44***	0.02***	0.30***	0.17***	0.01***	0.06***
	(0.04)	(0.04)	(0.01)	(0.02)	(0.11)	(0.03)	(0.05)	(0.01)	(0.03)	(0.02)	(0.00)	(0.01)
Post * Roads Built	−0.58***	−0.36**	0.14**	0.10	0.03	0.27	0.26	0.18***	0.43***	0.28***	0.01	0.20***
	(0.17)	(0.18)	(0.06)	(0.11)	(0.45)	(0.17)	(0.26)	(0.05)	(0.14)	(0.10)	(0.02)	(0.05)
Observations	93,266	93,266	93,266	93,266	93,266	93,266	93,266	93,266	93,266	93,266	93,266	93,266
R-Squared	0.09	0.08	0.04	0.04	0.07	0.05	0.07	0.02	0.06	0.04	0.00	0.05
Mean of Dep. Var.	2.82	2.77	0.89	1.47	10.01	1.66	2.07	0.07	1.30	0.81	0.04	0.36
Std Dev of Dep. Var.	1.33	1.59	0.71	1.29	3.48	1.31	1.55	0.26	0.91	0.64	0.19	0.48

Notes:
Standard errors in parentheses, clustered at the district level.
***, **, * indicate significance at 1, 5 and 10%.
All specifications have state fixed effects and household-level controls.
Household controls include household size, religion, social group, and occupation.
Estimates based on Equation (3).
Mean of % Connected is 0.154.
The dependent variable is the number of surveyed goods in each category that are consumed by the household.

A.6. Calculations for elasticity of demand for fertilizer

This appendix lays out the calculations for the elasticity of demand for fertilizer used in the discussion of the technology adoption results at the end of Section 6.5. These calculations are based on two assumptions: 1) When a farmer reports that a certain plot is being cultivated using fertilizer, then he/she is fertilizing the entire plot (and not part thereof), and 2) The fertilizer in question is Urea. Urea is the most widely used fertilizer in India, constituting over 80 percent of all Nitrogen fertilizers used, and well over half of total fertilizer used. The government of India strictly controls the market price of Urea (and other fertilizers) with a subsidy to manufacturers. Between 2002 and 2009, the controlled price of Urea did not change.

Table A6
Workings to calculate elasticity of demand for fertilizer.

Metric	Value	Source
Assumptions		
Average fertilizer use per hectare in 2002	100.3 kg	FAOStat
Transport cost on paved road	Rs. 0.30 per 100 kg per km	Mohapatra and Chandrasekhar, 2007
Transport cost on unpaved road	Rs. 2.00 per 100 kg per km	Mohapatra and Chandrasekhar, 2007
Average distance from town for treated villages	25 km	Table A2
Subsidized price of 50 kg bag of Urea	Rs. 276.00	Government of India
% increase in area under fertilizer (food crops)	9%	Table 9
Implications for treated villages		
Average transport cost for 1 ha. (100 kg) of fertilizer pre-PMGSY	Rs. 50.00	
Average transport cost for 1 ha. (100 kg) of fertilizer post-PMGSY	Rs. 7.50	
Village price of 100 kg Urea pre-PMGSY	Rs. 602.00	
Village price of 100 kg Urea post-PMGSY	Rs. 559.50	
⇒ % decrease in price of fertilizer	7%	
Price elasticity of demand for fertilizer	**1.29**	

A.7. Impact of road construction on household-size

Table A7
Impact of program intensity on household size & composition.

	Size of Household	Number of 18–40 year old males
Roads Built	0.01(0.128)	0.00(0.082)
Observations	3,62,237	3,62,237
R-Squared	0.04	0.03
Mean of Dep. Var.	5.36	1.50
Std Dev of Dep. Var.	2.65	1.31

Standard errors in parentheses, clustered at the district level.
***, **, * indicate significance at 1, 5 and 10%.
All specifications have time and district fixed effects, and household-level controls.
Mean of Roads Built: 0.081.

Appendix references

Food and Agriculture Organization of the United Nations (2014) “FAOSTAT Statistics Database”. FAO, Rome, Italy.

Mohapatra J.K. and B.P. Chandrasekhar (2007) “Rural Roads” in *India Infrastructure Report, 2007*. Indian Institute of Technology, Kanpur, India.

References

Aaronson, D., Mazumder, B., 2013. The impact of Rosenwald schools on black achievement. J. Polit. Econ. 119 (5), 821–888.

Adukia, A., Asher, S., Novosad, P., 2017. Educational Investment Responses to Economic Opportunity: Evidence from Indian Road Construction. Unpublished.

Aggarwal, S., Giera, B., Jeong, D., Robinson, J., Spearot, A., 2017. Market Access, Trade Costs, and Technology Adoption: Evidence from Northern Tanzania. Working Paper.

Aggarwal, S., Francis, E., Robinson, J., 2018. Grain Today, Gain Tomorrow: Evidence from a Storage Experiment with Savings Clubs in Kenya. NBER Working Paper 24391.

Aker, J., 2010. Information from markets near and far: mobile phones and agricultural markets in Niger. Am. Econ. J. Appl. Econ. 2 (3), 46–59.

Alder, S., 2017. Chinese Roads in India: the Effect of Transport Infrastructure on Economic Development. Unpublished.

Arkolakis, C., Demidova, S., Klenow, P.J., Rodríguez-Clare, A., 2008. Endogenous variety and the gains from trade. Am. Econ. Rev.: Pap. Proc. 98 (2), 444–450.

Asher, S., Novosad, P., 2016. Rural Roads and Structural Transformation. Working Paper.

Asturias, J., Garcia, M., Ramos, R., 2017. Competition and the Welfare Gains from Transportation Infrastructure: Evidence from the Golden Quadrilateral of India. Unpublished.

Atack, J., Bateman, F., Haines, M., Margo, R.A., 2010. Did railroads induce or follow economic growth? Urbanization and population growth in the American Midwest, 1850-60. Soc. Sci. Hist. 34, 171–197.

Atkin, D., 2016. Endogenous skill acquisition and export manufacturing in Mexico. Am. Econ. Rev. 106 (8), 2046–2085.

Atkin, D., Donaldson, D., 2015. WhoŠs Getting Globalized? the Size and Implications of Intranational Trade Costs. NBER Working Paper 21439.

Banerjee, A., Duflo, E., Qian, N., 2012. On the Road: Access to Transportation Infrastructure and Economic Growth in China. NBER Working Paper 17897.

Berger, T., Enflo, K., 2015. Locomotives of local growth: the short- and long-term impact of railroads in Sweden. J. Urban Econ. 98, 124–138.

Binswanger, H., 1974. A cost function approach to the measurement of elasticities of factor demand and elasticities of substitution. Am. J. Agric. Econ. 56, 377–386.
Bleakley, H., Lin, J., 2012. Portage and path dependence. Q. J. Econ. 127, 587–644.
Broda, C., Weinstein, D.E., 2010. Product creation and destruction: evidence and price implications. Am. Econ. Rev. 100 (3), 691–723.
Burde, D., Linden, L.L., 2013. The effect of village-based schools: evidence from a randomized controlled trial in Afghanistan. Am. Econ. J. Appl. Econ. 5 (3), 27–40.
Burgess, R., Jedwab, R., Miguel, E., Morjaria, A., Miquel, G.P., 2015. The value of democracy: evidence from road building in Kenya. Am. Econ. Rev. 105 (6), 1817–1851.
Casaburi, L., Glennerster, R., Suri, T., 2013. Rural Roads and Intermediated Trade: Regression Discontinuity Evidence from Sierra Leone. Unpublished.
Chaudhury, N., Hammer, J., Kremer, M., Muralidharan, K., Halsey Rogers, F., 2006. Missing in action: teacher and health worker absence in developing countries. J. Econ. Perspect. 20 (1), 91–116.
Cole, S., 2009. Fixing market failures or fixing elections? Agricultural credit in India. Am. Econ. J. Appl. Econ. 1 (1), 219–250.
Colmer, J. (n.d.). Urbanization, Growth, and Development: Evidence from India. Unpublished Review Paper.
Datta, S., 2012. The impact of improved highways on Indian firms. J. Dev. Econ. 99 (1), 46–57.
Dercon, S., Gilligan, D.O., Hoddinott, J., Woldehanna, T., 2009. The impact of agricultural extension and roads on poverty and consumption growth in fifteen Ethiopian villages. Am. J. Agric. Econ. 91 (4), 1007–1021.
Dholakia, R., Majumdar, J., 1995. Estimation of price elasticity of fertilizer demand at macro level in India. Indian J. Agric. Econ. 50 (1), 36–46.
Donaldson, D., 2013. Railroads of the Raj: Estimating the impact of transportation infrastructure. Am. Econ. Rev. (Forthcoming).
Donaldson, D., Hornbeck, R., 2016. Railroads and American economic growth: a market access approach. Quart. J. Econ. 131 (2), 799–858.
Dongre, A., Kapur, A., Tewari, V., 2014. How Much Does India Spend Per Student on Elementary Education?. Engaging Accountability: PAISA Report Series, Accountability Initiative.
Duflo, E., 2001. Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. Am. Econ. Rev. 91 (4), 795–813.
Duranton, G., Morrow, P.M., Turner, M.A., 2014. Roads and trade: Evidence from the US. Rev. Econ. Stud. 81 (2), 681–724.
Duranton, G., Turner, M.A., 2012. Urban growth and transportation. Rev. Econ. Stud. 79 (4), 1407–1440.
Faber, B., 2014. Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System. Rev. Econ. Stud. 81 (3), 1046–1070.
Feenstra, R.C., 1994. New product varieties and the measurement of international prices. Am. Econ. Rev. 84 (1-3), 157–177.
Ghani, E., Goswami, A., Kerr, W., 2016. Highway to success: The impact of the golden quadrilateral project for the location and performance of Indian manufacturing. Econ. J. 126, 317–357.
Gibson, J., Olivia, S., 2010. The effect of infrastructure access and quality on non-farm enterprises in rural Indonesia. World Dev. 38 (5), 717–726.
Gibson, J., Rozelle, S., 2003. Poverty and access to roads in Papua New Guinea. Econ. Dev. Cult. Change 52 (1), 159–185.
Gonzalez-Navarro, M., Quintana-Domeque, C., 2016. Paving streets for the poor: experimental analysis of infrastructure effects. Rev. Econ. Stat. 98 (2), 254–267.
Goyal, A., 2010. Information, direct access to farmers, and rural market performance in Central India. Am. Econ. J. Appl. Econ. 2 (3), 22–45.
Gulati, A., Narayanan, S., 2003. The Subsidy Syndrome in Indian Agriculture. Oxford University Press, New York, NY.
Handbury, J., Weinstein, D.E., 2011. Is New Economic Geography Right? Evidence from Price Data. NBER Working Paper No. 17067.
Hillberry, R., Hummels, D., 2008. Trade responses to geographic frictions: A decomposition using micro-data. Eur. Econ. Rev. 52 (3), 527–550.
Heath, R., Mushfiq Mobarak, A., 2015. Manufacturing growth and the lives of Bangladeshi women. J. Develop. Econ. 115, 1–15.
Jacoby, H., 2000. Access to markets and the benefits of rural roads. Econ. J. 110 (465), 713–737.
Jacoby, H., Minten, B., 2009. On measuring the benefits of lower transport costs. J. Dev. Econ. 89 (1), 28–38.
Jedwab, R., Kerby, E., Moradi, A., 2017. History, path dependence and development: evidence from colonial railroads, settlers and cities in Kenya. Econ. J. 127 (603), 1467–1494.
Jedwab, R., Moradi, A., 2012. Colonial Investments and Long-term Development in Africa: Evidence from Ghanaian Railroads. Unpublished. George Washington University.
Jensen, R., 2007. The digital provide: Information (technology), market performance, and welfare in the South Indian Fisheries Sector. Q. J. Econ. 122 (3), 879–924.
Jensen, R., 2012. Do labor market opportunities affect young women's work and family decisions? Experimental evidence from India. Q. J. Econ. 127 (2), 753–792.
Kazianga, H., Levy, D., Linden, L., Sloan, M., 2013. The effect of girl- friendly school construction: Evaluation of the BRIGHT school program in Burkina Faso. Am. Econ. J. Appl. Econ. 5 (3), 41–62.
Keller, W., Shiue, C.H., 2008. Tariffs, Trains, and Trade: the Role of Institutions versus Technology in the Expansion of Markets. CEPR Discussion Paper No. 6759.
Kennedy, G.L., Pedro, M.R., Seghieri, C., Nantel, G., Brouwer, I., 2007. Dietary diversity score is a useful indicator of micronutrient intake in non-breast-feeding Filipino Children. J. Nutr. 137 (2), 472–477.
Khandker, S.R., Bakht, Z., Koolwal, G.B., 2009. The poverty impact of rural Roads: Evidence from Bangladesh. Econ. Dev. Cult. Change 57 (4), 685–722.
Khandker, S.R., Koolwal, G.B., 2011. Estimating the Long-term Impacts of Rural Roads: a Dynamic Panel Approach. World Bank Policy Research Working Paper No. 5867.
Khemani, S., 2004. Political cycles in a developing economy: Effect of elections in the Indian States. J. Dev. Econ. 73, 125–154.
Kruger, D.I., 2007. Coffee production effects on child labor and schooling in rural Brazil. J. Dev. Econ. 82 (2), 448–463.
Lehne, J., Shapiro, J.N., Eynde, O.V., 2018. Building connections: political corruption and road construction in India. J. Develop. Econ. 131, 62–78.
McIntosh, C., Shumway, C., 1994. Multiproduct production choices and policy response. West. J. Agric. Econ. 16 (1), 291–303.
Ministry of Statistics and Programme Implementation, 2011. Key IIndicator of Household Consumer Expenditure in India, 2009-2010 Technical report. National Sample Survey Office, New Delhi, India.
Munshi, K., Rosenzweig, M., 2016. Networks and misallocation: insurance, migration, and the rural-urban wage gap. Am. Econ. Rev. 106 (1), 46–98.
Muralidharan, K., Prakash, N., 2017. Cycling to school: increasing secondary school enrollment for girls in India. Am. Econ. J. Appl. Econ. 9 (3), 321–350.
Muto, M., Yamano, T., 2009. The impact of mobile phone coverage expansion on market participation: Panel data evidence from Uganda. World Dev. 37 (12), 1887–1896.
Nelson, L.K., 2011. From Loans to Labor: Access to Credit, Entrepreneurship and Child Labor. Working Paper. Michigan State University.
Nguyen, K.-T., Do, Q.-A., Tran, A., 2012. One Mandarin Benefits the Whole Clan: Hometown Infrastructure and Nepotism in an Autocracy. Working Paper. Indiana University.
NSS, 2010. Migration in India: 2007-2008. Report 533. National Sample Survey Office, Delhi, India.
Rasul, I., Rogger, D., 2018. Management of bureaucrats and public service delivery: evidence from the Nigerian civil service. Econ. J. 128 (608), 413–446.
Ray, S., 1982. A translog cost function analysis of U.S. agriculture, 1939-77. Am. J. Agric. Econ. 64, 490–498.
Rogger, D., 2013. The Causes and Consequences of Political Interference in Bureaucratic Decision Making: Evidence from Nigeria. Unpublished. UCL.
Schady, N., 2004. Do macroeconomic crises always slow human capital accumulation? World Bank Econ. Rev. 18 (2), 131–154.
Shah, M., Steinberg, B.M., 2017. Drought of opportunities: contemporaneous and long term impacts of rainfall shocks on human capital. J. Political Econ. 125 (2), 527–561.
Shamdasani, Y., 2016. Rural Road Infrastructure and Agricultural Production: Evidence from India. Unpublished.
Shamdasani, Yogita, 2017. Rural Road Construction and Agricultural Production: Evidence from India. Working Paper.
Shumway, C., 1983. Supply, Demand, and Technology in a Multiproduct Industry: Texas Field Crops. Am. J. Agric. Econ. 65, 749–760.
Steinwender, C., 2018. The real effects of information frictions: when the states and the Kingdom became United. Am. Econ. Rev. 108 (3), 657–696.
Subramaniyan, G., Nirmala, V., 1991. A macro analysis of fertiliser demand in India (1966-67 to 1985-86). Indian J. Agric. Econ. 46 (1).
Suri, T., 2011. Selection and comparative advantage in technology adoption. Econometrica 79 (1), 159–209.
Tontisirin, K., Nantel, G., Bhattacharjee, L., 2002. Food-based strategies to meet the challenges of micronutrient malnutrition in the developing world. Proc. Nutr. Soc. 61, 243–250.
Volpe Martincus, C., Carballo, J., Cusolito, A., 2017. Roads, exports and employment: Evidence from a developing country. J. Dev. Econ. 125, 21–39.
World Bank, 2007. A Decade of Action in Transport: an Evaluation of World Bank Assistance to the Transport Sector, 1995 – 2005. Independent Evaluation Group, World Bank, Washington, D.C.
World Bank, 2009. Reshaping Economic Geography World Development Report. World Bank, Washington, D.C.