Effects of Population Density on Smallholder Agricultural Production and Commercialization in Rural Kenya

Milu Muyanga and T.S. Jayne

Department of Agriculture, Food and Resource Economics
Agriculture Hall, Michigan State University,
East Lansing, MI, 48824

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2012 AAEA Annual Meeting, Seattle, Washington, August 12-14, 2012.

EFFECTS OF POPULATION DENSITY ON SMALLHOLDER AGRICULTURAL PRODUCTION AND COMMERCIALIZATION IN RURAL KENYA

Milu Muyanga and T.S. Jayne[1]

Abstract
This study analyzes the implications of increasing population density in Kenya's rural areas on smallholder production and commercialization. Using data from five panel surveys on 1,146 small-scale farms over the 1997-2010 period, we use econometric techniques to determine how increasing rural population density is affecting farm household behavior and its implication to smallholder commercialization. We find that farm productivity and incomes tend to rise with population density up to 600-650 persons per km^2; beyond this threshold, rising population density is associated with sharp declines in farm productivity. Currently 14% of Kenya's rural population resides in areas exceeding this population density. The study concludes by exploring the nature of institutional and policy reforms needed to address these development problems.

Keywords: Land, population density, smallholder agriculture, food security, policy, Kenya

1. INTRODUCTION

The structural transformation process has long been considered the main route through which poverty and hunger in Africa would be overcome. A major feature of the structural transformation processes achieved in other parts of the world such as Asia was *broad-based* and *small farm-led* agricultural growth. A fundamental element of the structural transformation process was smallholder commercialization -- a transition from subsistence to market-oriented patterns of production and input use. Smallholder commercialization refers to a cycle in which farmers intensify their use of productivity-enhancing technologies on their farms, achieve greater output per unit of land and labor expended, produce greater farm surpluses, expand their participation in markets, and ultimately raise their incomes and living standards. Recent analysis has documented a declining trend in average farm size in sub-Saharan Africa, due to population pressures and an exhaustion of

[1] Milu Muyanga is Research Fellow, Tegemeo Institute, Egerton University and PhD candidate, Department of Agricultural, Food, and Resource Economics, Michigan State University. T.S. Jayne is Professor, International Development, Department of Agricultural, Food, and Resource Economics, Michigan State University, currently on leave with the Zambian Food Security Research Institute, Lusaka, Zambia. Correspondence: 207 Agriculture Hall, Department of Agricultural, Food, and Resource Economics Michigan State University East Lansing, MI 48824-1039 USA, Email: jayne@msu.edu.

the arable land frontier, especially in the productive highland regions (Table 1). In many countries, over 50% of the farms are below one hectare; 25% are virtually landless (Jayne et al., 2003). Small farm sizes and land market imperfections have contributed to land access problems in many areas, which when combined with low productivity and market access problems, results in low levels of smallholder commercialization. Farm household surveys across the region commonly find that a minority of smallholder farms produce a food surplus. Important policy issues therefore revolve around whether most farms are becoming, or have already become, "too small" to generate meaningful production surpluses and participate in broad-based inclusive agricultural growth processes given existing on-shelf production technologies.

In this paper we examine the impact of emerging land constraints and rising population density on smallholder agricultural productivity and commercialization in Kenya, and the factors that condition these relationships. Specifically, the study merges spatial remote sensing data on population density with panel survey data to estimate household input demand and output supply functions to understand the effects of localized population density on smallholder commercialization. We also examine the degree to which market access conditions and agro-ecological potential condition the relationship between population density and farmer input use and farm commercialization. We also address the potential endogeneity of population density through the using instrumental variables. The major methodological contribution of this paper is that it provides a useful method for dealing with an endogenous covariate, in this case localized population density, in the context of non-linear panel data models where the dependent variables are corner solution response variables. The potentially endogenous covariate necessitates the use of the two-step control function approach.

The study draws from various data sources. Fist, rural population data is available from the past five national censuses carried out in 1969, 1979, 1989, 1999, and 2009. More disaggregated data on rural population, land under agriculture, and unutilized land suitable for agriculture within 10 km^2 pixels are used from the Global Rural-Urban Mapping Project (GRUMP). We combine this spatial data on land use and population density with the nationwide Egerton University/Tegemeo Institute Rural Household Survey, a panel dataset tracking roughly 1,300 small-scale farm households in 5 survey waves over the 13-year period from 1997 to 2010. Over these 5 panel surveys, 1146 smallholder households were consistently interviewed for this analysis.

The preliminary results from this study indicate that smallholder landholding sizes are gradually declining in Kenya as in much of sub-Saharan Africa. Consequently, there is a rising strain on rural livelihoods in the densely populated rural areas due to land pressures and declining farm sizes. Smallholder productivity tends to rise with population density up to about 600 persons per km^2; beyond this threshold, rising population density is associated with sharp declines in input use per unit of land and farm commercialization. Currently 14 percent of Kenya's rural population resides in areas exceeding this population density. Another 20 percent of the rural population is approaching this population density threshold. The results also show that increased access to input markets, passable roads and other physical infrastructural facilities considerably influence the degree of smallholder commercialization.

2. CONCEPTUAL FRAMEWORK AND HYPOTHESES

There are several alternative ways to cast the issue of emerging land constraints within smallholder farming areas in Africa. One way is to ask how various rates of change in rural population density are affecting the evolution of farming systems, including technical and institutional responses to increased land constraints. Of course, the ways in which increasing population density affects farming systems and smallholder input demand and output supply behavior is primarily through factor and food prices. Hayami and Ruttan's (1971) theory of induced innovation has repeatedly shown that changes in person-land ratios cause farmers to adapt their farming system in ways that can be predicted. Other factors constant, rising labor-land ratios cause land values to rise compared to agricultural labor, and indirectly induce farmers to adopt new technologies that are land-saving. Other seminal works examining the ways that land-abundant agricultural systems evolve in response to growing population density include Boserup (1965), Binswanger and Ruttan (1978), and Pingali and Binswanger (1988). Binswanger and McIntire (1987) argued that increases in rural population density should induce a number of changes on tropical agricultural farming systems, including declining labor productivity, decreased fallows, increased landlessness, the development of land, labor and informal financial markets, and declining livestock tenancy. As rural communities become more heavily populated, farmers transition from shifting cultivation to annual cropping of the same plots. Fallows are reduced and more labor time is devoted to each unit of land produced, e.g., weeding labor per hectare rises. Farmers further search for land-saving technologies such as fertilizer and hybrid seed to raise the returns to the scarce factor of production (land). Given this kind of innovation, Binswanger and McIntire argue that through input intensification farmers can raise land productivity, and maintain or raise labor productivity growth even in the context of rising labor/land factor proportions. This literature has largely explained how many agricultural systems in Africa over the past 100 years have transitioned from one end of the continuum, shifting cultivation, to the other side of the continuum, intensive annual or multiple cropping with less and less land being held in fallow to restore soil productivity.

However, this literature for the most part has not considered what lies beyond the end of the continuum of annual and multiple cropping in the context of emerging land constraints and ever smaller farm sizes in increasingly densely populated rural areas. In the past two decades since these seminal articles were written, there is evidence of increased population pressures within many smallholder farming areas. Can land intensification be increased on ever smaller farms without incurring diminishing returns and scale-diseconomies? This leads to another set of research questions about appropriate and feasible smallholder-led agricultural strategies in the context of land constrained farming systems and limited off-farm employment opportunities to absorb redundant labor in densely populated rural areas. Important policy issues therefore revolve around whether most farms are becoming, or have already become, “too small” to generate meaningful production surpluses and participate in broad-based inclusive agricultural growth processes given existing on-shelf production technologies. This is the primary question that this study addresses. While we will not be able to fully address this question, our aim is to examine how densely populated farming areas are evolving compared to less densely populated areas, and to assess whether farm households in the densely populated areas are able to generate sufficient farm surpluses and incomes through agriculture (given existing technologies) to reduce rural poverty.

Our hypothesis is that farm households in the relatively densely populated areas will exhibit evidence of declining farm size, constraints on farm intensification, and lower surplus production leading to lower commercialization, incomes and asset wealth, especially per labor unit, than households in less land-constrained areas.

3. Data

Rural population data is available from the past five national censuses carried out in 1969, 1979, 1989, 1999, and 2009. More disaggregated data on rural population, land under agriculture, and unutilized land suitable for agriculture within 10 km^2 pixels are used from the Global Rural-Urban Mapping Project (GRUMP).[2] We also draw from the nationwide Egerton University/Tegemeo Institute Rural Household Survey, a panel dataset tracking roughly 1,300 small-scale farm households in 5 survey waves over the 13-year period from 1997 to 2010. The sampling frame for the panel was prepared in consultation with the Kenya National Bureau of Statistics (KNBS) in 1997. Twenty four (24) districts were purposively chosen to represent the broad range of agro-ecological zones (AEZs) and agricultural production systems in Kenya.[3] Next, all non-urban divisions in the selected districts were assigned to one or more AEZs based on agronomic information from

[2] See http://sedac.ciesin.columbia.edu/gpw/docs/UR_paper_webdraft1.pdf.

[3] Since the study was conducted, the administrative units under the New Constitution have been changed from Districts to Counties, although the physical boundaries are often different.

secondary data. Third, proportional to population across AEZs, divisions were selected from each AEZ. Fourth, within each division, villages and households in that order were randomly selected. In the initial survey in 1997, a total of 1,500 households were surveyed in 109 villages in 24 districts within eight agriculturally-oriented provinces of the country. Subsequent surveys were conducted in June of 2000, 2004, 2007 and 2010. Over these 5 panel surveys, 1243 household were able to be consistently located and surveyed. For this analysis, farms over 20 hectares (50 acres) were dropped to retain the study's focus on smallholder agriculture. Households in the coastal areas were also excluded because farming is found to account for a relatively small share of household incomes. This leaves a balanced panel of 1146 households surveyed consistently in each of the five years.

The surveys collect information on demographic changes, movements of family members in and out of the household since the prior survey, landholding size, land transactions and renting, farming practices, the production and marketing of farm products, and off-farm income-earning activities.[4]

We superimposed the longitude-latitude coordinates of the 109 villages in the Tegemeo survey on the $10km^2$ pixel population density estimates from the Global Urban-Rural Mapping Project database for 2009, to obtain population density estimates for each village. Population densities in the sample ranged from 44 persons per km^2 in the case of Laikipia West to 965 persons per km^2 in Vihiga District. We then stratified these 109 villages into five population density groups, or quintiles. Population densities range from 30 to 147 persons per km^2 in the lowest quintile, 148 to 313 in the second quintile, 315 to 470 in the third quintile, 475 to 655 in the fourth quintile, and 659 to 1135 persons per km^2 in the highest quintile. We then examine how the five groups are evolving differently over the 1997-2010 period in terms of three main features:

i. *Farming patterns:* changes in farm size, land values, rental rates, land-to-labor ratios, input intensity per unit of land cultivated and cropping patterns. The 2007 survey also contains a module exploring household members' inheritance of land and the amount of land controlled by their parents.
ii. *Farm production, assets and household incomes:* changes in incomes from crops, animal production and sales, and non-farm income as well as household asset holding.

4. Econometric Models

To study the effect of population density on specific behaviors or outcomes for household *i* in time *t* (y_{it}), we estimate a series of reduced form unobserved panel effects models for

[4] Each of these surveys instruments, which contain the details of the types of information collected and used in this study, can be viewed and downloaded at http://www.aec.msu.edu/fs2/kenya/index.htm.

the following dependent variables: farm size and area under crop cultivation; intensity of cash inputs use as a measure of the level of agricultural land intensification; and indicators of household welfare such as incomes per adult equivalent and asset holding. The models take the form:

$$y_{it} = \alpha_i + X_{it}\beta + W_{it}\eta + R_i\lambda + D_t\kappa + \mu_{it}\,,\ i = 1,2,...,N\,,\ t = 1,2,...,T \quad (1)$$

where X_{it} is a vector of household-level time-varying variables; W_{it} is a vector of village-level time-varying variables; R_i is a vector of village-level time-constant variables; and D_t is a vector of survey year dummies. The letter α_i represents the unobserved, time-constant heterogeneity that affect y_{it} while μ_{it} is the error term.[5] The vector X_{it} include variables such as distances to infrastructural facilities and services; W_{it} includes village-level population density (the main variable of interest), input prices (agricultural wage rates, land rental rates, and fertilizer prices), rainfall quantity (6-year moving average of annual rainfall prior to each survey) and rainfall variability (6-year moving average of the percentage of 20-day periods during the main growing season in which rainfall was less than 40 mm) indicators. The R_i vector includes land quality (potential kilocalories from $10km^2$ pixel land area) and agro-ecological dummies capturing other village-level time-constant characteristics. We also test for potential non-linear relationships between the dependent variables and population density by including squared, and if necessary, cubed density terms.

If the model outlined in (1) represents the true data generating mechanism, then the existence of correlation between independent variables and unobserved heterogeneity, if uncontrolled for, would result in inconsistent estimates in applied research. With panel data, there are two popular methods for estimating this model, fixed and random effects, each with their own benefits and costs. The main drawback of the random effects estimator is that it relies on the fairly strong, and in our case infeasible, assumption that the unobserved heterogeneity is uncorrelated with any of the observed independent variables. The fixed effects estimator relaxes this assumption, but at the cost of not being able to include any time-constant covariates, such as the locations where sampled households are situated. To overcome these shortcomings of both fixed and random effects estimators, Mundlak (1978) and Chamberlain (1984) propose a framework known as the *correlated random effects estimator* (CRE) or the Mundlak-Chamberlain device. In this approach,

[5] Omitted variables are the main source of unobserved heterogeneity, and they may fall into two categories: those that do not vary much across time (e.g., distance from the farm to the district town), which are easier to control for with panel analysis techniques as used here, and those that are time-varying (e.g., random shocks affecting households). For details on unobserved heterogeneity and methods for addressing it, see Wooldridge (2010).

rather than assuming the unobserved and observed explanatory variables are uncorrelated, α_i is modeled and the correlation is assumed to take the form:

$$\alpha_i = \delta + \overline{C}_i \lambda + \varsigma_i, \ \varsigma_i \mid C_i \sim N(0, \sigma_\varsigma^2) \tag{2}$$

where $\overline{C}_i$ represents the time-averaged value of all time varying variables (X_{it} and W_{it}) over the various panel periods. The main benefits of the CRE estimator are that (1) it controls for unobserved time-constant heterogeneity, and (2) because the assumption of correlation between the covariates and unobserved heterogeneity is modeled, the random effects estimator is applied, which allows also the measurement of the effects of time-invariant independent variables (see Wooldridge 2010 for details).

While equations (1) and (2) are linear in parameters, and thus easily estimated by any single equations estimator, the population density variable is potentially endogenous in equation (1). There is a possibility that some unobservables that influence agricultural production and household welfare are likely to influence population growth. When confronted by endogeneity, two methods are available to circumvent the problem. First is the usual instrumental variable (2SLS) method and the second is the control function (CF) approach (Wooldridge 2010). While the two methods yield the same results, the CF approach leads to a straightforward exogeneity test of the potentially endogenous variable. We therefore apply the CF approach in the paper. The CF approach involves two-step estimation procedure. In the first step, we estimate

$$d = z\pi + v, \ v \mid z \sim N(0, \sigma_v^2) \tag{3}$$

where d is the population density in village g at survey period t, z is a vector of exogenous variables (which includes unity as its first element), π's are the coefficients to be estimated; and v is a random error term. The vector z is supposed to contain at least one element that is not in equation (1) for identification purposes. In our case the vector z includes the population estimates from the 1969 and 1979 censuses for each village; factors measuring access to markets and infrastructure; and rainfall quantity and variability variables as well as small agro-ecological dummies to capture general agro-ecological potential in the villages where the households are found.

In the second step we estimate specifications (1) but this time plugging in the residual, $\hat{v}$, from (3) using the CRE approach. As Wooldridge (2010) shows, plugging $\hat{v}$ into equations (1) breaks the endogeneity link between the potentially endogenous variable and the error term in equation (1). The time-varying explanatory variables in both steps are lagged by one survey period for two reasons. First, while some explanatory variables may affect asset stocks contemporaneously, most of the variables are expected to influence asset stocks after a lag. For example, changes in the distance to infrastructural facilities and services often do not affect agricultural production and household asset accumulation immediately; these effects tend to appear with a lag. The second reason is to circumvent any other potential endogeneity problem arising from omitted variable problems. It is important to note that since the estimation of equation (1) involves generated regressor ($\hat{v}$), standard errors generated by most econometric software for the coefficients are bound to invalid since they ignore the sampling variation in the estimation of π in the first step. Disregarding the sampling error in the generated regressors ($\hat{v}$) is likely to underestimate the computed standard errors in equation (1). Consequently, we use the bootstrap approach with 500 replications to get a valid estimate of the standard errors. Inferences from equation (1) are made fully robust to arbitrary heteroskedasticity and serial correlation (Wooldridge 2010).

5. Descriptive Results

Table 2 presents information on farm size and farming practices by village population density quintiles over the four survey years. Landholding sizes per adult equivalent in the 20% most densely populated villages (0.31 hectares over the four survey years) are roughly one third of those in the low density quintile (0.92 hectares). The areas under cultivation have consistently declined for all five population density categories over the 10-year period by about 23 percent. The areas cultivated in the highest density (HD) quintile (0.89 hectares) are about half of those in the lowest density (LD) quintile (1.80 hectares). These differences between the top and bottom quintiles of farm size and area under cultivation are significant at the 95 percent confidence level. The proportion of farmland under fallow has also declined slightly over time for all population density quintiles. Family labor per hectare cultivated has generally increased over the 13-year period, and is significantly higher in the HD quintile than all other density quintiles. All of these indicators point to land being an increasingly constraining factor of production in smallholder agriculture in the high-density areas in Kenya.

Land constraints may also explain why the HD areas tend to be devoting a greater share of their cropped area to higher-valued crops and less to maize, a relatively low-value crop (data not shown to conserve space). Villages in the HD quintile put less than 3% of their land to monocropped maize, compared to 6% in the lowest density quintile. Maize intercrops account for 39% of cropped areas in the highest density areas compared to 42% in the lowest density areas. However, this difference is not statistically significant. By contrast, the HD areas devote a significantly greater share of their land to industrial cash crops such as tea, coffee, and sugarcane compared to the bottom two density quintiles. Similarly, fresh fruits and vegetables account for 26% of cropped area in the HD areas, compared to 13% in the 20 percent LD villages. Also, the percentage of households practicing zero-grazing increases with population density from a low of 4% in the first population density quintile, reaching a high of 56 percent in the 4th quintile and declining to 31 percent in the HD quintile. Similarly, Table 2 shows that the intensity of purchased inputs (mainly fertilizer, improved seed, and hired labor) per unit of land is an increasing function of population density up to the 4th density quintile, but then declines significantly from the fourth to the highest density quintile. The greater focus on high-value crops and more intensive land-saving dairy production in the densely populated regions maximizes revenue per scarce unit of land owned. This is a result that will be explored further in the econometrics section of the paper.

Land values, collected in 2010 were more than twice as high in the three highest population density quintiles than in the LD quintile. Conversely, agricultural wage rates in the LD villages were 30% lower than in the HD villages (Table 2). The overall picture from Table 3 is that farming practices in the areas of high population density are distinctly more land-intensive and are focused more on higher-value crops than in the low density areas.

Not shown in Table 3, but also of importance is how population density is related to the amount of land inherited from the previous generation. Respondents in the 2007 survey were asked how much land was owned by the father of the household head. The previous generation had considerably larger farms (3 times larger) than those of the current survey respondents themselves. The mean size of parents' farms varied from 7.80 hectares in the LD areas to 4.41 hectares in the HD areas. Survey respondents were also asked about the amount of land inherited by the household head from his father. This ranged from 1.49 hectares in the LD quintile to 0.89 hectares in the HD quintile. The mean amount of land inherited was roughly one-fifth of the total landholding size of the father. This might be explained by the fact that fathers in patriarchal Kenya tend to subdivide their land among sons. An important policy question might be how the current generation of adults in the high population density areas with 1.30 hectares of land or less are going to subdivide their land among their children when they reach their old age (the average age of household heads was 48 years in 2010) and whether farming can provide a viable livelihood for those remaining on the land. We speculate that, because farm sizes in the high density areas are already quite tiny and cannot be meaningfully subdivided much further, an increasingly smaller fraction of people born on farms in Kenya will be able to remain there. This may point to even higher rates of rural-to-urban migration in the future, or at least from agriculture to non-agriculture.

Table 3 presents trends in farm production, income, and asset wealth over the panel period by village population density quintiles. The value of net crop income (gross crop income minus input costs per hectare), a measure of partial land productivity, increases with population density up to the fourth density quintile and declines thereafter. As shown by results in Table 3, high population density areas are cultivating their scarce land more intensively by applying more labor and cash inputs per hectare cultivated, at least up to a certain threshold corresponding to the fourth highest population density quintile, which ranged from 531 to 678 persons per km^2. Similarly, the value of net farm income (from crops and animal products) per hectare also is an increasing function of population density up to a certain level corresponding to the fourth-highest quintile. By contrast, the value of farm income per family labor unit appears to be higher among the villages in the middle population density quintiles. This measure of partial labor productivity is perhaps the more meaningful of the two productivity measures because it more accurately reflects the implicit return to an individual. Table 3 also shows that off-farm income per adult equivalent is slightly higher for households in the low density areas, possibly reflecting a lower supply of labor in these areas (note also from Table 2 that agricultural wage rates were also higher in the low density areas than in high density areas).

Possibly the most important indicator discussed in this section is the value of asset wealth per adult equivalent. The list of productive assets consistently collected and valued in each of the four surveys includes ploughs, tractors and draft animal equipment, carts, trailers, cars, trucks, spray pumps, irrigation equipments, water tanks, stores, wheelbarrows, combine harvesters, cows, bulls, donkeys, and smaller animals. Recent studies in the poverty literature (e.g., Carter and Barrett, 2006; Krishna et al, 2004) argue that the value of assets more accurately measures wealth than income or consumption, as it is less susceptible to random shocks, and is likely to be a more stable indicator of household welfare. This is especially true in regions where rain-fed agriculture is a major source of annual income and where households rely greatly on their physical assets for their livelihoods. For these reasons, we consider asset holdings to be an important measure of household livelihood, productive potential, and safety net.

Table 3 shows that households' asset wealth per adult equivalent has been consistently higher (more than twice) in households located in the low population density areas. Family size in adults and adult equivalents is almost the same across all five population density quintiles, meaning that asset wealth per household is also substantially higher on average in the low density areas. Conversely, aggregate household income tends to rise with population density, once again up to the fourth quintile, and thereafter starts to decline.

Table 4 presents trends in smallholder farm commercialization indicators over the panel period by village population density quintiles. Just as expected, as implied by production trends in the previous two tables, the value of crop and livestock products (eggs, goat milk and cow milk, honey) sales and milk increase with the population density up to the 4th population density quintile and there after falls. The increasing smallholder commercialization with population can be explained from two fronts. First, when there are land constraints, the existence of alternative production possibilities creates a conflict of allocation of land resources to alternative crops. Thus, the decreasing farm sizes motivate farmers to alter production patterns to make the best out of their shrinking. Results not presented here due to lack of space show that households in landholding constrained regions devote a greater proportion of their land to fresh fruits and vegetable production. Second, the high population density has been shown to lead to improved economic infrastructures, owing to reduced average cost of infrastructure, which in turn leads to improved input and output market access. Improved market access motivates farmers to produce a surplus for the market and to engage in high value enterprises that are not possible with poor market access. The high value enterprises include production of perishable fresh fruits and vegetables, dairy and poultry products. Smallholder agricultural commercialization is a fundamental element of the structural transformation process - a transition from subsistence to market-oriented patterns of production and input use.

The bivariate relationships presented in Tables 2, 3 and 4, while providing a fairly consistent picture, do not control for the effects of other variables affecting farm productivity, incomes and asset wealth. However, these relationships do lead to an important hypothesis for more rigorous analysis in the next section. Specifically, are there threshold effects of population density that cause input use intensity, productivity, and incomes to decline beyond a certain point? And if this is found to be the case, what are the causes of this threshold effect?

6. Econometric Results

This section reports main results from the econometric regressions of the impact of increasing population density on some selected indicators of farm productivity and welfare. We first discuss the estimates from the first-stage model of population density determinants, followed by the second-stage control function results.

Table 5 presents the first-stage results of the drivers of population density growth. The variable capturing potential soil quality is measured at three different levels, namely, the potential kilocalories obtainable from the $10km^2$ location based on (i) existing cultivated land; (ii) existing cultivated land plus grasslands; and (iii) cultivated, grassland, and forest lands. The first and second indicators might reflect food production potential in the short- and medium-run, while the third indicator would reflect longer-term potential, and would obviously present major environmental trade-offs. Generally the results shown in Table 5 indicate that the major determinants of population density in 2009 include distances to infrastructural facilities, the population of the location in prior decades, and area sizes; village-level rainfall quantities, rainfall variability and soil quality; as well as the agro-ecological zones where these villages are located. For example, if households in Location A are one kilometer closer to motorable roads, water sources, and healthcare facilities than households in Location B, the population density in Location A is estimated to be 0.32, 0.57 and 0.17 percent higher than in Location B. If Location A's long-run average annual rainfall is 100mm higher than Location B, the population density of Location A is estimated to be ten percent higher than Location B. Increased rainfall variability is associated with lower population density. As expected, land quality as represented by the potential kilocalories obtainable from each 10km^2 pixel of both cultivated land increases population density by 7.2 percent.

Next we present the second-stage control function regression results of the impact of increasing population density on selected agricultural production and household welfare outcomes. Because of space limitations, we focus on only a few outcome variables of interest. By including the agro-ecological zones among the other controls, the results presented in this section have a "within zone" interpretation. This means that the relationship between village population density and outcome variables hold constant the variation in the outcome variables that might occur due to differences in agro-ecological potential. The same holds true for unobserved time effects through the inclusion of survey year dummy variables. Since the third land quality variable was not statistically significant in the first stage, we present the results using the first two land potential variables only.

Landholding size and cropped area

Tables 6 and 7 regression results indicate that landholding sizes and area in hectares cultivated per adult equivalent in the main season decline with population density. Controlling for all other variables shown on Table 6, an increase in population density by 100 persons per km^2 is associated with 9 percent smaller farm sizes. A similar increase in population density reduces area cropped per adult equivalent by about 8 percent. These relationships are highly statistically significant. A more complete presentation of these relationships is revealed when we look at the post-estimation simulations of the relationships between these outcome variables and population density, holding all other factors constant. Figures 1(a) and 4(b) show that landholding size and area cultivated per adult equivalent varies inversely with population density.

Input use intensity

Table 8 presents results on the cost of purchased inputs per hectare (fertilizers, seeds, hired labor, and land preparation costs), which is an indication of land intensification. The results show that the intensity of purchased inputs per hectare is a non-linear concave function of population density. Input intensity rises with population density to around 650 persons per km2; beyond this population density threshold, input intensification declines. Further, the intensity of purchased input use rises as land rental rates rise, and declines with increased distances to motorable roads, signalling increased input costs. The intensity of purchased input use also declines as we move from the relatively high-rainfall Central Highlands region (base region) to the more semi-arid lowlands.

Figures 1(c) and (d) show the simulated relationship between input use intensity on the y-axis and population density on the x-axis, controlling for all the other variables. The results show that both fertilizer use and the use of all purchased inputs per hectare is an increasing function of the population density up to roughly 660 persons per km^2, and then declines beyond that. Slightly less than 20% of the farm households in the sample are currently beyond this threshold (Figure 4c). As shown in Figure 4d, the general input use intensity starts to decline somewhere after 475 persons per km^2; about 35 percent of the households in the sample live in villages beyond this population density threshold.

What would explain these threshold effects? Market participation studies consistently show that farm sales are related to farm size (Barrett 2008). If farm sizes decline beyond a given point due to sub-division and land fragmentation caused by population pressures, households are less likely to generate the cash from crop sales that would allow them to purchase modern productivity-enhancing inputs. Less intensive input use then reinforces small farms' difficulties in producing a surplus. Furthermore, access to farm credit also tends to be restricted for farmers with limited land and other assets that could otherwise act as collateral. For these reasons, we feel that population density threshold effects are very plausible and may explain why a number of important farm productivity indicators tend to decline beyond a certain level of population density.

Household farm income

Tables 9 and 10 present the regression on net farm incomes per hectare and per adult equivalent, respectively. The CRE model estimates show that net farm incomes per hectare increase with population density up to about 680 persons per km^2, but fall off slightly thereafter. Net farm incomes per adult equivalent, by contrast, shows a more sharp decline at a lower population density threshold of about 550 persons per km^2, following a pattern very similar to input use intensification. All of these relationships are highly statistically significant. The subsequent post estimation simulation results are presented in Figures 1(e) and (f). The results also show that lower distances to motorable roads are associated with higher farm incomes. Higher farm wage rates, land rental rates, and fertilizer prices are all significantly associated with lower farm incomes per adult equivalent (Table 10); only the first two input prices are significantly associated with farm incomes per hectare. As expected, increased rainfall level (variability) is found to be associated with higher (lower) farm incomes.

These results apply to both crop and animal operations; results are similar when the dependent variable is net crop income or net animal income. Intensive animal operations such as zero grazing dairy is significantly more commonly practiced in the high density areas, producing higher levels of animal income per land unit. However, this is only possible up to a certain population density level, beyond which, the land sizes become too small for economical operations. This evidence of a decline in partial land productivity at high levels of rural population density is alarming, as it implies that land productivity growth cannot be sustained simply by using other inputs more intensively per unit of land. Animal income and milk production also show a similar plateau and drop off at 650 persons per km^2. As Kenya's rural population continues to grow,[6] a greater proportion of the country's rural areas will soon reach this apparent land productivity plateau. Currently, most of the districts with mean population density greater than 650 persons per km^2 are in Nyanza and Western Provinces, with most in Central Province approaching this threshold.[7] In 2009, the 16 districts with greater than 650 persons per km^2 accounted for 14.2% of Kenya's rural population and 1.3% of its rural land.

Household asset wealth and incomes

Lastly, we discuss the relationships between population density and household asset wealth (Table 11) and total income per adult equivalent (Table 12). The results show an unambiguously and statistically significant negative relationship between household assets and population density (Table 11). Holding constant differences in asset wealth due to differences in infrastructural conditions, input prices, rainfall quantity and variability, soil quality, agro-ecological potential and survey years, we find that an increase of 100 persons per km^2 is associated with a 7 percent decline in asset wealth per adult equivalent. This relationship is shown graphically in Figure 4(g).

By contrast, total household incomes tend to rise with population density up to a now familiar threshold and thereafter decline (Table 12). The post estimation simulations show a clearer picture of these relationships. Total household incomes per adult equivalent rise with population density up to roughly 550 persons per km^2 and decline thereafter, as shown in Figure 1(h). Higher population density is associated with smaller farm sizes, other factors constant. Small farm sizes may be associated with diseconomies of scale in input acquisition. Other factors constant, smaller farm sizes reduce the potential to produce surpluses, which may in turn cause capital constraints that impede the demand for purchased inputs and new technologies. These processes may explain why our results indicate adverse effects of population density, beyond some threshold, on indicators of farm intensification, farm income per unit of labor, and wealth.

[6] Fortunately, Kenya's rural population growth rate has been declining from its peak at 3.4% in 1984 to 2.3% in 2008 according to the 2009 official census.
[7] These districts include Emuhaya, Hamisi, Vihiga, Kisii Central, Gucha, Manga, Nyamira, Githunguri (in Central Province), Gucha South, Masaba, Kakamega South, and Kisii South. Median farm size in these districts covered in the Tegemeo sample (Vihiga, Kisii, and Kakamega) is 0.94 hectares per farm.

7. Conclusions and Implications for Institutional Reform

Generally, these findings suggest that increased access to land and markets will ensure that farm households are able to generate agricultural surpluses and consequently participate in agricultural output markets. Given the existing distribution of landholdings within Kenya's small farm sector, strategies to improve rural households' access to land will need to be not only on the country's land agenda, but also its food security and poverty reduction agendas. As the land frontier closes in many parts of Kenya and population continues to rise, smallholder farming areas will be producing fewer food surpluses in the future unless there is major productivity growth through technical innovation. There is also some scope for promoting equitable access to land through a coordinated strategy of public goods and services investments to raise the economic value of arable land in the country that is relatively remote and still unutilized. This would involve investments in road infrastructure, schools, health care facilities, electrification and water supply, and other public goods required to induce migration, settlement, and investment in these currently under-utilized areas. Through migration, such investments would also help to reduce population pressures in the densely populated areas, many of which are being degraded due to declining fallows associated with population pressure. The study also underscores the need for redoubled public investment in the national agricultural research and extension systems to focus on new farm technologies and practices appropriate for one-hectare farms or smaller. These technologies need to be land-saving. While improved land productivity can improve small farm livelihoods and food security in densely populated areas, this alone is unlikely to be a panacea for addressing land shortages that are emerging or worsening in many parts of rural sub-Saharan Africa.

Acknowledgements

This paper has been prepared under the Guiding Investments in Sustainable Agricultural Markets in Africa (GISAMA), a grant from the Bill and Melinda Gates Foundation to Michigan State University's Department of Agricultural, Food, and Resource Economics. The authors also acknowledge the long-term support that United States Agency for International Development (USAID) Kenya has provided to the Tegemeo Institute of Egerton University for the collection of panel survey data over the 10-year period on which this study draws. They also with to thank Margaret Beaver for her assistance in cleaning the data and generating the variables of interest as well as Steven Longabaugh and Jordan Chamberlin for their assistance several figures used in this article. Any errors and omissions are those of the authors only. The views expressed in this paper do not necessarily reflect the views of the Bill and Melinda Foundation, USAID, or any other organization.

References

Barrett, C.B. (2008). Smallholder Market Participation: Concepts and Evidence from Eastern and Southern Africa. *Food Policy,* 33(4): 299-317.

Binswanger, H. and Ruttan, V. (1978). *Induced Innovation: Technology, Institutions and Development*. Baltimore, MD: Johns Hopkins University Press.

Binswanger, H. andMcIntire, J. (1987). Behavioral and Material Determinants of Production Relations in Land-abundant Tropical Agriculture. *Economic Development and Cultural Change*, 36(1), 73-99.

Binswanger, H. and Pingali, P. (1988). Technological Priorities for Farming in Sub-Saharan Africa. *World Bank Research Observer*, 3 (1), pp. 81-98.

Boserup, E (1965). The Conditions of Agricultural Growth: The Economics of Agrarian Change under Population Pressure. London: Allen & Unwin.

Carter, M.R. and Barrett, C. (2006). The Economics of Poverty Traps and Persistent Poverty: An Asset Based Approach. *Journal of Development Studies*, 42(2), 178-99.

Chamberlain, G. (1984). Panel data. In Grilliches, Z., Intriligator, M.D., eds. *Handbook of Econometrics*, Vol. 2 North Holland, Amsterdam, 1247-1318.

Deininger, K. and Binswanger, H. (1995). Rent Seeking and the Development of Large-Scale Agriculture in Kenya, South Africa, and Zimbabwe. *Economic Development and Cultural Change*, 43(3) (April), 493-522.

Deininger, K. and D. Byerlee, with J. Lindsay, A. Norton, H. Selod, and M. Stickler. (2011). Rising Global Interest in Farmland. Report. The World Bank, Washington D.C.

Govereh, J. (1999). Impacts of Tsetse Control on Immigration and Household Accumulation of Capital: Zambezi Valley, Zimbabwe. PhD Dissertation, Michigan State University, East Lansing.

Hayami, Y. and Ruttan, V. (1971). *Agricultural Development: An International Perspective*. Baltimore, Johns Hopkins Press.

Johnston, B. F. and Kilby, P. (1975). *Agriculture and Structural Transformation: Economic Strategies in Late Developing Countries*. New York: Oxford University Press.

Juma, Calestous. 1996. Introduction. In Juma, C. and Ojwang, J.B. (eds.). In Land We Trust: Environment, Private Property and Constitutional Change. African Centre for Technology Studies (ACTS) environmental policy series: 7, Initiatives Publishers, Nairobi, pp. 1-5.

Jayne, T.S. Yamano, T., Weber, M., Tschirley, D., Benfica, R., Chapoto, A. and Zulu, B. (2003). Smallholder Income and Land Distribution in Africa: Implications for Poverty Reduction Strategies. *Food Policy*, 28(3), 253-275.

Jayne, T.S., Zulu, B., Kajoba, G., and Weber, M. (2009). Access to Land, and Poverty Reduction in Rural Zambia: Connecting the Policy Issues. Working Paper 34. Food Security Research Project, Lusaka, Zambia.

Kanyinga, K. (1998). Politics and Struggles of Access to Land: 'Grants from above' and 'squatters' in Coastal Kenya, *European Journal of Development Research*, 10(2), pages.

Krishna, A., Kristjanson, P., Radeny, M. and Nindo, W. (2004). Escaping Poverty and Becoming Poor in Twenty Kenyan Villages. *Journal of Human Development* 5(2), 211-26.

Mbaria, J. (2001). We Ignore Land Issues at Our Peril, *The Daily Nation*, 16 August 2001, Nairobi.

Mellor, J. (1976). *The New Economics of Growth*. Ithaca, NY: Cornell University Press.

Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica* 46, 69-85.

Munshifwa, E. (2002). Rural Land Management and Productivity in Rural Zambia: The Need for Institutional and Land Tenure Reforms. Paper presented at the Surveyor's Institute of Zambia Seminar, July 2002, Oxfam.

Namwaya, O. (2004). Who owns Kenya?, East Africa Standard, October 1, 2004. http://www.marsgroupkenya.org/pdfs/crisis/2008/02/large_landowners_in_Kenya.pdf

Okoth-Ogendo, H.W.O. (1976). African Land Tenure Reform. In Judith, H., Maitha, J.K. and Senga, W.M. (ed). *Agricultural Development in Kenya, and Economic Assessment*. Oxford University Press , Nairobi, 152-185.

Okoth-Ogendo, H.W.O. (1999). Land Policy development in East Africa, a survey of recent trends. A regional overview paper for the DFID workshop on "Land rights and sustainable development in Sub-Saharan Africa" held at Sunningdale Park conference centre, Berkshire, England, 16-19 February, 1999.

Platteau, J.(2004). Monitoring Elite Capture in Community-driven Development. *Development and Change*, 35,223-246.

Pingali, P.L. and Binswanger, H. (1988). Population Density and Farming Systems -- The Changing Locus of Innovations and Technical Change. in Lee, R. (Eds.) *Population, Food and Rural Development*. Oxford: Clarendon Press.

Republic of Kenya. (1965). *The Sessional Paper No. 10 of 1965 on African Socialism and its Application to Planning in Kenya.* Nairobi: Government Printer.

Republic of Kenya. (2004). Report of the Commission of Inquiry into Illegal/Irregular Allocation of Public Land, June 2004. Nairobi: Government Printer.
Republic of Kenya. (2010). *Constitution of Kenya*. Nairobi: Government Printer.

Stambuli, K. (2002). Elitist Food and Agricultural Policies and the Food Problem in Malawi. *Journal of Malawi Society - Historical & Scientific*, 55(2).

Wily, L.A. (2011). The tragedy of public lands: The fate of the commons under global commercial pressure. Report for International Land Coalition, IFAD, Rome. http://www.landcoalition.org/sites/default/files/publication/901/WILY_Commons_web_11.03.11.pdf. Accessed 30 October 2011.

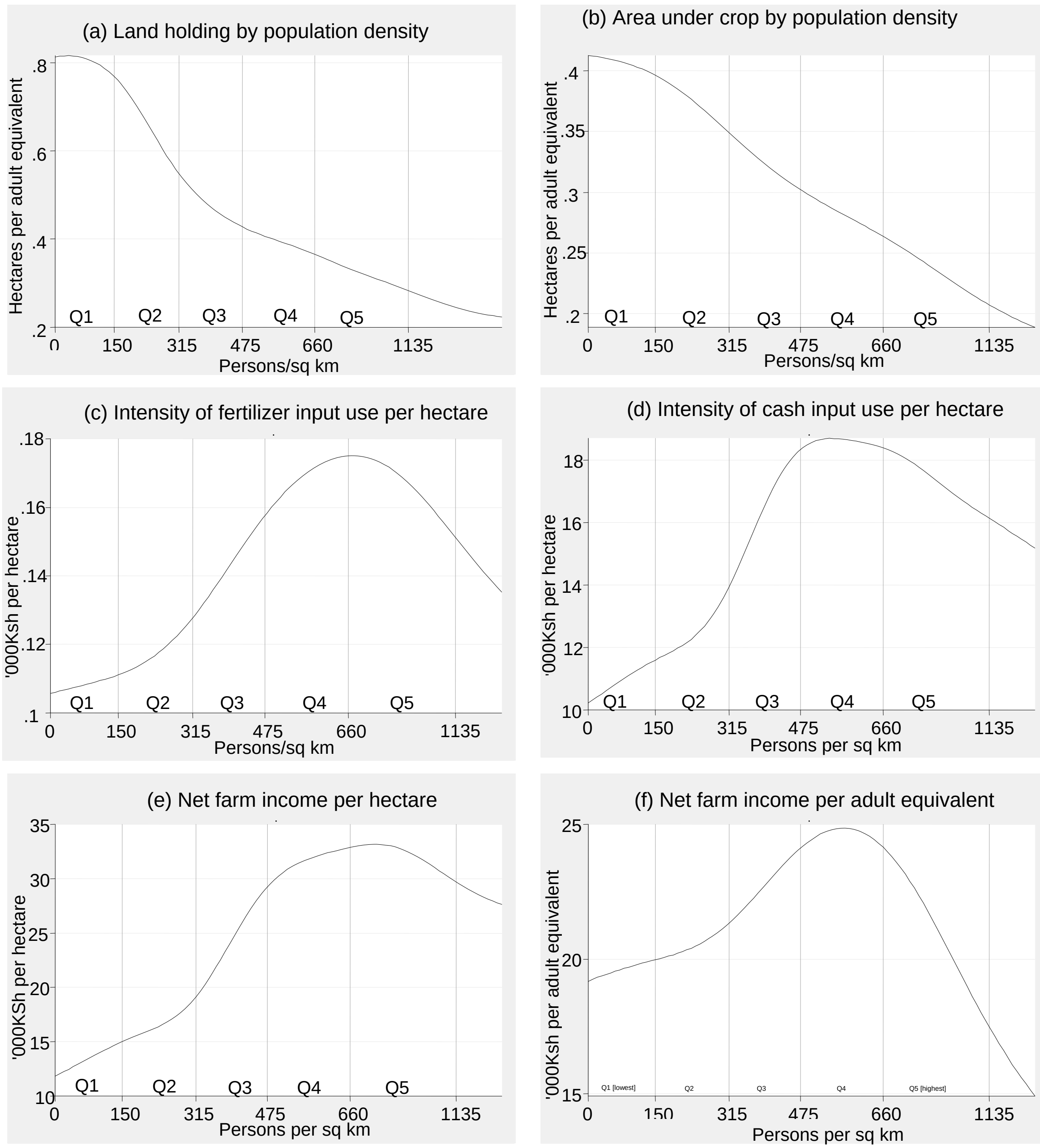
Woodhouse, P. (2003). African Enclosures: a Default Mode of Development. *World Development, 31*(10): 1705-1720.

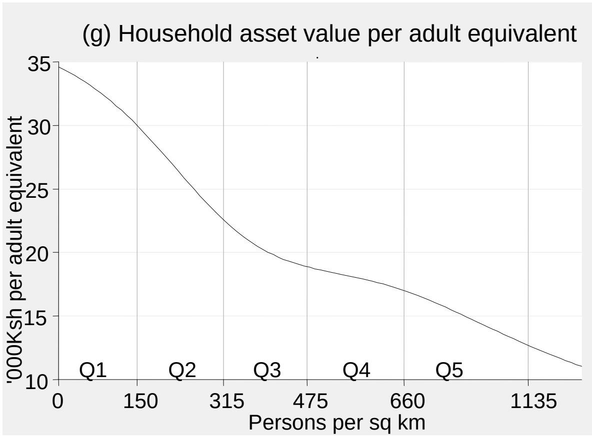
Wooldridge, J.M. (2010). *Econometric Analysis of Cross Section and Panel Data.* Second Edition, London: MIT Press.

World Bank. 2008. World Development Report 2008: Agriculture for Development, World Bank, Washington, DC.

Yamano, T., Place, F., Nyangena, W., Wanjiku, J. and Otsuka, K. 2009. Efficiency and equity impacts of land markets in Kenya. Chapter 5 in Holden, S.T., Otsuka, K. and Place, F.M. (eds), *The Emergence of Land Markets in Africa*. Washington, DC: Resources for the Future.

Fig. 1. Simulations from the econometric results showing the relationship between population density and variables of interest




(g) Household asset value per adult equivalent
'000Ksh per adult equivalent
35
30
25
20
15
10
Q1
Q2
Q3
Q4
Q5
0
150
315
475
660
1135
Persons per sq km

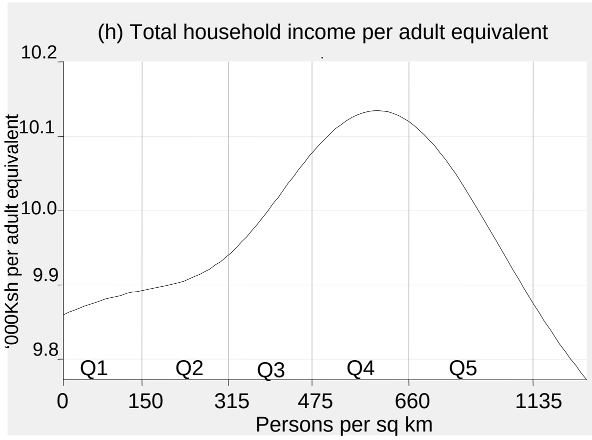

(h) Total household income per adult equivalent
'000Ksh per adult equivalent
10.2
10.1
10.0
9.9
9.8
Q1
Q2
Q3
Q4
Q5
0
150
315
475
660
1135
Persons per sq km

Table 1. Hectares of arable land per person in agriculture (10 year average) in selected countries

	1960-69	1970-79	1980-89	1990-99	2000-09[1]	2000-09 land-person ratio as % of 1960-69
Ethiopia	0.501	0.444	0.333	0.224	0.218	43.5%
Zambia	0.643	0.607	0.398	0.342	0.297	46.2%
Kenya	0.462	0.364	0.305	0.264	0.219	47.4%
Uganda	0.655	0.569	0.509	0.416	0.349	53.3%
Malawi	0.480	0.466	0.357	0.304	0.307	64.0%
Zimbabwe	0.613	0.550	0.452	0.420	0.469	76.5%
Rwanda	0.212	0.213	0.195	0.186	0.174	82.1%
Mozambique	0.356	0.337	0.320	0.314	0.294	82.6%
Ghana	0.646	0.559	0.508	0.492	0.565	87.5%
Nigeria	0.982	0.860	0.756	0.769	0.898	91.4%

Source: FAO STAT (2010)

Notes: [1]*Data on land utilization is only available for the period 2000 to 2008. Land-to-person ratio = (arable land and permanent crops)/(agricultural population). For the periods 1960-69 and 1970-79, agricultural population is estimated by multiplying rural population by an adjustment factor (mean agricultural population 1980-84/mean rural population 1980-84). This is because data on agricultural population was only collected from 1980 onward.*

Table 2. Farming practices and factor intensities, by population density quintile (all values in nominal terms)

	Pop. density quintile	Survey year				
		2000	2004	2007	2010	average
Landholding per adult equivalent (hectares)	5 [highest]	0.28	0.29	0.35	0.30	0.31
	4	0.34	0.37	0.47	0.36	0.38
	3	0.45	0.50	0.49	0.45	0.47
	2	0.55	0.61	0.56	0.63	0.59
	1 [lowest]	0.83	0.96	0.93	0.95	0.92
Area cultivated in the main season (hectares)	5 [highest]	0.99	0.94	0.85	0.79	0.89
	4	1.18	1.26	1.22	1.03	1.17
	3	1.54	1.36	1.16	0.99	1.27
	2	1.73	1.79	1.54	1.30	1.58
	1 [lowest]	1.98	1.87	1.74	1.59	1.80
Labor (number of adult members) per hectare cultivated	5 [highest]	6.04	7.15	5.94	6.43	6.39
	4	4.54	4.12	4.19	4.71	4.39
	3	5.14	5.18	4.81	4.67	4.96
	2	3.10	3.19	3.65	3.57	4.49
	1 [lowest]	3.06	3.11	3.34	3.15	3.16
Cost of purchased inputs per hectare (‘000 KSh)	5 [highest]	13.62	15.45	14.60	19.36	15.73
	4	17.13	21.26	18.98	26.63	21.07
	3	12.16	15.74	13.76	21.29	15.57
	2	5.71	12.34	13.57	17.60	12.65
	1 [lowest]	8.10	8.72	9.63	13.17	9.87
Land values /hectare (‘000KSh)	5 [highest]	-	-	-	703.02	703.02
	4	-	-	-	633.03	633.03
	3	-	-	-	723.67	723.67
	2	-	-	-	626.00	626.00
	1 [lowest]	-	-	-	271.82	271.82
Hired agricultural wage labor rate (KSh. per day)	5 [highest]	59.24	68.95	72.81	102.52	75.68
	4	71.20	93.49	95.53	137.74	100.07
	3	67.67	76.52	83.68	117.72	85.39
	2	69.38	92.56	96.07	134.21	99.97
	1 [lowest]	83.12	97.40	105.00	124.93	102.14

Source: Tegemeo Institute Rural Household Surveys.

Note: Population density quintiles are defined by ranking all households in the surveys by village-level population density and dividing them into five equal groups.

Table 3. Household farm income, income and wealth trends, by population density quintile ('000 KSh nominal)

	Pop den quintile	Survey year					
		1997	2000	2004	2007	2010	average
Net crop income per hectare	5 [highest]	27.17	55.47	48.67	57.67	64.69	50.83
	4	27.30	49.06	50.19	58.40	120.84	61.36
	3	22.35	33.71	37.88	45.98	75.44	42.35
	2	16.41	20.47	30.77	42.79	53.77	33.58
	1 [lowest]	17.40	21.74	20.06	19.13	13.85	18.51
Net crop income per unit of labor	5 [highest]	9.39	24.80	22.51	22.91	24.30	20.81
	4	11.80	27.68	26.22	29.14	60.35	30.87
	3	11.56	22.10	20.61	27.30	43.12	24.39
	2	15.73	18.08	21.98	33.73	45.86	27.55
	1 [lowest]	14.28	20.62	26.95	21.14	12.57	19.18
Net farm income per hectare owned	5 [highest]	46.75	80.66	83.66	59.52	69.76	68.22
	4	44.55	75.22	83.98	59.42	122.44	77.09
	3	30.71	44.24	54.45	46.86	77.78	50.25
	2	30.51	31.54	46.03	45.51	58.48	43.02
	1 [lowest]	25.13	31.81	35.61	21.16	14.91	25.85
Net farm income per unit of labor	5 [highest]	14.81	33.71	34.97	23.25	24.84	26.40
	4	18.23	39.47	39.45	29.56	60.96	37.32
	3	15.10	27.77	28.68	27.67	44.10	28.21
	2	25.54	26.11	33.35	37.08	49.76	34.78
	1 [lowest]	19.57	32.70	45.17	25.72	43.39	37.50
Value of off-farm income per adult equivalent	5 [highest]	7.84	9.18	13.36	13.86	19.34	12.72
	4	8.75	11.86	19.91	23.91	41.46	21.23
	3	6.68	9.34	14.25	17.03	22.99	13.86
	2	8.84	10.67	15.23	16.60	27.97	16.26
	1 [lowest]	7.88	13.59	15.84	20.57	26.01	16.75
Value of assets/wealth per adult equivalent	5 [highest]	8.37	8.60	10.21	13.65	12.40	10.66
	4	11.14	12.02	15.55	27.10	29.91	19.18
	3	9.06	9.14	15.26	18.54	24.58	15.10
	2	19.16	14.25	19.02	19.46	30.43	20.85
	1 [lowest]	22.20	26.31	43.95	49.35	57.12	39.59
Household aggregate annual income	5 [highest]	16.1	29.3	30.3	33.2	42.9	30.4
	4	19.1	34.6	46.1	51.7	93.9	49.2
	3	15.5	26.4	32.2	37.9	55.2	33.0
	2	22.4	24.4	34.0	39.8	62.8	37.6
	1 [lowest]	19.0	31.9	42.1	49.2	46.1	37.6

Source: Tegemeo Institute Rural Household Surveys.

Note: Population density quintiles are defined by ranking all households in the surveys by village-level population density and dividing them into five equal groups.

Table 4. Smallholder commercialization indicators, by population density quintile ('000 KSh nominal)

	Pop den quintile	Survey year					average
		1997	2000	2004	2007	2010	
Value of crops sales per hectare '000KSh	5 [highest]	18.93	30.77	21.98	28.77	35.07	27.03
	4	18.81	34.31	33.89	44.76	83.84	43.63
	3	15.82	24.59	23.75	29.93	45.12	28.22
	2	11.20	9.98	18.62	26.37	32.93	20.13
	1 [lowest]	13.00	15.57	10.90	9.00	4.45	10.59
Household crop commercialization index (HCCI*)	5 [highest]	0.49	0.49	0.38	0.46	0.43	0.45
	4	0.44	0.51	0.51	0.58	0.59	0.52
	3	0.41	0.50	0.43	0.45	0.45	0.45
	2	0.40	0.39	0.43	0.44	0.40	0.42
	1 [lowest]	0.43	0.41	0.44	0.39	0.23	0.38
Value of livestock product sales per hectare '000KSh	5 [highest]	-	11.17	16.76	10.73	28.65	16.86
	4	-	13.70	18.24	23.22	34.37	22.44
	3	-	4.08	4.54	9.78	16.22	8.63
	2	-	4.35	7.30	9.03	12.78	8.46
	1 [lowest]	-	6.66	7.60	9.82	12.96	9.23
Household animal commercialization index (HACI**)	5 [highest]	-	0.28	0.22	0.25	0.32	0.27
	4	-	0.39	0.43	0.48	0.50	0.45
	3	-	0.22	0.21	0.29	0.39	0.28
	2	-	0.30	0.24	0.31	0.32	0.29
	1 [lowest]	-	0.33	0.33	0.42	0.43	0.38
Value of milk sold per hectare '000KSh	5 [highest]	-	10.89	16.41	9.94	25.02	15.60
	4	-	13.37	17.79	22.90	33.74	22.02
	3	-	3.74	4.13	9.68	15.32	8.19
	2	-	3.76	5.54	7.41	9.65	6.66
	1 [lowest]	-	4.72	7.02	8.59	12.71	8.21
Household dairy commercialization index (HDCI***)	5 [highest]	-	0.27	0.21	0.24	0.30	0.25
	4	-	0.38	0.43	0.47	0.49	0.44
	3	-	0.22	0.19	0.27	0.36	0.26
	2	-	0.26	0.23	0.30	0.30	0.27
	1 [lowest]	-	0.31	0.32	0.40	0.42	0.36

Source: Tegemeo Institute Rural Household Surveys.

Note: Population density quintiles are defined by ranking all households in the surveys by village-level population density and dividing them into five equal groups.

* HCCI= gross value of crop sales /gross value of crop production

**HACI=gross value of animal product sales / gross value of animal products production

**HDCI=gross milk sales / gross milk production

Table 5. First-stage CRE estimation results for population density in 2009

Dep. Variable: Log of village-level population density for each household (persons per km^2)	[I]		[II]		[III]	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Distance to motorable road ('00km)	-0.324**	0.138	-0.324**	0.138	-0.324**	0.138
Distance to water source ('00km)	-0.571**	0.268	-0.571**	0.268	-0.571**	0.268
Distance to health center ('00km)	-0.173***	0.059	-0.173***	0.059	-0.173***	0.059
Distance to electricity ('00km)	-0.247***	0.068	-0.247***	0.068	-0.247***	0.068
Distance to public telephone ('00km)	-0.383***	0.067	-0.383***	0.067	-0.383***	0.067
Population in 1969 (‘000 persons)	0.009***	0.001	0.010***	0.001	0.009***	0.001
Area in sq. km in 1969	-0.001	0.002	0.001***	0.000	0.002***	0.000
Population in 1979 (‘000 persons)	-0.202***	0.018	-0.086***	0.019	-0.041**	0.019
Area in sq. km in 1979	0.002***	0.000	0.002***	0.000	0.002***	0.000
Population in 1989 (‘000 persons)	0.831***	0.036	0.677***	0.040	0.622***	0.043
Area in sq. km in 1989	-0.005***	0.001	-0.006***	0.001	-0.006***	0.001
Rainfall (‘00 mm)	0.098***	0.014	0.098***	0.014	0.098***	0.014
Rainfall stress	-0.035***	0.003	-0.035***	0.003	-0.035***	0.003
Potential calories (trillion) from $10km^2$ pixel:						
_arable cultivated land	0.072***	0.004	--	--	--	--
_arable cultivated and grasslands land	--	--	0.027***	0.004	--	--
_arable cultivated, and grass and forest lands	--	--	--	--	-0.003	0.003
Agro ecological zone dummies included						
Constant	-1.816***	0.118	-2.697***	0.162	0.733***	0.274
Number of obs.	4584		4584		4584	
Number of households	1146		1146		1146	
R Squared	0.986		0.986		0.986	

Note: All the time varying variables are lagged one survey period; S.E= Bootstrapped standard errors; ***= p<0.01, **= p<0.05, *= p<0.1.

Table 6. CRE estimation results of farm size per adult equivalent

Dep. Variable: log of land holding (ha) per adult equivalent	[I]			[II]		
	Coef.	S.E.	P>z	Coef.	S.E.	P>z
Population density (‘00persons/ km^2)	-0.088	0.007	0.00	-0.090	0.007	0.00
Distance to motorable road ('00km)	-0.810	0.800	0.31	-0.814	0.800	0.31
Distance to water source ('00km)	0.300	0.200	0.14	0.300	0.200	0.13
Distance to healthcare centre ('00km)	-0.468	0.373	0.21	-0.471	0.374	0.21
Distance to electricity supply ('00km)	-0.193	0.353	0.59	-0.197	0.353	0.58
Ag. wage rate ('00Ksh.)- village median	-0.024	0.087	0.78	-0.024	0.086	0.78
Land rent ('000Ksh.)- village median	-0.005	0.002	0.01	-0.005	0.002	0.01
DAP price (Ksh.)- village median	-0.017	0.005	0.00	-0.017	0.005	0.00
Rainfall (‘00 mm)	0.030	0.019	0.12	0.029	0.019	0.13
Rainfall stress	-0.178	0.226	0.43	-0.171	0.226	0.45
Calories from arable cultivated land (trillions/$10km^2$)	0.010	0.005	0.04	--	--	--
Calories from arable cultivated and grasslands land (trillions/$10km^2$)	--	--	--	-0.014	0.003	0.00
Zone dummies (Central highland is the base)						
Eastern lowlands	-0.149	0.096	0.12	0.129	0.102	0.20
Western lowlands	0.056	0.090	0.54	0.301	0.099	0.00
Western transitional	-0.017	0.099	0.86	0.108	0.098	0.27
High potential maize	-0.006	0.065	0.93	0.147	0.072	0.04
Western highlands	0.087	0.083	0.30	0.188	0.085	0.03
Marginal rain shadow	-0.588	0.093	0.00	-0.399	0.095	0.00
Survey year dummies (year 2010 is the base)						
year 2000	-0.017	0.031	0.58	-0.017	0.031	0.59
year 2004	0.042	0.041	0.31	0.041	0.041	0.31
year 2007	0.229	0.067	0.00	0.229	0.067	0.00
Residuals from first stage regression	0.089	0.024	0.00	0.092	0.024	0.00
Constant	1.361	0.365	0.00	1.546	0.364	0.00
Observations	5730				5730	
Households	1146				1146	
R-square	0.735			0.735		

Note: All the time varying variables are lagged one survey period; S.E= Bootstrapped standard errors; p-score is the measure of statistical significance.

Table 7. CRE estimation results of hectares under crop per adult equivalent

Dep. Variable: log of crop hectarage per adult equivalent	[I]			[II]		
	Coef.	S.E.	P>z	Coef.	S.E.	P>z
Population density ('00 persons/ km^2)	-0.080	0.007	0.00	-0.084	0.007	0.00
Distance to motorable road ('00km)	-0.875	0.355	0.01	-0.879	0.355	0.01
Distance to water source ('00km)	0.165	0.166	0.32	0.166	0.166	0.32
Distance to healthcare centre ('00km)	-0.617	0.802	0.44	-0.621	0.802	0.44
Distance to electricity supply ('00km)	0.200	0.348	0.57	0.194	0.347	0.58
Ag. wage rate ('00Ksh.)- village median	-0.155	0.089	0.08	-0.154	0.089	0.08
Land rent ('000Ksh.)- village median	-0.002	0.002	0.24	-0.002	0.002	0.24
DAP price (Ksh.)- village median	-0.009	0.005	0.07	-0.009	0.005	0.07
Rainfall ('00 mm)	0.038	0.020	0.06	0.037	0.020	0.07
Rainfall stress	-0.280	0.235	0.23	-0.271	0.235	0.25
Calories from arable cultivated land (trillions/$10km^2$)	0.004	0.005	0.50	--	--	--
Calories from arable cultivated and grasslands land (trillions/$10km^2$)	--	--	--	-0.013	0.003	0.00
Zone dummies (Central highland is the base)						
Eastern lowlands	0.096	0.101	0.34	0.314	0.102	0.00
Western lowlands	0.037	0.092	0.68	0.244	0.100	0.02
Western transitional	-0.070	0.100	0.48	0.019	0.098	0.85
High potential maize	-0.186	0.067	0.01	-0.054	0.073	0.46
Western highlands	0.013	0.085	0.88	0.097	0.086	0.26
Marginal rain shadow	-0.802	0.082	0.00	-0.646	0.086	0.00
Survey year dummies (year 2010 is the base)						
year 2000	0.076	0.034	0.02	0.076	0.034	0.02
year 2004	0.112	0.042	0.01	0.112	0.042	0.01
year 2007	0.210	0.063	0.00	0.209	0.063	0.00
Residuals from first stage regression	0.050	0.025	0.04	0.055	0.025	0.03
Constant	-0.191	0.371	0.61	0.071	0.375	0.85
Observations	5730				5730	
Households	1146				1146	
R-square	0.659			0.659		

Note: All the time varying variables are lagged one survey period; S.E= Bootstrapped standard errors; p-score is the measure of statistical significance.

Table 8. CRE estimation results for intensity of cash input use per hectare

Dep. Variable: log of cost of purchased inputs (KSh) per ha	[I]			[II]		
	Coef.	S.E.	P>z	Coef.	S.E.	P>z
Population density (‘00 persons/ km^2)	0.213	0.021	0.00	0.054	0.025	0.03
Population density square	-0.017	0.002	0.00	-0.004	0.002	0.10
Distance to motorable road ('00km)	-0.886	0.322	0.01	-0.864	0.318	0.01
Distance to water source ('00km)	-0.302	0.171	0.08	-0.277	0.170	0.10
Distance to healthcare centre ('00km)	-0.501	0.726	0.49	-0.504	0.727	0.49
Distance to electricity supply ('00km)	-0.381	0.296	0.20	-0.373	0.294	0.21
Ag. wage rate ('00Ksh.)- village median	0.080	0.095	0.40	0.084	0.095	0.37
Land rent ('000Ksh.)- village median	0.004	0.002	0.02	0.004	0.002	0.01
DAP price (Ksh.)- village median	0.007	0.006	0.28	0.006	0.006	0.32
Rainfall (‘00 mm)	-0.031	0.022	0.15	-0.033	0.022	0.13
Rainfall stress	-0.668	0.261	0.01	-0.671	0.260	0.01
Calories from arable cultivated land (trillions/10km^2)	-0.054	0.005	0.00	--	--	--
Calories from arable cultivated and grasslands land (trillions/10km^2)	--	--	--	-0.062	0.004	0.00
Zone dummies (Central highland is the base)						
Eastern lowlands	-0.635	0.108	0.00	-0.028	0.116	0.81
Western lowlands	-1.412	0.111	0.00	-0.716	0.118	0.00
Western transitional	0.028	0.117	0.81	0.173	0.114	0.13
High potential maize	-0.074	0.079	0.35	0.332	0.081	0.00
Western highlands	-0.498	0.096	0.00	-0.357	0.095	0.00
Marginal rain shadow	-0.015	0.089	0.86	0.382	0.090	0.00
Survey year dummies (year 2010 is the base)						
year 2000	0.245	0.038	0.00	0.262	0.038	0.00
year 2004	0.137	0.048	0.00	0.163	0.047	0.00
year 2007	0.405	0.076	0.00	0.434	0.076	0.00
Residuals from first stage regression	0.085	0.037	0.02	0.076	0.037	0.04
Square of residuals	-0.070	0.019	0.00	-0.049	0.019	0.01
Constant	4.956	0.419	0.00	7.261	0.422	0.00
Observations	5730				5730	
Households	1146				1146	
R-square	0.630				0.626	

Note: All the time varying variables are lagged one survey period; S.E= Bootstrapped standard errors; p-score is the measure of statistical significance.

Table 9. CRE estimation results for value of net farm income per hectare owned

Dep. Variable: log of net farm income (KSh) per hectare owned	[I]			[II]		
	Coef.	S.E.	P>z	Coef.	S.E.	P>z
Population density (‘00 persons/ km^2)	0.321	0.047	0.00	0.258	0.055	0.00
Population density square	-0.030	0.005	0.00	-0.024	0.005	0.00
Distance to motorable road ('00km)	-0.016	0.009	0.07	-0.016	0.009	0.08
Distance to water source ('00km)	0.001	0.004	0.88	0.001	0.004	0.88
Distance to healthcare centre ('00km)	-0.017	0.015	0.26	-0.017	0.015	0.26
Distance to electricity supply ('00km)	0.005	0.008	0.57	0.005	0.008	0.55
Ag. wage rate ('00Ksh.)- village median	-0.604	0.214	0.01	-0.596	0.215	0.01
Land rent ('000Ksh.)- village median	-0.018	0.004	0.00	-0.018	0.004	0.00
DAP price (Ksh.)- village median	-0.007	0.013	0.60	-0.007	0.013	0.59
Rainfall (‘00 mm)	0.066	0.022	0.00	0.063	0.022	0.00
Rainfall stress	-0.018	0.006	0.00	-0.018	0.006	0.00
Calories from arable cultivated land (trillions/10km^2)	0.013	0.012	0.28	--	--	--
Calories from arable cultivated and grasslands land (trillions/10km^2)	--	--	--	-0.024	0.008	0.00
Zone dummies (Central highland is the base)						
Eastern lowlands	-0.425	0.224	0.06	0.084	0.245	0.73
Western lowlands	-2.466	0.195	0.00	-2.046	0.236	0.00
Western transitional	-2.248	0.238	0.00	-2.011	0.246	0.00
High potential maize	-1.725	0.177	0.00	-1.487	0.190	0.00
Western highlands	-1.642	0.212	0.00	-1.538	0.218	0.00
Marginal rain shadow	-1.127	0.253	0.00	-0.831	0.257	0.00
Survey year dummies (year 2010 is the base)						
year 2000	0.015	0.085	0.86	0.014	0.085	0.87
year 2004	-0.077	0.109	0.48	-0.074	0.109	0.50
year 2007	0.267	0.163	0.10	0.268	0.163	0.10
Residuals from first stage regression	0.150	0.090	0.09	0.135	0.090	0.13
Square of residuals	-0.031	0.114	0.79	-0.045	0.113	0.69
Constant	7.100	0.986	0.00	7.360	1.019	0.00
Observations	5730				5730	
Households	1146				1146	
R-square	0.274				0.272	

Note: All the time varying variables are lagged one survey period; S.E= Bootstrapped standard errors; p-score is the measure of statistical significance.

Table 10. CRE estimation results of value of net farm income per adult equivalent

Dep. Variable: log of net farm income (KSh) per adult equivalent	[I]			[II]		
	Coef.	S.E.	P>z	Coef.	S.E.	P>z
Population density (‘00 persons/ km^2)	0.208	0.046	0.00	0.103	0.055	0.05
Population density square	-0.028	0.005	0.00	-0.018	0.005	0.00
Distance to motorable road ('00km)	-0.021	0.009	0.02	-0.021	0.009	0.02
Distance to water source ('00km)	0.001	0.004	0.96	0.001	0.004	0.95
Distance to healthcare centre ('00km)	-0.014	0.014	0.32	-0.014	0.014	0.32
Distance to electricity supply ('00km)	0.004	0.007	0.58	0.004	0.007	0.54
Ag. wage rate ('00Ksh.)- village median	-0.651	0.204	0.00	-0.634	0.204	0.00
Land rent ('000Ksh.)- village median	-0.021	0.004	0.00	-0.021	0.004	0.00
DAP price (Ksh.)- village median	-0.024	0.012	0.05	-0.024	0.012	0.04
Rainfall (‘00 mm)	0.069	0.020	0.00	0.066	0.020	0.00
Rainfall stress	-0.025	0.005	0.00	-0.025	0.005	0.00
Calories from arable cultivated land (trillions/10km^2)	0.022	0.012	0.07	--	--	--
Calories from arable cultivated and grasslands land (trillions/10km^2)	--	--	--	-0.041	0.008	0.00
Zone dummies (Central highland is the base)						
Eastern lowlands	-0.684	0.215	0.00	0.181	0.241	0.45
Western lowlands	-2.391	0.179	0.00	-1.675	0.221	0.00
Western transitional	-2.216	0.226	0.00	-1.820	0.235	0.00
High potential maize	-1.794	0.166	0.00	-1.390	0.180	0.00
Western highlands	-1.516	0.198	0.00	-1.338	0.203	0.00
Marginal rain shadow	-1.692	0.226	0.00	-1.190	0.232	0.00
Survey year dummies (year 2010 is the base)						
year 2000	0.068	0.077	0.38	0.068	0.077	0.38
year 2004	0.012	0.102	0.91	0.016	0.102	0.87
year 2007	0.522	0.146	0.00	0.524	0.146	0.00
Residuals from first stage regression	0.237	0.081	0.00	0.212	0.082	0.01
Square of residuals	0.110	0.111	0.32	0.090	0.110	0.42
Constant	9.015	0.928	0.00	9.505	0.949	0.00
Observations	5730				5730	
Households	1146				1146	
R-square	0.400				0.399	

Note: All the time varying variables are lagged one survey period; S.E= Bootstrapped standard errors; p-score is the measure of statistical significance.

Table 11. CRE estimation results for household assets value per adult equivalent

Dep. Variable: log of the household assets value (KSh)	[I]			[II]		
	Coef.	S.E.	P>z	Coef.	S.E.	P>z
Population density (‘00 persons/ km^2)	-0.071	0.016	0.00	-0.069	0.016	0.00
Distance to motorable road ('00km)	-0.011	0.014	0.44	-0.011	0.014	0.44
Distance to water source ('00km)	-0.002	0.004	0.55	-0.002	0.004	0.55
Distance to healthcare centre ('00km)	-0.006	0.007	0.37	-0.006	0.007	0.37
Distance to electricity supply ('00km)	-0.002	0.008	0.77	-0.002	0.008	0.77
Ag. wage rate ('00Ksh.)- village median	-0.419	0.198	0.04	-0.419	0.198	0.03
Land rent ('000Ksh.)- village median	-0.010	0.004	0.01	-0.010	0.004	0.01
DAP price (Ksh.)- village median	-0.022	0.014	0.11	-0.022	0.014	0.11
Rainfall (‘00 mm)	0.047	0.017	0.01	0.047	0.017	0.01
Rainfall stress	-0.124	0.048	0.01	-0.123	0.048	0.01
Calories from arable cultivated land (trillions/10km^2)	0.050	0.013	0.00	--	--	--
Calories from arable cultivated and grasslands land (trillions/10km^2)	--	--	--	-0.025	0.007	0.00
Zone dummies (Central highland is the base)						
Eastern lowlands	0.691	0.089	0.00	0.691	0.089	0.00
Western lowlands	0.919	0.110	0.00	0.918	0.110	0.00
Western transitional	0.865	0.184	0.00	0.864	0.184	0.00
High potential maize	-1.324	0.206	0.00	-0.593	0.223	0.01
Western highlands	-1.499	0.182	0.00	-0.934	0.209	0.00
Marginal rain shadow	-1.901	0.179	0.00	-1.526	0.165	0.00
Survey year dummies (year 2010 is the base)						
year 2000	-0.880	0.165	0.00	-0.425	0.181	0.02
year 2004	0.138	0.092	0.13	0.137	0.091	0.13
year 2007	0.000	0.323	1.00	0.008	0.323	0.98
Residuals from first stage regression	14.399	0.876	0.00	14.222	0.889	0.00
Constant	1.361	0.365	0.00	1.546	0.364	0.00
Observations	5730				5730	
Households	1146				1146	
R-square	0.639				0.640	

Note: All the time varying variables are lagged one survey period; S.E= Bootstrapped standard errors; p-score is the measure of statistical significance.

Table 12. CRE estimation results for household income per adult equivalent

Dep. Variable: log of household income (KSh) per adult equivalent	[I]			[II]		
	Coef.	S.E.	P>z	Coef.	S.E.	P>z
Population density (‘00 persons/ km^2)	0.161	0.042	0.00	0.139	0.051	0.01
Population density square	-0.018	0.004	0.00	-0.015	0.005	0.00
Distance to motorable road ('00km)	-0.019	0.014	0.18	-0.019	0.014	0.19
Distance to water source ('00km)	0.001	0.004	0.89	0.001	0.004	0.90
Distance to healthcare centre ('00km)	-0.002	0.008	0.77	-0.002	0.008	0.79
Distance to electricity supply ('00km)	0.008	0.010	0.38	0.009	0.010	0.37
Ag. wage rate ('00Ksh.)- village median	-0.458	0.181	0.01	-0.453	0.181	0.01
Land rent ('000Ksh.)- village median	-0.006	0.003	0.04	-0.006	0.003	0.05
DAP price (Ksh.)- village median	-0.023	0.011	0.03	-0.023	0.011	0.04
Rainfall (‘00 mm)	0.001	0.000	0.00	0.001	0.000	0.00
Rainfall stress	-1.013	0.463	0.03	-1.036	0.463	0.03
Calories from arable cultivated land (trillions/$10km^2$)	0.040	0.010	0.00	--	--	--
Calories from arable cultivated and grasslands land (trillions/$10km^2$)	--	--	--	-0.009	0.008	0.28
Zone dummies (Central highland is the base)						
Eastern lowlands	-0.775	0.218	0.00	-0.315	0.244	0.20
Western lowlands	-1.871	0.178	0.00	-1.572	0.215	0.00
Western transitional	-1.487	0.191	0.00	-1.232	0.194	0.00
High potential maize	-0.918	0.138	0.00	-0.760	0.151	0.00
Western highlands	-1.108	0.163	0.00	-1.028	0.163	0.00
Marginal rain shadow	-0.910	0.247	0.00	-0.667	0.250	0.01
Survey year dummies (year 2010 is the base)						
year 2000	0.178	0.071	0.01	0.171	0.071	0.02
year 2004	0.485	0.084	0.00	0.480	0.084	0.00
year 2007	0.830	0.132	0.00	0.832	0.132	0.00
Residuals from first stage regression	-0.045	0.070	0.52	-0.061	0.070	0.39
Square of residuals	-0.033	0.098	0.74	-0.058	0.098	0.55
Constant	10.458	0.758	0.00	9.973	0.767	0.00
Observations	5730				5730	
Households	1146				1146	
R-square	0.412				0.412	

Note: All the time varying variables are lagged one survey period; S.E= Bootstrapped standard errors; p-score is the measure of statistical significance.