WPS6190

Policy Research Working Paper 6190

Should African Rural Development Strategies Depend on Smallholder Farms?

An Exploration of the Inverse Productivity Hypothesis

Donald F. Larson
Keijiro Otsuka
Tomoya Matsumoto
Talip Kilic

The World Bank
Development Research Group
Agriculture and Rural Development Team
September 2012

Abstract

In Africa, most development strategies include efforts to improve the productivity of staple crops grown on smallholder farms. An underlying premise is that small farms are productive in the African context and that smallholders do not forgo economies of scale—a premise supported by the often observed phenomenon that staple cereal yields decline as the scale of production increases. This paper explores a research design conundrum that encourages researchers who study the relationship between productivity and scale to use surveys with a narrow geographic reach, when policy would be better served with studies based on wide and heterogeneous settings. Using a model of endogenous technology choice, the authors explore the relationship between maize yields and scale using alternative data. Since rich descriptions of the decision environments that farmers face are needed to identify the applied technologies that generate the data, improvements in the location specificity of the data should reduce the likelihood of identification errors and biased estimates. However, the analysis finds that the inverse productivity hypothesis holds up well across a broad platform of data, despite obvious shortcomings with some components. It also finds surprising consistency in the estimated scale elasticities.

This paper is a product of the Agriculture and Rural Development Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org.
The author may be contacted at DLarson@worldbank.org.

Should African rural development strategies depend on smallholder farms? An exploration of the inverse productivity hypothesis

Donald F. Larson[1,2], Keijiro Otsuka[3], Tomoya Matsumoto[3] and Talip Kilic[2]

Keywords: agriculture, inverse productivity, farm size, smallholders, Sub-Saharan Africa, technology choice

JEL classifications: O13, Q12, Q16, Q18

Sector Board: Agriculture and Rural Development

Acknowledgments

This research was supported in part by a grant from the Research Support Budget, entitled "Can Agricultural Growth Strategies rely on African Smallholders? The paper draws on data collected under the Living Standards Measurement Study - Integrated Surveys on Agriculture project, an ongoing initiative funded in part by the Bill and Melinda Gates Foundation. The authors would like to thank Amparo Palacios-Lopez for her help with the datasets for Malawi and Tanzania.

[1]Corresponding author: email: DLarson@worldbank.org; mail: World Bank, 1818 H Street, NW, Washington, DC 20433, USA.
[2] Development Research Group, World Bank.
[3] National Graduate Institute for Policy Studies (GRIPS)

Should African rural development strategies depend on smallholder farms? An exploration of the inverse productivity hypothesis

Donald F. Larson, Keijiro Otsuka, Tomoya Matsumoto, and Talip Kilic

Inventing and disseminating technologies that boost yields on smallholder farms is a central pillar of rural development strategies in Africa. Moreover, on-farm yields are seen as a measure of whether government policies have been successful and yields are used to evaluate new varietal research. To be sure, there are other components of rural development strategies, especially components related to health, education and nutritional safety nets. But agriculture remains the economic engine of rural Africa and the livelihoods of the rural poor depend on it. As a consequence, a dominant view among policy makers is that boosting yields is the best way to enhance rural incomes in the short-to-medium term and the goal of boosting yields on smallholder farms is advanced by the World Bank, the Food and Agriculture Organization of the United Nations, the CGIAR institutions, most African governments, and pan-African institutions. A key basis for the strategy is evidence that small farms are more productive than large farms in Africa when it comes to growing staple crops, a proposition known as the inverse-yield, or inverse-productivity relationship.[1]

Despite widespread institutional support, the debate over whether this emphasis on smallholder agriculture is wise remains active. Sometimes, the arguments take on normative tones. For example, Collier (2008) charges that the development community has stressed less-innovative smallholder agriculture over more-productive commercial agriculture because of an overly romantic view of peasant farming. Echoing Schultz (1964), Hazell et al. (2010) counter that promoting smallholder agriculture is a more equitable approach to rural development, as well as a more efficient one. Lipton (2006) argues that emphasizing smallholder development partly compensates for policies in rich and poor countries that are, on balance, urban-biased. Still, beyond issues of

[1] The following quotes are indicative of the institutional support for smallholder-led strategies: from the President of the World Bank "Eighty-six per cent of staples in poor areas come from local sources, so support for country-led efforts to bolster smallholder agriculture is critical." (Zoellick, 2011); from the Director General of FAO, "Sustainable intensification of smallholder crop production is one of FAO's strategic objectives." (Diouf, 2011); from the Director General of IFPRI, "G20 Ministers of Agriculture must focus on smallholder farmers to achieve food security and prevent food price volatility" (Fan, 2011); and from the pan-African Alliance for a Green Revolution in Africa, "AGRA works to achieve a food secure and prosperous Africa through the promotion of rapid, sustainable agricultural growth based on smallholder farmers." AGRA (2011).

motivation and equity, significant scope for positive debate remains because of uncertainty about the design features of the current body of empirical research, a topic which this paper explores.

On the face of it, emphasizing smallholder technologies as a means of boosting agricultural productivity seems a narrow strategy. However, proponents of smallholder-based rural development emphasize that the prevalence of small farms is not accidental, but an optimizing outcome in the face of imperfect land and labor markets. Moreover, the argument goes, decades of sustained effort are needed to build a more capable and better educated labor force and the strong institutions needed for secure land markets. Consequently, the only path to growth for this generation of rural poor lies with technologies that boost yields on smallholder farms.

A central component of this reasoning is that farmers are optimizing agents – albeit within the constraints of limited resources and weak land and labor markets – since this is needed to explain why farmers would willingly adopt yield-improving technologies, even as the sector as a whole seemingly fails to employ the larger-scale technologies associated with productive commercial farming on other continents.

While the appropriateness of a smallholder-led strategy for Africa is debated, there is less controversy about whether the strategy can be effective in general, largely because of Asia's Green Revolution experience. Like Africa, and unlike other regions, agriculture in Asia is smallholder based. Moreover, the scale of farming in many Asian countries has declined during recent decades. Moreover, this occurred while yields increased at a remarkable pace, especially for staple grains like wheat and rice.[2] What is more, when the Green Revolution began in Asia, poverty was as pervasive in rural areas as it is currently in many parts of Africa. As the new varieties of staple grains spread through Asia, associated productivity gains raised farm income and spilled over to spur growth in the non-farm sectors of rural communities; because of the large share of income spent on food by the poor, a resulting productivity-led decline in real food prices reduced urban poverty as well.[3]

---

[2] von Braun (2005) reports that the average farm size in both Africa and Asia is about 1.6 hectares, while farms average 27 hectares in Western Europe, 67 hectares in Latin America and the Caribbean, and 121 hectares in Canada and the United States. In Asia, farm sizes have been constant to declining in recent decades .For example, the average farm size in China was 0.6 hectares in 1980 and 0.4 hectares in 1999. In India, average farm size declines from 2.3 hectares in 1971 to 1.4 hectares in 1994-96. In Indonesia, the average farm size was 1.1 hectares in 1973 and 0.9 hectares in1993 (Nagayets 2005). At the same time, data from FAO (2012) show that average rice yields grew by 2.6 percent annually in China from 1961 to 2001 and at 2.1 percent and 2.7 percent annually in India and Indonesia. Annualized growth rates for wheat were 4.6 percent for China and 3.2 percent for India. In contrast, average maize yields in Sub-Saharan Africa grew at less than one percent per year from 1961 to 2010.

[3] The consequences for poverty were global. According the World Bank (2008, p.3), the primary reason why the global $1-a-day poverty rate fell from 28 percent in 1993 to 22 percent in 2002, was an 8 percent decline in rural poverty –

For these many reasons, the Asian experience factors heavily in the design of African rural development strategies (Otsuka and Larson, forthcoming). And there is much riding on the outcome; should the farm structure in Africa remain largely unchanged as it has in Asia, food output in Africa will stagnate as land frontiers close. If global food prices rise as a consequence, then the urban poor and the many rural poor who are not food self-sufficient will suffer. If productivity rises elsewhere and prices continue to decline, then the farm incomes of Africa's smallholders will fall, pushing them deeper into poverty.[4]

In this paper we revisit the inverse productivity phenomenon in the context of African maize, a key grain in Africa with an eye toward its implication for smallholder-led strategies in Sub-Saharan Africa.[5] To be relevant for policy, the inverse relationship between productivity and scale should hold broadly across large and diverse areas, since the strategy depends on national and regional efforts. However, credible tests of the hypothesis require fine and detailed data, which tend to arise out of more homogeneous settings. To explore this conundrum, we first developing a statistical model, based on Mundlak's (1988, 1993) insight that agricultural production data contain measured outcomes that arise from a mix of applied technologies, which are in turn endogenously determined by the differing circumstances framing on-farm decisions. In this, local conditions can be important and relevant information is therefore potentially lost as the unit of observation in the data moves from micro to macro. We then estimate the model using multiple datasets of varying granularity and geographic scope. This allows us to explore the consequence of trade-offs between desirable spatial diversity and desirable reductions in omitted information on tests of the inverse-yield hypothesis.

Related literature

An important component of the empirical work on the consequences of smallholder agriculture for productivity has to do with observations that small farms seemed to be more productive than larger farms in Asia and in Africa. Writing in 1946 about differences in rubber yields, P.T. Bauer notes that "Measured by long-period supply price, the smallholders are the more efficient class, though their methods necessarily differ from those of the estates (p. 391)." Using data from Uttar Pradesh, Mazumdar (1965) finds that the shadow price of family labor was lower for smallholders, leading to

---

mostly due to improved rural living conditions. Moreover, the largest decline was in East Asia, driven in part by Asia's Green Revolution (Ravallion, Chen and Sangraula, 2007).

[4] See Ivanic, Martin, and Zaman (2011), for a discussion about how food prices affect poverty.

5 See Smale, Byerlee and Jayne (forthcoming) for a discussion of maize and maize policies in Africa.

higher labor inputs and yields. Sen (1966) makes the same point, citing government studies from India. Carter (1984) does as well, based on observations from Haryana, India, even after adjusting for selection bias and for differences in village and soil characteristics. Also using data from India, Bhalla (1979) finds an inverse relationship, even though small farmers in his sample were less likely to adopt modern technologies. Bardhan (1973) notes that land and other market imperfections can work to generate inverse yield relationships; he also points out that differences can arise among crops because of crop-specific differences in labor and management practices.

Yotopoulos and Lau (1973) argue that Indian smallholders were more efficient because of owner-managers had an advantage in supervision and leadership. Taslim (1989) makes the same point using data from Bangladesh. Binswanger and Rosenzweig (1986) and Binswanger and McIntyre (1987) argue that family farms are the predominant organizational structure for agricultural sectors in places where containing transaction costs are more important than economies of scale. In one of the few conceptual treatments of the problem based on first principles, Feder (1985) shows that supervision costs and credit availability can generate systemic relationships between yields and farm size in a conceptual model. Lipton (2006) also notes that high transaction costs related to output markets that encourage self-sufficiency also favor small farms. In a similar way, Barrett (1996) argues that food security concerns induce smallholder farmers to supply added labor as a risk-mitigation strategy.[6]

Because the technology, crop and market characteristics driving farmer choices and influencing the organizational structure of agriculture vary, differences in empirical studies might be expected. In an early study that employed cross-sectional data, Deolalikar (1981) found differences in the relationships between farm size and productivity that were also associated with on-farm technology choices. Examining data from land-abundant Western Sudan, Kevane (1996) found that wealth rather than farm size mattered most for productivity and labor intensity. U-shaped relationships were reported by Carter and Wiebe (1990) for Kenya and Heltberg (1998) for Pakistan. Moreover, there is evidence that as the circumstances conditioning farmer decisions change, optimal scale outcomes shifts as well. For example, Foster and Rosenzweig (2010) show that increased mechanization in India may be shifting incentives toward larger scale farms.

[6] See Kimhi (2006) for a concise description of the set of arguments motivating applied studies of the inverse-yield phenomenon.

Another line of reasoning suggests that the inverse yield relationships are statistical artifacts. A compelling argument is that smallholder lands are naturally more productive since households are less likely to sell or rent-out their highest quality land. Consequently omitting measures of land quality can systematically biases empirical results (Bhalla and Roy, 1988; Benjamin, 1995; Lamb, 2003; Assuncao and Braido, 2007). A related argument is that inverse yield relationships are driven by measurement errors due to systematic tendencies in self-reported area and yields (Lamb, 2003). Empirical evidence is thin, but a study by Barrett et al. (2010) that includes soil measurements in Madagascar finds no evidence of systematic bias for measured yields. Similarly, a recent study that incorporates self-reported and GPS land measurements in Uganda concludes that measurement errors work against rather than in favor of the inverse yield hypothesis (Carletto et al., 2011).

While the large literature on the inverse-yield phenomenon encompasses diverse explanations, a unifying theme is that local features of markets and geography, and the transaction costs they engender, matter since they influence farmers' choices about what they produce and how they go about producing it. This creates a practical tension for applied work. On the one hand, a significant spatial variation in collected data is desirable in order to generate exploitable variation in the market conditions. At the same time, precise and comparable data is needed in order to generate robust and convincing empirical results. In general most research addresses the later need, and consequently most empirical results are based on data with limited geographic reach.[7] In this paper we explicitly recognize the trade-off between spatial diversity and measurement precision by using disparate datasets to estimate a common statistical model.

Conceptual and applied model

A starting point for the conceptual model is Mundlak's (1988) notion that applied technologies stem from endogenous choices and this should guide empirical applications. The idea is consistent with agricultural sectors comprised of multiple and disparate farming methods in developing countries and also the significant differences in applied technologies apparent across countries for the same crop. It is also conformable with the common sense notion that natural endowments, such as climate and soils, influence how farmers grow their crops.

In developing countries, the relative scarcity of land and labor is often key. For example, when land is abundant and labor scarce, shifting agriculture is the most cost-effective way to produce subsistence crops, conserving both labor and non-labor inputs. Conversely, fixed farming

[7] An early exception is Berry and Cline's (1979) cross-country study.

systems predominate in areas where populations are denser since population density is usually associated with land scarcity. It often means better market access as well, which creates the opportunity for farmers to produce beyond subsistence levels.[8] Layered on top of these fundamental structural drivers, the weak performances of land and labor markets have their own consequences for production choices.[9]

In the broadest sense, a technology is a mapping between planned inputs and expected outputs, and an applied technology reveals a strategic choice. This has conceptual and practical implications when the decision environment is heterogeneous – that is, when the endowments, prices and constraints farmers face vary. Since the set of optimization problems behind the data vary, outcomes measured in the data vary accordingly. For empirical work, this implies that outcomes from multiple applied technologies are implicit in the data. In turn, there must be an adequate accounting for the different underlying relationships, primal and dual, between inputs and outputs before inferences can be made, including inferences about scale effects.

To see this later point, let $\Phi(s) = max_x\, \pi(x, s) = pf(x, s) - wx$ characterize the full set of optimization problems faced by a given population of farmers, conditional on a set of available technologies, where the vector s fully describes exogenous (non-choice) state factors, including input and output prices, $w, p \in s$, that influence farmers' choices.[10] In the simple case, when farms and farmers are identical and face the same set of prices, supply and input demand schedules can be recovered in a straight-forward way at the solution values of x via the envelope theorem, and in turn, an empirical model can be based on a representative depiction of the farm problem. In an applied setting, the conditions farmers face are heterogeneous – that is the set of feasible technologies, relevant prices, endowments and other state variables that condition choice vary, which means that the state-variables should be sub-scripted. In this case, representative supply and input demand schedules are only available for classes of farmers, where within-class farmers face the same-valued state variables. With this in mind, the input demand schedules are given by $\frac{\partial \phi}{\partial w_i} = -x_i(s_i)$, and the supply function, given at solved values of x by $\frac{\partial \phi}{\partial p_i} = f(x, s) = y_i(s_i)$, where w_i and p_i are location-specific shadow prices for inputs and output.

[8] See Binswagner (1991), and Pingali, Bigot and Binswanger (1987) for good discussions of how land and labor scarcity affect technology choices.
[9] See Deininger and Feder (2001) for a review of land institutions and the performance of land markets. See Larson and Mundlak (1997) for a discussion about the slow pace of adjustment between agricultural labor and labor employed in other sectors.
[10] For a more detailed discussion, see Mundlak, Larson and Butzer (1999) and Mundlak, Butzer and Larson (2012).

For a particular household, the relevant set of variables in s_i needed to describe the decision problem can be large and include household constraints – for example, low levels of human capital or limited access to credit. Among all households, less constrained households might adapt more profitable technologies that are infeasible for constrained household. And in a competitive environment, economic incentives would encourage the spread of more profitable strategies through farm sales or the development of new intermediaries, resulting eventually in the emergence of a dominant technology and a dominant scale of farming. In the cases of Asia and Sub-Saharan Africa, there is a prevailing view that pervasive features of land, capital and labor markets work to limit the profitable scale of farming. Said differently, the set of s_i commonly faced by farmers in Sub-Saharan Africa and Asia is thought to explain agricultural sectors dominated by smallholder production. As such, the prevalence of smallholder farm structure is a conditionally optimal outcome and one consistent with the "inverse yield" phenomenon.

Turning to the applied model, it has become common practice in studies of yield and land-scale to assert a two-stage decision process in which a farmer first allocates land among activities and then chooses the rates of application for the remaining inputs.[11] This convenient device allows the production to be specified in terms of a yield, so that, written in logs, $y_i = \alpha_i \hat{y}_i(z) + \beta_i x_i^l$, where inputs other than land, x_i^l, has been scaled by land, that is $z_i^j = x_i^j / x_i^l$ for all j inputs where $j \neq l$. Because land is the only input that is not rescaled (and therefore not an element of z) we can dispense with the superscript.

For the applied model, we focus on the land-output relationship: $x_i = \gamma_i \hat{y}_i + \lambda_i z(s_i) + s_i + e_i$, where x_i is the natural log of the area planted to maize at location i, and where the e_i are random errors; the α_i and λ_i are technology-specific fixed parameters. The location-specific state variables are represented by fixed effects, s_i. Moreover, since the vector z contains choice outcomes (solutions) conditioned by s_i, their average impact on land use can be included in the fixed effects where $\acute{s}_i = \lambda_i z_i + s_i$ when the set of conditioning factors are complete. This is because farmers that face the same state variables face identical optimization problems and their solutions should be the same. When some state variables are not observed, different solutions can emerge due to differences in the conditioning state variables that the farmer recognizes but the researcher does not. When this is the case, there will be unexplained spreads in the average intensity of factor use: $z_i - \bar{z}$, where $\bar{z}$ is an average over i. This is an important point, since without an explicit mechanism to identify the

[11] See Antle (1983) and the related discussion in Kimhi (2006).

applied technology it may be impossible to completely distinguish among production functions using data generated by a mix of technologies. With this in mind, we estimate a weighted average value for α_i using[12]:

$$x_i = \gamma \hat{y}_i + \acute{s}_i + e_i \qquad 1)$$

Note that as the data used to estimate the relationships implied by $\Phi(s)$ becomes more granular, that is, as we move from country to regional to household to plot data, less information remains outside the fixed effects.

Before moving on to the empirical section, it is worth emphasizing that the notion of endogenous technology is a micro-economic concept. For applied work, there is a practical issue of whether less detailed data can contribute to empirical studies. This is a general problem that appears often when micro-models are used to guide empirical applications. The question is an important once, since there are significant differences in the scale-structure of farming among countries. To examine these differences empirically is challenging since detailed microeconomic surveys may not include a spread in the types of determinants – say institutions or climate – that are suspected determinants of observed cross-country variations. In the particular case of inverse-yield studies, given the large set of empirical findings and from diverse types of data, it is therefore worth asking, for example, whether farm-level averages, or even national averages of farms are representative of plot-level outcomes. We take up this issue in the following section and apply the microeconomic model across three types of data.

Empirical results

We estimate equation one using two pooled sets of data reporting maize production and yields that exemplify the types of data behind most empirical studies. The first set is a collection of household surveys covering ten countries in Sub-Saharan Africa that report production and yields at the household level. Taken together, the pooled dataset contains over 62,000 observations (Table 1). The country coverage for second set is described in the center columns of Table 1. The dataset pulls together observations from five countries from surveys that report plot-level data and contains over 8,000 observations.[13] Four surveys that contain detailed plot-level production data comprise the

[12] The weighting will depend on the estimation procedure. See Mundlak and Larson (1992) for a discussion in terms of panel data.
[13] The datasets are distinct; that is, plot-level observations were not aggregated and included in the household-level pool. Nor are the plot level observations included in the third.

third dataset, described in the far right columns of the table. We use this data to estimate separate models for each country setting.

Pooled regressions

The first two regressions focus on the two pooled datasets. Though the surveys contain limited information about farming methods employed, the datasets contain outcomes across a wide and divergent geography, with differing agro-climatic conditions, transportation networks and political institutions. The data also spans considerable time and this potentially implies a change in available technologies.

For each household, the total area under crop was calculated and used to classify the scale of farming into four categories (quantiles) from small to large. Summary statistics about farm size, area planted to maize, and yields by each of the four categories is given in Table 2. Although average yields were higher from the countries included in the plot-level pool, the scale of farming is quite similar, with more than three-quarters of the farms planting less than two hectares to crops.

Table 3 reports estimates from the fixed-effects model, given by equation 1, for each of the two pools. In the case of the household pool, yields represent an average based on an aggregation over plots. The fixed-effects used to represent the farmers' decision environment are constructed from country and survey year identifiers and therefore do not contain information about in-country differences. In contrast, household identifiers are used to capture differences among household decision environments in the pooled plot-level data. In this case, the model is used to explain the remaining variation in plot yields, and the omitted information has to do with unobserved plot-level differences, including soil fertility and applied inputs. In the empirical literature, this distinction can be lost, since studies typically use one type of data or the other. However, the difference may be relevant for policy, since studies based on household data deal with the average relationships between yields and the scale of production while the plot-based studies deal with the effects of fragmentation of farms on yields.

Nevertheless, despite significant difference in the granularity of the data, the regression outcomes are remarkably similar. In both cases, there is a negative correlation between land and yields; the correlation estimates are both statistically significant and are of similar magnitude (-0.161 versus -0.206). Results reported in the lower panel of the table under the heading "Core 90 percent" shows that the results are not especially affected by extreme values; tossing out the top and bottom

5 percent of yields in the samples lowers the estimated correlations, but the changes are quantitatively small.

Detailed datasets

In this section we turn our attention to the four datasets containing detailed production data at the plot level. The datasets are the result of two survey efforts. The first two datasets are derived from the Malawi Third Integrated Household Survey (IHS3) and the Tanzania National Panel Survey (TZNPS), products of collaboration between the World Bank and the Governments of Malawi and Tanzania as part of the Living Standards Measurement Study – Integrated Surveys on Agriculture initiative. The IHS3 was conducted from March 2010 to March 2011 by the Malawi National Statistical Office, covering 121,271 rural and urban households across Malawi. The TZNPS is a biannual survey conducted by the Tanzania National Bureau of Statistics (NBS). The first round implemented between September 2008 and October 2009 is used here. The TZNPS is based on a stratified two-stage cluster sample design. The total sample size is approximately 3,200 households. The data are nationally-representative, and are also representative at the level of the major agro-ecological zones and for Dar es Salaam, other mainland urban areas, mainland rural areas, and Zanzibar.

The second two datasets are the result of the REPEAT project which is an ongoing longitudinal rural household survey in Kenya and Uganda by National Graduate Institute for Policy Studies (GRIPS) and the Foundation for Advanced Studies on International Development (FASID) in Japan with collaboration of the TEGEMEO research institute in Kenya and Makerere University in Uganda. We use the survey data in 2004 and 2007 for Kenya covering 890 households and 2003 and 2005 for Uganda covering 940 households.[14]

For each country, surveyed farms were classified into one of four quantiles based on farm size; basic information on yields, farm size, plot size and number of plots is given in Table 4. As the surveys confirm, household farms are quite small in all four countries. In the case of Malawi, only about 22 percent of the surveyed farms exceeded one hectare. The surveyed farms were slightly larger in Kenya, Tanzania and Uganda, but the average size for the largest class of farms was less than five hectares for all three countries. In each survey farmed land is broken into still smaller plots. There is a tendency for plots to be larger on large farms, but multiple plots are common regardless of farm size. Looking at the last column of Table 4, the inverse yield relationship is not always

[14] Details about the REPEAT project and surveys are given by Yamano, et. al. 2004 and Yamano et. al. 2005.

readily apparent in the average yields across classes of farm size, except in Tanzania. Often, a U-shaped relationship is observed instead.

Malawi

The Malawi data contains observations on 8,848 households growing maize on 12,408 plots, a significant sub-set of the 12,000 households surveyed in IHS3. The survey reports on three types of labor – household provided labor, hired labor and exchange labor, a reciprocal arrangement where neighbors help each other at crucial times in the growing season. The amount of chemical fertilizers, herbicides and pesticides used per plot is recorded as well as whether or not organic fertilizer was applied. Table 5 reports average input use per hectare by farm-size quantile in two ways for Malawi and the other countries analyzed in this section. In the left-hand side of the table group-wise averages are reported, inclusive of zero-values. In the case of Malawi, no clear pattern emerges in the intensity of use for labor or for the remaining inputs. When the intensity of use is measured only when the decision has been made to use the input, the story is somewhat different; the per hectare use of exchange labor seems to decline as does the intensity of fertilizer use, while the intensity of herbicide use seems to increase with scale.

Table 6 provides information on the types of farming methods employed by scale of operation by country, with Malawi listed first. Surprisingly, the share of plots that are intercropped and the share of plots where hybrid seeds are used do not vary greatly with scale. Few households choose to employ farming technologies that rely on pesticides or herbicides, although farmers choose to use chemical fertilizers on most plots regardless of scale. The use of exchange labor is infrequent, but relatively constant across scale. As might be expected, larger farms are most likely to hire labor, although hired workers also work on the smallest farms and when employed are used more intensively (recall Table 5). Although few plots were rented in total, farmers with the smallest farms were most likely to rent plots.

Table 7 contains regression results based on the applied model. As with the pooled data, the log of land planted to maize per plot was regressed on the log of yield, fixed household effects, and two additional indicators of technology: whether or not hybrid seeds were used and whether intercropping techniques were used on the plot. The exercise was repeated using the per hectare intensity of the remaining inputs -- those inputs corresponding to the z in equation 1. To accommodate zero-valued observations, the variables were not converted to logs; however, to

facilitate comparisons, elasticities are reported in the table based on estimated parameters and mean values.

Turning to specific results from the regressions for Malawi, the estimated correlation between land and yield, at -0.226, is similar quantitatively to the pooled plot results in Table 3 and is also statistically significant. This is consistent with the observation that factor intensity declines with scale. The same message comes through from the remaining factor regressions. For each factor, the correlation is negative and, with the exception of exchange labor, statistically significant. As before, the regressions were rerun after clipping the upper and lower 5 percent of the observations based on yield. The yield correlation is unaffected quantitatively or qualitatively. There are some variations in the estimates for the remaining inputs although all correlations remain negative. In the truncated sample, the parameter on herbicide, an input rarely used, is no longer distinguishable from zero.[15]

Tanzania

The survey from Tanzania covers 1,786 maize plots on 1,272 farms, slightly more than half of the farms covered by TZNPS. Labor falls into two groups, household and hired, both measured in working days, and plot-level inputs are recorded for organic and chemical fertilizers and herbicides and pesticides. The intensity of input use, on average and conditional on use, are reported by type of input and scale of production, are reported in Table 5. In contrast to the corresponding measures for Malawi, the intensity of labor use declines noticeably with farm size. The patterns for the remaining inputs are less pronounced. While about half of the observed maize plots were intercropped in Malawi, about two-thirds of the maize plots from the Tanzania were intercropped – a pattern that was consistent across farm size (Table 6). Few farmers in Tanzania used hybrid seeds – about 14 percent of the surveyed plots – and the use of hybrid seed varied little with scale. As in Malawi, farmers with the smallest farms are more likely to rent in plots although the vast majority of small-scale farmers do not. About 27 percent of the smallest farms hired workers, a slightly smaller share than that of the largest category of farms, 36 percent of which entered the labor market. As a whole, few maize farmers relied heavily on either labor or rental markets and for about 64 percent of the plots, farmers relied on neither market.

[15] The estimated intercepts and the discrete impacts on plot technologies from the same regressions are given in Annex table 1. In each for Malawi, the use of hybrid seeds was land conserving, while the choice to intercrop resulted in a more extensive use of land. In each regression, the estimated parameters are statistically significant. The results from other countries were mixed.

Estimation results from the Tanzania data are given in Table 7. The estimated correlation between land and yield is statistically significant and, at -0.246, quite close to results based on the Malawi survey. With the exception of the elasticity for hired labor, the results show that the inputs are land-saving and are in line with the results from Malawi. With the exception of the herbicide-pesticide elasticity, the remaining estimates are statistically distinguishable from zero. The elasticity on hired labor is positive and significant at just over the 5 percent level. This suggests that increasing the per-hectare use of hired labor is associated with larger plots of maize. This is not unreasonable, but it is different from the results in Malawi and with the negative relationships between factor intensification and land use. As shown in the lower "Core 90 percent" panel of Table 7, this particular finding and the results from Tanzania in general, were not changed by dropping the extreme yield values from the sample.

Kenya

The Kenya data contains 3,152 maize plots of 814 maize producing households from 2 waves of the panel survey in Kenya in 2004 and 2007 covering 4 seasons (2 crop seasons per year for 2 years). Table 4 reports the average farm size, number of maize plots, size of maize plots, and average yield by farm-size quantile based on the cultivated land size of the households in 2004.

The intensity of input use, on average and conditional on use, are reported by type of input and scale of production in Table 5. The intensity level of household labor use, hired labor, and organic fertilizer use clearly declines with farm size while hybrid seed use and chemical fertilizer use do not show such a clear pattern. In Kenya, compared to other countries, hybrid seed and chemical fertilizer were used in higher proportion and their application level is high (Table 6). This observation indicates prevalence of the intensive farming method. More than 10 percent of plots are rented in for maize production on average, which is the highest among the four countries. About 90 percent of plots were intercropped. Farmers with the smallest farms are more likely to rent in plots. The larger farms are more likely to hire labor or draft animals. On the face of things, these facts indicate that rural labor market and land rental market are active relative to other countries.

Table 7 reports the results of the regressions. As before, the log of land planted to maize per plot was regressed on the log of yield, chemical fertilizer use per hectare, organic fertilizer use per hectare, total hours of family labor, and cost for the hired labor and drafted animal with other covariates: the household fixed effect, intercropping dummy, and a hybrid seed use dummy. In addition, since the data is a panel, a season-year dummy is included as well.

The elasticity of plot size is negative and statistically significant for both yield and all the inputs. Quantitatively, the elasticity at -0.49 is larger in absolute terms than the estimates for Malawi and Tanzania and this holds true of the elasticities for the factor intensities. It falls when the core dataset is used to -0.363. These results show that the yield declines with the plot size, which is consistent with the inverse yield relationship. They also show that the inverse relation would be due to the inverse input use relationship.

Uganda

The Uganda data contains 3,200 maize plots of 825 maize producing households from 2 waves of the panel survey in Uganda in 2003 and 2005 covering 4 seasons (2 crop seasons per year for 2 years). The intensity of input use, on average and conditional on use, are reported by type of input and scale of production in Table 5. The intensity of household labor use declines while hired labor cost increases in farm size. Other inputs do not show clear patterns. The table indicates that the intensity level of maize production is very low in Uganda unlike Malawi and Kenya. Chemical fertilizer was used in only 2 percent of maize plots and organic fertilizer was use in only 4 percent. On average 70 percent of maize plots were intercropped and the percentage decreases with farm size. Hybrid seed and hired labor are more likely to be used in larger farms. Again, the smallest farms were more likely to hire plots for maize, although rental land was used on few farms of any scale (Table 6).

Table 7 reports the results of the regression with household fixed effects. The plot size elasticity on yield and family labor use is negative and statistically significant while those of other inputs are not significant. At -0.557 the land-yield elasticity is the greatest in absolute terms of the country estimates. It remains significant but falls to -0.335 when we use only the core 90 percent samples for the regressions. The yield and family labor use declines with the plot size. It seems that the inverse yield relationship is associated with the decline in family labor use.

Summary

Taken together, the four data sets describe farming sectors where plots are small and households rely primarily on household labor and land that they own, conditions consistent with a conceptual depiction analyzed by Feder (1985). Only in Kenya are farmers likely to hire draft animals or employ workers. Other indicators of chosen technologies – for example, the use of intercropping, fertilizer or high-yielding seeds – vary across countries, but vary less within country by scale. Differences in

fertilizer use are indicative; a majority of maize farmers in Kenya and Malawi apply chemical fertilizers to their plots regardless of scale, while fertilizer is rarely used in Tanzania and Uganda.

Keeping in mind that what classifies as a large farm in Malawi Tanzania, Kenya or Uganda is small by standards in Europe, Central Asia or the Americas, within country variation comes not so much from differences in applied inputs, but rather the intensity of their application. This is most visible in the use of household labor in Tanzania, Kenya and Uganda. The difference in labor intensity is less obvious in Malawi, but then there is also little variation in the size of farms in Malawi.

The statistical analysis confirms the inverse yield relationship for maize at the plot level. The calculated elasticities, estimated independently for each sample, are distinguishable from zero with a high degree of statistical confidence in each case. Moreover, the elasticities are quantitatively similar to results based on the assorted pool of plot level data reported in Table 3 once the extreme tails of the yield distribution are excluded.

The analysis also suggests an inverse relationship between scale and household labor use, an empirical phenomenon that is related to how labor markets work in Sub-Saharan Africa and a potential determinate of why farms are small in Africa. In this case, there is greater variation in the quantitative estimates; however, the distinction falls according to survey groups – with the World Bank LSMS surveys producing lower household labor elasticities than the GRIPS REPEAT surveys. Since measuring household labor input is notoriously difficult, it may be that sampling methods play a role.[16]

Summary and conclusions

In this paper, we revisit the inverse-yield relationship, the observed phenomenon that cereal yields decline as the scale of production increases in Sub-Saharan Africa. The relationship is important for policy since it is taken as evidence that the small-scale structure of farming in Africa is a natural and efficient outcome, given the pervasive constraints on the sector, particularly those related to land and labor markets. In surn, this has had important implications for the evolution of African agricultural policies; currently, efforts to improve the productivity of smallholder farms are a ubiquitous core feature of the development strategies promoted by African governments and international development agencies.

[16] See Beegle et al. (2012) in the context of measuring household consumption in developing countries.

Early empirical evidence in favor of the inverse-yield relationship prompted criticism, since the early analyses omitted potentially important information about natural endowments. Moreover, systematic errors in the survey methods were thought to create biased results. As a consequence, empirical studies became increasingly focused on the variation in plot yields for a given farm and based on detailed data with a limited geographic reach. This created tension, since what matters for policy is the generalization that small farms are an appropriate and a natural outcome of the interplay of endowments and markets. In contrast, researchers look for settings that limit the scope for unmeasured influences that undermine the credibility of results, and thereby narrow the geographic relevance of the results.[17]

In this paper, we develop a model of endogenous technology choice in which farming methods (applied technologies) are an outcome of the economic environment in which farmers operate. One consequence is that heterogeneity in farming households, farm endowments and markets generate heterogeneity in applied technologies. Econometrically, the underlying production function associated with a given applied technology is identified by the relevant state variables, and incompleteness in the observed state variables introduces a greater likelihood of mis-identification. With this in mind, we develop an applied model to measure the inverse relationship between maize yields and production scale in Sub-Saharan Africa. To estimate the model, we progress from coarse data that is geographically and temporally diverse to more granular and place-specific data, with the presumption that the opportunity for omitting relevant information about the state variables is reduced. We find empirical evidence that supports the hypothesis that maize yields fall with increases in the scale of production in all settings. Moreover, once extreme yield values are expunged from the data, the estimated elasticities measuring the inverse relationship are remarkably similar across the assorted data platforms.

At the same time, yields and the relationship between yields and scale are narrow measures and the more detailed data provides a richer ground for describing how applied technologies differ by scale. For the four detailed datasets included in our study, evidence of an inverse relationship between household labor and scale is also evident. More generally the data suggests that variations in scale are associated with how intensively inputs are applied, rather than in the variation of adopted inputs.

[17] At least, this is the case in the narrow applications of individual studies. Collectively, we believe the accumulated evidence cited in the front of the paper is consistent with our findings.

Bibliography

AGRA. 2011. What is the Alliance for a Green Revolution in Africa? www.agra-alliance.org, downloaded June 16, 2011.

Antle, John M. 1983. Sequential decision making in production models. American Journal of Agricultural Economics 65, 282–290.

Assuncao, Juliano and Luiz Henrique Braido. 2007. Testing household-specific explanations for the inverse productivity relationship. American Journal of Agricultural Economics 89(4), 980-990.

Bardhan, Pranab K. 1973. Size, Productivity, and Returns to Scale: An Analysis of Farm-Level Data in Indian Agriculture. The Journal of Political Economy 81(6), 1370-1386.

Barrett Christopher B. 1996. On price risk and the inverse farm size-productivity relationship. Journal of Development Economics 51(2), 193-215.

Barrett, Christopher B., Marc F. Bellemare, Janet Y. Hou. 2011. Reconsidering conventional explanations of the inverse productivity–size relationship, World Development 38(1), 88-97.

Bauer P. T. 1946. The Working of Rubber Regulation. Economic Journal 56 (223), 391-414.

Beegle, Kathleen; Joachim De Weerdt, Jed Friedman, and John Gibson. 2012. Methods of Household Consumption Measurement through Surveys: Experimental Results from Tanzania. Journal of Development Economics 98 (1), 3-18

Benjamin, Dwayne. 1995. Can unobserved land quality explain the inverse productivity relationship .Journal of Development Economics 46 (1), 51–84.

Berry, R. Albert, and William Robert Cline. 1979. Agrarian Structure and Productivity in Developing Countries. Baltimore, MD: Johns Hopkins Univ. Press.

Bhalla Surjit.1979. Farm size, productivity and technical change in Indian agriculture, appendix. In R. Albert Berry and William Robert Cline (eds.). Agrarian Structure and Productivity in Developing Countries. Baltimore, MD: Johns Hopkins University Press.

Bhalla, Surjit S. and Prannoy Roy. 1988. Mis-Specification in farm productivity analysis: the role of land quality. Oxford Economic Papers, 40(1), 55–73.

Binswanger, Hans. Brazilian policies that encourage deforestation in the Amazon. World Development 19(7), 821-829.

Binswanger, Hans and John McIntyre. 1987. Behavioral and material determinants of production relations in land abundant tropical agriculture. Economic Development and Cultural Change 36, 75–99.

Binswanger, Hans P. and Mark R Rosenzweig. 1986. Behavioral and material determinants of production relations in agriculture. Journal of Development Studies 22(3). 503–539.

Carletto, Calogero, Sara Savastano and Alberto Zezza. 2011. Fact or artefact: the impact of measurement errors on the far size-productivity relationship. Policy Research Working Paper 5908. Washington: World Bank.

Carter, Michael R. 1984. Identification of the inverse relationship between farm size and productivity: an empirical analysis of peasant agricultural production. Oxford Economic Papers 36(1), 131-145.

Carter, Michael R. and Keith D. Wiebe. 1990. Access to capital and its impact on agrarian structure and productivity in Kenya. American Journal of Agricultural Economics 72(5), 1146–1150.

Collier, Paul. 2008. The politics of hunger. Foreign Affairs 87(6), 67–79.

Deininger, Klaus and Gershon Feder. 2001 Chapter 6: Land institutions and land markets. In: Bruce L. Gardner and Gordon C. Rausser, Editor(s), Handbook of Agricultural Economics, Volume 1, Part A. Amsterdam: Elsevier.

Deolalikar, Anil B. 1981. The inverse relationship between productivity and farm size: a test using regional data from India. American Journal of Agricultural Economics 63(2), 275–279.

Diouf, Jacques. 2011. Foreword to Save and Grow: A Policymaker's Guide to the Sustainable Intensificaion of Smallholder Crop Production. Rome: Food and Agriculture Organization of the United Nations.

Eastwood, Robert, Michael Lipton and Andrew Newell. 2005. Farm size. In: R. Evenson and P. Pingali, Editors, Handbook of Agricultural Economics vol. 3A. Amsterdam: North Holland Press.

Fan, Shenggen. 2011 Press Statement June 15, 2011. Washington: International Food Policy Research Institute.

FAO. 2012. FAOSTAT data base. Food and Agriculture Organization of the United Nations.

Feder, Gershon. 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.

Foster, Andrew D. and Mark R. Rosenzweig. 2010. Barriers to farm profitability in India: mechanization, scale and credit markets. Draft.

Hazell, Peter, Colin Poulton, Steve Wiggins, and Andrew Dorward. 2010. The future of small farms: trajectories and policy priorities. World Development 38(10), 1349–1361.

Heltberg, Rasmus. 1998. rural market imperfections and the farm size-productivity relationship: evidence from Pakistan. World Development 26, 1807–182.

Ivanic, Maros, Will Martin and Hassan Zaman. 2011. Estimating the short-run poverty impacts of the 2010-2011 surge in food prices. Policy Research Working Paper 5633. Washington: World Bank.

Kevane, Michael. 1996. Agrarian structure and agricultural practice: typology and application to Western Sudan. American Journal of Agricultural Economics 78(1), 236–245.

Kimhi, Ayal. 2006. Plot size and maize productivity in Zambia: is there an inverse relationship?. Agricultural Economics 35(1), 1–9.

Lamb, Russell L. 2003. Inverse productivity: land quality, labor markets, and measurement error. Journal of Development Economics 71 (1), 71–95.

Larson, Donald, and Yair Mundlak. 1997. On the intersectoral migration of agricultural labor. Economic Development and Cultural Change 45(2), 295-319.

Lipton, Michael. 2006. Can small farmers survive, prosper, or be the key channel to cut mass poverty? The Electronic Journal of Agricultural and Development Economics 3(1), 58–85.

Mazumdar, Dipak. 1965. Size of Farm and Productivity: A Problem of Indian Peasant Agriculture Economica 32(126), 161-173.

Mundlak Yair, Donald F. Larson and Rita Butzer. 1999. Rethinking within and between regressions: the case of agricultural production functions. Annales d'Économie et de Statistique 55/56, Économétrie des Données de Panel, 475-501.

Mundlak, Yair and Donald F. Larson. 1992. On the Transmission of World Agricultural Prices World Bank Economic Review 6(3): 399-422

Mundlak, Yair, Rita Butzer and Donald F. Larson. 2012. Heterogeneous technology and panel data: The case of the agricultural production function. Journal of Development Economics 99(1), 139-149.

Mundlak, Yair. 1988. Endogenous Technology and the Measurement of Productivity. In Susan M. Capalbo and John M. Antle (eds) Agricultural Productivity: Measurement and Explanation. Washington: Resources for the Future.

Mundlak, Yair. 1993. On the empirical aspects of economic growth theory. American Economic Review 83, 415-420.

Nagayets, Oksana. 2005. Small farms: current status and key trends. In The Future of Small Farms. Proceedings of a Research Workshop. Washington: International Food Policy Research Institute.

Otsuka, Keijiro and Donald F. Larson. forthcoming. An African Green Revolution: Finding Ways to Boost Productivity on Small Farms. Dordrecht: Springer.

Pingali, Prabhu, Yves Bigot, and Hans P. Binswanger. 1987. Agricultural Mechanization and the Evolution of Farming Systems in Sub-Saharan Africa. Baltimore: Johns Hopkins University Press.

Ravallion, Martin, Shaohua Chen, and Prem Sangraula. 2007. New evidence on the urbanization of global poverty. Background paper for the WDR 2008.

Schultz, Theodore W. 1964. Transforming Traditional Agriculture. New Haven, London: Yale University Press.

Sen, Amartya K. 1966. Peasants and dualism with or without surplus labor. The Journal of Political Economy 74 (5), 425–450.

Smale, Melinda, Derek Byerlee and Thom Jayne. Maize revolutions in Africa. Forthcoming. In Keijiro Otsuka and Donald F. Larson (eds.) An African Green Revolution: Finding Ways to Boost Productivity on Small Farms. Dordrecht: Springer.

Taslim, M.A. 1989. Supervision problems and the size-productivity relation in Bangladesh agriculture. Oxford Bulletin of Economics and Statistics 51(1), 55-71.

von Braun, Joachim, 2005. Science and technology policies for agricultural productivity and growth in developing countries (PowerPoint Presentation), Agricultural Outlook Forum 2005 32857, United States

Department of Agriculture, Agricultural Outlook Forum. Available on the Internet at: http://ageconsearch.umn.edu/handle/32857.

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

Yamano, Takashi, Dick Sserunkuuma, Keijiro Otsuka, George Omiat, John Herbert Ainembabazi, and Yasuharu Shimamura. 2004. The 2003 REPEAT Survey in Uganda: Results, available at: http://www3.grips.ac.jp/~globalcoe/j/data/repeat/REPEATinUgandaReport.pdf.

Yamano, Takashi, Keijiro Otsuka, Frank Place, Yoko Kijima, and James Nyoro. 2005. The 2004 REPEAT Survey in Kenya (First Wave): Results, available at: http://www3.grips.ac.jp/~globalcoe/j/data/repeat/ReportKenya2004F.pdf

Yotopoulos, P.A. and L. J. Lau. 1973. A test for relative economic efficiency: some further results. American Economic Review 63(1), 214–223.

Zoellick, Robert. 2011. Free markets can still feed the world. Financial Times January 5, 2011.

Tables

Table 1: Household and plot datasets

Pooled household		Pooled plot		Detailed plot data	
Country	**Observations**	**Country**	**Observations**	**Country**	**Observations**
Ghana	677	Ethiopia	4,681	Kenya	3,109
Kenya	6,347	Madagascar	388	Malawi	12,408
Madagascar	755	Mali	176	Tanzania	1,787
Mozambique	1,306	Tanzania	2,003	Uganda	3,195
Malawi	12,099	Uganda	1,033		
Nigeria	1,760				
Rwanda	3,941				
Tanzania	27,786				
Uganda	5,170				
Zambia	2,195				
Total	62,036	Total	8,281	Total	20,499

Note: Included surveys were conducted between 1999 and 2009.

Table 2: Farm and plot characteristics by quantile.

		Averages from plot pool		
Quantile	Farm size	Plot maize area	Maize yield	Number of plots
1	0.37	0.23	1,479	1.23
2	0.92	0.37	1,470	1.38
3	1.67	0.45	1,427	1.56
4	10.62	0.78	1,128	1.92
Full sample	3.38	0.46	1,376	1.52
		Averages from household pool		
Quantile	Farm size	Farm maize area	Maize yield	
1	0.41	0.29	1,070	
2	0.96	0.50	1,006	
3	1.78	0.74	1,012	
4	7.28	1.52	1,037	
Full sample	2.45	0.74	1,030	

Note: Farm size is given by cultivated area for all crops on all plots.

Table 3: Fixed-effects model results from pooled datasets

	Household pool		Plot pool	
	Coefficient	**Std. Err.**	**Coefficient**	**Std. Err.**
Full sample				
Log yield	-0.161*	0.003	-0.206*	0.020
Constant	0.364*	0.020	0.215	0.135
Rho	0.285+		0.606+	
Core 90 percent				
Log yield	-0.135*	0.004	-0.178*	0.030
Constant	0.188*	0.027	0.016	0.203
Rho	0.310+		0.611+	

Note: Rho is the fraction of variance due to the fixed effects.* denotes that the associated t-score is significant at the 1 percent level. + denotes that the associated F-score is significant at the 1 percent level. The highest and lowest 5 percent of observed yields were dropped from the full samples to create the "Core 90" samples.

Table 4: Sample averages for Malawi, Tanzania, Kenya and Uganda

	Farm size	**Number of maize plots**	**Plot area in maize**	**Maize yield**
Quantile		Malawi		
1	0.15	1.46	0.27	1,502
2	0.45	1.97	0.30	1,368
3	0.74	2.20	0.38	1,423
4	2.47	2.20	0.60	1,591
Full sample	0.95	1.96	0.39	1,472
Quantile		Tanzania		
1	0.47	1.50	0.32	889
2	1.20	2.05	0.49	846
3	2.10	1.74	0.73	758
4	5.35	1.96	1.22	788
Full sample	2.24	1.79	0.68	821
Quantile		Kenya		
1	0.40	1.22	0.20	1,575
2	0.94	1.42	0.40	1,672
3	1.67	1.51	0.60	1,383
4	4.19	1.55	0.92	1,588
Full sample	1.74	1.42	0.52	1,567
Quantile		Uganda		
1	0.43	1.21	0.21	1,199
2	1.03	1.23	0.30	1,119
3	1.88	1.42	0.51	1,181
4	4.92	1.51	0.82	1,123
Full sample	2.01	1.35	0.48	1,154

Note: Farm size in Malawi and Tanzania is the area owned by the household. In Kenya and Uganda, farm size is adjusted by net rentals.

Table 5: Malawi: Plot-level average use per hectare by farm size.

	Average use per hectare, including zero values					Average use per hectare when input is used (non-zero use)				
	Quantile					Quantile				
	1	2	3	4	Full sample	1	2	3	4	Full sample
Malawi										
Household labor (hours/ha)	2,379	2,344	2,754	2,425	2,476	2,440	2,360	2,771	2,459	2,509
Hired labor (hours/ha)	8.45	5.75	4.92	8.06	6.81	33.42	29.89	23.78	24.91	27.88
Exchange labor(hours/ha)	2.49	2.23	2.34	1.56	2.16	24.00	20.89	19.64	15.75	20.12
Herbicide and pesticide use (kg/ha)	0.06	0.08	0.14	0.48	0.19	7.29	15.69	30.71	33.41	23.57
Chemical fertilizer use (kg/ha)	191.46	168.85	154	165.52	170.07	255.12	224	203.7	209.13	222.93
Tanzania										
Household labor (days/ha)	239.96	170.22	115.83	90.1	156.64	242.34	171.11	116.34	90.72	157.71
Hired labor (days/ha)	9.01	7.87	7.41	6.42	7.72	33	22.46	22.54	18.02	23.81
Herbicide and pesticide use (kg/ha)	1.01	0.21	1.85	1.12	1.08	9.66	2.59	15.26	7.91	9.56
Chemical fertilizer use (kg/ha)	16.92	20.59	15.85	16.86	17.43	114.84	141.56	92.07	101.17	110.31
Organic fertilizer use(kg/ha)	22.68	27.4	18.42	23.45	22.81	160.36	234.42	117.51	168.36	163.49
Kenya										
Household labor (hours/ha)	1,287.9	921.0	810.8	648.8	930.3	1,294.0	922.9	834.1	679.8	947.4
Hired labor and animal cost (Ksh/ha)	9,530	6,174	5,617	5,754	6,802	17,049	9,828	8,356	7,780	10,534
Chemical fertilizer use (kg/ha)	90.07	91.42	85.97	86.55	88.79	118.70	119.81	118.61	120.32	119.40
Organic fertilizer use (kg/ha)	2,763.1	1,646.8	1,366.8	884.5	1,679.8	5,101.0	3,187.2	2,754.2	2,146.4	3,399.4
Uganda										
Household labor (hours/ha)	1,017	909	809	714	847	1,032	912	825	734	862
Hired labor and animal cost (Ush/ha)	18,527	27,032	28,503	38,827	29,084	78,741	92,242	73,677	84,378	81,774
Chemical fertilizer use (kg/ha)	0.68	0.49	0.66	0.90	0.69	36.48	29.16	23.02	47.14	32.43
Organic fertilizer use (kg/ha)	34.5	31.7	15.8	8.5	21.4	554.7	519.2	445.1	317.5	479.3

Table 6: Share of plots by farming techniques

Quantile	1	2	3	4	Full sample
Malawi					
Inter-cropped	0.49	0.54	0.48	0.43	0.48
Hybrid seed used	0.53	0.45	0.44	0.42	0.46
Labor hired	0.25	0.19	0.21	0.32	0.24
Exchange labor used	0.10	0.11	0.12	0.10	0.11
Plot rented	0.22	0.02	0.02	0.01	0.07
Herbicide pesticide used	0.01	0.01	0.01	0.01	0.01
Chemical fertilizer used	0.75	0.75	0.76	0.79	0.76
Organic fertilizer used	0.11	0.12	0.14	0.15	0.13
Tanzania					
Inter-cropped	0.67	0.65	0.67	0.67	0.67
Hybrid seed used	0.16	0.14	0.12	0.14	0.14
Labor hired	0.27	0.35	0.33	0.36	0.32
Plot rented	0.16	0.04	0.02	0.02	0.06
Herbicide pesticide used	0.10	0.08	0.12	0.14	0.11
Chemical fertilizer used	0.15	0.15	0.17	0.17	0.16
Organic fertilizer used	0.14	0.12	0.16	0.14	0.14
Kenya					
Inter-cropped	0.91	0.91	0.87	0.86	0.89
Hybrid seed used	0.57	0.60	0.58	0.61	0.59
Labor or animal hired	0.56	0.63	0.67	0.74	0.65
Plot rented	0.14	0.14	0.10	0.09	0.12
Chemical fertilizer used	0.76	0.76	0.72	0.72	0.74
Organic fertilizer used	0.54	0.52	0.50	0.41	0.49
Uganda					
Inter-cropped	0.78	0.75	0.66	0.64	0.70
High yielding seed used	0.15	0.23	0.22	0.24	0.22
Labor or animal hired	0.24	0.29	0.39	0.46	0.36
Plot rented	0.20	0.08	0.08	0.05	0.09
Chemical fertilizer used	0.02	0.02	0.03	0.02	0.02
Organic fertilizer used	0.06	0.06	0.04	0.03	0.04

Note: For Uganda, high-yielding seeds include hybrid and open pollinated varieties.

Table 7: Input elasticities from household fixed effects estimations.

	Full data set		Core 90 data	
	Elasticity	t-score	Elasticity	t-score
Malawi				
Yield	-0.226[a]	-18.04	-0.222[a]	-12.69
Household labor	-0.043[a]	-6.95	-0.064[a]	-6.85
Hired labor	-0.019[a]	-5.64	-0.006[c]	-1.78
Exchange labor	-0.002	-1.37	-0.003	-1.01
Herbicide-pesticide	-0.013[a]	-11.34	-0.001	-1.19
Chemical fertilizer	-0.135[a]	-18.10	-0.093[a]	-8.76
Organic fertilizer*	-0.053	-1.45	-0.080[b]	-2.14
Tanzania				
Yield	-0.246[a]	-6.06	-0.218[a]	-3.88
Household labor	-0.100[a]	-7.23	-0.092[a]	-6.21
Hired labor	0.033[c]	1.91	0.035[b]	2.04
Herbicide-pesticide	-0.004	-0.80	-0.001	-0.17
Chemical fertilizer	-0.021[b]	-2.55	-0.033[b]	-2.15
Organic fertilizer	-0.013[a]	-2.75	-0.010[c]	-1.70
Kenya				
Yield	-0.490[a]	-16.27	-0.363[a]	-14.80
Household labor	-0.666[a]	-11.18	-0.631[a]	-9.20
Hired labor and animal cost	-0.681[a]	-6.15	-0.794[a]	-6.10
Chemical fertilizer	-0.184[a]	-5.26	-0.199[a]	-5.24
Organic fertilizer	-0.415[a]	-6.03	-0.439[a]	-5.37
Uganda				
Yield	-0.557[a]	-15.06	-0.335[a]	-10.82
Household labor	-0.601[a]	-9.27	-0.524[b]	-6.51
Hired labor and animal cost	0.147	1.27	0.205	1.44
Chemical fertilizer	0.169	0.64	0.166	0.47
Organic fertilizer	-0.107	-0.71	-0.169	-0.89

Note: Elasticities were calculated at mean values.* denotes a discrete variables and the associated parameter is the percentage change in yield from adopting the input. Superscripts a, b and c denote significance at the 1, 5 and 10 percent level. For Kenya, time dummies (year-season dummies) are included in addition to household effects; for Uganda, year-season dummies are included.

Annex Tables

Table A1: Malawi: technology impact parameters associated with hybrid seed use and intercropping, from household fixed effects model.

Yield	Parameter	Std. Err	t-score	Pr>|t|
Hybrid seed used	-0.149	0.020	-7.280	0.000
Plot intercropped	0.142	0.030	4.660	0.000
Constant	0.390	0.121	3.210	0.001
Household labor				
Hybrid seed used	-0.156	0.021	-7.490	0.000
Plot intercropped	0.126	0.031	4.080	0.000
Constant	-1.070	0.021	-51.350	0.000
Hired labor				
Hybrid seed used	-0.158	0.021	-7.530	0.000
Plot intercropped	0.125	0.031	4.010	0.000
Constant	-1.127	0.019	-58.780	0.000
Exchange labor				
Hybrid seed used	-0.158	0.021	-7.520	0.000
Plot intercropped	0.124	0.031	3.970	0.000
Constant	-1.130	0.019	-59.240	0.000
Herbicide				
Hybrid seed used	-0.158	0.021	-7.500	0.000
Plot intercropped	0.125	0.031	3.990	0.000
Constant	-1.132	0.019	-60.040	0.000
Chemical fertilizer				
Hybrid seed used	-0.153	0.021	-7.370	0.000
Plot intercropped	0.120	0.031	3.900	0.000
Constant	-1.041	0.021	-48.620	0.000
Organic fertilizer*				
Hybrid seed used	-0.158	0.021	-7.500	0.000
Plot intercropped	0.126	0.031	4.040	0.000
Constant	-1.124	0.019	-57.990	0.000

Note: The technology impact parameters measure the percentage change in land due to a discrete switch from non-hybrids to hybrids and from mono-cropping to inter-cropping.

Table A2: Tanzania: Technology impact parameters associated with hybrid seed use and intercropping

Yield	Parameter	Std. Err	t-score	Pr>|t|
Hybrid seed used	0.352	0.185	1.910	0.057
Plot intercropped	0.062	0.080	0.780	0.437
Constant	0.644	0.260	2.480	0.013
Household labor				
Hybrid seed used	0.241	0.181	1.330	0.183
Plot intercropped	0.059	0.079	0.750	0.453
Constant	-0.761	0.062	-12.210	0.000
Hired labor				
Hybrid seed used	0.220	0.190	1.160	0.246
Plot intercropped	0.091	0.082	1.100	0.271
Constant	-0.912	0.064	-14.190	0.000
Herbicide				
Hybrid seed used	0.255	0.190	1.340	0.182
Plot intercropped	0.095	0.083	1.150	0.251
Constant	-0.882	0.063	-14.010	0.000
Chemical fertilizer				
Hybrid seed used	0.281	0.189	1.480	0.139
Plot intercropped	0.086	0.082	1.040	0.297
Constant	-0.863	0.063	-13.680	0.000
Organic fertilizer*				
Hybrid seed used	0.222	0.189	1.180	0.240
Plot intercropped	0.103	0.082	1.260	0.209
Constant	-0.875	0.063	-13.970	0.000

Note: The technology impact parameters measure the percentage change in land due to a discrete switch from non-hybrids to hybrids and from mono-cropping to inter-cropping.

Table A3: Kenya technology impact parameters associated with hybrid seed use and intercropping

Log of yield (kg/ha)	Parameter	Robust Std. Err	t-score	Pr>|t|
Hybrid seed used	0.153	0.053	2.900	0.004
Plot intercropped	0.059	0.066	0.890	0.372
Constant	6.477	0.086	75.690	0.000
Household labor (hours/ha)				
Hybrid seed used	-59.214	75.792	-0.780	0.435
Plot intercropped	53.489	111.291	0.480	0.631
Constant	270.593	126.942	2.130	0.033
Hired labor and animal cost (Ksh/ha)				
Hybrid seed used	-681.922	791.789	-0.860	0.389
Plot intercropped	-2669.899	1618.241	-1.650	0.099
Constant	2521.652	1768.819	1.430	0.154
Chemical fertilizer (kg/ha)				
Hybrid seed used	27.922	5.321	5.250	0.000
Plot intercropped	22.512	7.053	3.190	0.001
Constant	50.711	8.627	5.880	0.000
Organic fertilizer (kg/ha)				
Hybrid seed used	364.314	166.481	2.190	0.029
Plot intercropped	503.425	293.222	1.720	0.086
Constant	287.268	339.011	0.850	0.397

Note: The technology impact parameters measure the change in land due to a discrete switch from non-hybrids to hybrids and from mono-cropping to inter-cropping.

Table A4: Uganda technology impact parameters associated with high-yielding variety (HYV) seed use and intercropping

Log of yield (kg/ha)	Parameter	Robust Std. Err	t-score	Pr>|t|
HYV seed used	0.293	0.067	4.380	0.000
Plot intercropped	-0.356	0.053	-6.700	0.000
Constant	5.861	0.082	71.680	0.000
Household labor (hours/ha)				
HYV seed used	-266.816	113.502	-2.350	0.019
Plot intercropped	62.104	58.561	1.060	0.289
Constant	254.319	105.077	2.420	0.016
Hired labor and animal cost (Ksh/ha)				
HYV seed used	1568.799	9666.487	0.160	0.871
Plot intercropped	-14557.370	5442.326	-2.670	0.008
Constant	51578.530	7118.770	7.250	0.000
Chemical fertilizer (kg/ha)				
HYV seed used	1.479	0.521	2.840	0.005
Plot intercropped	-0.690	0.322	-2.140	0.032
Constant	0.538	0.499	1.080	0.281
Organic fertilizer (kg/ha)				
HYV seed used	-1.386	9.692	-0.140	0.886
Plot intercropped	13.225	6.190	2.140	0.033
Constant	14.005	9.078	1.540	0.123

Note: The technology impact parameters measure the change in land due to a discrete switch from non-hybrids to hybrids and from mono-cropping to inter-cropping.