The Effects of Improved Roads on Wages and Employment

Evidence from Rural Labor Markets in Indonesia[1]

Futoshi Yamauchi[2]

World Bank

Current version: September 2015

Abstract

This paper examines the impact of road quality on labor supply and wages using household panel data from rural Indonesia. The analysis uses fixed-effect instrumental variable estimation by first differencing two-round panel data. First, road projects are found to increase the transportation speed. Second, the empirical results from intra-village variations of household endowments and labor-market behavior show that an increase in transportation speed raised wages in both non-agricultural and agricultural employment, and was associated with a decline in working time in agricultural employment, for the households whose members are relatively educated. The findings support potential complementarity between road quality and education, implying that the government's public investments in roads and education should be coordinated to capture cross-augmenting positive impacts in the long run.

Keywords: Complementarity, Roads, Transportation speed, Labor markets, Schooling, Indonesia

JEL classifications: I24, I25, J24, J31, O15

1. Introduction

It has been increasingly recognized that transportation infrastructures critically determine allocative efficiency of economic activities, and therefore influence household resource allocation and outcomes (e.g., Binswanger et al. (1993) and Fan et al. (2004)).[3] The improvement of road quality expands the geographic area individuals can reach, and resources can move faster. For instance, improved transportation networks enable households to reach new employment opportunities and also invite new capital investments, altering the spatial distribution of economic activities and demands for skills. New market opportunities can change returns to household resources, such as human capital and land, and thus, their initial endowments affect benefits they can have from new roads.

Household endowments potentially matter in the impacts of improved roads on household outcomes. For example, when the demand for skilled labor is concentrated in urban labor markets, the educated in rural areas will benefit more from improved spatial connectivity to urban areas (e.g., Fafchamps and Shilpi, 2003, 2005). To utilize these new opportunities created by improved spatial connectivity, knowledge and/or experience are important since the process involves learning as well as risk taking.[4]

In this paper, I examine the effects of improved roads on labor supply and earnings using household panel data from Indonesia by focusing on potential heterogeneity in the impacts related to the initial household endowments such as educational attainment. The analysis also compares agricultural and non-agricultural labor markets to investigate possible differences in household labor supply and wage responses to road improvements. The impact on wages (returns to labor) could be different between agriculture and non-agriculture since the improvement of roads potentially opens a faster access to urban labor markets. Accordingly, labor can be reallocated between agriculture and non-agriculture if wages change differentially between the two sectors and sector-specific skill requirements do not significantly constrain the household decision.

The following three points are important in the analysis. First, as mentioned, this paper examines how the intra-village distribution of household labor earnings alters in response to village-level changes in transportation speed. Thus, changes in the intra-village distribution of household labor earnings are compared across villages with and without changes in transportation speed. By focusing on the intra-village distributional issues, I will minimize the effect of endogenous placement of road projects. Second, I use an important characteristic of road projects that is useful to instrument change in transportation speed.[5] That is, during the project, the local demand for construction workers will certainly increase, but the effect ceases when the project is completed, whereas the impact of road improvement on transportation speed will sustain as long as the road quality is maintained.

Third, the primary source of our data is the village and household surveys conducted in 2007 and 2010 for 98 villages in 7 provinces. Indonesia, at least in the survey sites, had established key road networks before 2007 so most village road projects aimed to improve and rehabilitate existing roads during the above mentioned period (i.e., improvements in road quality), rather than opening new road paths.[6] To measure road quality, I estimated transportation speed by dividing distance to an employment site (measured in kilometers) by the required number of hours.

In the literature, Yamauchi et al. (2011) showed from Indonesia that improvements in road quality in neighboring areas increased incomes of relatively educated households and promoted their transition to non-agricultural employment between 1995 and 2007. The findings of the current study are also consistent with their work. Gachassin (2012) showed similar evidence on complementarities between road access and education using panel data from Tanzania. In contrast with agriculture, non-agricultural economic activities generally require more advanced skills and knowledge to handle more complicated technologies (e.g., Yamauchi, 2004; Fafchamps and Shilpi, 2005). To realize larger returns to human capital in the context of rural labor markets, it seems critically important to have access to non-agricultural labor markets (Fafchamps and Shilpi, 2003).[7]

In the context of Indonesian road networks, three papers in particular are close to the current study: Gertler, et al. (2014), Gibson and Olivia (2010) and Yamauchi, et al. (2011). Gibson and Olivia (2010) used panel data from the Indonesian Family Life Surveys (IFLS) to confirm a positive linkage between non-farm enterprises and infrastructure developments such as roads and electricity. Using the IFLS but departing from conventional measures of road quality, Gertler, at al. (2014) adopted International Roughness Index to investigate the effects of roughness-induced travel time on various economic outcomes.[8] Yamauchi, et al. (2011) used the proportion of inter-village roads made of asphalt, concrete, or cone blocks computed from village census data (PODES), and showed a positive effect of changes in road quality on non-agricultural labor supply.

Table 1 to be inserted

Table 1 shows the proportion of inter-village roads using asphalt, concrete, or cone blocks in 1996 and 2006.[9] As confirmed in Gibson and Olivia (2010) citing Parikesit (2006), the quality of existing roads substantially varies across regions. The table demonstrates that, in addition to variations in quality at a given point in time (e.g., 100% in Jakarta to 40% in Nusa Tenggara Timor), the changes also substantially vary across provinces. The average trend is, however, clearly upward in the country. Changes in percentage point in the 7 sample provinces from 1995 to 2006 are: +5% (Lampung), +10% (Central Java), +8% (East java), -4% (NTB), +4% (South Kalimantan), no change (North Sulawesi) and +10% (South Sulawesi). In 2006, one year prior to the initial survey round, the least developed province was Lampung (47% asphalt), while the most developed ones were North Sulawesi (76%) and Central Java (74%). In general, Java has been the most developed area in terms of rural road conditions.

In sum, the contribution of this paper is three-fold. First, the data enable me to directly compute the average transportation speed and the impacts on household labor supply and income are estimated. Second, agricultural and non-agricultural labor markets are distinguished. Third, the analysis shows distributional effects across households within a village, which offers direct implications for local income distributions. The next section describes the empirical strategy.

2. Empirical Strategy

This section describes the specification and estimation strategy used, and discusses identification issues. The analysis uses household-level panel data to analyze labor supply and wages that potentially respond to changes in road quality.

Transportation speed is a function of road quality and transportation modes, the latter of which is again a function of road quality and household income.[10] In this setting, as a reduced form, road quality determines transportation speed. The primary goal of this paper is to estimate the effect of change in transportation speed on labor supply and wages, but transportation speed is endogenous in the sense that endogenous allocation of road projects, which affects transportation speed, is potentially correlated with location characteristics.

To mitigate potential endogeneity problems related to the placement of road projects to villages, the focus of this paper is on heterogeneous impacts of road quality improvement (measured by changes in transportation speed) associated with household endowments such as education, gender and age (i.e., interaction terms of change in transportation speed and the initial household characteristics). Combined with first differencing between the two survey rounds (2007 and 2010), the current procedure therefore uses triple differencing. Furthermore, the number of road projects that took place between 2007 and 2010 is used as an instrument for change in transportation speed to wipe out potential bias that comes from a possible correlation between change in transportation speed and household-level labor supply or wage shocks.[11]

The following first-differenced equation is estimated,

$$\Delta y_{ij(0,1)} = \alpha + \beta_1 \Delta s_{j(0,1)} + \beta_2 \Delta s_{j(0,1)} x_{ij0} + x_{ij0}'\gamma + \Delta v_{j(0,1)} + \Delta \varepsilon_{ij(0,1)} \qquad (1)$$

where $\Delta y_{ij(0,1)}$ is change in the number of working days (or months) or wages for household *i* in village *j*, from time *0* to *1*, $\Delta s_{j(0,1)}$ is change in the average transportation speed between 2007 and 2010, x_{ij0} is a vector of initial household characteristics, $\Delta v_{j(0,1)}$ is a village-level shock, and $\Delta \varepsilon_{ij(0,1)}$ is the error term, that is, difference in year-specific shocks (assume that ε_{ijt} is an ex-post shock after household decisions are made in each year). Note that β_1 is the effect of change in the average transportation speed on the dependent variable, and β_2 captures how the initial household characteristics affect the impact of change in the average transportation speed. The estimated village-level average mobility speed is interacted with household characteristics x_{ij0}: the average years of schooling, the average age and number of household members defined in the prime working age 20-55 in *t=0*.

Since $\Delta s_{j(0,1)}$ is correlated with village-level shocks in the labor market conditions (e.g., an increase in income enabling households to purchase motorbikes), $\Delta v_{j(0,1)}$ is controlled for using village fixed effects. The focus of estimation is on identifying β_2, rather than β_1. That is,

$$\Delta y_{ij(0,1)} = \alpha + \beta_2 \Delta s_{j(0,1)} x_{ij0} + x_{ij0}'\gamma + FE_j + \Delta \varepsilon_{ij(0,1)} \qquad (2)$$

where FE_j is the village fixed effects, which capture both $\beta_1 \Delta s_{j(0,1)}$ and $\Delta v_{j(0,1)}$. With the inclusion of FE_j, the inference is based on *intra-village* variations in the above estimation.

The estimation uses instruments to remove potential bias due to correlations between initial period wage shocks and changes in the average transportation speed. For example, a positive income shock in the initial period enables the household to choose a more expensive but speedy transportation mode (e.g., purchasing a motorbike), which increases the average speed prevailing in *t=1*. To wipe out potential bias due to the correlation between $\Delta s_{j(0,1)}$ and ε_{ij0}, the number of road projects at the village level ($\Delta R_{j(0,1)}$) is used as an instrument for $\Delta s_{j(0,1)}$. However, since the village fixed effects are included in Eq (2), the actual instruments are also the interactions between $\Delta R_{j(0,1)}$ and x_{ij0}. Road projects in the current empirical setting aimed to improve rural roads including rehabilitation and maintenance, not to construct new roads. The village survey data provide information on start and completion dates (year/month) of a road project, as part of the village history.[12]

Two features of road projects deserve special attention. First, since road projects are mostly public investments and their preparation normally requires some time (say, a few years), their installation is unlikely to be correlated with individual decisions influenced by the initial period labor market shocks (ε_{ij0}).[13] Second, the direct impact of road projects on labor demand is by definition temporary, so the direct effect ceases after project completion. However, the impact of road improvements on transportation speed lasts as long as the quality of the road is maintained. Section 4 shows some evidence supporting this point (Table A1). For the above two reasons, $E\left[\Delta R_{j(0,1)} x_{ij0}' \varepsilon_{ij0}\right] = 0$. The next section describes panel data used in the analysis.

3. Data

In this section, I describe the survey data from Indonesia used in this analysis in detail. The data come from two rounds of household survey conducted in Indonesia (Indonesian Center for Agricultural Socio Economic and Policy Studies and International Food Policy Research Institute). First, the primary source of our data is the village and household surveys conducted in 2007 and 2010 for 98 villages in 7 provinces (Lampung, Central Java, East Java, West Nusa Tenggara, South Sulawesi, North Sulawesi, and South Kalimantan) under the Japan Bank for International Cooperation (JBIC; currently JICA)'s Study of Effects of Infrastructure on Millennium Development Goals in Indonesia (IMDG).[14] Figure 1 shows locations of surveyed villages.

Figure 1 to be inserted

The selection of the survey sites needs some explanations. The 2007 survey was designed to overlap with villages covered in the 1994/95 PATANAS survey conducted by ICASEPS to build household panel data from 1995 to 2007. The 1994/95 PATANAS survey focused on agricultural production activities in 48 villages chosen from different agro-climatic zones in seven provinces.[15] The sample was representative for major agro-climatic zones in Indonesia. In 2007, the team visited those villages to expand the scope of research as a general household survey.

The 2007 round added 51 new villages to the 1994/95 PATANAS sample. These new villages were selected with the following criteria. First, the same districts (kebupaten) where PATANAS villages are located were chosen. Second, villages that had received relatively large amounts of government infrastructure projects during the period of 1995 to 2005 were listed (mainly funded by either the Japan Bank for International Cooperation or the World Bank).[16] Finally, the new villages were randomly sampled from the list. The above setting guarantees variations in infrastructure investments in the sample though their allocations were not random.

In the revisited original-sample villages in 2007, 20 households were resampled per village from the 1994/95 sample and the split households (defined below) were included. In the new villages, 24 households were sampled in each village. In 2010, the survey team revisited all the 98 sample villages to re-interview sample households and their splits. Out-migrants who moved out of the villages were also tracked through either direct or phone interviews in 2010.

Between the surveys, some household members separated from the 2007 original households to start their own new households or to join other households. In the survey, a split household is defined as a new household in which (i) a member of the original household became the *head* of a new household, and (ii) the household resides in the same village. The above two conditions define split households in this study; household members who moved out of their original village or simply joined other households within the village are not called splits in the survey. From 2007 to 2010, 9.25% of the 2010 households were split households under the above definition. 12.97% of the sample individuals recorded as household members in 2007 migrated out, i.e., either moved out of their village or joined other households. The proportion is 11.82% once the group is restricted to those aged 20-55 in 2007.

The omission of out-migrants could potentially downwardly bias the estimated transportation speed impacts on labor supply and earnings if improved spatial connectivity is likely to make able individuals migrate to look for outside employment opportunities. However the opposite could hold if more able individuals can commute when roads are improved. Though out-migrants are most likely to be engaged in non-agricultural works in the current empirical context, this will also lead to an under-estimation of the impacts on labor supply in non-agricultural sectors.

4. Road Infrastructure and Transportation Speed

This section shows some descriptive evidence for the relationship between road projects and transportation speed. The rate of change in speed (i.e., change in the log of speed) is related to the number of road projects completed in the period from 2007 to 2010. Table 2 tabulates the number of road projects by village. More than 61 percent of the sample villages had some road projects completed during the period. However, most villages fall into only one project or zero, which implies that a large portion of the variations come from whether they had a road project or not.

Table 2 to be inserted

Both the 2007 and 2010 surveys have employment modules that capture the nature of employment, such as wages (both cash and in-kind), duration, contract types for each employment case in the past year (therefore, an individual can have multiple records of employment experience). An important feature of the IMDG surveys is that they captured details of commuting in each case, e.g., distance, time, cost and main mode of transportation. Some cases include temporary migration--such as living and working in a distant location for some months and returning to the original household-- since our definition of household membership uses 6 months of residence in the past one year. However even in this case, it is useful to compute the speed to capture road and transportation quality in the surrounding areas.

Figures 2a/2b and 3a/3b to be inserted

Figure 2a shows the distributions of speed estimates from individual non-agricultural employment cases in 2007 and 2010. The distributions are skewed in both 2007 and 2010. The distribution has slightly shifted rightward. Figure 2b displays the distributions in agricultural employment. Changes in the distribution are quite small. Though both distributions are skewed, the speed in accessing agricultural employment is much lower than that of non-agricultural employment. Next, Figures 3a and 3b show the relationship between the estimated speed and distance for non-agriculture and agriculture, respectively. It is found that the speed has increased evenly in a large range of distance in non-agriculture, but the change is not little in agriculture. The difference most likely comes from modes of transportation actually used in reaching employment. In agriculture, 69% of reported cases walked (on foot) in 2007, whereas only 42% walked in non-agriculture, followed by private motorbike (25%) and mini bus (13%). Walking speed is less responsible to change in road quality. Therefore, Figures 2 and 3 point to the asymmetry between the two sectors: transportation speed has increased more in accessing non-agricultural employment.

Figures 4 to be inserted

Figures 4 shows the relationship between the speed for non-agricultural employment (village-average) in 2007 and (a) the proportion of asphalt/concrete/cone-blocked and (b) the proportion of 4-wheel passable roads in inter-village roads at sub-district (kecamatan) in 2006 (both constructed from the 2006 PODES village census).[17] The graphs show that they are positively correlated, which implies that the overall quality of inter-village roads in the neighboring areas seems to explain speed in accessing non-agricultural employment. The above observation is reasonable since the actual speed available to villagers is determined by both their own inter-village roads as well as those in their neighboring areas.

Figure 5 to be inserted

Figure 5 shows the distribution of change in log village mean speed (averaging non-agricultural and agricultural employment). The change is symmetrically distributed across villages.

Table 3 to be inserted

Next, the change of log speed is regressed on the number of road projects completed in 2007-2010, the village-mean log speed in 2007, and their interactions (Table 3, Columns 1 and 2). The access speed for either non-agricultural or agricultural employment could not be calculated for 6 villages, thus not enabling to compute the mean village-level speed, which renders 92 villages in the regressions. It is found that the number of completed road projects significantly has increased the village-mean speed.

The initial level of speed is negatively associated with the subsequent change of speed, and the interaction term of the initial speed and the number of road projects has a significant negative effect on the change in speed, which implies that a potential gain from improved roads is larger if the initial level of speed is low (i.e., road conditions are bad in the initial period).

In Column 3, the number of road projects is regressed on the initial village-level speed and the proportions of asphalt roads and four-wheel transportable roads in the sub-district. The incidence of road constructions/rehabilitations is not correlated with the mobility speed in the initial period,[18] but is positively related to the proportion of asphalt roads within the sub-district, implying that potential benefits of the road improvement could be greater if the village is surrounded by relatively better roads (this is consistent with Figures 4).

Note that road projects could have increased labor demands during the project period (e.g., construction period) but, by definition, such a temporary increase in employment in road constructions does not sustain after the project completion. Though working time and wages in 2010 could reflect some parts of the recent road project impacts, the major portion of the impacts is not directly related to the dependent variables. Table A1 shows the effect of road projects on the change in working time, measured in months for non-agricultural and days in agricultural employment, respectively. The number of road projects does not affect the change in working time in both cases due to the fact that road construction can only temporarily alter labor demands during the project period and the effect ceases upon the completion.

Figure 6 to be inserted

Variations in number of projects could be explained by locations and political importance too, but these are village-specific factors and therefore will be controlled by village dummies. Again, the focus of this analysis is change in intra-village distribution, jointly attributable to changes in transportation speed (and road projects) and household endowments. Before showing the estimation results, the intra-village schooling distributions are compared between villages with and without road projects. Figure 6 uses the residuals of the average years of schooling completed after controlling village dummies. The two distributions are strikingly similar, which indicates that the intra-village schooling distributions are not different between road project and non-project villages, though the placement of projects could be biased to a certain group of villages.

5. Empirical Results

This section summarizes estimation results. The estimation includes village dummies to absorb village-level shocks (especially, general price changes). Village average change in speed is interacted with household-level average years of schooling completed, average age (age 20-55) and the number of household members aged 20-55, all measured in 2007. Instruments are: indicators for the number of road projects completed interacted with years of schooling, average age and the number of household members aged 20-55 in 2007, and other exogenous variables.

Table 4 to be inserted

Before reporting the main results, it is important to analyze household split behavior since it potentially affects the estimates of the transportation speed impact on labor supply and earnings. On the mechanism of household splits (in the above definition), the incidence of splits is regressed on the explanatory variables used in the main outcome equations (Table 4).[19] Column 1 uses province dummies, and Column 2 uses village dummies. The results show that changes in the average transportation speed, and its interactions with the initial household characteristics are all insignificant, which indicates that transportation factors are not driving household splits (specifically defined in

Section 3). Second, among the household characteristics, only the average age and the number of household members, both defined only over the prime-working age in 20 to 55, are significantly related to the incidence of splits. Namely, older households tend not to split, while larger households tend to be split. Based on the above results, I aggregate the original households and their split households in 2010 (both residing in the same village) to match with the original household in 2007. Though the decision to start a new household in the same village could be potentially affected by changes in inter-village road quality, the magnitude of any such influence seems to be substantially smaller than that of migration decisions since the major factor affecting split decisions (rather than migration) remains demographic, such as marriage (see Table 4). [20]

Tables 5 and 6 show the main results from the household-level panel analysis. Table 5 shows the effects of transportation speed change on working time. Columns 1 and 2 show the results on non-agricultural employment. In both no-instrument and instrument cases, changes in the speed have no significant effect (via schooling, age and household size). Columns 3 and 4 show the results on agricultural employment working time. Working time in non-agricultural employment decreased if the average educational attainment is high (Column 3, no instrument). The above results point out that there is a tendency that the household decreases their labor supply to agricultural works, especially among the educated members, but the effect on non-agricultural work was not clear. There is a possibility that agricultural producers allocate more of their time to their production activities, by reducing labor supply to agricultural labor works on their farms.

Tables 5 and 6 to be inserted

Table 6 summarizes estimation results for wages. Monthly and daily incomes are used as dependent variables for non-agricultural and agricultural employment, respectively. Columns 1 and 2 show results for the monthly wage received in non-agricultural employment with and without instruments, respectively. First, in both cases, an increase in the village-average speed significantly raises returns to schooling at the household level. In other words, relatively more educated households experience increased wages (and total income) from non-agricultural employment when road and transportation quality improve in the neighboring areas. The effect is larger in the instrumental variable estimation than the OLS estimation. The effect of the average age becomes smaller (negative effect). Household size (age 20-55) does not have the above effect. The test of exogeneity shows that the IV estimates significantly differ from the OLS estimates (p = 0.015).

Columns 3 and 4 show the results from agricultural employment daily wages with and without instruments, respectively. Interestingly, it is found that the interaction term of changes in village-average speed and household schooling is significantly positive (in contrast to the cross-sectional wage regressions). The estimate of the speed-schooling complementarity is larger in the instrumental variable estimation than that of the OLS estimation, though the test of exogeneity favors the non-instrument case.

Tables 7 and 8 to be inserted

Next, Tables 7 and 8 report the results using the sample of villages where road projects were completed before 2009. In other words, the sample does not include villages where road projects were still implemented in 2009 and/or 2010 to avoid potential effects of the projects on labor demand. Since the employment module asks employments that took place in the past 12 months, the 2010 round data may contain employments created by road constructions if we include road projects that continued to 2009 and/or 2010.

The results are similar to those reported in the previous tables. That is, the instrumental variable estimation shows that change in transportation speed significantly raised wages in non-agricultural employment among relatively educated households. In contrast, an increase in transportation speed significantly decreased non-agricultural wages among relatively old and large households. In other equations, the OLS estimation results were preferred, showing that the effect of increased transportation speed is not statistically significant.

6. Conclusions

This paper examined the impact of road quality on labor supply and wages using household panel data from rural Indonesia. The findings are two-fold. First, an increase in the average transport speed increased labor market wages in both non-agricultural and agricultural employment, while decreasing working time in agricultural employment, especially for the households whose members are relatively educated. The findings support the complementarity between road quality and education. However, the question of whether better roads attract new investments and create more job opportunities in the same place remains unanswered here. Consistent with the previous studies (e.g., Yamauchi et al., 2011) that showed that the income sources of the educated tend to shift from agriculture to non-agriculture by increasing labor supply to non-agricultural employment when road quality improves in the local economy, the current study also showed that the improvement of road quality, measured by increased transportation speeds significantly augmented returns to schooling in non-agricultural labor markets.

Second, this paper also has shown that road projects significantly increased the average transportation speed around villages. The average speed also depends on spatial connectivity in the neighboring areas, measured by the proportion of inter-village roads made of asphalt, concrete, or cone-blocks inter-village roads within the sub-district. This simple correlation informs us of the importance of spatial coordination in public road investments.

The complementarity between education and road quality has a clear implication for the government's public investment decisions. Public investments in rural road constructions can generate greater labor-market returns if the population is, on average, more educated. That is, unless educational attainment is achieved, we cannot expect large gains in labor market incomes from improved roads. Labor-market employment transition to non-agriculture is promoted only if the population is relatively educated.[21] Besides, improved roads increase school enrollment and attendance unless improved access to work opportunities substantially discourages children to study. For the reasons mentioned above, greater positive impacts are expected in the long run if the government's public investments are coordinated to improve local transportation and develop spatial connectivity as well as support rural education (through, e.g., school constructions).

---

[1] I would like to thank Elisabeth Goller, Arif Haryana, Megumi Muto, Will Martin, seminar participants at the Indonesian Center for Agriculture Socio Economic and Policy Studies (ICASEPS) and anonymous referees for useful comments and suggestions. Special thanks to Ali Subandoro, Dini Inayati, Iis Surtina, and Sony Sumaryanto for their superb assistance to conduct the IMDG 2010 survey. The data are available from the author upon request. Any remaining errors are mine.

[2] The World Bank, 1818 H Street, NW, Washington DC 20433; Email: fyamauchi@worldbank.org; Phone: 202-473-2065

[3] The previous studies estimated the returns to road constructions using aggregate data. To identify the impact of road infrastructure, however, it is necessary to use household-level data with a sufficiently long span of period in which road conditions change or explicit information on new road projects. This study takes the latter approach.

[4] This is similar to the availability of a new technology (Schultz, 1975), which increases returns to schooling. The hypothesis that the educated have an advantage in learning has been tested by Foster and Rosenzweig (1995) and others.

[5] Fafchamps and Shilpi (2005) uses walk time and physical characteristics to instrument actual transportation time in their Nepal study.

[6] Indonesia widely introduced decentralization in 2000, road projects can be implemented at the district (kebupaten) or sub-district (kecamatan) level. Road rehabilitation and maintenance are therefore managed at that level, which potentially raises the issue of coordination problems between districts (or sub-districts). Note that the administrative hierarchy under province goes down from district, sub-district and finally to village (desa).

[7] In agricultural commodity markets, Jacoby (2000) and Jacoby and Minten (2009) showed evidence that farmers who transport their produces to local markets can potentially increase their profits (reflected in the value of their farm land).

[8] The survey data used in this analysis do not provide information required to construct International

Roughness Index.

[9] The village census, PODES, reports materials used in inter-village roads for each village.

[10] Potentially, transportation speed also depends on the hour of the day ride is undertaken. However, since traffic congestion is very rarely observed in the current empirical setting, this factor is assumed out here.

[11] Transportation speed is computed from hours and kilometers required for a household member to reach employment. Employment cases are recorded in both dry and rainy seasons, which raises the issue of measurement errors. The use of instruments therefore also minimizes potential attenuation bias due to measurement errors.

[12] The survey does not provide information on the magnitude of a project, such as budget size and area coverage (meters). Using the information on start and completion dates, I prepare different samples to check robustness of the main results.

[13] Note that villages with a relatively better endowment, e.g., in educational attainment, could be possibly in a good position to plan road projects (since they know potential returns to such a project are high in their villages). However, these potential village-specific effects are all controlled by village fixed effects.

[14] See also Yamauchi (2012) that used the same survey data from 2007 and 2010 to analyze determinants of birthweight seasonality and its impacts on child human capital formation.

[15] In addition, the initial sample included Aceh.

[16] The analysis uses data from 2007 and 2010. Public investments (externally funded by the two international organizations) prior to the 2007 round are part of the initial conditions, not directly affecting road projects during 2007-2010. As mentioned, placement of road projects is not random. Thus, the focus of the current analysis is intra-village distribution, not inter-village distribution, driven by change in transportation speed (partly explained by the number of road projects).

[17] One village in North Sulawesi (currently in Gorontalo) could not be identified in the 2006 PODES. This village is not included when we use the 2006 PODES (Figure 4 and Table 1). However, since we do not use the 2006 PODES in the main analysis, this does not affect the main results.

[18] As the next section shows, non-agricultural wages are positively related to the mobility speed in 2007.

[19] The sample size is 1904 since 6 villages were omitted due to collinearity.

[20] In a preliminary analysis, Eq(2) was estimated without including households who split (i.e., both main and split households). The results remain qualitatively similar, which confirms that in this particular period, potential bias due to the split process is small (consistent with the results in Table 4). In split and migration decisions, local factors such as arable land and population pressure are potentially important but these are controlled by village dummies.

[21] In general, improved rural roads can support the evolution of self-employed activities (i.e., small enterprises). Though many of such business/enterprise activities do not necessarily contribute to poverty reduction, they can be an important source of rural employments. Educational attainment, among many others, could be an important key factor that determines new enterprise activities. For example, agro-processing can be more profitable if they are connected through improved roads to local markets. Note that non-agricultural employment in this paper includes such employment opportunities created by new enterprise activities (but not including self-employment itself), so the above issue is partly covered in the analysis. More generally, whether or not improved rural roads diversify employment (and income) opportunities is an important issue.

References

Binswanger, H., Khandker, S.R., and Rosenzweig, M.R. (1993) How infrastructure and financial institutions affect agricultural output and investment in India. *Journal of Development Economics*, 41(2), pp.337-366.

Fafchamps, M and Shilpi, F. (2003) Spatial division of labor in Nepal. *Journal of Development Studies*, 39(6), pp.23-66.

Fafchamps, M. and Shilpi, F. (2005) Cities and specialization: Evidence from south Asia. *Economic Journal*, 115(503), pp.477-504.

Fan, S., Zhang, L., and Zhang, X. (2004) Reforms, investment, and poverty in rural China. *Economic Development and Cultural Change*, 52(2), pp.395-421.

Foster, A.D. and Rosenzweig, M.R. (1995) Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of Political Economy*, 103(6), pp.1176-1209.

Gachassin, M.C. (2012) Roads and education complementarity: Empirical evidence from the Kagera region, Tanzania, Mimeograph, Centre d'Economie de la Sorbonne, Université Paris 1 Panthéon-Sorbonne.

Gertler, P., Gonzalez-Navarro, M., Gracner, T., and Rothenberg, A.D. (2014) The effects of road quality on labor markets: Evidence from Indonesia's roads, Mimeograph, University of California at Berkeley.

Gibson, J. and Olivia, S. (2010) The effect of infrastructure access and quality on non-farm enterprises in rural Indonesia. *World Development*, 38(5), pp.717-726.

Jacoby, H. (2000) Access to markets and the benefit of rural roads. *Economic Journal*, 110(465), pp.713-737.

Jacoby, H. and Minten, B. (2009) On measuring the benefits of lower transport costs. *Journal of Development Economics*, 89(1), pp.28-38.

Minten, B. and Kyle, S. (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(2), pp.467-495.

Parikesit, D. (2006) Infrastructure support for improving rural investment climate, Background Paper for the Indonesia Rural Investment Climate Assessment.

Schultz, T.W. (1975) The value of the ability to deal with disequilibria. *Journal of Economic Literature*, 13(3), pp.827-846.

Yamauchi, F. (2004) Are experience and schooling complementary? Evidence from migrants' assimilation in the Bangkok labor market. *Journal of Development Economics*, 74(2), pp.489-513.

----------- (2012) Prenatal seasonality, child growth and schooling investments: Evidence from rural Indonesia. *Journal of Development Studies*, 48(9), pp.1323-1341.

Yamauchi, F., Muto, M., Chowdhury, S., Dewina, R., and Sumaryanto, S. (2011) Are schooling and roads complementary? Evidence from income dynamics in rural Indonesia. *World Development*, 39(12), pp.2232-2244.

Figure 1 Locations of surveyed villages

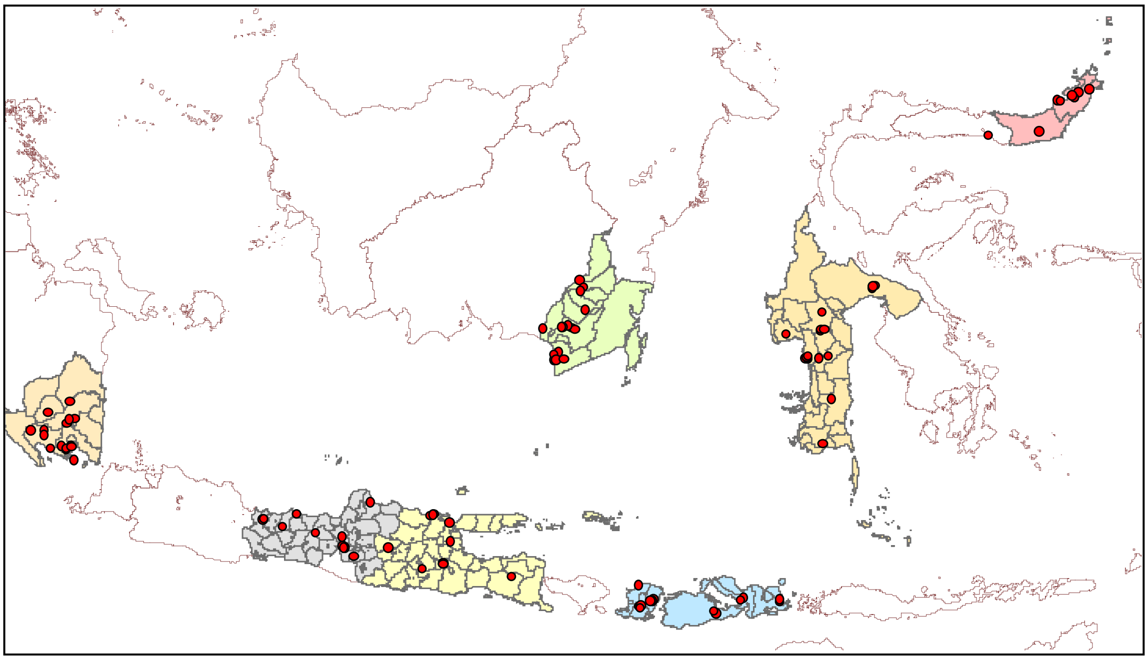


Figure 2a Density of transportation speed in non-agricultural employment

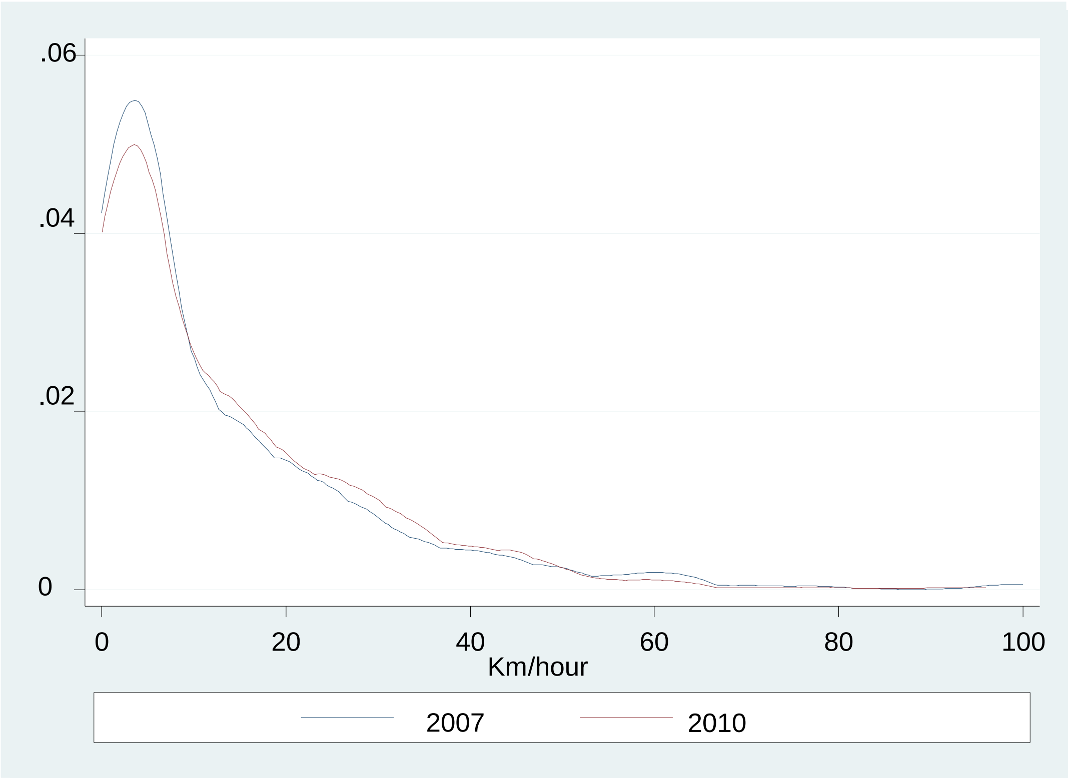


Figure 2b Density of transportation speed in agricultural employment

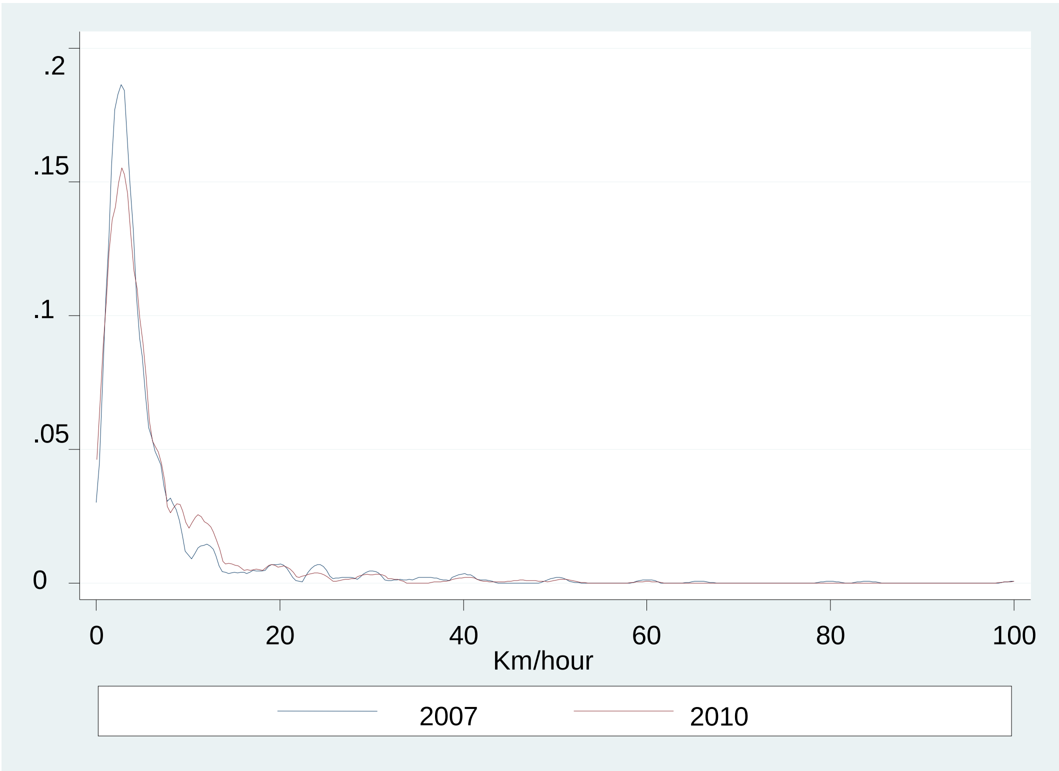


Figure 3a Speed (km/hour) vs distance (km) in non-agricultural employment

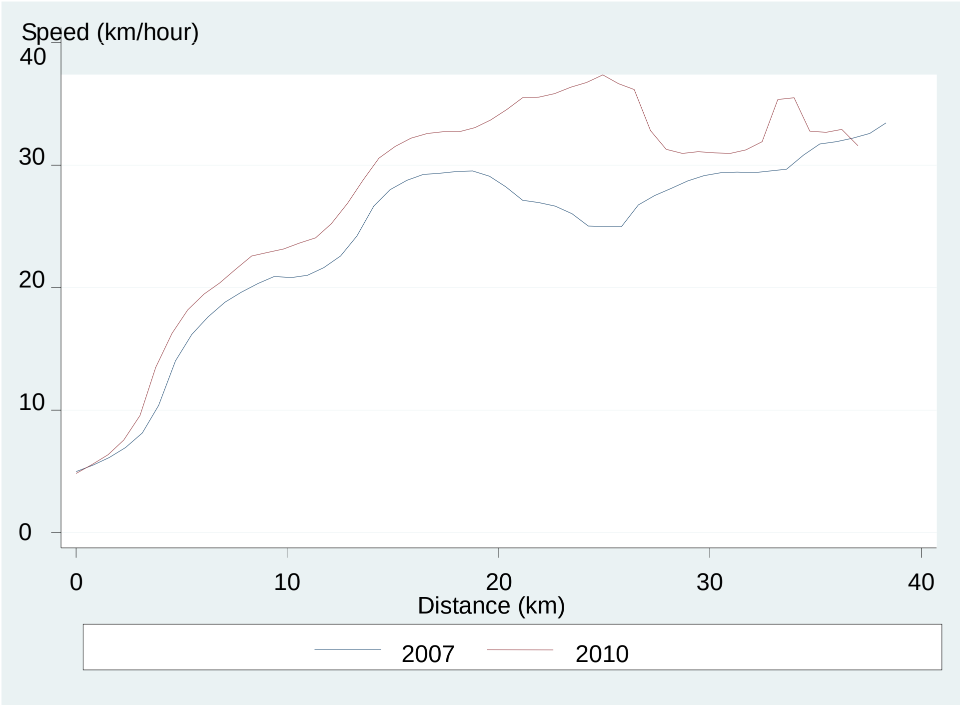


Figure 3b Speed (km/hour) vs distance (km) in agricultural employment

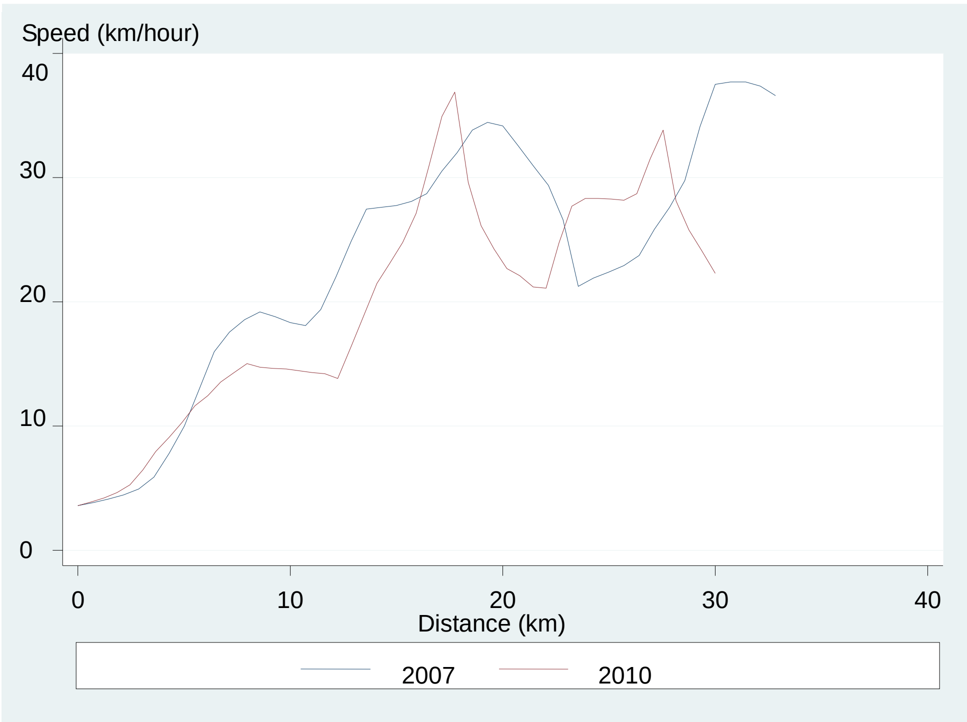
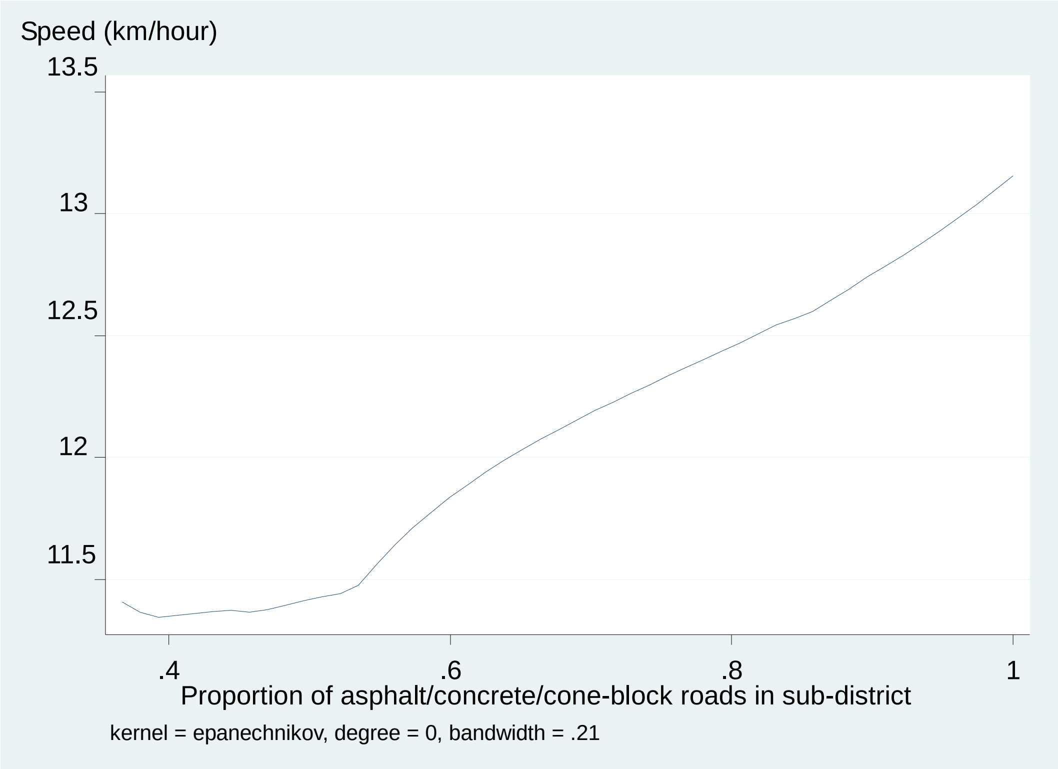


Figure 4 Relationship between non-agriculture speed (km/hour) in 2007 and (a) the proportions of asphalt/concrete/cone-block used in inter-village roads and (b) the proportion of 4-wheel passable in inter-village roads within sub-district (PODES 2006)



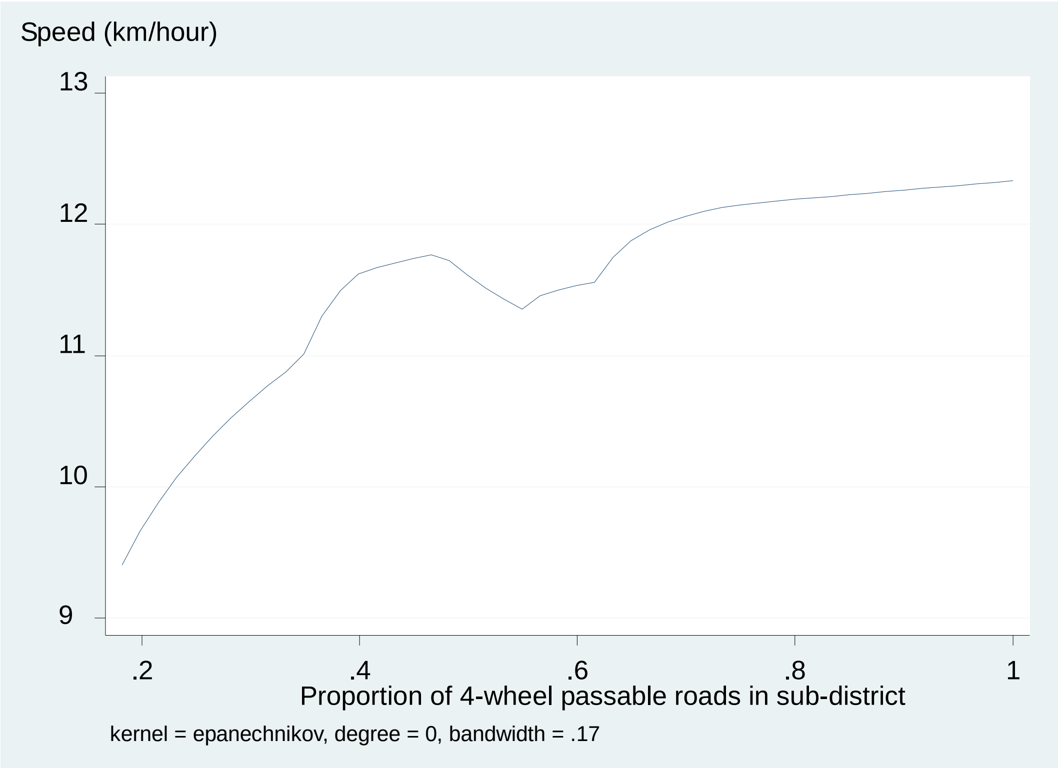


Figure 5 Histogram of change in (log) mean village speed (km/hour)

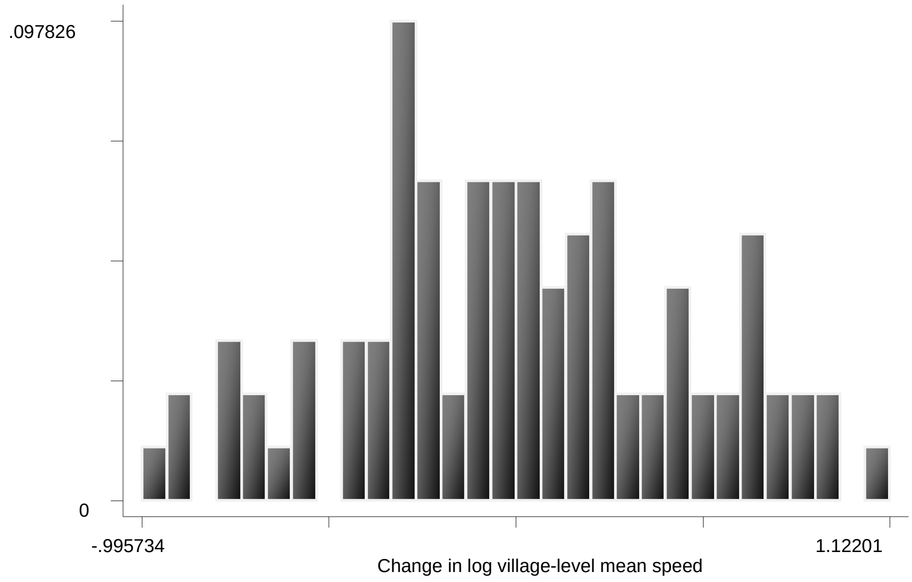


Figure 6 Density of intra-village schooling distributions with and without road projects

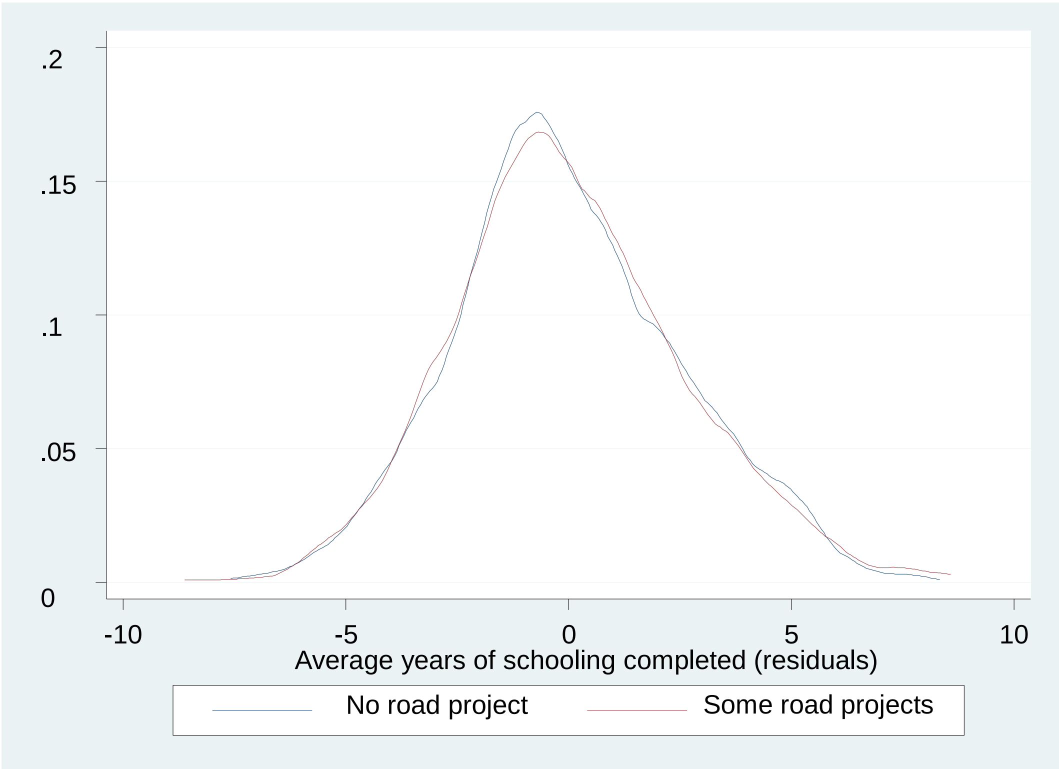


Table 1 Asphalt/concrete/cone-block road proportion in inter-village roads (province-wise average)

Province code	1996	2006
Aceh	0.46	0.39
Sumatera Utara	0.49	0.53
Sumatera Barat	0.69	0.93
Riau	0.40	0.48
Jambi	0.61	0.74
Sumatera Selatan	0.63	0.69
Bengkulu	0.74	0.73
Lampung *	0.52	0.47
Jakarta Raya	0.99	1.00
Jawa Barat	0.69	0.66
Jawa Tengah *	0.64	0.74
Yogyakarta	0.81	0.79
Jawa Timur *	0.56	0.68
Bali	0.98	0.99
Nusa Tenggara Barat*	0.82	0.78
Nusa Tenggara Timur	0.44	0.40
Kalimantan Barat	0.41	0.47
Kalimantan Tengah	0.36	0.44
Kalimantan Selatan *	0.63	0.67
Kalimantan Timur	0.32	0.49
Sulawesi Utara *	0.76	0.76
Sulawesi Tengah	0.58	0.63
Sulawesi Selatan *	0.50	0.60
Sulawesi Tenggara	0.52	0.55
Maluku	0.57	0.64
Maluku Utara	0.25	0.44

Sources: PODES 1996 and 2006 (Yamauchi, et al. 2011, Table 7). * marks the sample provinces.

Table 2 Split decision

Dependent: = 1 if split, and 0 otherwise		
Change in log speed (vlg)	0.1470	
	(1.09)	
Change in log speed (vlg) * average yrs of schooling	-0.0067	-0.0049
	(1.49)	(1.08)
Change in log speed (vlg) * average age (age 20-55)	-0.0033	-0.0015
	(1.24)	(0.57)
Change in log speed (vlg) * household size (age 20-55)	0.0096	0.0100
	(0.53)	(0.49)
Average years of schooling	-0.0074	-0.0085
	(0.98)	(0.97)
Years of schooling squared	0.0001	0.0003
	(0.24)	(0.67)
Average age (age 20-55)	-0.0049	-0.0041
	(3.96)	(3.67)
Household size (age20-55)	0.0643	0.0677
	(4.88)	(6.16)
Province dummies	yes	
Village dummies		yes
Number of observations	1904	1904
R squared	0.0778	0.1443

Numbers in parentheses are absolute t values using robust standard errors using province clusters (Column 1) or village clusters (Column 2).

Table 3 Road projects in 2007 - 2010

Number of road projects completed	Villages	Frequency
0	38	38.78
1	35	35.71
2	16	16.33
3	4	4.08
4	4	4.08
6	1	1.02

Table 4 Road projects and change in log speed: village-level data

Dependent:	Change in log mean village-level speed		Number of road projects
Number of road projects completed	0.2525	0.2729	
	(2.29)	(2.69)	
Log speed in 2007	-0.3139	-0.3260	-0.1269
	(3.07)	(3.20)	(0.56)
Road projects * log speed 2007	-0.1033	-0.1117	
	(1.76)	(2.02)	
Sub-district proportion asphalt/concrete/cone-block		-0.2193	0.7541
		(1.08)	(1.80)
Sub-district proportion 4 wheel		0.6457	-0.0778
		(2.32)	(1.10)
Constant	0.6448	0.2187	0.8424
	(3.16)	(0.88)	(1.20)
R squared	0.2423	0.2744	0.0397
F stat	12.22	9.47	1.29
Prob > F	0.0000	0.0000	0.2817
Number of observations	92	92	92

Numbers in parentheses are absolute t values using robust standard errors. The access speed for either non-agricultural or agricultural employment could not be calculated for 6 villages, thus not enabling to compute the mean village-level speed.

Table 5 Change in working time

Dependent:	Non-agricultural Change in total months		Agricultural Change in total days	
	No IV	IV	No IV	IV
Change in log speed (vlg) * average yrs of schooling	0.2918	-0.1319	-6.1628	-11.031
	(1.50)	(0.23)	(2.64)	(1.16)
Change in log speed (vlg) * average age (age 20-55)	0.0912	-0.0265	-1.6961	1.2959
	(1.35)	(0.09)	(1.61)	(0.42)
Change in log speed (vlg) * household size (age 20-55)	-0.6118	-1.9355	-3.0854	-14.780
	(1.29)	(1.51)	(0.33)	(1.07)
Average years of schooling	-0.2791	-0.2511	5.2784	6.1037
	(0.92)	(0.80)	(1.13)	(1.25)
Schooling squared	0.0222	0.0218	-0.1741	-0.2075
	(1.17)	(1.11)	(0.72)	(0.83)
Average age	-0.0056	0.0025	1.3305	1.2364
	(0.17)	(0.07)	(2.57)	(2.21)
Household size	0.1653	0.3054	0.5806	1.4742
	(0.68)	(0.09)	(1.11)	(0.29)
Village dummies	yes	yes	yes	yes
Durbin-Wu-Hausman: Chi-sq (3) [p-value]		1.219 [0.748]		1.186 [0.756]
Number of observations	1943	1943	1943	1943
R squared	0.0770	0.0704	0.1136	0.1063

Numbers in parentheses are absolute t values using robust standard errors using village clusters. Instruments are: indicators for the number of road projects completed interacted with years of schooling, years of schooling squared, average age and the number of household members aged 20-55 in 2007, and other exogenous variables.

Table 6 Change in monthly/daily income

	Non-agricultural employment Monthly wage		Agricultural employment Daily wage	
	No IV	IV	No IV	IV
Change in log speed (vlg) * average yrs of schooling	30509.71	100304.9	3626.14	8086.12
	(1.97)	(1.91)	(2.30)	(1.91)
Change in log speed (vlg) * average age (age 20-55)	-5667.51	-94143.59	-109.93	-155.62
	(0.62)	(2.21)	(0.23)	(0.19)
Change in log speed (vlg) * household size (age 20-55)	-21347.45	-176045.5	-2563.98	-975.01
	(0.39)	(0.88)	(1.02)	(0.17)
Average years of schooling	-28959.77	-49664.92	-583.56	-1170.06
	(1.0)	(1.38)	(0.32)	(0.58)
Schooling squared	3134.57	4217.41	-47.31	-22.93
	(1.75)	(2.01)	(0.39)	(0.19)
Average age	3226.41	7297.64	-76.62	-75.94
	(0.50)	(1.01)	(0.48)	(0.42)
Household size	6845.74	28235.47	-1131.18	-1337.00
	(0.27)	(0.86)	(1.00)	(0.86)
Village dummies	yes	yes	yes	yes
Durbin-Wu-Hausman: Chi-sq (3) [p-value]		10.428 [0.015]		1.586 [0.663]
Number of observations	1943	1943	1943	1943
R squared	0.0641	n.a	0.070	0.0619

Numbers in parentheses are absolute t values using robust standard errors using village clusters. Instruments are: indicators for the number of road projects completed interacted with years of schooling, years of schooling squared, average age and the number of household members aged 20-55 in 2007, and other exogenous variables.

Table 7 Change in working time: Instruments using road projects completed before 2009

Sample: Excluding villages where projects completed in 2009/2010	Non-agricultural employment Monthly wage		Agricultural employment Daily wage	
	No IV	IV	No IV	IV
Change in log speed (vlg) * average yrs of schooling	0.1090	-0.2140	-3.9358	-8.5804
	(0.42)	(0.39)	(1.21)	(0.95)
Change in log speed (vlg) * average age (age 20-55)	0.0671	-0.1448	-0.5199	0.2782
	(0.90)	(0.46)	(0.47)	(0.11)
Change in log speed (vlg) * household size (age 20-55)	0.0664	-1.4875	-7.0564	-29.400
	(0.12)	(1.15)	(0.47)	(1.84)
Average years of schooling	-0.1852	-0.2097	3.7962	3.1914
	(0.58)	(0.61)	(0.66)	(0.51)
Schooling squared	0.0183	0.0208	-0.1247	-0.0710
	(0.87)	(0.92)	(0.40)	(0.21)
Average age	-0.0126	-0.0008	0.9915	1.0309
	(0.36)	(0.02)	(1.60)	(1.59)
Household size	0.2833	0.4439	6.1284	7.9809
	(1.0)	(1.25)	(0.78)	(1.05)
Village dummies	yes	yes	yes	yes
Durbin-Wu-Hausman: Chi-sq (3) [p-value]		2.300 [0.5124]		1.825 [0.6096]
Number of observations	1335	1335	1335	1335
R squared	0.0696	0.0615	0.1204	0.1135

Numbers in parentheses are absolute t values using robust standard errors using village clusters. Instruments are: indicators for the number of road projects completed interacted with years of schooling, years of schooling squared, average age and the number of household members aged 20-55 in 2007, and other exogenous variables.

Table 8 Change in monthly/daily income: Instruments using road projects completed before 2009

Sample: Excluding villages where projects completed in 2009/2010	Non-agricultural employment Monthly wage		Agricultural employment Daily wage	
	No IV	IV	No IV	IV
Change in log speed (vlg) * average yrs of schooling	35318.79	130982.8	1881.07	3952.87
	(1.53)	(2.50)	(1.37)	(1.25)
Change in log speed (vlg) * average age (age 20-55)	-15855.6	-63117.8	-659.69	-895.51
	(1.49)	(1.96)	(1.01)	(1.03)
Change in log speed (vlg) * household size (age 20-55)	6577.60	-254052.5	-2030.39	-7982.88
	(0.11)	(1.76)	(0.70)	(1.49)
Average years of schooling	3343.27	-10305.4	-608.82	-968.04
	(0.09)	(0.28)	(0.28)	(0.41)
Schooling squared	1406.97	2016.6	1.6495	19.330
	(0.57)	(0.78)	(0.01)	(0.14)
Average age	5359.26	7690.55	27.929	52.673
	(1.27)	(1.47)	(0.14)	(0.25)
Household size	-4804.77	23406.3	136.31	680.26
	(0.18)	(0.70)	(0.10)	(0.39)
Village dummies	yes	yes	yes	yes
Durbin-Wu-Hausman: Chi-sq (3) [p-value]		11.779 [0.008]		0.976 [0.8069]
Number of observations	1335	1335	1335	1335
R squared	0.0687	0.0102	0.0494	0.0456

Numbers in parentheses are absolute t values using robust standard errors using village clusters. Instruments are: indicators for the number of road projects completed interacted with years of schooling, years of schooling squared, average age and the number of household members aged 20-55 in 2007, and other exogenous variables.

Table A1 Road projects and working time

Dependent:	Change in months Non-agricultural	Change in days Agricultural
The number of road projects	-0.0295	-1.7147
	(0.12)	(0.42)
Average years of schooling	-0.4627	4.2634
	(1.57)	(1.99)
Years of schooling squared	0.0289	-0.2003
	(1.58)	(1.20)
Average age (age 20-55)	-0.0068	1.1178
	(0.13)	(2.67)
Household size (age20-55)	0.1281	0.1002
	(0.87)	(0.03)
Province dummies	yes	yes
Number of observations	2080	2080
R squared	0.0102	0.0193

Numbers in parentheses are absolute t values using robust standard errors using province clusters.