Land Misallocation and Productivity†

Diego Restuccia
University of Toronto*

Raül Santaeulàlia-Llopis
Washington University in St. Louis**
Universitat de València
CEMFI

November 2015

Abstract

Using detailed household-farm level data from Malawi, we measure real farm total factor productivity (TFP) controlling for a wide array of factor inputs, land quality, and transitory shocks. We find that factor inputs are roughly evenly spread among farmers: operated land size and capital are essentially unrelated to farm TFP implying a strong negative effect on aggregate agricultural productivity. A reallocation of factors to their efficient use among existing farmers would increase agricultural productivity by a factor of 3.6-fold. We relate factor misallocation to severely restricted land markets as the vast majority of land is without a title and a very small portion of operated land is rented in. The gains from reallocation are 2.6 times larger for farms with no marketed land than for farms that operate marketed land. The efficient reallocation of factor inputs in the agricultural sector would trigger a profound process of structural change setting the farm size and the agricultural employment share of Malawi to industrialized levels.

Keywords: misallocation, land, productivity, agriculture, Malawi, micro data.
JEL codes: O1, O4.

†For helpful comments we thank Nadege Azebaze, Chang-Tai Hsieh, Leandro M. de Magalhaes, Matthias Doepke, Pascaline Dupas, Chad Jones, Pete Klenow, Evelyn Korn, Rody Manuelli, Nancy Qian, José-Víctor Ríos-Rull, Paul Romer, Nancy Stokey, Kei-Mu Yi, and seminar participants at Edinburgh, ENSAI, Stanford Institute for Theoretical Economics, SED Annual Meetings, NBER Summer Institute, University of Toronto, IMF Workshop on Macroeconomic Policy and Income Inequality, CIDE Mexico, Queen's University, University of Marburg, Universitat d'Alacant, CEMFI, RIDGE workshop, Wilfried Laurier University, HEC Montreal, Oxford CSAE, the World Bank Land & Poverty Conference and the ECAMA-IPFRI-IPA. All remaining errors are our own. Restuccia gratefully acknowledges the financial support from the Social Sciences and Humanities Research Council of Canada. Santaeulàlia-Llopis gratefully thanks the Weidenbaum Center on the Economy, Government, and Public Policy for financial support.

*150 St. George Street, Toronto, ON M5S 3G7, Canada. E-mail: diego.restuccia@utoronto.ca.
**Seigle Hall 339, Campus Box 1208, St. Louis, MO, 63130-4899, USA. E-mail: rauls@wustl.edu.

1 Introduction

A fundamental question in the field of economic growth and development is why some countries are rich and others poor. The literature has offered many useful perspectives but in this paper we build on two. First, agriculture is important in accounting for productivity differences between rich and poor countries. This is because poor countries are much less productive in agriculture and allocate most of their labor to this sector than rich countries. Moreover, low productivity in agriculture together with a subsistence constraint for food explains the large fraction of employment in agriculture in poor countries. As a result, the key question is why agricultural labor productivity is so low in poor countries.[1] Second, (mis)allocation of factors of production across heterogeneous production units is important in explaining differences in measured productivity across countries (e.g. Restuccia and Rogerson, 2008). In this paper, we use unique household-farm level data from Malawi — one of the poorest countries in the world — to measure real farm total factor productivity (TFP) controlling for a wide array of factor inputs, land quality, and transitory shocks. We show that operated land size and capital are essentially unrelated to farm TFP which provides direct and solid evidence of misallocation. Quantitatively, factor misallocation has a dramatic negative effect on agricultural productivity. The reallocation of factors to their most efficient use implies agricultural productivity gains for Malawi that are three times larger than those obtained for the manufacturing sector of more developed countries such as China and India relative to the United States in Hsieh and Klenow (2009). Importantly, we provide evidence that the bulk of productivity losses are directly associated with restricted land markets.

Malawi represents an interesting case to study for several reasons. First, Malawi is an extremely poor country in Africa, featuring very low agricultural productivity, a large share of employment in agriculture, and extremely low farm operational scales. Second, the land market in Malawi is largely underdeveloped. Most of the land in Malawi is customary and without a title. In our sample, more than 83% of household farms do not operate any marketed land (either purchased or rented-in). For these households, land is typically granted by a local leader, transmitted by inheritance, or received as

[1] See, for instance, Restuccia et al. (2008).

bride price. The data and our analysis allow us to precisely connect land constraints at the household-farm level and factor misallocation in agriculture, providing an important departure from the existing macro development literature.

We assess the quantitative importance of factor misallocation using detailed household-farm micro data from Malawi. The first key element of the analysis is the measure of real farm TFP which requires detailed information on outputs, inputs, and other factors. The second key element is that we can directly identify from the data which household-farm is operating land acquired from the market and which farms are not. The data is the 2010/11 Integrated Survey of Agriculture (ISA) for Malawi collected by the World Bank and consists of a large nationally-representative sample of more than 12 thousand households.[2] Using the detailed micro data we measure real farm productivity controlling for a wide array of factor inputs, land quality, and transitory shocks. Whereas the dispersion in our measure of farm TFP is quantitatively similar to the findings in previous studies for farm and plant-level TFP in developed and developing countries, our key finding is that factors of production are roughly evenly spread among farmers. That is, the striking fact is that operated land size and capital are essentially unrelated to farm productivity. These patterns imply that the marginal products of capital and land are strongly positively related with farm TFP which, under a standard framework of farm size constitutes evidence of factor misallocation.

To assess the role of misallocation on agricultural productivity in Malawi, we consider as a benchmark the efficient allocation of factors across existing household-farms in the data, taking as given the total amounts of land and capital. We calculate the aggregate output loss as the ratio of actual to efficient output, where the efficient output is the aggregate agricultural output resulting from the efficient allocation of capital and land across farms. Since the total amounts of capital, land, and number of farmers are kept the same under the efficient allocation, the output loss is equivalent to a TFP loss. Our main finding is that the output loss is 0.28 in the full sample. That is, if capital and land are reallocated across farms in Malawi to their efficient uses, agricultural productivity would increase by a factor of 3.6-fold. Importantly, we show that the large agricultural productivity gains are not due to

[2] The ISA data are considered a *gold standard* for the study of poor countries; see a detailed analysis in de Magalhaes and Santaeulàlia-Llopis (2015). We provide further discussion in Section 2.

TFP dispersion — which is similar to that of the manufacturing sector in China and India reported in Hsieh and Klenow (2009) — but to the severe misallocation of factors across household farms in Malawi, which we relate to restricted access to land markets.

A relevant limitation of the empirical literature on factor misallocation is the weak link between misallocation and the policies and institutions that cause it, a caveat that also applies to previous studies of land allocations. This limitation has inspired a promising quantitative literature studying specific policies and institutions, but the findings are yet elusive in explaining the bulk of TFP differences across countries.[3] In this context, an important contribution of our paper is to provide a strong empirical connection between factor misallocation and the limited market for land, showing large productivity losses associated with restricted land markets. The evidence comes from contrasting the output gains among farmers that have no marketed land to those that operate marketed land. Reallocating factors among farms with no marketed land to reduce the dispersion in marginal products to the same extent to those farms with all marketed land increases agricultural output and productivity by a factor of 2.6-fold. Our results provide an important first step in directly identifying (lack of) land markets as an important institution generating misallocation and productivity losses. These results have important implications for the design of policies and institutions to promote a better allocation of factors. We recognize that even with the detailed and excellent micro data from ISA Malawi that we use, measuring precisely farm-level TFP is an enormous difficult task. However, we find it reassuring that the gains from factor reallocation are still large even when comparing groups of farmers in the same economy with and without marketed land from which the potential biases of mis-measurement, unobserved heterogeneity, and randomness are mitigated.

In order to put our evidence of factor misallocation in Malawi and its large negative impact on agricultural productivity in perspective relative to the existing literature, we note that our evidence is closely linked to the seminal work of Hsieh and Klenow (2009) for the manufacturing sector in China and India. Our analysis contributes to this work by providing evidence of misallocation in agriculture in a very poor country that is less subject to concerns of measurement and specification errors that has casted doubt on the extent of misallocation in poor countries. The evidence of misallocation we

[3] See for instance the survey in Restuccia and Rogerson (2013).

provide is strong because of several reasons. First, it is easier to measure productivity in agriculture than in other sectors; essentially output (in Kgs.) and most inputs (e.g., land size from GPS) are easily measured in quantities, see Beegle et al. (2012) and Carletto et al. (2013). Moreover, there is much less product heterogeneity in the agricultural sector than in other sectors, and such heterogeneity can be controlled for in our data by focusing on individual crops. Indeed, the prevalence of maize production in Malawi allows us to corroborate the extent of misallocation for farms that produce the same crop, eliminating artificial TFP dispersion potentially generated from relative price differentials. Second, the micro data provide precise measures of the quality of inputs (e.g., eleven dimensions of land quality) as well as transitory shocks such as rain and health shocks that we explicitly control for to measure farm-level TFP. Third, the micro data provide information on land and capital rental payments that we use to compute land and capital shares for the agricultural sector in Malawi that generate virtually identical results to our benchmark with U.S. values of these shares. Fourth, while Hsieh and Klenow (2009) assess the reallocation gains in China and India relative to the gains in the United States, we are able assess the reallocation gains relative to a set of farmers in Malawi whose operated land is all marketed, providing a more direct benchmark comparison. Whereas the gains from reallocation to the efficient benchmark are 3.6-fold, the gains from reallocation for farmers with no marketed land relative to those farmers with all marketed land are 2.6-fold. To put it differently, the output gains of reducing the dispersion of marginal products among farmers with no marketed land to the dispersion in marginal products of those farmers with all marketed land in Malawi is 160 percent, which represents a gain three times larger than for manufacturing plants in China and India relative to U.S. plants. A relative reallocation gain of 160 percent is substantial and also a lower bound as it is unlikely that farmers with marketed land in Malawi are operating their desired optimal amount of land input given the severity of land market constraints in the country. Finally, the large sample size allows us to show that our results are robust to within narrow geographic areas and other relevant characteristics. There is also a high response rate with very few missing observations, making the analysis less subject to measurement and specification errors. For all these reasons, we argue that our analysis constitutes the most direct and comprehensive evidence of misallocation in a poor country.

Our misallocation results have important macroeconomic implications for structural change. The large productivity impact of misallocation would unravel a substantial process of structural change with broader implications for aggregate outcomes. For instance, in the context of a simple two-sector model, the increase in agricultural TFP in Malawi resulting from reduced misallocation, would reduce the share of employment in agriculture from 65 percent to 4 percent, and would increase average farm size by a factor of 16.2-fold. These effects would be even larger in more elaborate models that include endogenous investments such as that in Goldstein and Udry (2008), ability selection across sectors such as that in Lagakos and Waugh (2013), among other well-studied channels in the literature.

In terms of the inequality implications of reallocation, one reason to rationalize the current land distribution in poor countries is for egalitarian purposes. However, even though the actual allocation of factors is evenly spread among heterogeneous farmers, we find that factor equalization is ineffective at equalizing incomes in Malawi, in addition to its negative effect on agricultural productivity already discussed. Taking the actual allocation of factors as endowments and decentralizing the efficient allocation via perfectly competitive rental markets, we show that the income distribution associated with the efficient allocation features not only much lower overall income inequality, but also that the largest gains are accrued by the poorest farmers. That is, the introduction of rental markets where operational scales can deviate from land ownership substantially improves aggregate productivity, reduces poverty, and alleviates income inequality.

Our paper closely relates to work in macroeconomics studying misallocation in the agricultural sector as first emphasized by Adamopoulos and Restuccia (2014). Adamopoulos and Restuccia (2014) study the potential role of misallocation in the agricultural sector across countries due to policies and institutions that affect farm size. Our analysis differs from this previous work in that, by focusing on a very poor country in Africa, we are able to exploit detailed micro household-farm level data to measure TFP at the farm level, provide direct and compelling evidence of misallocation, and assess the quantitative aggregate productivity impact of factor misallocation in agriculture. Moreover, our analysis is able to connect misallocation to restrictions in land markets, providing an important direct link to policy analysis. Within the macroeconomics literature on misallocation in agriculture our paper also relates to Adamopoulos and Restuccia (2015) who study the impact of a specific policy —land reform—on productivity exploiting panel micro data from the Philippines; and Chen (2015) studying the impact of land titles on agricultural productivity across countries. Our paper is also related to a growing literature in macroeconomics using micro data to study macro development such as Hsieh and Klenow (2009), Gollin et al. (2014), Buera et al. (2014), among many others. There is an important literature in micro development analyzing the role of tenancy and property rights for agricultural productivity such as Shaban (1987), Besley (1995), Banerjee et al. (2002), among others. Whereas this literature has focused on the impact of endogenous investments and effort incentives on farm productivity, our analysis focuses on the impact of resource allocation on aggregate agricultural productivity, taking as given farm-level TFP. Integrating the role of land-market institutions on both factor misallocation and endogenous farm-level productivity, is an important and promising area of study that we leave for future research. Within the micro development literature, closer to our paper are Foster and Rosenzweig (2011) and Udry (1996). Foster and Rosenzweig (2011) provide evidence of inefficient farm sizes in India that may be connected to the low incidence of tenancy and land sales limiting land reallocation to efficient farmers. Udry (1996) focuses on the intra household-farm reallocation of factors across wives and husbands for a relatively small sample of farms, obtaining a small role of misallocation in the agricultural sector. Our results indicate a large role for misallocation. The main difference is that we focus on reallocations across farms instead of within farms. By exploiting large representative data to study nationwide factor reallocation across farms, our analysis delivers, arguably, a more accurate picture of the macroeconomic gains from reallocation.

In the next section, we describe the important elements of the micro data for our analysis. Section 3 constructs our measure of farm TFP and reports facts on resource allocation across farmers in Malawi. In Section 4, we assess the productivity impact of misallocation by comparing aggregate agricultural output in the actual data to the hypothetical efficient allocation of factors. We discuss output losses within geographical areas, institutions, specific skills, as well as across the extent of access to marketed land. We also report some robustness results. Section 5 discusses the broader implications of reallocation for structural change. In Section 6, we explore the income inequality implications of misallocation. We conclude in Section 7. An on-line Appendix is available at: `https://www.dropbox.com/s/je5wv90trv2nmq3/RS_Online_Appendix.pdf?dl=0`.

2 Data

We use a new and unique household-level data set collected by the World Bank, the Malawi Integrated Survey on Agriculture (ISA) 2010/11, see de Magalhaes and Santaeulàlia-Llopis (2015). The survey is comprehensive in the collection of the entire agricultural production (i.e., physical amounts by crop and plot) and the full set of inputs used in all agricultural activities at the plot level, all collected by a new and enlarged agricultural module that distinguish ISA from previous Living Standards Measurement Study (LSMS) surveys. The data are nationally-representative with a sampling frame based on the Census and an original sample that includes 12,271 households (and 56,397 individuals) of which 81% live in rural areas.[4]

The survey provides information on household-farm characteristics over the entire year and we focus our attention to agricultural activities related to the rainy season. The detail on household-farm agricultural production is excruciating. Information on agricultural production is provided by each and all crops produced by the household. This is an economy largely based on maize production that uses 80% of the total land. The total quantity of each crop harvested by each household is available per plot. We value agricultural production using median at-the-gate prices per region. In crop production, each household potentially uses different quantities of intermediate inputs such as fertilizers, herbicides, pesticides and seeds. This information is also provided by plot. We also apply common median prices to value these intermediate inputs. As a result, our benchmark measure of household-farm output is a common-price measure of real value added constructed as the value of agricultural production (of all crops) minus the costs of the full set of intermediate inputs.[5]

We measure household land as the sum of the size of each cultivated household plot. This includes rented-in land, which consists of 12.5% of all cultivated land. On average, household farms cultivate

[4]The Malawi ISA is part of a new initiative funded by the Bill & Melinda Gates Foundation (BMGF) and led by the Living Standards Measurement Study (LSMS) Team in the Development Research Group (DECRG) of the World Bank. For further details on the Malawi ISA data and the construction of agricultural output and inputs see Appendix A.

[5]We use common prices of output and intermediate inputs to construct real value added in the same spirit as the real measures of output from the International Comparison Project and the Penn World Table.

1.8 plots.[6] Plot size is recorded in acres using GPS (with precision of 1% of an acre) for 98% of plots (for the remaining 2% of plots, size is estimated). The operational scale of farms is extremely small. For each household, we compute the amount of land used for agricultural production regardless of the land status (whether land is owned, rented, etc.). Hence, we focus on the operational scale of the household-farm. We find that 78.3% of households operate less than 1 hectare (henceforth, Ha.), 96.1% of households operate less than 2 Ha., and only 0.3% of households operate more than 5 Ha., see the first column in Table 1. The average farm size is 0.83 Ha.[7] The data contains very detailed information on the quality of land for each plot used in every household. There are 11 dimensions of land quality reported: elevation, slope, erosion, soil quality, nutrient availability, nutrient retention capacity, rooting conditions, oxygen availability to roots, excess salts, topicality, and workability. This allows us to control for land quality to measure household-farm productivity.

Regarding land, we emphasize that the land market is largely underdeveloped in Malawi. The proportion of household-farms that do not operate any marketed land is 83.4%. These are households whose land was granted by a village chief, was inherited or was given as bride price. The remaining 16.6% of farm households operate some land obtained from the market, either rented or purchased, and the proportion of household-farms whose entire operated land was obtained in the market is 10.4%. Disaggregating the main types of marketed land, we find that 3.0% of household-farms rent-in land informally (e.g., land borrowed for free or moved in without permission), 9.5% rent-in land formally (e.g., leaseholds, short-term rentals or farming as a tenant), 1.8% purchase land without a title and 1.3% purchase land with a title.

In terms of agricultural capital, we have information on both equipment and structures. Capital equipment includes implements (hand hoe, slasher, axe, sprayer, panga knife, sickle, treadle pump,

[6] In rural Malawi, land is the largest household asset representing 44% of household total wealth. House structure is 30%, livestock is 13%, and agricultural equipment and structures (e.g. tools and barns) is 3%, see de Magalhaes and Santaeulàlia-Llopis (2015).

[7] To make a comparison of operational scales in Malawi with other countries, we report the distribution of farm sizes from the World Census of Agriculture in Adamopoulos and Restuccia (2014). These data include all land used for agricultural production and corresponds to the year 1990. Note that despite the 20 years difference between ISA 2010/11 and the Census 1990 the distribution of land across the two sources is very similar. Comparatively with other countries, the operational scale of farms in Malawi is extremely small, ∼0.70-0.83 Ha., whereas average farm size is 187 Ha. in the United States and 16.1 Ha. in Belgium. Belgium is a good reference for a developed country since the land endowment (measured as land per capita) is similar to that of Malawi (land per capita is 0.56 Ha. in Malawi and 0.5 Ha. in Belgium, whereas land per capita is 1.51 Ha. in the United States).

watering can, and so on) and machinery (e.g. ox cart, ox plough, tractor, tractor plough, ridger, cultivator, generator, motorized pump, grain mill, and so on), while capital structures includes chicken houses, livestock kraals, poultry kraals, storage houses, granaries, barns, pig sties, among others. To proxy for capital services after conditioning for its use in agricultural activities, we aggregate across the capital items evaluated at the estimated current selling price.[8]

In Malawi, a large proportion of the households members, beside the household head, contribute to agricultural work. Household size is 4.6 with extended families in which several generations live together in a single household. We use the survey definition of household members as individuals that have lived in the household at least 9 months in the last 12 months. In terms hours, data are collected at the plot level for each individual that participates in agriculture and by agricultural activity (i.e., land preparation/planting, weeding/fertilizing, and harvesting). The data provides information of weeks, days per week, and hours per day employed per plot, activity and individual. To compute household-farm hours we aggregate the hours of all plots, activities and individuals. Further, the same information is provided for hired labor and labor received in exchange (for free), but most household-farm hours consist of family hours. We add hours by hired labor and free exchange of labor to our measure of total household hours.

Since our data comprises a single cross section of households for 2010-11, it is important to control for temporary output shocks that may explain variation in output and hence productivity across households in the data. The single most important temporary shock for farmers is weather. We use the annual precipitation which is total rainfall in millimetres (mm) in the last 12 months. In further robustness exercises, we net household-productivity from additional transitory shocks in the form of health, deaths or food security risks suffered by the household in the last 12 months.

Geographic and institutional characteristics are also recorded for each household-farm. We use several partitions of these characteristics to conduct robustness exercises. In particular, we use geographi-

[8]This selling price for agricultural capital items, rarely available in previous LSMS data, helps capture potential differences in the quality and depreciation of capital across farms. We provide robustness of our results to alternative physical measures of agricultural capital such as asset indexes for agricultural equipment and structures in Appendix B. Our results are robust to using these alternative indexes to proxy for agricultural capital.

cal information on the region, districts, and enumeration area to which household-farms belong and institutional characteristics such as the traditional authority (TA) governing the household-farm or ethno-linguistic characteristics. TAs are relevant for our exercise as chiefs appointed by TAs perform a variety of functions that include resolving issues related to land and property.[9]

Finally, we note that the survey response is very high with very few missing observations. Conditioning on households that produce agricultural output and for which all factor inputs, including the 11 dimensions on land quality, are available and further trimming about 1% of the household-farm productivity distribution, our sample consists of 7,157 households.[10]

To summarize, the data allows us to obtain precise measures of real household-farm productivity. The ISA data represents a substantial improvement with respect to previous LSMS questionnaires. The detailed information on quantity inputs and outputs reduces substantially the possibility of measurement error and composition bias, making the dataset ideal for our purpose of measuring productivity at the farm level, assessing the extent of factor misallocation in the Malawian economy, and assessing the extent to which productivity losses due to misallocation are related to imperfections or frictions in the land and other input markets.

[9]Traditional Authorities are part of the administrative structure of Malawi. Any one district is only in one region, any one TA is only in one district. Traditional authorities (and their sub-ordinate chiefs) are 'town chiefs' formally recognized under the Chiefs Act (1967) and receive a monetary honorarium by the government. The Chiefs Act states that chieftaincies are hereditary and hierarchical, being the highest level in the TA or the Paramount Chief (PC) and TAs cover the entire country. There can be several TAs within each ethno-linguistic group. The sub-structure is such that below each TA there are Sub-TAs, Group Village Headmen (GVH), and Village Headmen (VH). All villages have a VH, and several villages will be grouped under one GVH. These chiefs perform a variety of functions: cultural affairs, administration and management of various sorts, oversight of issues related to land and property, resolving disputes, an involvement in politics, and promoting economic and social development. The TAs that we use in our reallocation exercise are the ones provided by Malawi ISA 2010/11 which are a subset of those used by the National Statistical Office to organize the data collection of the Malawi Census. There are about 380 populated TAs and wards in the country with an average population of just under 26,000 persons; TAs are found in rural areas, while urban-equivalent administrative wards are in the four urban centers of Blantyre, Lilongwe, Mzuzu, and Zomba. District and local government authorities may recognize additional TAs or sub-chiefs used in order to create reasonably sized administrative units within large TAs.

[10]See further details on the trimming strategy in Appendix B.

3 Measuring Farm Productivity

We use the micro data from ISA 2010/11 described in the previous section to measure productivity at the farm level. Constructing a measure of farm total factor productivity (TFP) is essential in assessing the extent to which factors are misallocated in the agricultural sector. The detailed micro data for Malawi presents a unique opportunity to assess factor misallocation in agriculture and its aggregate productivity implications.

We measure farm productivity by exploiting the detailed micro data where not only we obtain real measures of output and value added in each farm but also control for land quality and a wide array of transitory shocks. We measure farm-level total factor productivity (TFP) s_i as the residual from the following farm-level production function,

$$y_i = s_i \zeta_i k_i^{\theta_k} (q_i l_i)^{\theta_l},$$

where y_i is real value added, k_i is capital, l_i is the amount of land in operation, ζ_i is a rain shock, q_i is land quality, and $\theta_{k,l}$ are the input elasticities. In our analysis, we focus on the allocation of capital and land across farms. As a result we abstract from differences in the labor input other than the productivity of the farmer. For this reason, the measures of value added, capital, and land are in per hour. We note, however, that hours in farms are uncorrelated with farm TFP and, as a result, our characterization and analysis are unaffected by treating hours as an explicit input in the farm production function. In constructing our measure of farm TFP, we choose $\theta_k = 0.36$ and $\theta_l = 0.18$ from the capital and land income shares in U.S. agriculture reported in Valentinyi and Herrendorf (2008). We later discuss the sensitivity of our results to these factor shares using our micro data for Malawi.

Our measure of output is real value added which takes into account the real amount of intermediate inputs used in production. This is relevant in Malawi because intermediate inputs are subsidized via the "Malawi Input Subsidy Program" and the subsidy allocation is based on farmer's income, with poorer farmers receiving higher subsidies. According to our preferred measure of productivity, less productive farmers, which are also poorer farmers, receive a subsidy that is more than 60 percent their output, whereas more productive farmers (richer farmers) receive a subsidy that is less than 10 percent their output. This implies that a measure of productivity that uses actual expenditures in intermediate inputs instead of actual quantities of intermediate inputs substantially underestimates productivity dispersion across farms. In our data, not accounting for distortions in intermediate input prices underestimates the dispersion in farm TFP by 23 percent.

We are interested in permanent measures of productivity at the farm level, associated with the productivity of the farm operator. As a result, two key elements for this measure are to distinguish between the productivity of the farmer and the productivity of the land under operation and to abstract from temporary variations in output due to whether shocks. We deal with each of these components in turn.

There are important differences in the land characteristics of farms in our sample. For the full sample, more than 34 percent of land is high-altitude plains while around 20 percent are low plateaus and 19 percent mid-altitude plateaus. These characteristics also differ by region where the Center region is mostly high-altitude plains whereas the South region is mostly low plateaus. We control for land quality in our measure of productivity by constructing an index as follows. Our benchmark land quality index is q_i^0 is constructed by regressing log output in each farm on the full set of land quality dimensions described in Section 2.[11] Figure 1, panel (a), reports the distribution of the land quality index q across households in the full sample and across regions. There is substantial variation in q across households and across regions. In Table 2 we report the dispersion (variance of the log) of land quality indexes versus the dispersion in the quantity of land across households in our data. We find that the dispersion in land quality is large and slightly above that of land size nationwide, respectively .859 and .749 in terms of the variance of the log. Not surprisingly, the dispersion in land quality decreases with the size of geographic area and the average dispersion is .833 within regions, .568 within districts and .147 within enumeration areas.

[11] We have also considered alternative land quality indexes for subsets of the land quality dimensions and with different forms of land size and capital controls in Appendix A.

In Malawi, weather shocks, in particular rain, is the single most important source of transitory shocks to agricultural production—we find that most cropland is rain-fed and 84% of household-farms do not have any alternative irrigation system. Figure 1, panel (b), reports the density of rain for the entire population of farmers and, separately, for farmers within regions. We note what appear substantial differences across household-farms and regions with median values for the Center and South region that fall below the range of of the distribution of rain in the North region. To quantify the actual dispersion in rain we report the variance of logs of rain in the bottom of Table 2 for our entire sample and regions. We find that the dispersion of rain is very small compared to the dispersion in land size and even smaller relative to output. In particular, the dispersion in annual precipitation is 5.2% that of land size and 2.1% that of agricultural output. The small role of rain variation further appears within regions, districts, and enumeration areas, as well as using alternative definitions of precipitation measures, with mean cross-sectional variances that decline with the size of the geographical area. In short, rain is not a substantial contributor to output variation across farm households in Malawi.

Having constructed our measure of farm TFP, Figure 2 documents its distribution across all households in our sample. There are large differences in farm productivity across households that remain even within regions. The dispersion in farm productivity compares in magnitude to that of the variation in physical productivity (TFPQ) across plants in the manufacturing sector in the United States, China, and India reported in Hsieh and Klenow (2009), see Table 3. Whereas the ratio of physical productivity between the 90 to 10 percentile is a factor of around 15-fold in China and India and around 9-fold in U.S. manufacturing, the 90-10 ratio across farms in our sample is 10.8-fold. Similarly, the 75-25 ratio is 3.2-fold in our sample of farms whereas it is around 4.5-fold across manufacturing plants in China and India.

Table 4 reports the variance decomposition of farm output per hour using our assumed production function. We note that farm productivity s_i and to a lesser extent inputs of capital and land are the key determinants of output variation across farm households in Malawi, with rain and land quality playing a quantitatively minor role. Our conclusion of a small quantitative role of land quality for output differences across household farms is robust to variations in the index. We discuss these variations in

Appendix A.

Figures 3 and 4 report several variables of interest across farms by farm-level TFP s_i. The patterns are striking. The data show that the actual allocations of land and capital in farms are unrelated to farm productivity. Figure 3, panel (a), shows the amount of land in each farm by farm productivity using our benchmark measure of productivity that adjusts for rain and quality as described previously. Land in farms is essentially uncorrelated to farm TFP, the correlation between land size and productivity is 0.05. Figure 3, panel (b), documents the relationship between the amount of capital in farms by farm productivity. Capital and productivity across farms are essentially unrelated, the correlation between the two variables is -0.01. Figure 3, panel (c), shows that the capital to land ratio is essentially unrelated to farm productivity, with a correlation between these variables of -0.03. This finding suggests that larger farms do use more capital, but that the capital to land ratio is unrelated to farm TFP. Our findings on the allocations of capital and land in farms have implications for the the pattern of capital and land productivity across farms. Figure 4, panels (a) and (b), show that the marginal product of land (proportional to the yield or output per unit of land) and capital (proportional to output per unit of capital) are strongly positively correlated with farm productivity. The correlation between these measures of factor productivity and farm TFP are .77 and .76 in each case. These patterns suggest substantial gains of factor reallocation across farms.

4 Quantitative Analysis

We assess the extent of factor misallocation across farms in Malawi and its quantitative impact on agricultural productivity. We do so without imposing any additional structure other than the farm-level production function assumed in the construction of our measure of household-farm productivity. We then study the connection between factor misallocation and land markets.

4.1 Efficient Allocation

As a benchmark reference, we characterize the efficient allocation of capital and land across a fixed set of heterogeneous farmers s_i. A planner chooses the allocation of capital and land across a given set of farmers with productivity s_i to maximize agricultural output given fixed total amounts of capital K and land L. The planner solves the following problem:

$$Y^e = \max_{\{k_i, l_i\}} \sum_i s_i (k_i^\alpha l_i^{1-\alpha})^\gamma,$$

subject to

$$K = \sum_i k_i, \qquad L = \sum_i l_i.$$

The efficient allocation equates marginal products of capital and land across farms and has a simple form. Letting $z_i \equiv s_i^{1/(1-\gamma)}$, the efficient allocations are given by simple shares of a measure of productivity $(z_i / \sum z_i)$ of capital and land:

$$k_i^e = \frac{z_i}{\sum z_i} K, \qquad l_i^e = \frac{z_i}{\sum z_i} L.$$

We note for further reference, that substituting the efficient allocation of capital an land into the definition of aggregate agricultural output renders a simple constant returns to scale aggregate production function for agriculture on capital, land, and agricultural farmers given by

$$Y^e = Z N_a^{1-\gamma} [K^\alpha L^{1-\alpha}]^\gamma, \tag{1}$$

where $Z = (\sum z_i g(z_i))^{1-\gamma}$ is total factor productivity as the average productivity of farmers and N_a is the number of farms. For our benchmark, we choose α and γ consistent with the production function used in our measure of farm-level productivity so that $\alpha\gamma = 0.36$ and $(1-\alpha)\gamma = 0.18$ are the capital and land income shares in the U.S. economy (Valentinyi and Herrendorf, 2008) — and we do sensitivity with factor shares directly computed from our Malawi micro data (see section 4.5). This implies $\gamma = 0.54$. The total amount of capital K and land L are the total amounts of capital and land across the farmers in the data. Farm level productivity s_i's are given by our measure of farm TFP from data as described previously.

We illustrate the extent of factor misallocation in Figure 5, where we contrast the actual allocation of land and capital and the associated factor productivity by farm TFP against the efficient allocation of factors and the associated factor productivities across farms. In the efficient allocation, operational scales of land and capital are strongly increasing in farm productivity so that factor productivity is constant across farms. The efficient allocation contrasts sharply with the actual allocation in Malawi.

To summarize, the actual allocation of land across farmers in Malawi is unrelated to farm productivity which is consistent with our characterization of the land market where the amount of land in farms is more closely related to inheritance norms and redistribution and access to land is severely restricted in rental and sale markets so more productive farmers cannot grow their size. Capital is also unrelated to farm productivity with the capital to land ratio being roughly constant across farm productivity. Our interpretation of this fact is that restrictions to land markets are also affecting capital allocations echoing de Soto (2000) findings that land market restrictions and insecure property rights of farmers limit the ability to raise capital for agricultural production.[12] Our findings constitute strong evidence of land and capital misallocation across farmers in Malawi.

4.2 Aggregate Output Loss

To summarize the impact of misallocation on productivity, in what follows we report the aggregate output loss defined as

$$\frac{Y^a}{Y^e} = \frac{\sum y_i^a}{Y^e},$$

where Y^a is actual agricultural output aggregated from farm-level output using $y_i = s_i \left(k_i^{\alpha} l_i^{1-\alpha}\right)^{\gamma}$, where k and l are actual allocations and; and Y^e is efficient aggregate agricultural output as defined previously. Because the efficient allocation takes as given the total amounts of capital, land, and the

[12]See the related discussion in Besley and Ghatak (2010). We find further direct evidence of the relationship between land markets and access to credit in our micro data for Malawi, see Appendix C for a discussion.

number of farmers observed in the data, the output loss is also a TFP loss.[13]

Table 5, panel (a), reports the main results. For the full sample in the entire Malawi, the output loss is 0.28, that is actual agricultural output is 28 percent of that of efficient agricultural output at the aggregate level. We can also express the results in terms of the gains from reallocation from the actual distribution of factors to the efficient, which is the inverse of the output loss. If capital and land were reallocated efficiently in Malawi to maximize agricultural output, output and hence productivity would increase by a 3.6-fold factor. This is a very large increase in productivity as a result of a reduction in misallocation, at least compared to the results in the misallocation literature when evaluating specific policies which has found increases on the order of 5 to 30 percent. Even when eliminating all wedges in manufacturing in China and India in Hsieh and Klenow (2009), the increases range between 100-160 percent. Given that the real productivity dispersion across farms in Malawi is similar to the productivity dispersion of manufacturing plants reported in Hsieh and Klenow (2009), what this result suggests is that the actual allocation of resources in Malawi is much more distorted than that of those other countries or to put it differently that factors are more severely misallocated in Malawi.[14] Because we have a large sample, our mean output loss is tightly estimated. In Table 5, panel (a), we also report the 5 and 95 percentiles of Bootstrap estimates, which provide a narrow interval between 0.25 and 0.32 for the aggregate output loss.

Factor inputs are severely misallocated in Malawi, implying the large aggregate productivity losses just discussed. There are two features of factor misallocation, factor inputs are dispersed among farmers with similar productivity (misallocation in factor inputs within s_i productivity types) and factor inputs are misallocated across farmers with different productivity (which lowers the correlation of factor inputs with farm productivity). We argue that factor input variation within a farm productivity type is not due to measurement error as factor inputs such as land are measured with tight precision.

[13]The computation of actual output at the farm level abstracts from rain and land quality as it is the case in the efficient allocation. We also emphasize that the empirical measure of farm productivity s_i is purged of rain and land quality effects as described in Appendix A.

[14]Note that a more direct comparison with the analysis of Hsieh and Klenow (2009) for the case of China and India would be to report dispersion in log revenue productivity (TFPR). However, unlike for the case of manufacturing firms, most of the production of farms in Malawi is not sold in markets and hence revenue of farms would have to be imputed. More importantly, the micro data allow us to directly measure real productivity across farms and hence assess the aggregate productivity implications without the need to characterize the associated market wedges.

Nevertheless, we can assess the magnitude of the aggregate output loss associated with lower dispersion in factor inputs within a productivity type. We remove within-s variation in factor inputs by regressing separately log land and capital on a constant and log farm productivity s and use the estimated relationship to construct measures of factor inputs that partially or fully remove residual variation. The results are in Table 5, panel (b). At the extreme, with no within-s variation of factor inputs, there is only gain from reallocation across productivity types. Even in this case, the output loss is still substantial .40 (versus .28 in the baseline) which implies a substantial reallocation gain of 2.5-fold. To put it differently, our finding is that almost 70 percent of the output loss is due to misallocation of factor inputs across farmers with different productivity, whereas the other 30 percent owes to misallocation in factor inputs of farms within the same productivity type.

We also report the output loss for narrower definitions of geographical areas and institutions such as language classifications. Table 5, panel (c), reports these results. In the fist column, we report the output loss for the average of regions (i.e., output loss of reallocating factors within a region for each region and then averaged across regions), for the average across districts (a narrower geographical definition than region), for the average across enumeration areas. Enumeration areas are a survey-specific geographical description and is the narrower definition available. Each enumeration area amounts to about 16 households in the data, which given the average land holdings amounts to about 30 acres of geographically connected land. The table reports results by Traditional Authority (TA) and language. Columns 2 to 4 report the Median, Min, and Max in each case.

What is striking about the numbers in Table 5, panel (c), is that the gains from reallocation are large in all cases, even within narrowly defined geographical areas, traditional authority and language. In particular, reallocating capital and land within a region generates output losses of the same magnitude as in the aggregate (mean output loss of 0.29).[15] Similarly, within districts, the average output loss is 0.34, but the output losses can be as large as 0.14 and the lowest 0.48. Even for the narrowest geographical definition, enumeration areas, just reallocating factors within 16 households, average output losses are 0.60 and can be as large as 0.09 in some areas. Considering reallocation within

[15] The idea is that narrower geographic definitions provide additional land quality control, see Larson et al. (2014) for a similar strategy.

enumeration areas also serves as a further check on the relevance of land quality since within such a narrow geographical definition it is unlikely that land quality plays much if any role. We also report the distribution of output losses in each enumeration area in the survey in Figure 6. The output loss within traditional authorities is 0.32 in average, with the largest 0.18 and median 0.29. Within languages, the average output loss is 0.30, the largest 0.17 and median 0.26.

4.3 The Role of Land Markets

We have found strong evidence that capital and land are severely misallocated in the agricultural sector in Malawi. We connect factor misallocation directly to the limited market for land in Malawi. To do so, we use plot-level information about how each plot was acquired and group household-farms by the share of marketed land that they operate. We find that 83.4 percent of all household farms operate only non-marketed land and the remaining 16.6 percent operate some marketed land, with 10.4 percent operating exclusively marketed land. Using this classification of household farms, we explicitly assess the output loss across farms that differ in the extent of marketed land in their operation.

Table 6 reports the aggregate output loss for farms that operate with no marketed land, with some marketed land as either rented in or purchased, and with only marketed land. The output loss is much higher for those farms with no marketed land than with some marketed land. For instance, the output loss for those farms with no marketed land, which comprises 83.4 percent of the sample of household farms, is 0.24, slightly larger than the 0.28 output loss for the entire sample. But for farms with some marketed land, which comprises the remaining 16.6 percent of the sample farms, the output loss is 0.51. The striking finding is that the output loss of factor misallocation is reduced by more than half when farms have access to some marketed land. The output loss is even smaller among the set of farmers whose entire operated land is either rented in or purchased. These farmers comprise 10.4 percent of the sample and have an output loss of 0.64 versus 0.24 for farmers with no marketed land. These results provide direct and solid evidence of the connection of misallocation to land markets.

Table 6 also decomposes the output losses among farmers by the type of operated marketed land: (a) rented-in informally, for example land borrowed for free or moved-in without permission; (b) rented-in formally such as leaseholds, short-term rentals or farming as tenant; (c) purchased as untitled; and (d) purchased as titled. The vast majority of farmers with some marketed land are formally renting-in, 9.5 percent of the sample out of the 17 percent with some marketed land. The output loss is fairly similar for farmers operating land rented-in formally or informally, an output loss of roughly 0.58 for both. The lowest output loss is recorded for farms with operated land that was purchased with a title, 0.72 versus 0.24 for farms with no marketed land. Interestingly, farmers with purchased land without a title record a fairly large output loss, 0.20, somewhat larger than the output loss for farmers without marketed land of 0.24. We note however that the sample of farms with purchased land is fairly limited, only 3.1 percent of the sample: 1.8 percent of farms with purchased untitled land and 1.3 percent of farms with purchased titled land.

Even though we have showed a strong connection between land markets and the degree of output loss generated by misallocation, farms with marketed land are still far from operating at their efficient scale which suggests that land markets are still limited even for farms with access to marketed land. For instance, the correlation between land and farm TFP is .14 for farms with no marketed land, .25 for farms with some marketed land, and .30 for farms that operate all marketed land. Having access to marketed land implies that farmers can command more inputs and produce more output. Also, operating marketed land is associated with greater access to other markets (e.g., credit) and with other indicators of economic development. Farmers with marketed land are substantially more educated than farmers with no marketed land, women in these farms are more empowered in terms of labor force participation and market wages, more of these farmers are migrants, a fewer proportion live in rural areas, and a larger proportion of these farmers invest in intermediate inputs and technology adoption. We report these dimensions of farm differences by the type of operated land in Appendix C. One potential concern with our characterization is the possibility that access to marketed land is driven by farm productivity. We find however very small differences in productivity between farms with and without marketed land. These differences are too small compared to what it would be if access to marketed land was purely driven by frictionless selection in farm TFP. Even though on average farms with some marketed land are 25 percent more productive than farms with no marketed land, the distribution of household-farm productivity is very similar between the group of household-farms with marketed land compared to the group of farms without marketed land, with respective variances of 1.19 and 1.17. Figure 7 shows the close overlap of these distributions. Nevertheless, we can assess the gains from reallocation holding constant farm-level TFP. To do so, we regress household-farm output gains from reallocation on household-farm productivity and a dummy variable that controls for whether the farm operates marketed land. Specifically, we estimate using OLS the following relationship: $\ln \frac{y_i^e}{y_i^a} = cons + \psi_1 \ln s_i + \psi_2 \mathbf{1}_{market}$, where ψ_2 is captures the effect of marketed land on the farm output gain controlling for farm productivity s_i. We use actual output as weights so as to reproduce the output gains of reallocation in our nationwide benchmark. We find that $\psi_2 = -.39$, that is, operating marketed land decreases the output gain by 39 percent. Considering that the total output gain from farms without marketed land compared to farms with marketed land is $(1 - .2411/.5081) \times 100 = 52$ percent, this implies that operating marketed land alone accounts for $(39/52) \times 100 = 75$ percent of the total output gains.

Finally, to gain further perspective of the productivity gains from reallocation, we note that Hsieh and Klenow (2009) report the gains from reallocation for the manufacturing sectors in China and India compared with those in the United States. The idea is that the U.S. data is not necessarily absent of frictions, measurement and specification errors but that more misallocation in China and India than in the United States is indicative of more distortions. Our decomposition of the degree of misallocation across groups of households farms with or without marketed land provides a more direct reference for the gains of reallocation, that is, reallocation gains attained by household farms without marketed land relative to farms operating marketed land. Taking advantage of the fact that we can identify which household farms have marketed land, we restrict the gains of reallocation of household-farms that do not operate marketed land to be at most the gains attained by households farms that operate marketed land. We find that output (productivity) could increase by a factor of 2.6-fold if factor misallocation is reduced among farms with no marketed land to the same extent as those farms operating only marketed land. This implies a reallocation gain that is three times the one obtained in Hsieh and Klenow (2009) for the manufacturing sector of China and India relative to the United States. Further, this gain may be a lower bound, as it is likely that the U.S. farming sector allocates inputs more efficiently than the sector of household farms in Malawi that operates marketed land. Nevertheless, the fact that we use a group of household farms within Malawi as benchmark reference mitigates the effects of differential institutional features and frictions (beyond land markets access) that are more prevalent when using the United States or another country as reference in comparison with Malawi.

4.4 Farm Size vs. Farm Productivity

A common theme in the development literature is the view that small farms are more efficient than large farms. This view stems from a well-documented inverse relationship between yields (output per unit of land) and farm size, e.g. Berry and Cline (1979). An important concern for the inverse yield-to-size relationship is whether the yield by farm size is a good proxy for farm productivity. Our data for Malawi provide an opportunity to contrast this view of farm size and productivity with a direct measure of farm TFP.

Figure 8 documents the relationship between land productivity (yield) by farm size in our data. Land productivity is slightly declining with farm size, which conforms with the finding in the inverse yield-to-size literature. For instance, the land productivity of the largest 10 percent of farms relative to the smallest 10 percent of farms is 0.34 (a 3-fold yield gap). This characterization contrasts sharply with our earlier documentation that land productivity is strongly increasing with farm TFP, see Figure 4 panel (a): land productivity of the 10 percent most productive farms relative to the 10 percent least productive farms is a factor of 58-fold.

Whereas the pattern of land productivity by size suggests reallocation of land across farm sizes—from large to small farms— the pattern of land productivity by farm TFP suggests reallocation across farmers with different productivity—from less productive to more productive farmers. We find that the aggregate output and productivity gains of these two forms of reallocation are dramatically different. As discussed earlier, reallocating factors across farms with different productivity can increase agricultural output and TFP by a factor of 3.6-fold. In contrast, reallocating factors across farms by size, taking the yield-by-size as a measure of productivity, accrues a gain in output of only 26 percent. And this gain may be an upper bound as the implementation of reallocation by size may lead to further amounts of misallocation. For instance, land reforms are often pursued to implement reallocation of land from large to small holders, that are also accompanied by further restrictions in land markets in order to favor redistribution to smallholders and landless individuals. This redistribution can lead to productivity losses instead of gains as it was the case in the comprehensive land reform in the Philippines, see for instance Adamopoulos and Restuccia (2015).

It is also interesting to directly contrast the yield as a measure of farm productivity commonly used in studies of agricultural development, e.g., Binswanger et al. (1995) with our actual measure of farm TFP. In Malawi, the yield and farm TFP are highly correlated only because factors are extremely misallocated. In the efficient allocation, land productivity and farm TFP are uncorrelated. This implies that changes in policies and institutions that allow a better allocation of land would tend to, other things equal, reduce land productivity of the expanding farms and increase land productivity of the contracting farms, potentially masking the productivity benefits of the reforms.

To summarize, because farm size and farm TFP are essentially uncorrelated in Malawi, it is misleading to characterize farm-level observations and diagnose policy recommendations based on the more common measure of farm size—i.e., size is not a good proxy for productivity. This result emphasizes the role of policies and institutions driving a wedge between size and productivity at the establishment level in developing countries.

4.5 Further Robustness Results

We conduct further robustness results to alternative factor shares, specific crop type production, human capital and specific skills, and transitory health and other shocks.

Factor Shares From Malawi Micro Data It is important to note that by setting our benchmark agricultural factor shares to those in the United States we aim at avoiding that the shape of our production function inherit land and capital market distortions that we are studying for Malawi. However, the U.S. agricultural sector is not necessarily absent of frictions which, in turn, might differ from those in Malawi. To address this issue, we use household-farm rented-in land and capital payments reported in our micro data to compute, respectively, the land and capital share of income for the agricultural sector in Malawi. Note that this can only be done from household-farms that rent-in inputs. This implies recognizing as efficient benchmark production the one attained by the household-farms that are less subject to market frictions in Malawi. Then, we use these new measures of factor shares, i.e., from actual rental markets in Malawi, to conduct our main reallocation exercise.

Our results are reported in Table 7. Our benchmark results that use the U.S. values of the agricultural sector, are reproduced in the first column. The second and third columns provide two alternatives. In the second column, Full Sample, we measure factor shares of output using the average rental rates of land and capital. We compute these rates as the ratio of factor rental payments to the factor stocks from the sample of farms that rent-in all of their capital and land. Then we use these rental rates to impute the land and capital rental income for each and all farms in the entire sample. This implies an average capital share of income of .190 and an average land share of income of .391. Withe these factor shares, the output loss in the full sample is .2799 which is not substantially different from our benchmark, .2788. In the third column, Renting Sample, we assume the land and capital shares of output are the averages from the renting sample only and that we then apply to all farms in the sample. In this case, factor shares are lower which imply a smaller but still substantial output loss of .329. In all, we find that our results do not depend on the size of the decreasing returns in the U.S. economy, as we get similar results if we use the returns from the households that have access to rental markets in Malawi.

Crop Type Production In our sample, close to 80% of land is devoted to maize production, see also FAO (2013).[16] Since optimal farm operational scale may differ across crop types, we investigate

[16] Maize, cassava, and potatoes are the main crops produced in Malawi in terms of volume with roughly equal amount in tonnes, but it is maize that accumulates the vast majority of cropland. The importance of maize further appears in terms of the household diet in Malawi where about 50% the average daily calorie intake is obtained from maize, 8.4% from potatoes, and 5.8% from cassava, see FAO (2013).

whether the output losses are different across farms producing different crops. Since most farms produce both maize and non-maize crops, we consider reallocation among farms that produce mostly maize and those farms that produce mostly non-maize. We report the results in Table 8 where the first column is the output losses in the nationwide benchmark, the next two columns are farms that produce mostly maize (whether the maize share is positive or above the median of all farms), and the last two columns for those farms that produce mostly non-maize (wether the non-maize share is positive or above the median). The output loss among farms that produce maize is somewhat smaller (0.37 for farms with maize production above the median versus 0.28 in the benchmark), which implies that the farms with non-maize production have output losses that are larger (0.24 for farms that produce non-maize above the median versus 0.28 in the benchmark). We conclude that crop composition and the dependence of maize production in Malawi are not driving our main output-loss results.

Human Capital and Specific Skills Table 9 reports robustness results with respect to human capital and specific terrain skills. In each case, we perform the efficient reallocation exercise discussed earlier but within each education and skill type. With respect to schooling levels of the household head, the output losses are large at all schooling levels and roughly constant across the human capital spectrum. For no schooling farmers the output loss is 0.31 and between 0.26 to 0.28 for farmers with primary and more than primary schooling. Output losses may also be related to specific skills such as how to operate farms in different terrain roughness. Similarly to our findings for human capital, the output losses are large even within specific terrain roughness, 0.27 for high altitude plains and 0.36 for mid-altitude mountains.

Other Transitory Output Variation Table 10 reports the output losses associated with a measure of farmer productivity that further adjusts for transitory health and other shocks. We regress our benchmark measure of farm productivity on a set of controls for transitory shocks which includes health and death shocks, food security risk, marital status, distance to markets and the availability of other income sources. In each case, we use the residual from each of these regressions as the new measure of farm productivity. Using these alternative measures of productivity, we perform the efficient reallocation exercise discussed earlier. Output losses are large in all cases, from 0.28 in the benchmark specification to 0.27 when using all controls.

Permanent Heterogeneity in Farm TFP The ISA conducted a second wave of interviews for roughly one-fourth of the original sample in 2013. That is, there is a panel for almost 3,000 household farms in Malawi for 2010 and 2013. This provides us with the opportunity to recover for each household an alternative measure of permanent productivity defined as the estimated fixed effect of individual farm TFP across waves. We find that the fixed effects from the panel roughly capture 87 percent of the original cross-sectional TFP variation. Importantly, this new measure of TFP based on fixed effects is, again, essentially unrelated to land and capital. Not surprisingly, the efficient reallocation in 2010 increases aggregate agricultural productivity by a factor of 3.1-fold. That is, our result for large output gains is robust to alternative measures of productivity that exploit the panel component of the data.

5 Implications for Structural Change

We assess the broader implications of reallocation by considering the effects of increased productivity in agriculture on the movement of factors across sectors. We argue that a TFP increase of a 3.6-fold factor in the agricultural sector would produce a process of substantial structural change in the Malawian economy. This reallocation would produce broad impacts in the economy through well-known features such as the potential selection effects associated with the movement of labor from agriculture to non-agriculture and dynamic investment effects such as additional investments farmers would make to exploit increased farm size, capital and human capital accumulation, among many others.

To provide a simple characterization of these broader implications of reallocation, we consider an extension of the previous analysis to allow for a non-agricultural sector.[17] Recall that the aggregate

[17]For a more elaborate analysis of misallocation in agriculture in the context of a two-sector economy see Adamopoulos

production function for agriculture in the efficient allocation is given by equation (1). To simplify the analysis, without much loss of generality, we abstract from capital by assuming that $\alpha = 0$. Hence, the aggregate production function in agriculture is given by:

$$Y_a = ZL^{\gamma} N_a^{1-\gamma},$$

where N_a is the fraction of employment in agriculture and $\gamma = 0.54$. The non agricultural production function is given by $Y_n = A(1 - N_a)$. We assume that preferences are such that consumers not only have a minimum consumption of agricultural goods $\bar{a}$, but also this level is a satiation point so per capita consumption of agricultural goods is given by $\bar{a}$ and any income above the one required for this amount of agricultural consumption is spent on non-agricultural goods.[18] We continue to consider a benevolent social planner that chooses the allocation of labor across sectors to maximize consumer's welfare. Given our assumptions about preferences, the solution to this allocation problem has a simple form and is given by:

$$N_a = \left(\frac{\bar{a}}{ZL^{\gamma}} \right)^{\frac{1}{1-\gamma}}.$$

Note that the solution is such that an increase in productivity in the agricultural sector Z reduces the share of employment in agriculture. In particular, the change in employment in agriculture can be easily calculated using the above equation from the change in TFP in agriculture raised to the power $1/(1-\gamma)$. We report the results in Table 11. With $\gamma = 0.54$, an increase in productivity of 3.6 implies a decrease in the employment share in agriculture of 16.2-fold. In other words, the share of employment in agriculture in Malawi would decrease from the actual 65 percent to only 4 percent which is close to the average for rich countries, see for instance Restuccia et al. (2008). This tremendous reallocation of labor from agriculture to non-agriculture implies that average farm size would increase by a factor of 16.2-fold (recall that in our previous one-sector analysis of misallocation, by construction the gains in agricultural productivity of moving to the efficient allocation involved no change in average farm size). It is also easy to see that the increase in labor productivity in agriculture Y_a/N_a is given by a factor of 16.2-fold.[19] Hence, for our quantitative experiment, the entire increase in agricultural labor

and Restuccia (2014).

[18]See for instance Gollin et al. (2002) for a model with such preferences.

[19]Note that agricultural output per unit of land or yield Y_a/L remains the same because for the preferences we

productivity is reflected in an increase in average farm size and none in an increase in the yield.

We emphasize that the effects just described abstract from other potential sources of amplification. For instance, the potential role of selection into the ability of farmers that stay in agriculture. This feature that has been found quantitatively important in amplifying the productivity increases in the agricultural sector (e.g., Lagakos and Waugh 2013). Also, complementary investments such as mechanization, improvements in land quality, or the adoption of modern technologies could further increase productivity in the agricultural sector. Overall, we find that the increase in agricultural productivity due to the more efficient allocation of factors can unravel a substantial process of structural change and productivity growth in agriculture that can dramatically change the face of the Malawian economy.

6 Implications for Economic Inequality

Unlike the actual distribution of factors in the Malawian economy which is fairly flat across farmers with different productivity, the efficient allocation implies a substantial increase in the dispersion of operational scales, both in terms of the distribution of capital and land across farmers. This massive redistribution of factors across farmers may lead to concerns over distributional implications, specially since the actual allocation of factors reflects policy choices and institutional features in place partially to alleviate poverty and distributional concerns, see de Magalhaes and Santaeulàlia-Llopis (2015).[20]

In this section, we study the distributional implications of redistributing operational scales towards the efficient allocation. Table 12 reports the actual and efficient distribution of factors across farmers by productivity and Table 13 reports actual and efficient output and income. As discussed previously, whereas the actual distribution of land across farm TFP is fairly flat, with most farms operating less than 2 acres of land, the efficient distribution implies that the top quintile of farm TFP operates on average almost 6 acres and 97 percent of the land whereas the lowest quintile only operates less than a

consider, the increase in productivity is exactly offset by the decrease in agricultural labor so total agricultural output remains the same.

[20] For example, de Magalhaes and Santaeulàlia-Llopis (2015) show that the bottom 20% and top 20% of the land distribution hold a similar share of aggregate consumption in rural Malawi, respectively 20% and 25%.

percent of the land. This implies a substantial redistribution of factors and operational scales to achieve the higher levels of aggregate productivity. But it is important to emphasize that despite relative equalization of factor inputs across farmers, the actual distribution of income is widely dispersed, in fact as dispersed as the distribution of productivity. For instance, taking agricultural output as a measure of farm income, the ratio of top to bottom quintile of income is a factor of 34-fold even though the ratio of capital and land factors is within a factor of 1 to 2-fold. To put it differently, equalizing access to land across households does not necessarily translate into income equalization as these farmers differ substantially in their productivity of working the farm.

To gauge the income effects of factor redistribution we pursue the following counterfactual. We consider the actual distribution of factors as endowments and allow the efficient allocation of factors to be achieved via perfectly competitive rental markets for capital and land, that is, with perfect markets a decentralized competitive allocation corresponds to the efficient allocation. Given the competitive rental rates of capital and land in this decentralized solution, we compute the income associated with the efficient allocation as:

$$\text{endowment income} = r_k(k^a - k^e) + r_l(l^a - l^e) + y^e,$$

where (k^a, l^a) are the actual allocation of capital and land which we take as endowment and (k^e, l^e) are the efficient allocations, r_k and r_l are the prices that support the efficient allocation as a competitive equilibrium, and y^e is efficient output. Table 13 reports the results for this measure of endowment income and compare the income inequality to that of the actual income, which we assume is approximated by actual output y^a.

Not only farmers in the lower end of the productivity (and income) distribution benefit the most from the increase in the average yield and aggregate agricultural output, but also overall inequality declines. For instance, the overall variance of the log income decreases from 1.8 with actual income to 1.2 to efficient income. More dramatic are the changes in income across the richest and poorest households. Whereas the ratio of income between farmers in the top and bottom quintiles is a factor of 34-fold in the actual allocation, this ratio is 3.4 in the efficient allocation, that is income inequality among these farmers falls by a factor of 10-fold. Moreover, the ratio of efficient to actual income increases for all household farms but this increase is much larger for the households at the bottom of the productivity distribution, 23.7-fold for the first quintile, 3.9-fold for the second quintile and only 2-fold for the top quintile.

Well-functioning rental markets for capital and land to achieve the efficient allocation of operational scales can lead to substantial increases in agricultural productivity as well as dramatic reductions in inequality levels and poverty.

7 Conclusions

Land misallocation in the agricultural sector is enormous in Africa. Using detailed nationally representative household-farm level data for Malawi, we show that a reallocation of land (and capital) to their efficient uses increases agricultural output and productivity by a factor of 3.6-fold. We show these large gains in the agriculture sector in Malawi are not due to differentials in TFP dispersion compared with other sectors or more developed countries, but to a severe misallocation of factor inputs. That is, land size (and capital) are essentially unrelated to household-farm level TFP, which is perhaps not surprising given the egalitarian nature of land-use distributions and weak property rights over land. These restrictions imply that 83% of the total land in Malawi is not marketed. Indeed, our analysis provides a strong empirical connection between factor misallocation and restricted land markets. The productivity gains from reallocation are 2.6 times larger for farms with no marketed land than for farms with marketed land, which is roughly three times the reallocation gains found in the manufacturing sector of China and India in Hsieh and Klenow (2009).

We show that the increase in aggregate productivity from the efficient reallocation of factors could trigger of a profound process of structural change by which the agricultural sector in Malawi could approach levels of farm size and sectoral employment shares of the industrialized world. This result implies that factor misallocation in the agricultural sector is of first order importance to understand productivity differences and economic development across countries.

Further, in a counterfactual exercise we show that the introduction of rental markets where operational scales can deviate from land ownership not only increases aggregate productivity to its efficient level but also decreases income inequality; large but unproductive farmers are better off by renting-out land to small but productive farmers. These findings point to a pressing need to facilitate the reallocation of land to the more productive farmers without necessarily altering the ownership structure. This requires the development of well-functioning rental markets and well-defined property rights over land. What policies and institutions are best in promoting a better allocation of resources across farmers is of crucial importance for future research.

Looking ahead, while our analysis takes the distribution of land and productivity across farmers as given and asked about the efficiency gains of reallocation, it may also be of interest to study the dynamic implications of misallocation for productivity whereby a reduction in misallocation encourages the more productive farmers to grow, utilize modern inputs (mechanization, chemical seeds, intermediate inputs), and invest in farm productivity. See Restuccia and Rogerson (2013) and Restuccia (2013) for a discussion of the importance of the dynamic implications of misallocation. Further, while the gains from the reallocation of agricultural inputs across wives and husbands within farm households are small compared to the gains from reallocation across households, the quantitative role of women — who are potentially more restricted in land (and capital) inputs than men in many parts of the world — in understanding productivity differences across farms remains a very open but important question. We leave these interesting and important extensions of our analysis for future research.

References

Adamopoulos, T. and Restuccia, D. (2014). The Size Distribution of Farms and International Productivity Differences. *American Economic Review*, 104(6):1667–97.

Adamopoulos, T. and Restuccia, D. (2015). Land Reform and Productivity: A Quantitative Analysis with Micro Data.

Banerjee, A. V., Gertler, P. J., and Ghatak, M. (2002). Empowerment and efficiency: Tenancy reform in west bengal. *Journal of political economy*, 110(2):239–280.

Beegle, K., Carletto, C., and Himelein, K. (2012). Reliability of Recall in Agricultural Data. *Journal of Development Economics*, 98(1):34–41.

Berry, A. and Cline, W. (1979). *Agrarian Structure and Productivity in Developing Countries*. Baltimore: Johns Hopkins University.

Besley, T. (1995). Property Rights and Investment Incentives: Theory and Evidence from Ghana. *Journal of Political Economy*, pages 903–937.

Besley, T. and Ghatak, M. (2010). *Property Rights and Economic Development*, volume 5 of *Handbook of Development Economics*, chapter 0, pages 4525–4595. Elsevier.

Binswanger, H. P., Deininger, K., and Feder, G. (1995). Power, Distortions, Revolt and Reform in Agricultural Land Relations. In Chenery, H. and Srinivasan, T., editors, *Handbook of Development Economics*, volume 3 of *Handbook of Development Economics*, chapter 42, pages 2659–2772. Elsevier.

Buera, F. J., Kaboski, J. P., and Shin, Y. (2014). Macro-Perspective on Asset Grants Programs: Occupational and Wealth Mobility. *American Economic Review*, 104(5):159–64.

Carletto, C., Savastano, S., and Zezza, A. (2013). Fact or Artifact: The Impact of Measurement Errors on the Farm SizeProductivity Relationship. *Journal of Development Economics*, 103(C):254–261.

Chen, C. (2015). Land Tenure, Occupational Choice and Agricultural Productivity,. Unpublished manuscript, University of Toronto.

de Magalhaes, L. and Santaeulàlia-Llopis, R. (2015). The Consumption, Income, and Wealth of the Poorest: Cross-Sectional Facts of Rural and Urban Sub-Saharan Africa for Macroeconomists. World Bank Policy Research Working Paper, WPS7337.

de Soto, H. (2000). *The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else*. Basic Books. New York.

FAO (2013). Malawi BEFS Country Brief. Technical report, Food and Agriculture Organization of the United Nations.

Foster, A. D. and Rosenzweig, M. R. (2011). Are Indian Farms Too Small? Mechanization, Agency Costs, and Farm Efficiency. Unpublished Manuscript, Brown University and Yale University.

Goldstein, M. and Udry, C. (2008). The Profits of Power: Land Rights and Agricultural Investment in Ghana. *Journal of Political Economy*, 116(6):981–1022.

Gollin, D., Lagakos, D., and Waugh, M. E. (2014). Agricultural Productivity Differences across Countries. *American Economic Review*, 104(5):165–70.

Gollin, D., Parente, S., and Rogerson, R. (2002). The Role of Agriculture in Development. *American Economic Review*, 92(2):160–164.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics*, 124(4):1403–1448.

Lagakos, D. and Waugh, M. E. (2013). Selection, Agriculture, and Cross-country Productivity Differences. *The American Economic Review*, 103(2):948–980.

Larson, D. F., Otsuka, K., Matsumoto, T., and Kilic, T. (2014). Should African Rural Development Strategies Depend on Smallholder Farms? An Exploration of the Inverse-Productivity Hypothesis. *Agricultural Economics*, 45(3):355–367.

Restuccia, D. (2013). Factor Misallocation and Development. *The New Palgrave Dictionary of Economics, Online Edition, Eds. Steven N. Durlauf and Lawrence E. Blume, Palgrave Macmillan.*

Restuccia, D. and Rogerson, R. (2008). Policy Distortions and Aggregate Productivity with Heterogeneous Plants. *Review of Economic Dynamics*, 11(4):707–720.

Restuccia, D. and Rogerson, R. (2013). Misallocation and Productivity. *Review of Economic Dynamics*, 16(1):1–10.

Restuccia, D., Yang, D. T., and Zhu, X. (2008). Agriculture and Aggregate Productivity: A Quantitative Cross-Country Analysis. *Journal of Monetary Economics*, 55(2):234–250.

Shaban, R. A. (1987). Testing Between Competing Models of Sharecropping. *The Journal of Political Economy*, pages 893–920.

Udry, C. (1996). Gender, Agricultural Production, and the Theory of the Household. *Journal of Political Economy*, 104(5):1010–46.

Valentinyi, A. and Herrendorf, B. (2008). Measuring Factor Income Shares at the Sector Level. *Review of Economic Dynamics*, 11(4):820–835.

Table 1: Size Distribution of Farms (% of Farms by Size)

	ISA 2010/11	World Census of Agriculture 1990		
	Malawi	Malawi	Belgium	USA
Hectares (Ha):				
≤ 1 Ha	78.3	77.7	14.6	–
$1-2$ Ha	17.8	17.3	8.5	–
$2-5$ Ha	3.7	5.0	15.5	10.6
$5-10$ Ha	0.2	0.0	14.8	7.5
10+ Ha	0.0	0.0	46.6	81.9
Average Farm Size (Ha)	0.83	0.7	16.1	187.0

Notes: The first column reports the land size distribution (in hectares) for household farms from the Malawi 2010/11 Integrated Survey of Agriculture (ISA). The other columns report statistics from the World Census of Agriculture 1990 for Malawi, Belgium, and United States documented in Adamopoulos and Restuccia (2014a).

Table 2: Dispersion of Output, Land Size, Land Quality, and Rain, Malawi ISA 2010/11

		Within Geographic Areas		
	Full Sample	Regions	Districts	Enum. Area
Output, y_i:	1.896	1.867	1.778	1.649
Land Size, l_i:	.749	.746	.719	.671
Land Quality:				
▷ Index, q_i	.852	.833	.568	.147
▷ Index Subitems:				
Elevation	.439	.349	.075	.001
Slope (%)	.657	.635	.453	.093
Erosion	.480	.496	.472	.427
Soil Quality	.608	.605	.514	.458
Nutrient Avail.	.387	.402	.248	.023
Nutrient Ret. Cap.	.329	.365	.180	.016
Rooting Conditions	.342	.372	.302	.029
Oxygen Avail. to Roots	.097	.094	.105	.010
Excess Salts	.037	.045	.048	.004
Toxicity	.027	.038	.033	.003
Workability	.475	.474	.335	.033
Quality-Adjusted Land Size, $q_i l_i$:	1.571	1.531	1.243	.808
Rain, ζ_i:				
▷ Annual Precip. (mm)	.039	.025	.014	.001
▷ Precip. of Wettest Qrter (mm)	.026	.013	.005	.000

Notes: Output y_i, land size (in acres) l_i, land quality index q_i, quality-adjusted land size $q_i l_i$, and rain ζ_i, are continuous variables for which we use the variance of logs as a measure of dispersion. Subitems of land quality, slope (in %) and elevation (in meters), are also continuous variables and we also report the variance of logs as dispersion measure. The other subitems of land quality are categorical variables such as soil quality (with categories 1 good, 2 fair, 3 poor), erosion (with categories 1 none, 2 low, 3 moderate, 4 high) and nutrient availability, nutrient retention, rooting conditions, oxygen to roots, excess of salts, toxicity and workability that take values from four categories (1 low constraint, 2 moderate constraint, 3 severe constraint and 4 very severe constraint). For all categorial variables we use the proportion of non-mode values in the sample as measure of dispersion. The construction of the land quality index q_i is discussed in Section 3. For the last three columns referring to within geographic areas, we report the averages across geographic area under consideration (e.g., the variance of output in the 'Regions' column is the average across regions of the variance of output by region).

Table 3: Dispersion of Productivity across Farms and Manufacturing Plants

	Farms		Manufacturing Plants		
	Malawi	USA	USA	China	India
Statistic	ISA 2010/11	1990	1977	1998	1987
SD	1.19	0.80	0.85	1.06	1.16
75-25	1.15	1.97	1.22	1.41	1.55
90-10	2.38	2.50	2.22	2.72	2.77
N	7,157	AR(2014)	164,971	95,980	31,602

Notes: The first column reports statistics for the household-farm productivity distribution from the micro data in Malawi. The second column reports statistics for farm productivity in the United Sates from the calibrated distribution in Adamopoulos and Restuccia (2014) to U.S. farm-size data. The other columns report statistics for manufacturing plants in Hsieh and Klenow (2009). SD is the standard deviation of log productivity; 75-25 is the log difference between the 75 and 25 percentile and 90-10 the 90 to 10 percentile difference in productivity. N is the number of observations in each dataset.

Table 4: Variance Decomposition of Agricultural Output, Malawi ISA 2010/11

	Benchmark		$(\zeta_i = 1, q_i = 1)$	
	Level	%	Level	%
$var(y)$	1.896	100.0	1.896	100.0
$var(s)$	1.435	75.7	1.457	76.8
$var(\zeta)$	.039	2.1	–	–
$var(f(k, ql))$	.383	20.2	.343	18.1
$2cov(s, \zeta)$	-.044	-2.3	–	–
$2cov(s, f(k, ql))$	.034	1.8	.096	5.1
$2cov(\zeta, f(k, ql))$	.048	2.5	–	–

Notes: The variance decomposition uses our benchmark production function that adjusts for rain ζ_i and land quality q_i across household farms, $y_i = s_i \zeta_i f(k_i, q_i l_i)$ with $f(k_i, q_i l_i) = k_i^{.36}(q_i l_i)^{.18}$. All variables have been logged. The variables are output y_i, household-farm productivity s_i, rain ζ_i, structures and equipment capital k_i, and quality-adjusted land size, $q_i l_i$. The first two columns report results from our benchmark specification where rain and land quality are controlled for. The last two columns report the results abstracting from rain and land quality, i.e. we set $\zeta_i = 1$ and $q_i = 1$ $\forall i$. In each case, the column 'Level' reports the variance and the column '%' reports the contribution to total output in percentage points. The construction of the land quality index q_i and rain ζ_i and their effects on farm productivity are discussed in detail in Section 3.

Table 5: Agricultural Output Loss (Y^a/Y^e)

(a) Main Results

		Bootstrap Simulations		
	Full Sample	Median	5th pct.	95th pct.
Nationwide	.2788	.2817	.2455	.3211

(b) Within Productivity-s_i Variation

	Benchmark	95%	90%	80%	0%
Output Loss	.2788	.2939	.3080	.3330	.4032

(c) By Geographical Areas and Institutions

	Average	Median	Min	Max
Geographic Areas:				
Regions	.2943	.2921	.2757	.3485
Districts	.3410	.3568	.1409	.4821
Enum. Areas	.6077	.6245	.0911	.9624
Institutions:				
Traditional Authority	.3213	.2872	.1791	.7220
Language	.2980	.2576	.1681	.8646

Notes: In panel (a), bootstrap median and confidence intervals are computed from 5,000 simulations obtained from random draws with 100 percent replacement, i.e., each simulation consists of a sample of the same size as the original sample. See further discussion in Appendix B. Panel (b), see text in Section 4 for details. Panel (c) reports the ratio of actual to efficient output when the reallocation exercise is conducted within three narrower definitions of geographical areas (3 regions, 31 districts and 713 enumeration areas) and two measures of institutional settings/cultural identity (53 traditional authorities and 13 languages). We drop enumeration areas with less than 5 household-farm observations.

Table 6: Land Markets, Output Loss for Farms with Marketed vs. Non-marketed Land

	By Marketed Land Share			By Marketed Land Type			
	No	Yes	All	Rented		Purchased	
	(0%)	(> 0%)	(100%)	Informal	Formal	Untit.	Titled
Output Loss	.2411	.5081	.6378	.5809	.5782	.1951	.7192
Observations	5,962	1,189	746	215	682	126	97
Sample (%)	83.4	16.6	10.4	3.0	9.5	1.8	1.3

Notes: The output loss is calculated as in our benchmark full sample reallocation but aggregate output losses are computed separately for subsamples of farm households defined by the share of marketed land used and its type. This share of marketed land is defined from the household-farm level information on how land was acquired, see Section 2. Each column refers to a particular subsample. The first column reports the output losses (and its inverse, the gains) for the subsample of household farms that do not operate any marketed land. The household farms in this subsample operate land that was either granted by a chief, inherited or as bride price. The second column refers to the subsample of household farms operating a strictly positive amount of marketed land, either purchased or rented-in. The third column refers to the subsample of household farms for which all their operated land is marketed land. The last four columns disaggregate the results by the main types of marketed land: (1) rented informally, i.e. land borrowed for free or moved in without permission; (2) rented formally, i.e. leaseholds, short-term rentals or farming as a tenant; (3) purchased without a title; and (4) purchased with a title. There is 1% of households with marketed land whose type is missing in the Malawi ISA data.

Table 7: Reallocation Results with Factor Shares From Malawi Micro Data

	Benchmark, U.S.	Micro Malawi ISA Data	
	VH (2008)	Full Sample	Renting Sample
	θ_k=.360, θ_l=.180	θ_k=.190, θ_l=.391	θ_k=.091, θ_l=.332
Output Loss	.2788	.2799	.3290

Notes: VH (2008) refers to the US values for the agricultural sector provided by Valentinyi and Herrendorf (2008). Full Sample measures land and capital shares using the average rental rates of land and capital computed as the ratio of factor rental payments to factor stocks from the sample of farms that rent in all of their capital and land. We use these rental rates to impute the land and capital rental income per farm in the entire sample by multiplying these rates times the respective farm level stocks. Renting Sample measures factor shares using the average ratios of factor rental payments to value added using the renting sample only and then applying these averages to the entire sample. The sample of farm-households used is the same in all specifications.

Table 8: Reallocation Results by Crop Type

		Maize Share		Non-Maize Share	
	Bench.	>0%	> Median	>0%	> Median
Output Loss	.2788	.3534	.3687	.2618	.2419

Notes: This table reports the efficient reallocation exercise within households producing the main crop in Malawi, maize, and within households producing non-maize crops. Because most households (74%) produce both maize and non-maize crops, the table also reports the reallocation results within household-farms whose production of maize is above median production. The median share of maize production within the group of household-farms that produce a strictly positive amount of maize is 83.1% and the median share of non-maize production within the group of household-farms that produce a strictly positive amount of non-maize crops is 66.2%.

Table 9: Robustness to Human Capital and Specific Skills

Output Loss within Schooling and Terrain-specific Groups

Schooling Groups:	No Schooling	Dropouts	Primary	> Primary
Output Loss	.3104	.3005	.2628	.2795
Terrain-roughness Skills:	High Altitude Plains	Low Plateaus	Mid-Altitude Plateaus	Mid-Altitude Mountains
Output Loss	.2712	.2697	.3297	.3655

Notes: The table reports robustness exercises within schooling groups and within types of terrain roughness for which specific skills might be required. The sample distribution across education groups defined as highest educational degree completed is: no schooling 25%, primary school dropouts 45%, primary school graduates 23%, and more than a primary school degree 7%.

Table 10: Robustness to Health and other Transitory Risks

Output Loss when Productivity is Net of Health and Other Transitory Risks

		Productivity Specification					
	Bench.	(1)	(2)	(3)	(4)	(5)	(6)
Further Transitory Risks Controls:							
Health Risk (last 12m)	–	✓	✓	✓	✓	✓	✓
Death Risk (last 2 years)	–	–	✓	✓	✓	✓	✓
Food Security Risk (last 12m)	–	–	–	✓	✓	✓	✓
Marital Status	–	–	–	–	✓	✓	✓
Distance to Markets	–	–	–	–	–	✓	✓
Other Income Sources	–	–	–	–	–	–	✓
Output Loss	.2788	.2763	.2764	.2679	.2678	.2685	.2674

Notes: This table reports measures of household-farm productivity that controls for transitory effects on output such as health and other individual shocks. Our measure of health risks includes illnesses/injuries in the last 2 weeks, hospitalizations in the last 12 months (formal and at traditional healer locations), health expenditures in the last 12 months (including prevention expenditures and treatment) and malaria conditions in the last 12 months (including the use of bed nets and insecticides). Death risk includes deaths in the family over the past two years depending on the age of the deceased. Food security states whether the household suffered an episode of not being able to feed family members in the last 12 months. In higher order specifications we also include marital status, distance to roads, markets and main suppliers, and variation in other sources of income. See Appendix D for details.

Table 11: Implications for Structural Transformation: Reallocation with Two Sectors

	Benchmark	Increased Productivity
Productivity in Agriculture (Z)	1.00	3.60
Share of Employment in Ag. (N_a)	0.65	0.04
Average Farm Size (L/N_a)	1.00	16.2
Labor Productivity in Ag. (Y_a/N_a)	1.00	16.2

Notes: The "Benchmark" refers to the actual allocation in Malawi. The "Increased Productivity" refers to the efficient reallocation which in the full sample increases total factor productivity in agriculture by a factor of 3.6-fold. The table reports the effects of increased TFP in agriculture on output per unit of land in agriculture (yield), the share of employment in agriculture, average farm size, and aggregate labor productivity in agriculture when factors are allowed to be reallocated across agriculture and non-agriculture.

Table 12: Actual and Efficient Distribution of Land, Capital, MPL and MPK

Malawi ISA 2010/11 Productivity Partition	Bottom(%)			Quintiles					Top(%)				
	0-1	1-5	5-10	1st	2nd	3rd	4th	5th	10-5	5-1	1	Total	$var(\ln x)$
Productivity s_i:	.07	.23	.41	.75	2.10	3.72	6.39	21.50	32.36	48.06	112.01	6.89	1.435
Land, l_i:													
Actual	1.35	.94	1.50	1.19	.87	1.01	1.03	1.99	2.36	2.86	2.78	1.22	.749
Efficient	.00	.00	.00	.00	.01	.04	.14	5.91	11.40	21.70	88.66	1.22	6.782
Capital, k_i:													
Actual	204.01	98.10	64.34	55.93	25.35	21.84	24.70	26.71	28.50	32.44	34.35	30.91	1.820
Efficient	.00	.00	.01	.04	.32	1.10	3.60	149.52	288.46	548.94	2,242.11	30.91	6.782
MPL ($\propto$ Yield):													
Actual	.23	.90	.96	.75	2.10	3.72	6.39	21.50	56.37	69.81	152.09	21.27	1.485
Efficient	76.30	76.30	76.30	76.30	76.30	76.30	76.30	76.30	76.30	76.30	76.30	76.30	.000
MPK:													
Actual	.00	.01	.04	.75	2.10	3.72	6.39	21.50	9.35	12.34	24.66	1.68	2.154
Efficient	6.03	6.03	6.03	6.03	6.03	6.03	6.03	6.03	6.03	6.03	6.03	6.03	.000

Notes: Households are ranked from bottom to top according to household-farm productivity s_i. Land (acres) and capital (USD) are expressed in per 100 hours terms. Regarding with the marginal product of land, MPL, and the marginal product of capital, MPK, we compute for each productivity bin, MPL $= .18 \frac{\sum_i y_i}{\sum_i l_i}$ and MPK $= .36 \frac{\sum_i y_i}{\sum_i k_i}$, and the variance is computed with respect to household-farm specific ratios $\text{MPL}_i = .18 \frac{y_i}{l_i}$ and $\text{MPK}_i = .36 \frac{y_i}{k_i}$.

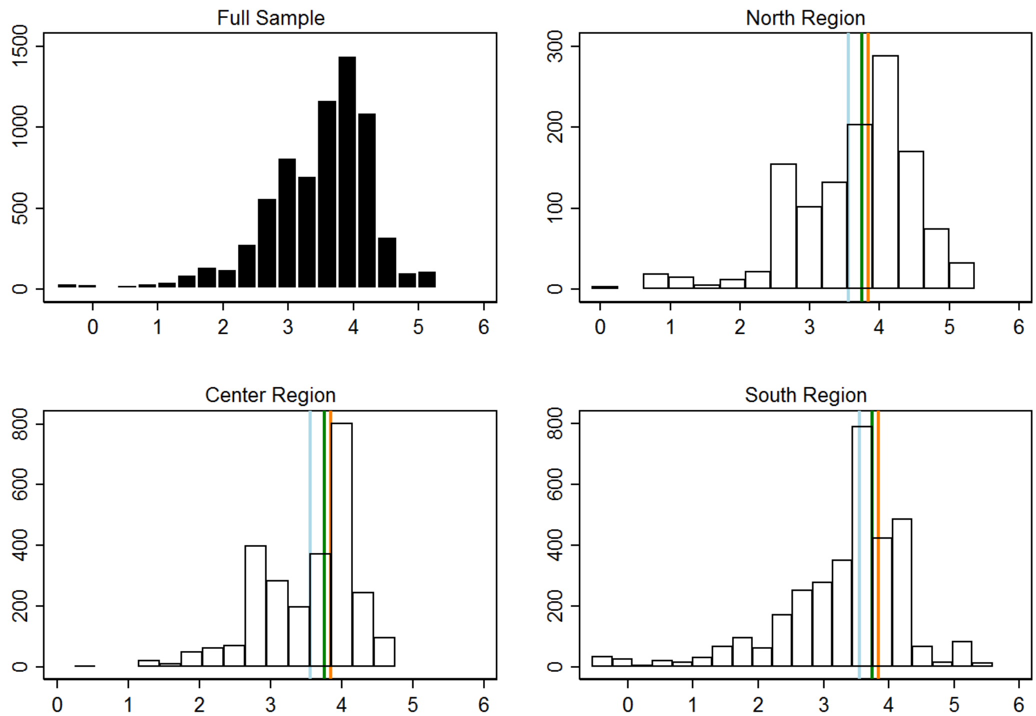
Table 13: Actual and Efficient Distribution of Operational Scale vs. Income

Malawi ISA 2010/11 Productivity Partition	Bottom(%)			Quintiles					Top(%)				
	0-1	1-5	5-10	1st	2nd	3rd	4th	5th	10-5	5-1	1	Total	$var(\ln x)$
Output, y_i:													
▷ Level:													
Actual	.01	.04	.08	.14	.39	.69	1.20	4.78	7.41	11.12	23.53	1.44	1.824
Efficient	.00	.00	.00	.00	.05	.18	.60	25.06	48.35	92.01	375.83	5.18	6.782
▷ % of Total:													
Actual	.01	.16	.55	2.01	5.46	9.57	16.67	66.26	51.24	38.42	16.16	100.00	–
Efficient	.00	.00	.00	.02	.20	.71	2.33	96.71	93.22	88.57	71.95	100.00	–
Income:													
▷ Level:													
Actual	.01	.04	.08	.14	.39	.69	1.20	4.78	7.41	11.12	23.53	1.44	1.824
Efficient	13.34	6.64	5.02	4.28	2.22	2.17	2.56	14.65	25.76	46.47	177.08	5.18	1.228
Ratio: Actual/Eff.	155.78	58.03	39.73	23.70	3.88	2.27	1.58	1.97	2.58	3.71	9.86	6.68	–
▷ % of Total:													
Actual	.01	.16	.55	2.01	5.46	9.57	16.67	66.26	51.24	38.42	16.16	100.00	–
Efficient	2.59	6.41	9.71	16.55	8.58	8.41	9.88	56.56	49.68	44.73	33.90	100.00	–

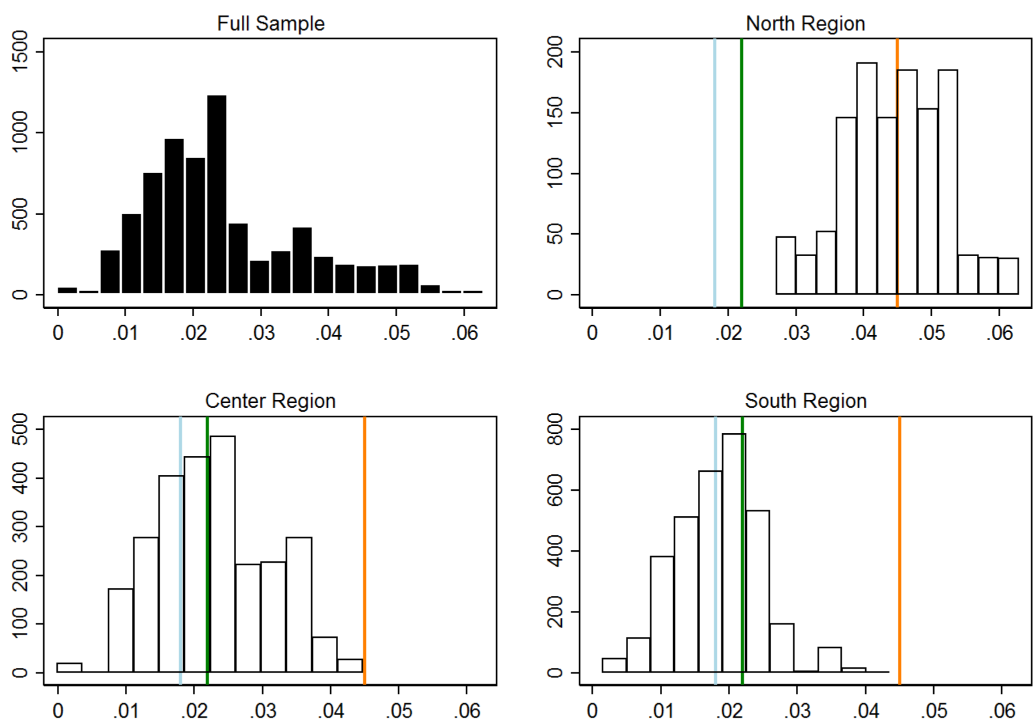
Notes: Households are ranked from bottom to top according to household-farm productivity s_i. The definition of income refers to a scenario in which the ownership of household-farm endowments of land and capital does not change and there is reallocation entirely through rental markets, see Section 6.

Figure 1: Histograms of Land Quality Index and Rain, Malawi ISA 2010/11

(a) Land Quality Index, q_i (in logs)

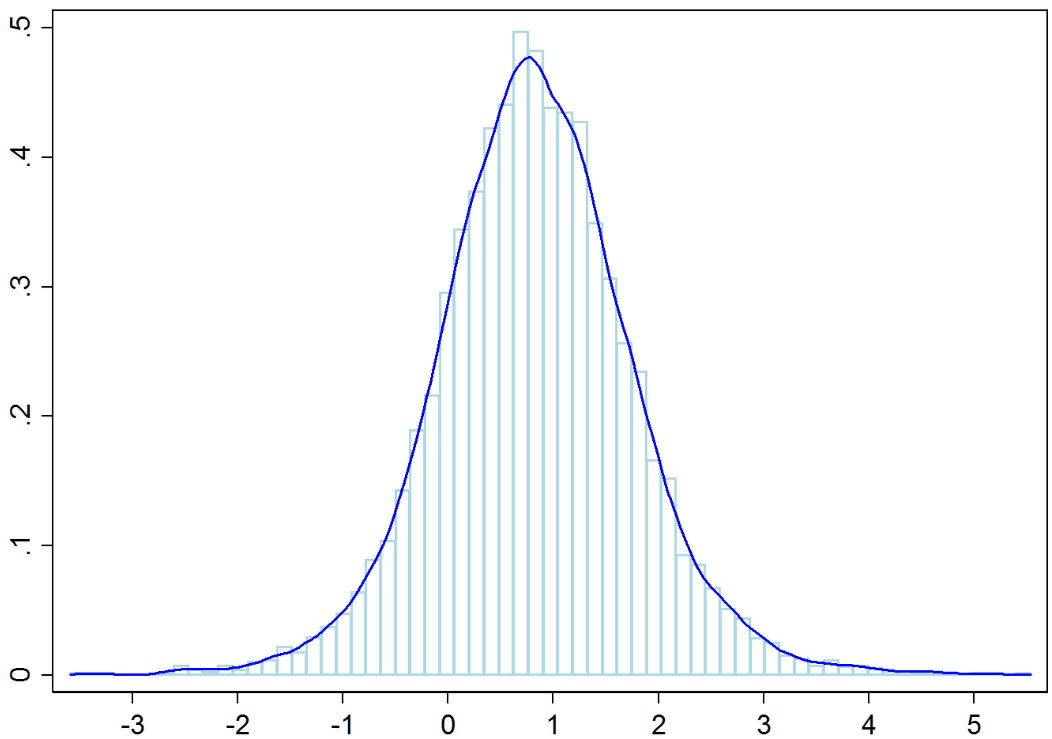


(b) Rain, ζ_i (in logs)



Notes: The land quality index is computed in Section 3. Median values by region are depicted with vertical lines: Orange (North), green (Center) and light blue (South).

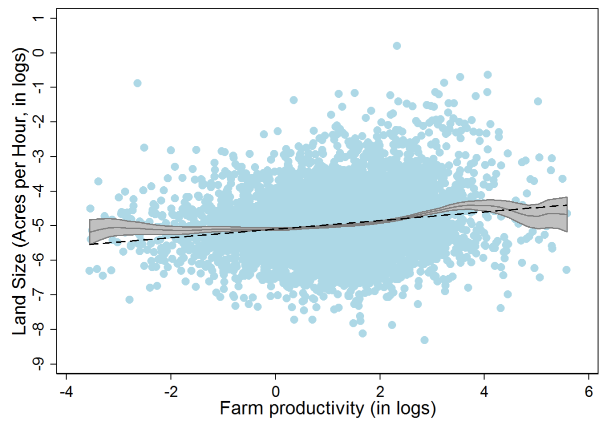
Figure 2: Density of Farm Productivity s_i (in logs), Malawi ISA 2010/11



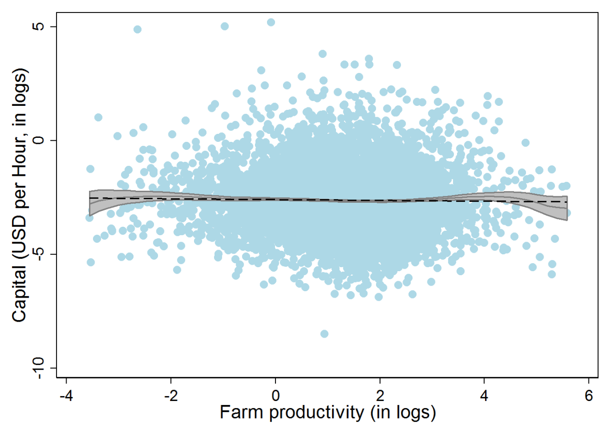
Notes: Household-farm productivity s_i is measured using our benchmark production function, adjusting for rain ζ_i and land quality q_i , $y_i = s_i \zeta_i f(k_i, q_i l_i)$ with $f(k_i, q_i l_i) = k_i^{.36} (q_i l_i)^{.18}$, where y_i is farm output, k_i is capital, and l_i is land. All variables have been logged.

Figure 3: Land and Capital by Farm Productivity, Malawi ISA 2010/11

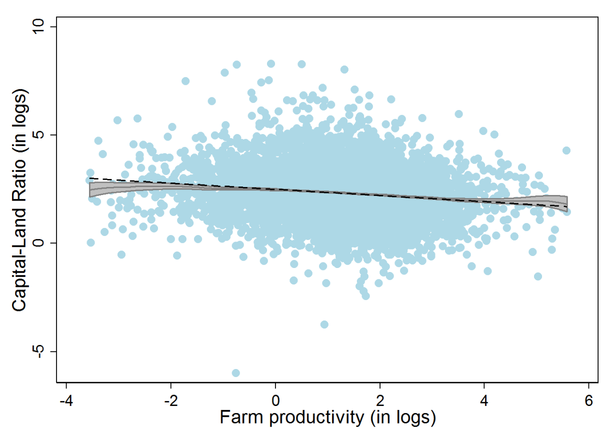
(a) Land Size by Farm Productivity



(b) Capital by Farm Productivity



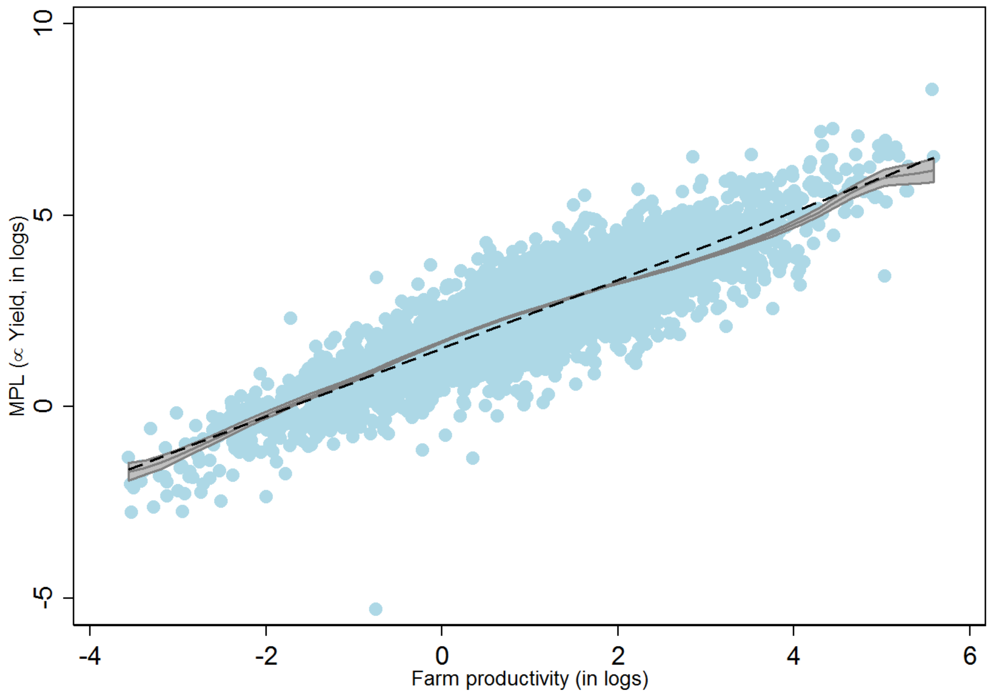
(c) Capital-Land Ratio by Farm Productivity



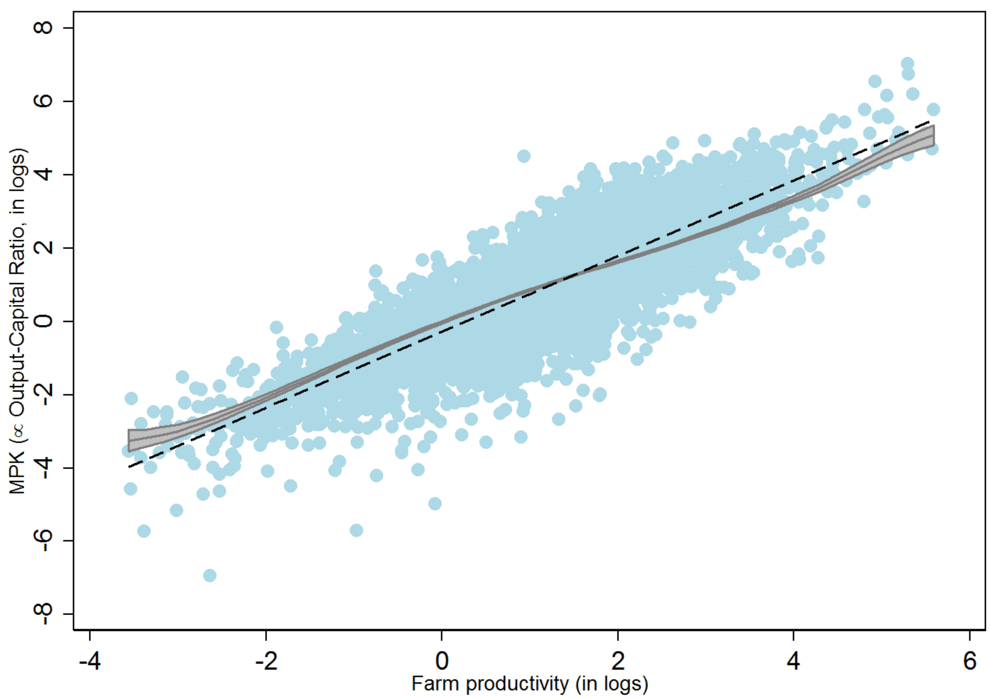
Notes: Panel (a) reports the relationship between operated farm land size (in acres per hour) l_i and farm productivity s_i. Panel (b) reports the relationship between farm capital (in USD per hour) k_i and farm productivity s_i. Panel (c) reports the relationship between the farm capital to land ratio k_i/l_i and farm productivity s_i. The computation of farm productivity, s_i, is discussed in Section 3. All variables have been logged.

Figure 4: Marginal Product of Land and Capital by Farm Productivity, Malawi ISA 2010/11

(a) MPL ($\propto$ Yield) vs. Farm Productivity



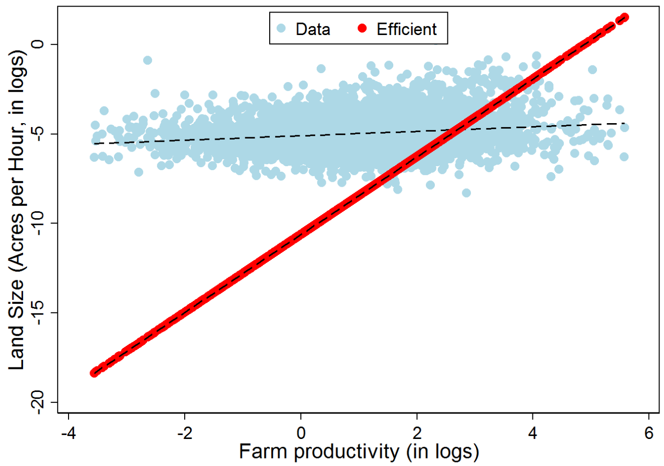
(b) MPK ($\propto$ Output-Capital Ratio) vs. Farm Productivity



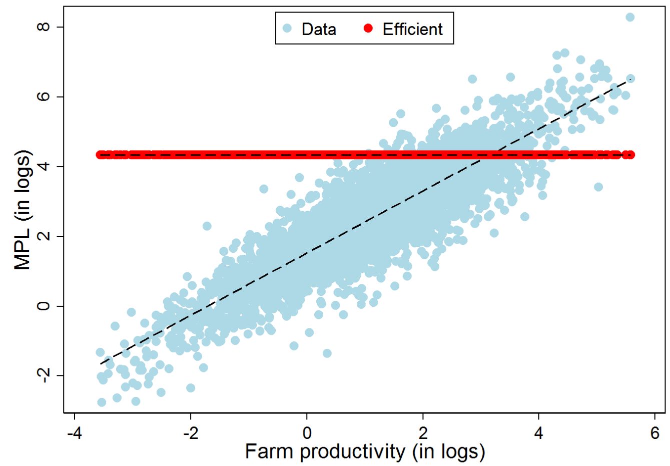
Notes: Panel (a) reports the relationship between the marginal product of land constructed using our benchmark production function, i.e., MPL = .18 y_i/l_i (which is proportional to the yield) and farm productivity s_i. Panel (b) reports the relationship between the marginal product of capital constructed using our benchmark production function, i.e., MPK= .36 y_i/k_i (which is proportional to the output-capital ratio) and farm productivity, s_i.. The computation of farm productivity s_i is discussed in Section 3. All variables have been logged.

Figure 5: Land Size, Capital, MPL and MPK: Actual and Efficient Allocations

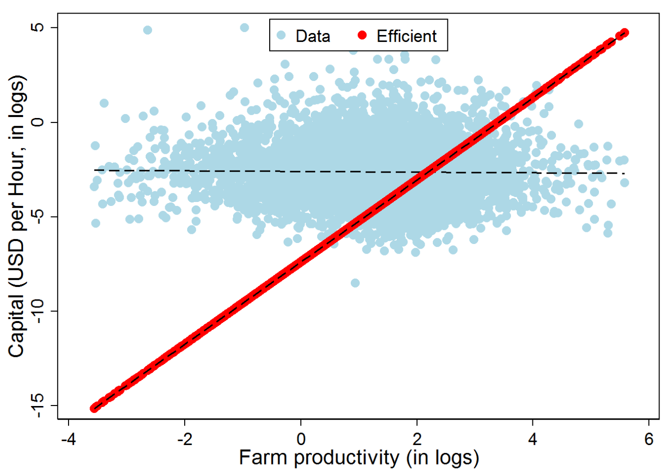
(a) Land Size vs. Farm Productivity



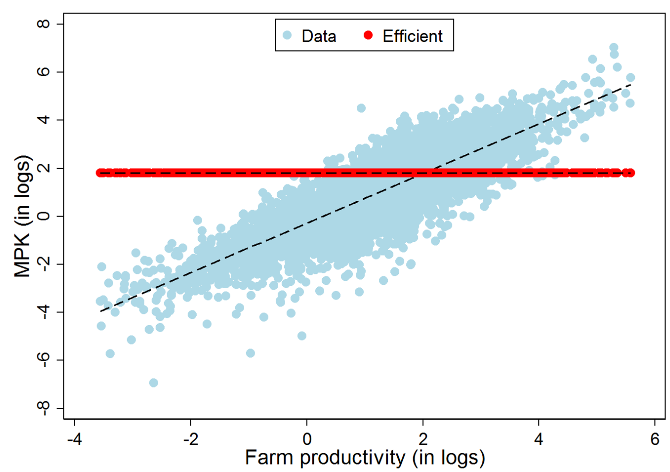
(b) MPL vs. Farm Productivity



(c) Capital vs. Farm Productivity

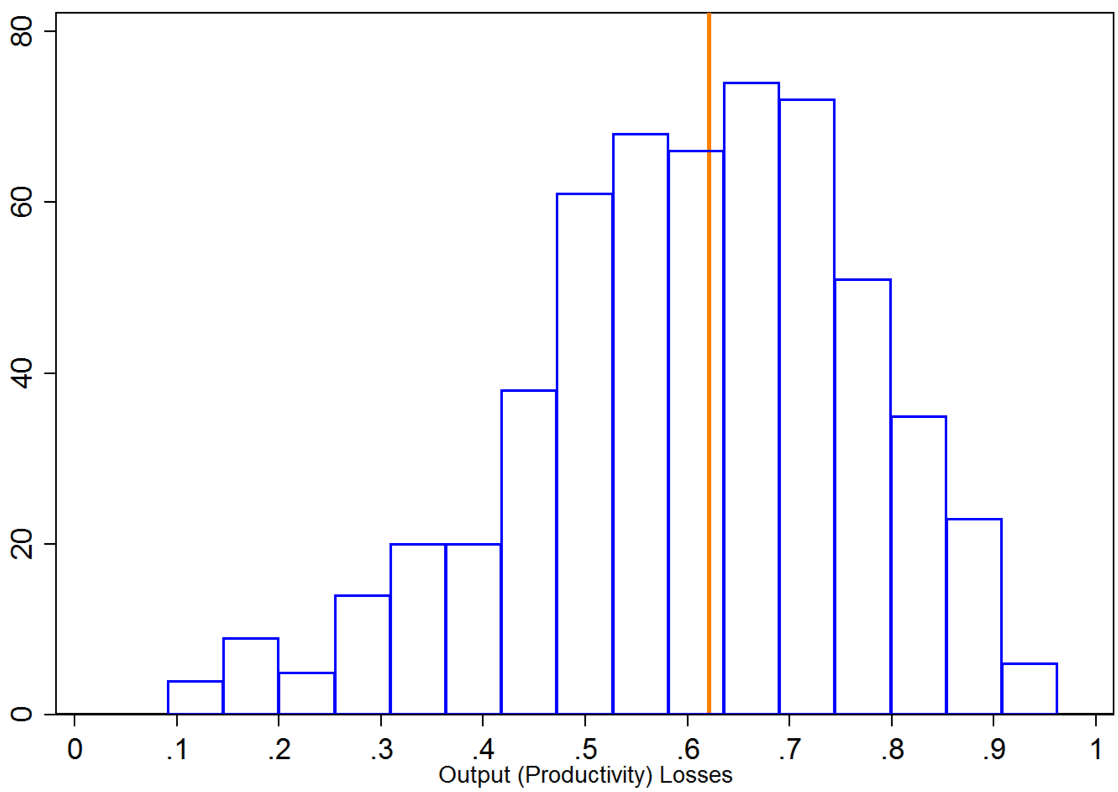


(d) MPK vs. Farm Productivity



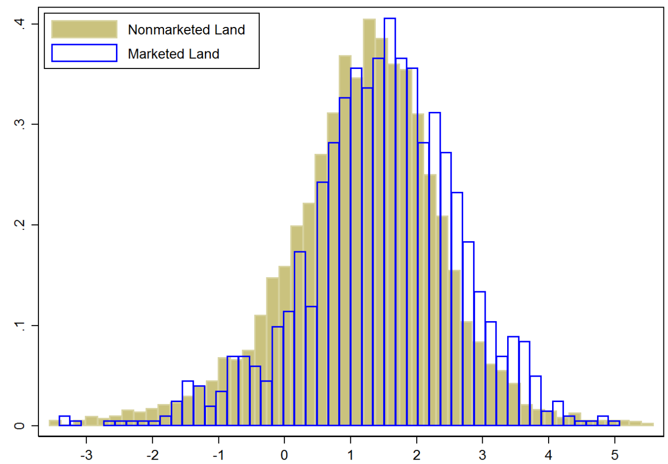
Notes: Panel (a) scatters the actual and efficient relationship between land size, l_i, and farm productivity, s_i. Panel (b) scatters the actual and efficient relationship between the marginal product of land, MPL, and farm productivity, s_i. Panel (c) scatters the actual and efficient relationship between capital, k_i, and farm productivity, s_i. Panel (d) scatters the actual and efficient relationship between the marginal product of capital, MPK, and farm productivity, s_i. The computation of farm productivity, s_i, is discussed in Section 3. All variables have been logged.

Figure 6: Histogram of Output (Productivity) Losses within Enumeration Areas



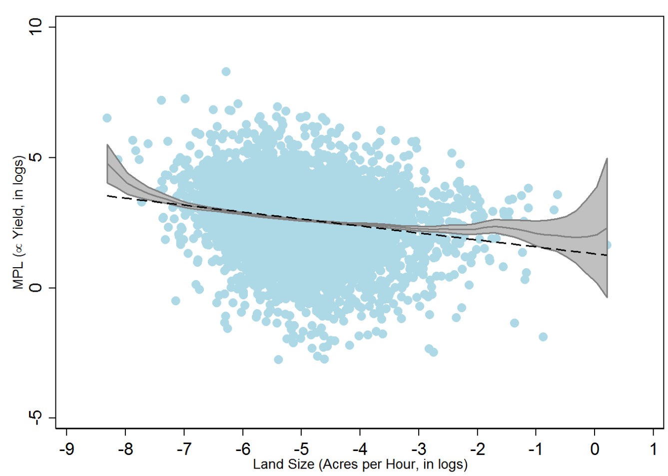
Notes: The figure reports the output loss (actual to efficient output) by enumeration area. Enumeration areas are a survey definition of groups of about 16 households each, the narrowest geographic area in the survey.

Figure 7: Density of Farm Productivity s_i (in logs) by Marketed Land, Malawi ISA 2010/11



Notes: Household-farm productivity s_i is measured using our benchmark production function, adjusting for rain ζ_i and land quality q_i , $y_i = s_i \zeta_i f(k_i, q_i l_i)$ with $f(k_i, q_i l_i) = k_i^{.36} (q_i l_i)^{.18}$, where y_i is farm output, k_i is capital, and l_i is land. All variables have been logged. The sample is divided between those that operate land acquired in the market and those that do not.

Figure 8: Land Productivity by Farm Size, Malawi ISA 2010/11



Notes: Household-farm land size is the sum of all plots per household farm and marginal product of land (MPL) is constructed using our benchmark production function, i.e., MPL = .18 y_i/l_i. All variables have been logged.