CSAE WPS/2007-01

The impact of roads and agricultural extension on consumption growth and poverty in fifteen Ethiopian villages

Stefan Dercon, Daniel O. Gilligan, John Hoddinott and Tassew Woldehanna

May 23, 2006
This version, February 14, 2007

Abstract

This paper investigates whether public investments that led to improvements in road quality and increased access to agricultural extension services led to faster consumption growth and lower rates of poverty in rural Ethiopia. Using a Generalized Methods of Moments – Instrumental Variables – Household Fixed Effects estimator, we find evidence of positive impacts with meaningful magnitudes. Access to all-weather roads increases consumption growth by 16 per cent and, reduces the incidence of poverty by 6.7 per cent. Receiving at least one visit from an extension agent raises consumption growth by 7 per cent and reduces poverty incidence by nearly 10 per cent. These results are robust to changes in model specification and estimation methods.

Acknowledgements:
We thank the following organizations for funding the survey work on which this paper is based: SIDA, USAID and the World Food Programme. Analysis was supported by funding received from USAID under BASIS grant LAG-4213-A-00-6016-00 and from IFPRI's Ethiopian Strategy Support Program (ESSP). We have benefited from superb research assistance from Yisehac Yohannes and comments from seminar participants at Addis Ababa University, the Ethiopian Development Research Institute, the 2006 ESSP Policy Conference and IFPRI. Errors are ours.

Addresses for correspondence:
Dercon, Department of Economics, University of Oxford. Email: Stefan.Dercon@economics.oxford.ac.uk
Gilligan and Hoddinott, International Food Policy Research Institute, Washington DC. Email: D.Gilligan@cgiar.org; J.Hoddinott@cgiar.org
Woldehanna, Department of Economics, Addis Ababa University. Email: wtassew@econ.aau.edu.et.

In many African countries, improving growth rates in agriculture is seen as critical for sustained poverty reduction. Such a view stems from the fact that while Africa is urbanizing, the vast majority of people still live in rural areas and derive livelihoods from agricultural activities. Nowhere is this truer than in Ethiopia where agriculture accounts for 85 per cent of employment nationally and 96 per cent of employment in rural areas (World Bank 2005).

Public investments can play several roles in creating the enabling environment necessary to stimulate agricultural growth. One of these is through the provision of infrastructure, most notably improved roads. Better roads lower transactions costs associated with agricultural activities and in so doing have the potential to reduce the costs of acquiring inputs, increase output prices, reduce the impact of shocks and permit entry into new, more profitable activities. A second is through facilitating technology transfer. For example, by providing agricultural extension services, governments can make farmers aware of new agricultural technologies, advise them on best farming practice and assist farmers in dealing with adverse shocks such as insect infestations or plant diseases. While governments frequently are involved in other dimensions of agricultural activities, there is an *a priori* strong case for governments undertaking these investments given the public goods nature of roads and technology transfer.

However, uncovering evidence of impact of investments such as these is not straightforward.[1] One approach has been to use country-level or regional level data and relate this to changes in agricultural productivity. Antle's (1983) early, and still widely cited study is an example of this approach as is the work by Fan, Hazell and Thorat (2000) and Fan, Zhang and Zhang (2002). A strength of approaches such as these is that they permit the construction of benefit:cost ratios, and thus allow researchers to compare investments in infrastructure with other forms of public spending. However, these approaches do not tell us what component of infrastructure spending generates these benefits so that, for example, where investments are measured in terms of all public spending on roads, increased salaries to workers have the same impact on agricultural productivity as money spent building new roads. They do not inform discussions as to whether it is the quantity of infrastructure that matters or its quality nor, in the absence of distributional data, do they show the impact of these investments on poverty. By contrast, household level studies, such as Ahmed and Hossain (1990), Jalan and Ravallion (2002) and Jacoby (2000) can uncover the impact of infrastructure on poverty at the household level and, depending on the data available, take into account differences in infrastructural quality but cannot generate benefit:cost ratios. And all studies of the impacts of infrastructure need to confront issues of endogenous program placement (Rosenzweig and Wolpin 1986).

In the case of agricultural extension, Birkhaeuser, Evenson and Feder (1991) note that "attention should be given to the possibility that the allocation of extension efforts is not random across areas or communities. Such tendencies could distort results." Consider the following simple linear regression model: $Y_{ijt} = \gamma \,.\, X_{ijt} + B \,.\, Z_{ijt} + e_{ijt}$ where Y_{ijt} is a measure of output for farm i, situated in locality j at time t, X_{ijt} is a measure of contact with extension services, Z_{ijt} is a vector of other relevant characteristics that affect farm output, B and γ are parameters to be estimated and e_{ijt} is a disturbance term of the form, $e_{ijt} = v_j + v_i + v_{ijt}$. Here, v_j captures fixed characteristics of the locality not incorporated into Z_{ijt}, v_i captures fixed characteristics of the farm not incorporated into Z_{ijt} and v_{ijt} is a white noise disturbance term. As with roads, one form of bias follows from endogenous program placement. Suppose governments decide to concentrate extension resources in highly productive areas and that this fixed locality characteristic is not controlled for in the linear regression. Consequently, via the correlation between X_{ijt} and v_j, X_{ijt} and e_{ijt} will be correlated, yielding biased estimates of γ. Concerns regarding such placement effects resonate in the Ethiopian context. Efforts by the government extension service to encourage farmers to adopt a fertilizer-improved seed-credit package in the 1990s under the PADETES program were seen as leading to improvements in yields in some parts of the country. However, because these efforts were concentrated in areas with higher agricultural potential, placement effects may account for these improvements (World Bank 2005). The second bias is a form of selection bias. If better able or better skilled farmers are more likely to seek out extension services, or if extension agents prefer to seek out such individuals, and if this farm level characteristic is not taken into account, again X_{ijt} and e_{ijt} will be correlated (this time, via the correlation between X_{ijt} and v_i) and again γ will be a biased estimate of the impact of extension.

This paper contributes to literature on the impact of public investments in rural areas of poor countries – specifically the effect of improvements in road access and agricultural extension. Our specific focus is on their welfare impacts: growth in consumption and poverty status. To do so, we draw on data collected as part of the Ethiopian Rural Household Survey, a unique longitudinal survey of approximately 1400 households found in 15 villages. Using a Generalized Methods of Moments, Instrumental Variables, Household Fixed Effects estimator – one that accounts for the issues of endogenous placement and selection bias described above – we find that access to all-weather roads increases short run annual consumption growth by 15 per cent and reduces the likelihood of a household being poor by 6-7 per cent. These effects are well measured and robust to changes in model specification. Receiving at least one extension visit has smaller effects on consumption growth, increasing it by around 7 per cent but the effect on levels of poverty remain marked, with receipt of at least one extension visit reducing the likelihood of being poor by 10 per cent.

The paper begins by sketching out a growth model that is used to inform our approach to our data and our estimation strategy. After describing the survey data, we present descriptive statistics outlining trends in consumption and poverty as well as changes in access to these public investments. We then present our empirical results and robustness checks before offering some concluding remarks.

Theoretical framework

The framework used is a standard empirical growth model, allowing for transitional dynamics, inspired by Mankiw et al. (1992). We observe i households (i = 1, …, N) across periods t (t = 1, …, T). Growth rates for household i ($\ln y_{it} - \ln y_{it-1}$) are negatively related to initial levels of income ($\ln y_{it-1}$). Let δ represent sources of growth common to all households and X reflect fixed characteristics of the household, such as location, that also affect growth. Other sources of growth from t to t-1 are exogenous levels of capital stocks and access to technologies (k_{it-1}) observed at t-1 both of which are time varying. Lastly, while standard growth models do not allow for transitory shocks such as changes in rainfall ($\ln R_t - \ln R_{t-1}$), we know from previous work with our data (Dercon 2004; Dercon, Hoddinott and Woldehanna 2005) that such events do have growth effects. One way of thinking about these events is that initial efficiency (the technological coefficient in the underlying production function) may be influenced by period-specific conditions (Temple 1999) which cause growth rates to deviate from long term trend.

Mindful of the numerous reasons why one should be careful in applying this framework to any context, given the theoretical and empirical assumptions implied by this model (for example, see the reviews by Temple 1999, or Durlauf and Quah 1998), and dropping the i subscripts, our basic model is:

(1) $$\ln y_t - \ln y_{t-1} = \delta + \alpha \ln y_{t-1} + \beta \ln k_{t-1} + \gamma(\ln R_t - \ln R_{t-1}) + \lambda X$$

Data and setting

Ethiopia is a federal country divided into 11 regions. Each region is sub-divided into zones and the zones into woredas which are roughly equivalent to a county in the US or UK. Woredas, in turn, are divided into Peasant Associations (PA), or kebeles, an administrative unit consisting of a number of villages. Peasant Associations were set up in the aftermath of the 1974 revolution. Our data are taken from the Ethiopia Rural Household Survey (ERHS), a unique longitudinal household data set covering households in 15 areas of rural Ethiopia. Data collection started in 1989, when a survey team visited 6 Peasant Associations in Central and Southern Ethiopia. The survey was expanded in 1994 to encompass 15 Peasant Associations across the country, yielding a sample of 1477 households. As part of the survey re-design and extension that took place in 1994, the sample was re-randomized by including an exact proportion of newly formed or arrived households in the sample, as well as by replacing households lost to follow-up by others considered broadly similar to them in demographic and wealth terms by village elders and officials. The nine additional PAs were selected to better account for the diversity in the farming systems found in Ethiopia. The sampling in the PAs newly included in 1994 was based on a list of all households that was constructed with the help of the local Peasant Association officials.[2] The sample was stratified within each village to ensure that a representative number of landless households were also included. Similarly, an exact proportion of female headed households were included via stratification.

Table 1 gives the details of the sampling frame and the actual proportions in the total sample and Table 2 provides some basic characteristics of these localities. Using Westphal (1976) and Getahun (1978) classifications, Table 1 also shows that population shares within the sample are broadly consistent with the population shares in the three main sedentary farming systems – the plough based cereals farming systems of the Northern and Central Highlands, mixed plough/hoe cereals farming systems, and farming systems based around enset (a root crop also called false banana) that is grown in southern parts of the country. Note too that in 1994, the Central Statistical Office collected a national data set as part of the Welfare Monitoring System. Many of the average outcome variables, in terms of health and nutrition were very similar to the results in the ERHS, suggesting that living conditions in our sample did not differ greatly from those found more generally throughout rural Ethiopia (see Collier et al. 1997).

For these reasons, it can be argued that the sampling frame to select the villages was stratified in the main agro-ecological zones and sub-zones, with one to three villages selected per strata. Further, sample sizes in each village were chosen so as to approximate a self-weighting sample, when considered in terms of farming system: each person (approximately) represents the same number of persons found in the main farming systems as of 1994. However, results should not be regarded as nationally representative. The sample does not include pastoral households or urban areas.[3] Also, the practical aspects associated with running a longitudinal household survey when the sampled localities are as much as 1000km apart in a country where top speeds on the best roads rarely exceed 50km/hour constrained sampling to only 15 communities in a country of thousands of villages. Therefore, while these data can be considered broadly representative of households in non-pastoralist farming systems as of 1994, extrapolation from these results should be done with care.

Following the first survey round on the 15 village sample in 1994, an additional round was conducted in late 1994, with further rounds in 1995, 1997, 1999 and 2004. These surveys were conducted, either individually or collectively, by the Economics Department at Addis Ababa University, the Centre for the Study of African Economies, University of Oxford or the International Food Policy Research Institute. Sample attrition between 1994 and 2004 is low in part because of this institutional continuity, with a loss of only 12.4 percent (or 1.3 percent per year) of the sample over this ten year period.[4] This continuity also helped ensure that questions asked in each round were identical, or very similar, to those asked in previous rounds and that the data were processed in comparable ways.

Descriptives

Outcomes

We consider two outcomes: real consumption per adult equivalent; and whether the household is poor.

Consumption is defined as the sum of values of all food items, including purchased meals and non-investment non-food items. The latter are interpreted in a limited way, so that contributions for durables and non-durables, as well as health and education expenditures are excluded (Hentschel and Lanjouw 1996). Although there are good conceptual reasons for including use values for durables or housing (Deaton and Zaidi 2002), we do not do so here; the heterogeneity in terms of age and quality of durables owned by our respondents, together with the near complete absence of a rental market for housing would make the calculation of use values highly arbitrary. Because comparisons of productive and consumer durable holdings between 1994 and 2004 show rising holdings of these durables[5] and comparisons of school enrollment data show significant increases in enrollment, *ceteris paribus*, our consumption estimates may understate the actual increases in household welfare and bias downwards impacts of public investments on consumption and poverty. Consumption is expressed in per adult equivalent terms; see Dercon and Krishnan (2000) for the conversion factors used to express these. Lastly, it is deflated by a food price index, calculated as a Laspeyres index, based on local (peasant association) prices collected specifically for this purpose and using average expenditure shares in 1994 as the weights.

Estimating levels and changes in poverty requires first setting a poverty line. Here, we use a cost-of-basic-needs approach. Based on the 1994 data, a food poverty line is constructed using a bundle of food items that would provide 2300 Kcal per adult per day. To this, we add a non-food bundle using the method set out in Ravallion and Bidani

(1994). Dercon and Krishnan (1996; 2003) provide further information on the construction of the poverty line, including details of the food basket and its sensitivity to different sources of data on prices used to value the food basket.

Figures 1 and 2 show the evolution of these three outcome variables over time. In examining these, it is important to note that the timing of the rounds in 1995, 1999 and 2004 was approximately the same as that of the first 1994 round. However, the 1997 round was collected at a much different point of time in the agricultural year – the immediate post-harvest period – and this seasonal consideration together with the fact that 1997 was, in agricultural terms, atypically good has the effect of making the 1997 outcomes look particularly high. Mindful of this, there is significant growth in real consumption between 1994 and 1999. Distributionally, this growth is pro-poor as evidenced by the reduction in headcount poverty from 48 to 36 per cent in this period and a fall in the Gini coefficient for consumption (not shown) from 0.44 to 0.41. However, these improvements slow after 1999. Median real per adult equivalent consumption, which grew by 24 per cent between 1994 and 1999, grows by only 9 per cent in the following five years. Growth is less distributionally neutral with mean consumption growth growing faster than median growth and there being no change in the headcount poverty measure.[6]

Public investments: Trends at the household level

As noted in the introduction, the primary focus of this paper is the impact of two forms of public investments – roads and extension - on consumption growth and poverty. Here, we explain how these are defined in our household data set and show how these have evolved over time.

All households in this sample have access to some sort of road or path. However, the quality of this road varies significantly from all-weather roads suitable for vehicular traffic to mud tracks that at best can support foot traffic. As noted in the introduction, the benefits to roads are perceived to operate through four channels: reducing the costs of acquiring inputs; increasing output prices; reducing the impact of shocks and permitting entry into new, more profitable activities. Given this, and given the data available to us in the survey, we define road access as a dummy variable equaling one if the household has access to a road capable of supporting truck (and therefore trade) and bus (and therefore facilitating the movement of people) traffic in both the rainy and dry seasons.

The household survey instrument asked households how many times they had been visited by an extension agent during the last main cropping season. Using these data, we create a dummy variable equaling one if the household had received at least one such visit, zero otherwise.

Figure 3 shows how access to these forms of public investments has changed over the ten year period covered by the ERHS. Initial levels of access to all-weather roads was around 40 per cent, with significant improvements being recorded between 1997 and 1999 and 1999 and 2004. The percentage of households receiving at least one visit from an extension agent triples over this ten year period and it is worth noting that this increase is widely distributed with 13 of our 15 villages recording an increase in the number of households receiving at least one visit. However, the starting level in 1994 – 5.6 per cent – was stunningly low and most of this improvement occurs between 1994 and 1999.

Model and results

The empirical model

Before estimating equation (1) using the data described in sections 3 and 4, there are a number of empirical issues that require consideration. First, note that we do not have evenly spaced observations over time. This can be thought of as a missing data problem – that is, how do we estimate (1) when we are missing data for 1996, 1998 and 2000-2003? To see how this might affect our model, writing (1) for growth between t-1 and t-2 would give:

(2) $$\ln y_{t-1} - \ln y_{t-2} = \delta + \alpha \ln y_{t-2} + \beta \ln k_{t-2} + \gamma(\ln R_{t-1} - \ln R_{t-2}) + \lambda X$$

Suppose now we only observe t-2 and t. Then adding up (1) and (2) and dividing by two gives us:

(3) $$(\ln y_t - \ln y_{t-2})/2 = \delta + \alpha(\ln y_{t-1} + \ln y_{t-2})/2 + \beta(\ln k_{t-1} + \ln k_{t-2})/2 + \gamma(\ln R_t - \ln R_{t-2})/2 + \lambda X$$

Our left hand side is the average growth rate (also equal to $\ln y_t^{1/2} - \ln y_{t-2}^{1/2}$) while the right hand side consists of a number of complicated terms, with the exception of the last term, which is the yearly average of the rainfall change (or the total change divided by two). Extending this to p-periods in between this becomes:

(4) $$(\ln y_t - \ln y_{t-p})/p = \delta + \alpha(\ln y_{t-1} + \ldots + \ln y_{t-p})/p + \beta(\ln k_{t-1} + \ldots + \ln k_{t-p})/p + \gamma(\ln R_t - \ln R_{t-p})/p + \lambda X$$

This presents problems for estimation for the lagged dependent variable, and all time varying 'level' variables (such as infrastructure at k). However, if one is willing to acknowledge that changes are still relatively slow so that ln $y_{it-1} \approx$ ln y_{it-p} and similarly for k, then the p-period average is approximated by the initial level at t-p. Then the regression to be estimated is:

(5) $(\ln y_t - \ln y_{t-p}) /p = \delta + \alpha \ln y_{t-p} + \beta \ln k_{t-p} + \gamma(\ln R_t - \ln R_{t-p})/p + \lambda X$

All changes are expressed in averages per period (divided by p) and all level variables remain as they are, defined at t-p. The constant (and the fixed effects) are not affected.

Our next step is to introduce a disturbance term, ε_{it}, into (5). ε_{it} has two parts, a time invariant component (μ_i) and a time varying component (υ_{it}). The time invariant component can be thought of as capturing all characteristics of the village and household not observed by us which do not change over time while υ_{it} is white noise disturbance. Including these yields

(6) $(\ln y_t - \ln y_{t-p}) /p = \delta + \alpha \ln y_{t-p} + \beta \ln k_{t-p} + \gamma(\ln R_t - \ln R_{t-p})/p + \lambda X + \varepsilon_{it-p}$

However, this disturbance term introduces further complications. First, there are good *a priori* reasons to believe that $E(\ln y_{t-p}\, \varepsilon_{it-p}) \neq 0$. To see why, note that ln y_t reflects growth in ln y between periods t and t-1 and that ln y_{t-1} reflects growth in ln y between periods t-1 and t-2. ε_{it-1} enters into the growth regression for $(\ln y_t - \ln y_{t-1})$ and ε_{it-2} enters into the growth regression for $(\ln y_{t-1} - \ln y_{t-2})$. If there is any serial correlation in the disturbance terms, $E(\varepsilon_{it-1}\, \varepsilon_{it-2}) \neq 0$ and so $E(\ln y_{t-p}\, \varepsilon_{it-p}) \neq 0$. Making matters worse, note that a standard question in estimates of models like (6) is whether there is conditional convergence in the household data: a negative estimate for α would suggest convergence, allowing for underlying differences in the steady state. Unobserved village or household characteristics play a role in determining these steady states so that there is correlation between ln y_{t-p} and μ_i and therefore between ln y_{t-p} and ε_{it-p}. Second, as discussed in the introduction, E($k_{t-p}\ \varepsilon_{it-p}$) ≠ 0 because E($k_{t-p}\ \mu_i$) ≠ 0. In the case of roads, it is simply not tenable to believe that they are randomly scattered across the countryside. In the case of extension, if government extension services are targeted to more productive areas – for example, based on unobserved land fertility or entrepreneurship – or if better farmers were seeking out extension agents, this would be reflected in our model by correlation between observed characteristics of these farmers, such as k_{t-p} and ln y_{t-p}, and their unobserved characteristics.

Given these concerns, we estimate (6) using an instrumental variables (IV) – household fixed effects (HFE) estimator. We do so using Generalized Methods of Moments (GMM) so that these estimates are both consistent and efficient (Wooldridge 2002).[7] The household fixed effects dimension of our estimator means that we difference all left and right hand side components of (6) by their mean. In so doing, unobserved, time invariant household characteristics are differenced out, as are the observed, time invariant household characteristics, *X*. We instrument ln y_{t-p} using time varying household characteristics observed at time t-p. These are log fertile land holdings, log number of adult equivalents and log number of livestock units. Again thinking of the discussion of Solow-type growth models, these can be thought of as household characteristics which influence how close the household is to the steady state. There are two additional advantages of this approach. First, should there be any attrition bias brought about by the influence of time invariant household characteristics on attrition, household fixed effects estimation will also address this (though, as noted in footnote 3, we do not believe such attrition is a concern). Second, the use of IV will reduce attenuation bias from measurement error in the regressors.

Basic results

Table 3 presents the results of estimating (6) using a GMM-IV-HFE for the outcomes considered here, growth in consumption and whether the household is poor. The last outcome is not, strictly speaking, a direct product of a growth regression. Rather, it can be thought of as an extension of the consumption growth regressions. While these growth regressions show the average effect of public investments across the whole sample, the poverty regressions give us insights into the distributional effects of these investments – specifically whether they are of sufficient magnitude to pull poor households out of poverty.

The first two columns provide a basic set of results that exclude the impacts of rainfall and other shocks; in other words, growth models without allowances for transitory shocks. These show that access to good roads increase consumption growth by approximately 15 per cent; thus effect is significant at the 5 per cent level. These impacts also appear to be pro-poor with access to good roads reducing the likelihood that a household is poor by 7 percentage points. Receiving a visit from an agricultural extension officer boosts consumption growth and reduces the likelihood that the household is poor, but the former effect is relatively small and not significant at the 10 per cent level.

The third and fourth columns report estimates of these models with a set of time varying shocks included as additional controls. The first of these, rainfall shocks, is the log change in annual rainfall since the previous round. The next four are household self-reports of input price, output price, death or illness ‘shocks’ that, when they occurred, lead to reductions in household income or consumption or led to asset losses.[8] This model includes all components of the growth regression described by (6) and thus, represents our preferred specification.

We begin by considering the results of the specification tests associated with our results. The first stage F statistic and the Cragg-Donald F statistic provide information on the relevance of the variables used to instrument the endogenous lagged dependent variable. Where lagged log consumption is treated as endogenous, these test statistics indicate that we have good instruments in the relevance sense; using Table 1 of Stock and Yugo (2004) as a guide, with 95 per cent confidence, our IV estimates have less than 5 per cent of OLS bias. The Hansen J test is an over-identification test. It shows that we cannot reject the null that our instruments can be excluded from the second stage regressions for consumption growth and poverty despite the fact that, as Baum, Shaffer and Stillman (2003, p. 17) and Hoxby and Paserman (1998) both note, this test tends to *over-reject* this null in presence of intra-cluster correlation.[9]

The results in columns (3) and (4) are nearly identical to those reported in columns (1) and (2), with the notable exception that the impact of extension of consumption growth is now significant at the 10 per cent level. Improvements in access to good roads, and the receipt of at least one visit by an extension agent increase consumption and reduce poverty by meaningful amounts.

Robustness checks

Table 4 shows the results of a series of robustness checks designed to assess whether changes in model specification, estimation or sample affect these core results.

While GMM estimates are both consistent and efficient, they are vulnerable to the influence of outliers because the optimal weighting matrix that underpins them is a function of fourth moments (Baum, Schaffer and Stillman 2003, p. 11; Hayashi 2000, p. 215). We address this concern in two ways. First, we report the results of 1% trim estimates that drop the top and bottom 1% of observations of consumption growth. Second, we use a limited information maximum likelihood (LIML) estimator that is not vulnerable to this concern and has the added advantage of being a superior estimator when instruments are weak (see Stock and Yugo 2004, p. 31). Doing so has no meaningful effect on the regressions on consumption growth or poverty status.

Rows (3) and (4) of Table 4 show how sensitive the results are to changes in model specification. In row (3), we replace self-reported output price shocks with the weighted value of changes in output prices. Row (4) reports results when additional characteristics of the household (lagged average number of grades of schooling of household members older than 15) and lagged sex of head are included as additional determinants of growth and poverty status. Generally, these do not lead to major changes in parameter estimates or significance. This is also true if we try other specifications not reported here such as replacing livestock units with livestock values in the instrument set, adding other public investments such as access to piped water or adding lagged rain as an additional instrument.

Next, we consider the impact of also treating access to extension as endogenous. While our estimation strategy controls for many observable and unobservable factors that might otherwise be correlated with access to extension – including time invariant household characteristics and time varying household shocks – it still might be possible that there is some unobserved, time varying household characteristic correlated with access to extension. To do so, we add an additional instrumental variable, the number of agricultural extension officers available within the Peasant Association. When we do so, results from the first stage F statistic and the Cragg-Donald F statistics again indicate that we have strong instruments for the consumption growth and poverty regressions. And also as before, our Hansen J test results indicate that we satisfy the uncorrelatedness condition for the consumption growth and poverty results. Mindful of these test statistics, the results reported in row (5) indicate that, when instrumented, the coefficient on extension increases dramatically. Because the number of agricultural extension officers changes only slowly over time – and in some localities does not change at all – it is possible that these results are picking up a *local* treatment effect (LATE) as opposed to an average treatment effect over the full sample (Imbens and Angrist 1994; Card 2001). For this reason, we are cautious about taking the results shown in row (5) at face value. What they do suggest, however, is that our estimates of the impact of visits by extension services are conservative.

Lastly, we estimated our growth and poverty models as reduced forms, fixed effects regression when endogenous dummy variables are dropped and their instruments included in their place. The impact of access to roads is unchanged in these regressions as is the impact of access to extension on poverty status. Access to extension continues to reduce poverty; however, the magnitude of its effect on consumption growth is slightly smaller than in the IV regressions and the parameter estimate is no longer statistically significant.[10]

Further discussion

The results presented here give estimates of the impact of public investments that lead to better roads and greater access to extension services. They do not tell us why we observe these effects. There is other data in the ERHS that can help address this question.

Better roads in these localities make it easier for households to access local market towns which in turn are linked to larger urban centres. Dercon and Hoddinott (2005) document the myriad economic links between these survey sites and these market towns. They show that in 2004, roughly half of households purchasing inputs for crops in the Meher (long rain) and Belg (short rain) seasons do so in local market towns. About 40 per cent of households purchase inputs for livestock such as feed in these localities. For four crops grown widely in this sample (teff, wheat, maize and eucalyptus), there is considerable variation in location of sale: ranging from 24 per cent (eucalyptus) to 59 per cent (wheat) being sold in local market towns. Most notably, the vast majority of livestock and livestock products are sold in the local market towns. Artisanal products made by villagers (particularly by women) such as handicrafts are typically sold in local market towns. Lastly, more than half the purchases of goods for consumption occur in local market towns. Dercon and Hoddinott (2005) also show that improvements in road quality increase the likelihood of purchasing crop inputs (by 29 to 34 per cent, depending on the season) and, for women, selling artisanal products (by 39 per cent).

Understanding why agricultural extension has positive impacts is trickier because, apart from the 1999 survey round, we have little direct information on exactly what information is imparted by agents to farmers. The 1999 survey asked farmers to describe the two most important activities of extension agents. Being a source of information about the usage of modern inputs was ranked by 62 per cent of respondents as being the most important activity and a further 10 per cent of respondents listed this as their second most important activity. A source of knowledge about new cultivation practices was listed by 16 per cent of households as extension agents while 46 per cent listed this as their second most important activity. Further, amongst households using a modern input such as fertilizers, 56 per cent reported that they were encouraged to do so by extension agents. We also computed Pearson correlation coefficients for the use of fertilizer and receipt of at least one visit by an extension agent. In 1994, this relationship was weak, with the Pearson correlation coefficient equaling 0.07. However, by 2004 this association appeared much stronger with the Pearson correlation coefficient equaling 0.27 and being significant at the 1% level. Given this, drawing implications of our results on agricultural extension should be done cautiously. Some of the effect may represent transfers of technology or knowledge while some of the effect may reflect the influence that extension agents have in terms of increased use of fertilizer and other inputs.

Conclusions

Public investments have the potential to play important roles in facilitating increased growth and faster poverty reduction. In this paper we have investigated whether two forms of public investment have played such a role in rural Ethiopia. Using longitudinal household data, we find that public investments that led to improvements in road quality and increased access to agricultural extension services led to faster consumption growth and lower rates of poverty. The magnitudes of these effects are meaningful. Access to all-weather roads increases consumption growth by 16 per cent and reduces the incidence of poverty by 6.7 per cent. Receiving at least one visit from an extension agent also increases consumption growth by 7 per cent and reduces poverty incidence by nearly 10 per cent. These results are obtained using a GMM-IV-HFE estimator so they take into account all fixed household characteristics as well as the impact of transitory shocks. The results for consumption growth and poverty are robust to changes in model specification and estimation methods. Lastly, we caution that while some of the effects of agricultural extension represent transfers of technology or knowledge, others may capture the influence that extension agents have in terms of increased use of fertilizer and other inputs.

References

Ahmed, A., and M. Hossain. 1990. *Development impact of rural infrastructure in Bangladesh*. IFPRI Research Report 83, International Food Policy Research Institute, Washington DC.

Angrist, J. and V. Lavy. 2002. "The effect of High School Matriculation Awards: Evidence from randomized trails." Cambridge, MA: National Bureau of Economic Research Working Paper 9389.

Antle, J. 1983. "Infrastructure and aggregate agricultural productivity: International evidence." *Economic Development and Cultural Change* 31: 609-619.

Baum, C., M. Shaffer and S. Stillman. 2003. "Instrumental variables and GMM: Estimation and testing." Department of Economics, Boston College Working paper 545.

Bevan, P. and A. Pankhurst. 1996. "Report on the sociological dimensions of the Ethiopian Rural Economies Project." Unpublished, Centre for the Study of African Economies, Oxford University.

Birkhaeuser, D., R. Evenson and G. Feder.1991. "The economic impact of agricultural extension: A review." *Economic Development and Cultural Change* 39(3): 507-521.

Card, D. 2001. "Estimating the return to schooling: Progress on some persistent econometric problems." *Econometrica* 69: 1127-60.

Collier, P., S. Dercon, and J. Mackinnon. 1997. *Social Sector Review - Per II, Ministry of Finance*. Addis Ababa: Government of Ethiopia.

Deaton, A. 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*, Baltimore and London: The Johns Hopkins University Press for the World Bank.

Deaton, A. and S. Zaidi. 2002. "Guidelines for constructing consumption aggregates for welfare analysis." Living Standards Measurement Study Working Paper: 135, Washington, D.C.: The World Bank.

de Janvry, A., F. Finan, E. Sadoulet and R. Vakis. 2006. "Can conditional cash transfer programs serve as safety nets in keeping children at school and from working when exposed to shocks?" *Journal of Development Economics* 79(2): 349-373.

Dercon, S. 2004. "Growth and shocks: Evidence from rural Ethiopia." *Journal of Development Economics* 74(2): 309-329.

Dercon, S., and J. Hoddinott. 2004. "The Ethiopian Rural Household Surveys: Introduction." Unpublished, International Food Policy Research Institute, Washington DC.

Dercon, S. and J. Hoddinott. 2005. "Livelihoods, growth, and links to market towns in 15 Ethiopian villages." FCND Discussion Paper 194, International Food Policy Research Institute, Washington DC.

Dercon, S., J. Hoddinott and T. Woldehanna. 2005. "Consumption and shocks in 15 Ethiopian Villages, 1999-2004." *Journal of African Economies*, 14: 559-585.

Dercon, S., and P. Krishnan. 1996. "A consumption based measure of poverty in Ethiopia: 1989-1994" in, M. Taddesse and B. Kebede, *Poverty and economic reform in Ethiopia*, Proceedings Annual Conference of the Ethiopian Economics Association.

Dercon, S., and P. Krishnan. 2000. “Vulnerability, poverty and seasonality in Ethiopia.” *Journal of Development Studies* 36(6): 25-53.

Dercon S., and P. Krishnan. 2003. “Changes in poverty in rural Ethiopia 1989-1995” in A.Booth and P.Mosley, (eds.) *The New Poverty Strategies*, Palgrave MacMillan, London.

Durlauf, S. and D. Quah. 1998. The new empirics of economic growth. CEP Discussion paper 384.

Fan, S., and C. Chan-Kang. 2005. *Road development, economic growth, and poverty reduction in China*. IFPRI Research Report 138, International Food Policy Research Institute, Washington DC.

Fan, S., P. Hazell and S. Thorat. 2000. “Government spending, agricultural growth and poverty in rural India.” *American Journal of Agricultural Economics* 82(4): 1038-1051.

Fan, S., L. Zhang and X. Zhang. 2002. *Growth, inequality and poverty in rural China: The role of public investments*. IFPRI Research Report 125, International Food Policy Research Institute, Washington DC.

Getahun. 1978. Report on framing systems in Ethiopia. Ministry of Agriculture, Government of Ethiopia, Addis Ababa.

Hayashi, F. 2000. *Econometrics*. Princeton University Press, Princeton NJ.

Hentschel, J. and P. Lanjouw. 1996. “Constructing an indicator of consumption for the analysis of poverty: Principles and illustrations with reference to Ecuador.” Living Standards Measurement Study Working Paper: 124, Washington, D.C.: The World Bank.

Hoxby, C. and M. Paserman. 1998. Overidentification tests with grouped data. Cambridge, MA: National Bureau of Economic Research technical Paper 223.

Hyslop, D. 1998. "State dependence, serial correlation, and heterogeneity in intertemporal labor force participation of married women." *Econometrica* 67(6): 1255-94.

Imbens, G. and J. Angrist. 1994. Identification and estimation of local average treatment effects. *Econometrica* 62**:** 467-76.

Islam, N. 1995. "Growth empirics: A panel data approach." *Quarterly Journal of Economics*, 110(4): 1127-1170.

Jacoby, H. 1990. "Access to markets and the benefits of rural roads." *Economic Journal* 110(465): 713-737.

Jalan, J., and M. Ravallion. 2002, "Geographic poverty traps? A micro model of consumption growth in rural China." *Journal of Applied Econometrics* 17(4): 329-346.

Kish, L. 1965, *Survey Sampling*, New York: John Wiley & Sons.

Mankiw, G., D. Romer and D. Weil. 1992. "A contribution to the empirics of growth." *Quarterly Journal of Economics* 107: 409-437.

Moulton, B. 1990. "An illustration of a pitfall in estimating the effects of aggregate variables on micro units." *Review of Economics and Statistics* 72(2): 334–338.

Ravallion, M. and B. Bidani 1994. "How robust is a poverty profile?" *World Bank Economic Review* 8(1): 75-102.

Rosenzweig, M. and K. Wolpin. 1986. "Evaluating the effects of optimally distributed public programs." *American Economic Review* 76(3): 470-487.

Stock, J. and M. Yugo. 2004. Testing for weak instruments in linear IV regression. Unpublished, Department of Economics, Harvard University.

Temple, J. 1999. “The new growth evidence.” *Journal of Economic Literature*, 37(1): 112-156.

Westphal, E. 1976. *Farming Systems in Ethiopia*. Food and Agriculture Organization of the United Nations.

World Bank. 2005. *Poverty and Well-Being: The Role of Agriculture and Shocks*: Washington, D.C.: World Bank.

Wooldridge, J. 2002. *Econometric analysis of cross section and panel data*. MIT Press, Cambridge MA.

Wooldridge, J. 2003. “Cluster-sample methods in applied econometrics.” *American Economic Review Papers and Proceedings* 93(2): 133–8.

Table 1: The distribution of households in the Ethiopian Rural Household Survey, by agroecological zone

	Population share in 1994	**Sample share in 1994**	**Number of villages**
	(percent)	(percent)	
Grain plough complex: Northern Highlands	21.2%	20.2%	3
Grain plough complex: Central Highlands	27.7	29.0	4
Grain plough: Arsi/Bale	9.3	14.3	2
Sorghum plough/hoe: Hararghe	9.9	6.6	1
Enset (with or without coffee/cereals)	31.9	29.9	5
Total	100	100	15

Source: Dercon and Hoddinott (2004).
Note: Percentages of population share relate to the rural sedentary population; they exclude pastoralists who account for about 10 percent of total rural population.

Table 2: Characteristics of the sample sites

Survey site	Location	Description	Main crops	Perennial crops?	Mean Rainfall mm
Haresaw	Tigray	Poor and vulnerable area.	Cereals	no	558
Geblen	Tigray	Poor and vulnerable area; used to be quite wealthy.	Cereals	no	504
Dinki	N. Shoa	Badly affected by 1984/85 famine; not easily accessible even though near Debre Berhan.	Millet, teff	no	1,664
Debre Berhan	N. Shoa	Highland site. Near town.	Teff, barley, beans	no	919
Yetmen	Gojjam	Near Bichena. Ox-plough cereal farming system of highlands.	Teff, wheat, and beans	no	1,241
Shumsha	S.Wollo	Poor area in neighborhood of airport near Lalibela.	Cereals	no	654
Sirbana Godeti	Shoa	Near Debre Zeit. Rich area. Much targeted by agricultural policy. Cereal, ox-plough system.	Teff	no	672
Adele Keke	Hararghe	Highland site. Drought in 1985/86	Millet, maize, coffee, chat	yes, no food	748
Korodegaga	Arssi	Poor cropping area in neighborhood of rich valley.	Cereals	no	874
Turfe Kechemane	S.Shoa	Near Shashemene. Ox-plough, rich cereal area. Highlands.	Wheat, barley, teff, potatoes	yes, some	812
Imdibir	Shoa (Gurage)	Densely populated *enset* area.	*Enset*, chat, coffee, maize	yes, including food	2,205
Aze Deboa	Shoa (Kembata)	Densely populated. Long tradition of substantial seasonal and temporary migration.	*Enset*, coffee, maize, teff, sorghum	yes, including food	1,509
Addado	Sidamo (Dilla)	Rich coffee producing area; densely populated.	Coffee, *enset*	yes, including food	1,417
Gara Godo	Sidamo (Wolayta)	Densely packed *enset*-farming area. Famine in 1983/84. Malaria in mid-88.	Barley, *enset*	yes, including food	1,245
Doma	Gama Gofa	Resettlement Area (1985); Semi-arid; experienced droughts throughout the 1980s; remote.	*Enset*, maize	yes, some	1,150

Source: Community survey ERHS, Bevan and Pankhurst (1996), and Dercon and Hoddinott (2004).

Table 3: Instrumental Variables – Household Fixed Effects determinants of consumption growth and poverty status: Basic results

	Consumption growth	Poor	Consumption growth	Poor
	(1)	(2)	(3)	(4)
Lagged endogenous variables				
Log consumption	-0.376	-0.147	-0.365	-0.148
	(9.43)**	(3.67)**	(9.03)**	(3.64)**
Public investments				
Access to all-weather road	0.166	-0.069	0.159	-0.067
	(5.42)**	(2.17)**	(5.14)**	(2.11)**
Received visit from extension officer	0.059	-0.095	0.071	-0.098
	(1.59)	(2.48)**	(1.91)*	(2.54)**
Other controls				
Rainfall shocks			0.173	-0.057
			(4.37)**	(1.70)*
Input price shocks			-0.116	-0.046
			(0.89)	(0.39)
Output price shocks			0.069	-0.079
			(0.41)	(0.48)
Death shocks			-0.149	0.071
			(1.80)*	(0.98)
Illness shocks			-0.108	0.084
			(1.14)	(0.99)
Diagnostic statistics				
F stat on first stage instruments	90.10**	90.10**	87.64**	87.64**
Cragg-Donald F stat	104.52**	104.52**	103.00**	103.00**
Hansen J test	4.68	0.82	4.59	0.89
Sample size	4781	4781	4771	4771

Notes:
1. Lagged endogenous variables are expressed in real per adult equivalent terms. 2. Instruments for lagged endogenous variables are lagged log livestock units per adult equivalent, lagged log number of adult equivalents and lagged log cultivable land per adult equivalent. 3. A dummy variable if survey conducted in post harvest period is included but not reported. 4. Absolute values of z stats in parentheses; 5. * significant at 10% level; ** significant at 5% level

Table 4: Selected robustness checks on basic results

	Specification	Growth in Consumption: Access to all-weather road	Growth in Consumption: Received visit from extension officer	Poor: Access to all-weather road	Poor: Received visit from extension officer
(0)	Basic results	0.159	0.071	-0.067	-0.098
		(5.14)**	(1.91)*	(2.11)**	(2.54)**
(1)	1% trim of dependent variables	0.163	0.061	-0.069	-0.095
		(6.07)**	(1.75)*	(2.14)**	(2.45)**
(2)	LIML estimation	0.161	0.069	-0.067	-0.096
		(5.10)**	(1.91)*	(2.08)**	(2.62)**
(3)	Include weighted value of changes in output prices	0.163	0.071	-0.070	-0.098
		(5.34)**	(1.96)**	(2.22)**	(2.58)**
(4)	Include additional characteristics of household and head	0.149	0.071	-0.062	-0.097
		(4.83)**	(1.92)*	(1.94)*	(2.52)**
(5)	Instrument access to extension	0.152	0.559	-0.060	-0.574
		(4.91)**	(2.88)**	(1.86)*	(2.36)**
(6)	Reduced form regression	0.170	0.043	-0.061	-0.106
		(4.13)**	(0.93)	(1.95)*	(3.01)**

Specification notes: (0): Reports, for purposes of comparison, results found in columns (3) – (4) of Table 3. (1): Reports 1% trim estimates that drop the top and bottom 1% of observations of consumption growth (for growth in consumption and poverty status); (2) Reports results of using Limited Information Maximum Likelihood (LIML) estimation rather than GMM; (3) Reports results when the output price shocks variable is dropped and replaced with the weighted value of changes in output prices; (4) Reports results when additional characteristics of the household (lagged average number of grades of schooling of household members older than 15) and lagged sex of head are included as additional determinants of growth and poverty status; (5) Reports results when "received visit from extension officer" is treated as endogenous and instrumented by number of extension offices within PA; and (6) Reports results of estimated a reduced form fixed effects regression when endogenous dummy variables are dropped and their instruments included. For additional notes, see Table 3.

Figure 1: Trends in consumption

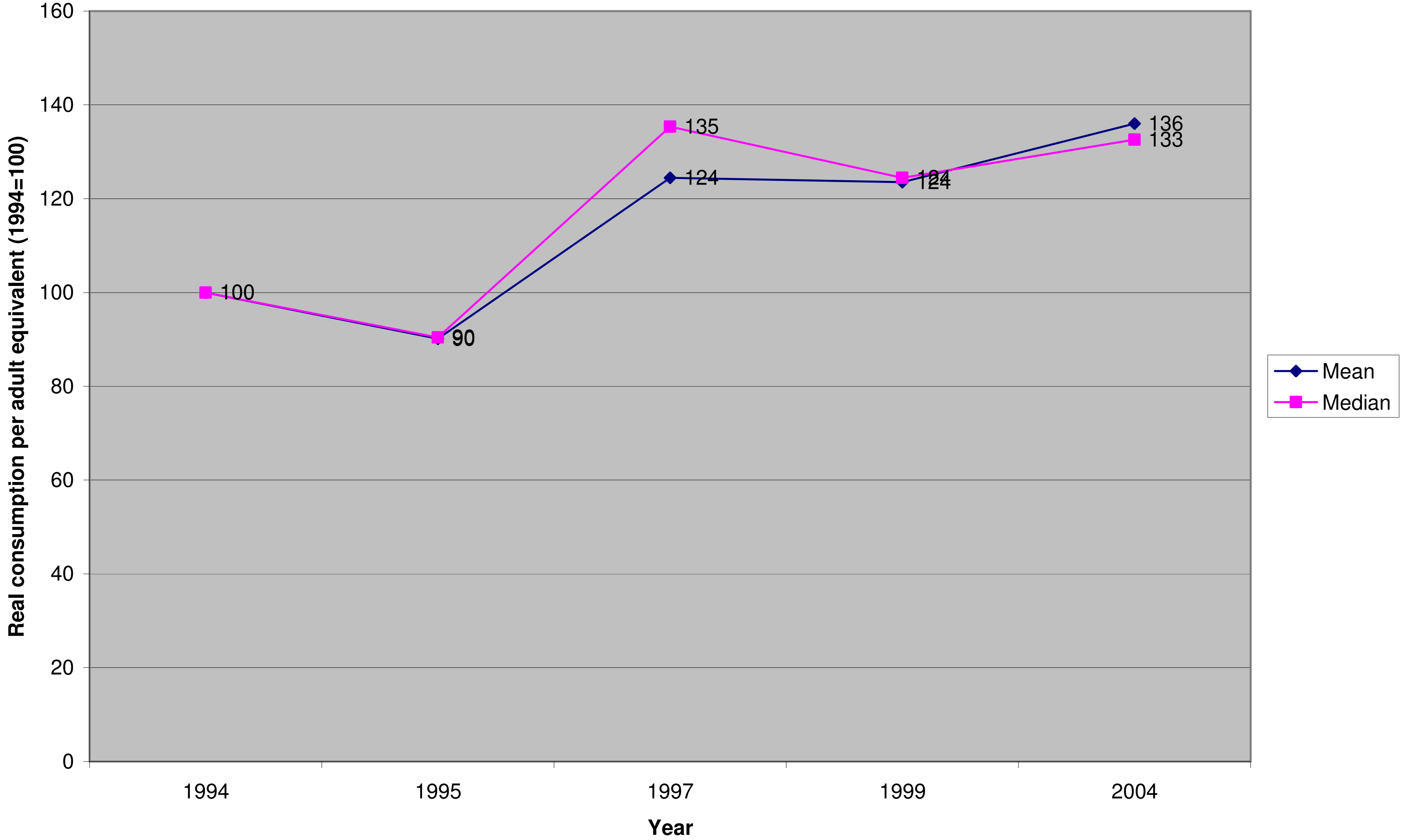


Figure 2: Percent households poor

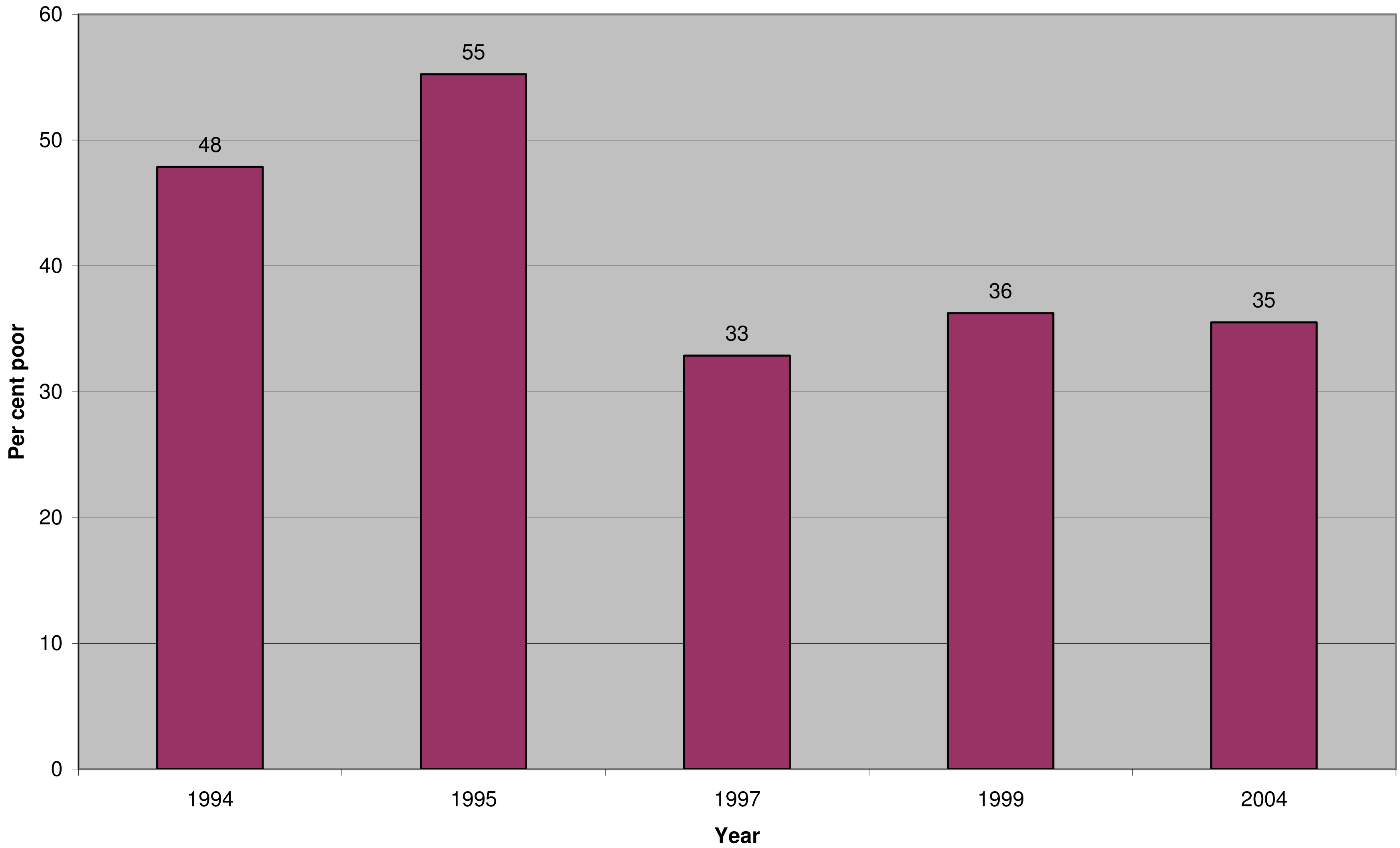
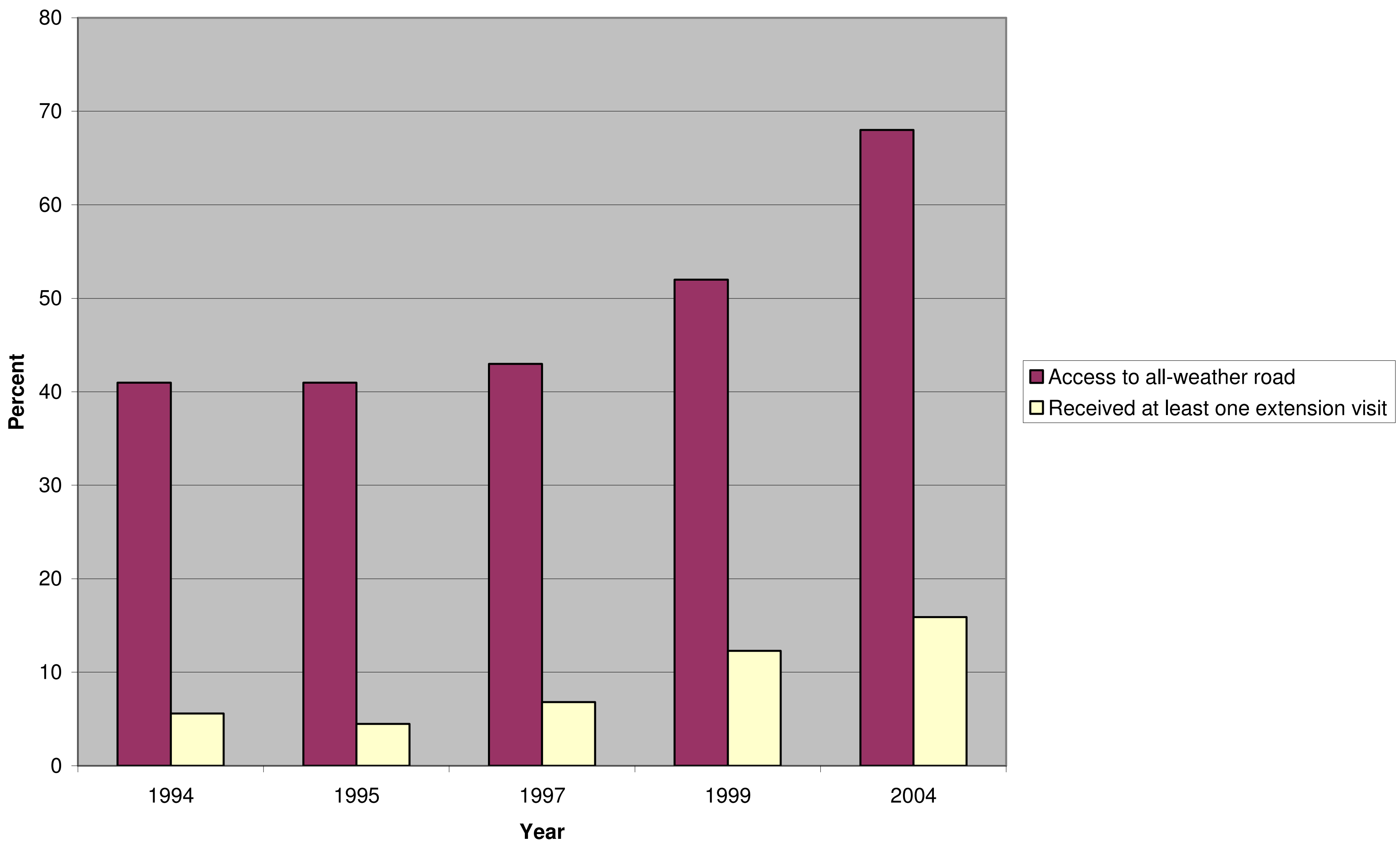


Figure 3: Access to public investments



[1] Fan and Chan-Kang (2005) provide a recent review of the literature on the links between infrastructure development, agricultural growth and poverty.

[2] The PA was responsible for the implementation of land reform following 1974 and held wide ranging powers as a local authority. All land is owned by the government. To obtain land, households have to register with the PA and, thus, lists are maintained of the households who have been allocated land. These household lists were a good source of information for the construction of a sampling frame.

[3] Pastoral areas were excluded, in part, because of the practical difficulties in finding and resurveying such highly mobile households over long periods of time.

[4] We examined whether this sample attrition is non-random. Over the period 1994-2004, there are no significant differences between attriters and non-attriters in terms of initial levels of characteristics of the head (age, sex), assets (fertile land, all land holdings, cattle), or consumption. However, attriting households were, at baseline, smaller than non-attriting households. Between 1999 and 2004, there are some significant differences by village with one village, Shumsha, having a higher attrition rate than others in the sample. Our survey supervisors recorded the reason why a household could not be traced. Using these data, we examined attrition in Shumsha on a case-by-case basis, but could not find any dominant reason why households attrited.

[5] For example, the percentage of households reporting owning hoes rises from 59 to 79 percent; owning ploughs rises from 79 to 87 per cent; and owning beds rises from 49 to 58 per cent.

[6] These trends in poverty are very similar to those reported for the country as a whole in World Bank (2005, Figure 1.1, p. 10).

[7] Note that in the case of the poverty status regressions, we estimate these as linear probability models. While it is technically feasible to estimate logit models with fixed effects, doing so carries several costs. First, the estimation of these automatically drops all observations where poverty status does not change leading to a selected sample. We cannot use GMM so estimates derived from these models are not fully efficient and the estimated coefficients are not readily interpretable in terms of their marginal effects. Given this, we follow the lead of de Janvry, Finan, Sadoulet and Vakis (2006) and Hyslop (1999) and use a linear probability model.

[8] See Dercon, Hoddinott and Woldehanna (2005) for a detailed description of these.

[9] GMM standard errors are robust to heteroscedasticity of unknown form. However, Moulton (1990, p. 334) has noted, “It is reasonable to expect that units sharing an observable characteristic, such as industry or location, also share unobservable characteristics that would lead the regression disturbances to be correlated.” These correlations, if positive, may cause the estimated standard errors to be biased downwards. In the statistics literature, this issue is referred to as the design effect, see Kish (1965) and Deaton (1997). While it is possible to correct for this intra-cluster correlation, work by Angrist and Lavy (2002) and Wooldridge (2003), suggests that doing so here would be invalid because we have only a relatively small number of clusters. For this reason, we have not reported cluster robust standard errors here. As a check, however, we did re-estimate columns (3) and (4) with this correction. It made no substantive difference to the results for consumption growth or poverty.

[10] We also undertook two additional robustness checks that we do not report in Table 4. First, we included a full set

of survey round by village dummy variables as additional regressors. While doing so carries a cost – we can no longer identify the effect of changes in road quality – it does allow us to assess whether the impact of extension is robust to a full set of fixed and time varying locality controls. When we do so, we obtain a coefficient on extension that is only slightly lower than that reported in Tables 3 and 4, suggesting that the results of extension are robust even with the inclusion of these additional controls. Second, we interacted selected household characteristics observed at the start of the survey – land holdings and literacy – with access to improved roads and extension to see whether the impact of these public services varied by household type. We find some evidence of interaction effects – land size interacted with road quality and literacy interacted with access to extension both have positive coefficients, but the prob values associated with these coefficients tend to hover around 0.10-0.15.