Road to Specialization in Agricultural Production:

Evidence from Rural China

Yu Qin yq38@cornell.edu

The Dyson School of Applied Economics and Management

Cornell University, Ithaca, NY, 14850

Xiaobo Zhang X.ZHANG@cgiar.org

International Food Policy Research Institute

2033 K St. NW Washington, DC, 20006-1002

Selected Paper prepared for presentation at the International Association of Agricultural Economists (IAAE) Triennial Conference, Foz do Iguacu, Brazil, 18-24 August, 2012.

Abstract:

We used a primary panel survey at the household level conducted in 18 remote natural villages over three waves in China to study how road access shapes farmers' agricultural production patterns and input uses. Our results show that access to roads is strongly associated with specialization in agricultural production. In natural villages with better road access, farmers plant fewer numbers of crops, purchase more fertilizer, and invest more money on labor hiring. In combination of these factors, road connections improve household agricultural income, and in particular cash income. However, better access to rural roads does not appear to bring about significant changes in non-agricultural income.

To get rich, build road first.

------*An old Chinese proverb*

Introduction

In developing countries, the rural poor often live in isolated areas. Being far away from the market, the poor are more likely to rely on self-sufficient subsistence farming to survive. Spatial poverty traps are a silent feature in rural landscape (Jalan and Ravallion, 2002). Rural roads have been argued as one key instrument to overcome the spatial poverty traps in developing countries (Calderón 2009; Escobal and Ponce 2002; Fan and Hazell 2001; Jacoby and Minten 2008). However, because rural roads are often remote and hard to reach, it may be costly to build. Therefore, rigorous impact assessments on rural roads in lagging areas are critically needed for policy interventions.

There have been a limited number of studies evaluating the returns to road investment in developing countries, but many of them are conducted at the aggregate level (Fan and Hazell, 2011; Zhang and Fan, 2004). However, the studies at the aggregate level have been criticized for failing to uncover the mechanisms driving how road connections shape household production and consumption behavior (Jacoby, 2000). The studies at the household level, on the other hand, often rely on cross-sectional data due to difficulties in obtaining long-term time series data in poor areas. However, the issue of road placement endogeneity complicates such a reliance on cross-sectional data. To overcome this problem of endogeneity, Jacoby (2000) develops an innovative approach to evaluate the impact of road access on agricultural land values, which uses the discounted stream of maximal profits from cultivation to measure the value of land. Yet, because there are a large number of landless laborers in many developing countries, the approach is not adequate to evaluate the impact on overall rural welfare. In the context of China, the method is also inapplicable because farmers hold only the right for land cultivation but not the ownership. In the absence of agricultural land markets, it is impossible to uncover the true farmland values.

In this paper, we make use of a primary collected panel household dataset in a remote and poor area in China to investigate the impact of road connections on rural welfare through the lenses of agricultural specialization and input use. Road connections reshape the production choice setof farmers in isolated areas and can potentially affect agricultural production—the major livelihood of the poor—in at least two ways.

First, with lower transportation costs, farmers may shift their agricultural production from an autarky-based self-sufficient subsistence farming to a more market-oriented specialized production (Limao and Venables 2001; Renkow et al. 2003). Yang and Ng (1993) develop a theoretical model, showing that producers will choose to specialize in an activity according to their comparative advantage and simply purchase other goods and services from the market, provided that transaction costs are sufficiently small. Using a simulation approach, Omamo (1998) finds that as distance shortens, small-scale farm tends to shift away from diversified cropping patterns in favor of cultivating only one crop. However, the empirical findings are mixed. For example, Stifel et al. (2003) show that in Madagascar the concentration level of agricultural production in the least remote areas is around 1.5 times that of the most remote areas, suggesting that improved road access facilitates specialization in agricultural production. However, Gibson and Rozelle (2003) find that in Papua New Guinea, each extra hour to reach the nearest road induces a 2.6 percent reduction in the number of activities, in contrary to the theoretical prediction.

Second, as improved road access reduces transportation costs, the prices of modern inputs such as fertilizer are more likely to drop (Khandker et al. 2006). Consequently, farmers may apply more modern inputs to improve agricultural productivity. At the same time, farmers may hire more labor to take care of the land as agricultural production becomes more specialized as road access improves. Gollin and Rogerson (2010) develop a theoretical model and calibrate it with Uganda data, showing that as transportation cost decreases, farmers will use more intermediate inputs, which in turn contribute to agricultural output growth. The empirical findings on the impact of rural roads on modern input use, however, are inconclusive. For example, Benziger (1996) finds that better road access leads to increasing fertilizer use in villages in Hebei, China. In the context of Madagascar, Stifel et al. (2003) show that farmers in more isolated regions use less fertilizer than those in places with better road access. However, Dorosh et al. (2010) paint a more complicated story: Input use depends on not only distance to roads but also the density of road networks. For example, in East Africa, reducing travel time significantly increases adoption of high-input/high-yield technology, while the impacts of roads are insignificant in West Africa, where the density of road networks is relatively higher at the beginning.

One challenge to empirically evaluate the impact of road access on agricultural production is limited data. Most of the empirical studies rely on cross-sectional data, making it hard to control for unobserved factors. In this paper, we use a primary household panel dataset collected throughout 18 natural villages over three waves in Guizhou province, China, to study how road access shapes farmers' cropping patterns and input use.

Our dataset possesses two advantages in studying the impact of access to road networks in isolated villages. First, as a non-recall panel data, our study provides relatively accurate and credible information with respect to household agricultural production. Second, the three waves of data allow us to conduct a difference-in-difference analysis, which helps to mitigate estimation biasesas a result of omitting variables and reverse causality commonly seen in regressions based on cross-sectional datasets. To the best of our knowledge, this is the first paper to empirically document the causal impact of road access on agricultural specialization and input use in China.

We find that access to roads fosters household agricultural specialization. The impact is economically significant and is about one-fourth of the standard deviation of the standard Herfindahl-Hirschman specialization index. In addition, better road connectivity induced farmers to apply more fertilizer. Thanks to these two channels, road access is shown to boost farmers' agricultural income. However, the introduction of road access does not seem to improve farmers' non-agricultural income in this remote area.

The findings may have some policy implications for China as well as other developing countries. In the past several decades, the Chinese government has made significant investments to build a nationwide highway system. As the highway density increases, the marginal returns to highway investment are likely to decrease. Fan and Chan-Kang (2005) argue that it may make more economic sense to gear investment towards rural roads. However, rural roads carry less traffic and are harder to maintain, often making it more costly to build rural roads in remote areas. Therefore, it is important to gather more evidence as to how rural roads affect agricultural patterns and rural livelihoods in lagging regions.

One should be cautious in explaining the findings. Our sample focuses only on the mountainous rural areas in southwestern China, where smallholder farming is the dominant mode of agricultural production. As China is a large and spatially diverse country, the findings drawn from this sample may be not applicable to China as a whole.

Despite this limitation, our study may shed some light on other developing countries. Many parts of rural Africa have limited road access, mirroring the survey area in Guizhou province. Governments and donors face the similar tradeoff of whether to build roads where people live or invest in areas where jobs are located. The public investment strategy largely hinges upon the precise estimates of benefit–cost ratios among different types of investments, including social and productive investments (for example, social safety nets and rural roads). However, due to lack of data covering multiple time periods, rigorous impact evaluations on rural roads in Africa are lacking. Our study has documented the mechanism as to how road connections might affect farming practices and rural livelihoods in isolated and impoverished regions.

Description of Data

As shown in Figure 1, Guizhou is located in the southwest area of China. As one of the poorest provinces, Guizhou has the least road length per capita due in part to its mountainous terrain. Figure 2 depicts the road system in Guizhou as of 2004. Highway networks are sparse in Guizhou, with only four roads reaching from the provincial capital (Guiyang Shi) to major big cities in the province. Although there are more numbers of national roads and province roads, the density is much lower than the national mean. In the mountainous villages, many households in practice subsistence farming, whereas households in relatively flatter areas are much better off due to better road connections. The large variation in road access in our sample thus provides us with a valuable opportunity to study the road impact on agricultural production in China.

The survey site is in Puding County, consisting of 11 townships, 317 administrative villages, and a total population of 448,000 people by the end of 2008.[1] A highway and a national road bypass the county border and one provincial road cuts through the county. In 2008 the average household income in Puding County was around 5,800 *yuan*, which was slightly above the provincial median but below the provincial mean. As depicted in Figure 3, in terms of per capita rural income, Puding is in the middle tercile, suggesting Puding is a rather representative county in Guizhou Province.

Three administrative villages that represent different levels of economic development of Puding were chosen in the survey. There are 18 natural villages in the three administrative villages. A census-type survey of all 805 households in all the natural villages was first administered in early 2005. The second survey wave was conducted in early 2007 and included 833 households. The third survey wave was undertaken in early 2010 and included

[1] An "administrative village" is a bureaucratic entity comprised of several "natural villages" (hamlets). A typical natural village includes about 30-50 households. It is too small to form an administrative unit. As a result, some nearby natural villages are artificially put together to create an administrative village. However, in the mountainous area, sometimes it takes a few hours to walk from one natural village to another one within the same administrative village.

873 households.[2] The surveys collected detailed information on household characteristics, demographics, income, agricultural production, and consumption.

The natural villages exhibit large variations in the degree of road access. We define a road being accessible if tractors can drive through during the rainy season. Table 1 summarizes road access in the 18 natural villages in 2004, 2006, and 2009. The information was collected from the records of village offices. As shown in Table 1, Administrative Village III, which is right next to the county seat, has the best road access of the three administrative villages. All the natural villages of Administrative Village III had already had road access prior to the first wave of the survey. Four natural villages in Administrative Village I constructed roads during our survey periods. However, until our most recent survey, some natural villages, such as Natural Village 1 and Natural Village 3, still had yet to gain road access. In Administrative Village II, one natural village built a new road during 2004 and 2006, while two other natural villages still lacked road access at the time of our most recent survey.

As we are interested in the impact of road access on agricultural specialization, we constructed a Herfindahl-Hirschman index to measure varying degrees of specialization at the household level. The Herfindahl-Hirschman index is defined as the sum of squares of agricultural income shares derived from different production activities.[3] The specialization index ranges between 0 and 1. The greater the value, the higher is the degree of specialization.

[2]There are 782, 815 and 834valid observations households in the three waves, respectively.

[3]For example, if the household produces maize and fruit, with an income of 2,000 *yuan* and 3,000 *yuan*, then the specialization index is calculated as $(2000/5000)^2+(3000/5000)^2=0.52$.

Table 2 reports income sources from several major agricultural activities. Maize is the predominant crop, generating the largest share of agricultural income, ranging from 39 to 46 percent in the three survey years. As the second most important crop, rapeseed provides 16 to 21 percent of household agricultural income. Livestock rank third in terms of agricultural income generation. In our sample, only approximately 27-30 percent of households were engaged in livestock production, comparing to an approximate 90 percent participation rate in maize and rapeseed production.

It is worth noting that the categories of agricultural income decomposition are slightly different in the three waves of surveys due to change in questionnaire design. For example, in the surveys of 2004 and 2009, there are nine sub-categories of agricultural income, while there are ten sub-categories in the 2006 survey. Additionally, there is no data available for vegetable income in 2009. To address these problems, we construct two alternative measures of the specialization indices as robustness checks. In the first alternative measurement, we impute the vegetable income in2009 based on the vegetable seed cost available in 2009 and the past relationship between vegetable seed cost and income observed in the first two waves of survey (variable denoted as HHI (2)). In so doing, we obtain comparable household vegetable income for all the three waves. For the second alternative measurement, we reclassify the non-overlapping subcategories and the rest of the income as "other." After the adjustment, there are nine comparable subcategories of agricultural income across the three waves (variable denoted as HHI (3)).

Table 3 presents the summary statistics for the key variables used in the analysis. Average household income almost doubled from 6,246 *yuan* in 2004 to 11,996*yuan* in 2009. Income generated from non-agricultural activities played a key role in overall income growth. Nonfarm income grew from 2,267 *yuan* in 2004 to 6,541*yuan* in 2009. By comparison, average household agricultural income grew at a slower pace, from 3,978 *yuan* to 5,454 *yuan*, during the five-year period. The lackluster performance in the agricultural sector is not surprising given limited arable land in this area. After all, Guizhou ranks among the provinces with least amount of arable land in China. On average each person in our survey village cultivates only 0.74 *mu*[4] of land in 2009, only about half of the national average of 1.4 *mu* per capita.

Lastly, the mean level of household agricultural specialization index is 0.46, 0.41, and 0.49 in 2004, 2006, and 2009, respectively. The drop in specialization index in 2006 is perhaps due to the severe drought in that year. In 2006, the share of corn income dropped to 39 percent, lower than that of 2004 (46 percent) and 2009 (42 percent). The drought may thus explain the dip in the trend of HH index.

The summary statistics reveal stark differences between households with and without road access. As shown in Table 4, mean household income in villages with road access is almost double that of villages without road access. Both the agricultural and non-agricultural incomes per capita in households with road access are higher than their counterparts without roads. In terms of agricultural production, the villages with roads were more specialized than those without connecting access to roads. In general, households with road access tend to have smaller household size, larger areas of cultivated land, higher levels of education, and slightly lower numbers of primary age laborers.

[4] 1 *mu*=0.066667 hectare.

Empirical Model

Our empirical question is the following: does road access have any impact on extent of specialization and input use in agricultural production? In this paper, we adopt a difference-in-difference method to answer this question. The specification is as follows:

$$Y_{i,t} = \alpha_0 + \beta_1 Road_i + \beta_2 Road_i * Beforeafter_{i,t} + Z_{i,t} + \varphi_{village} + \psi_{year} + \varepsilon_{i,t}, \quad (1)$$

where $Y_{i,t}$ is a dependent variable for household i in time t; $Road_i$ denotes whether the village to which household i belongs has a road by the end of our last wave of survey (year 2009); $Beforeafter_{i,t}$ denotes whether the village household i belongs to has road access or not in year t; $Z_{i,t}$ represents a series of control variables, including cultivated land area, number of primary age labor force aging between 16 to 60 in the family, household size, the highest year of education within the household, and whether there is village cadre in the household; $\varphi_{village}$ stands for natural village fixed effects; ψ_{year} controls for year fixed effects; $\varepsilon_{i,t}$ is the error term.

Our coefficient of interest is β_2, which represents the impact of road access on the outcome variables. The main dependent variables in our estimation are: (i) the household agricultural specialization index (HHI (1), HHI (2) and HHI (3)); (ii) fertilizer use measured by the natural log of the monetary value of fertilizer use per *mu* of land; (iii) logged expenditures on hired labor; (iv) the natural log of agricultural income, non-agricultural income, and total income per capita in the household; and (v) logged agricultural cash income as a measure of agricultural product marketization. If road access significantly leads to agricultural specialization, we expect β_2 to be positive and significant. Similarly, β_2 is expected to be positive and significant as well if the outcome variable is either fertilizer use, the cost of hired labor, or household income.

Because we have a panel dataset, we can remedy many of the problems plaguing cross-sectional analyses. For instance, we can include household characteristics, natural village fixed effects, and year fixed effects to largely control for omitted variable bias. Since the cropping and input use decisions largely depend upon the existing road conditions, reverse causality from specialization and input use on road placement is unlikely. It is hard to imagine farmers would change their cropping patterns in anticipation of a new road in the next several years. Perhaps the biggest challenge is road placement. As suggested by recent impact evaluation literature (Duflo and Pande, 2007), the non-random program placement may bring about endogeneity problems in economic estimations. A typical solution is to carry out a Two-stage Least Square analysis by instrumenting the policy with a set of exogenous variables. However, the road variable varies at the natural village level and there are only 18 natural villages in the dataset, making it impossible to implement the first stage regression with such a small number of observations. Since the objective of our paper is to examine how households respond to road connections in their production decisions, the potential endogeneity problem of road placement, if any, is minimal.

Empirical results

Table 5 reports the main regression results on specialization, fertilizer use, and cost of hired labor. Natural village fixed effects and year fixed effects are included in all the regressions to control for village-specific factors, such as village growth potential, and common temporal trends such as investment policy. The first column under each heading lists the most parsimonious specification, while household characteristics are added in the second column.

For specialization, we use three indexes: HHI (1), HHI (2), and HHI (3). Regardless of the two different specifications, road access is shown to have positive and significant impact on agricultural specialization. The results are robust to three slightly different specialization indices. On average, road access will positively affect the specialization index by four percentage points, which is approximately one fourth of one standard deviation of HH indexes.

As shown in Table 5, better road connections also induce farmers to apply more fertilizer. After improvements in road connections, fertilizer use (*yuan* per *mu*) rose by 29.7 percent (after translating the log form coefficient 0.26 into a real growth rate). Similarly, road access boosted households' expenditure on hired labor by 37.7 percent (after translating the log form coefficient 0.32 into a real growth rate).

As farmers specialize in their agricultural production, apply greater amounts of modern inputs, and hire more skilled professional workers, we expect their agricultural income to increase as well. Table 6 summarizes the regressions on household income, including agricultural income, non-agricultural income and total income per capita. As expected, road access enhances agricultural income by 29.7 percent (after translating the log form coefficient 0.26 into a real growth rate), which is significant at the 0.05 level. However, roads do not appear to play a major role in shaping non-agricultural household income. In this area, most young people migrate outside of the province to work in the nonfarm sector. Road conditions are not a significant factor in their decision to migrate. Overall, the impact of roads on total income is positive but not significant.

As a robustness check to the results on agricultural income, we conduct further regressions on agricultural cash income. As road access facilitates agricultural specialization, which turns the village away from autarky and toward a more integrated economy, farmers are likely to sell more surplus agricultural product to the market, thereby earning a larger share of cash income. As shown in Columns 7-8 in Table 6,the coefficient for the road variable is 0.32 (significant at the 0.05 level). This means that road access would lift agricultural cash income by 37.7%.

Robustness Checks

As suggested by Cameron et al. (2008), small numbers (5-30) of clusters may lead to over-rejection of the null hypothesis. As we only have 51 clusters (natural village by year) in the dataset, it is important to check if our results are robust to the alternative calculation of standard errors. Cameron et al. (2008) suggest using cluster bootstrap to refine conventional clustered standard errors in the presence of small number of clusters.

The cluster bootstrap procedure is as follows. Do B iterations of this step. For the b^{th} iteration: 1) Form a sample of G clusters of both dependent and independent variables by resampling with replacement G times from the original sample. 2) Calculate the Wald test statistic $w_b^* = \frac{\beta_{1,b}^* - \beta_1}{s_{\beta_{1,b}^*}}$; where $\beta_{1,b}^*$ is the coefficient from OLS estimation using the b^{th}

pseudo-sample, with $s_{\beta^*_{1,b}}$ as its standard error. β_1 is the OLS estimate from the original sample. 3) Use the empirical distribution of w^*_1, w^*_2, $\ldots w^*_B$ to determine critical values and p-values. In our case, G=51 (17 natural villages * 3 survey years) and B=1000.

Table 7 and Table 8 display the estimation using cluster bootstrap. As we expected, the significance level decreases for several coefficients, including fertilizer use, hired labor cost, agricultural income, and agricultural cash income. But the significance level only disappears for one coefficient in the regression on agricultural income at the 10 percent level. Therefore, our estimation is robust given small number of clusters.

Conclusions

In this paper, through the use of primary census-type household surveys in remote villages in China, we examine the impact of road access on agricultural production, and particularly on specialization and intermediate input use. We find that better access to roads facilitates agricultural specialization, induces farmers to use more fertilizer, and prompts the hiring of more labor. Considering these factors together, road access played a positive role in increasing agricultural income in our sample. However, its impact on non-agricultural income is rather minimal.

A possible reason for the insignificant impact on the non-agricultural sector is that the era of rural surplus labor in China has been over since the mid-2000s (Zhang et al., 2011). Under this circumstance, farmers increasingly rely on remittance as the major nonfarm income. In this remote area, farmers' migration decisions may have little to do with local infrastructure conditions. However, road access may help facilitate the market integration of the agricultural economy, therefore enlarging the production scale of products with comparative advantage. For example, in Natural Village 4 of Administrative Village II, the natural endowment is suitable for growing peaches. Before improvements in road connections, peaches were often damaged after being carried by shoulder to the nearest market for a long walk. After building a road, farmers are able to sell their peaches at a collection point right in their natural village. As a result, peach production has boomed in this area.

We shall add a cautious note that the findings on the positive impact of road investment on agricultural production do not necessarily mean that roads should be built to connect to all the remaining natural villages, as the marginal cost of building roads to the more remote communities may far outweigh the benefit. Thus, a benefit–cost analysis is needed when considering such rural road projects.

References

Benziger, V. 1996.“Urban Access and Rural Productivity Growth in Post-Mao China.” *Economic Development and Cultural Change*, 44(3): 539–570.

Calderón, C. 2009.“Infrastructure and Growth in Africa.” *World Bank Policy Research Working Paper Series No. 4914*.

Cameron, A. Colin, J B. Gelbach, and D L. Miller. 2008. “Bootstrap-Based Improvements for Inference with Clustered Errors.” *Review of Economics and Statistics* 90 (3): 414-427.

Chen, S., R. Mu, and M. Ravallion. 2009. “Are There Lasting Impacts of Aid to Poor Areas?” *Journal of Public Economics* 93 (3): 512–528.

Dorosh, P., H.G. Wang, L.You, and E. Schmidt. 2010. “Crop Production and Road

Connectivity in Sub-Saharan Africa: A Spatial Analysis.”*World Bank Policy Research Working Paper Series No.5385.*

Duflo, E. and R. Pande. 2007.“Dams.”*Quarterly Journal of Economics*, 122(2):601–646.

Escobal, J. and C. Ponce. 2002. “The Benefits of Rural Road: Enhancing Income Opportunities for the Rural Poor.”*Grupo de Analisispara el Desarrollo,GRADE Working Paper No. 40*, Lima.

Fan, S. and C. Chan-Kang. 2005. “Road Development, Economic Growth, and Poverty Reduction in China.” *IFPRI Research Report 138.*International Food Policy Research Institute, Washington D.C.

Fan, S. and P.Hazell. 2001. “Returns to Public Investments in the Less-favored Areas of India and China.”*American Journal of Agricultural Economics*, 83(5): 1217–1222.

Gibson, J. and S. Rozelle. 2003.“Poverty and Access to Roads in Papua New Guinea.”*Economic Development and Cultural Change*, 52(1): 159-85.

Gollin, D.and R. Rogerson. 2010.“Agriculture, Roads, and Economic Development in Uganda.”*NBER Working paper* No.15863.

Jacoby, Hanan.2000. “Access to Markets and the Benefits of Rural Roads.”*The Economic Journal*, 110 (465), 713-737.

Jacoby, H.G. and B. Minten. 2008. “On Measuring the Benefits of Lower Transport Costs.” *Journal of Development Economics,* 89: 28-38.

Jalan, J. and Martin Ravallion.2002.“Geographic Poverty Traps? A Micro Model of Consumption Growth in Rural China.”*Journal of Applied Econometrics*, 17: 329–346.

Khandker, S, Z.Bakht and G.Koolwal.2006. “The Poverty Impact of Rural Roads: Evidence from Bangladesh." *World Bank Policy Research Working Paper 3875*.

Limao, N. and A.J.Venables.2001."Infrastructure, Geographical Disadvantage, Transport Costs and Trade."*World Bank Economic Review,* 15(3): 451-479.

Mu, R. and D. van de Walle. 2011. "Rural Roads and Local Market Development in Vietnam."*Journal of Development Studies*, 47(5): 709-734.

Omamo, S. W. 1998. "Farm-to-Market Transaction Costs and Specialization in Small-Scale Agriculture: Explorations with a Non-Separable Household Model."*Journal of Development Studies*, 35(2):152-63.

Renkow,M., D.G. Hallstrom, and D. D. Karanja.2004."Rural Infrastructure, Transactions Costs and Market Participation in Kenya." *Journal of Development Economics*, 73(1): 349– 367.

Stifel, D., B. Minten, and P. Dorosh. 2003. "Transaction Costs and Agricultural Productivity: Implications of Isolation for Rural Poverty in Madagascar." Washington, D.C.: International Food Policy Research Institute, MSSD Discussion Paper No. 56.

Yang, X. and Y-k.Ng. 1993."Specialization and Economic Organization, a New Classical Micro economic Framework."Amsterdam, North-Holland.

Zhang, X., J. Yang, and S. Wang. 2011. "China Has Reached the Lewis Turning Point."*China Economic Review,* forthcoming.

Figure 1. Map of China

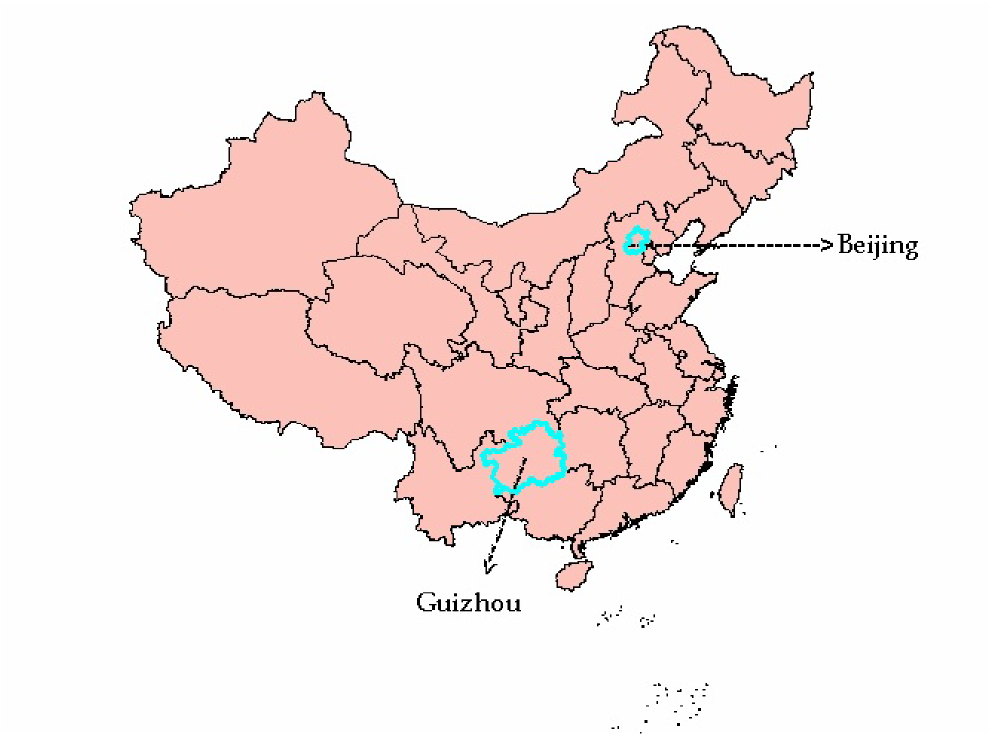

Figure 2. Map of Guizhou province with road network

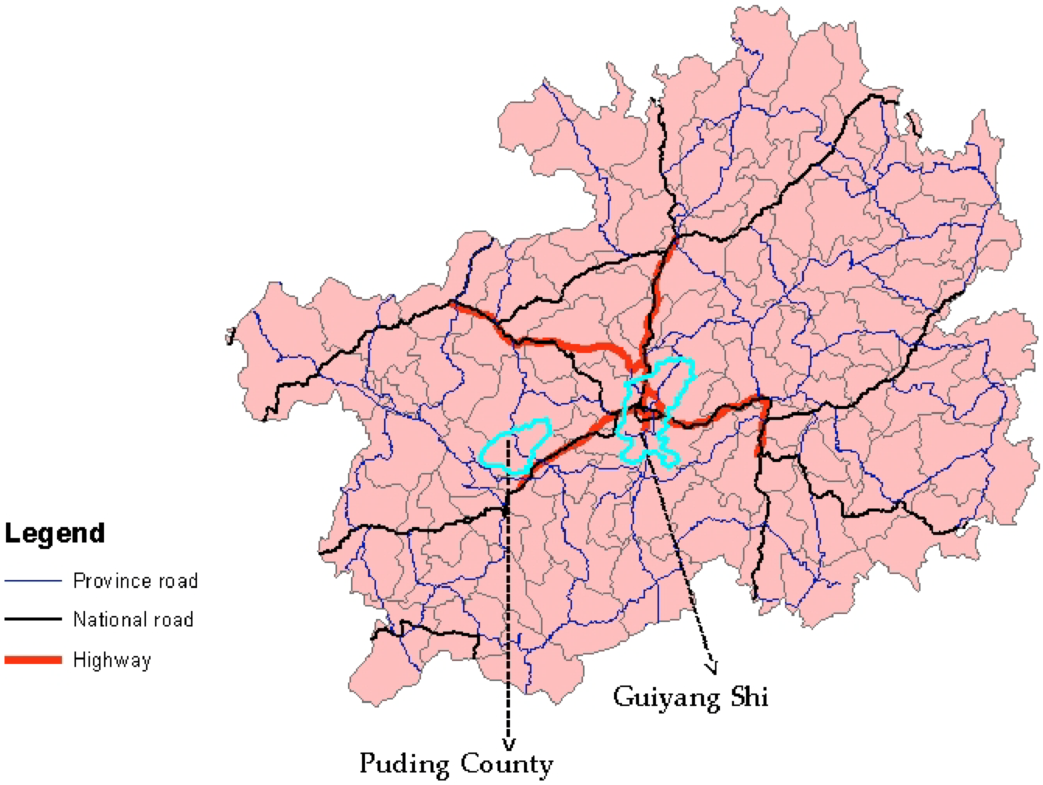

Figure 3.Income per capita of counties in Guizhou (Year 2008)

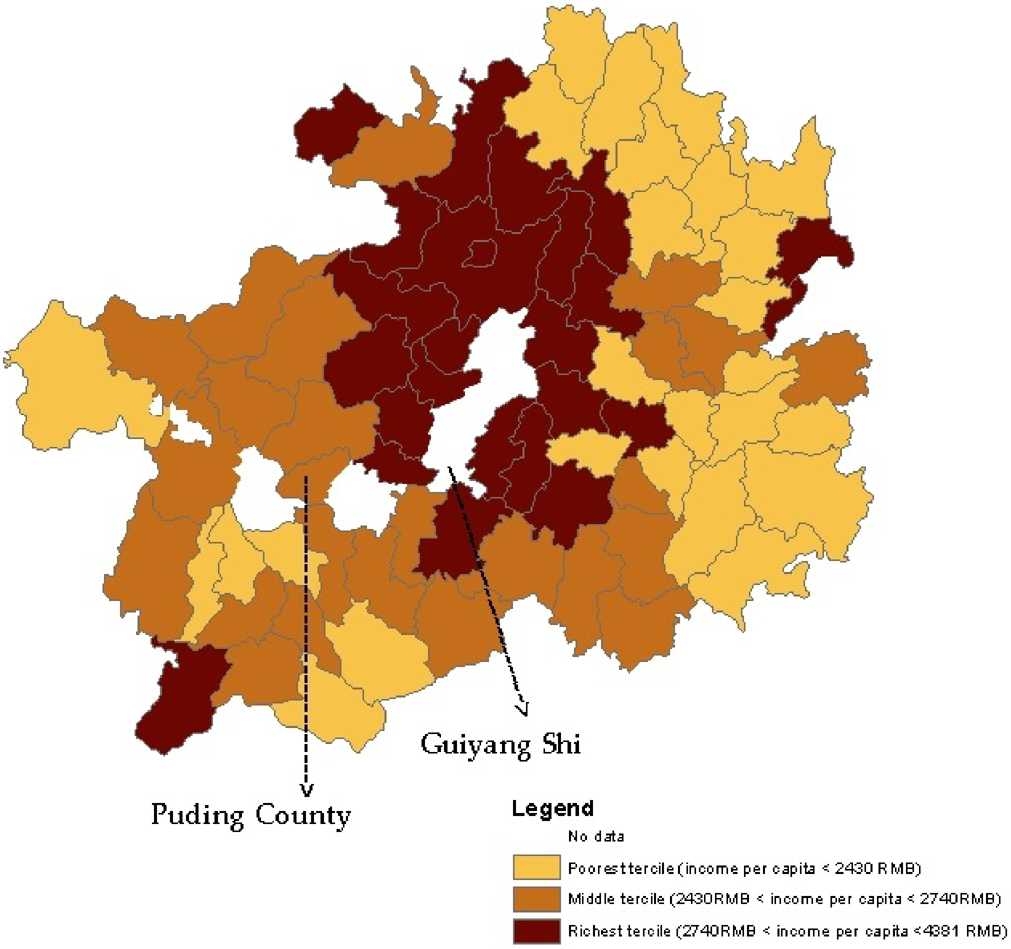

Table 1. Road Access in Three Surveyed Villages

	2004	2006	2009
Administrative Village I			
Natural Village 1	0	0	0
Natural Village 2	1	1	1
Natural Village 3	0	0	0
Natural Village 4	0	1	1
Natural Village 5	0	0	1
Natural Village 6	0	0	1
Natural Village 7	1	1	1
Natural Village 8	0	0	1
Natural Village 9	1	1	1
Administrative Village II			
Natural Village 1	0	0	0
Natural Village 2	1	1	1
Natural Village 3	0	0	0
Natural Village 4	0	1	1
Administrative Village III			
Natural Village 1	1	1	1
Natural Village 2	1	1	1
Natural Village 3	1	1	1
Natural Village 4	1	1	1

Source: Authors' Survey (2005, 2007, 2010).

Note: "1" denotes there is road access in the village in the specific year; "0" otherwise.

Table 2. Household Agricultural Income Shares in Different Categories

First Wave: Year 2004 (N=758)			
Income sources (%)	Mean	Std	Share>0
Corn	46.24	22.03	98.28
Paddy	11.87	17.68	46.04
Rapeseed	20.89	15.41	88.26
Vegetable	3.36	9.08	33.91
Fruit	2.20	7.02	20.71
Poultry	2.43	8.98	25.99
Livestock	11.93	21.21	29.95
Forestry	0.23	3.81	1.45
Fishing	0.84	6.46	2.11
Second Wave: Year 2006 (N=765)			
Income sources (%)	Share	Std	Share>0
Corn	38.70	21.48	94.90
Paddy	11.60	16.85	47.32
Rapeseed	19.28	14.34	92.81
Other grain	1.98	4.51	33.33
Vegetable	11.43	15.02	85.49
Fruit	3.38	10.72	23.27
Poultry	1.54	6.37	20.92
Livestock	10.14	19.06	27.32
Forestry	0.33	2.71	2.88
Fishing	1.62	10.43	3.14
Third Wave: Year 2009 (N=759)			
Income sources (%)	Share	Std	Share>0
Corn	42.04	24.34	94.99
Paddy	10.00	17.76	43.87
Rapeseed	16.00	12.40	93.15
Bean	4.00	7.62	82.35
Fruit	4.00	13.23	29.78
Poultry	2.00	10.60	37.94
Livestock	18.00	27.91	30.30
Forestry	0.00	5.00	5.27
Fishing	1.00	9.68	3.03

Source : Authors' survey (2005, 2007, 2010).

Table 3. Summary of Statistics

First Wave: Year 2004 (N=782)					
Variables	Obs	Mean	Std	Min	Max
Household income	782	6644.11	5156.50	0	50000
Agricultural income	782	3978.44	3861.81	0	37165
Non-agricultural income	782	2665.66	3472.30	0	50000
Household HH Index of agricultural production	758	0.46	0.16	0.19	1
Household size (migrants excluded)	782	3.69	1.55	0	8
Household land cultivated (*mu*)	781	3.66	2.76	0	20
Number of Labor (age 16-60)	782	2.53	1.42	0	7
Highest education in the household (year)	780	5.43	3.31	0	14
Village cadre in the household (dummy)	782	0.04	0.20	0	1
Second Wave: Year 2006 (N=815)					
Variables	Obs	Mean	Std	Min	Max
Household income	815	7618.57	10413.26	0	223080
Agricultural income	815	3825.31	3822.14	0	33148
Non-agricultural income	815	3793.26	9666.20	0	223000
Household HH Index of agricultural production	765	0.41	0.16	0.18	1
Household size (migrants excluded)	815	3.35	1.66	0	10
Household land cultivated (*mu*)	810	3.90	2.99	0	20.5
Number of Labor (age 16-60)	815	2.52	1.51	0	9
Highest education in the household (year)	811	6.12	3.47	0	18
Village cadre in the household (dummy)	811	0.02	0.13	0	1
Third Wave: Year 2009 (N=834)					
Variables	Obs	Mean	Std	Min	Max
Household income	834	11994.57	13933.61	0	191265
Agricultural income	834	5453.96	6613.53	0	67155
Non-agricultural income	834	6540.62	11868.99	0	182620
Household HH Index of agricultural production	759	0.49	0.18	0.22	1
Household size (migrants excluded)	834	3.85	1.83	0	12
Household land cultivated (acre)	806	3.10	2.92	0	32.5
Number of Labor (age 16-60)	834	2.47	1.48	0	7
Highest education in the household (year)	833	6.17	3.59	0	18
Village cadre in the household (dummy)	834	0.04	0.19	0	1

Source: Authors' survey (2005, 2007, 2010).

Table 4. Summary of Statistics by Road Access

Without Road Access (N=557)					
Variables	Obs	Mean	Std	Min	Max
Household income	557	5345	4357	0	37883
Agricultural income	557	3207	3217	0	38012
Non-agricultural income	557	2100	2504	0	20460
Household HH Index of agricultural production (HHI (1))	535	0.44	0.16	0.18	1
Alternative measure (HHI (2))	536	0.44	0.16	0.18	1
Alternative measure (HHI (3))	536	0.44	0.16	0.18	1
Household size (migrants excluded)	557	3.79	1.81	0	12
Household land cultivated (*mu*)	551	3.23	2.75	0	20.5
Number of Labor (age 16-60)	557	2.58	1.50	0	9
Highest education in the household (year)	556	5.24	3.40	0	16
Village cadre in the household (dummy)	555	0.05	0.21	0	1
With Road Access (N=1874)					
Variables	Obs	Mean	Std	Min	Max
Household income	1874	9669	11931	0	223080
Agricultural income	1874	4791	5455	0	68806
Non-agricultural income	1874	4883	10395	0	223000
Household HH Index of agricultural production (HHI (1))	1747	0.46	0.17	0.18	1
Alternative measure (HHI (2))	1753	0.45	0.17	0.18	1
Alternative measure (HHI (3))	1753	0.45	0.17	0.19	1
Household size (migrants excluded)	1874	3.59	1.66	0	12
Household land cultivated (*mu*)	1846	3.65	2.95	0	32.5
Number of Labor (age 16-60)	1874	2.48	1.46	0	8
Highest education in the household (year)	1868	6.12	3.47	0	18
Village cadre in the household (dummy)	1872	0.03	0.16	0	1

Source: Authors' survey (2005, 2007, 2010). All the prices are deflated to year 2003.

Table 5. Impact of Road Access on Agricultural Production

	Dependent Variables: Agricultural specialization indices and input use (fertilizer and labor input)									
	HHI (1)		HHI (2)		HHI (3)		Fertilizer use (per *mu*)		Hired Labor Cost (*yuan*)	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Road	-0.12***	-0.14***	-0.12***	-0.14***	-0.12***	-0.13***	0.03	-0.02	-0.04	0.21
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.30)	(0.29)	(0.27)	(0.26)
Road*beforeafter	0.05***	0.04***	0.05***	0.04***	0.05***	0.04***	0.28***	0.26***	0.23**	0.32***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.08)	(0.08)	(0.10)	(0.10)
Land		-0.01***		-0.01***		-0.01***		-0.07***		0.16***
		(0.00)		(0.00)		(0.00)		(0.01)		(0.03)
Number of primary age population (age 16-60)		0.00		0.00		0.00		0.04**		0.01
		(0.00)		(0.00)		(0.00)		(0.02)		(0.03)
Household size		-0.01***		-0.01***		-0.01***		0.04***		-0.04
		(0.00)		(0.00)		(0.00)		(0.02)		(0.03)
Highest education (year)		0.00		0.00		0.00		0.01		0.00
		(0.00)		(0.00)		(0.00)		(0.01)		(0.02)
Village cadre (dummy)		-0.04**		-0.04**		-0.04**		0.16		0.38
		(0.02)		(0.02)		(0.02)		(0.14)		(0.25)
Year fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Natural vllage fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
R-squared	0.09	0.13	0.08	0.11	0.07	0.10	0.07	0.10	0.20	0.23
AIC	-2007.7	-2084.7	-2064.1	-2126.7	-2080.3	-2145.4	6061.7	5982.3	9402.7	9322
N	2138	2138	2138	2138	2138	2138	2138	2138	2138	2138

Notes : 1. Households with no land are dropped from the regression for sample consistency through all the regressions. In addition, seven observations with self consumption ratio larger than one are dropped which are possibly generated from recording errors. Main findings remain the same after including those dropped observations; 2. *significant at 0.10 level; ** significant at 0.05 level; *** significant at 0.01 level. 3. Robust standard errors are clustered at natural village * year level.

Table 6. Impact of Road Access on Income

	Dependent Variable: Agricultural income, nonfarm income, income per capita and agricultural cash income							
	Agricultural income (log)		Nonfarm income (log)		Income per capita (log)		Agricultural cash income (log)	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Road	-0.29	0.03	-0.52	-0.28	0.12	0.20	-0.10	0.20
	(0.21)	(0.16)	(1.34)	(1.29)	(0.20)	(0.17)	(0.33)	(0.25)
Road*beforeafter	0.14	0.26**	-0.35	-0.22	0.04	0.06	0.20	0.32**
	(0.10)	(0.10)	(0.29)	(0.27)	(0.08)	(0.07)	(0.15)	(0.14)
Land		0.13***		-0.01		0.08***		0.15***
		(0.01)		(0.02)		(0.01)		(0.01)
Number of primary age population (age 16-60)		0.06***		0.22***		0.08***		0.06***
		(0.01)		(0.05)		(0.02)		(0.02)
Household size		0.08***		0.21***		-0.17***		0.07***
		(0.01)		(0.04)		(0.01)		(0.01)
Highest education (year)		0.00		0.04**		0.02**		0.00
		(0.01)		(0.02)		(0.01)		(0.01)
Village cadre (dummy)		0.20*		0.64*		0.31***		0.30**
		(0.12)		(0.35)		(0.10)		(0.15)
Year fixed effect	YES	YES	YES	YES	YES	YES	YES	YES
Natural vllage fixed effect	YES	YES	YES	YES	YES	YES	YES	YES
R-squared	0.09	0.30	0.06	0.09	0.13	0.30	0.10	0.21
AIC	5622.9	5067.8	10546.5	10469.6	5151.8	4714.4	7421.8	7147.4
N	2138	2138	2138	2138	2138	2138	2138	2138

Notes : 1. Households with no land are dropped from the regression for sample consistency through all the regressions. In addition, seven observations with self consumption ratio larger than one are dropped which are possibly generated from recording errors. Main findings remain the same after including those dropped observations; 2. *significant at 0.10 level; ** significant at 0.05 level; *** significant at 0.01 level. 3. Robust standard errors are clustered at natural village * year level.

Table 7. Impact of Road Access on Agricultural Production (Robustness Test)

	Dependent Variables: Agricultural specialization indices and input use (fertilizer and labor input)									
	HHI (1)		HHI (2)		HHI (3)		Fertilizer use (per *mu*)		Hired Labor Cost (*yuan*)	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Road	-0.07***	-0.08***	-0.07**	-0.07***	-0.07***	-0.07***	0.20	0.21	2.28***	2.30***
	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.34)	(0.32)	(0.25)	(0.26)
Road*beforeafter	0.05***	0.04**	0.05***	0.04**	0.05***	0.04**	0.28**	0.26*	0.23	0.32*
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.13)	(0.14)	(0.17)	(0.19)
Land		-0.01***		-0.01***		-0.01***		-0.07***		0.16***
		(0.00)		(0.00)		(0.00)		(0.01)		(0.03)
Number of primary age population (age 16-60)		-0.00*		0.00		0.00		0.04**		0.01
		(0.00)		(0.00)		(0.00)		(0.02)		(0.03)
Household size		-0.01***		-0.01***		-0.01***		0.04***		-0.04
		(0.00)		(0.00)		(0.00)		(0.02)		(0.03)
Highest education (year)		0.00		0.00		0.00		0.01		0.00
		(0.00)		(0.00)		(0.00)		(0.01)		(0.02)
Village cadre (dummy)		-0.04**		-0.04**		-0.04**		0.16		0.38
		(0.02)		(0.02)		(0.02)		(0.13)		(0.24)
Year fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Natural vllage fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
R-squared	0.10	0.13	0.08	0.11	0.08	0.11	0.07	0.10	0.20	0.23
AIC	-2007.7	-2084.7	-2064.1	-2126.7	-2080.3	-2145.4	6061.7	5982.3	9402.7	9322
N	2138	2138	2138	2138	2138	2138	2138	2138	2138	2138

Notes : 1. Households with no land are dropped from the regression for sample consistency through all the regressions. In addition, seven observations with self consumption ratio larger than one are dropped which are possibly generated from recording errors. Main findings remain the same after including those dropped observations; 2. *significant at 0.10 level; ** significant at 0.05 level; *** significant at 0.01 level. 3. Robust standard errors are clustered at natural village * year level.

Table 8. Impact of Road Access on Income (Robustness Test)

	Dependent Variable: Agricultural income, nonfarm income, income per capita and agricultural cash income							
	Agricultural income (log)		Nonfarm income (log)		Income per capita (log)		Agricultural cash income (log)	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Road	0.22	0.26	0.75	0.82	0.58***	0.53***	0.47	0.51
	(0.27)	(0.19)	(1.26)	(1.23)	(0.22)	(0.19)	(0.41)	(0.38)
Road*beforeafter	0.14	0.26*	-0.35	-0.22	0.04	0.06	0.2	0.32
	(0.17)	(0.16)	(0.53)	(0.51)	(0.13)	(0.12)	(0.26)	(0.26)
Land		0.13***		-0.01		0.08***		0.15***
		(0.01)		(0.02)		(0.01)		(0.01)
Number of primary age population (age 16-60)		0.06***		0.22***		0.08***		0.06***
		(0.01)		(0.05)		(0.02)		(0.02)
Household size		0.08***		0.21***		-0.17***		0.07***
		(0.01)		(0.04)		(0.01)		(0.01)
Highest education (year)		0.00		0.04**		0.02***		0
		(0.01)		(0.02)		(0.01)		(0.01)
Village cadre (dummy)		0.20*		0.64*		0.31***		0.30**
		(0.11)		(0.34)		(0.10)		(0.14)
Year fixed effect	YES	YES	YES	YES	YES	YES	YES	YES
Natural vllage fixed effect	YES	YES	YES	YES	YES	YES	YES	YES
R-squared	0.09	0.3	0.06	0.09	0.13	0.3	0.1	0.21
AIC	5622.9	5067.8	10546.5	10469.6	5151.8	4714.4	7421.8	7147.4
N	2138	2138	2138	2138	2138	2138	2138	2138

Notes : 1. Households with no land are dropped from the regression for sample consistency through all the regressions. In addition, seven observations with self consumption ratio larger than one are dropped which are possibly generated from recording errors. Main findings remain the same after including those dropped observations; 2. *significant at 0.10 level; ** significant at 0.05 level; *** significant at 0.01 level. 3. Robust standard errors are clustered at natural village * year level.