PRODUCTIVITY AND MARKET PARTICIPATION: CAMBODIAN RICE FARMERS

Channary Khun
Researcher
Southern Illinois University Carbondale
1263 Lincoln Dr, Carbondale, IL 62901, USA
khunchannary@gmail.com

ABSTRACT

With Cambodia's agriculture policies focusing on promoting rice production and exports, the study attempts to understand the underlying issues of low productivity and whether progress made on this front will lead to greater market participation. Employing a control function to address the endogeneity issue and using the latest survey of Cambodia's agriculture sector, the paper finds that rice productivity strongly affects commercialization. That is, a ton increase in rice yields leads to about 20 percent rise in the probability of market participation. This has an important implication; policies and reforms targeting productivity not only boost rice production, but also promote commercialization and possibly exports. In addition, enhancing productivity hinges, among others, on improvement in general education, expansion of irrigation and pesticide usage, as well as increased adoption of aromatic paddies, sticky paddies and modern varieties.

Keywords: Market participation, rice yields, productivity, commercialization

I. INTRODUCTION

Economic development and poverty reduction in low-income economies remain heavily dependent upon the performance of the agriculture sector. In 2018, agriculture value added constitutes about a quarter of GDP and two thirds of employment in low income economies (World Development Indicators, 2021). Given population growth and increasing constraint on farmland expansion, enhancing productivity of poor farmers is key to rapid and sustainable progress to poverty alleviation (Datt & Ravallion, 1998; Minten & Barrett, 2008; Palmer-Jones & Sen, 2003) and integrating them into local and global supply chain represents an effective way to improve productivity and rural income (Govereh & Jayne, 2003; Ogutu & Qaim, 2019; Strasberg et al. 1999). It enhances productivity by increasing specialization, improving access to modern inputs, achieving greater economies of scale, and promoting technological adoption (Asfaw et al., 2012; Barrett, 2008; Govereh & Jayne, 2003). Rising farm productivity directly leads to increasing production and expanding employment, which may in turn leads to greater market participation (Abu et al., 2016; Rios et al., 2009). It also indirectly boosts relative wages and reduces food prices (Datt & Ravallion, 1998).

Recognizing the importance of agriculture productivity and market participation in tackling the longstanding problem of poverty and food security (i.e., Ashley & Maxwell, 2001; Binswanger & Braun, 1991; Braun, 1995; Irz et al., 2001; Rahman & Westley, 2001), the study set out to determine key factors that influence farmers' productivity and, more importantly, to examine how productivity affects farmers' decision to participate in the market. The study analyzes the case of Cambodia's rice sector where national policies have primarily focused on production and exports (Eliste & Zorya, 2015). Cambodia's agriculture production is dominated by paddy rice and its cultivated area has steadily increased. In 2019, paddies occupied about three fourths of the total cultivated area three times that of other crops combined (Food and Agriculture Organization Corporate Statistical Database, 2021). Rice production in the country, to a large extent, is cultivated by low-input, low-productivity farmers, the majority of whom are poor and subsistence or semi-subsistence. Albeit some progress, Cambodia's rice yields remain relatively low compared to that of its neighboring countries, indicating potential role for yield augmentation.[1]

By analyzing factors contributing or hampering their ability to improve efficiency and its link to commercialization, the study attempts to understand key challenges in Cambodian rice sector and provide a new perspective on potential policies and reforms. Progress made on these fronts will contribute to the nation's poverty alleviation, food and nutrition security, and rural development while supporting national agriculture goals.

There is a growing body of empirical studies examining agriculture productivity in developing countries. They predominantly involve the puzzling negative relationship between productivity and farm size (i.e., Barrett et al., 2010; Kimhi, 2006; Larson et al., 2012; Mazumdar, 1965), the impact of credit access or constraints (i.e., Akudugu et al., 2012; Ali et al., 2014; Foltz, 2004; Lawal et al., 2009), and the role of agriculture inputs such as irrigation (i.e., Ahmed & Sampath, 1992; Bidzakin et al., 2018; Huang et al., 2006; Kuppannan et al., 2017; Nonvide, 2017; Wang et al., 2017). Of particular interest is the research on rice productivity in Cambodia. To be precise, Yu and Fan (2011) estimate rice production response to prices using the Cambodia Socioeconomic Surveys conducted in 2004-2007. Chun (2014), on the other hand, employs a farm investment climate assessment survey to determine key factors contributing to rice production and

[1] Yield and productivity are used interchangeably; they are defined as output in kg per harvest, per ha of land.

commercialization of small farms in three provinces: Battambang, Kampong Thom, and Takeo. Applying the stochastic frontier model, Kea et al. (2016) measure technical efficiency and examine the determinants of rice production in 25 provinces between 2012 and 2015. In addition, Mishra et al. (2018) investigate the impact of flood and drought as well as access to capital on rice production in Cambodia by utilizing survey data from the International Rice Research Institute covering four provinces surrounding the Tonle Sap Lake: Kampong Thom, Pursat, Battambang, and Siem Reap. Finally, Chhim et al. (2020) attempt to determine factors augmenting rice production and efficiency in Takeo Province. Most of these studies find that rice production in Cambodia is significantly below its potential and have identified key factors that are crucial to enhanced rice production and farmers' income. Collectively, they include irrigation, fertilizer, pesticide, machinery, farm land, production technique, seeds, labor, domestic milling, education, credit, drought, and soil fertility. One key issue in some of these studies is that they look at factors influencing rice production rather than yields. We contend that productivity is better measured by yield than total output which can be increased by expanding cultivated areas.

Empirical literature on agriculture commercialization comprises studies examining determinants in various developing countries (i.e., Adepoju et al., 2019; Kabiti et al., 2016; Kondo et al., 2019; Mariyono, 2019; Mutyasira & Sukume, 2020; Osmani & Hossain, 2015;). Some studies focus the analysis on an indicator of interest. For instance, Chirwa and Matita (2012) investigates the relationship between food security and commercialization among smallholder farmers in Malawi. Kim et al. (2016) examines the effect of market orientation on market participation in Ethiopia. Sher et al. (2020) estimates the effect of interest-free agriculture credit on commercialization and urban-rural linkages of rice growers in Pakistan. A few studies investigate the role of productivity as parts of a broad objective of examining the determinants of market participation or factors of interest (Abu et al., 2016; Achandi & Mujawamariya, 2016; Kim et al., 2016; Namazzi et al., 2015; Olwande et al., 2015). These studies, however, ignore the endogeneity issue of productivity. The exception includes that of Rios et al. (2009) who analyze a bidirectional relationship between farm productivity and market participation based on data from Tanzania, Vietnam and Guatemala. They utilize household's age structure and its access to irrigation opportunities as instrumental variables. Additionally, Alhassan et al. (2020) study the impact of productivity on market participation simultaneously with the effect of credit on productivity in Ghana. They apply a conditional mixed process estimation technique to address the issue of endogeneity. Despite rich literature on agriculture commercialization and considerable studies on market participation of rice farmers (i.e., Achandi & Mujawamariya, 2016, Sher et al., 2020), to the author's knowledge, no studies have undertaken to provide empirical estimations of market participation of Cambodian crop farmers in general, or rice farmers in particular.

This study attempts to fill the void in existing literature by contributing in three aspects. First, it builds upon the existing studies of rice productivity in Cambodia by using a newly released Cambodia Inter-Censal Agriculture Survey in 2019 (CIAS19), allowing an investigation of the driving force behind rice productivity in 25 provinces. Second, the paper represents the first study examining the role of rice productivity in promoting market participation using Cambodia data in general or the CIAS19 in particular. Third, to determine its causal impact, the study attempts to address the endogeneity issue of agriculture productivity by employing a control function approach utilizing agriculture shocks as an instrumental variable. The data, methodology and instrument employed separate the current study from the existing ones. The results from the analysis indicates an unambiguous positive effect of productivity on commercialization. After accounting for the endogeneity of productivity, we find that a ton increase in rice yields raises the probability of market participation by 20 percent. Moreover, the level of wealth owned by a household, the share of certified modern varieties adopted, and the number of parcels and total area cultivated by a household play a significant role in improved market participation. As far as productivity is concerned, the analysis reveals evidence of the positive impact of general education; adopting aromatic paddies, sticky paddies, or modern varieties; growing on a parcel; and applying pesticide or irrigation while the agriculture shocks, number of children or parcels, and growing mixed varieties adversely affect productivity.

In a nutshell, the central finding of the current study is that commercialization of rice farming in Cambodia partly depends on the level of productivity. It suggests that a potential avenue to achieve national agriculture goal of increasing rice productions and exports, by extension commercialization, is to tackle the issue of relatively low productivity in the country. Potential policies and reforms targeting productivity may focus on education, family planning, high yield varieties, land consolidation, irrigation, and access to pesticide. Increasing these investments to enhance productivity will also result in greater market participation among Cambodian rice farmers. The paper is organized as follows: in the next section, conceptual framework is presented. In section III, we explain the data and variables. Section IV presents the methodology. Results are presented in section V, followed by the conclusion in the last section.

II. CONCEPTUAL FRAMEWORK

From the existing empirical studies on crop productivity and market participation in developing countries, a fairly large number of factors have been identified as important to productivity enhancement and successful commercialization. The study attempts to incorporate as many relevant factors suggested by the literature to the extent allowed by the data, yet within the scope of the study. Accordingly, relevant agriculture literature, not limited to rice studies, of variables included in the analysis are discussed below.

A. Rice Productivity

There is a vast literature on determining rice productivity in developing countries and studies have identified a multitude of factors that can improve rice yields and famers' income. Farmers' characteristics such as age and education have been found to be related to agriculture productivity. In particular, the relationship between age and productivity is conceivably non-linear. As farmers grow more mature and gain more experiences, productivity may increase. Once the middle age has passed, lower productivity likely ensues. In our study, however, *Age* is a variable of four categories: 20, 35, 55 and 65 with less than one percent of the farmers in the 20 age group. Hence, its relationship with productivity in our study is expected to be linear and negative. Education is related to the ability and intelligence of farmers. It is expected to improve farmers' ability to make use of yield improving agriculture technology or inputs. Wiebe et al. (2001) find that educational status of farmers is an important determinant of productivity growth.

As mentioned earlier, the constructive role of credit in improving agriculture productivity has been extensively documented in the literature. It works to reduce capital and liquidity constraints and enables farmers to acquire necessary inputs and adopt yields enhancing technology (Alhassan et al., 2020; Duong & Izumida, 2002; Misra et al., 2016). Improvements in farm size dynamics and allocative efficiency of land is another channel through which financing affects productivity (Chun, 2014). Like agriculture credits, farmers' wealth represents available resources, which can be invested in improving productivity. Thus, we expect credit and wealth to have a positive effect on productivity. Just as fertility negatively affecting workers' performance in other sectors, it may lower rice farmers' productivity as raising children requires time and resources.

Different types of paddies have been documented to results in different yield level. For instance, aromatic paddies produce higher yields and more profitable, even though non-aromatic paddies dominate rice production in Cambodia (Bunthan et al., 2018). It is estimated that the former occupies about 10 percent of the annual cultivated area, but accounts for 30 percent of total production (Eliste & Zorya, 2015). However, aromatic rice requires suitable land condition for cultivation, thus hindering its adoption. Seeds of modern varieties have higher yield potential and their investment offer one of the highest returns (Chun, 2014). However, Cambodian farmers predominantly use farm saved seeds and traditional rice varieties are preferred in the wet season. The characteristics of farmlands may also affect the level of productivity. Fragmented lands, specifically, reduce efficiency of agriculture inputs and prevent farmers from achieving economies of scale. Tube wells, for instance, are capable of drawing sufficient groundwater for a large irrigated area; however, if lands are fragmented, households need to install tube wells in every parcel, causing significant inefficiency if all the parcels are to be irrigated (Sareth et al., 2020). Thus, we would expect higher number of fragmented lands intended for rice production to be associated with lower productivity.

One of the indispensable factors in crop production is agriculture inputs. Yu and Fan (2010) show that rice production in Cambodia is far below its potential output and that fertilizer and irrigation are the main factors determining paddy supply response to the increase in paddy price. Kea et al. (2016), similarly, find that fertilizer and machinery are the main factors raising rice production. They further suggest that technical efficiency can be improved with irrigation, production technique and agricultural labor. However, they find that provinces with greater utilization of pesticide experience lower rice output. On the contrary, Chun (2014) finds that pesticide is one of the worthwhile investments providing the highest returns among other inputs. Although we expect productivity to be positively correlated with fertilizer and irrigation, its relationship with pesticide is ambiguous in the context of Cambodian rice farming.

External factors such as increases in temperature, weather-related disasters, geographical location, agronomic conditions, and distance to market are partially accountable for the level of rice efficiency. Higher temperatures adversely affect agricultural productivity (Burke et al., 2015; Lesk et al., 2016). Mishra et al. (2018) show that drought significantly affects rice production in four provinces in Cambodia. Binswanger et al. (1993) observes that improved roads contribute directly to growth in agricultural output in India. Chun (2014), on the other hand, ascertains that production inefficiency in Battambang, Kampong Thom, and Takeo is due to the absence of domestic milling. To the extent that these factors are fixed within a city or province, they can be accounted for by incorporating provincial dummies, the variables that are available in our data.

B. Market Participation

Market participation of staple farmers is generally associated with output surplus in excess of households' consumption demand, suggesting that farming households' productivity is one of the key factors impacting the probability of market participation. Rios et al. (2009) examines the relationship between productivity and market participation using comparable household data from Tanzania, Vietnam and Guatemala. They find that, controlling for differences in market access and the underlying determinants of market participation, increased household's productivity results in enhanced market participation. Abu et al. (2016), similarly, conclude that maize and groundnut farmers with higher productivity have greater participation in agricultural markets. Olwande et al. (2015) estimate output supply functions for maize, kale and milk farmers in Kenya using panel data expanding over a decade. They argue that extensive smallholder market participation can only be realized by raising productivity. This is in line with the finding from Kim et al. (2016) who determine that higher level of crop production is linked to greater likelihood of commercialization. The expectation, therefore, is that participation in the market place is positively linked to the increased level of rice productivity.

Other underlying factors that are generally found to be associated with commercialization are household characteristics. The age of a household's head, for instance, is negatively linked to the probability of market participation (Kim et al., 2016). Older farmers may be more risk-averse and concerned about food security while younger farmers are more forward-looking and better able to take advantages of the market (Abu et al., 2014, Randela et al., 2010). Male headed households are expected to participate more in the market relative to those headed by their female counterparts (Reyes et al., 2012). Education is another important factor. The literacy level of household heads, for instance, impacts their ability to interpret market signals and to capitalize on market opportunities (Namazzi et al., 2015). Households with more dependents generally consume more of their outputs, depleting their marketable excess (Ehui et al., 2003; Kim et al., 2016; Leavy & Poulton, 2007). Like labor inputs, households with more land have the capacity to generate greater surpluses as well as to expand their production to ensure adequate supply to the market.

The lack of working and investing capital may hinder farmers' ability to join the market and agriculture credit can help fulfill that need and facilitate the participation (Kim et al., 2016). Asset accumulation is an equally important contributor. The degree of commercialization is higher for wealthier households due to their ability to afford transportation cost relative to poorer households (Fafchamps & Hill, 2005). In Cambodia, varietal choices generally indicate the tendency of market participation. Fragrant rice, for instance, is produced mainly to meet market demand, particularly that from the international market, while non-aromatic rice is commonly used for home consumption and local market (Bunthan et al., 2018).

Local infrastructure, agro-ecological zones, distance to market, etc. are factors that are beyond farmer's control but contribute to their commercialization decision. Kim et al. (2016), for instance, find that access to all-weather roads enhances farmers' market participation. In addition, households cultivating in favorable agro-ecological areas have a higher degree of commercialization. Gebremedhin and Jaleta (2010) find that distance to market negatively influences market participation. Remote areas with poor transport and market infrastructures, for example, contribute to higher transaction costs, thus creating barriers to market participation (Key et al., 2000; Goetz, 1992). Chun (2014) observes that rice commercialization in Cambodia is hindered by the absence of domestic milling. Again, we account for these time invariant, locational specific factors by incorporating provincial dummies.

III. DATA AND VARIABLES

The data for this study are from CIAS19 conducted by Cambodia's National Institute of Statistics of the Ministry of Planning and the Ministry of Agriculture, Forestry and Fisheries, with assistance from the Food and Agriculture Organization, the United States Agency for International Development and the Bill and Melinda Gates Foundation. It is the country's first large-scale agriculture survey since Cambodia Agriculture Census in 2013 (CAC13). The sampling procedure is a two-stage stratified sampling based on CAC13 and the reference period is between July 2018 and June 2019. It collects data on households' characteristics, crop cultivation, raising livestock and poultry, and aquaculture and capture fishing operations from a sample of around 16,000 agricultural households and 186 large-scale agricultural enterprises in 25 provinces. The survey includes households with any size of cultivated land area or any amount of livestock or poultry inventory. However, households in core urban areas, including those living in six districts in Phnom Penh, a district in Preah Sihanouk, and a district in Siem Reap are excluded from the survey.[2]

The analysis here uses three datasets from the survey: *Main*, *Parcels_Homelots*, and *Members*. (1) *Main* includes household or holding level data of various agriculture activities such as crop cultivation, raising livestock and poultry, and aquaculture and capture fishing operations. There are 15,985 observations in the data. (2) *Parcels_Homelots* covers paddies and other crop cultivation activities at parcel or home lot level. In the survey, an agriculture holding cultivates up to 16 parcels in addition to home lots, resulting in 30,221 observations.[3] (3) *Members* contains data on the characteristics and agriculture activities of each member in a holding. One or more households constitute a holding, which comprises of up to 18 members.[4] There are 63,029 individuals identified in the dataset.

To examine the relationship between productivity and various important factors, the analysis is carried out at the parcel level using *Parcels_Homelots* dataset, supplemented with certain indicators from *Main* and *Members*.[5] It is based on a subsample of 11,434 farmlands intended for paddy production. Agriculture productivity have been defined differently in the existing studies. Rios et al. (2009) use technical efficiency as a measure of productivity. Govereh and Jayne (1999) and Strasberg et al. (1999) measure productivity as the gross value of crop production per acre. Still, others use gross production as a measure of productivity. In this study, rice productivity/yield is defined as the amount of rice output in kilograms per harvest, per hectare of cultivated land. It is a widely used measure of productivity both in qualitative and quantitative analyses. To obtain *Yield*, we count on three indicators: the total quantity of rice harvested in kg (q), the number of harvests (n), and the total cultivated area in ha (a). That is, $Yield = q/an$.[6] Within our sample, the majority of the parcels has one harvest, approximately 86%, while two harvests

[2] Reference CIAS19 full report for a detailed description of the survey methodology.

[3] In *Parcels_homelots*, 156 observations with identical values in all variables were entered twice, and 12 were entered 3 times. These can be the results of duplication errors or the data are coincidently identical. In either case, these duplicated observations are inadvertently excluded during our sample selection process.

[4] Because an overwhelming majority of the holdings are operated by a household, the word "holdings", "households", and "farmers" are used interchangeably.

[5] Because there are one or more members in each holding, only the responses from a household head in *Members* are appended to *Parcels_homelots*. In few cases where there are more than one reported heads, the observation is arbitrarily selected.

[6] There are some data inconsistencies we encounter and have made necessary adjustments. In particular, the number of harvests recorded ranges from zero to four. In addition to six recordings of four harvests, there are nine observations of continuous harvests during the reference period. There are two main rice cropping seasons in Cambodia. Additionally, farmers may plant early wet season rice to supplement existing income or stocks of rice for household consumption (Sareth et al., 2020). Thus, four or continuous harvests are unlikely; we opt to exclude these from the calculation, dropping 15 observations. After the adjustments, the average cultivated area is about 1.7 ha and paddy output stand at around 2,186 kg, significantly lower than 3,250 kg reported by the World Bank during 2018-2019 harvesting season. The difference can be attributed to the fact that our sample encompasses only agricultural households while excluding large-scale agricultural enterprises.

account for about 10% reflecting dramatic seasonal variations in rice production in Cambodia with most of the cultivation occurring in the wet season and a fraction produced in the dry season.

Table A in the Appendix provides information on relevant survey questions used in the study. Descriptive statistics of the variables in productivity equation are presented in table 1. An elaboration of these variables is needed. One of our important indicator variable is *Agriculture shocks*, identifying farmlands that have experienced shocks triggered by flood, drought, insects, crop disease, etc., during the reference period. These disasters result in lower-than-expected production, impacting rural livelihoods and food security. Approximately 37 percent of the farmlands in our sample have experienced severe agriculture shocks, of which 76 percent do not have any mitigation strategies, making them especially susceptible to climate events.

Table 1: Descriptive Statistics for Productivity at the Parcel Level

	Obs.	Mean	Std. Dev.	Min	Max
Yield in kg/hvt/ha	11419	2186	1395	0	25000
Harvest in kg	11419	4575	14835	0	700000
Area in ha	11419	1.723	3.648	0.001	150
Number of harvest	11419	1.101	0.416	0	3
Agriculture shocks	11419	0.367	0.482	0	1
Holding characteristics					
Age	11419	48.54	11.51	20	65
Male	11419	0.768	0.422	0	1
Married	11382	0.869	0.337	0	1
Illiterate	11280	0.209	0.407	0	1
Primary	11280	0.504	0.500	0	1
Secondary	11280	0.198	0.399	0	1
High school+	11280	0.089	0.284	0	1
Agriculture training	11419	0.261	0.626	0	7
Number of children	11419	0.968	1.063	0	7
Household labor	11419	894.8	1002	0	10872
Number of seniors	11419	0.274	0.565	0	3
Agriculture loan	11419	0.232	0.422	0	1
Wealth	11343	4.591	1.097	1	6
Crop characteristics					
Aromatic paddy	11419	0.171	0.377	0	1
Non-aromatic paddy	11419	0.812	0.391	0	1
Sticky paddy	11419	0.016	0.127	0	1
Mixed variety	11419	0.104	0.306	0	1
Modern variety	10412	9.822	29.18	0	100
Agriculture inputs					
Fertilizer	11419	0.881	0.324	0	1
Irrigation	11419	0.440	0.496	0	1
Pesticide	11419	0.641	0.480	0	1
Hired workers	11419	0.489	0.500	0	1
Parcel	11419	0.941	0.236	0	1
Number of parcels	11265	2.312	1.391	0	6

Note: These numbers are author's calculations based on CIAS19.

Variables representing holding characteristics are appended mostly from *Members* and two from *Main*. They vary over households but are constant across farmlands within the household. The average age of a household head in our sample is about 49. About 77 percent and 87 percent of them are male and married, respectively. Educational level of a household head is categorized into four binary variables: *Illiterate*, *Primary*, *Secondary*, and *High school+*.[7] Around half of the farmers have gone through primary education; 21 percent are illiterate; 20 percent have secondary education; and less than 10 percent have attained high school or higher education. Agriculture training is uncommon among Cambodian farmers; about 87 percent of rice farmers have never received such training. Thus, instead of agriculture training of a household head, the number of household members who have received such training is arguably a better indicator as other members also contribute to rice production. Still, around 79 percent of our sample do not report any agriculturally trained members, while 14 percent have one and 5 percent have two members.

The variable, *Number of children*, includes household members that are 14 or under. About 44 percent of the households do not have any children; 27 percent have one; and 21 percent have two. As dependents can be either children or the elderly, we also account for the *Number of Senior*, representing those members who are 65 or older. However, about 79 percent of the holdings do not have any senior members. Younger adults of farming households contribute most of the labor inputs. We define *Household labor* as the total number of hours all household members who have worked on the farms during the wet season. On average, they spend about 895 hours on the farms. Agriculture loans and farmers' wealth, proxied by residential wall materials, are analogous in a sense that they represent resources available to farmers. Accessing to agriculture loans in Cambodia remains limited; only 23 percent of the holdings report having one. *Dwelling's wall materials*, on the other hand, is an ordinal categorical variable of six, ascendingly ordered according to perceived values of residential wall materials. Approximately 76 percent have wall materials in group five or six, and about 19 percent in category three.

Rice farms producing three different varieties of paddy have been surveyed and thus included in the study: non-aromatic, aromatic, and sticky paddies, which account for about 81 percent, 17 percent, and 1.6 percent of the selected sample, respectively. Farmers may grow more than one varieties of paddy on a parcel. The indicator variable, *Mixed Variety*, shows merely 10 percent of the parcels fall in this category, while the overwhelming majority, 90 percent, grow only one variety. Often times, contract farming firms demand rice with certain standards in terms of varietal purity and yields. *Modern Variety*, on the other hand, is a continuous variable, indicating the percentage of certified modern varieties grown on a parcel. Despite the higher yield potential of modern varieties, Cambodian farmers in Takeo province usually retains their own seeds and rarely purchase rice seeds in the wet season (Chhim, 2020). This observation also emerges in our data with 89 percent forgoing modern varieties, 9 percent with full adoption, and the rest in between.

Input indicator variables include *Fertilizers*, *Irrigation*, and *Pesticide*. Approximately 88 percent, 44 percent and 64 percent of the paddy parcels surveyed are reported to have been fertilized, irrigated and applied pesticide, respectively. In terms of irrigation, it should be noted that the predominant rice producing ecosystem in Cambodia is rain-fed lowland and the use of irrigation is very limited with partial irrigation required during the wet season and full irrigation needed during the dry season. Due to the binary nature of the variable, the extent of irrigation utilization can be overstated. For instance, pond-water or groundwater may be used to supplement the water needs of rice seedlings when rainfall is inadequate early in the wet season. While such

[7] *High school+* refers to farmers with high school, bachelor's, master's, PhD, technical diploma, or others.

parcels may be recorded as irrigated in the survey, it may well be characterized as purely rain-fed paddy in other studies (i.e., Sareth et al., 2020). Similarly, the amount of pesticide and fertilizer usage in Cambodia remains low relative to similar agro-ecological zones in Thailand and Vietnam although their applications have been on the rise over the years (Kean, 2012; Theng & Koy, 2011). Because the data lacks the information on the nature and the intensity of input application, we should be more cautious when making inferences.

Parcel is an indicator variable taking value one for a parcel and zero for a home lot. About 94 percent of the cultivated lands in our selected sample are parcels. The variable is introduced to account for different association between the types of farmlands and productivity. We also control for the number of farmlands cultivated by a holding. It represents the number of fragmented lands intended for rice and/or other crops cultivation. Another important agriculture input that needs accounted for is labor. We utilize dummy variables indicating whether there are hired workers. Around 49 percent of the parcels reports having hired labor. Thanks to the rising cost of rural labor, mechanization services become more available and more widely utilized, reducing the impact of labor on paddy production.

In order to estimate the *Market Participation* regressions, we use the household-level data instead. The sample consists of 4,607 aggregate household data from *Main* dataset. It is appended with the characteristic variables of household heads from *Members*. It is also supplemented with crop and input variables re-defined and re-aggregated over all farmlands within a holding from *Parcels_Homelots* dataset.[8] The sample consists of holdings led by household heads whose main tasks during the wet/main season are rice and other crops cultivation and whose agriculture income accounts for more than half of total income. It is important to note that about half of these holdings are producing rice only while others are growing rice alongside other crops.

Agriculture commercialization is typically defined as the extent of output market participation, referred to as household commercialization index and measured as the proportion of the gross value of crop sales to that of crop production (Govereh et al., 1999; Ochieng et al. 2016; Strasberg et al., 1999). With the household commercialization index, the amount of crop sold is observed only for participating households and zero marketed amount may incorrectly suggest non-participation, especially in cross-sectional data. However, before deciding on the extent of commercialization, households presumably make a decision whether to participate in the market (Goetz, 1992). While some studies take these decisions as being simultaneous, other assume sequential (Bellemare & Barrett, 2006). The current study focuses on the first-stage participation decision. CIAS19 contains a relevant question, *"What is the main intended destination of your agricultural production?"* There are two response options: *Mainly for home consumption* and *mainly for sale*. This particular question seeks to understand farmers' intention irrespective of the amount sold during the referenced period. We argue that a variable based on the question can serve as a better indicator of farmer status. We define a binary dependent variable, *Cfarmer*, taking value one for households whose agriculture production is mainly for sale and zero otherwise. Within the selected sample, approximately 43 percent of households grow rice and other crops to meet market demand while 57 percent retains most of the outputs for consumption.

Descriptive statistics of the covariates in *Market Participation* equations are provided in table 2. It presents the number of observations, mean, and standard deviation of variables for each category of farmers. Independent sample t-test and Chi-squared test are conducted to compare the characteristics of subsistence and commercial farmers. Except for *Primary, High school+*, *Number*

[8] Refer to table 1 in the Appendix for more information on the calculation of variables.

of Children, and *Mixed paddy* the results indicate that there is a statistically significant difference between the characteristics of market participating and non-participating households. For a subsistence farmer, the average rice yield is about 2,000 kg as compared to about 2,300 kg for a commercial one. The calculation of households' yield is based on total harvests from all farmlands and the aggregate of the cultivated areas multiplied by the number of harvests. On average, the latter is significantly higher in market participating holdings than non-participating ones, 4.3 and 1.9, respectively. That is because the former cultivates on a farmland twice the size of the latter, 3.4 ha and 1.6 ha, respectively. In addition, a higher proportion of commercial holdings experiences agriculture shocks compared to that of their subsistence counterparts, about 41.3 percent as compared to 36 percent.

Table 2: Descriptive Statistics for Market Participation at the Household Level

	Subsistence			Commercial		
	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.
Yield in kg/ha***	2634	1999	1168	1967	2293	1415
Total harvest in kg***	2634	3457	11051	1967	9716	21546
Sum (area x no. of harvests)***	2638	1.862	6.917	1969	4.266	8.038
Agriculture shocks***	2638	0.359	0.480	1969	0.413	0.493
Holding characteristics						
Age***	2638	48.62	11.36	1969	47.64	11.16
Male***	2638	0.758	0.428	1969	0.861	0.346
Married***	2633	0.864	0.343	1964	0.929	0.257
Illiterate***	2604	0.215	0.411	1953	0.170	0.376
Primary	2604	0.509	0.500	1953	0.522	0.500
Secondary***	2604	0.192	0.394	1953	0.230	0.421
High school+	2604	0.084	0.278	1953	0.078	0.268
Agriculture training***	2638	0.221	0.548	1969	0.382	0.751
Number of children	2638	1.019	1.048	1969	0.969	1.060
Household labor***	2638	900.0	907.9	1969	1270	1119
Number of seniors**	2638	0.266	0.566	1969	0.220	0.511
Agriculture loan***	2638	0.222	0.416	1969	0.370	0.483
Wealth***	2617	4.565	1.091	1963	4.705	0.971
Crop characteristics						
Aromatic paddy***	2638	0.111	0.314	1969	0.201	0.401
Non-aromatic paddy***	2638	0.785	0.411	1969	0.688	0.464
Mixed paddy	2638	0.104	0.306	1969	0.111	0.314
Modern variety***	2521	7.694	24.11	1754	14.56	32.19
Agriculture inputs						
Fertilizer***	2638	1.482	0.663	1969	1.470	0.700
Irrigation***	2638	0.711	0.862	1969	0.920	0.895
Pesticide***	2638	0.996	0.842	1969	1.243	0.806
Hired workers***	2638	0.501	0.500	1969	0.561	0.496
Number of parcels***	2599	2.412	1.374	1926	2.750	1.496
Total area***	2638	1.867	4.127	1969	4.395	7.022
Pfarmers***	2638	0.357	0.175	1969	0.522	0.195

Note: Chi-square test is carried out for the categorical variables; t test is used for the continuous variables. The asterisks *, **, and *** represent significance at the 90%, 95%, and 99% confidence level, respectively.

The age of commercial farmers is, on average, one year younger than that of subsistence farmers, 48 and 49 years old, respectively. Also, 86 percent and 93 percent of the former are male and married while the numbers for the latter are 76 percent and 86 percent, respectively. Regardless of their status, the majority of farmers surveyed, 51 percent, have primary education. Both groups also have similar share, 8 percent, in the category of *High school+*. However, commercial farmers have higher percentage of primary (52 percent) and secondary schooling (23 percent) and lower share of illiterate (17 percent). *Agriculture training* is the total number of members in a household that have been agriculturally trained. Its average is higher among commercial farmers (0.38) compared to that of subsistence ones (0.22).

The number of dependents in a household may impact farmers' ability to join the market. We control for the *Number of children* (14 or under) and the *Number of seniors* (65 or older) in the regression analysis. The average number of children is about one and it is not statistically different between the two types of farmers. Irrespective of farmers' status, 81 percent in our sample do not report any senior members, merely 14 percent have one and 5 percent have two members in that age group. Other members may contribute to *Household labor*, the total number of hours that all the members have spent working on the farms. On average, subsistence households spend about 900 hours during the wet season while commercial households spend about 1,270 hours. Despite rapid growth in microfinance in Cambodia, the impact on smallholder farmers remains limited. According to table 2, the percentage of subsistence farmers receiving agriculture loans is considerably lower compared to that of commercial farmers, 22 percent and 37 percent, respectively. Besides agriculture loans, farmers' wealth may facilitate the transition from low-input agricultural systems to more productive commercial ones. The variable *Wealth* is proxied by farmers' residential wall materials of six ascending categories. The level of asset among both groups of farmers are similar, 4.7 for commercial farmers and 4.6 for subsistence farmers.

Non-aromatic paddy is a primary variety grown in Cambodia, 80 percent of the holdings in our data are cultivating this variety, and it is more prevalent among subsistence holdings. Table 2 shows that about 79 percent of subsistence farmers adopt the variety, a 10 percentage points higher than that of commercial farmers. On the other hand, the opposite is true in terms of fragrant rice, approximately 11 percent for the former and 20 percent for the latter. At the holding level, 10 percent of subsistence farmers and 11 percent of commercial farmers grow aromatic, non-aromatic, and/or sticky paddy, referred to as *Mixed Paddy*.[9] It is no secret that commercial farmers have better access to higher-quality seeds, for instance through contract farming. Within our sample, the average share of certified *Modern Variety* adopted is about 14 percent for commercial farmers and only 8 percent for subsistence farmers.

Typically, commercial farming is characterized by higher application of agriculture inputs relative to subsistence farming. Despite lack of data on the intensity of agriculture inputs, this observation, to a certain extent, is reflected in our data. Note that while *Fertilizer, Irrigation*, and *Pesticide* are dichotomous variables in *Productivity* equation, they are categorical variables with three categories at household level.[10] Commercial farmers, on average, utilize more irrigation (0.92) and pesticides (1.24), but less fertilizer (1.47) than subsistence farmers: 0.71, 1, and 1.48, respectively. Although the increased use of mechanization services in commercial households may displace the need for more labor inputs, the share of commercial farmers with hired labor is about

[9] There is one household that grows sticky paddy only in our sample; thus the indicator was dropped from the estimation.

[10] i.e. 0 means no application on any parcels; 1 means application on some parcels; and 2 means application on all parcels.

56 percent, 6 percent higher than that of subsistence ones. In addition, the number of farms cultivated by commercial farmers are relatively higher than subsistence farmers: 2.75 as compared to 2.41. Finally, *Pfarmer* represents the proportion of commercial farmers to all farmers in each province. The average is about 0.52 for commercial farmers and 0.36 among subsistence farmers.

IV. METHODS

Given the dichotomous nature of the dependent variable, we specify a discrete choice model to estimate the effect of rice yield on farmers' market participation. The underlying regression for a continuous but unobservable response variable, $Cfarmer_i^*$, representing farmer i's propensity to participate in the market, is assumed to be a linear function of the explanatory variables and the unobservable error term u_i.

$$Cfarmer_i^* = \alpha Yield_i + \beta X_i + u_i \qquad (1)$$

where $Yield_i$ is the amount of rice yields in kg/ha of cultivated land for each harvest and X_i is a vector of explanatory variables including household characteristics, crop attributes, agriculture inputs, and fixed effects. The latent response variable $Cfarmer_i^*$ is not directly observed, instead we observe that

$$Cfarmer_i = \begin{cases} 1 & if\ Cfarmer_i^* > 0 \\ 0 & if\ Cfarmer_i^* \leq 0 \end{cases}$$

That is, when $Cfarmer_i^*$ crosses 0, the observed discrete response $Cfarmer_i$ transitions from subsistence to commercial farmer. If the errors u_i are independently distributed according to a unit-normal distribution, $u_i \sim N(0,1)$, the Probit model is given by

$$Pr(Cfarmer_i = 1) = \Phi(\alpha Yield_i + \beta X_i) \qquad (2)$$

where $\Phi(.)$ is the normal distribution function. The assumption of independent u_i is violated due to the endogeneity of $Yield_i$. Potential reverse causality and unobserved factors simultaneously affecting yields and commercialization create a correlation between $Yield_i$ and u_i. Intelligent farmers, for instance, are more likely to adopt new technology and also reap the benefit that the market has to offer. Well-rounded local extension services can be conducive to both yield improvement and commercialization.

Given the binary dependent variable and a continuous endogenous regressor, we utilize control function approach to address the issue of endogeneity (i.e., Blundell and Smith, 1989; Petrin & Train, 2010; Rivers & Vuong, 1988; Train, 2003). This approach to estimation is comparable to the two-stage least square in a linear regression. The method involves, first, regressing $Yield_i$ against observed factors and instruments:

$$Yield_i = \gamma Shock_i + \delta X_i + v_i \qquad (3)$$

where X_i is the vector of variables defined earlier. $Shock_i$ is a dummy variable indicating whether the household has experienced agriculture shocks such as flood, drought, insects, or crop diseases. It is used as an instrumental variable because we reasonably believe that it strongly affects $Yield_i$

but does not directly affect market participation decision except through its impact on $Yield_i$. That is, it is independent of both error terms (i.e., v_i and u_i). The instrument can be criticized on the basis that farmers might be poorer or more risk averse and less likely to participate in the market in shock prone areas. While we cannot establish with certainty that the instrument, *Shock*, does not correlate with market participation and recognize its potential weaknesses, we argue that the effect is more likely due to its impact on yields. The error term in equation (3), v_i, are factors that affect $Yield_i$ but are not captured by the regressors. The predicted values of v_i are obtained and included in the second-stage equation (2).

Recall that there are two main exercises undertaken in the study. First, we examine factors contributing to productivity of rice farms; OLS is employed to estimate equation (3) based on a sample of 11,434 farms. Second, the paper attempts to ascertain the causal impact of productivity on farmers' market participation using observations of 4,607 households. To that end, we initially treat productivity as exogenous and estimate equation (2) using Maximum Likelihood (ML) estimators. The exogenous assumption of productivity is, subsequently, relaxed and the control function approach is employed. That is, equation (2) is re-estimated using equation (3) as the control function; Wald test of exogeneity is conducted to check if the function is necessary. For the purpose of sensitivity analysis, both ML and Newey's two-step estimations are carried out on the full samples as well as on a smaller sample of household based on the alternative definition of the dependent variable. The first-step results from Newey's two-step estimators serve as a robustness check on *Productivity* estimations. For each of these estimations, different specifications are estimated to further validate the results.

V. RESULTS

A. Rice Productivity

Table 3 reports OLS estimation results of *Productivity* equation of rice parcels. Beside agriculture shocks and provincial fixed effects, three categories of explanatory variables are included: household characteristics, crop attributes, and agriculture inputs. Three alternative specifications are presented to establish the consistency of the estimated coefficients. Column (2) excludes agriculture inputs while the first column additionally leaves out crop attributes. Inferences are drawn mainly from the last column where all the variables are accounted for.

First of all, the effect of agriculture shocks on rice productivity is negative and significant at one percent level. That is, farms that experience flood, drought, or other natural disasters, on average, produce about 247kg lower in rice yields in comparison with those that happen to evade the disasters. This represents more than 10 percent loss in yields considering the fact that the average rice yields in the sample is about 2,186kg. The negative effects of severe climate events on agriculture production is well-documented and this result further echo its impact on farmers' livelihood as severe climate events become more frequent.

Some of the household characteristics are of statistical significance. Specifically, education, be it primary, secondary or high school/higher education of household head, is an important factor. Rice parcels headed by a farmer with primary education, on average, experience approximately 180kg higher in yields than those headed by a non-educated one. The difference is even greater for farmers with secondary schooling, around 277kg more relative to the base group. On the other hand, attaining high school/higher education is similarly associated with better agriculture outcome, although it does not appear to be more beneficial relative to secondary education. This is consistent with the findings of Foster and Rosenzweig (1996) who finds that educated individuals can better manage new technologies than their less educated counterparts. *Agriculture training* represents the number of household members who have been agriculturally trained. It is statistically significant in the specifications without agriculture inputs. However, once agriculture inputs are accounted for, its significance disappears. Note that the majority of farmers lack such training: about 81 percent of the households. The situation is comparable for gender indicator. It turns insignificant in the last specification while age and marital status are not significant in any specification. The role of these individual characteristics has been documented in earlier studies. However, we fail to detect their relationship with rice productivity possibly due to limited variations of our variables in the data.

Table 3: Results for Rice Productivity at the Parcel Level

	(1)		(2)		(3)	
Dependent variable: *Yield*	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Agriculture shocks	-281.4***	(27.63)	-293.7***	(29.02)	-247.3***	(29.24)
Household characteristics						
Age	1.023	(1.344)	0.325	(1.399)	-0.147	(1.391)
Male	99.79**	(40.92)	98.79**	(43.41)	64.99	(43.38)
Married	-56.73	(49.36)	-48.47	(51.43)	-31.80	(51.59)
Primary	180.8***	(32.20)	191.1***	(33.43)	180.3***	(33.49)
Secondary	272.3***	(40.11)	285.5***	(41.71)	276.9***	(41.96)
High school+	216.6***	(51.87)	202.6***	(54.20)	203.8***	(54.29)
Agriculture training	50.22**	(20.63)	47.12**	(21.83)	28.20	(22.12)
Number of children	-49.99***	(12.65)	-53.24***	(13.22)	-48.32***	(13.15)
Household labor	0.007	(0.014)	0.006	(0.015)	0.027*	(0.015)
Number of seniors	8.854	(23.81)	-1.120	(24.63)	9.367	(24.51)
Agriculture loan	95.14***	(30.91)	95.30***	(32.44)	92.69***	(32.50)
Wealth	-10.78	(12.21)	-11.98	(12.69)	2.096	(12.59)
Crop attributes						
Aromatic paddy			88.97**	(36.89)	114.9***	(36.89)
Sticky paddy			426.4**	(191.7)	520.8***	(202.7)
Mixed variety			-174.6***	(47.37)	-144.9***	(49.16)
Modern variety			2.111***	(0.522)	1.502***	(0.506)
Agriculture inputs						
Fertilizer					32.84	(51.63)
Irrigation					393.5***	(29.47)
Pesticide					210.4***	(36.52)
Hired workers					89.64***	(29.30)
Parcel					170.5***	(60.60)
Number of parcels					-59.96***	(9.855)
Area					-10.23***	(3.828)
Constant	1906***	(114.0)	2008***	(120.3)	1614***	(140.4)
Province dummies	yes		yes		yes	
Adjusted R-squared	0.120		0.128		0.156	
Observations	11172		10194		10053	

Note: Robust standard errors are in parentheses. The asterisks *, **, and *** represent significance at the 90%, 95%, and 99% confidence level, respectively.

Two indicators for the number of dependents are included in the analysis: the minors and the seniors. While the effect of the former is negative and statistically very significant, the latter is neither consistent in sign nor significant. In particular, households with more children have lower rice productivity; an additional minor lowers rice yields by about 48kg. The aggregate household labor devoted to farming is positive and statistically significant at 10 percent level in column (3), albeit insignificant in others. Binary variable, *Agriculture loan*, reveals to be important determinant while farmers' wealth, proxied by residential wall materials, is not a meaningful contributor. That is, farms having a loan which is partly used for agriculture purposes observe about 93kg higher in rice yields compared to those without one. This finding further reaffirms the favorable effect of credit on agriculture productivity in the literature (i.e., Akudugu et al., 2012; Ali et al., 2014; Foltz, 2004; Lawal et al., 2009).

All indicators for crop attributes including *Aromatic paddy*, *Sticky paddy*, *Mixed variety*, and *Modern variety* are all statistically significant with or without controlling for agriculture inputs. Aromatic and sticky paddies produce higher yields than non-aromatic ones. On average, they offer about 115kg and 521kg higher in rice yield relative to non-aromatic rice, respectively. This is partly in line with the finding by Bunthan et al. (2018) who contend that aromatic rice gives higher yields in comparison with non-aromatic rice albeit higher production cost. The results also indicate that certified modern varieties are more productive. A one percentage point increase in the share of certified modern varieties raises rice yields by about 1.5kg. On the other hand, growing mixed varieties in a parcel results in about 145kg lower in rice yields relative to cultivating a single variety. In addition, the regression results show that agriculture inputs are essential elements determining yield level. The estimated coefficients of agriculture inputs are statistically very significant except for *Fertilizer*. Irrigation usage, pesticide application and hired labors, on average, improve rice yields by about 394kg, 210kg, and 90kg, respectively. *Fertilizer* and *Pesticide* are highly correlated which is likely to cause multicollinearity. Although *Fertilizer* is not significant in the last specification, it actually turns significant when *Pesticide* is excluded, implying the role of fertilizer in yield enhancement. Our results, to certain extent, further confirm the role of fertilizer, irrigation and pesticide as suggested by Yu and Fan (2011) and Chun (2014). The positive coefficient of *Parcel* indicates that it is more efficient to cultivate on parcels than on home lots, the difference being about 171kg. The analysis additionally reveals that fragmented land is linked to lower productivity; that is, increasing the number of parcels by one unit reduces rice yields by about 60kg. In accordance with the conclusion from many existing studies, our results confirm the inverse productivity–size relationship. A one-hectare increase in the cultivated area lowers rice yields by about 10kg. This has been attributed to cross-sectional variation in household-specific shadow prices due to factor market imperfections and the omission of soil quality variables (Barrett et al., 2010).

B. Market Participation of Cambodian Rice Farmers

With the assumption of exogeneity of productivity, we estimate the Probit model of market participation using ML estimators and report the results in table 4. The dependent variable represents rice farmers who cultivate mainly during the wet/main season regardless of their activities during the dry/low season and takes value one for those whose rice production is mainly for sale and zero otherwise. The first three columns present similar specifications as in table 3, while the last column reports average marginal effects of the third specification and serves as the basis for inferences. Our variable of interest, *Yield,* has the expected sign. Across specifications, its sign and coefficients are consistent and statistically significant at one percent level. Given the specification and assumption, we find that as rice yields increase by 1,000kg, the probability of market participation, on average, goes up by 4 percent.

Table 4: Results for Market Participation

	(1)		(2)		(3)		Marginal effects	
Dependent var.: Cfarmer	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Yield	0.0001***	(0.000)	0.0001***	(0.000)	0.0001***	(0.000)	0.00004***	(0.000)
Household characteristics								
Age	-0.004*	(0.002)	-0.003	(0.002)	-0.004*	(0.002)	-0.001*	(0.001)
Male	0.234***	(0.064)	0.265***	(0.068)	0.219***	(0.071)	0.066***	(0.021)
Married	0.108	(0.086)	0.084	(0.089)	0.059	(0.093)	0.018	(0.028)
Primary	0.080	(0.058)	0.090	(0.061)	0.040	(0.062)	0.012	(0.019)
Secondary	0.208***	(0.070)	0.195***	(0.074)	0.126*	(0.077)	0.038*	(0.023)
High school+	0.059	(0.091)	0.054	(0.096)	-0.055	(0.099)	-0.017	(0.030)
Agriculture training	0.170***	(0.033)	0.149***	(0.036)	0.134***	(0.036)	0.040***	(0.011)
Number of children	-0.058***	(0.022)	-0.057**	(0.023)	-0.053**	(0.024)	-0.016**	(0.007)
Household labor	0.0001***	(0.000)	0.0001***	(0.000)	0.0001***	(0.000)	0.00004***	(0.000)
Number of seniors	-0.047	(0.040)	-0.053	(0.042)	-0.048	(0.043)	-0.014	(0.013)
Agriculture loan	0.277***	(0.047)	0.282***	(0.050)	0.226***	(0.053)	0.068***	(0.016)
Wealth	0.083***	(0.020)	0.099***	(0.022)	0.094***	(0.022)	0.028***	(0.007)
Crop attributes								
Aromatic paddy			0.115*	(0.067)	0.085	(0.069)	0.025	(0.021)
Mixed paddy			0.144**	(0.072)	0.063	(0.078)	0.019	(0.023)
Modern variety			0.004***	(0.001)	0.004***	(0.001)	0.001***	(0.000)
Agriculture inputs								
Fertilizer					-0.077*	(0.043)	-0.023*	(0.013)
Irrigation					0.155***	(0.028)	0.046***	(0.008)
Pesticide					0.178***	(0.035)	0.053***	(0.010)
Hired workers					0.170***	(0.051)	0.051***	(0.015)
Number of parcels					0.092***	(0.018)	0.028***	(0.005)
Total area					0.027**	(0.014)	0.008**	(0.004)
Pfarmer					1.657	(1.945)	0.496	(0.583)
Constant	-0.880***	(0.187)	-1.131***	(0.203)	-2.535**	(1.071)		
Province dummies	yes		yes		yes			
Pseudo R-squared	0.176		0.190		0.222			
Observations	4499		4175		4100			

Note: Robust standard errors are in parentheses. The asterisks *, **, and *** represent significance at the 90%, 95%, and 99% confidence level, respectively.

The regression results indicate that a male headed household is about 6.6 percent more likely to participate in the market than a household led by a female counterpart and the probability is decreasing as they are aging. In particular, a 10-year older household head has a one percent lower in the chance of market participation compared to a younger one. Note that while the coefficients of *Male* are significant across specifications, *Age* is insignificant in column (2). With about 90 percent of household heads are married, the effect of marital status on commercialization is not established in table 4. However, there is a 4 percent increase in the likelihood of commercialization among household heads with secondary schooling relative to those with no prior education. Interestingly, the same magnitude of impact can also be brought about by having an additional member agriculturally trained. Other general educational indicators do not appear to exhibit significant effect. The impact of the number of dependents in a household is different depending on their age group. Just as in table 3, table 4 shows that minor dependents have negative significant effect on commercialization while senior dependents do not display significant, albeit negative, effect. On average, having an additional child lower commercialization likelihood by about 1.6 percent.

Furthermore, it is evident from our results that household and hired labor are important contributory factors: a one-hundred hours invested in rice cultivation generate about 0.4 percent boost in the probability while having hired workers augments it by 5.1 percent. Agriculture loan and wealth are both found to be positive and statistically significant at one percent level. Households with agriculture loans have around 6.8 percent higher chance of commercialization than those without any loans while a category increase in wealth induces 2.8 percent rise. This is in line with the conclusion by Kim et al. (2016) and Fafchamps and Hill (2005), who examine the role of credit and wealth, respectively. The only crop attribute that is statistically significant is the share of certified modern variety. A ten percentage point increase in the share of certified modern variety leads to about one percent increase in the likelihood of being commercial farmers. With a deviation from the assertion that in Cambodia aromatic rice is intended mainly for sale and non-aromatic rice for home consumption (Bunthan et al., 2018), the analysis, thus far, reveals no differences in cultivating various paddies (i.e., aromatic, non-aromatic, or mixed paddies) as far commercialization is concerned. Given the assumption of exogenous productivity, all agriculture input variables are shown to be essential determinants of commercialization. An increase in irrigation and pesticide used, for instance, contributes to about 4.6 percent and 5.3 percent boost in the probability, respectively. In contrast, the coefficient of *Fertilizer* is negative and statistically significant at 10 percent level. Due to the high correlation between *Pesticide* and *Fertilizer*, excluding the former rendering the latter insignificant. Thus, due caution is needed when interpreting the variable. Farmlands cultivated also positively affect commercialization decision: an additional farm lot raises the likelihood by about 2.8 percent while expanding the total area by one hectare increases it by 0.8 percent. Finally, we fail to uncover the external effect of collective presence of commercial farmers in the province/city on farmers' market participation.

The assumption of exogeneity of productivity can be violated due to potential reverse causality and unobserved factors simultaneously affecting yields and commercialization. Table 5 reports the results from ML estimation of the probability of market participation when the assumption is relaxed and the control function is incorporated. The indicator of agriculture shocks is used as an instrumental variable. The specifications and presentation in table 5 are analogous to the preceding table. The sign and coefficients of productivity are consistent with the expectation and statistically significant at one percent level in all specifications. Here, we find that the previous estimators underestimate the effect of productivity. With the control function, the probability of market participation now goes up by 20 percent as a results of a 1,000kg increase in rice yields, a sizable increase. Except for age, secondary education, number of children, fertilizer, and irrigation, other variables retain their signs and statistical significance although they have a smaller effect on rice yields in most cases. A male headed household, for instance, is now about 4.1 percent more likely to participate in the market than that headed by a female while an additional member receiving agriculture training raises the probability by approximately 3.2 percent.

We will spare repeated elaboration of significant variables and offer plausible explanations for some whose impacts have dissipated. Precisely, *Fertilizer* and *Irrigation* have turned insignificant after addressing the endogeneity of productivity. As agriculture inputs contribute directly to productivity, accounting for its corrected effect on market participation may have filtered out the actual role of *Fertilizer* and *Irrigation* as a contributor to market participation. That is, there is a limited role of these inputs in rice commercialization rather than their impact through productivity.

Table 5: Results for Market Participation with the Control Function

	(1)		(2)		(3)		Marginal effects	
Dependent var.: Cfarmer	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Yield	0.0004***	(0.000)	0.0005***	(0.000)	0.001***	(0.000)	0.0002***	(0.000)
Household characteristics								
Age	-0.003	(0.002)	-0.003	(0.002)	-0.003	(0.002)	-0.001	(0.001)
Male	0.177**	(0.075)	0.191**	(0.080)	0.145*	(0.081)	0.041*	(0.024)
Married	0.153*	(0.086)	0.131	(0.087)	0.107	(0.089)	0.030	(0.025)
Primary	0.024	(0.066)	0.016	(0.069)	-0.041	(0.070)	-0.012	(0.020)
Secondary	0.129	(0.087)	0.090	(0.091)	0.007	(0.095)	0.002	(0.027)
High school+	-0.009	(0.099)	-0.030	(0.101)	-0.145	(0.100)	-0.041	(0.028)
Agriculture training	0.156***	(0.036)	0.127***	(0.038)	0.112***	(0.039)	0.032***	(0.011)
Number of children	-0.040	(0.025)	-0.032	(0.027)	-0.025	(0.028)	-0.007	(0.008)
Household labor	0.0001***	(0.000)	0.0001***	(0.000)	0.0001***	(0.000)	0.00003***	(0.000)
Number of seniors	-0.046	(0.039)	-0.048	(0.041)	-0.049	(0.041)	-0.014	(0.012)
Agriculture loan	0.264***	(0.049)	0.248***	(0.056)	0.183***	(0.061)	0.052***	(0.018)
Wealth	0.084***	(0.020)	0.100***	(0.021)	0.086***	(0.023)	0.025***	(0.007)
Crop attributes								
Aromatic paddy			-0.007	(0.091)	-0.060	(0.098)	-0.017	(0.028)
Mixed paddy			0.088	(0.076)	-0.042	(0.088)	-0.012	(0.025)
Modern variety			0.004***	(0.001)	0.004***	(0.001)	0.001***	(0.000)
Agriculture inputs								
Fertilizer					-0.032	(0.049)	-0.009	(0.014)
Irrigation					0.046	(0.064)	0.013	(0.018)
Pesticide					0.127***	(0.048)	0.036***	(0.014)
Hired workers					0.131**	(0.055)	0.037**	(0.016)
Number of parcels					0.111***	(0.017)	0.032***	(0.005)
Total area					0.026**	(0.012)	0.007**	(0.003)
Pfarmer					0.931	(1.845)	0.265	(0.527)
Constant	-1.372***	(0.312)	-1.751***	(0.305)	-2.790***	(1.003)		
Province dummies	yes		yes		yes			
Log pseudolikelihood	-40789		-37746		-36927			
Observations	4499		4175		4100			

Note: For column 3, Wald test of exogeneity is chi2(1) = 3.55 Prob > chi2 = 0.060, rejecting the null hypothesis of no endogeneity. Robust standard errors are in parentheses. The asterisks *, **, and *** represent significance at the 90%, 95%, and 99% confidence level, respectively.

C. Wet and Dry Season Farmers

The specifications in table 5 are re-estimated using a sample of rice farmers who cultivate in both wet and dry seasons. The goal is to investigate if earlier results still hold among farmers whose livelihood primarily depends on crop cultivation. This leads to a considerable reduction in the number of observations, from 4,100 to 1,638. Table 6 reports the findings. Like earlier inferences, we base our interpretation on the estimation of the marginal effects. The indicator for productivity, instrumented by agriculture shocks, retains its sign, significance, and magnitude. Regardless of the sample size, increasing rice yields by 1,000kg raises the likelihood of commercialization by 20 percent.

Redefining the dependent variable has resulted in considerable adjustment of some covariates with regard to their significant level. The estimation results of *Wealth*, *Modern variety*, *Number of parcels*, and *Total area* are largely in agreement with those in table 5, making further elaboration unnecessary. *Primary*, *High school+*, *Aromatic paddy*, and *Mixed paddy* emerge as significant elements. However, it is rather counter intuitive that farmers with primary or high school/higher education have a lower likelihood of market participation compared to those without schooling. Similarly, growing aromatic or mixed paddy result in a lower chance of market participation relative to cultivating non-aromatic paddy. This runs counter to the finding by Bunthan et al. (2018). On the other hand, *Male*, *Agriculture training*, *Household labor*, *Agriculture loan*, *Pesticide*, and *Hired workers* are no longer important factors as far as intensive crop cultivating farmers are concerned.

Table 6: Results for Wet and Dry Season Farmers

	(1)		(2)		(3)		Marginal effects	
Dependent var: Cfarmer	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Yield	0.001***	(0.000)	0.001***	(0.000)	0.001***	(0.000)	0.0002***	(0.000)
Household characteristics								
Age	-0.003	(0.003)	-0.001	(0.004)	-0.001	(0.003)	-0.0002	(0.001)
Male	0.007	(0.101)	0.020	(0.121)	-0.061	(0.095)	-0.018	(0.028)
Married	0.225**	(0.117)	0.202*	(0.122)	0.117	(0.117)	0.034	(0.034)
Primary	-0.148*	(0.083)	-0.097	(0.095)	-0.151**	(0.075)	-0.044**	(0.022)
Secondary	-0.029	(0.110)	0.009	(0.121)	-0.077	(0.099)	-0.022	(0.029)
High school+	-0.257**	(0.120)	-0.287**	(0.122)	-0.333***	(0.117)	-0.097***	(0.034)
Agriculture training	0.055	(0.050)	0.027	(0.055)	0.001	(0.050)	0.0003	(0.015)
Number of children	0.017	(0.037)	0.021	(0.041)	0.034	(0.038)	0.010	(0.011)
Household labor	0.0001	(0.000)	0.0001	(0.000)	0.000003	(0.000)	0.000001	(0.000)
Number of seniors	0.028	(0.054)	0.022	(0.059)	0.031	(0.055)	0.009	(0.016)
Agriculture loan	0.217***	(0.070)	0.222***	(0.077)	0.117	(0.076)	0.034	(0.022)
Wealth	0.073***	(0.028)	0.084***	(0.029)	0.067**	(0.029)	0.020**	(0.009)
Crop attributes								
Aromatic paddy			-0.298***	(0.107)	-0.308***	(0.093)	-0.090***	(0.027)
Mixed paddy			-0.386***	(0.117)	-0.500***	(0.143)	-0.146***	(0.041)
Modern variety			0.005***	(0.001)	0.004***	(0.001)	0.001***	(0.000)
Agriculture inputs								
Fertilizer					0.008	(0.072)	0.002	(0.021)
Irrigation					-0.097	(0.085)	-0.028	(0.025)
Pesticide					0.006	(0.092)	0.002	(0.027)
Hired workers					0.087	(0.070)	0.025	(0.021)
Number of parcels					0.123***	(0.025)	0.036***	(0.007)
Total area					0.023**	(0.012)	0.007**	(0.003)
Pfarmer					0.672	(2.231)	0.196	(0.650)
Constant	-1.434***	(0.319)	-1.637***	(0.386)	-2.436**	(1.243)		
Province dummies	yes		yes		yes			
Log pseudolikelihood	-16894		-15366		-14973			
Observations	1830		1670		1638			

Note: For column 3, Wald test of exogeneity is chi2(1) = 7.18 Prob > chi2 = 0.007, rejecting the null hypothesis of no endogeneity. Robust standard errors are in parentheses. The asterisks *, **, and *** represent significance at the 90%, 95%, and 99% confidence level, respectively.

D. Newey's Two-Step Estimations

In this section, we use Newey's (1987) minimum $\chi 2$ estimations to obtain productivity and market participation results. The specifications in tables 7 and 8 are comparable to those in tables 3 and 5, respectively. However, while the results in table 3 are based on parcel data, those in table 7 use the aggregate data at the household level with some deviations in the definition and inclusion of certain variables as elaborated earlier. Because the previous estimations are based on a much larger sample with more variations among some indicators, they provide better estimates and are considered as the benchmark results while the latter as a robustness check.

Table 7: Results for Productivity from the Newey's Two-Step Estimation

	(1)		(2)		(3)	
Dependent var: Cfarmer	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Agriculture shocks	-274.2***	(42.16)	-258.3***	(44.15)	-211.5***	(44.48)
Household characteristics						
Age	-0.429	(1.885)	-1.040	(1.945)	-1.168	(1.943)
Male	166.4***	(57.07)	146.0***	(58.60)	112.9**	(58.71)
Married	-181.7***	(74.13)	-154.0**	(75.36)	-131.3*	(75.07)
Primary	187.5***	(49.96)	184.3***	(51.34)	178.8***	(51.53)
Secondary	249.5***	(61.10)	246.8***	(62.83)	242.7***	(63.11)
High school+	244.6***	(79.51)	226.6***	(81.88)	226.6***	(82.44)
Agriculture training	31.94	(28.72)	35.09	(30.34)	20.78	(30.44)
Number of children	-54.05***	(18.84)	-55.40***	(19.35)	-49.43***	(19.36)
Household labor	0.031	(0.020)	0.029	(0.020)	0.038*	(0.020)
Number of seniors	7.827	(35.64)	3.478	(36.46)	16.55	(36.29)
Agriculture loan	20.09	(41.69)	46.38	(43.44)	51.76	(43.73)
Wealth	-17.52	(18.05)	-25.98	(18.66)	-9.089	(18.64)
Crop attributes						
Aromatic paddy			357.9***	(58.08)	353.2***	(58.05)
Mixed paddy			163.5***	(62.21)	260.6***	(63.45)
Modern variety			-1.188*	(0.717)	-1.607**	(0.722)
Agriculture inputs						
Fertilizer					-85.47**	(36.09)
Irrigation					204.7***	(24.27)
Pesticide					71.67***	(28.85)
Hired workers					52.81	(42.07)
Number of parcels					-68.54***	(13.82)
Total area					-5.573	(3.533)
Pfarmer					944.9	(1679)
Constant	2087***	(159.2)	2123***	(168.3)	1586*	(921.6)
Province dummies	yes		yes		yes	
Adjusted R-squared	0.136		0.146		0.170	
Observations	4499		4175		4100	

Note: robust standard errors are in parentheses. The asterisks *, **, and *** represent significance at the 90%, 95%, and 99% confidence level, respectively.

Table 7 confirms the majority of the results from previous estimations. For instance, the effect of agriculture shocks remains negative and significant. With the aggregate data, it now lowers rice yields by about 212kg, a minor discrepancy. Similar results are also observed for educational indicators, the number of children, and the number of parcels in a household. The key differences are that, at the household level, the role of *Agriculture loan*, *Hired workers* and *Total area* disappear while that of age and marital status emerges. Male headed households, on average, generate about 113kg higher in rice yields than that headed by a female. Agriculture holdings led by married farmers, in contrast, experience about 131kg lower in rice yields than those headed by non-married counterparts. It should be noted that non-married farmers include the widowed, single, separated, or divorced. In addition, the nature of some variables has changed as they are aggregated, warranting particular attention. For example, there are households growing diverse paddies, entailing the inclusions of *Mixed paddy*. On the other hand, there is only one household growing sticky paddy, necessitating its exclusion from the regressions. Indicators for growing more than one variety and on parcels/home-lots are also excluded because it is not important determinant of commercialization. Households growing aromatic or mixed paddies are more productive than those growing non-aromatic paddies; the difference is about 353kg and 261kg, respectively. *Modern variety* is the average of the variable over various parcels for each household. It turns negative and statistically significant at five percent level. The unexpected sign may be attributable to measurement errors of the variable at the holding level. It is important to note that *Fertilizer*, *Irrigation* and *Pesticide* are variables of 3 categories, instead of dummies at the parcel level; thus their coefficients are not comparable with those in table 3. The analysis indicates that increasing irrigation and pesticide application by one category raises yields by about 205kg and 72kg, respectively while an increase of fertilizer usage lower productivity by 85kg, which is in contrast with the finding of existing studies. Again, this points to the problem of multicollinearity.

Table 8: Results for Market Participation from the Newey's Two-Step Estimation

	(1)		(2)		(3)	
Dependent var: Cfarmer	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Yield	0.0004**	(0.000)	0.001***	(0.000)	0.001**	(0.000)
Household characteristics						
Age	-0.003	(0.002)	-0.003	(0.002)	-0.004	(0.003)
Male	0.186***	(0.074)	0.209***	(0.079)	0.165**	(0.083)
Married	0.161*	(0.097)	0.143	(0.101)	0.122	(0.107)
Primary	0.026	(0.069)	0.017	(0.076)	-0.046	(0.084)
Secondary	0.136	(0.086)	0.099	(0.094)	0.008	(0.106)
High school+	-0.010	(0.104)	-0.033	(0.112)	-0.165	(0.124)
Agriculture training	0.164***	(0.034)	0.139***	(0.038)	0.128***	(0.039)
Number of children	-0.043*	(0.024)	-0.035	(0.027)	-0.029	(0.029)
Household labor	0.0001***	(0.000)	0.0001***	(0.000)	0.0001***	(0.000)
Number of seniors	-0.049	(0.042)	-0.053	(0.045)	-0.056	(0.047)
Agriculture loan	0.279***	(0.048)	0.272***	(0.052)	0.209***	(0.056)
Wealth	0.089***	(0.022)	0.110***	(0.024)	0.099***	(0.025)
Crop attributes						
Aromatic paddy			-0.007	(0.098)	-0.068	(0.116)
Mixed paddy			0.096	(0.080)	-0.048	(0.103)
Modern variety			0.005***	(0.001)	0.005***	(0.001)
Agriculture inputs						
Fertilizer					-0.037	(0.052)
Irrigation					0.053	(0.066)
Pesticides					0.145***	(0.042)
Hired workers					0.150***	(0.055)
Number of parcels					0.127***	(0.026)
Total area					0.030***	(0.005)
Pfarmer					1.063	(2.092)
Constant	-1.448***	(0.411)	-1.920***	(0.485)	-3.186***	(1.196)
Province dummies	yes		yes		yes	
Wald Chi-squared	836.0		792.6		851.3	
Number of obs	4499		4175		4100	

Note: For column 3, Wald test of exogeneity is chi2(1) = 3.50 Prob > chi2 = 0.06, rejecting the null hypothesis of no endogeneity. Robust standard errors are in parentheses. The asterisks *, **, and *** represent significance at the 90%, 95%, and 99% confidence level, respectively.

The estimated results between ML estimators in table 5 and Newey's two-step estimators in table 8 are analogous. In spite of minor variations in the estimated coefficients, the sign and significance are largely compatible. Note that table 8 does not present the marginal effects because they cannot be directly calculated and obtaining corresponding standard errors is a rather complex task. In addition, the marginal effects are deemed to be redundant because the purpose of obtaining two-step estimations is to carry out sensitivity analysis, which is adequate with the results in table 8. The Wald test of exogeneity of the instrumented variable is provided in the note to the tables with the control function. Tests are carried out based on the last specifications where all the relevant controls are accounted for. In all cases, they reject the null hypothesis of exogeneity, confirming the endogeneity of *Yield*.

VI. CONCLUSION

By utilizing the data from the latest 2019 survey of Cambodian agriculture, this paper attempts to investigate two important issues in Cambodia's rice sector: productivity and commercialization. Using a fairly large sample of 11,434 rice farms, it first sets out to determine key factors that influence smallholder farmers' rice productivity. Second, it investigates the effect of productivity on farmers' decision to participate in the market based on a sample of 4,607 rice farming households. The endogenous issue of productivity is addressed by incorporating the control function and utilizing agriculture shocks as an instrumental variable.

With productivity defined as the amount of rice yields in kg per harvest, per ha of cultivated land, we find that general education; adopting aromatic or sticky paddy as opposed to non-aromatic paddy; growing modern varieties, utilizing irrigation or applying pesticide; and farming on rice parcels as opposed to home lots are significant contributors to enhanced rice productivity. On the other hand, agriculture shocks, the number of children in a household, growing mixed varieties on a parcel, and the number of parcels cultivated by a household hinder it. These results are robust to alternative specifications and samples. In our investigation of the effect of productivity on commercialization, we find that the impact is statistically significant and unambiguously positive across alternative specifications, assumptions, estimation methods, and samples. After accounting for the endogeneity of productivity, the results show that the probability of market participation goes up by 20 percent as a results of a ton increase in rice yields. Moreover, the level of wealth owned by a household, the share of certified modern variety adopted, and the number of parcels and total area cultivated by a household play a significant role in improved market participation.

For Cambodia, the importance of improved rice productivity and market participation cannot be overstated given the fact that rice sector is a major player in Cambodia's agriculture. In addition, it has been plagued by chronically low rice productivity and increasing rice exports is one of the nation's top agenda. Progress made in the sector will likely contribute to poverty alleviation and rural development while supporting national agriculture goals. There are several important findings from the study that might offer fresh perspective on the issues and possibly contribute to potential policies and reforms:

- General education, be it primary, secondary, or high school or higher education, does make a difference in enhanced rice productivity. Thus, ensuring that it is accessible to the next generation of potential farmers should be part of the policy to improve rice productivity.
- Household composition matters, and so does family planning. The number of children in a household put a downward pressure on rice productivity, so family planning aiming at balancing the burden of child rearing can play a part in improved productivity.

- Aromatic or sticky paddy offers better yields than non-aromatic paddy. Encourage farmers to shift from growing non-aromatic to aromatic or sticky paddy where land condition is suitable and consumption preference is flexible.
- Support farm consolidation as fragmented farms reduce efficiency.
- Improve access to agriculture inputs such as irrigation and pesticide.
- Focus on enhancing productivity, such as investment in irrigation system or improved seeds, to expand rice commercialization and exports.

VII. ACKNOWLEDGMENT

This research was financially supported by the 50x2030 Initiative through the International Fund for Agricultural Development (IFAD). We gratefully acknowledge Dr. Smriti Tiwari for her constructive review and comments.

VIII. REFERENCES

Abu B. M., Osei-Asare, Y. B., & Wayo, S. (2014). Market participation of smallholder maize farmers in the upper west region of Ghana. *African Journal of Agricultural Research, 9*(31), 2427-2435. https://doi.org/10.5897/AJAR2014.8545

Abu, B. M., Issahaku, H., & Nkegbe, P. K. (2016). Farmgate versus market centre sales: A multi-crop approach. *Journal of Agricultural and Food Economics, 4*(21). doi:10.1186/s40100-016-0065-6

Achandi, E. L., & Mujawamariya, G. (2016). Market participation by smallholder rice farmers in Tanzania: A double hurdle analysis. *Studies in Agricultural Economics*, *118*(2), 112–115. https://doi.org/10.7896/j.1528

Adepoju, A. O., Oyegoke, O., & Amusan, A. (2019). Productivity and market participation of fluted pumpkin farmers under tropical conditions. *International Journal of Vegetable Science*, *26*(4), 323–332. https://doi.org/10.1080/19315260.2019.1631238

Ahmed, A. U., & Sampath, R. K. (1992). Effects of irrigated-induced technological change in Bangladesh rice production. *American Journal of Agricultural Economics*, *74*(1), 144–157.

Akudugu, M. A., Guo, E., & Dadzie, S. K. (2012). Adoption of modern agricultural production technologies by farm households in Ghana: What factors influence their decisions? Journal of Biology, Agriculture and Healthcare, 2(3), 1–13.

Alhassan, H., Abu, B. M., & Nkegbe, P. K. (2020). Access to credit, farm productivity and market participation in Ghana: A Conditional Mixed Process approach. *Margin—The Journal of Applied Economic Research*, *14*(2), 226–246. https://doi.org/10.1177/0973801020904490

Ali, D. A., Deininger, K., & Duponchel, M. (2014). Credit constraints and agricultural productivity: Evidence from rural Rwanda. The Journal of Development Studies, 50(5), 649–665.

Asfaw, S., Shiferaw, B., Simtowe, F., & Lipper, L. (2012). Impact of modern agricultural technologies on smallholder welfare: Evidence from Tanzania and Ethiopia. *Food Policy*, *37*(3), 283–295. https://doi.org/10.1016/j.foodpol.2012.02.013

Ashley, C., & Maxwell, S. (2001). Rethinking rural development. *Development Policy Review*, *19*(4), 395–425. https://doi.org/10.1111/1467-7679.00141

Barrett, C. B. (2008). Smallholder market participation: Concepts and evidence from Eastern and Southern Africa. *Food Policy*, *33*(4), 299–317. https://doi.org/10.1016/j.foodpol.2007.10.005

Barrett, C. B., Bellemare, M. F., & Hou, J. Y. (2010). Reconsidering conventional explanations of the inverse productivity–size relationship. *World Development*, *38*(1), 88–97.

Bellemare, M. F., & Barrett, C. B. (2006). An ordered Tobit model of market participation: Evidence from Kenya and Ethiopia. *American Journal of Agricultural Economics*, *88*(2), 324–337. https://doi.org/10.1111/j.1467-8276.2006.00861.x

Bidzakin, J. K., Fialor, S. C., Awunyo-Vitor, D., & Yahaya, I. (2018). Impact of irrigation ecology on rice production efficiency in Ghana. *Advances in Agriculture, 10*. https://doi.org/10.1155/2018/5287138

Binswanger, H. P., & Braun, J. V. (1991). Technological change and commercialization in agriculture: The effect on the poor. *The World Bank Research Observer*, *6*(1), 57–80. https://doi.org/10.1093/wbro/6.1.57

Blundell, R. W., & Smith, R. J. (1989). Estimation in a class of simultaneous equation limited dependent variable models. *The Review of Economic Studies*, *56*(1), 37–57. https://doi.org/10.2307/2297748

Braun, J. V. (1995). Agricultural commercialization: Impacts on income and nutrition and implications for policy. *Food Policy*, *20*(3), 187–202. https://doi.org/10.1016/0306-9192(95)00013-5

Bunthan, S., Takahashi, Y., & Izumida, Y. (2018). A study on Cambodian rice farming: Comparative analysis on aromatic and non-aromatic rice farming in Voatkor commune, Battambang province. *International Journal of Environmental and Rural Development*, *9*(1).

Burke, M., Hsiang, S., & Miguel, E., (2015). Global non-linear effect of temperature on economic production. *Nature, 527*, 235-239. https://doi.org/10.1038/nature15725

Chhim, C., Theng, V., & Nou, K. (2020). The production, marketing, and export of rice in Takeo. In R. Cramb (Ed.), *White gold: The commercialization of rice farming in the lower Mekong basin* (pp. 247–259). doi.org/10.1007/978-981-15-0998-8

Chirwa, E., & Matita, M. (2012). *From subsistence to smallholder commercial farming in Malawi: A case of NASFAM commercialization initiatives* (FAO Working Paper No. 037). https://www.researchgate.net/publication/236945557_From_Subsistence_to_Smallholder_Commercial_Farming_in_Malawi_A_Case_of_NASFAM_Commercialisation_Initiatives

Chun, N. (2014). *Improving rice production and commercialization in Cambodia: Findings from a farm investment climate assessment*. Retrieved from Asian Development Bank website: https://www.adb.org/publications/improving-rice-production-and-commercialization-cambodia

Duong, P. B., & Izumida, Y. (2002). Rural development finance in Vietnam: A microeconometric analysis of household surveys. *World Development, 30*(2), 319–335.

Ehui, S., Benin, S. & Paulos, A. (2003). Policy options for improving market participation and sales of smallholder livestock producers: A case of Ethiopia. Paper presented at the *International Conference on African Development Archives*. Retrieved from https://scholarworks.wmich.edu/africancenter_icad_archive/77

Eliste, P., & Zorya, S. (2015). *Cambodian agriculture in transition: opportunities and risks* (Report No. 96308). Retrieved from World Bank Group website: http://documents.worldbank.org/curated/en/805091467993504209/Cambodian-agriculture-in-transition-opportunities-and-risks

Fafchamps, M., & Hill, R. V. (2005). Selling at the farm gate or traveling to market. *American Journal of Agricultural Economics*, *87*(3), 717–734. https://doi.org/10.1111/j.1467-8276.2005.00758.x

Foltz, J. D. (2004). Credit market access and profitability in Tunisian agriculture. European Review of Agricultural Economics, 30(3), 229–240.

Food and Agriculture Organization Corporate Statistical Database. (n.d.). Retrieved August 1, 2021, https://www.fao.org/faostat/en/#home

Foster, A. & Rosenzweig, M. (1996). Technical change and human-capital returns and investments: Evidence from the green revolution. *American Economic Review 86*(4):931–953.

Gebremedhin, B., & Jaleta, M. (2010). *Commercialization of smallholders: Does market orientation translate into market participation?* (Improving Productivity and Market Success of Ethiopian Farmers Project Working Paper No. 22). Retrieved from A Repository of Agricultural Research Outputs website: https://cgspace.cgiar.org/handle/10568/3015

Goetz, S. J. (1992). A selectivity model of household food marketing behavior in sub-Saharan Africa. *American Journal of Agricultural Economics*, *74*(2), 444–452. https://doi.org/10.2307/1242498

Govereh, J., & Jayne, T. S. (2003). Cash cropping and food crop productivity: Synergies or trade-offs? *Agricultural Economics*, *28*(1), 39–50. https://doi.org/10.1111/j.1574-0862.2003.tb00133.x

Govereh, J., Jayne, T.S., & Nyoro, J. (1999). *Smallholder commercialization, interlinked markets and food crop productivity: Cross-country evidence in Eastern and Southern Africa*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.151.5366&rep=rep1&type=pdf

Huang, Q., Rozelle, S., Huang, J., Lohmar, B., & Wang, J. (2006). Irrigation, agricultural performance and poverty reduction in China. *Food Policy*, *31*(1), 30–52. https://doi.org/https://doi.org/10.1016/j.foodpol.2005.06.004

Irz, X., Lin, L., Thirtle, C., & Wiggins, S. (2001). Agricultural productivity growth and poverty alleviation. *Development Policy Review*, *19*(4), 449–466. https://doi.org/10.1111/1467-7679.00144

Jaleta, M., Gebremedhin, B., & Hoekstra, D. (2009). *Smallholder commercialization: Processes, determinants and impact* (ILRI Discussion Paper No. 18). Retrieved from A Repository of Agricultural Research Outputs website: https://cgspace.cgiar.org/handle/10568/27

Kabiti, H. M., Raidimi, N. E., Pfumayaramba, T. K., & Chauke1, P. K. (2016). Determinants of agricultural commercialization among smallholder farmers in Munyati resettlement area, Chikomba District, Zimbabwe. *Journal of Human Ecology*, *53*(1), 10–19. https://doi.org/10.1080/09709274.2016.11906951

Kea, S., Li, H., & Pich, L. (2016). Technical efficiency and its determinants of rice production in Cambodia. *Economics*, *4*(22), 1–17. https://doi.org/https://doi.org/10.3390/economies4040022

Kean, S. (2012). *The rice situation in Cambodia* (Project No. TA 7495-REG). Retrieved from Asian Development Bank website: https://www.adb.org/sites/default/files/project-document/73081/43430-012-reg-tacr-01.pdf

Key, N., Sadoulet, E., & Janvry, A. D. (2000). Transactions costs and agricultural household supply response. *American Journal of Agricultural Economics*, *82*(2), 245–259. https://doi.org/10.1111/0002-9092.00022

Kim, C. S., Abafita, J., & Atkinson, J. (2016). Smallholder commercialization in Ethiopia: Market orientation and participation. *Journal of Administrative and Business Studies*, *2*(1). https://doi.org/10.20474/-jabs2.1.4

Kimhi, A. (2006). Plot size and maize productivity in Zambia: is there an inverse relationship? *Agricultural Economics 35*(1), 1–9. https://doi.org/10.1111/j.1574-0862.2006.00133.x

Kondo, E. (2019). Market participation intensity effect on productivity of smallholder cowpea farmers: Evidence from the Northern Region of Ghana. *Review of Agricultural and Applied Economics*, *22*(1), 14–23. https://doi.org/10.15414/raae.2019.22.01.14-23

Kuppannan, P., Ramarao, R. C., Samiappan, S., & Malik, R. P. (2017). Estimating technical and irrigation water productivities in rice production in Tamil Nadu, India. *Irrigation and Drainage*, *66*(2), 163–172. https://doi.org/10.1002/ird.2099

Larson, D. F., Otsuka, K., Matsumoto, T., & Kilic, T. (2012). Should African rural development strategies depend on smallholder farms? An exploration of the inverse productivity hypothesis (Policy Research Working Paper No. 6190). Washington, DC: The World Bank.

Lawal, J. O., Omonona, B. T., Ajani, O. I. Y., & Oni, O. A. (2009). Determinants of constraints to credit access among cocoa farming households in Osun State. Pakistan Journal of Social Sciences, 6(3), 159–163.

Leavy, J., & Poulton, C. (2007). Commercializations in agriculture. *Ethiopian Journal of Economics*, *16*(1), 3–42.

Lesk, C., Rowhani, P. & Ramankutty, N. (2016). Influence of extreme weather disasters on global crop production. *Nature 529*, 84–87. https://doi.org/10.1038/nature16467

Mariyono, J. (2019). Stepping up from subsistence to commercial intensive farming to enhance welfare of farmer households in Indonesia. *Asia & the Pacific Policy Studies*, *6*(2), 246–265. https://doi.org/10.1002/app5.276

Mazumdar, D. (1965). Size of farm and productivity: A problem of Indian peasant agriculture. Economica, 32(126), 161–173.

Minten, B., & Barrett, C. B. (2008). Agricultural technology, productivity, and poverty in Madagascar. *World Development*, *36*(5), 797–822. https://doi.org/10.1016/j.worlddev.2007.05.004

Mishra, A. K., Bairagi, S., Velasco, M. L., & Mohanty, S. (2018). Impact of access to capital and abiotic stress on production efficiency: Evidence from rice farming in Cambodia. *Land Use Policy*, *79*, 215–222. https://doi.org/10.1016/j.landusepol.2018.08.016

Misra, R., Chavan, P., & Verma, R. (2016). Agricultural credit in India in the 2000s: Growth, distribution and linkages with productivity. *Margin—The Journal of Applied Economic Research*, *10*(2), 169–197.

Mutyasira, V. and Sukume, C. (2020). *Agricultural commercialization pathways, input use, and crop productivity: Evidence from smallholder farmers in Zimbabwe* (APRA Future Agricultures Consortium Working Paper No. 28). Retrieved from Institute of Development Studies website: https://opendocs.ids.ac.uk/opendocs/handle/20.500.12413/15144

Namazzi, S., Ekere, W., Kyazze, F. B., & Bareeba, F. B. (2015). Determinants of participation of smallholder farmers in marketing of grain amaranth in Kamuli District, Uganda. *Journal of Development and Agricultural Economics, 4*(5), 75–82.

Newey, W. K. (1987). Efficient estimation of limited dependent variable models with endogenous explanatory variables. *Journal of Econometrics*, *36*(3), 231–250. https://doi.org/10.1016/0304-4076(87)90001-7

Nonvide, G. M. (2017). Effect of adoption of irrigation on rice yield in the municipality of Malanville, Benin. *African Development Review*, *29*(S2), 109–120. https://doi.org/10.1111/1467-8268.12266

Ochieng, J., Kirimi, L., & Mathenge, M. (2016). Effects of climate variability and change on agricultural production: The case of small scale farmers in Kenya. *NJAS: Wageningen Journal of Life Sciences*, *77*(1), 71–78. https://doi.org/10.1016/j.njas.2016.03.005

Ogutu, S. O., & Qaim, M. (2019). Commercialization of the small farm sector and multidimensional poverty. *World Development*, *114*, 281–293. https://doi.org/10.1016/j.worlddev.2018.10.012

Olwande, J., Smale, M., Mathenge, M. K., Place, F., & Mithöfer, D. (2015). Agricultural marketing by smallholders in Kenya: A comparison of maize, kale and dairy. *Food Policy*, *52*, 22–32. https://doi.org/10.1016/j.foodpol.2015.02.002

Osmani, A. G., & Hossain, E. (2015). Market participation decision of smallholder farmers and its determinants in Bangladesh. *Ekonomika Poljoprivrede*, *62*(1), 163–179. https://doi.org/10.5937/ekopolj1501163g

Palmer-Jones, R., & Sen, K. (2003). What has luck got to do with it? A regional analysis of poverty and agricultural growth in rural India. *Journal of Development Studies*, *40*(1), 1–31. https://doi.org/10.1080/00220380412331293647

Petrin, A., & Train, K. (2010). A control function approach to endogeneity in consumer choice models. *Journal of Marketing Research*, *47*(1), 3–13. https://doi.org/10.1509/jmkr.47.1.3

Rahman, A., & Westley, J. (2001). The challenge of ending rural poverty. *Development Policy Review*, *19*(4), 553–562. https://doi.org/10.1111/1467-7679.00152

Randela, R., Alemu, Z. G., & Groenewald, J. A. (2010). Factors enhancing market participation by small-scale cotton farmers. *Agricultural Economics Research*, *Policy and Practice in Southern Africa 47*(4), 451–469. https://doi.org/10.1080/03031853.2008.9523810

Reyes, B., Donovan, C., Bernsten, R. H. & Maredia, M. K. (2012, August). Market participation and sale of potatoes by smallholder farmers in the central highlands of Angola: A Double Hurdle approach. Paper presented at the *International Association of Agricultural Economists (IAAE) Triennial Conference*. Retrieved from https://ideas.repec.org/p/ags/iaae12/126655.html

Rios, A. R., Shively, G. E., & Masters, W. A. (2009, August). Farm productivity and household market participation: Evidence from LSMS data. Paper presented at the *International Association of Agricultural Economists Conference.* Retrieved from https://ideas.repec.org/p/ags/iaae09/51031.html

Rivers, D., & Vuong, Q. H. (1988). Limited information estimators and exogeneity tests for simultaneous probit models. *Journal of Econometrics, 39*(3), 347–366. https://doi.org/10.1016/0304-4076(88)90063-2

Sareth, C., Cramb, R., & Fukai, S. (2020). The role of irrigation in the commercialization of rice farming in southern Cambodia. In R. Cramb (Ed.), *White gold: The commercialization of rice farming in the lower Mekong basin* (pp. 261–289). https://doi.org/10.1007/978-981-15-0998-8_13

Sher, A., Mazhar, S., Azadi, H., & Lin, G. (2020). Smallholder commercialization and urban-rural linkages: Effect of interest-free agriculture credit on market participation of rice growers in Pakistan. *Land, 10*(1), 7. https://doi.org/10.3390/land10010007

Strasberg, P. J., Jayne, T. S., Yamano, T., Nyoro, J., Karanja, D., & Strauss, J. (1999). *Effects of agricultural commercialization on food crop input use and productivity in Kenya* (Policy Synthesis Working Paper No. 41). https://doi.org/10.22004/ag.econ.11463

Theng, V., & Koy, R. (2011). *Review of agriculture policy and policy research* (Special Report No. 6). Retrieved from Cambodian Development Resources Institute website: https://www.cdri.org.kh/publication/review-of-agricultural-policy-and-policy-research

Train, K. (2003). *Discrete choice methods with simulation.* New York City, New York: Cambridge University Press.

Wang, Y., Huang, J., Wang, J., & Findlay, C. (2017). Mitigating rice production risks from drought through improving irrigation infrastructure and management in China. *Australian Journal of Agricultural and Resource Economics, 62*(1), 161–176. https://doi.org/10.1111/1467-8489.12241

Wiebe, K. D., Soule, M. J., & Schimmelpfennig, D. E. (2001). *Agricultural productivity for sustainable food security in Sub-Saharan Africa* (FAO Corporate Document Repository). Retrieved from www.fao.org/docrep/003/X9447E/x9447e06.htm

Yu, B., & Fan, S. (2010). Rice production response in Cambodia. *Agricultural Economics, 42*(3), 437–450. https://doi.org/10.1111/j.1574-0862.2010.00522.x

APPENDIX

Table A: Survey questions and variables

Dataset	Variable ID	Questions	Response	Defined Variables: Yield at parcel level	Defined Variables: Market participation at holding level
Parcels_homelots	s03q07a1	What was the total quantity of harvested during the last 12 months?	0 to 2,000,000 kg	Harvest in kg	Harvest in kg (Sum harvest by holding_id)
Parcels_homelots	s03q05a1	What area was planted? (in hectares)	0.0001 to 300	Area in ha	Area x nhvt (sum by holding_id)
Parcels_homelots	s03q06a	How many harvests did you have for the crop in Ref.Period?	Continuous harvest, Four harvests, No harvest, One harvests, Two harvests	Number of harvest	
Main	s05q17	Did any severe shocks hit the holding or household during Ref.Period?	Yes, No	Agriculture shocks (0/1)	Agriculture shocks (0/1)
Member	s06q03f	Age, in completed years	0-14 years, 15-24 years, 25-44 years, 45-64 years, 65 years and older	Age of 4 categories: 20, 35, 55, 65. Number of children (sum of those aged 0-14), number of senior (sum of those aged 0-14)	Age of 4 categories: 20, 35, 55, 65. Number of children (sum of those aged 0-14), number of senior (sum of those aged 0-14)
Member	s06q03a	Sex	Male, female	Male (0/1)	Male (0/1)
Member	s06q03g	Marital status	Married, Widowed, Single, Seperated or divorced	Married (0/1)	Married (0/1)
Members	s06q03h	What is the highest level of education completed?	None; Primary; Secondary; High school; Bachelor's, Master's and PhD; Other and technical diploma	Illiterate (0/1), Primary (0/1), Secondary (0/1), High school+ (0/1)	Illiterate (0/1), Primary (0/1), Secondary (0/1), High school+ (0/1)
Members	s06q03j	Has ever received any formal training on agriculture?	Yes, No	Agriculture training (the number of members who are agriculturally trained	Agriculture training (the number of members who are agriculturally trained
Main	s08q04	Did you or any member of your household have a loan?	Yes, No	Agriculture loan (0/1): Interaction between the indicators	Agriculture loan (0/1): Interaction between the indicators
Main	s08q04b	Was any part of the loan used for agricultural purposes?	Yes, No		
Main	s08q01a	What is the type of wall material used in the holder's dwelling?	Earth, Bamboo/thatch/grass/reeds, Wood/plywood, Concrete/brick/stone, Galvanised iron/aluminum/other metal, Asbestos cement sheets, Salvaged/improvised materials, Other (specify)	Dwelling's wall materials with 6 categories: (1) Earth, (2) Bamboo/thatch/grass/reeds, (3) Galvanised iron/aluminum/other metal, (4) Asbestos cement sheets, (5) Wood/plywood, (6) Concrete/brick/stone	Dwelling's wall materials with 6 categories: (1) Earth, (2) Bamboo/thatch/grass/reeds, (3) Galvanised iron/aluminum/other metal, (4) Asbestos cement sheets, (5) Wood/plywood, (6) Concrete/brick/stone
Parcels_homelots	crops_id	What crops were produced on this parcel during the last 12 months?	Non-aromatic paddy, Aromatic paddy, Sticky paddy, and 38 other crops categories	Non-aromatic paddy (0/1), Aromatic paddy (0/1), Sticky paddy (0/1)	Non-aromatic paddy (0/1), Aromatic paddy (0/1), Mixed Paddy = 1 for a holding having parcels with a combination of any

					two types of paddy, and 0 otherwise
Parcels_homelots	s03q04d	How many varieties of the crop were used?	More than one variety, NA, One variety	Mixed variety (1/0)	NA
Parcels_homelots	s03q04e	What share of the crop seed consisted in certified modern varieties?	From 0 to 100%	Share of certified modern variety	Share of certified modern variety (average by holding_id)
Parcels_homelots	s03q04a	Were fertilizers used on the crop?	Yes, No	Fertilizer (0/1)	Fertilizer: 0 = no parcels fertilized, 1 = some parcels, 2 = all parcels
Parcels_homelots	s03q04c	Was the crop irrigated during Ref.Period?	Yes, No	Irrigation (0/1)	Irrigation: 0 = no parcels irrigated, 1 = some parcels, 2 = all parcels
Parcels_homelots	s03q04b	Were pesticides used on the crop?	Yes, No	Pesticide (0/1)	Pesticide: 0 = no parcels applied pesticide, 1 = some parcels, 2 = all parcels
Main	s07q04	Holding has any paid/unpaid workers who were not part of the hh(s) of holder(s)	Yes, No	Hired workers (0/1)	Hired workers (0/1)
Parcels_homelots	homelot_parcel	Is this parcel a homelot or other parcel?	Homelot, Parcel	Parcel (0/1)	NA
Main	s03q01	How many parcels did you use for agricultural production?	From 0 to 6 and more parcel	Number of parcels	Number of parcels
Members	s07q01d	What were the main tasks on the holding during the main season?	Crop cultivation, Raising livestock and/or poultry, Aquaculture or capture fishing activity, Non-agricultural activities	NA	Cfarmer, Proportion of cfarmers (cfarmer/total farmers)
Members	s07q02d	What were the main tasks on the holding during the dry/low season?	Crop cultivation, Raising livestock and/or poultry, Aquaculture or capture fishing activity, Non-agricultural activities		
Main	s05q09	Share of agricultural income (crops, livestock, poul., aqua) in total hh income ?	None/close to 0 (Less than 10%), Less than half (10%-39%), About half (40%-59%), Most/almost all (60%-99%), All (100%)		
Main	s02q23	What is the main intended destination of your agricultural production?	Mainly for home consumption, Mainly for sale		
Members	s07q01a	Number of months worked on the holding during the wet season	From 0 to 6	Household labor (s07q01a*s07q01b*s07q01c)	Household labor (s07q01a*s07q01b*s07q01c)
Members	s07q01b	Average number of days worked per month during the wet season	From 1 to 31		
Members	s07q01c	Average number of hours worked per day during the wet season	From 1 to 12		

Main	province_id	Identified by interviewer	A list of 25 cities and provinces: Phnom Penh, Svay Rieng, Prey Veng, Takeo, Kampong Cham, Kandal, Tboung Khmum, Banteay Meanchey, Battambang, Kampong Chhnang, Kampong Thom, Pursat, Siemreap, Otdar Meanchey, Pailin, Kampot, Koh Kong, Preah Sihanouk, Kep, Mondul Kiri, Preah Vihear, Ratanak Kiri, Stung Treng, Kratie, Kampong Speu	Province dummies	Province dummies

Note: (0/1) indicates dichotomous variable equal to 1 for the included category, otherwise equal to 0 for the base category