Roads, Participation in Markets, and Benefits to Agricultural Households: Evidence from the Topography-based Highway Network in Nepal

Slesh A. Shrestha*
National University of Singapore

March 30, 2018

Abstract

This paper identifies the role of roads in improving agricultural livelihoods, and examines the key market mechanisms through which improved connectivity translates into economic gains for agricultural households. I use a rigorous identification strategy based on the rugged terrain that significantly influences the design and costs of constructing roads in Nepal, together with a new geospatial data, to find a positive impact of road on farmland values. A 1 percent decrease in distance to a road raises the market price of an agricultural plot by 0.1 to 0.25 percent. This increase in land value is underpinned by increased participation by households in agricultural markets, and improved farm production and incomes. The results also suggest that a decrease in the distance to a road contributes to the commercialization of agriculture, and it increases the use of fertilizer in agricultural production and reduces the unit cost of fertilizers.

*Shrestha: Department of Economics, National University of Singapore, 1 Arts Link, AS2 Level 6, Singapore 117570, slesh@nus.edu.sg. This article benefited greatly from the comments of the editor and anonymous referees and would not have been possible without the support of the Central Bureau of Statistics, National Planning Commission of Nepal for its assistance with the data. I am grateful to Dean Yang, Raj Arunachalam, David Lam, Susan Dynarski, Charles Brown, Nethra Palaniswamy, and seminar and conference participants at the University of Michigan, National University of Singapore, and Northeast Universities Development Consortium Conference for their helpful comments and to Pei Zhi Chia, Ruijie Tan, and Pankaj Bajracharya, who provided outstanding research assistance.

1 Introduction

The role of infrastructure in reducing poverty has long been a subject of empirical investigation (Lipton and Ravallion, 1995; Datt and Ravallion, 1998; Gibson and Rozelle, 2003; Lokshin and Yemtsov, 2005; Khandker et al., 2009). This literature highlights the relevance of the agricultural sector in improving incomes of the poor (Gollin et al., 2002; Gollin, 2010); and the role of improved market access for agricultural households in achieving these pro-poor outcomes (Binswanger et al., 1993; Minten and Kyle, 1999; Minot and Hill, 2007).

Roads were an early choice for large-scale public investments in poor and rural countries, and they continue to be a popular investment for both governments and aid agencies (van de Walle, 2002; World Bank, 2013). While there are sound economic arguments as to why investments in road should benefit agrarian households, estimating this economic gain is not straightforward. Roads are not randomly placed, but they are instead endogenously determined by various socio-economic factors that are likely to confound the estimate based on a simple comparison across regions with various degree of road infrastructure. I use the rugged geography of the terrain in Nepal to overcome this endogeneity problem to estimate the effects of road on household's decision to participate in agricultural markets; measure the scope of this participation; and examine if these roads also bring welfare gains through improved agricultural production and incomes. I also estimate the effect of roads on farmland values, to quantify these immediate and inter temporal economic gains from improved market integration in agriculture.

Road investments are typically costly; and these costs can escalate rapidly in remote areas and with difficult geographies. Although the costs of investment in roads are often large, there is limited empirical evidence on whether the economic gains justify the levels of expenditure. In Nepal, road construction is the largest public investment program, accounting for 14 percent of its total development budget over the last five decades (Government of Nepal, 2002a). During this period, its road network has expanded by more than forty-fold from 376 kilometers in 1951 to 15,308 kilometers in 2002. Nevertheless, more than 20 percent of its agrarian households still do not participate in either input or output agricultural markets, and more than 80 percent of its poor households draw income from agriculture. This makes it an ideal case to examine the effect of public investments in roads on market choices made

by poor agrarian households; and to examine if these investments bring gains in economic welfare to economies that continue to be dominated by subsistence agriculture.

More than 80 percent of Nepal's terrain is covered by mountains. This rugged terrain significantly influences the design and costs of constructing road networks. In this setting, I estimate the economic impact of roads using the 2010 Nepal Living Standards Survey, which collected information on household's agricultural activities, its connectivity and participation in the agricultural market, information on every agricultural plot it owned, and its GIS location. I merge this data with a larger set of geographic and climatic maps collected from different sources to construct a unique geospatial data of socio-economic and environmental variables for Nepal.

I develop an algorithm that uses the newly constructed geospatial data and predicts the most cost-effective design to link all 75 district headquarters into a single national road network based on three cost factors: land gradient, river crossing, and surface distance. I then use the household's distance to this predicted road network to examine the impact of road on agriculture using three different instrumental variable (IV) strategies. The IV approach that relies on the spatial variation identified by the predicted network is valid if this network is not correlated with land quality that directly affects land values. Each IV model therefore seeks to address this plausible correlation in different ways.

The first IV model uses the distance to the predicted road to instrument for household's distance to the actual road, and controls for observed land quality and climatic variables to capture any land heterogeneity that might be correlated with the predicted road network. Alternatively, I use the interaction term of district-level road completion rate and distance to the predicted road to instrument for road accessibility. This second IV model includes the distance to the predicted road variable as a separate control to capture the spatial correlation between unobserved land quality and the predicted road network. The third IV model utilizes the cross-sectional differences in the growth of road networks over time (for a subset of panel households) as predicted by the cost-based algorithm, and the longitudinal data allows to control for any time-invariant spatial heterogeneities in land quality correlated with the predicted road.

It has been long established that many agricultural households in poor countries often do not participate in markets and stay in subsistence farming, due to high transaction costs

(Strauss, 1984; de Janvry et al., 1991; Renkow et al., 2004).[1] Decreasing these transaction costs can, however, improve the marketable surplus of agricultural households at both the extensive and intensive margins (Goetz, 1992; Key et al., 2000). Public investments in roads that reduce transportation costs have also been demonstrated to improve access to agricultural markets. This evidence suggests that the economic gain from market access can manifest itself at the household level in a decrease in the costs of agricultural inputs and an increase in the "effective" prices that farmers receive for their produce,[2] both of which are likely to improve the agricultural production, incomes and the livelihoods of agricultural households (Dorosh et al., 2012; Binswanger et al., 1993; Gertler et al., 2015). Jacoby (2000) finds large economic gains for agrarian households, and examines the distributional impacts from improved market access through roads using farmland values in Nepal. In this paper, I use a rigorous identification strategy along with new geographic data to measure the economic gains that accrue to households from improved access to roads; and I identify the key market mechanisms through which this access translates into improved economic welfare for agricultural households.

The estimated impacts of roads on farmland values are positive and similar in magnitude across all three IV models as well as in the OLS. The results suggest an increase in the market price of an agricultural land of 0.1 to 0.25 percent, for every 1 percent decrease in its distance to a road. For a median household that resides 8 kilometers from the nearest road, shortening this distance by 1 kilometer is estimated to generate economic benefits of around $185 per hectare of agricultural land, which represents a two-fold return in the agriculture sector from Nepal's public road investments. I find that this increase in land value is underpinned by increased participation in agricultural markets; and that increased farm productivity and revenues accompany it. The results also suggest that a decrease in the distance to

[1] de Janvry et al. (1991), among others, examine the output and market supply responses of agrarian households to changes in various price and non-price variables, and they find low supply elasticity in developing countries. They argue that high transaction costs in these countries mean that either the markets are completely missing; or while some farmers participate in the market, many households choose to remain self-sufficient to economize on transaction costs. Renkow et al. (2004) find that farmers in Kenya face fixed transaction costs that are on average equivalent to a 15% ad valorem tax.

[2] Fafchamps and Hill (2005) find that access to markets in Uganda influences coffee farmers' decision to sell their produce directly to the market where they get a higher price for their produce compared to selling at farm-gate to intermediate traders. Likewise, Minten and Kyle (1999) find that in Former Zaire the distance to the urban center decreases the effective price that the rural farmers get for selling their produce in the urban market. Moveover, on bad roads, the effective price for the farmer decreases faster than transportation cost increases with the distance to the urban market, suggesting other market transaction costs can also arise with bad road quality. Casaburi et al. (2013) on the other hand find that market prices of local crops like rice and casava decrease with improvements in roads in Sierra Leone.

a road contributes to the commercialization of agriculture, along with the increased use of fertilizers in agriculture production, and a decline in per unit costs of fertilizer. These results, taken together, suggest that poor road infrastructure significantly raises the market transaction costs for many agrarian households, and these costs can severely limit their livelihood options.

This paper also relates to the role of geography in determining household welfare through its influence on local public goods (Jalan and Ravallion, 2002). Duflo and Pande (2007), Dinkelman (2011), and Lipscomb et al. (2013) use geography to explain the placement and timing of infrastructure investments in irrigation and electricity. Jacoby and Minten (2009) compare the benefits that accrue to households on different points on a transport cost gradient, where this gradient is implicitly defined by the regions impenetrable mountain geography. Dorosh et al. (2012) and Emran and Hou (2013) formalize and quantify the role of geography in determining households' access to roads, and estimate its impacts on crop production and consumption. Other transportation studies like Michaels (2008), Faber (2014), and Donaldson (forthcoming) focus primarily on price convergence and other regional outcomes from the viewpoint of a trade framework. By using micro data on land prices and the production and market decisions of agricultural households, the results in this paper seek to understand the role of road infrastructure in improving household incomes in the agricultural sector.

A large body of literature also examines the variation in economic activities across space, mainly as a function of its geographical proximity to cities (Fujita et al., 1999; Kanbur and Venables, 2005; Fafchamps and Shilpi, 2005), which could have important implications for spatial inequality in developing countries. In Nepal, recent studies have found that proximity to (and size of) urban centers can explain the spatial differences in the household's division of labor allocation and crop choices (Fafchamps and Shilpi, 2003; Emran and Shilpi, 2012). Roads that connect or come closer to remote areas could influence spatial inequality; and the results in this paper suggest that roads could hold potential to reduce such inequality in Nepal.

2 Background

Nepal is largely an agrarian society with an extremely poor transportation infrastructure. An estimated 40 percent of Nepal's 24 million citizens live at least two hours' walk from the nearest all-season road, and the country's road density is one of the lowest in the world at 14 kilometers of roads per 100 square kilometers (Government of Nepal, 2007; Meyer, 2008).

Road construction in Nepal requires different design parameters and priorities than those conventionally adopted for less severe flat or rolling terrains. Mountains cover almost 80 percent of the country's land surface, and the rugged terrain significantly raises the cost of road projects, mainly through sharp changes in elevation. The steepness of the terrain raises the total quantity of soil and hard rock that must be excavated and transported during road construction, and the placement of roads on slopes steeper than 30 degrees requires the construction of embankment retaining walls. Such terrain hazards mean that the estimated construction cost per kilometer is one of the highest in the world (Overseas Development Administration, 1997).

Despite the engineering challenges inherent to constructing high mountain roads, the Nepalese Government invested heavily in developing its transportation infrastructure. Starting 1951, after the democratization of the country's political system, road construction became a top priority for development (Government of Nepal, 2002a). Since then, Nepal's highway road network has expanded from a virtually non-existent 115 kilometers to 5,030 kilometers in 2006, made up of 15 national highways (Government of Nepal, 2007).

The expansion began with the construction of the East-West Highway (EWH) in the mid-1950s. Although many road projects have followed since, the EWH is still the longest and the only highway that spans the entire country, connecting its eastern and western borders. It is 1,027 kilometers in length, and accounts for 30 percent of the country's total highway network.

All other highways eventually feed into the EWH, making it the backbone of Nepal's road network system. Most of these other highways are north-south corridor roads that link district headquarters and their hinterlands with the EWH and, therefore, to the national road network. Until recently, integrating district headquarters into the national road network

was the main reason for all major road constructions (Government of Nepal, 2002a,b, 2007).[3]

Connecting them with roads requires careful consideration of geology, slope, and cost of construction, and the shortest road alignment is not necessarily the easiest or the cheapest option. For instance, a district's headquarters may link to a neighboring district's headquarters not by the most direct route but by connecting both of them to the EWH. In this way, the EWH and the networks of north-south corridor roads play an important role in overcoming unique geographic constraints of the terrain to link district headquarters.

As of 2007, road projects connecting district headquarters to EWH had been completed in 52 of the 75 districts. At the same time, construction work to connect the remaining 23 district headquarters was underway with varying degrees of completion, except in three districts that had no proposed work to connect their district headquarters. In 11 of the districts with on-going work, their district headquarters already had limited access to a seasonal dirt road (Government of Nepal, 2007).

3 Data and Variable Construction

The data for empirical analysis come from multiple sources. I obtained household information on agriculture and access to infrastructure from the 2010 Nepal Living Standards Survey (NLSS), a nationwide survey conducted by the Nepal Central Bureau of Statistics. The survey collected data on all agricultural plots owned or leased by the household: size, quality, market value, and the net rent received by the household if the plot was leased out. In addition, it measured the household's proximity to the nearest paved road along with its visits to the market center. For households involved in cultivating their owned or rented land, it also measured their use of various farm inputs, total agricultural output, and income from their agricultural activity.

Table 1, Panels A and B provide summary statistics for 15,717 agricultural plots and 4,989 agrarian households respectively. The median plot size is 0.1 hectare, with market value of $12,088 per hectare. Almost 90 percent of plots are self-cultivated, and 97 percent of them have legal land titles. An average household cultivates 0.616 hectare of land, with 18

[3]Most district headquarters are former fort towns, there were only recently converted into administrative centers. They are often situated on the tops of mountains, mainly for historical military advantage.

percent of the cultivated area rented-in (11 percent under sharecropping contract). About 70 percent of agrarian households use chemical fertilizer in their farms, and even fewer, 47 percent sell their produce. This is perhaps not surprising because only 30 percent of households visit the market center at least once every month, and the average distance to the nearest paved road is 8 km.

For 350 households interviewed in 2010, information about their land-holdings and access to roads was also collected in 1996. I construct a two-wave household panel and calculate the change in their distance to the nearest paved road and the change in their total and types of land-holdings, along with the change in the average market price of their owned land between the two rounds.

A unique feature of the NLSS is that it collected GIS information on the location of households. I use this GIS data to merge the geospatial data on terrain characteristics and environment from different sources. Mountain Environment and Natural Resources Information System (MENRIS) data from International Center for Integrated Mountain Development (ICIMOD) contain information on soil quality and annual precipitation, together with the locations of district headquarters, the EWH, and walking trails. The elevation and temperature data come from the Global Climate Database complied by WorldClim.[4] Appendix Figure A1 provides the spatial maps of elevation, precipitation, temperature, and soil type. I use elevation map to construct the measure of land gradient and calculate the location of rivers, both of which I use as inputs for predicting the road network, as discussed later in the paper.[5] Lastly, I collect district-level statistics on the actual road network from the Department of Roads.

3.1 Value of an Agricultural Land

I use land market values to estimate the economic benefits from improved connectivity to markets. The standard asset-pricing model suggests that the present market value of a plot is equal to the discounted sum of its future rents, where rent normally captures the economic profits from the plot. In Nepal, this link between the observable land values and

[4]The environmental and topography data can be downloaded at the Worldclim website, www.worldclim.org, and the ICIMOD website, www.geoportal.icimod.org.

[5]The identification of a river is based on the Strahler Stream Order Classification. The order of 4 or greater defines a river.

mostly unobservable land rents might however be tenuous for two main reasons.

First, the self-reported land value is measured by the following survey question: if you wanted to buy/sell a plot exactly like this, how much would it cost/fetch you? If land sale markets are mostly inactive, the owner's expectation and prior knowledge can influence the self-reported valuation of land, and it can significantly deviate from the true market value. In my data, about 8.7 percent of land-owning households either bought or sold land during the previous year, and 24.7 percent purchased some of their acquired land, suggesting a fairly active land market. Moreover, among 3,006 households that cultivated their owned land, the correlation between their self-reported land value and land productivity is positive and high, with the estimated elasticity of 0.738.

Second, the distortions in the credit market or the weak laws on property rights in developing countries mean that the market value of land could be affected by its collateral value, or due to the private costs of enforcing property rights. Therefore, I test the relationship between rent and market value of a plot, as described by the asset-pricing model. This relationship is given by

$$log(V) = log(r) - log(b) \tag{1}$$

where V is the present market value, r is the land rent, and b is the constant discount rate.

I regress the log of land values on the log of rents and test whether the coefficient on rents equals one. For this purpose, I use 1,162 plots in the 2010 NLSS data that were rented out either in both agricultural seasons or in the wet season. I calculate their net rents by summing the rents across both seasons and adding the value of in-kind payments, while deducting the costs of inputs provided to the tenants. The average rent-to-value ratio is 0.044, which can be interpreted as an estimate of the discount rate.

The estimated coefficients on rents using different specifications are presented in Appendix Table A1. All estimates are close to the magnitude of one and accurate but not statistically different from one. Therefore, the validity of the asset-pricing model cannot be rejected.[6]

[6] Jacoby (2000) conducts a similar test using the 1996 NLSS data, and finds that the validity of the asset-pricing model can not be rejected.

3.2 The Instrument: Construction Cost Minimization

The instrument is based on the national road network that is predicted by the model to minimize construction costs resulting from geographic characteristics. I utilize three main features of road construction in Nepal: first, the existence of the EWH; second, the emphasis on integrating all district headquarters into the national road network; and third, geography-based cost considerations. The model takes as inputs the location of the EWH and the location of the 75 district headquarters and goes on to build a road network based on two geographic characteristics of the terrain: river and land gradient.

While the length of the road is a standard determinant of its construction cost, geographical factors such as steepness and hydrology are significant factors that impact the design and alignment of roads in mountainous regions. For example, river crossings require the construction of bridges, which account for 20 percent of a typical road project's total budget (Government of Nepal, 2002a). The per kilometer construction cost of a two-lane bridge in the Mountain belt of Nepal is estimated at $13.6 million, compared to the $0.2 million per kilometer cost of constructing a paved road (Overseas Development Administration, 1997; World Bank, 2012).

Furthermore, most roads in Nepal are constructed by cutting into the slope of the mountain, which requires excavation of soil and rocks from the mountain face. The excavation quantity from the cut-and-throw practice increases with slope, from 147.62 cubic meters per kilometer in a terrain with gradient of 0-10 degrees to 3053.26 cubic meters per kilometer in a terrain with a 30-40 degree gradient (Shrestha, 2010). In addition, steepness also increases environmental hazards such as landslide and soil erosion that could undermine the construction and maintenance of the road. Estimates from Papua New Guinea suggest that such hazards increase exponentially with slope angle for all types of soil, while any slope greater than 30 degrees requires an embankment retaining wall, all of which add considerably to the total cost of construction (Overseas Development Administration, 1997).

For each district headquarter, the decision to connect it either with any one of the other district headquarters or with the EWH and the exact path of that connection is jointly determined based on minimizing the total cost of building the road, where the cost of

building a road across each 1x1 square kilometer area is given by:

$$River\ Factor\ *\ Vertical\ Factor\ *\ Surface\ Distance \tag{2}$$

The river factor takes a value of 2 if there is a river in the area and 1 otherwise. The vertical factor is based on the inverse symmetric linear function of land gradient in the area, which is calculated using an elevation map of the region. The cost function and the value of the parameters in the above model are based on the standard functions available in ArcGIS mapping software. While alternative values would only affect the strength of the first stage regression in the instrumental variable strategy discussed later, the IV results that are based on the prediction of the model are robust to these choices.[7]

After the first round, each district headquarters will have a road connecting it to another district headquarters or to the EWH, but there could be districts or clusters of districts that are not yet integrated into the national road network. If so, in the second round, for each "fragmented" cluster, the model would determine the cheapest road that would connect any one of the districts inside the "fragmented" district cluster to either another district headquarters outside the cluster or to the EWH using the same cost minimizing rule as above. The model would repeat the rounds until all 75 district headquarters are directly or indirectly connected to each other. In the above example, all 75 district headquarters were connected into a single integrated network of roads after the second round.

Figure 1 provides the cross-section maps of land gradient, river networks, district headquarters and the EWH, along with the road network predicted by the model. Most of the country is characterized by steep terrain and extreme elevation changes, punctuated by deep river gorges, expect for a thin southern strip of flat land on which the EWH runs. The cost-minimizing model generally connects the district headquarters in the north-south direction, instead of the east-west direction.

The road network map predicted by the model is merged with the household location data from NLSS to calculate the straight-line distance of each household to the predicted road network. I use the distance as the crow flies instead of surface distance because the latter is more likely to be correlated with the vertical position of the household, relative

[7] The results are robust to the choice of parameters for the river factor and the choice of gradient function for the vertical factor that are used in the algorithm that construct the predicted road network.

to the base of the mountain, which could directly affect land quality. I also calculate the distances to the nearest river and historical walking trails to use as additional controls in the empirical strategy discussed in the following section.

4 Empirical Strategy

Let V_{pid} be the market value of plot p owned by household i living in district d. T_{id} is the measure of a household's proximity to a paved road.[8] If T_{id} was randomly assigned, the average treatment effect of improved access to transportation could be estimated by the following ordinary least squared specification:

$$log(V_{pid}) = \alpha + \eta_d \ + \gamma \ log(T_{id}) + \Gamma \bar{X}_{id} + \Upsilon \bar{P}_{pid} + \epsilon \tag{3}$$

where η_d is a district fixed effect, $\bar{X}_{id}$ is the vector of household-level controls, and $\bar{P}_{pid}$ is the vector of plot-level controls. These controls include dummies for 12 types of soil, elevation and elevation squared, gradient and gradient squared, elevation interacted with gradient, annual precipitation, average temperature, distance to the nearest river, and an indicator of whether the plot is suitable for rice plantation. I also include 0.25x0.25 degrees latitude-longitude grid dummies to capture any unobserved land and climate characteristics that could vary across these geographical lines.

Alternatively, I instrument for road placement using the road network design predicted by my cost-minimizing model. The instrumental variable approach (IV Model 1) involves estimating the following equation:

$$log(V_{pid}) = \alpha + \eta_d \ + \gamma \ log(\hat{T}_{id}) + \Gamma \bar{X}_{id} + \Upsilon \bar{P}_{pid} + \varepsilon \tag{4}$$

where $log(\hat{T}_{id})$ is an instrumented connectivity measure, estimated on the basis of my model predicting the proximity to actual roads in the first stage:

$$log(T_{id}) = \kappa + \mu_d + \delta \ log(G_{id}) + \Phi \ \bar{X}_{id} + \Psi \ \bar{P}_{pid} + \vartheta \tag{5}$$

[8] I use the distance to the paved road because such blacktop roads are all-season roads unlike unpaved roads that are often closed during monsoon and winter seasons.

G_{id} is the household's straight-line distance to the road predicted by the cost-minimizing model. The empirical strategy requires that G_{id} not be correlated with land quality. In the above specification, I control for observed land quality by including the vector of all the land and climatic covariates listed above, including the latitude-longitude grid dummies. In addition, I also include the distance to walking trails as a control to address the potential correlation between G_{id} and historical trading routes as historical farming practices associated with being near these routes could affect current land quality.[9]

To address the possibility that unobserved land quality could still be correlated with G_{id}, I use two alternative IV estimation strategies. The first strategy utilizes the differences in the timing of project completion across districts, in addition to G_{id}, to jointly predict the proximity to an actual road. The second strategy uses the two-wave household panel data to control for any time-invariant unobserved land quality, and use G_{id} to instrument for the change in the household's road connectivity between the two rounds.

In 2007, just prior to the survey, 23 of the 75 district headquarters were not connected by all-season paved road, although road construction had already began in many "late" districts. If $log(Gid)$ is correlated with unobserved land quality, $log(G_{id})$ is likely to capture this spurious correlation in these "late" districts. The interaction of "early" district dummy and $log(G_{id})$ could therefore be used as an instrument for improved access to roads, while including $log(G_{id})$ separately in the regression to control for correlation between land quality and the predicted road network. In addition, district dummies would control for any endogeneity in the timing of road construction in different districts. This new empirical strategy (IV Model 2) would involve estimating the following 2SLS model:

$$log(V_{pid}) = \alpha' + \eta'_d \ + \gamma' \ log(\hat{T}_{id}) + \theta' \ log(G_{id}) + \Gamma' \bar{X}_{id} + \Upsilon' \bar{P}_{pid} + \varepsilon' \tag{6}$$

where $log(\hat{T}_{id})$ is predicted by $E_d * log(G_{id})$ in the first stage:

$$log(T_{id}) = \kappa' + \mu'_d + \sigma' \ log(G_{id}) + \delta' \ E_d * log(G_{id}) + \Phi' \ \bar{X}_{id} + \Psi' \ \bar{P}_{pid} + \vartheta' \tag{7}$$

[9]Stiller (1976) explains that in Nepal transportation and communication between district headquarters were historically carried out through numerous foot trails that connected any two cities via a shortest distance. According to Shrestha (2001), as of 1984, such trails stretched 10000 kilometers if laid end-to-end, almost twice long the sum of government-built roads; the government made no attempts to incorporate them into its transportation development plan. Banerjee et al. (2012) acknowledge that ignoring this correlation could violate the exclusion restriction of infrastructure instruments.

E_d indicates whether district d is an early district where its district headquarter is connected by paved road. The estimation strategy allows early and late districts to be different in observed and unobserved characteristics. It also allows $log(G_{id})$ to be correlated with land quality. The key assumption is that this correlation is identical in early and late districts.

Next, I use the household panel data from two rounds of NLSS conducted in 1996 and 2010. During this 15-year period, 450 kilometers of new roads were added annually to its network of national highways.[10] The rapid expansion of road network implies that the households located near the predicted road network, as measured by $log(G_{id})$, should have experienced the largest improvement in their distance to the actual road.

Using the two-wave household panel, I calculate the change in these households' distance to the nearest paved road and the change in their total and types of land-holdings, along with the change in the average market price of their owned land. The first-difference approach (IV Model 3) involves aggregating equation (4) at the household-level and rewriting in first differences:

$$\Delta log(V_{id10}) = log(V_{id10}) - log(V_{id96}) = \alpha'' + \gamma'' \Delta log(\hat{T}_{id10}) + \Gamma'' \Delta \bar{X}_{id10} + \varepsilon'' \tag{8}$$

where $\Delta log(\hat{T}_{id10})$ is estimated by $log(G_{id})$ in the first stage:

$$\Delta log(\hat{T}_{id10}) = \kappa'' + \delta'' \; log(G_{id}) + \Phi'' \; \Delta \bar{X}_{id10} + \vartheta'' \tag{9}$$

The empirical strategy controls for all observed and unobserved time-invariant land qualities. The key assumption instead is that the land qualities that initially determine the placement of the predicted road do not change differentially over time. I include controls for household's baseline land holdings and region dummies to control for different trajectories of regional land values, as well as to allow initially different land holdings based on observed land and climate characteristics to evolve differentially over time irrespective of road construction.

[10] The national highways, also called the Strategic Road Network, exclude within-city and within-village roads. In 1996, there was 4,254 kilometers of Strategic Road Network laid out in the entire country. By 2010, it had expanded to slightly more than 10,800 kilometers (Government of Nepal, 2007, 2012).

5 Results

5.1 Effects of road access on agricultural land value

Table 2 presents the OLS and IV estimates of the impact of road on farmland values.[11] In Column 1, I estimate γ from equation (3) using the plot-level cross-sectional data from 2010 NLSS. The results suggest that a 1 percent decrease in distance to a road raises land market prices by 0.081 percent. This OLS estimate is statistically significant at the 1 percent level.[12]

Column 2 estimates δ from IV Model 1 described in equations (4) and (5). The instrument—household's distance to the predicted road network, $log(G_{id})$—is a strong predictor of households' actual distance to the nearest road. The F-statistic on this excluded variable in the first stage regression is 126.20.[13] The IV results suggest that a 1 percent decrease in distance to a road raises land market prices by 0.221 percent. The estimate is statistically significant at the 1 percent level.

Column 3 presents the results from IV Model 2 described in equations (6) and (7), in which the distance to a road is predicted by the interaction of distance to the predicted road network and the district-level indicator variable of project completion. The result in Column 3 suggests that a 1 percent decrease in distance to a road increases land market prices by 0.110 percent. The estimate is statistically significant at the 10 percent level, and the F-statistics on the excluded variable $E_d * log(G_{id})$ in the first stage regression is 97.0.

Lastly, Column 4 uses the household panel data between 1996 and 2010 NLSS rounds, and estimates IV Model 3 described in equations (8) and (9). I report the estimated coefficient of the main regressor of interest: change in log of household's distance to the nearest road. The dependent variable is the change in log of average per-hectare market price of agricultural

[11] In Table 2, Columns 1-3, the errors are clustered at the household level.

[12] Non-parametric estimates of the impact of road on agricultural land values is presented in Appendix Figure A2. It plots the relationship between distance to the nearest paved road and per-hectare market price of an agricultural plot using a difference-based semi-parametric estimation strategy based on Lokshin (2006). The semi-parametric estimation strategy controls for district dummies and area of the plot. The scatter plot of the estimates are depicted in grey color. It suggests a negative and a log-linear relationship between distance to the nearest paved road and market value of land in Nepal. The slope at numerous highlighted points in Appendix Figure A2 decreases exponentially with distance. All the reported slope estimates are negative and statistically different from zero based on bootstrapped clustered errors.

[13] Relevant statistics from the first-stage regression are reported at the bottom of each column. More detailed results from the first-stage regressions and other test statistics are presented in Appendix Table A2.

plots owned by the household. The former variable is predicted in the first stage regression by the household's distance to the predicted road ($log(G_{id})$), and the F-statistics on this excluded variable in the first stage is 21.41. The results suggest a 0.275 percent appreciation in the market price of agricultural land owned by the household for every 1 percent decrease in distance to a road (statistically significant at the 5 percent level).

The empirical findings are robust to using different empirical strategies. Across all four methods (OLS and three IV models), the estimated economic gains from proximity to a road are positive and also similar in magnitude. Moreover, the estimated distance elasticities range between -0.1 to -0.25, and they are inline with the travel-time elasticity estimated by Jacoby (2000) in Nepal.

I use the above elasticity to estimate economic gains for an agrarian household. The distance to the nearest road for a median agrarian household in 2010 was 8 kilometers, and the average per-hectare market price of land owned by these households was $13,475. Based on this information, for a median household, shortening the distance to a road by 1 kilometer is estimated to generate economic benefits of around $185 per hectare of land.

Moreover, between the 1996 and 2010 NLSS, the travel time to the nearest road for a median household decreased by one hour, which roughly translates to a decrease of four kilometers to the nearest road. During this period, 6,580 kilometers of road were added to the country's highway network. If improvement in road accessibility was uniformly distributed across all cultivated land in the country (approximately 30 percent of total land area), the total economic benefit to the agricultural sector from this road expansion is estimated to be about $3 billion. Assuming a road construction cost of $0.2 million per kilometer (Overseas Development Administration, 1997), this roughly translates to a twofold economic return on the country's road infrastructure investments.[14]

[14] Dillon et al. (2011) estimate a benefit-cost ratio of 3.7 for the public expenditure on roads in Nepal. In the above calculation, the investment cost excludes the environmental costs of constructing roads and the maintenance cost. The total investment = $6580*0.2 million. Given that the total area of Nepal is 14.08 million hectares and agricultural land accounted for 29.4 percent of total land in 1996, the present value of total benefits = $4*14.08*0.294*185 million.

5.2 Effects of road access on agrarian households

To shed light on a potential mechanism through which economic benefits from road construction are realized by agrarian households, I extend the empirical strategy to analyze the impact of a household's proximity to a paved road on its decisions related to market participation and agriculture production.

For this purpose, I take 4,989 agrarian households from the 2010 NLSS, and I estimate IV Model 2, using household-level outcomes as the dependent variable. Such outcomes however are likely to be affected by factors such as household's preferences that are independent of transportation costs or land quality. If household preferences are correlated with the instrument $E_d * log(G_{id})$—either due to household migration or differential selection into agriculture, it could bias the IV estimates. I address this concern in two ways.

First, I examine whether $E_d * log(G_{id})$ is correlated with household preferences, by estimating equation (7) with household characteristics as the dependent variable. Table 3 presents the results on various household demographic characteristics, migration history, wealth measures, and characteristics of its agricultural activity such as farm size and land tenure system. For all the reported 18 measures, the estimated coefficient on $E_d * log(G_{id})$, δ', is small and not statistically significant. The results therefore provide no evidence of selective migration or differential selection into agriculture based on observed household characteristics, across areas with varying degree of road infrastructure as predicted by the instrument.

Second, for all the households that ever moved from their original location, I replace their values for $log(G_{id})$ and E_d (along with other location-specific environmental variables) with the values associated with the "farthest" household in their original district. In each district, the farthest household is identified based on the largest value of G_{id} among the sample of non-migrant households. This adjustment is likely to provide a lower bound on the IV estimates as migrant households are often positively selected nearer to the road.

5.2.1 Market participation, commercialization, and agricultural income

Table 4 presents the IV results on household's participation in agricultural markets. Based on Column 1, a 1 percent decrease in distance to a road raises the proportion of agrarian households that visit a market center at least once every month by 6.5 percentage points. The estimate is statistically significant at the 1 percent level, and it suggests a 20.1 percent increase over the mean market participation rate among agrarian households.

Columns 2-3 estimate the impact of a road on commercialization of agriculture. A 1 percent decrease in distance to a road increases the share of produce sold by agrarian households by 2.9 percentage points (Column 2). This signifies a 23.6 percent increase over the mean share, and the estimate is statistically significant at the 5 percent level. The results also suggest that a 1 percent decrease in distance to a road increases the proportion of households that sell their produce by 4.9 percentage points. This impact at the extensive margin is not small, and amounts to a 9.2 percent decrease in subsistence farming over the mean level.

Column 4 estimates the impact on household's income from its agriculture production. Among households selling their produce in a market, a 1 percent decrease in distance to a road increases their total agricultural income by 0.363 percent. The estimate is statistically significant at the 10 percent level. It signifies a $1.08 per hectare increase in farm-based revenue compared to the mean income of $298 per hectare, for every 1 percent improvement in road connectivity.

5.2.2 Agricultural input use and costs, and productivity

Table 5 examines the impact of road on the use of two types of agricultural inputs and farm output. Column 1 shows that a 1 percent decrease in distance to a road increases the proportion of households using chemical fertilizer by 9.1 percentage points, and the IV estimate is statistically significant at 1 percent level. Given that 70 percent of agrarian households use fertilizer, this signifies a 13.1 percent increase in the number households using chemical fertilizers in their farms.

One of the main reasons fertilizer use would increase with road connectivity is because

it improves household's access to agricultural input markets, and lowers the costs of such purchases. I examine this in Column 2 by estimating the impact of proximity to a road on the per unit cost of fertilizer among a subset of households that used chemical fertilizer in their farms. The results show that a 1 percent decrease in distance to a road decreases the "effective" price for the farmer by \$0.029. The estimate is statistically significant at the 5 percent level, and the size of the effect is also not trivial, signifying a 8.2 percent decline over the mean price.

Column 3 and 4 examine the impacts of road on the use and the price of hired laborer in agriculture production. The results suggest a 1 percent decrease in distance to a road increases the share of agrarian households that hire paid laborer by 11.2 percentage points (statistically significant at the 1 percent). In addition, the impact of better road connectivity on the daily wage of laborer is positive, but it is not statistically significant.

Lastly, Column 5 estimates the effect of a road on agricultural output. A 1 percent decrease in distance to a road increases agrarian household's total output per area by 0.092 percent.[15] The impact on agricultural productivity is statistically significant at the 10 percent level.

These results provide a direct evidence of the benefits of road in agriculture. They also document important mechanisms through which economic gains captured by farmland values are, in large part, realized by agrarian households.

6 Conclusion

Transportation infrastructure plays a crucial role in development policy, but little is known about the benefits to agrarian households, let alone the economic returns on such investments in the agriculture sector. I address the endogenous placement of roads by relying on geography that significantly influences the design and the cost of constructing roads in

[15] Tenancy contracts like sharecropping, which are often associated with missing insurance markets, can negatively affect the productivity of agricultural land. In the NLSS data, the incidence of sharecropping is low, accounting for 11 percent of all cultivated agricultural plots, and the results in Table 3 show that the instrument ($E_d * log(G_{id})$) is also not correlated with sharecropping. This implies that the above IV estimate does not pick up the effects of road connectivity on farm production that might come from a move away from sharecropping contracts that may be observed in other settings (insofar as road connectivity improves the access to insurance markets for tenants).

the mountainous terrain of Nepal. I find that decreasing distance to a road by 1 percent increases the market price of an agricultural land by 0.1 percent. Assuming any increase in agricultural profits due to improved access to roads are capitalized in farmland values, the estimated distance elasticity suggests an economic return of roughly twofold from road investments made in the last 15 years.

This is a significant finding, in part because transportation investment is one of Nepal's largest public spending programs. Insofar as households residing farthest from the existing road network are also more economically disadvantaged, infrastructure spending could be one of the most effective means of helping the poor. In addition, it can also have a desirable distributional effect (as highlighted by Jacoby (2000)).

The household-level analyses point to some of the serious bottlenecks in agriculture that include high market transaction costs associated with poor transportation infrastructure. Such constraints could also seriously undermine the implementation, targeting, and long-term effectiveness of other agricultural development programs that rely on market transactions.

The benefits of roads, however, are not limited to agricultural production; they can also improve access to schools, health facilities and other services. The estimated positive impact on agricultural land values may reflect benefits from such amenities insofar as farmers live near their farms. On the other hand, infrastructure-induced in-migration could also lead to congestions of other public services, thereby undermining some of the economic gains (Dinkelman and Schulhofer-Wohl, 2015). Although I do not find any significant in-migration into regions with better roads in context of this study, these wider sets of externalities could be important in determining the net benefits of new infrastructure investments in general.

References

Banerjee, Abhijit, Esther Duflo, and Nancy Qian. 2012. "On the Road: Transportation Infrastructure and Economic Growth in China." NBER Working Paper 17897.

Binswanger, Hans P., Shahidur R. Khandker, and Mark R. Rosenzweig. 1993. "How Infrastructure and Financial Institutions Affect Agricultural Output and Investment in India." *Journal of Development Economics* 41:337–66.

Casaburi, Lorenzo, Rachel Glennerster, and Tavneet Suri. 2013. "Rural Roads and Intermediated Trade: Regression Discountinuity Evidence from Sierra Leone." MIT Sloan Working Paper.

Datt, Gaurav and Martin Ravallion. 1998. "Why Have Some Indian States Done Better Than Others at Reducing Rural Poverty." *Economica* 65:17–38.

de Janvry, Alain, Marcel Fafchamps, and Elisabeth Sadoulet. 1991. "Peasant Household Behaviour with Missing Markets: Some Paradoxes Explained." *Economic Journal.* 101:1400–17.

Dillon, Andrew, Manohar Sharma, and Xiaobo Zhang. 2011. "Estimating the Impact of Rural Investments in Nepal." *Food Policy* 36:250–58.

Dinkelman, Taryn. 2011. "The Effects of Rural Electrification on Employment: New Evidence from South Africa." *American Economic Review* 101:3078–108.

Dinkelman, Taryn and Sam Schulhofer-Wohl. 2015. "Migration, Congestion Externalities, and the Evaluation of Spatial Investments." *Journal of Development Economics* 114:189–202.

Donaldson, David. Forthcoming. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." *American Economic Review.*

Dorosh, Paul, Hyoung Gun Wang, Liangzhi You, and Emily Schmidt. 2012. "Road Connectivity, Population, and Crop Production in Sub-Saharan Africa." *Agricultural Economics* 43:89–103.

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

Emran, M. Shahe and Zhaoyang Hou. 2013. "Access to Markets and Rural Poverty: Evidence from Household Consumption in China." *Review of Economics and Statistics* 95:682–97.

Emran, M. Shahe and Forhad Shilpi. 2012. 'The Extent of the Market and Stages of Agricultural Specialization." *Canadian Journal of Economics* 45:1125–53.

Faber, Benjamin. 2014. "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System." *Review of Economics and Statistics* 81:1046–70.

Fafchamps, Marcel and Ruth Vargas Hill. 2005. "Selling at the Farmgate or Traveling to Market." *American Journal of Agricultural Economics* 87:717–34.

Fafchamps, Marcel and Forhad Shilpi. 2005. "Cities and Specialisation: Evidence from South Asia." *Economic Journal* 115:477–504.

Fafchamps, Marcel and Forhad Shilpi. 2003. "The Spatial Division of Labour in Nepal." *Journal of Development Studies* 39:23–66.

Fujita, Masahisa, Paul Krugman, and Anthony Venables. 1999. *The Spatial Economy: Cities, Regions and International Trade.* MIT Press, Cambridge, Mass.

Gertler, Paul, Marco Gonzalez-Navarro, Tadeja Gracner, and Alexander Rothenberg. 2015. "Road Quality, Local Economic Activity, and Welfare: Evidence from Indonesia's Highways." Working paper.

Gibson, John and Scott Rozelle. 2003. "Poverty and Access to Roads in Papua New Guinea." *Economic Development and Cultural Change* 52:159–85.

Goetz, Stephan. 1992. "A Selectivity Model of Household Food Marketing Behavior in Sub-Saharan Africa." *American Journal of Agricultural Economics* 74:444–52.

Gollin, Douglas. 2010. "Agricultural Productivity and Economic Growth." *Handbook of Agricultural Economics* 4:3825–66.

Gollin, Douglas, Stephen Parente, and Richard Rogerson. 2002. "The Role of Agriculture in Development." *American Economic Review Papers and Proceedings* 92:160–4.

Government of Nepal. 2012. *Statistics of Strategic Road Network 2011-12.* Department of Roads, Ministry of Physical Planning and Works.

Government of Nepal. 2007. *Sector Wide Road Programme and Priority Investment Plan.* Department of Roads, Ministry of Physical Planning and Works.

Government of Nepal. 2002a. *Twenty Year Road Plan.* National Planning Commission.

Government of Nepal. 2002b. *National Transportation Policy.* Ministry of Physical Planning and Works.

Jacoby, Hanan. 2000. "Access to Markets and the Benefits of Rural Roads." *Economic Journal* 110:713–37.

Jacoby, Hanan G. and Bart Minten. 2009. "On Measuring the Benefits of Lower Transport Costs." *Journal of Development Economics* 89:28–38.

Jalan, Jyotsna and Martin Ravallion. 2002. "Geographic Poverty Traps? A Micro Model of Consumption Growth in Rural China." *Journal of Applied Econometrics* 17:329–46.

Kanbur, Ravi and Anthony Venables. 2005. *Spatial Inequality and Development.* Oxford University Press, Oxford, London.

Key, Nigel, Elisabeth Sadoulet, and Alain de Janvry. 2000. "Transactions Costs and Agricultural Household Supply Response." *American Journal of Agricultural Economics* 82:245–59.

Khandker, Shahidur R., Zaid Bakht, and Gayatri B. Koolwal. 2009. "The Poverty Impact of Rural Roads: Evidence from Bangladesh." *Economic Development and Cultural Change* 57:685–722.

Lipscomb, Molly, Mushfiq Mobarak, and Tania Barkam. 2013. "Development Effects of Electrification: Evidence from the Topographic Placement of Hydropower Plants in Brazil." *American Economic Journal: Applied Economics* 5:200–31.

Lipton, Michael and Martin Ravallion. 1995. "Poverty and Policy." *Handbook of Development Economics* 3:2467-3047.

Lokshin, Michael. 2006. "Difference-based Semiparametric Estimation of Partial Linear Regression Models." *The Stata Journal.*

Lokshin, Michael and Ruslan Yemtsov. 2005. "Has Rural Infrastructure Rehabilitation in Georgia Helped the Poor?" *World Bank Economic Review* 19:311–33.

Meyer, Walter. 2008. *Roads to Prosperity Partnership Results- Asia Brief.* South Asia Division, SDC Swiss Agency for Development and Cooperation.

Michaels, Guy. 2008. "The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System." *Review of Economics and Statistics* 90:683–701.

Minot, Nicholas and Ruth Vargas Hill. 2007. "Developing and Connecting Markets for Poor Farmers." *2020 Focus Brief.*

Minten, Bart and Steven Kyle. 1999. "The Effect of Distance and Road Quality on Food Collection, Marketing Margins, and Traders' Wages: Evidence from the Former Zaire." *Journal of Development Economics* 60:467–95.

Overseas Development Administration. 1997. *Principles of Low Cost Road Engineering in Mountainous Regions with Special Reference to the Nepal Himalaya.* Transport Research Laboratory, UK Overseas Development Administration, London.

Renkow, Mitch, Daniel G. Hallstrom, and Daniel D. Karanja. 2004. "Rural Infrastructure, Transactions Costs and Market Participation in Kenya." *Journal of Development Economics* 73:349–67.

Shrestha, Hare Ram. 2010. "Road Versus Hill Environment: The Trend of Road Construction in Nepal." Transport in Mountains International Workshop, Nepal.

Shrestha, Nanda. 2001. *The Political Economy of Land, Landlessness and Migration in Nepal.* Nirala Series, Kathmandu.

Stiller, L. F. 1976. *The Silent Cry.* Sahayogi Prakashan, Kathmandu.

Strauss, John. 1984. "Marketed Surplus of Agricultural Households in Sierra Leone." *American Journal of Agricultural Economics.* 66:321–31.

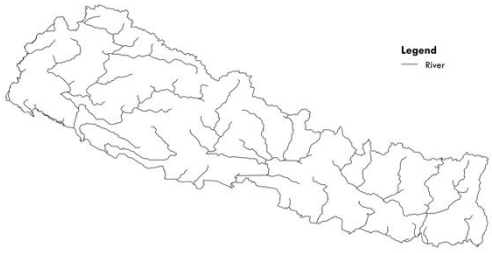
van de Walle, Dominique. 2002. "Choosing Rural Road Investments to Help Reduce Poverty." *World Development* 30:575–89.

World Bank. 2013. *World Bank Annual Report Lending Data and Organizational Information.* World Bank, Washington DC.

World Bank. 2012. *Nepal - Bridges Improvement and Maintenance Program Project.* South Asia Regional Office, Washington DC.

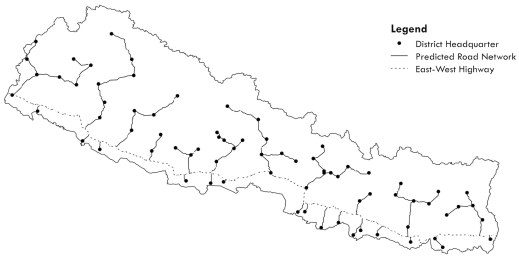
Figure 1: Road Network Prediction from the Cost Minimizing Model

(a) Input: Land Gradient



(b) Input: River Channels

(c) Input: District HQs and EWH



(d) Output: Predicted Road Network

Note: The algorithm connects the district headquarters and the EWH, shown in sub-figure (c), to form a single, integrated national road network. The variables that determine the design of this network are land gradient, river channels, and surface distance. Sub-figure (d) presents the design of the road network predicted by this cost-minimizing algorithm.

Table 1: Descriptive Statistics

	Mean	Median	Std. dev.
Panel A: Plot characteristics			
Number of plots	15717	-	-
Plot size *(Ha.)*	0.200	0.102	0.322
Market value of plot *($ per ha.)*	92668	12088	2805609
Suitable for rice *(%)*	49.71	-	50.00
Year round irrigation *(%)*	21.93	-	41.61
Self-cultivation *(%)*	86.81	-	34.10
Rent-out for cultivation *(%)*	7.72	-	26.72
Rented-out under sharecropping *(%)*	6.84	-	25.25
Land title[a] *(%)*	96.68	-	17.91
Panel B: Household characteristics			
Number of agrarian households	4928	-	-
Total land owned *(Ha.)*	0.636	0.406	0.979
Total livestock *($)*	1175	1020	974.6
Durable asset index	-0.483	-0.784	1.369
Housing quality index	0.355	0.224	1.054
Migrant household *(% ever moved)*	11.4	-	18.1
Household size	4.05	3.90	1.886
Household head education*(years)*	3.32	0.00	4.205
Household head female *(%)*	24.79	-	43.20
Household head age	47.84	47.00	13.77
Household head married *(%)*	87.80	-	32.8
Total area cultivated *(Ha.)*	0.616	0.401	0.864
Share of area rented-in *(%)*	17.90	0.00	0.316
Share of area sharecropped *(%)*	11.00	0.00	0.257
Used fertilizer *(%)*	71.20	-	45.29
Cost of fertilizer *($ per kg.)*	0.355	0.346	0.152
Hired casual labor *(%)*	48.00	-	49.96
Wage paid to hired labor *($ per day)*	2.81	2.46	2.78
Total agricultural output *($ per ha.)*	658.73	373.13	1387.64
Share of output sold *(%)*	12.30	0.00	19.98
Subsistence household *(%)*	53.30	-	49.89
Total agricultural income *($ per ha.)*	297.97	113.02	1587.89
Monthly visits to the market *(%)*	30.58	-	46.08
Distance to the nearest paved road *(Km.)*	21.39	8.00	35.12

Notes: The plot and household data come from 2010 Nepal Living Standards Survey (NLSS). Agrarian households make up 69.55 percent of the entire NLSS sample. [a] *Land title* variable is constructed from the 1996 NLSS data for a subset of plots (1508 plots) owned by the panel households in 1996.

Table 2: Impact of a Road on Market Value of an Agricultural Land

	OLS	IV		
		Model 1	Model 2	Model 3
	(1)	(2)	(3)	(4)
Log distance to road	-0.081***	-0.221***	-0.110*	
	(0.0072)	(0.0356)	(0.0601)	
Δ Log distance to road				-0.275**
				(0.1238)
Instrument		$log(G_{id})$	$E_d * log(G_{id})$	$log(G_{id})$
First-stage statistics:				
R-Squared		0.586	0.592	0.446
F-Statistics		126.6	97.26	21.41
Probability > F		0.000	0.000	0.000
Dependent variable	Log land value	Log land value		Δ Log land value
Level of observation	Plot level	Plot level		Household level
Data source	2010 cross-section	2010 cross-section		1996 & 2010 panel
No. of observations	15717	15717	15717	350

Notes: The table reports the estimated impact of road on agricultural land market value using the OLS and IV strategies. In Columns 1-3, the land and climatic controls include dummies for 12 types of soil, elevation and elevation squared, gradient and gradient squared, elevation interacted with gradient, annual precipitation, average temperature, distance to the nearest river, and 0.25x0.25 degrees latitude-longitude grid dummies. The specification in Column 4 controls for regional dummies, 1996 land holdings, the change in land holdings between the two rounds, and land and climatic variables. Standard errors are adjusted for within-household correlation between plots and are reported in parentheses; *p<0.1, **p<0.05, ***p<0.01.

Table 3: Relationship between Household Characteristics and $E_d * log(G_{id})$

	Coef.	Std. error	t-value
Dependent variable:			
Household demographics			
Household size	0.020	(0.1150)	0.177
Household head education	-0.421	(0.2830)	-1.491
Household head female	0.017	(0.0316)	0.553
Household head age	-0.304	(0.9470)	-0.321
Household head married	0.029	(0.0236)	1.236
Household wealth			
Total land owned	-0.002	(0.0525)	-0.031
Total value of livestock	38.44	(61.590)	0.624
Durable asset index	-0.081	(0.0669)	-1.212
Housing quality index	-0.093	(0.0708)	-1.312
Migration history			
Ever moved	0.033	(0.0227)	1.452
Moved since 1996	0.027	(0.0195)	1.396
Activity in land market			
Bought or sold land	-0.016	(0.0198)	-0.802
% owned land bought or sold	-0.070	(0.0844)	-0.835
Agricultural enterprise			
Total area cultivated	-0.012	(0.0516)	-0.229
Share of area rented-in	-0.009	(0.0140)	-1.105
Any rented-in area	-0.009	(0.0140)	-1.358
Share of area under sharecropping	-0.009	(0.0140)	-0.481
Any sharecropped area	-0.009	(0.0140)	-0.604

Notes: The table reports the estimated value of δ', the coefficient on $E_d * log(G_{id})$, from equation (7) using household characteristics as dependent variables. The sample includes 4928 agrarian households from the 2010 NLSS data. Additional controls include district dummies, distance to the predicted road ($log(G_{id})$), dummies for 12 types of soil, elevation and elevation squared, gradient and gradient squared, elevation interacted with gradient, annual precipitation, average temperature, distance to the nearest river, and 0.25x0.25 degrees latitude-longitude grid dummies. Standard errors are reported in parentheses; *p<0.1,**p<0.05, ***p<0.01.

Table 4: Impact of a Road on Household Market Participation

	Monthly market visits (1)	% output sold (2)	Any output sold (3)	Agriculture income[a] (4)
Log distance to road	-0.065**	-0.029**	-0.049	-0.363*
	(0.0322)	(0.0120)	(0.0364)	(0.2040)
First-stage statistics				
R-Squared	0.601	0.601	0.601	0.542
F-Statistics	51.86	51.86	51.86	15.13
Probability > F	0.000	0.000	0.000	0.000
No. of observations	4928	4928	4928	2288
Mean of dep. var. (in levels)	0.312	0.123	0.467	298

Notes: The table reports the estimated value of γ' from IV Model 2 described in equations (6) and (7), using household outcomes as dependent variables. The sample includes all agrarian households from the 2010 NLSS data. Additional controls include original district dummies, distance to the predicted road ($log(G_{id})$, dummies for 12 types of soil, elevation and elevation squared, gradient and gradient squared, elevation interacted with gradient, annual precipitation, average temperature, distance to the nearest river, and 0.25x0.25 degrees latitude-longitude grid dummies. For households who have moved from their original location, their values for $log(G_{id})$ and E_d (along with other location-specific land and climatic variables) with the values associated with the non-migrant household that resides farthest from the predicted road network in their original district. [a] *Agriculture income* is the log of total farm revenue per hectare of land cultivated by the household. Standard errors are reported in parentheses; *p<0.1,**p<0.05, ***p<0.01.

Table 5: Impact of a Road on Household Agriculture Production

	Fertilizer use (1)	Cost of fertilizer (2)	Hired labor (3)	Daily wage (4)	Agriculture output[a] (5)
Log distance to road	-0.091***	0.029**	-0.112***	-0.467	-0.092*
	(0.0345)	(0.0142)	(0.0379)	(0.3060)	(0.0557)
First-stage Statistics					
R-Squared	0.601	0.575	0.601	0.564	0.601
F-Statistics	51.86	25.95	51.86	33.38	51.86
Probability > F	0.000	0.000	0.000	0.000	0.000
No. of observations	4928	3500	4928	2352	4928
Mean of dep. var. (in levels)	0.786	0.355	0.480	0.346	658.73

Notes: The table reports the estimated value of γ' from IV Model 2 described in equations (6) and (7), using household outcomes as dependent variables. The samples in Columns 1, 3, and 4 include all agrarian households from the 2010 NLSS data. In Columns 2 and 4, the sample includes a subset of households that have used chemical fertilizer, and that have hired casual laborer respectively. The specification in all columns controls for original district dummies, distance to the predicted road ($log(G_{id})$, dummies for 12 types of soil, elevation and elevation squared, gradient and gradient squared, elevation interacted with gradient, annual precipitation, average temperature, distance to the nearest river, and 0.25x0.25 degrees latitude-longitude grid dummies. For households who have moved from their original location, their values for $log(G_{id})$ and E_d (along with other location-specific environmental variables) with the values associated with the original household that resides farthest from the predicted road network in their original district. [a] *Agriculture output* is the log of total value of output per hectare of land cultivated by the household. The value of output is calculated by multiplying each agricultural good with its corresponding median national market price in 2010 and then aggregating it at the household-level. Standard errors are reported in parentheses; *p<0.1,**p<0.05, ***p<0.01.

Table A1: Relationship between Plot Values and Rents

	(1)	(2)	(3)
$log(r)$	0.816***	0.851***	0.762***
	(0.1235)	(0.1439)	(0.1116)
P-value of F-test:			
H_o: Coefficient of $log(r)$=1	0.137	0.301	0.448
First-stage statistics:			
F-statistics	22.48	24.79	14.59
Probability > F	0.000	0.000	0.000
Controls:			
District FE	No	Yes	No
Household FE	No	No	Yes
No. of plots	1162	1162	739

Notes: This table reports the coefficients of the log rent ($log(r)$) from equation (1), where the dependent variable is the log of plot market value, $log(V)$. The former variable is instrumented by plot area. Standard errors are reported in parentheses; *p<0.1,**p<0.05, ***p<0.01.

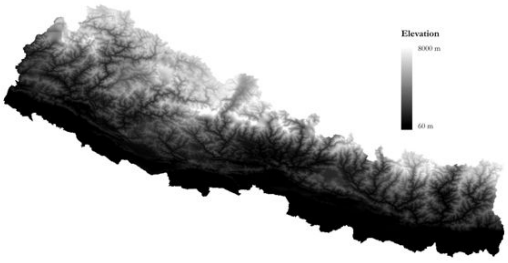
Table A2: First-stage Regressions and IV Tests

	(1)	(2)	(3)
$Log(G_{id})$	0.842***		1.21***
	(0.0748)		(0.2614)
$E_d * Log(G_{id})$		1.25***	
		(0.1270)	
Dependent variable	$Log(T_{id})$		$\Delta Log(T_{id})$
Level of observation	Plot level		Household level
Data source	2010 cross-section		1996 & 2010 panel
Test of endogeneity :			
P-value (H_o: Exogenous)	0.000	0.566	0.017
Weak instrument robust inference:			
P-value of Anderson-Rubin test (H_o: IV estimate=0)	0.000	0.078	0.011
R-Squared	0.586	0.592	0.446
No. of observations	15718	15718	350

Notes: The table reports the first-stage results for IV Models 1-3 presented in Table 2. Statistics from the Hausman test and the Anderson-Rubin test of weak instrument robust inference for the three IV Models in Table 2, Columns 2, 3, and 4 are presented at the bottom of Columns 1, 2, and 3 respectively. Standard errors are adjusted for within-household correlation between plots and are reported in parentheses; *p<0.1, **p<0.05, ***p<0.01.

Figure A1: Spatial Distribution of Land and Climatic Variables

(e) Elevation

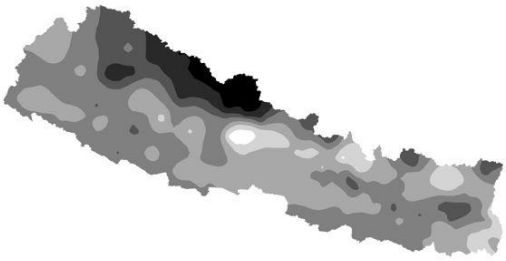


Mean: 885.30, median: 872, s.d. 711.69.
Source: Global Climate Database, WorldClim.

(f) Temperature

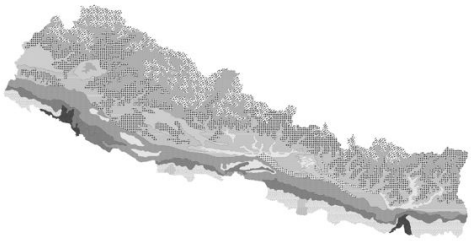
Mean: 20.55, median: 20.8, s.d. 3.69.
Source: Global Climate Database, WorldClim.

(g) Precipitation



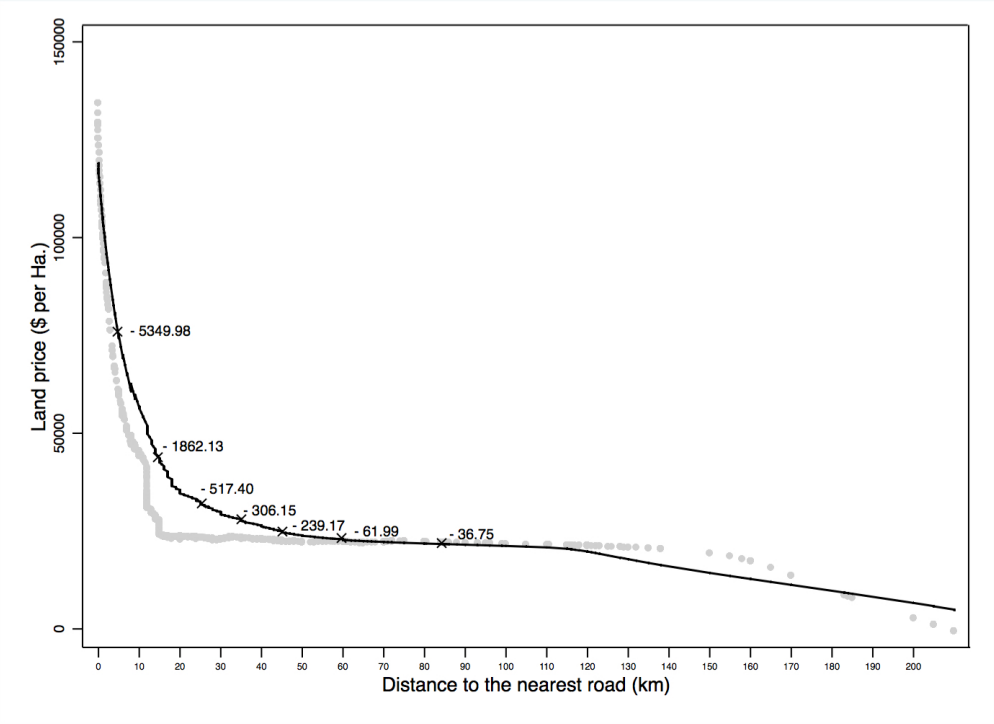
Mean: 1898.49, median: 1800, s.d. 601.62.
Source: Mountain Environment and Natural Resource Information System Database, International Center for integrated Mountain Development.

(h) Soil type



Notes: 12 categories of soil type.
Source: Mountain Environment and Natural Resource Information System Database, International Center for integrated Mountain Development.

Figure A2: Relationship between a Road and Market Value of an Agricultural Land



Note: The figure presents the scatter plot of the the non-parametric relationship between the nearest road and per-hectare market price of an agricultural plot predicted using a first-difference estimation strategy developed by Lokshin (2006). It controls for district dummies and area of the plot. The locally weighted scatter plot smoothing (LOWESS) is depicted in a solid line. The LOWESS uses a bandwidth of 0.8. The values reported in the figure represent the slope estimated at the highlighted points on the LOWESS curve.