

Rural Roads and Development: Evidence from Ethiopia

Kifle A. Wondemu^a and John Weiss^b

University of Bradford, Bradford, UK

Previous empirical studies on Ethiopia identified adequate access to rural roads as one of the key determinants of rural income. However, most of these studies ignored the effect of spatial dimension of road infrastructure on rural income. As a result, the productivity of road reported by these studies is likely to be underestimated by the extent that local equilibrium prices adjust to the stock of local road infrastructure. This paper attempts to partially address the problem by specifying the link between rural roads and household income in the context of a spatial equilibrium framework. The results show that road-induced rural income growth is substantially higher than what was reported by previous studies that used the same dataset. Road-induced factor productivity and returns to land and labour are also found to be the main channels by which better road access enhances rural income.

Key Words: Ethiopia, Rural, Household, Income and Roads

1. Introduction

In recent years, considerable attention has been given to the role of infrastructure in promoting economic growth and the Aid community has identified this sector as a priority for funding. Rural road projects are often considered as critical income earning opportunities for the poor by integrating them into regional and national markets. Whilst this view of the benefits of road investment is widespread, quantification of the benefits is often difficult and raises important methodological issues. The conventional methods of road project appraisals measure benefits in terms of vehicle operating cost savings for normal and diverted traffic plus some often vague allowance for generated traffic to pick up induced production effects. Such approach not only lacks theoretical rigour, but is also usually associated with inadequate quantification of the induced benefits. Recent attempts to improve upon such an approach include the use of structural and reduced form econometric models. For example, Menon and Warr (2008) apply a

^a Richmond Road, BD7 1DP, UK, T: +441274945898, E: kasfaw@yahoo.com/kiflea@gmail.com

^b Richmond Road, BD7 1DP, UK, T: +441274235274, E: J.Weiss@Bradford.ac.uk

CGE model approach; Fan and Chan-Kang (2005) apply a simultaneous equation approach; Van der Walle and Mu (2007) apply a double-difference and propensity score method and Khandker et al (2009) use a double-difference household-level fixed effects regression model¹.

A key limitation of the most recent attempts is that most of these studies assessed impacts by ignoring the spatial dimensions of the effects of road infrastructure (Anriquez and Valdes, 2006; Dercon, 2002; Dercon, et al. 2007). To the extent that local equilibrium prices adjust to the stock of local road infrastructure, the productivity of road reported by these studies is likely to be underestimated. This paper attempts to contribute to the debate by partially circumventing such a limitation, and measuring the impact of roads on rural income by specifying the link between the two in the context of a spatial equilibrium framework. The use of such a framework will help better capture the dynamic impact of roads on income as in such framework income, prices, employment and input demand are a function of the stock of infrastructure and vary in response to change in the stock of local infrastructure (Costa, 1998; Haughwout 1998, 2002).

For the purpose, while we use household level data from a number of villages in Ethiopia, we postulate a reduced form demand for road infrastructure and derive the impact of road quality on rural income as an area below the demand curve. The result reveals an impact that is very substantial and which appears to support the view that rural roads can in some circumstances be an important catalyst for development. The paper is organized as follows. In the next section, we describe the data. The third section outlines the theoretical framework the empirical model will base on. After the estimation strategy is discussed, the estimation results will be presented. The final section concludes.

2. The Data

Ethiopia is a land locked country where the major share of passenger and freight movement is by means of road transport and where the transport network is recognised as a major bottleneck. In the 1990's, due to civil war, financial constraints and limited capacity for planning and maintenance, much of the road infrastructure deteriorated. Recognizing the seriousness of the problem, since 1997 the government launched a road sector development program with the aim of expanding the road density to 67, 300 km as well as increasing the share of good quality roads from its level of less than 50% at the start of the program to 65% by the end of 2015. Despite recent improvements the road density still stands at about 33 km/1000 km² for the entire classified road network and around 22km/1000 km² for roads in good or fair condition. This is well below the average road density of sub-Saharan countries. Moreover, since over 90% of the road network covered under the program is accounted by federal roads, which are inter-regional trunk roads, a program to expand the rural road network has only been launched recently (MoFED, 2002). Although there has been a significant increase in rural roads construction, most of these roads are dry weather roads, with less improvement in all-weather roads, and as a result, substantial share of the rural population has yet to travel about six hours to reach a road. This may have far-reaching implications for grain marketing, crop prices, and on the rate of rural poverty reduction (World Bank, 2004).

The data for this study is drawn from the longitudinal Ethiopian Rural Household Surveys (ERHS) that were conducted by Addis Ababa University jointly with International Food Policy Research Institute (IFPRI) and the University of Oxford. The surveys were conducted in villages drawn from Amhara, Oromiya, Tigray and the Southern Nations and Nationalities Regional states. The survey is claimed to be broadly representative of rural households that are in sedentary agriculture (Dercon and Hoddinott 2004). Although the survey covers both the ox and

¹ Van Der Walle (2009) surveys approaches using reduced form models.

the hoe-plough agriculture systems, the present study focuses on households that are under ox-plough agriculture. While the total number of households surveyed under ox-plough agriculture was 1200, due to missing data and reporting inconsistency, only the data of 841 households could be used. Similarly, although the survey on these villages had been undertaken for six rounds, the required village level data however is only available for the 1997, 1999 and 2004 round surveys. This paper will therefore only use the survey of these years.

The household survey collects quantitative information on a set of variables that capture both household characteristics and geographic characteristics of the villages, including average size of land holdings, number of oxen owned, fertilizer input usage per hectare, average age and years of schooling of household head, number of household members, altitude above sea level, average temperature and distance from the village to the nearest town.

In addition, the survey data includes qualitative information on respondents' perceptions of the magnitude and reliability of rainfall, the quality of soil in the area and on the accessibility of the villages at different times of the year. In the survey, four questions are related to the reliability of rainfall, its magnitude, the time it stopped, and whether it was raining at harvest time. An answer to each question that implies a negative impact on production is given a score of 1, and the total for all four is added to give a 'rainfall shock index.' Thus, if a household experienced all shocks it has a score of 4 and if it experienced none its score is zero. Similarly respondents are asked about soil quality. Here three possible values are used in a 'soil quality index', where 0 is for poor quality, 1 for semi fertile and 2 for top soil.

A key aspect of the study is a measure of the quality of road access. In the survey, the road quality of the sampled villages was compiled through a structured community level questionnaire. Community leaders were asked to attach a value of 1 to 6 depending on how the road allows accessibility to and from the village during the rainy season. Respondents attached a value of 1 to 6 respectively for a road that allows easy access to any vehicles (6), reasonable access to any vehicles (5), good access to trucks and buses (4), reasonable access to trucks and buses (3), access to carts and animals (2) and finally access only for foot traffic (1). Due to inconsistency regarding the reported road quality of the villages, namely between what is reported under the qualitative and quantitative surveys of some of the villages, and partly to compare our results with the result of similar studies that used the same data set, the road quality of the villages is categorized into two groups (Dercon, et al. 2007). The first one is 'good road access' and covers roads that allow accessibility from any vehicle to good access to trucks and buses. The second one is 'poor road access' and it represents roads that allow reasonable access only to trucks and buses and to foot traffic. Therefore, while estimating the empirical model, a value of 1 is given for villages that have good road access and 0 for villages with poor road access.

The sample villages, as can be observed from the table in Appendix 1, also differ substantially from each other not just in road access, but also in relation to aspects like average income, size of land holding, soil quality, temperature, rainfall intensity, altitude and use of fertilizer. Any attempt to isolate the impact of road access must control adequately for these differences. In order to establish whether we can sensibly compare outcomes across villages with different road access, we calculated the overlapping coefficients for key determinants of rural households' income (Goldstein 1994). The coefficient measures the degree of overlap between the distributions of these variables across different road access. The magnitude of the coefficient ranges between zero and one; and it is one if the distributions of the variables compared are completely alike or overlap, and zero if they are completely different or if there is no overlap between them. In order to generate the coefficients, we first conducted a variance ratio test to determine whether the variance of each variable under different road access differs. The test shows that the differences are not statistically significant. Next assuming equal variances, we calculated the overlapping coefficient for each variable by splitting the sample into poor and

good road access. While the average coefficient is 0.82, the level ranges from 0.61 for size of land holding to 0.93 for a number of adults in the household. Since these coefficients are significantly different from zero, it suggests that villages or households that have different road access share similar characteristics that influence income. We also supplemented this by plotting the probability distribution of the variables that determine households' income (see Appendix II). As can be observed from the diagrams, there are significant overlaps between the distributions of these variables, suggesting that one can sensibly compare outcomes across different road access.

3. Farm Household Model

In order to empirically estimate the impact of road access on rural income, we use a farm household model as our theoretical framework (Jacoby and Minten, 2009). Here, the sampled households are assumed to maximize a utility function (U) defined over net revenue (M) and leisure (l). The income of the household is assumed to come from crop production and off-farm employment. The level of crop production (Q) is also considered to be a function of fixed inputs K (such as land, oxen and farm implements), labour (L), commercial fertilizer (F) and other fixed household (Z) and village characteristics (V). Each crop requires a different level of transport service for its production and marketing.

Selling crops entails a transport cost (τ) that varies by distance and the quality of the road that connects the farm household with the main nearest market. The effective (farm gate) price farm households receive for the crops they supply, therefore, will be $P_f = P_M - \tau_j$ where, P_M is the price of a similar crop at the terminal market. Similarly, while fertilizer is assumed to be the only commercial input purchased from a town, the effective price a household pays for fertilizer inputs will be $\alpha = \omega + \tau_j$, where, ω is the market price of fertilizer at the source of supply.

For the purpose of this analysis, it is assumed that there is no active land market and the size of land the household owns is fixed. Although there is an active labour market, the numbers of days household members can work off-farm is assumed to be limited to H days and from the total labour time available (T), L units are used on the family farm, N units are hired out for off-farm employment and the balance is assumed to be used for leisure activities, (l). Family and hired labour are also assumed to be perfect substitutes. For simplicity, the market wage rate (w) is assumed to be unaffected by the level of transport cost. However, the magnitude of transport cost, mainly through its effect on farm productivity, is assumed to have an impact on the availability of off-farm employment opportunities.

The utility maximization problem of a representative household can, therefore, be set up as:

$$\text{Max } U(M, l) \tag{1}$$

The household is assumed to maximize the above utility function subject to the following constraints:

$$M = P_f Q + wN - \alpha F \text{ (income constraint assuming a labour market constraint)} \tag{2}$$

$$Q = f(K, L, F, Z, V) \text{ (production technology constraint)} \tag{3}$$

$$T = l + L + N \text{ (time constraint assuming a labour market constraint)} \tag{4}$$

$$N < H \text{ (labour market constraint)} \tag{5}$$

Setting up the Lagrangian (\mathcal{L}) for the household's maximization problem yields:

$$\text{Max?} = U[M, l] + \lambda [P_f f(K, L, F, Z, V) + wN - \alpha F - M] + \mu_1 (T - L - N - l) + \mu_2 (H - N) \quad (6)$$

Taking the first order condition and only considering those variables that appear in the net-income function, the equilibrium level of the choice variables, i.e. the level of net-income (M), the total labour time spent on own farm (L), the number of hours worked off-farm (N) and the level of fertilizer used (F), will be:

$$M^* = P_f f(K, L^*, F^*, Z, V) + wN^* - \alpha F^* \quad (7)$$

$$L^* = f(P_f, \alpha, w, K, T, Z, V, H) \quad (8)$$

$$N^* = f(P_f, \alpha, w, K, T, Z, V, H) \quad (9)$$

$$F^* = f(P_f, \alpha, w, K, T, Z, V, H) \quad (10)$$

Since our interest here is to determine the factors that shift the net income position of the household, assuming that factors such as land size and quality as well as oxen inputs are fixed in the short term, equation (7) can be totally differentiated and becomes:

$$\frac{dM}{d\tau} = f(\bullet) \frac{\partial P_f}{\partial \tau} + (P_f f_F - \alpha) \frac{\partial F}{\partial \tau} - F \frac{\partial \alpha}{\partial \tau} + w \frac{\partial N}{\partial \tau} \quad (11)$$

Where, $f(\bullet) = f(K, L, F, Z, V) = Q$, $f_F = \frac{\partial f(\bullet)}{\partial F}$

From equation (11), it is clear that, depending on the position of the household in the goods and factor markets, a change in transport cost affects the short run net income position of the household through its influence on output price, the level of fertilizer demand, the cost of fertilizer and the number of household members that can be engaged in off-farm employment². Assuming zero income effect of leisure consumption, the reduction in transport cost, therefore, rotates the budget constraint and thereby allows the household to attain a higher level of welfare.

In order to measure the implied welfare improvement in monetary term, following Jacoby and Minten (2009), the concept of equivalent/compensating variation can be used. Accordingly, let $\mu(\tau, K, T, Z, V)$ represents the additional income level that would make the households that reside in villages with good road access indifferent to market situations that prevail in villages with poor road access. The magnitude of the compensation required, which should be at least equal to the additional revenue the household will forgo due to the market prices it would face in villages with poor road access, depends on the level of the transport cost, the level of household endowments and other geographic characteristics. Given the indirect objective function of the household, i.e., $v[(P_f, \alpha, w, T, K, Z, V)]$, the equilibrium condition can be implicitly defined as:

$$v[(P_f^D, \alpha^D, w^D, T, K, Z, V) + \mu(\tau, T, K, Z, V)] \equiv v[(P_f^A, \alpha^A, w^A, T, K, Z, V)] \quad (12)$$

Where the superscripts D and A respectively represent market prices that prevail in villages with poor and good road access.

Differentiating equation (12) with respect to transport cost (τ) and applying the envelope theorem, gives

² The impact of a road will be particularly high if the marginal value of fertilizer input is greater than its marginal cost, i.e. if the input is under-utilized.

$$\frac{\partial V}{\partial \tau} \equiv \mu_{\tau}(\tau, K, T, Z, V) - Q - F \equiv 0 \quad (13)$$

Letting $m(\tau) = Q + F$ and rearranging the terms, we have

$$\mu_{\tau}(\tau, K, T, Z, V) = m(\tau) \quad (14)$$

As argued by Jacoby and Minten (2009), $m(\tau)$ resembles the composite commodity theorem where a group of commodities that have constant relative prices can be treated as a single commodity, and when applied to the present case, the term can be assumed to represent the demand for freight transport where τ acts as a price. The demand for transport services in this context covers transporting crops (Q) to market and purchased fertilizer input (F) from the market. The higher the transport cost, the lower the demand for transporting crops to higher priced but distant markets and hence the household will receive a lower price for the crops it supplies. Similarly, the higher the transport cost, the lower the transport demand to move fertilizer from its source of supply to the farm and hence the household either applies a lower quantity of fertilizer or pays a higher price, which in both cases reduces the net-income of households.

Considering equation (14) as a partial differential equation in μ and integrating both sides at the average transport cost level that prevails under poor and good road conditions, the benefit of transport cost reduction due to road improvement becomes,

$$\mu(\tau_1, K, T, Z, V) = \int_0^1 m(\tau) d\tau \quad (15)$$

Equation (15), which represents the transport demand function of a representative household, can be integrated at 0, which represents the average transport cost households incur in village with good road access, and at 1, which represents the average transport cost farm households incur in villages with poor road access. As the objective is to generate the average benefit of road improvement on net income, equation (15) becomes:

$$E[\mu(\tau_1)] = \int_0^1 m(\tau) d\tau \quad (16)$$

Assuming an infinitely elastic labour supply at the village level, although we previously assumed that a reduction in transport cost will not influence the market wage rate, the change in transport cost however was acknowledged to increase village level farm productivity. The improvement in farm productivity, by shifting the demand for labour and expanding employment opportunities, is expected to increase off-farm employment income of the household. In order to accommodate that in the benefit measure, following Jacoby and Minten (2009), the difference in the average off-farm income earned in villages with good [$e(0)$] and poor road access [$e(1)$] can be added to equation (16) so it becomes

$$E[\mu(\tau_1)] = \int_0^1 m(\tau) d\tau + E[e(0)] - E[e(1)] \quad (17)$$

Equation (17) serves as the theoretical model on which the empirical analysis is based. After controlling for the effect of other household and village specific factors, the road quality dummy in the net revenue model is assumed to approximate the benefit level implied by equation (17), i.e., the magnitude of income increase farm households could gain due to an improvement in the quality of a rural road that allows all weather road access.

4. Empirical Model Specification

Based on the first order condition, the reduced form of the net-revenue function for a given farm household becomes:

$$M_{hjt} = f(P_{fjt}, w_{jt}, \alpha_{jt}, K_{hjt}, T_{hjt}, Z_{hjt}, V_{jt}) \quad (18)$$

Where h represents household, j village, t time, K the size of the household's land holding, T total labour time available (in terms of total number of adult productive family members), Z other household level characteristics and V village level characteristics.

Assuming that the local goods and factor markets are spatially integrated, spatial equilibrium requires that farm households cannot get excess net revenue because of their location (Haughwout 1998, 2002). As a result, the zero-profit spatial equilibrium condition for households in village j will be:

$$C_{ijt}(K_{hjt}, Z_{hjt}, V_{jt}, w_{jt}, \alpha_{jt}) = P_{fjt} \equiv P_m(1 - \tau) \quad (19)$$

Where i, j and t respectively represent the type of crop, village and time; C is the marginal cost of crop production in village j; P_m is the crop price at the main market and τ is the magnitude of the transfer cost traders incur to move crops and fertilizer across spatial markets.

Equation (19) states that at equilibrium the marginal cost of crop production at the surplus (deficit) market should equal to the terminal market price less (plus) the transfer cost that traders incur in transporting the crop between markets. In a reduced form, this implies that local demand and production conditions, the price level of similar crops at the terminal market and the level of transfer cost are the main determinants of local market equilibrium prices. Following other empirical studies, the magnitude of transfer cost (τ) can be assumed to be a function of the quality of road infrastructure (Rd) and the distance of the village from the nearest spatial market (Dm) (Minten, 1999). Following similar reasoning for fertilizer price, but assuming that local wage formation is more influenced by local conditions, the local market equilibrium prices for crops, labour and fertilizer would then become a function of the level of transfer cost as well as local production and demand shifting factors. The level of transfer cost is also assumed to be a function of the road quality of the village, the distance of the village from the nearest spatial markets as well as other factors that alter the spatial market condition for crops and fertilizer input. Assuming that the impacts of spatial market prices (P_m and ω) on household net-income are governed by distance and road quality of the village, the reduced form of equation (18) becomes³:

$$M_{hjt} = f(RD_{jt}, DM_j, K_{hjt}, T_{hjt}, Z_{hjt}, V_{jt}) \quad (20)$$

Where RD is the road quality of the village, DM is distance of the village from a town and the remaining variables are as defined above.

The road variable in the above function is expected to capture the effect of road quality on net income that works through shifting the production function, altering farm gate price ratios and influencing the resource allocation and crop choice behaviour of a farm household. The vector of household specific characteristics (Z) is assumed to include the number of oxen the household owns (Ox), the schooling (Ed) and age of the household head (Age), the soil quality of the land the household cultivates (SQ), whether or not the household has participated in government

³ Minten (1999) postulated that local prices are a function of the terminal market price, distance and infrastructure, however when he empirically estimated the model, he removed the terminal market assuming that the strength of the link between the two prices is determined by state of infrastructure and distance alone.

sponsored extension services and the frequency of rainfall shocks the household has experienced (R) (Rahman, 2003; Parikh, et al, 1995). The population size of the village (PoP) is also included as a demand shifting factor as well as the determinants of income through its influence on labour market opportunities and on the average per-capita land size households can have access to.

The sampled farm households cultivate more than one crop and two identical households could generate different levels of income per unit of land and labour employed just because they cultivate different types of crops. In order to control for the effect of cropping patterns on income, since teff is the crop that fetches the highest price, and also the crop that is not uniformly cultivated in all sampled villages, a dummy variable is introduced and a value of 1 is given to farmers that have cultivated teff and 0 otherwise.

Finally, assuming that the production function takes a Cobb-Douglas form, where output has a logarithmic relationship with some and a semi-log relationship with the other household and geographic characteristics, and also assuming that the effect of distance from the market is captured by the household specific effect term, the estimable empirical model becomes⁴:

$$\ln M_{ijt} = \alpha_0 + \beta_1 \ln Ld_{ijt} + \beta_2 \ln Ox_{ijt} + \beta_3 \ln Lb_{ijt} + \beta_4 R_{ijt} + \beta_5 Rd_{jt} + \beta_6 SQ_{ijt} + \beta_7 \ln Ed_{ijt} + \beta_8 \ln Age_{ijt} + \beta_9 \ln PoP_j + \beta_{10} Teff_{ijt} + v_{ij} + \varepsilon_{ijt} \dots\dots\dots(21)$$

Where M is the net income of household i in village j at time t , (measured as income from crop production and off-farm employment less expenditure on hired labour and fertilizer inputs), Ld is the land size in hectare, Ox is number of oxen the household owns, Lb is the number of adult productive household members, R is the frequency of rainfall shock, Rd is dummy for the road quality of the village, SQ is the soil quality dummy of the cultivated land, Ed and Age respectively are the schooling and age of the household head, PoP is the population size of the village and $Teff$ is a dummy for teff cultivation.

While the results of the above model would inform us whether or not roads increase rural income, the results do not allow us to draw conclusions as to whether providing better road access increases the income of the lower income quantile faster than the upper income quantile. In order to address such distributional issue, as we have no baseline data for households with good road access, we use the propensity score matching technique⁵ (Brand and Xie, 2007). The land size that the household cultivates, which is the only variable that meets the balancing requirement, is used as a control variable to generate the propensity score of each household (Becker and Ichino, 2002). The estimated propensity score is used to classify households into different stratum and for each stratum, including the other covariates of income, the average impact of a road on income is estimated. The estimation result provides information as to whether the impact of road access on income is heterogeneous across the income distribution.

5. Estimation Strategy

As three years of panel data is used to estimate the net-income model, the use of panel data model estimators is necessary. The presence of fixed effects in the data was tested based on a Lagrange multiplier test proposed by Breusch and Pagan (1980), and this test confirms the presence of fixed effects at the 5% level. The Hausman (1978) specification test also shows that the

⁴ A number of interaction terms, as well as variables that reflect village level initial conditions, are also included when estimating the model but for brevity they are not included here

⁵ The propensity score matching technique allows us to match and compare households that have good road access with those that have poor road access but have similar characteristics before the formers had had good road access.

use of the fixed effect model is appropriate for the data. The presence of heteroskedasticity and autocorrelation is also confirmed at less than the 1% level⁶.

Although all of the above deviations will not lead to biased coefficient estimates, they would understate the standard error estimates and thus could lead to biased statistical inference. In order to address the problems and ensure the validity of our results, which are robust to heteroscedasticity and general forms of cross-sectional and temporal dependence, we estimate the fixed effect model with the Driscoll and Kraay (1998) standard errors (Hoechle, 2007).

Further, while the unit of analysis is the household, the presence of village level variables, which are constant within a village, makes the regressors and hence the random error term of each household within the village to be correlated. As a result, the standard error estimates are likely to be underestimated and hence renders the inference on the basis of the t-statistics unreliable (Kloek 1981; Moulton 1990). In order to see if there is bias due to the cluster effect, although we did not report the result here, we implemented the Moulton (1990) proposed correction factor⁷. The result confirmed the presence of intra-village correlation, but the road variable is still positive and significant⁸.

All studies such as this risk 'placement bias', where a road infrastructure variable is endogenous, because higher income potential and existing infrastructure in an area influence decisions on road improvements or the location of new roads. If that is the case, as these factors influence the productivity of a road and also since these factors are subsumed in the error term, the coefficient estimate of the road variable is expected to be overestimated as it picks up the effect of these local specific factors.

However, it is questionable whether placement bias is a serious issue for this sample. The sample villages that have all-weather road access are the one that are closest to a trunk road or where a trunk road crosses such villages. The investment on these roads was undertaken in the past (most of the roads were constructed many years ago)⁹. As a result, as argued by Fan and Chan-Kang (2005) and Fan, et al, (2002) in the context of China, reverse causality should not exist as the road stock is likely to be unrelated to the current or recent past output level of the villages. Moreover, even if it is argued that such roads are placed in politically or economically important regions, where such villages are located, this impact will be captured in the village specific effect and hence the use of a fixed effect (within) estimator will be sufficient¹⁰. Using the same data set, Dercon, et al, (2007) also addressed the endogeneity of road by using the fixed effect estimator.

⁶ Moreover, due to data problems (mainly to determine the spatial weight matrix) we could not prove the presence of spatial correlation, although we suspect its presence as all sampled villages are exposed to common rainfall or macro related shocks. Although we attempted to perform Pesaran's (2004) test of cross sectional dependence, we could not derive the test result as one or more observations are missing for a number of households

⁷ Unfortunately the Moulton proposed correction factor is implemented on OLS, which generates biased coefficient estimates. We have not found a panel data estimator that simultaneously takes into account inter-class correlation, heteroscedasticity and general forms of cross-sectional and temporal dependence.

⁸ As the Moulton correction factor is applied on OLS, the coefficient estimates reported are expected to be biased as correlation between the unobserved specific effects and the regressors is not accounted for. Among the explanatory variables, the intra-village correlation coefficients are the highest for road quality and village population size, which are respectively 0.72 and 0.64. Such high level of correlation is expected to make the error terms of each households within the village correlated and thus the standard error estimates will be biased.

⁹ From discussion with the designers of the original survey in Addis Ababa University, it does not appear that road improvement was based on the villages' growth potential or an induced demand for infrastructure.

¹⁰ Zhang and Fan (2001) tested the two directions of causality between productivity growth and road capital. To avoid the reverse causality of road development leading to productivity growth, they used an instrumental variable approach, and found that the coefficient of their road variable changed very little when compared with

As noted by Jalan and Ravallion (1998), initial village conditions generate positive or negative externalities and thus could significantly influence the productivity of private and public capital. Such initial conditions, through their interaction effects, can create increasing returns to capital and as a result the coefficient of the road variable could be biased unless the effects of these externalities are properly accounted for. Here we attempted to do that by introducing variables that can capture the initial condition of the villages and to that effect, since the main economic activity of the sampled villages is farming, we included the level of rainfall, temperature, altitude and distance of the village as variables that represent the initial condition of the village. However, since some of these variables are time invariant and also even for time varying variables, such as rainfall and temperature, the data we have is only for a single period, the fixed effect estimator cannot generate the parameter estimates of these variables. In order to ascertain whether the magnitude and sign of the coefficient estimate of the road variable changes when one controls for the effect of initial conditions of the village, we also estimate the income model using the Hausman-Taylor (1981) estimator. Although this estimator cannot address the correlation between the random error term and the explanatory variables, in addition to allowing us generate the coefficient estimates of time invariant variables, it addresses the endogeneity bias that arises due to the correlation between the village level effects and the explanatory variables¹¹.

6. Results

Table 1 gives the results of the net income model of equation (21). The model is estimated using the fixed effect and the Hausman-Taylor estimators. The results of the fixed effect estimator are robust to heteroscedasticity, temporal and spatial correlation. The results of the Hausman-Taylor estimator, on the other hand, are robust to endogeneity bias, which mainly arises due to the correlation between household specific effects and the road variable, as well as to a bias that could occur due to the interaction effect of initial conditions on the productivity of private and public capital. As the Hausman-Taylor estimator provides the coefficient estimate of time invariant variables, our discussion of results will focus on this estimator.

As the result table shows, under both estimators, albeit with a different degree of elasticity, the road quality of the village has a significant impact on rural income. On the basis of the fixed effect estimator, the result suggests that improving the quality of rural roads to a level that allows all weather road access raises average household income by as much as 63%. The estimated income improvement, however, is significantly lower under the Hausman-Taylor estimator, which is only 37%. Although the difference between the two estimators could be partly attributed to differences in distributional assumptions, high coefficient estimate of the road variable under the fixed effect estimator could be because this estimator does not take into account the role of initial conditions of the villages. Since initial conditions mediate the link between road and income, the fixed effect estimator will undoubtedly give an upward biased estimate of the productivity of road (Jalan and Ravallion, 1998).

The level of impact we reported here is also slightly higher than what is reported by Dercon, et al, (2007). While using the same survey data, except that their sample includes both households from Ox and Hoe-plough agriculture systems, they reported that providing all-weather roads access increases short run annual consumption growth by 15% and reduces the likelihood of a household being poor by 6-7%.

the original model. One of the reasons, Fan and Chan-Kang, (2005) argue, was that road capital, such as the current length of roads, was a result of past government investment.

¹¹ The other advantage of the estimator is that there is no need to identify exogenous variables that can serve as instruments. Instead, after they passed various tests, the explanatory variables included in the model are used to estimate their own coefficients and serve as instruments for endogenous variables.

One possible reason for the difference in the estimated impact could be partly due to difference in the modelling approach followed. In order to measure the productivity of road, they used a production function specification. Since the production function specification is normally based on the assumption that market prices are fixed or exogenously determined, if village level equilibrium prices endogenously adjust to the quality of road infrastructure, such specification would be inadequate to fully capture the income effect of road. Particularly it fails to capture the impact of road on income that comes through its influence on the level and volatility of spatial prices and subsequently, through the price mechanisms, by altering micro level decisions, such as how to and what to produce, that have significant impact on the level of income that can be generated from land and labour resources. Although the coefficient estimate of the road dummy is expected to measure its direct impact on income that comes through narrowing spatial price gaps, reducing price volatility and minimizing the adverse effect of geography, the significance of the interaction term of road quality and private resources (such as land and labour) suggests the presence and significance of such indirect impact.

From our result, important policy conclusions can be drawn. First, given that almost all rural households in Ethiopia are operating on their land frontier and also given that the option of pushing the land frontier outwards is very limited, the result could suggest that providing good quality rural roads could be one instrument by which government could counteract the adverse effect of a shrinking land size on rural income and welfare. This holds because better road access, by enhancing the profitability and reducing the market risk, enhances the adoption of land-augmenting technologies.

Similarly, a positive and significant sign of the interaction term for road quality and the number of adult family members also suggests that providing good road access enhances the productivity of family labour. This holds because good road access, by enhancing the application of fertilizer, which is a labour augmenting input, increases the farm productivity of family labour. Similarly, high degree of fertilizer application, by shifting the aggregate production at the village level and increasing village level demand for labour, expands off-farm employment opportunities and hence increases earnings from off-farm employment. For instance, using the same data set, Wondemu (2011) reported that, for a 10% increase in per-hectare fertilizer use, on average, the local wage rate increases by 1.6%. The magnitude of the impact however would have been higher had the fertilizer input were efficiently applied. In the sampled villages fertilizer input is employed 22% less than the optimal level and as a result the elasticity of the wage to fertilizer application is expected to be understated. After controlling for other indirect impact, he also reported that providing villages with good quality road access directly increases the local equilibrium wage rate by an average of 8%. The result therefore suggests that providing good road access could be one intervention that the government could implement to expand remunerative employment opportunity for land less and poor farm households even in the face of high population growth and growing land scarcity.

We also attempted to assess how different groups of households are affected by a change in road access conditions. Using the propensity score matching technique and classifying the households into different stratum, we estimated the income model¹². For each stratum, using a fixed effect estimator and properly accounting for the problem of heteroscedasticity and spatial and temporal correlations, the average impact of roads on income is estimated. The results, reported in Table 2, not only confirms the heterogeneity in the impact of roads, but also shows that the impact is regressive in that the productivity effect of road access on average is positively related to the land size that a household has access to. In fact, the impact of a road on income for households in the

¹² The key characteristic in determining the matching was size of land holding households have access to. This is a good proxy for income but one which is unlikely to be endogenous to the model.

lowest strata is nil, whilst it increases at an average of 17% as the rank of the propensity score stratum increases by one unit¹³.

Controlling for all factors that influence income, for households that are in the second strata, improved road access generates a 15% higher income compared to households in the same stratum that have poor road access. The return to road access is highest in the highest stratum, which means, for this group with the same level of land holding, having all-weather road access allows a representative farm household to generate 82% higher income than would be the case with poor access. It must be acknowledged however that these distributional results can be no more than suggestive. While the results indicate the heterogeneity of the impact of road access, one cannot indisputably conclude whether the impact of a road is more productive for higher income groups, as the control variable is only land and the effects of other observable and unobservable variables are not accounted for. A clear statement on such issues can only be made if we can have income figures for a period prior to the construction of all the roads covered here.

Apart from road access condition, the result shows that other village level conditions are also significant determinants of income. Among these, the average rainfall level of the village is an important one. The coefficient estimate of this variable represents the impact of rainfall on income that works through increasing farm level productivity, expanding off-farm employment opportunities and earnings. Using the same data set, for instance, Wondemu (2011) reported that for a 10% mm increase in rainfall, the village wage rate increases by 5%.

Among the other village characteristics, although its impact is only significant at just above 5% level, the distance of the household from the market also has an adverse impact on income. The altitude and temperature of the village however are not significant determinants of income. Among household specific observable factors, the size of land the household cultivates, the types of crop (mainly teff) cultivated, the number of adults in a household, and the frequency of rain shocks the household had experienced are significant determinants of income. For this sample, the age and schooling of the head of the household however are not significant determinants of income. The participation of the household in government sponsored extension services however has significant impact on income.

7. Conclusions

Our analysis has aimed to contribute to the growing literature on the impact of road investments. It shows that in a low income environment like Ethiopia improving road access can have strong developmental effects. Its broader interest outside the narrow context of investments in one poor country is two-fold. First it applies an econometric approach based a rigorous theoretical model to the analysis of household survey data that can be replicated elsewhere given the availability of such data. It is well known that impact studies of this type can be subject to various biases and econometric problems and we have highlighted where our analysis faces such difficulties as well as outlining our response to these. There is a growing literature on ways of assessing impact from road investment and we trust that our analysis is a contribution to this.

Second both theoretically and empirically we have tried to show the complex mechanisms through which road projects in poorly connected areas impact on economic activity. Such interventions improve rural income by altering farm gate price ratios and subsequently influencing micro level production, resource allocation and marketing decisions. By allowing for changes in product and factor prices our analysis moves very considerably beyond simple analyses of vehicle operating cost savings or approximate estimates of induced traffic in a way

¹³ The coefficient estimates are generated by regressing the strata-specific treatment effects on strata rank using variance-weighted least squares (Becker and Ichino, 2002).

which is backed by economic theory. Such mechanisms of changing prices operate in context well beyond Ethiopia our country of study and provide one of the keys to understanding the benefits of 'connectivity' which are often asserted rather than demonstrated. Similar analyses can be undertaken on other country contexts.

However approaching benefits in this way makes it clear that while expanding rural road access may often be essential as a means of raising rural incomes in poorly connected areas, it is not a panacea in that such intervention should be complemented by other policy and institutional measures that enhance the capacity of farmers to reap the benefits induced by road infrastructure.

As the empirical results show, apart from the condition of road access, household and location specific factors are also important determinants of rural income. Furthermore, since the key mechanism is the interaction through price changes, the effectiveness of markets and the institutions that support them will be critical in determining producer responses to the incentives created by road improvements. For instance, although road infrastructure significantly impacts micro level welfare, its potential impact is mediated by the degree of efficiency with which spatial arbitrage activities are undertaken and the degree of symmetry in spatial price adjustments. This means to maximize the developmental impact of road projects may also require reforms that address the institutional constraints that give rise to inefficient spatial integration of markets and asymmetric spatial price adjustments.

In addition, rural road investment is often seen as a central component of a 'pro-poor' or 'inclusive growth' strategy. Clearly building or improving roads in areas where the poor live should help lower poverty, but our results illustrate that we cannot automatically assume that within rural communities such programmes are automatically progressive in distributional terms, since access to land will be a key factor in how far rural households benefit. Thus, rationalizing the current land policy will be essential in order to maximize the developmental effectiveness of rural road infrastructure or other interventions that aim to enhance farm level income.

Despite concerted efforts by the government to enhance the farm level application of modern inputs, principally due to the low profitability and high risk of applying such inputs, their rate of adoption in Ethiopia still remains very low. The fact that in villages that have good road access the rate of fertilizer application is high across all the periods considered here, suggests that providing good road access should be a priority for government extension programs. Since good road access normally enhances the spatial integration of markets and reduces the responsiveness of farm gate prices to local supply shocks, it is expected to reduce the risk but enhances the profitability of the application of such inputs. Moreover, since in villages that have good road access prices are generally responsive to spatial market conditions, providing good road access could also enhance the effectiveness of other policy measures aimed at altering micro-level economic incentives and creating a shift in resource allocation.

From the analysis it also emerged that the crop pattern the household cultivates is an important determinant of farm level income. In addition to its micro level impact, farmers' crop choice is also expected to have macro level ramifications. Suboptimal crop choice at the farm level entails an efficiency loss and makes the economy produce and consume below its potential. Exploring both the extent to which the existing crop pattern in each agro-zone is consistent with comparative advantage and also to what extent providing good road access promotes specialization according to comparative advantage is an important future area of research.

Finally, the household level benefits reported here often cannot be captured through traditional transport surveys. The findings of this paper therefore challenge such an approach and underline the need to use household survey data in an attempt to reflect more fully changes as they affect h

ouseholds. This is because the standard project evaluation approach simply focuses on the short-run or immediate impact, usually proxied by transport cost savings, and typically neglects the structural change that road infrastructure brings about. This is also the case in the context of Europe where low volume roads or rural areas are often neglected in favour of roads in more urbanized and high-traffic density areas (Johansson, 2004). Since priority for funding is typically determined by conventional rate of return calculations that fail to pick up induced or generated structural shifts, this creates a bias against rural or low traffic density roads.

Table 1. The Impact of Road Quality on Income: The Estimation Results

Variables	Fixed Effect	H-Taylor
Land size	0.29***	0.37***
Number of Oxen	0.14***	0.18***
Number of Adults	0.26***	0.21***
Rain Shock Dummy	-0.06***	-0.04***
Extension Dummy	0.11***	0.18*
Soil Quality	-0.02	0.21**
Age of the Head	-0.04**	-0.12
Schooling of the Head	0.01	0.01
Teff Dummy	0.29***	0.27***
Village Population Size	-0.72***	0.12
Time Trend	0.33***	0.49***
Road Quality Dummy	0.63***	0.37*
Road Quality*Land	0.14*	0.16**
Road Quality*Time	-0.26***	-0.16***
Road Quality*Education	0.00	0.15
Road Quality*Number of Adults	-0.04	0.18*
Level of Annual Rainfall (in mm)		0.18***
Altitude (in Meter)		-0.20
Temperature (Celsius)		0.06
Distance from the Nearest Market Town		-0.05#
Constant	10.31***	5.30***
Number of Observations	2459	2459
R2	0.20	
rho (fraction of variance due to vi)		0.20
Wald chi2(22)		1556
F-statistics(18, 840)	36.41	

*** p<.001 ** p<.01, *<.05

#The distance variable is marginally significant at 5.5%. H-Taylor is the the result of the Hausman-Taylor Estimator (1981).

NOTE: All variables, except the time trend, Rain Shock, Soil Quality, Teff Dummy, schooling of the head, Road Quality and Extension are in natural logs. The fixed effect model is with Driscoll-Kraay standard errors. For both estimators, the unit of analysis is household.

Table 2. The Distributional Effect of Road Access

By Propensity score Strata	Coef.	z	P>z
1	-0.04	-0.35	0.73
2	0.15	2.08	0.04
3	0.22	6.25	0.00
4	0.82	5.84	0.00
Linear trend			
slope	0.17	3.72	0.00
constant	-0.26	-1.99	0.05

References

- Becker, S. and Ichino, A. (2002). Estimation of Average Treatment Effects Based on Propensity scores, *the STATA Journal* 2, pp. 358–377.
- Brand, J. Xie, Yu (2007). Who Benefits Most From College? Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education, California Centre for Population Research On-Line Working Paper Series no. 033.
- Breusch, T. and Pagan, A. (1980). the Lagrange Multiplier Test and Its Application to Model Specification in Econometrics, *Review of Economic Studies*, 47, pp. 239-253.
- Dercon, S. and Hoddinott, J. (2004). The Ethiopian Rural Household Surveys: Introduction, IFPRI and University of Oxford.
- Dercon, S., D.Gilligan, J.Hoddinott and T. Woldehanna. (2007). The Impact of Roads and Agricultural Extension on Consumption and Poverty in Fifteen Ethiopian villages, *ESSP Policy Conference Brief No. 5*, June.
- Driscoll, J., and Kraay, A. (1998). Consistent Covariance Matrix Estimation with Spatially Dependent Data, *Review of Economics and Statistics* 80(4).pp. 549-560.
- Fan, S., Zhang, L. and Zhang, X. (2002). Growth, inequality, and poverty in rural China: The role of public investments, Research Report No. 125, International Food Policy Research Institute.
- Fan, S., and Chan-Kang, C. (2005). Road development, Economic Growth, and Poverty Reduction in China, Research Report no. 138, International Food Policy Research Institute.
- Goldstein, R. (1994). the overlapping coefficient and an “improved” rank-sum statistic, *Stata Technical Bulletin* 22, pp.12-15.
- Haughwout, A. (1998). Aggregate Production Functions, Interregional Equilibrium, and the Measurement of Infrastructure Productivity, *Journal of Urban Economics*, 44, 216–227.
- Haughwout, A. (2002). Public infrastructure investments, productivity and welfare in fixed geographic areas, *Journal of Public Economics* 83, pp. 405–428.
- Hausman, J. (1978). Specification Tests in Econometrics, *Econometrica*, 46, pp. 1251-1271.
- Hausman, J. and Taylor, W. (1981). Panel Data and Unobservable Individual Effects, *Econometrica*, 49(6). 1377-1398.
- Hoechle, D. (2007). Robust Standard Errors for Panel Regressions with Cross-Sectional Dependence, *The Stata Journal*, 7(3), pp. 281-312.
- Jacoby, H. and Minten, B. (2009). On Measuring the Benefits of Lower Transport Costs, *Journal of Development Economics*, 89(1). pp. 28-38.
- Jalan, J. and Ravallion, M. (1998). Are there dynamic gains from a poor-area development program? *Journal of Public Economics* 67, 65–85.
- Johansson, S. (2004). Socio-Economic Impacts of Road Conditions on Low Volume Roads, Roadscanners Sweden AB.
- Khandker, S., Bakht, Z. and Koolwal, G. (2006). the Poverty Impact of Rural Roads: Evidence from Bangladesh, Policy Research Working Paper, WPS3875, World Bank.
- Khandker, S. Bakht, Z. and Koolwal, G (2009) Poverty Impact of Rural Roads: evidence from Bangladesh, *Economic Development and Cultural Change*, 57, 4, pp 685-722.

Kloek, T. (1981). OLS Estimation in a Model Where a Microvariable Is Explained by Aggregates and Contemporaneous Disturbances are Equicorrelated, *Econometrica* 49, pp. 205-207.

Menon, J. and Warr, P. (2008). Roads and Poverty: a General Equilibrium Analysis for Lao PDR. In *Infrastructure and Trade in Asia*, Cheltenham. Edited by Brooks, D. and Menon, J., pp.115-142.

Minten, Bart, (1999). Infrastructure, market access, and agricultural prices, MTID discussion papers 26, International Food Policy Research Institute (IFPRI).

MoFED (2002). Ethiopia: Sustainable Development and Poverty Reduction Program, Ministry of Finance and Economic Development, Addis Ababa.

Moulton, B. (1990). an Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units, *The Review of Economics and Statistics*, 72(2).

Parikh, A, Ali, F. and Shah, M. (1995). Measurement of economic efficiency in Pakistan agriculture, *American Journal of Agricultural Economics*, 77, pp. 675 - 685.

Pesaran, M. (2004). General Diagnostic tests for Cross Section Dependence in Panels, University of Cambridge, Faculty of Economics, Cambridge Working Papers in Economics No. 0435.

Rahman, S. (2003). Profit Efficiency among Bangladesh Rice Farmers, Paper presented on the 25th International Conference of Agricultural Economists (IAAE). Durban, South Africa.

Van de Walle, D. (2002). Choosing Rural Road Investments to Help Reduce Poverty, *World Development* 30(4): 575-589.

Van der Walle, D. and Mu, R. (2007) Fungibility and the flypaper effect of project aid: micro-evidence from Vietnam, *Journal of Development Economics*, 84, 2, pp 667-684

Wondemu, Kifle (2010). Road Infrastructure and Rural Poverty in Ethiopia, PhD thesis, University of Bradford, UK.

Wondemu, Kifle (2011). Road Infrastructure and Rural Income in Ethiopia: The Transmission Mechanisms, Lambert Publishing, Germany.

World Bank (2004). Can Agriculture Lead Growth in Ethiopia? The Importance of Linkages, Markets and Tradability, Ethiopia Country Economic Memorandum, the World Bank, Washington, DC.

Zhang, X. and Fan, S. (2001). How Productive is Infrastructure? New Approach and Evidence from India, EPTD Discussion Paper no. 84, IFPRI.

Appendix I

Household and Village Characteristics of the Sample by Types of Road Access

Key Characteristics	POOR ROAD ACCESS		GOOD ROAD ACCESS	
	Mean	Stdev.	Mean	Stdev.
Land Size (hectare)	1.21	1.20	2.11	1.62
Land Quality	0.80	0.36	0.88	0.23
Number of Oxen Owned	0.95	0.93	1.54	1.17
Human Capital (Years of School Attended)	0.51	0.92	0.92	1.38
Fertilizer Used in kg.	22.19	39.09	124.26	110.05
Fertilizer Used per Hectare in kg.	21.38	47.39	75.04	81.93
Share of Households used Fertilizer	0.38	0.49	0.88	0.32
Crop Output in Quintal	9.85	12.22	21.16	21.22
Crops Marketed in Quintal	2.77	14.68	5.63	12.34
Share of Households Supplied Crops to Market	0.54	0.50	0.65	0.48
Daily Wage Rate (Birr/day)	4.69	1.29	5.85	1.40
Fertilizer Price (Birr/ kg)	3.06	0.31	2.70	0.25
Crop Price (Birr/kg)	1.38	0.47	1.56	0.46
Family Size	5.87	3.13	6.43	3.26
Years of Schooling of Head of the Household	0.66	1.66	1.04	2.29
Distance from Zonal Town (km)	30.89	15.27	13.37	1.63
Distance from Addis Ababa (km)	467.90	250.58	179.08	70.32
Crop Output per Labour Input	0.13	0.24	0.28	0.62
Crop Output per Hectare of Land	8.81	12.98	12.15	12.16
Share of Income from Off-farm Activities	0.38	1.51	0.06	0.38
Average Annual Household Income (in Birr)	1110	1584	2952	3228
Share of Households Participated in Off-farm Activities	0.53	0.50	0.14	0.35
Share of Households by Types of Road Access (1997)	0.47			0.53
Share of Households by Types of Road Access (2004)	0.27			0.73
Number of Plots cultivated	3.39	1.68	6.24	3.81
Number of Crops Cultivated	2.15	1.16	3.36	1.51

SOURCE: Own Calculation from the Survey Data

Appendix II:











