Revisiting the Farm Size-Productivity Relationship Based on a Relatively Wide Range of Farm Sizes: Evidence from Kenya

Milu Muyanga and T. S. Jayne

This paper revisits the inverse farm size-productivity relationship in Kenya. The study makes two contributions. First, the relationship is examined over a much wider range of farm sizes than most studies, which is particularly relevant in Africa given the recent rise of medium- and large-scale farms. Second, we test the inverse relationship hypothesis using three different measures of productivity including profits per hectare and total factor productivity, which are arguably more meaningful than standard measures of productivity such as yield or gross output per hectare. We find a U-shaped relationship between farm size and all three measures of farm productivity. The inverse relationship hypothesis holds on farms between zero and 3 hectares. The relationship between farm size and productivity is relatively flat between 3 and 5 hectares. A strong positive relationship between farm size and productivity emerges within the 5 to 70 hectare range of farm sizes. Across virtually all measures of productivity, farms between 20 and 70 hectares are found to be substantially more productive than farms under 5 hectares. When the analysis is confined to fields cultivated to maize (Kenya's main food crop) the productivity advantage of relatively large farms stems at least partially from differences in technical choice related to mechanization, which substantially reduces labor input per hectare, and from input use intensity.

Key words: Africa, agriculture, farm size, inverse relationship, Kenya, medium-scale farms, productivity.

JEL codes: O13, Q12, D24, Q15, R14.

Ever since the critical acclaim given to the Asian green revolution starting in the 1980s, it has been widely accepted that a smallholder-led growth strategy would also be the pathway for achieving economic transformation and mass poverty reduction in Africa. Over 80% of farms in India, Bangladesh, Indonesia, China, Japan, and Viet Nam were smaller than two hectares at the beginning of the Green Revolution (Johnston and Kilby 1975; Hayami and Ruttan 1985; Anriquez and Bonomi 2007). Because small-scale farms also constitute the vast majority of farms in Africa, agricultural economists have for decades generally accepted that a smallholder-led strategy also holds the best prospects for agricultural development in Africa (e.g., Mellor 1995; Lipton 2006; Hazell et al. 2007).

However, recent studies have questioned the viability of a smallholder-led growth strategy in Africa (e.g., Collier and Dercon 2014; Dercon and Gollin 2014). Many African governments have actively promoted the development of medium- and large-scale farms in recent years as a means to transform their countries' agricultural sectors into more capitalized and commercialized engines of economic dynamism (African Development Bank 2017).

These initiatives seem incongruous, at least on the face of it, with the literature on the

Milu Muyanga is an assistant professor, and T.S. Jayne is a University Foundation Professor, both at Michigan State University. Correspondence to be sent to: muyangam@msu.edu.

Amer. J. Agr. Econ. 101(4): 1140–1163; doi: 10.1093/ajae/aaz003
Published online March 25, 2019

inverse farm size-productivity relationship (IR), which has generally found that small farms are more productive than larger farms. Most of these studies use data in which the vast majority of observations are between zero and five hectares. However, parts of Sub-Saharan Africa are witnessing rapid changes in farm size distributions. “Medium-scale” farm landholdings of five to 100 hectares now account for a substantial and growing share of African farmland in many countries (Jayne et al. 2016).[1] Major concerns have arisen as to whether labor-intensive small-scale farms are being marginalized in the process (Anseeuw et al. 2016). Tests of the relationship between farm size and productivity in Africa covering a wide range of farm scales take on even greater policy importance in light of these dynamic changes in the distribution of farm sizes.[2]

This study revisits the IR hypothesis in Kenya. Several recent studies have examined the IR hypothesis in Africa based on data considered statistically representative of smallholder farms between zero and 10 hectares or so (Carletto, Savastano, and Zezza 2013; Larson et al. 2013). Their findings that the IR is upheld are important, but such studies cannot address productivity differences between small-scale and medium- or large-scale farms because these farm sizes exceed the range of farm sizes reliably covered in these studies.

Our study makes two contributions to the IR debate. First, we explore the IR hypothesis over a much wider range of farm sizes by including a representative sample of farms in Kenya’s high-potential zones between zero and 70 hectares. Our analysis replicates the conventional IR result found in the literature using a restricted sample of small-scale farms, but we find that the farm size-productivity relationship changes dramatically when utilizing the full sample that includes medium-scale farms. For these reasons, we conclude that empirical studies of the relationship between farm size and productivity should be explicit about the range of farm sizes over which their results apply to avoid potential overgeneralization in policy discussions.

The study’s second contribution is that it examines the relationship between farm size and productivity for three alternative measures of productivity, both to examine robustness of findings and to utilize measures of productivity that are arguably more meaningful than the measures used in most IR studies—yield or gross output per hectare—which take no account of production costs. In addition to gross output per hectare, we also examine the farm size-productivity relationship based on the net value of crop output per unit of area planted (a measure of profits that account for labor and other input costs) and total factor productivity (TFP). We also examine the robustness of our findings to alternative valuations of family labor, to potential differences in crop composition across farm size categories, and to potential measurement error in respondent-reported plot sizes.

The study focuses on areas of relatively favorable agro-ecological potential, which account for 70% of agricultural output in Kenya. Our conclusions are confined to such areas and cannot be considered nationally representative, but it is in these high-potential areas where medium-scale farms are growing most rapidly and where the competition for farmland is most acute (Jayne, Chamberlin, and Headey 2014).

Overview of Prior IR Literature and Extensions

Chayanov (1926) was the first to observe that small farms tended to obtain higher crop yields and output values per unit of land than large farms. Literature from Asia and Latin America has generally reinforced this finding of an inverse relationship (IR) between farm size and productivity (Schultz 1964; Hayami and Otsuka 1993; Binswanger, Deininger, and Feder 1995; Vollrath 2007; Kagin, Taylor, and Yúnez-Naude 2015).

Numerous explanations have been advanced to support the IR hypothesis. Some of these explanations result from the omission of input costs, particularly labor, as is often the case in IR studies. Other explanations coalesce around imperfect factors markets, the costs of supervising hired labor relative to own family labor, systematic measurement errors, and omitted variable issues.

[1] Consistent with the taxonomy of many African governments, this paper defines “small-scale farms” as those between 0–5 hectares of farmland. Medium-scale farms in this paper are defined as farms between 5–100 hectares of land, while large farms are those over 100 hectares. These definitions may not correspond exactly to those used by all national governments in the region.

[2] As we argue later, farm productivity is only one of several important criteria upon which land allocation and agricultural policy decisions could be based.

Imperfect factor markets may occur when factor prices facing small and large farms systematically diverge. In developing countries, agricultural wage rates in particular tend to be higher on large farms than for small farms (Carter 1984; Kagin, Taylor, and Yúnez-Naude 2015). Because family labor is intensively used on small-scale farms in developing countries, a low imputed opportunity cost of family labor contributes to findings of relatively high net value of output per hectare compared to large farms (Mazumdar 1965; Taslim 1989; Binswanger-Mkhize, Bourguignon, and van den Brink 2009; Hazell et al. 2010). Still, "large farms" in these studies are generally not more than five to ten hectares in size. Owner-operators have an advantage in supervision (Yotopoulos and Lau 1973). As residual claimants to farm profits, owner-operators tend to exert more effort than hired large farm managers (Frisvold 1994). Owner-operators also have better knowledge of local soil and climatic conditions, which are often accumulated over generations and tends to give small farms an edge over non-family operated large farms (Rosenzweig and Wolpin 1985).

The failure to fully capture land quality was also found to contribute to the inverse farm size-productivity relationship (Bhalla and Roy 1988; Benjamin 1995; Lamb 2003; Assunção and Braido 2007). It is sometimes argued that small farms are more productive because owner-operators tend to farm their highest quality land and sell or rent out less fertile land (Larson et al. 2013). The IR relationship often still holds even when land quality and unobserved effects are controlled for using panel data estimation techniques (Carter 1984; Benjamin 1995; Assunção and Braido 2007; Barrett, Bellemare, and Hou 2010; Foster and Rosenzweig 2017).

Systematic measurement errors associated with respondent-reported plot sizes have also been found to be important. Recent and novel studies by Carletto, Savastano, and Zezza (2013) and Dillon et al. (2016) using Living Standards Monitoring Surveys (LSMS) data and GPS plot measurements in Uganda and Nigeria, respectively, found that the correction of measurement errors in plot size further reinforces the IR hypothesis, at least within the range of farm sizes examined in these studies. Larger farms were more likely to under-report the size of their landholdings, while small farms tend to overstate them, further reinforcing the conclusions in support of the IR. However, as mentioned earlier, IR studies based on LSMS data generally contain no more than a few observations on farms over 10 hectares.[3]

While most studies on the farm size-productivity relationship have found that the IR hypothesis holds, some studies have found evidence to the contrary. Advantages associated with intensive use of modern technology (e.g., mechanization, precision farming) available to large farms may in some cases be reversing the historical inverse farm size-productivity relationship (Zaibet and Dunn 1998; Foster and Rosenzweig 2011, 2017). High transactions costs associated with smallholders' participation in input and output markets may provide advantages to relatively large farms (Dorward 1999). Food security concerns may lead small farms to specialize in less profitable staple crops (Fafchamps 1992; Omamo 1998; Dorward 1999; Lipton 2006). And small farms might obtain higher yields than larger farms due to more intensive use of labor, but obtain relatively low net output value after deducting the cost of labor. Studies examining the relationship between farm size and profits have generally either reversed the IR or weakened it (Carter 1984; Carter and Wiebe 1990; Rosenzweig and Binswanger 1993; Lamb 2003). Others have found a U-shaped relationship between farm size and productivity. For example, in Zambia, maize yield was found to decline with operated farm size up to about three hectares, which constitutes 86% of the study sample, and rise thereafter (Kimhi 2006). Foster and Rosenzweig (2017), using data from India, find that a U-shaped relationship may arise due to two factors: (*a*) fixed transaction costs in labor markets, which can result in an under-provision of hired farm labor, with the depressing effect on productivity becoming more severe as farms increase in size and need more labor (this finding would reinforce the IR hypothesis at low farm sizes); and (*b*) economies of scale in

[3] For example, the 2010/11 Tanzania LSMS contains 11 farms cultivating between 20–50 hectares, and only one farm between 50–100 hectares. In the Uganda LSMS, there are 12 farms between 20–50 hectares and none over 50 hectares. The Malawi 2010/11 LSMS contains one farm observation between 10-20 hectares, 1 farm between 20–50 hectares, and zero farms over 50 hectares. These surveys obviously do not contain a sufficient sample size to draw any meaningful conclusions about farms over 20 hectares. This conclusion is also acknowledged in the World Bank's recent 2018 Myths and Facts book relying on the use of LSMS data (Christiaensen and Demery 2018).

the utilization of machinery, which might result in higher levels of productivity beyond the minimum farm size at which mechanization becomes attractive. This study utilizes data to examine both of these hypotheses based on the case of high-productivity areas of Kenya.

Description of the Data and Variables

Data Sources

The study uses data from two main sources. The first is the nationwide Egerton University/Tegemeo Institute Rural Household Survey, a panel dataset tracking roughly 1,300 small-scale farm households in five survey waves over the 13-year period from 1997 to 2010. The sampling frame for the panel was prepared in consultation with the Central Bureau of Statistics (CBS; now Kenya National Bureau of Statistics) in 1997. The sample is representative of all the broad range of agro-ecological zones (AEZs) and agricultural production systems in Kenya. However, for reasons explained below, this study uses data only from the 2010 survey from the high-potential agro-ecological regions of the Rift Valley and Western Provinces. This zone accounts for 70% of Kenya's agricultural output. Hence, while the areas covered in this survey are not nationally representative, they do represent the country's most important agricultural zones. The surveys collect information on household socio-demographic characteristics, landholding size, farming practices, and plot-level output and input levels, and input costs including local labor wages.

Population-based data sets such as the Tegemeo surveys are not ideal for testing the farm size-productivity relationship because medium and large farms constitute a low proportion of the population (and hence a low probability of being included in the sample) even though they may account for a sizeable proportion of national farmland (Christiaensen and Demery 2018). Jayne et al. (2016) indicate that farms larger than 5 hectares account for over one-quarter of Kenya's farmland, and even this may be grossly underestimated.[4] To augment the sample of farms cultivating over 10 hectares, a supplemental survey of 200 medium-scale farm households was carried out in 2012 that randomly selected villages from the same high-potential regions that were surveyed in 2010. Lists of all farmers owning over 5 hectares in these villages were developed, and farms were randomly selected from these lists. The two surveys in 2010 and 2012 were conducted by the same research institute using the same questionnaire design, in the same villages, at the same time of year, and inputs and production in the two surveys were valued exactly in the same manner. This approach at least partially controls for unobserved time-constant geographic and community variables that might otherwise confound our study's findings. Unobserved time-varying differences are more problematic and our approach for addressing this is described below. We ended up with a sample of 479 households in which 282 (59%) were from the Tegemeo 2010 survey and 197 (41%) were from the 2012 survey. Of the total sample, 253 households were small-scale holdings (cultivating zero to five hectares), while 226 were medium-scale holdings (ranging from five to a maximum of 113 hectares cultivated). About 87% of the smallholder subsample came from the 2010 Tegemeo data, while 73% of the medium-scale sub-sample came from 2012 survey.

We examined whether the resulting pooled sample of small- and medium-scale farms was in the same proportion as in the overall population from the high-potential zone. Based on the most recently available Population Based Survey undertaken by Tegemeo Institute and the Kenyan National Bureau of Statistics in 2016, we find that 66%, 19%, and 15% of the farms cultivated between 0–5 hectares, 5–20 hectares, and over 20 hectares, respectively. The distribution in our pooled sample was 53%, 31% and 16%, respectively. We corrected for over/under sampling using inverse proportional weights computed by dividing the population shares by corresponding sample shares for each of the three farm types.

To control for other spatial factors, the medium- and small-scale farmers' survey instruments both recorded geographic positioning system (GPS) coordinates of each household. This made it possible to extract geographic information system (GIS) data to control for localized differences in soil type, elevation, slope, and length of growing period.[5] The GIS data was extracted at the 10×10 km pixel

[4] See Namwaya (2004).

[5] We thank Jordan Chamberlin of CIMMYT for extracting the GIS data used in this study.

level. The length of the growing period (LGP) combines information on temperature and available moisture to determine the locations' average duration of adequate crop growth (Fischer et al. 2000). Elevation and slope variables were extracted from Shuttle Radar Topography Mission (SRTM) data.[6]

Data on growing season rainfall came from the Climate Prediction Center and are part of the USAID/Famine Early Warning System (FEWS) project. This data interpolates rainfall estimates based on data from meteorological stations as well as satellite data such as cloud cover and cloud top temperatures. The FEWS rainfall estimates were then matched to surveyed households using their GPS coordinates. The rainfall variability variable is defined as the percentage of 20-day periods during the main growing season with less than 40 mm of rainfall.

Dependent and Explanatory Variables

This study utilizes three measures of agricultural productivity, namely (*a*) the gross value of output per hectare planted (the "standard IR measure"); (*b*) the net value of output per hectare planted, and (*c*) total factor productivity, computed following Zhang and Carter (1997) and Assunção and Braido (2007). While each measure may have its advantages, net output per hectare accounts for the costs of inputs—presumably an important consideration in productivity analysis—and TFP measures the returns on all measured inputs in addition to land. Agricultural production data come from respondents' reports on each field and sub-field on the farm.

The net value of crop production per hectare is computed as the value of total crop production minus variable and fixed input costs, divided by area planted. Variable costs include the cost of buying seed, fertilizer, land preparation (manual, draft power and tractor hire services, etc.) and labor (family and hired) costs. Family labor costs is the product of the total days worked by adult members of the household and the village median agricultural wage per day. Fixed costs constitute the estimated annual use value of all family agricultural assets and their depreciation. Each asset's annual use value is the asset's current value divided by its useful life. Where asset rental rates were available from the survey data, we used these as the best available indication of seasonal cost. Both variable and fixed costs are consistently defined across the two surveys. Output and production costs for the 2012 survey were computed using the 2010 village median prices to ensure comparability in valuation.

The major explanatory variable of interest in this study is area planted by the farm, as reported by farmer respondents. Other explanatory variables include household demographic variables (age and sex of the household head, and household size), distance to the nearest motorable road, amount of rainfall received in the cropping year and rainfall variability, community variables (length of growing period, and elevation). In addition to models with these exogenous variables, we also ran production function models that include input use variables (fertilizer, family and hired labor days, hybrid seed use, and own tractor use.

Estimation Strategy

Our analysis is based on the neo-classical production function approach following earlier studies (Feder 1985; Assunção and Braido 2007; Barrett, Bellemare, and Hou 2010; Ali and Deininger 2015). In this framework, farm output or productivity (Q) depends on labor, land, and capital inputs involved in production,

$$(1) \qquad Q = (L, A, K)$$

where (A) is farm size, (L) is labor, and (K) is capital inputs. The full household model to be estimated is specified as

$$(2) \qquad Q_i = \alpha + \beta A_i + \mathbf{X}\delta + \mathbf{W}\tau + \mathbf{Z}\pi + \varepsilon_i$$

where the dependent variable, Q_i, is the measure of agricultural productivity or profitability of household i. The explanatory variables are planted area (A_i); a vector of other labor and capital input variables such family and hired labor, fertilizer use, hybrid seed choice, and use of mechanization ($\mathbf{X}$); a vector of exogenous household socio-demographic characteristics ($\mathbf{W}$); and exogenous community variables ($\mathbf{Z}$) that are hypothesized to influence productivity, such as the length of growing period, elevation and slope of the farm, rainfall in the growing season, and market

[6] Available at: http://srtm.csi.cgiar.org/.

access conditions. The greek letters represent the parameters to be estimated, while ε_i is the random error term.

We first estimate models containing A_i and only exogenous variables in the **W** and **Z** vectors. We then estimate production function models that include input variables in the **X** vector. The inclusion of growing season rainfall and rainfall variability controls at least partially for time-varying differences. We also include survey year dummies in all models to examine whether any systematic difference in the dependent variables persists after including household and community controls.

Robustness Tests

Deriving unbiased estimates of the relationship between farm productivity and farm size requires addressing a number of challenges.

Measures of productivity. To examine the robustness of our results to the choice of productivity measure, we estimated all models separately for the three alternative measures of productivity described earlier.

Functional form. Our baseline farm size/productivity models are estimated in levels. Estimation in levels provides the most straightforward test of the relationship between farm size and farm productivity.[7] Alternative log-log models were estimated and because they produce highly consistent results, they are not reported here but are available upon request. Especially if the relationship between farm size and productivity is non-linear—as we find in this study—the partial derivative of a model estimated in levels will provide this relationship evaluated at any given level of farm size. We attend to functional form issues by testing for quadratic farm-size terms using F-test criteria to guide specification.

Comparability of 2010 and 2012 surveys. Pooling two data sets from different years (2010 and 2012) warrants thorough tests of comparability to ensure that unobserved differences between the two surveys and two years are controlled for. We assess the appropriateness of combining the two data sets in two different ways. First, we established the sub-sample in which there is overlap or common support between the two datasets. Following Khandker, Koolwal, and Samad (2010), we used propensity-score matching to match observation in the 2012 survey with observations in the 2010 survey. The propensity score used in matching were generated from a logit regression of a survey-year indicator variable ($1 = 2012$) on landholding sizes, demographic characteristics, distances to infrastructure, and spatial characteristics of the household location. This allowed us to match households in the 2012 survey with similar households in the 2010 survey based on these variables but not the outcomes of interest. About 94.6% of the pooled sample observations (96.1% in 2010 and 92.4% in the 2012 surveys) were found to lie within the common support area, with farm sizes ranging from 0.1 to 83 hectares. Observations outside the common support area were dropped. OLS models were estimated on this sub-sample and compared with results with those from the full sample.

As an alternate robustness check, we tested for systematic differences in productivity in the two survey years by introducing a survey dummy ($1 = 2012$ survey) in all regression models. A statistically significant sign on the survey dummy would indicate the presence of unobserved time-varying heterogeneity that could bias our results.

Alternative valuation of family labor. For models using the net value output per hectare as the dependent variable, we valued family labor alternatively at the observed village median wage rate and at the derived family labor shadow price computed as the marginal product of family labor from production function estimations on the largest maize field, for which detailed family labor data is available. The largest maize field accounted for 53.6% of total cultivated area for farms of zero to 5 hectares, and 47.7% of area on farms over 20 hectares, so that the shadow price of family labor on this field can be considered a reasonable approximation of that for the farm. The mean family labor shadow wage was found to be \$3.36, with a statistically significant coefficient at $\alpha = 0.05$, considerably higher than the mean ruling local agricultural wage rate of \$1.41.

Endogeneity of area. The variable of interest, area planted, could be endogenous in

[7] Log-log estimation, while commonly applied in IR studies, produces coefficient estimates of the relationship between percentage changes in farm size and percentage changes in productivity. One could easily envision cases whereby increases in farm size are associated with increasing levels of productivity, but declining percentage changes in productivity (or the reverse) at the margin. For this reason, we present results from level-level models.

equation (2). To address this concern, we explored the two-stage control function (CF) approach suggested by Wooldridge (2010) using the amount of land owned by the father of the household head before subdivision, the number of years that the household has engaged in farming, and years in the current settlement as instrumental variables in the first stage. The null hypothesis of endogeneity of A_i in equation (4) was rejected and thus the CF approach was dropped.

Potential for correlation between farm size and types of crops grown. It is possible, for example, that large farms may devote a greater portion of their area to high-valued crops, and if so, we would want to control for such crop composition differences in tests between farm size and farm productivity. To account for these possibilities, we computed crop category shares of total area cultivated for each household for maize, wheat, other staples, industrial cash crops such as tea and coffee, fresh fruits and vegetables, pulses, and fodder crops. In addition, given that maize fields account for roughly half of all area cultivated in the sample and maize is grown by 96.5% of the sampled farms, we estimate separate models for maize fields. Typically, these are intercropped fields so we include the value of all crops grown on these fields.

Measurement error in respondent-reported vs. actual plot sizes. Finally, use of self-reported landholding sizes may result in systematic error that could lead to bias in productivity estimation. Carletto, Gourlay, and Winters (2015) compared self-reported landholding sizes to Global Positioning Systems (GPS) plot measurement using Living Standards Measurement Study (LSMS) data from the World Bank for four African countries and found that farmers tend to overestimate and underestimate smaller and large plots, respectively. As a robustness check, we adjust the respondent-reported field sizes by generating area correction factors from the Tanzania LSMS. We defined these correction factors as the median of the GPS field size measurement/self-reported field size measurement ratio for all fields within various field size categories (0–0.5ha, 0.5–1ha, 1–2ha, 2–5ha, 5–10ha, 20–20ha, over 20 ha). To correct for the self-reporting bias, we multiply the respondent-reported hectares planted in the current study by these bias correction factors. While we would expect some differences in the degree to which farmers in one sample over- or under-estimate their field sizes compared to GPS field measurements in another sample, we feel that utilizing comparable information from neighboring Tanzanian households is a reasonable way to test the sensitivity of our results to potential area measurement bias in the absence of direct GPS measurement in our survey.

Descriptive Results

Tables 1 and 2 present bivariate descriptions of the data, disaggregated by three categories of farm size: (*a*) 0.1 to five hectares; (*b*) five to 20 hectares; and (*c*) over 20 hectares. We highlight three salient relationships emerging from table 1.

First, the unconditional productivity measures are consistently highest among farms over 20 hectares, next highest among farms 5–20 hectares, and lowest among farms 0–5 hectares. Gross output per hectare and total factor productivity on farms over 20 hectares are over 50% higher than on farms 0–5 hectares. The mean net value of output per hectare on farms over 20 hectares is over double that of farms under 5 hectares. Looking at maize fields in particular, which account for roughly half of all cultivated area, the net output per hectare on farms over 20 hectares are 30% higher on average than farms 5–20 hectares, which are in turn over double that of farms under five hectares.

Second, input costs per hectare are lowest among the farms over 20 hectares and highest on farms under five hectares, regardless of whether family labor is valued at observed mean agricultural wage rates of $1.41 per day, as shown in table 1 or at shadow wages. Input costs per hectare are roughly 15% higher on farms cultivating under five hectares than on farms over 20 hectares. When family labor is valued at the higher shadow wage, the cost difference becomes even more pronounced. The small farms use family labor relatively intensively, so scenarios that raise the value of family labor create cost disadvantages for small farms. The bigger farms are more intensive users of hired labor, so factors that would raise the price of hired labor relative to family labor will create cost advantages for small farms. The major sources of cost advantage for the relatively large farms are lower labor costs per hectare and mechanization costs. As shown in table 1, most small farms in Kenya's high potential

Table 1. Descriptive Statistics for Variables Used in Analysis

	Smallholders	Medium-Scale Farms		Full Sample
	$ha \leq 5$	$5 < ha \leq 20$	$ha > 20$	
Sample size (n)	253	149	77	479
% of sample from 2010 survey	87.3	34.9	11.7	58.9
		——— mean values ———		
Dependent variables				
Gross value of crop output/ha planted ('000KSh)	57.31	64.21	90.05	63.12
Net value of crop output/ha planted ('000KSh)	20.06	26.83	56.12	26.59
Total factor productivity ('000KSh)	2.95	3.34	5.06	3.41
Area variables				
Household landholding (ha)	1.89	13.97	50.06	13.39
Area planted (ha)	1.51	7.38	24.34	7.01
Area to maize crop (ha)	0.81	2.75	11.62	3.15
Input variables				
Fertilizer use/ha planted (kgs)	160.43	198.60	306.46	195.78
Family labor days/ha planted	43.74	8.23	3.20	26.18
Hired labor days/ha planted	21.41	15.31	11.72	17.95
Use hybrid maize seed (proportion)	0.95	0.94	0.99	0.95
Land preparation (proportion using)				
manual	0.17	0.07	0.01	0.11
oxen	0.21	0.03	0.01	0.12
tractor	0.62	0.91	0.97	0.77
Community level variables				
Length of growing period (days)	264.26	266.91	272.98	266.49
Elevation: '000 meters above sea level	1,940	1,964	1,953	1,950
Distance to motorable road (km)	4.48	4.83	3.89	4.49
Rainfall ('000mm) in main growing season	448.9	458.4	484.4	457.6
Rainfall stress (fraction of 20-day periods in main growing season with <40mm rain)	4.57	4.39	4.55	4.51
Demographic variables				
Age of household head (years)	59.91	64.51	63.37	61.90
Male headed households (proportion)	0.67	0.85	0.79	0.75
Household size (persons)	6.13	6.17	7.72	6.40
Household head's education attainment				
informal	21.6%	11.7%	11.0%	16.8%
primary	45.6%	35.7%	29.3%	39.8%
secondary	29.3%	23.4%	28.0%	27.3%
post-secondary	3.5%	29.2%	31.7%	16.2%
Total costs per hectare (000 Kenya shillings)	39.85	38.01	34.59	38.40
mechanization + oxen /ha planted	10.22	6.40	6.22	8.35
fertilizer /ha planted	7.34	11.10	13.10	9.44
seed /ha planted	6.75	4.07	4.08	5.49
total labor /ha planted	11.38	10.57	7.58	10.52
family labor /ha planted	9.32	1.63	0.37	5.49
hired & salaried/ha planted	2.06	8.94	7.21	5.03
fixed costs/ha planted	4.16	5.87	3.61	4.60

Source: TAPRA 2010 surveys and Medium-scale Farms' 2012 survey. *Note*: Family and hired labor valued at observed village agricultural labor wage.

zones rent mechanization services, but their costs per hectare are somewhat higher than for larger farms for which there are economies of scale in use of mechanization. Only if family labor is valued at close to zero would small farms have lower costs of production per hectare than farms in the two larger size categories.

To summarize, the main sources of unconditional productivity advantages of farms over 20 hectares are substantially higher value of output per hectare, and secondarily, slightly lower input costs per hectare.

Third, there are important differences in socio-economic characteristics between the three farm size categories. The heads of

Table 2. Household Landholding and Land Use

	Smallholders	Medium-Scale Farms		Full Sample
	$ha \leq 5$	$5 < ha \leq 20$	$ha > 20$	
N	253	149	77	479
Mean landholding (ha)	1.89	13.97	50.06	13.39
Hectares planted	1.51	7.38	24.34	7.01
% of total land under crop	79.89	52.83	48.62	52.35
Land under fallow (ha)	0.05	0.27	1.71	0.39
Idle/grazing land (ha)	0.33	6.32	24.01	5.99
% of area planted under monocrop	32.3%	74.5%	89.0%	55.1%
Mean number of fields	4.10	3.58	3.35	3.82
Mean field size (ha)	0.49	2.13	9.18	2.40
Mean number of crops grown/field	8.51	5.07	3.75	6.68
% of field with zero harvests	4.45	3.47	1.45	3.74
% of area with zero harvest	10.35	5.86	1.68	7.78

Source: TAPRA 2010 surveys and Medium-scale Farms' 2012 survey.

households of the farms over 20 hectares have substantially higher levels of educational attainment. The majority (68%) of farms 0–5 hectares are headed by persons with no more than primary school education. Over half of the medium-scale farm owners have secondary or post-secondary educations. Over 80% of the medium-scale farms are controlled by male-headed households compared to 67% among small-scale farms. Moreover, households' farm sizes are inversely related to the number of years the family has stayed in the current settlement. This finding is supported by a recent multi-country study showing that many medium-scale farmers in Africa were not born in the village where their farms are currently located; rather, many of these farms were recently purchased or otherwise acquired from local authorities (Jayne et al. 2016). These socio-demographic differences could be important sources of farm productivity differences between small- and medium-scale farms, hence the need to control for them in regression analysis.

Table 2 presents information on operated farm size and land use by production scale category. Households' farm sizes, area under fallow, and idle/grazing land are positively related to the amount of land owned by households. The proportion of area under crops is inversely related to households' landholding sizes. Small-scale farmers cultivated 80% of their total landholdings while farms over 20 hectare cultivate only 49% of their land under crops. This would indicate that relatively large farms are using their land less intensively than small farms, a finding that generally conforms to the literature; however, we cannot firmly conclude this because our survey unfortunately did not collect data on land devoted to pasture and other productive uses.[8] Results in table 2 also show that small-scale farms have a relatively high proportion (54%) of their plots under intercrops, growing about six different crops in each field, on average. Conversely, medium-scale farms in the 5 to 20 hectares category use mono-cropping on over 60% of their planted area. Mono-cropping accounts for roughly 85% of the area planted by farms over 20 hectares.

Bivariate non-parametric regressions. An important observation in the relationship between farm size and farm productivity emerges when examining the bivariate relationship between farm size and productivity in continuous terms (figure 1). In figure 1, we restrict the sample to farms cultivating less than six hectares, which constitute the vast majority of farm sizes analyzed in most IR studies to date using Asian and African data, and run bivariate non-parametric regressions. A U-shape relationship is observed in which all three measures of crop productivity decline with area cultivated between zero to four hectares. This is the classic IR result found by most studies examining this general range of farm sizes (e.g., Carletto, Savastano, and Zezza 2013; Larson et al. 2013). In this data set, the relationship between farm size

[8] The lack of data on non-crop land utilization precludes us from considering selectivity corrections on farmers' decisions to utilize their land.

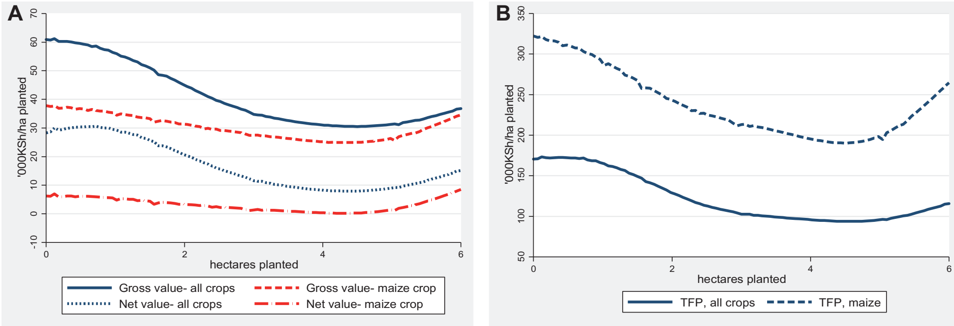

Figure 1. Non-parametric regression results of alternative measures of farm productivity by farm size using the sub-sample of farms planting under six hectares

Note: Non-parametric regression using Nadaya-Watson Approach, bandwidth = 0.8.

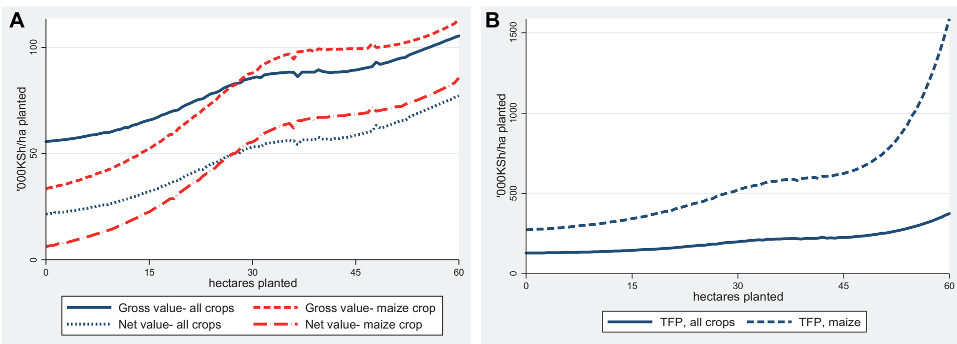

Figure 2. Non-parametric regression results of alternative measures of farm productivity by farm size utilizing the full sample of farms

Note: Non-parametric regression using Nadaya-Watson Approach, bandwidth = 0.8.

and productivity is relatively flat between four to five hectares, and then turns positive beyond five hectares.

Figure 2 presents the exact same non-parametric regressions but this time over the full sample of farms ranging from zero to 113 hectares. We now observe a positive relationship between farm size and each of the three productivity measures for all crops as well as for maize. This positive relationship is quite steep over the 10 to 35 hectare range for the gross and net value of crop output per hectare as shown in figure 2 panel (a), and much less so for total factor productivity as shown in figure 2 panel (b). However, it is possible that other household and/or community effects may be influencing these bivariate relationships and hence we now move to multivariate analysis.

Econometric Results

This section is organized as follows. We first present results using the "standard IR measure" as the dependent variable, that is, the gross value of output per hectare planted (table 3). As with all tables showing econometric results, we present these baseline results first for the sub-sample of farms under six hectares in the left panel of the table, and then for the full sample in the right-hand side panel. Table 4 replicates that of table 3 except that the measure of productivity is the net value of production per hectare planted. Table 5 shows analogous results using total factor productivity as the dependent variable. The purpose of first running models on the sub-sample of farms cultivating less than six hectares is to determine whether the standard

Table 3. OLS Estimation Results for Gross Value of Crop Production per Hectare Planted (‘000 Kenyan Shillings)

Dependent Variable:	Small-Scale Farms (6 ha and Below)				Full Sample (Small- and Medium-Scale Farms)			
	Model I(a)	Model I(b)	Model I(c)	Model I(d)	Model II(a)	Model II(b)	Model II(c)	Model II(d)
Ha planted	−35.11*	−41.96**	−16.68**	−14.19*	1.61***	0.97***	1.26***	1.54***
Sq. ha planted	5.79*	6.87**	2.89**	2.50*	−0.01***	−0.01***	−0.01***	−0.01***
Length of growing period		0.95**	0.54***	0.53***		0.95*	0.57***	0.55***
Elevation ‘00m		6.88	0.44	0.76		5.58	−0.32	−0.04
Distance to motorable road		−0.21	0.21	0.19		−0.72	0.03	0.01
Rainfall mm		0.28*	0.12	0.13		0.31**	0.16**	0.17**
Rainfall variability		−17.39	−6.07	−6.65		−20.96*	−9.09	−9.69*
Age of the household head		−0.33	−0.21	−0.18		−0.52**	−0.32*	−0.30*
Sex of the household head		9.20	1.26	0.31		5.29	−0.40	−0.95
Education attainment of hh head (base=informal)								
primary		12.20	6.10	6.59		7.31	4.29	4.74
secondary		9.55	8.77	8.92		−3.87	3.28	3.49
post-secondary		10.92	8.16	8.13		1.19	4.99	5.57
Household size		0.65	−0.34	−0.53		0.31	−0.32	−0.45
Fertilizer (‘000kg/ha planted)			16.34	19.63			2.15	4.84
Family labor (adult equiv. days/ha planted)			0.54***	0.54***			0.57***	0.57***
Hired labor labor (days/ha planted)			0.01	−0.32*			−0.03	−0.35**
Sq. hired labor labor ‘00				0.28*				0.30**
Hybrid seed use [1=yes]			7.82	7.59			11.37	11.16
Own tractor * 5<=ha<20							23.99***	22.79***
Own tractor * ha>=20							−10.87	−17.92
Household location dummies	Yes	Yes	Yes		Yes	Yes	Yes	Yes
Sample (1=2012; 0=2010)	7.53	−2.89	2.63	3.52	−1.02	−10.84	−1.85	−0.44
cons	106.47***	−309.69	−105.00	−108.47	77.62***	−289.01	−103.55	−105.94
Observations	343	343	343	343	479	479	479	479
R Square	0.16	0.24	0.61	0.61	0.10	0.18	0.57	0.58
Turning point for hectare planted	3.03	3.06	2.89	2.84	78.79	95.84	109.95	92.14

Note: Asterisks indicate the following: ***= 1% significance; **= 5% significance; *= 10% significance.

Table 4. OLS Estimation Results for Net Value of Crop Production per Hectare Planted ('000 Kenyan Shillings)

Dependent Variable:	Small-Scale Farms (6 ha and Below)				Full Sample (Small- and Medium-Scale Farms)			
	Model I(a)	Model I(b)	Model I(c)	Model I(d)	Model II(a)	Model II(b)	Model II(c)	Model II(d)
Ha planted	−21.26***	−23.49***	−22.02***	−19.25***	2.41***	2.06***	1.67***	2.01***
Sq. ha planted	3.77***	4.09***	3.79***	3.36***	−0.02***	−0.02***	−0.01***	−0.01***
Length of growing period		0.39**	0.45**	0.43**		0.45**	0.47***	0.45**
Elevation '00m		−0.34	0.53	0.01		−0.89	−0.08	0.25
Distance to motorable road		0.26	0.26	0.23		−0.02	0.05	0.03
Rainfall mm		0.14*	0.17**	0.18**		0.17**	0.20***	0.21***
Rainfall variability		−5.74	−8.58	−9.24		−9.44*	−11.37**	−12.09**
Age of the household head		−0.29	−0.30	−0.27		−0.41**	−0.41**	−0.38**
Sex of the household head		0.05	0.31	−0.75		−2.05	−2.23	−2.90
Education attainment of hh head (base=informal)								
primary		−2.41	−1.80	−1.25		−4.12	−3.48	−2.94
secondary		0.53	1.75	1.91		−7.30	−4.11	−3.86
post-secondary		−1.64	−0.67	−0.70		−4.82	−4.24	−3.55
Household size		−0.14	−0.17	−0.39		−0.19	−0.29	−0.45
Fertilizer (kg/ha planted)			−0.01	−0.01			−0.02	−0.02
Family labor (adult equiv. days/ha planted)			−0.04**	−0.04**			−0.01	−0.02
Hired labor labor (days/ha planted)			−0.12	−0.48**			−0.17**	−0.56***
Sq. hired labor labor '00				0.32**				0.36***
Hybrid seed use [1=yes]			7.08	6.82			9.26	9.01
Own tractor * 5<=ha<20							26.13***	24.68***
Own tractor * ha>=20							−15.24	−23.74
Household location dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample (1=2012; 0=2010)	−0.04	−2.90	−3.40	−2.40	−4.86	−8.16	−8.21	−6.52
cons	61.45***	−65.19	−98.50	−102.37	42.51***	−72.86	−97.60	−100.49
Observations	343	343	343	343	479	479	479	479
R Square	0.28	0.31	0.33	0.34	0.24	0.28	0.31	0.33
Turning point for hectare planted	2.82	2.86	2.90	2.86	64.45	66.45	101.71	88.24

Note: ***= 1% significance; **= 5% significance; *= 10% significance. Family labor valued at village agricultural wage rate.

Table 5. OLS Estimation Results for Crop Total Factor Productivity

	Small-Scale Farms (6 ha and Below)				Full Sample (Small- and Medium-Scale Farms)			
	Model I(a)	Model I(b)	Model I(c)	Model I(d)	Model II(a)	Model II(b)	Model II(c)	Model II(d)
Ha planted (ha)	−1.62**	−1.84***	−1.16***	−1.03**	0.11***	0.09***	0.09***	0.11***
Sq. ha planted ‘000	271.56**	304.11***	195.23***	175.35**	−0.71***	−0.51***	−0.37***	−0.53***
Length of growing period		0.02*	0.02	0.02		0.03*	0.02*	0.02
Elevation ‘00m		15.46	2.05	3.65		10.63	−1.73	−0.11
Distance to motorable road		0.03	0.04	0.04		0.01	0.03	0.03
Rainfall mm		0.01*	0.01	0.01		0.01**	0.01**	0.01**
Rainfall variability		−0.46	−0.24	−0.27		−0.69*	−0.43	−0.46
Age of the household head		−0.02	−0.02	−0.01		−0.03	−0.02**	−0.02**
Sex of the household head		0.60	0.41	0.36		0.39	0.23	0.20
Education attainment of hh head (base=informal)								
primary		0.22	0.08	0.10		0.04	−0.02	0.01
secondary		−0.12	−0.12	−0.11		−0.69	−0.41	−0.40
post-secondary		0.16	0.11	0.11		−0.24	−0.12	−0.08
Household size		−0.01	−0.04	−0.05		−0.02	−0.04	−0.05
Fertilizer (‘000kg/ha planted)			−0.03	0.13			−0.71	−0.56
Family labor (adult equiv. days/ha planted)			0.01***	0.01***			0.01***	0.01***
Hired labor labor (days/ha planted)			−0.02	−0.18***			−0.01	−0.02***
Sq. hired labor labor ‘00				0.15***				0.02***
Hybrid seed use [1=yes]			−0.03	−0.04			0.17	0.16
Own tractor * 5<=ha<20							1.35***	1.28***
Own tractor * ha>=20							−0.83	−1.24
Household location dummies	−0.94	−0.58	−0.24	−0.25	−0.63	−0.45	−0.10	−0.12
Sample (1=2012; 0=2010)	0.26	−0.08	0.05	0.10	−0.14	−0.48	−0.25	−0.17
cons	5.51***	−5.92	−1.46	−1.63	4.13***	−5.80	−1.58	−1.72
Observations	343	343	343	343	479	479	479	479
R Square	0.22	0.27	0.38	0.39	0.17	0.21	0.36	0.37
Turning point for hectare planted	2.99	3.03	2.97	2.95	76.63	89.41	121.77	101.61

Note: ***= 1% significance; **= 5% significance; *= 10% significance.

IR conclusion is upheld in this data set before moving to the full sample containing a broader range of farm sizes.[9]

Robustness checks on these main results are contained in tables A.1–A.4. Robustness tests are reported on variants of the models specified in table 4 for the net value of output per hectare, though robustness tests were performed on all models, which produced highly consistent results.

Because the analysis combines surveys from two different years, it is necessary to determine whether there is any systematic difference between the productivity measures in the two years. Table A.1a summarizes key results that replicate model results from tables 3–5 but based on the sub-sample of 453 farms within the common support area. Table A.1b provides descriptive statistics on the common support sub-sample for both 2010 and 2012. Table A.2 replicates the model results for the net value of output per hectare planted in table 4, but values labor at the shadow wage instead of village wage rates. Table A.3 replicates table 4 results but includes variables for the share of area under different crop categories, as well as model results where the sample is confined to maize fields. Finally, to examine the sensitivity of results to potential measurement error in field sizes, table A.4 replicates table 4 results after adjusting area measurements following the procedure described in the section on robustness checks.

Relationship between Farm Size and Productivity

We find robust evidence for a U-shaped relationship between farm size and productivity. In almost every model, the inclusion of a quadratic area term rejects the F-test. The turning point for hectares planted is reported at the bottom of each set of results.

Across all models presented in tables 3–5, when the sample is confined to farms under six hectares, the IR is upheld between zero to roughly three hectares. The relationship starts to turn positive at around 4 hectares and beyond. Models based on the full sample show a distinctly positive relationship between farm size and productivity driven by the strength of the positive relationship over the 10 to 60 hectare range, which overrides the IR observed when the sample is confined to farms at the low end of the farm size distribution. Results from the full sample models also indicate a slight tapering-off of productivity levels beyond 70 to 80 hectares, although relatively small sample sizes beyond this range warrant caution in interpretation. Across every robustness model presented in the appendix tables, this same pattern is upheld.

These differences in productivity across farm sizes are clearly not related to location or distance to markets. The sampling design required that medium-scale farms were drawn randomly from lists in the same villages as the sample containing the small-scale farms. As shown in table 1, the rainfall characteristics, length of growing period, elevation, and distance to tarmac road are very comparable between all three farm size groups. Setting all exogenous variables at the full sample means and predicting the values of the dependent variables based on differences in farm size only reveals that the mean of predicted net output per hectare planted among farms over 20 hectares remains 1.6 times greater than farms 5-20 hectares and 2.4 times greater than farms under five hectares. Similar differences are observed for the TFP and gross output per hectare productivity measures.

Comparability of 2010 and 2012 Surveys

As can be seen in tables 4–5, the survey dummy was generally not even close to being statistically significant after basic household covariates and time-varying growing season rainfall levels and rainfall stress are included in the models. This indicates that observed productivity differences cannot be attributed to systematic differences across the two survey years.

Similarly, our main finding of a U-shaped relationship between farm size and productivity continues to hold after restricting the sample to the common support area. Each of the nine model results shown in table A.1 produced statistically significant linear and squared area terms and similar turning points as in the main model results. However, it is noted that the common support area contains small sample sizes for some farm size categories in 2010. For example, the common support sample contains only nine observations

[9] In addition to these models with a cut-off of six hectares, we also ran alternative sub-sample models with 8 and 10 hectares and derived highly consistent results.

of farms cultivating over 20 hectares in 2010. There are also sizeable differences in the productivity measures between 2010 and 2012 (table A.1b). The net value of crop output per hectare for farms over 20 hectares in 2010 is slightly more than half that of the 2012 measure. So, even though the model results in the common support sub-sample are highly consistent with the main results, small sample issues are acknowledged, especially for the largest farm size category in 2010.

Robustness to Family Labor Valuation

Table A.2 replicates the main results in table 4 except that family labor is valued at its derived household-specific shadow wage, which was found to be higher in this sample than the ruling local agricultural wage rate. Because small farms tend to use family labor more intensively than larger farms, valuation of family labor at the higher shadow wage considerably narrows the range of farm sizes over which the IR is observed, that is, the turning point is now at 0.50 hectares or below, meaning that the IR holds only on farms between 0.1 and 0.50 hectares. These results indicate that as the value of family labor increases, the relatively productivity of labor-intensive small farms declines relative to larger farms. This may have important implications for anticipating future relationships between farm size and productivity in much of Sub-Saharan Africa where the opportunity cost of family labor may be rising in response to rising real per capita incomes and a declining share of the labor force engaged in farming (Barrett et al. 2017; Yeboah and Jayne 2018).

Robustness to Potential Differences in Crop Production Patterns

Table A.3 reports results for the net value of output per hectare and includes a set of variables specifying the share of households' area planted to the various categories of crops described earlier to control for potential differences in crop composition across the range of farm sizes. The main differences are that farms over 20 hectares devote a larger share of their total area to maize and wheat, while farms under five hectares devote a larger share to fruits, vegetables, and pulses. The crop categories with the highest value per hectare are industrial cash crops and fresh fruits and vegetables; the share of cropped area for these two crop categories are highest among farms 0–5 hectares. So, if anything, it is the small-scale farms that are devoting the highest share of their land to relatively high-valued crops. Nevertheless, the results in the right-hand side panel of table A.3 are highly consistent with the main results, with a strong positive relationship emerging between farm size and net output per hectare when the full sample is used. Therefore, the main results are not affected by any systematic differences in crop composition across the farm size distribution. As with the earlier reported results when the full sample is used, the squared area term, being negative and significant, indicates that productivity reaches a maximum somewhere between 66 and 130 hectares and then declines. So there is some evidence that productivity may taper off at the high end of this sample. However, consistent with the descriptive results, the mean predicted net output per hectare of farms over 20 hectares, other factors being held constant, remains 1.4 times higher than farms 5–20 hectares and 2.2 times higher than farms under 5 hectares.

In addition, table A.3 presents results on the relationship between farm size and the net value of output on fields containing maize—Kenya's primary staple crop—which is grown by 462 of the 479 farms in the sample. In these models, shown in Models 1a through 1d, the squared area term is not significant at the 0.10 level and is hence dropped. The linear area term is positive and highly significant in the first two columns of results (models 1a and 1b), but statistically insignificant in the second two models (1c and 1d) when inputs applied to the maize fields are included. These results imply that the bivariate productivity advantages in maize production of the farms over 20 hectares (which, as reported earlier, are three times higher per hectare than on farms under five hectares) appear to be driven by differences in technology, especially the use of mechanization, which contributes positively to net output per hectare. As shown in table 1, farms over 20 hectares use roughly twice as much fertilizer per hectare as the farms under five hectares, which translates into much greater output per unit land. Use of mechanization on farms between five and 20 hectares also appears to be a source of the bivariate productivity advantage of the relatively large farms. When these differences in technology and input use levels are controlled for, as in models 1(c) and 1(d) in table A.3, the relationship between farm size and the productivity of maize plots shrinks to insignificance, indicating that these variables are accounting for part of the productivity advantages of the relatively large farms.

Robustness to Potential Systematic Measurement Error in Field Size

Our last robustness check is to adjust respondent-reported field sizes to account for potential systematic measurement error in field sizes as described in the section on robustness checks. In computing the correction factors from the 2009 Tanzanian LSMS as described earlier, we had to compute the GPS/respondent field size ratios on the basis of relatively few observations in the 10–20ha and over 20ha categories. So, while the results in table A.4 are highly consistent with the main findings—upholding the IR between zero and roughly three hectares when the sample is restricted to farms under six hectares, and overturning it when utilizing the full sample of farms—we acknowledge that this robustness check relies on small samples for the area correction factors for some farm size categories.

Other Salient Observations from the Regression Results

While we observed large differences in educational attainment between the three farm size categories as mentioned earlier, we found no significant effects of differential educational attainment on any of the measures of farm productivity, suggesting that formal school learning may not be a suitable measure of farm management ability but it may be correlated with other factors that are associated with farm productivity.

Family and hired labor are found to be imperfect substitutes. Family labor contributed positively and statistically significantly to gross output per hectare and TFP and negatively to the net value of output per hectare, regardless of whether family labor is valued at the village wage rates or at household shadow prices. This indicates that while additional family labor contributes to crop output, it reduces profits when costed—even at observed ruling wage rates in this study that are below the conventional $1.90 per day poverty line.

Hired labor's contribution to all three productivity measures is found to be consistently non-linear. At low to moderate levels of hired labor, which tends to be piecework casual labor from another adult in the village, additional hired labor is associated with declining productivity per hectare, consistent with Foster and Rosenzweig (2017). As shown at the bottom of tables 3–5, the turning point for hired labor occurs between 60 and 78 days per hectare. More than 75% of the farms in the full sample hire less labor than 60 days per hectare, and they are mostly in the zero to five hectare category. Among those farms that hire more than 60 days of labor (where the partial derivative turns positive), they are mainly among farms in the five to 20 hectare category. Holding all else constant, the impact of moving from zero hired labor days (the value at the 25th percentile of hired labor) to 50 days (close to the 75th percentile) is associated with a 46% decline in the net value of output per hectare. A similar difference in the use of hired labor is associated with a 29% decline in mean TFP across the full sample. The imperfect functioning of hired labor markets appears to account for some of the observed lower productivity levels of small farms compared to more capital-intensive medium-scale farms.

Use of mechanization was found to contribute positively and significantly to all measures of productivity for farms between 5 and 20 hectares. Farms in this size category using mechanization obtained roughly 41% greater TFP and almost 100% higher net output per hectare than comparable farms not using mechanization. The mechanization dummy was not statistically significant for farms over 20 hectares, even though over 97% of these farms utilized mechanization primarily for land preparation. While the story is complex, it appears that labor market imperfections, small farms' more intensive use of family and hired labor, and medium-scale farms' more intensive use of fertilizers and mechanization are partially accounting for the U-shaped relationship between farm size and productivity observed in this study.

Conclusions and Policy Implications

This study examines the relationship between farm size and farm productivity over a broader range of family farm sizes than has typically been examined in Africa. To our knowledge, this is the first study to examine the farm size/productivity relationship that includes a matched sample of medium-scale farms in the 5 to 70 hectare range along with small-scale farms under five hectares. Most prior farm size-productivity studies rely on data for which there are very few observations over 10 hectares. This study is therefore motivated by the need to understand whether the well-established inverse farm size-productivity relationship holds when a broader range of farm sizes are considered. While our study covers farms between 0.1 and 113 hectares of cultivated land, our conclusions are confined to the range between zero and 70 hectares due to limited observations above this range. Results also pertain to the high potential areas of the country, which is unsurprisingly where most medium-scale farms are being established and where the competition for land is most acute.

The farm size-productivity issue has taken on special importance in recent years in light of the pace at which land has been acquired by medium-scale local investors (Jayne et al. 2016) as well as foreign investors (Deininger et al. 2011). Many African governments are promoting large and capital-intensive agricultural projects (African Development Bank 2017) that could further alter the distribution of farm size over time. Research findings on the relative productivity of small-scale, medium-scale, and larger farms can therefore inform and guide African governments' agricultural, land use, and land tenure policies in the region.

The overall picture emerging from this study is that there is a U-shaped relationship between farm size and the three measures of farm productivity. Regardless of whether productivity is defined in terms of gross output per hectare cultivated, gross output minus input costs including labor per hectare cultivated, or total factor productivity, the inverse relationship is indeed found to hold for farms cultivating between zero and roughly three hectares, consistent with most prior studies. Beyond five hectares the relationship between farm size and farm productivity is unambiguously positive, at least up to roughly 70 hectares.

For all three measures of productivity, farms over 20 hectares are substantially more productive on average than farms zero to five hectares. The unconditional mean TFP of farms over 20 hectares is 51% greater than farms 5 to 20 hectares, and 72% greater than farms under 5 hectares. The productivity advantages of farms in the 20 to 70 hectare category are even greater with the net output per hectare indicator. The farms 20–70 hectares obtain much higher levels of gross output per hectare and also obtain 15% lower costs of production per hectare.

The conditional productivity measures narrow only slightly after controlling for exogenous rainfall, community, and household characteristics. Therefore, the productivity advantages of the medium-scale farms are not due to advantageous agro-ecological or market access conditions compared to small-scale farms. When all variables including endogenous input use levels are included in the models and again set at full sample means, the productivity differences between the three groups, while still generally statistically significant, shrinks considerably, indicating that an important source of productivity difference between the three farm size categories are differences in technical choice and input use intensity.

These findings are robust to the three alternative measures of farm productivity, a restricted sub-sample of farms based on common support tests, alternative valuation of family labor, the inclusion of crop category variables to account for differences in the types of crops grown across the distribution of farm sizes, and measurement error in respondent-reported vs. actual plot sizes. When the analysis is restricted to fields cultivated to maize and when land preparation technology, fertilizer use, and other input use variables are included in the model, the productivity advantage of farms cultivating 20 to 70 hectares shrinks to insignificance, suggesting that their unconditional productivity advantage is largely driven by mechanization and input use intensity.

The findings of this study suggest that conventional findings of an inverse farm size-productivity relationship in Africa, based mainly on samples of small-scale farms, may need to be nuanced in light of changes taking place in family farm size distributions in many African countries. At a minimum, our study indicates that empirical studies of the relationship between farm size and productivity should be explicit about the range of farm sizes over which their results apply to avoid potential misinterpretation in policy discussions.

Much more evidence is needed from the heterogeneous conditions of Sub-Saharan Africa before any generalized conclusions could be reached about the relative productivity of small-scale vs. medium-scale farms. The region is highly diverse, and productivity advantages in high-potential areas may or may not apply to other agro-ecologies. However, the study's findings do raise several important policy questions.

First, we found that the competitiveness of small farms are strongly inversely related to the opportunity cost of family labor because they are relatively intensive users of family labor. If incomes and returns to labor in Africa continue to rise as they have over the past 15 years (Barrett et al. 2017), small farms may become relatively less competitive over time compared to more capital-intensive medium-scale and large farms. Mechanization is found to be a major source of productivity advantage for farms cultivating 5 to 20 hectares. Farms within the 5 to 20, and 20 to 70 hectare range apply roughly twice as much fertilizer per hectare as farms 0 to 5 hectares, which contributes to the much higher output values per hectare of these farms.

Second, findings that medium-scale farms are more productive than small-scale farms does not necessarily mean that policy actions are warranted to promote their development over small-scale farms or to support land transfers to medium-scale farms because productivity is but one criterion of importance to African governments and societies. Other relevant development criteria include which scale of farming provides the greatest contribution to gainful employment per unit land cultivated, and which scale of farming provides the greatest upstream and downstream multiplier effects to support broader economy-wide transformation. A stylized fact from Asia's experience is that the agricultural growth linkages with the rest of the economy are stronger with a uni-modal (smallholder-led) rather than bi-modal farm structure (Johnston and Kilby 1975; Mellor 1976), although these and other authors also found certain synergies resulting from a mix of large and commercialized small farms in the same vicinity. It is possible that the general equilibrium effects resulting from agricultural growth under alternative agrarian structures may be most decisive in influencing the pace of rural development and poverty reduction in these regions. For these reasons, we regard the findings of this study as an important but very incomplete contribution toward understanding the complex developmental effects of alternative farm-size distributions.

Acknowledgments

The authors gratefully acknowledge funding for this study from the USAID-funded Food Security Policy Innovation Lab led by Michigan State University, from the CGIAR Research Program on Policies, Institutions, and Markets (PIM), and from the DFID-funded Agricultural Policy Research in Africa (APRA) Project led by the Institute of Development Studies. The study has benefited from prior discussions with Michael Carter and Jordan Chamberlin, and from the detailed comments of four anonymous reviewers.

References

African Development Bank. 2017. Feed Africa: The Road to Agricultural Transformation in Africa. High Five Brief #2, African Development Bank.

Ali, D., and K. Deininger. 2015. Is There a Farm Size–Productivity Relationship in African Agriculture? Evidence from Rwanda. *Land Economics* 91 (2): 317–43.

Anriquez, G., and G. Bonomi. 2007. Long-Term Farming Trends: An Inquiry Using Agricultural Census Data. Agricultural Development Economics Division Working Paper 07-20. Food and Agricultural Organization of the United Nations, Rome.

Anseeuw, W., T. Jayne, R. Kachule, and J. Kotsopoulos. 2016. The Quiet Rise of Medium-Scale Farms in Malawi. *Land* 5 (19): 1–22.

Assunção, J., and L. Braido. 2007. Testing Household-Specific Explanations for the Inverse Productivity Relationship. *American Journal of Agricultural Economics* 89 (4): 980–90.

Barrett, C.B., M.F. Bellemare, and J.Y. Hou. 2010. Reconsidering Conventional Explanations of the Inverse Productivity-Size Relationship. *World Development* 38 (1): 88–97.

Barrett, C.B., L. Christiaensen, M. Sheahan, and A. Shimeles. 2017. On the Structural Transformation of Rural Africa. *Journal of African Economies* 26 (supp. 1): i11–35.

Benjamin, D. 1995. Can Unobserved Land Quality Explain the Inverse Productivity Relationship? *Journal of Development Economics* 46 (1): 51–84.

Bhalla, S.S., and P. Roy. 1988. Misspecification in Farm Productivity Analysis: The Role of Land Quality. *Oxford Economic Papers* 40 (1): 55–73.

Binswanger, H.P., K. Deininger, and G. Feder. 1995. Power Distortions Revolt and Reform in Agricultural Land Relations. In *Handbook of Development Economics*, Volume III, ed. J. Behrman and T.N. Srinivasan, 2659–772. Amsterdam, the Netherlands: Elsevier Science B.V.

Binswanger-Mkhize, H.P., C. Bourguignon, and R. van den Brink. 2009. *Agricultural Land Redistribution: Toward Greater Consensus*. Washington DC: World Bank.

Carletto, C., S. Gourlay, and P. Winters. 2015. From Guesstimates to GPStimates: Land Area Measurement and Implications for Agricultural Analysis. *Journal of African Economies* 24 (5): 593–628.

Carletto, C., S. Savastano, and A. Zezza. 2013. Fact or Artifact: The Impact of Measurement Errors on the Farm Size Productivity Relationship. *Journal of Development Studies* 103: 254–61.

Carter, M. 1984. Identification of the Inverse Relationship between Farm Size and Productivity: an Empirical Analysis of Peasant Agricultural Production. *Oxford Economic Papers* 36 (1): 131–45.

Carter, M., and K. Wiebe. 1990. Access to Capital and Its Impact on Agrarian Structure and Productivity in Kenya. *American Journal of Agricultural Economics* 72 (5): 1146–50.

Chayanov, A.V. [1926] 1991. *The Theory of Peasant Co-Operatives*. Columbus, OH: Ohio State University Press.

Christiaensen, L., and L. Demery, ed. 2018. *Agriculture in Africa: Telling Myths from Facts. Directions in Development*. Washington DC: World Bank.

Collier, P., and S. Dercon. 2014. African Agriculture in 50 Years: Smallholders in a Rapidly Changing World? *World Development* 63: 92–101.

Deininger, K., D. Byerlee, with J. Lindsay, A. Norton, H. Selod, and M. Stickler. 2011. *Rising Global Interest in Farmland: Can It Yield Sustainable and Equitable Benefits?* Washington DC: World Bank.

Dercon, S., and D. Gollin. 2014. Agriculture in African Development: A Review of Theories and Strategies. *Annual Review of Resource Economics* 6 (1): 471–92.

Dillon, A.S., Gourlay, K. McGee, and G. Oseni. 2016. Land Measurement Bias and Its Empirical Implications: Evidence from a Validation Exercise. Washington DC: World Bank Working Paper 7597.

Dorward, A. 1999. Farm Size and Productivity in Malawian Smallholder Agriculture. *Journal of Development Studies* 35 (5): 141–61.

Fafchamps, M. 1992. Cash Crop Production, Food Price Volatility, and Rural Market Integration in the Third World. *American Journal of Agricultural Economics* 74 (1): 90–9.

Feder, G. 1985. The Relation between Farm Size and Farm Productivity: The Role of Family Labor, Supervision and Credit Constraints. *Journal of Development Economics* 18 (2–3): 297–313.

Fischer, G., H. van Velthuizen, and F.O. Nachtergaele. 2000. *Global Agro-Ecological Zones Assessment: Methodology and Results*. Interim Report IR-00-064. International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria; and Food and Agriculture Organization of the United Nations (FAO), Rome, Italy.

Foster, A., and M.R. Rosenzweig. 2011. *Are Indian Farms Too Small? Mechanization, Agency Cost, and Farm Efficiency*. New Haven, CT: Economic Growth Center, Yale University.

———. 2017. Are There Too Many Farms in the World? Labor-Market Transaction Costs, Machine Capacities and Optimal Farm Size. NBER Working Paper 23909, October, 2017.

Frisvold, G.B. 1994. Does Supervision Matter? Some Hypothesis Tests Using Indian Farm-Level Data. *Journal of Development Economics* 43 (2): 217–38.

Hayami, Y., and K. Otsuka. 1993. *The Economics of Contract Choice: An Agrarian Perspective.* Oxford: Oxford University Press.

Hayami, Y., and V. Ruttan. 1985. *Agricultural Development*. 2nd ed. Baltimore, MD: Johns Hopkins University Press.

Hazell, P., A. Dorward, C. Poulton, and S. Wiggins. 2007. The Future of Small Farms for Poverty Reduction and Growth. Discussion Paper No. 42, International Food Policy Research Institute 2020, IFPRI, Washington, DC.

Hazell, P., C. Poulton, S. Wiggins, and A. Dorward. 2010. The Future of Small

Farms: Trajectories and Policy Priorities. *World Development* 38 (10): 1349–61.

Jayne, T.S., J. Chamberlin, and D. Headey. 2014. Land Pressures, the Evolution of Farming Systems, and Development Strategies in Africa: A Synthesis. *Food Policy* 48: 1–17.

Jayne, T.S., J.B. Chamberlin, N. Sitko, L.N. Traub, F. Yeboah, M. Muyanga, W. Answeeuw, et al. 2016. Africa's Changing Farm Size Distribution Patterns: The Rise of Medium-Scale Farms. *Agricultural Economics* 47 (S1): 197–214.

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

Kagin, J., E. Taylor, and A. Yúnez-Naude. 2015. Inverse Productivity or Inverse Efficiency? Evidence from Mexico. *Journal of Development Studies* 52 (3): 396–411.

Khandker, S.R., G.B. Koolwal, and H.A. Samad. 2010. *Handbook on Impact Evaluation. Quantitative Methods and Practices*. Washington DC: The International Bank for Reconstruction and Development/The World Bank.

Kimhi, A. 2006. Plot Size and Maize Productivity in Zambia: Is there an Inverse Relationship? *Agricultural Economics* 35 (1): 1–9.

Lamb, R.L. 2003. Inverse Productivity: Land Quality, Labor Markets, and Measurement Error. *Journal of Development Economics* 71 (1): 71–95.

Larson, D., K. Otsuka, T. Matsumoto, and T. Kilic. 2013. Should African Rural Development Strategies Depend on Smallholder Farms? An Exploration of the Inverse Productivity Hypothesis. *Agricultural Economics* 45: 1–13.

Lipton, M. 2006. Can Small Farmers Survive, Prosper, or Be the Key Channel to Cut Mass Poverty? *Electronic Journal of Agricultural and Development Economics* 3 (1): 58–85.

Mazumdar, D. 1965. Size of Farm and Productivity: A Problem of Indian Peasant Agriculture. *Economica* 32 (126): 161–73.

Mellor, J. 1976. *The New Economics of growth: A Strategy for India and the Developing World*. Ithaca: Cornell University Press.

Mellor, J.W. 1995. *Agriculture on the Road to Industrialization*. Baltimore, MD: Johns Hopkins University Press.

Namwaya, O. 2004. Who Owns Kenya? East Africa Standard, October 1, 2004. http://www.jaluo.com/wangwach/200709/Otsieno_Namwaya092807.html (Accessed Mar 6, 2018).

Omamo, S.W. 1998. Transport Costs and Smallholder Cropping Choices: An Application to Siaya District, Kenya. *American Journal of Agricultural Economics* 80 (1): 116–23.

Rosenzweig, M.R., and H.P. Binswanger. 1993. Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments. *Economic Journal* 103 (416): 56–78.

Rosenzweig, M.R., and K.I. Wolpin. 1985. Specific Experience, Household Structure, and Intergenerational Transfers: Farm Family Land and Labor Arrangements in Developing Countries. *Quarterly Journal of Economics* 100 (Supplement): 961–87.

Schultz, T.W. 1964. *Transforming Traditional Agriculture*. New Haven, CT: Yale University Press.

Taslim, M.A. 1989. Supervision Problems and the Size-Productivity Relation in Bangladesh Agriculture. *Oxford Bulletin of Economics and Statistics* 51 (1): 55–71.

Vollrath, D. 2007. Land Distribution and International Agricultural Productivity. *American Journal of Agricultural Economics* 89 (1): 202–16.

Wooldridge, J. 2010. *Econometric Analysis of Cross-Section and Panel Data*. 2nd Ed. Cambridge, MA: MIT Press.

Yeboah, K., and T.S. Jayne. 2018. Africa's Evolving Employment Structure. *Journal of Development Studies* 54 (5): 803–32.

Yotopoulos, P.A., and L.J. Lau. 1973. A Test for Relative Economic Efficiency: Some Further Results. *American Economic Review* 63 (1): 214–23.

Zaibet, L.T., and E.G. Dunn. 1998. Land Tenure, Farm Size, and Rural Market Participation in Developing Countries: The Case of the Tunisian Olive Sector. *Economic Development and Cultural Change* 46 (4): 831–48.

Zhang, B., and A.C. Carter. 1997. Reforms, the Weather, and Productivity Growth in China's Grain Sector. *American Journal of Agricultural Economics* 79 (4): 1266–77.

Appendix

Table A.1a. Summary of OLS Estimation Results for Productivity Indicators within the Common Support Area

	Model	Hectares Planted	Hectares Planted sq.	Exogenous Variables Included in Model?	Inputs & Management Practice Variables Included in Model	N	R Square	Turning Point for Hectares Planted
(a) Common support sub-sample								
Gross value of crop production/ha planted ‘000KSh	I(a)	1.76***	−0.01***			453	0.10	61.49
	I(b)	1.03*	−0.01*	Yes		453	0.17	60.08
	I(c)	2.51***	−0.03***	Yes	Yes	453	0.58	39.17
Net value of crop production/ha planted ‘000KSh	II(a)	2.93***	−0.03***			453	0.23	45.83
	II(b)	2.55***	−0.03***	Yes		453	0.27	43.29
	II(c)	2.85***	−0.03***	Yes	Yes	453	0.29	42.06
Total factor productivity [ha planted square term ‘000]	III(a)	0.11***	−0.85***			453	0.18	62.92
	III(b)	0.08***	−0.70***	Yes		453	0.22	59.78
	III(c)	0.14***	−1.57***	Yes	Yes	453	0.39	44.64

Note: ***= 1% significance; *= 10% significance. Family labor valued at village agricultural wage rates.

Table A.1b. Means of Key Variables for Sub-Sample of Households in the Common Support Area

	Farm Size Category						Full Sample	
	$ha \leq 5$		$5 < ha \leq 20$		$ha > 20$			
Survey year:	2010	2012	2010	2012	2010	2012	2010	2012
N:	210	32	52	97	9	53	271	182
Age of the household head (years)	59.46	63.81	62.88	65.39	62.56	63.13	60.22	64.46
Gender of the household head (1=male)	0.70	0.59	0.85	0.86	1.00	0.74	0.74	0.77
Household size	6.28	5.42	6.09	6.21	6.07	7.64	6.24	6.49
Hectares owned	1.78	2.64	12.34	14.84	30.67	31.80	4.76	17.63
Hectares planted	1.20	3.70	5.45	5.80	12.24	18.19	2.39	9.04
Length of growing period (days)	264.42	264.40	263.07	268.96	259.31	274.42	263.99	269.75
Elevation: ‘000 meters above sea level	1.93	1.99	2.01	1.94	2.12	1.95	1.95	1.95
Distance to motorable road (km)	4.04	4.18	6.18	4.11	6.53	3.54	4.53	3.95
Rainfall (mm) in main growing season	449.53	454.04	403.36	487.91	431.28	494.06	440.07	483.74
Rainfall stress (fraction of 20-day periods in main growing season with <40mm rain)	4.61	4.39	3.77	4.73	3.80	4.65	4.42	4.65
Gross value of crop output/ ha planted (‘000KSh)	53.77	52.46	58.58	64.87	59.31	88.85	55.01	69.67
Net value of crop output/ha planted (‘000KSh)	20.20	18.48	27.28	26.59	30.16	56.40	21.89	33.84
Total factor productivity (‘000KSh)	2.98	2.83	3.33	3.35	3.51	4.93	3.07	3.72

Table A.2. OLS Estimation Results for Net Value Crop Output per Hectare Planted (‘000KSh), Family Labor Valued at Shadow Wage

	Small-Scale Farms (6 ha and Below)				Full Sample (Small- and Medium-Scale Farms)			
	Model I(a)	Model I(b)	Model I(c)	Model I(d)	Model II(a)	Model II(b)	Model II(c)	Model II(d)
Ha planted (ha)	−13.01*	−11.79	−30.26***	−27.14***	2.47***	2.45***	1.28***	1.69***
Sq. ha planted	2.52*	2.24	5.11***	4.62***	−0.02***	−0.02***	−0.01***	−0.01***
Length of growing period		−0.03	0.31	0.29		0.07	0.32*	0.30
Elevation (m)		−0.04	0.01	0.01		−0.04	0.01	0.01
Distance to motorable road		0.64	0.35	0.32		0.49	0.17	0.14
Rainfall mm		0.03	0.17**	0.18**		0.06	0.18**	0.20***
Rainfall variability		1.63	−8.94	−9.67		−1.77	−10.86*	−11.74**
Age of the household head		−0.18	−0.28	−0.25		−0.25	−0.40**	−0.36**
Sex of the household head		−8.25	−1.85	−3.04		−9.25	−5.00	−5.85
Education attainment of hh head (base=informal)								
primary		−8.24	−3.67	−3.05		−1.05	−6.43	−5.73
secondary		1.41	2.93	3.12		−8.47	−5.12	−4.73
post-secondary		−5.81	−2.25	−2.29		−3.61	−7.95	−7.07
Household size		−1.29	−0.53	−0.78		−6.00	−0.71	−0.92
Fertilizer (kg/ha planted)			−0.03	−0.03			−0.04**	−0.03*
Family labor (adult equiv. days /ha planted)			−0.42***	−0.42***			−0.38***	−0.39***
Hired labor labor (days/ha planted)			−0.05	−0.45**			−0.15*	−0.62***
Sq. hired labor labor ‘hundreds of days				0.36**				0.45***
Hybrid seed use [1=yes]			10.12	9.83			9.23	8.98
Own tractor * 5<=ha<20							33.18***	31.76***
Own tractor * ha>=20							−12.30	−22.34
Household location dummies	YES	YES	YES	YES	YES	YES	YES	YES
Sample (1=2012; 0=2010)	−0.90	1.23	−1.31	−0.20	−3.25	−1.97	−6.73	−4.65
cons	33.04**	90.72	−73.87	−78.21	20.58***	67.30	−72.82	−76.65
Observations	343	343	343	343	479	479	479	479
R Square	0.12	0.16	0.49	0.50	0.13	0.17	0.44	0.46
Turning point for hectare planted	2.58	2.63	2.96	2.94	62.24	60.88	101.09	117.76

Note: ***= 1% significance; **= 5% significance; *= 10% significance.

Table A.3 OLS Estimation Results for Net Value of Maize and Crop Output per Hectare Planted, with Crop Composition Dummies Included

	Net Value of ***Maize Production***/ha Planted '000KSh				Net Value of Crop Production/ha Planted '000KSh			
	Model I(a)	Model I(b)	Model I(c)	Model I(d)	Model II(a)	Model II(b)	Model II(c)	Model II(d)
Ha planted (ha)	1.13***	1.04***	0.08	0.33	2.29***	2.19***	1.38***	1.98
Sq. ha planted					−0.02***	−0.02***	−0.01***	−0.01
Length of growing period		0.37***	0.25***	0.28***		0.29	0.27	0.26
Elevation '00m		1.57	2.08**	2.22**		−0.99	−0.91	0.26
Distance to motorable road		−0.38	−0.36	−0.40*		0.07	0.16	0.12
Rainfall mm		0.22***	0.16***	0.19***		0.20***	0.20***	0.25***
Rainfall variability		−12.30***	−7.64**	−9.47***		−10.94**	−10.66*	−14.03**
Age of the household head		−0.10	−0.10	−0.12		−0.41**	−0.43**	−0.37**
Sex of the household head		−2.77	−4.04	−3.99		−2.31	−2.05	−3.71
Education attainment of hh head (base=informal)								
primary		0.75	1.45	0.45		−6.73	−6.68	−5.40
secondary		0.77	0.46	−0.23		−9.17	−7.41	−5.37
post-secondary		16.64**	14.30**	13.88**		−8.07	−7.63	−6.81
Household size		0.68	0.68	0.66		−0.25	−0.27	−0.58
Fertilizer ('000kg/ha planted)			4.82	3.41			−29.72	−20.47
Family labor (adult equiv. days/ha planted) '000			−280.77***	−957.67***			−33.78	−195.26
Sq. family labor '000				7.28***				0.22
Hired labor (days/ha planted) '000			41.32**	−8.51			2.33	−597.56***
Sq. hired labor (per ha planted) '000				0.50				3.70***
Hybrid seed use [1=yes]			−9.47	−9.32			7.48	9.28
Own tractor * 5<=ha<20			13.74*	23.83**			28.70***	26.79***
Own tractor * ha>=20			28.62***	31.13***			−2.60	−20.26
Household location dummies								
Proportion of area under different crops included	36.32***	24.38***	18.30***	19.14***	−2.96	−6.27	−5.15	−4.94
Sample (1=2012; 0=2010) cons	−6.88***	−177.94***	−136.29***	−148.23***	36.78***	−36.04	−36.90	−57.56
Observations	471	471	471	471	479	479	479	479
R Square	0.26	3,204.00	0.37	0.35	0.27	0.30	0.32	0.35
Turning point for hectare planted					66.74	67.84	129.78	91.09

Note: ***= 1% significance; **= 5% significance; *= 10% significance. Share of crop area variables included for maize, wheat, other staples, fruits and vegetables, industrial cash crops, fodder crops, and pulses. Family labor valued at village agricultural wage rates.

Table A.4. OLS Estimation Results for Net Value Crop Output per Bias-Corrected Hectare Planted

Dependent Variable: Net Value of Crop Output/ha Planted (Bias Corrected) '000KSh	Small-Scale Farms				Full Sample (Small- and Medium-Scale Farms)			
	Model I(a)	Model I(b)	Model I(c)	Model I(d)	Model II(a)	Model II(b)	Model II(c)	Model II(d)
Ha planted	−28.38***	−30.43***	−28.98***	−25.70***	2.00***	1.59**	0.94**	1.33**
Sq. ha planted	4.99***	5.26***	4.96***	4.44***	−0.02***	−0.01**	−0.01**	−0.01**
Length of growing period		0.44*	0.51**	0.49**		0.51**	0.53**	0.51**
Elevation ('00m)		−0.48	0.52	0.94		−1.17	−0.06	0.34
Distance to motorable road		0.32	0.32	0.29		0.01	0.15	0.13
Rainfall mm		0.16*	0.20**	0.22**		0.21**	0.24***	0.26***
Rainfall variability		−7.09	−10.69	−11.47		−12.19*	−14.45**	−15.30**
Age of the household head		−0.36	−0.36	−0.33		−0.48**	−0.49**	−0.46**
Sex of the household head		−0.35	0.13	−1.12		−2.77	−2.18	−2.99
Education attainment of hh head (base=informal)								
primary		−4.84	−4.15	−3.51		−7.09	−6.93	−6.26
secondary		−0.71	0.80	0.99		−9.72	−6.65	−6.29
post-secondary		−3.23	−1.80	−1.84		−6.22	−5.27	−4.44
Household size		−0.29	−0.31	−0.56		−0.39	−0.51	−0.70
Fertilizer ('000kg/ha planted)			−0.02	−0.02			−32.46	−0.03
Family labor (adult equiv. days/ha planted)			−0.05**	−0.06***			−0.02	−0.02
Hired labor labor (days/ha planted)			−0.13	−0.55**			−0.21**	−0.67***
Sq. hired labor labor ('00)				0.37**				0.42***
Hybrid seed use [1=yes]			9.03	8.73			10.02	9.77
Own tractor * 5<=ha<20							38.15***	36.81***
Own tractor * ha>=20							−17.74	−27.27
Household location dummies	Yes	Yes	Yes	Yes	Yes	Yes	−0.17	Yes
Sample (1=2012; 0=2010)	1.04	−2.03	−2.09	−0.91	−4.40	−8.32	−7.25	−5.27
cons	75.04***	−61.69	−101.05	−105.61	51.41***	−74.46	−101.37	−105.00
Observations	343	343	343	343	479	479	479	479
R Square	0.29	0.32	0.34	0.35	0.22	0.27	0.31	0.32
Turning point for hectare planted	2.85	2.89	2.92	2.89	53.45	53.44	112.69	85.11

Note: ***= 1% significance; **= 5% significance; *= 10% significance. Bias corrected (BC) hectares planted=self-reported hectares/correction factor as computed from Tanzania LSMS 2010/11; Correction factor=GPS measurement/Self-reported measurement. Family labor valued at village agricultural wage rates.