2023

Productivity and market participation: Cambodian rice farmers

Channary Khun

Sokchea Lim

Journal of Asian Economics 88 (2023) 101646

Contents lists available at ScienceDirect

Journal of Asian Economics

journal homepage: www.elsevier.com/locate/asieco

Productivity and market participation: Cambodian rice farmers☆

Channary Khun [a], Sokchea Lim [b,*]

[a] International Fund for Agricultural Development, Italy
[b] John Carroll University, USA

ARTICLE INFO

JEL classification:
O12
O13
Q12
Q18

Keywords:
Market participation
Rice yields
Productivity
Commercialization
Cambodia

ABSTRACT

Rice is the staple food in many Asian countries. The study attempts to understand the underlying issues of low rice farm 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 farmers are more likely to commercialize their crops when productivity rises. The probability rises by about 20 % with a ton increase in rice yields. This has an important implication; policies and reforms targeting farm 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.

1. Introduction

Economic development and poverty reduction in low-income economies remain heavily dependent on the agriculture sector's performance. 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).

Cambodia is one of the major rice exporters; however, empirical studies on rice productivity in Cambodia are scant. These studies employ different datasets and examine different regions of the country (see ; Chea et al., 2020; Chhim et al., 2020; Chun, 2014; Kea et al., 2016; Mishra et al., 2018; Yu & Fan, 2011). 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

☆ We are grateful to an anonymous referee and the associate editor for constructive comments. We also thank Smriti Tiwari for the comments on the earlier drafts of the paper. Channary Khun acknowledges the financial support of the International Fund for Agricultural Development (IFAD) through the 50 × 2030 Initiative.

* Corresponding author.
E-mail addresses: khunchannary@gmail.com (C. Khun), slim@jcu.edu (S. Lim).

https://doi.org/10.1016/j.asieco.2023.101646
Received 12 August 2022; Received in revised form 23 June 2023; Accepted 27 July 2023
Available online 28 July 2023

irrigation, fertilizers, pesticides, machinery, farm size, farming technique, seed types, labor, domestic milling, education, access to 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 production which can be increased by expanding cultivated areas.

Another issue of interest is the link between farm productivity and market participation (or crop commercialization). 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 reverse causality when trying to estimate the impact of farm productivity on market participation except for a few studies. Rios et al. (2009) utilize household's age structure and its access to irrigation opportunities as instrumental variables for productivity in the data for Tanzania, Vietnam, and Guatemala. Alhassan et al. (2020) apply a conditional mixed process estimation technique to address the issue in their study on Ghana. To our knowledge, no studies have been undertaken to provide empirical estimations for the impact of farm productivity on market participation of Cambodian farmers in general, or rice farmers in particular.

The objective of the study is two-fold. First, we set out to determine key factors that influence farmers' productivity. Then, most importantly we examine to what extent an increase in land productivity affects the likelihood for a farmer to participate in the market or transition from subsistence to commercialization. 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 agricultural 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). On the other hand, rice production in the country, to a considerable extent, is cultivated by low-input, low-productivity farmers, the majority of whom are poor and subsistence or semi-subsistence.[1]

This study contributes to the existing literature 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). Second, the paper represents the first study examining the casual effect of rice productivity on market participation. While previous studies fail to address the endogeneity of agricultural productivity, we address the problem by employing a control function approach. We are able to use agricultural shocks as an instrumental variable. Finally, unlike many studies in the literature which use gross value of crop production as a measure of productivity, we calculate rice productivity as the crop yield per hectare per harvest, accounting for multiple harvests on land.

The findings from the analyses indicate that Cambodian rice farmers are more likely to participate in the market when they productivity increases. The probability of commercialization rises by 20 % with a tone increase in rice yields. As far as productivity is concerned, the results also show that yields rise with the use of pesticides or irrigation. Farmers with general education are more productive. Aromatic paddies, sticky paddies, and modern varieties provide higher yields. On the other hand, farmers with child dependents and more parcels are less productive. These results have important policy implications. A potential avenue to achieve the national agricultural goal of increasing rice production and exports, by extension of commercialization, is to tackle 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 pesticides.

The remainder of the paper is organized as follows: Section II presents the literature review which is followed by the description of the data and empirical method in Section III. The results are presented in Section IV which is followed by the robustness check in Section V. Section VI concludes the findings and discusses potential policy implications.

2. Literature review

There is a vast literature on determining rice productivity in developing countries and studies have identified a multitude of factors including education, access to credit, modern crop varieties, irrigation, pesticides, and fertilizers, that can improve rice yields. Wiebe et al. (2001) find that educational status of farmers is an important determinant of productivity growth. Access to credit reduces capital and liquidity constraints and enables farmers to acquire necessary inputs and adopt yield-enhancing technology (Alhassan et al., 2020; Duong & Izumida, 2002; Misra et al., 2016). In the assessment of farm investment climate in Cambodia, Chun (2014) concludes that seeds of modern varieties have the potential to produce higher yield and thus the investment offers one of the highest returns. Kea et al. (2016) find that fertilizers and irrigation are the main factors improving rice production in Cambodia (see also Chea et al., 2020; Yu & Fan, 2011). However, they find that provinces with greater utilization of pesticides experience lower rice output while Chun (2014) finds the contrary.

Commercialization of crops is generally associated with output surplus, suggesting that farm productivity is one of the key factors impacting the probability of market participation of the farmers. Controlling for differences in market access and the underlying determinants of market participation, Rios et al. (2009) find that increased farm productivity results in market participation of farmers in Tanzania, Vietnam, and Guatemala. A similar result is documented in Abu et al. (2016) for Ghanian maize and groundnut farmers, in Olwande et al. (2015) for Keyan maize, kale and milk farmers, and in Kim et al. (2016) for smallholder Ethiopian farmers. Lyons and Thompson (1981) also show that an increase in corn yields results in substantial excess in corn production which leads to increased exports. Other underlying factors that are generally found to be associated with commercialization are household characteristics which include age, education levels, number of dependents, farm size, or number of farmlands (see Abu et al., 2014; Ehui et al., 2003; Kim

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

Table 1
Descriptive statistics for farm productivity at the parcel level.

	Obs.	Mean	Std. Dev.	Min	Max
Yield	11,419	2186	1395	0	25,000
Total harvest	11,419	4575	14,835	0	700,000
Total area	11,419	1.723	3.648	0.001	150
Number of harvests	11,419	1.101	0.416	0	3
Agricultural shocks	11,419	0.367	0.482	0	1
Household characteristics					
Age	11,419	48.54	11.51	20	65
Male	11,419	0.768	0.422	0	1
Married	11,382	0.869	0.337	0	1
Illiterate	11,280	0.209	0.407	0	1
Primary	11,280	0.504	0.500	0	1
Secondary	11,280	0.198	0.399	0	1
High school+	11,280	0.089	0.284	0	1
Agriculture training	11,419	0.261	0.626	0	7
Number of children	11,419	0.968	1.063	0	7
Number of seniors	11,419	0.274	0.565	0	3
Household labor	11,419	894.8	1002	0	10,872
Agriculture loan	11,419	0.232	0.422	0	1
Wealth	11,343	4.591	1.097	1	6
Crop characteristics					
Aromatic paddy	11,419	0.171	0.377	0	1
Non-aromatic paddy	11,419	0.812	0.391	0	1
Sticky paddy	11,419	0.016	0.127	0	1
Mixed variety	11,419	0.104	0.306	0	1
Modern variety	10,412	9.822	29.18	0	100
Agricultural inputs					
Fertilizer	11,419	0.881	0.324	0	1
Irrigation	11,419	0.440	0.496	0	1
Pesticide	11,419	0.641	0.480	0	1
Hired workers	11,419	0.489	0.500	0	1
Parcel	11,419	0.941	0.236	0	1
Number of parcels	11,265	2.312	1.391	0	6

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

et al., 2016; Leavy & Poulton, 2007; Namazzi et al., 2015; Randela et al., 2010; Reyes et al., 2012) and infrastructure, which includes local markets, road conditions, and transportation (see Gebremedhin & Jaleta, 2010; Goetz, 1992; Key et al., 2000; Kim et al., 2016).

3. Data and empirical method

3.1. Data

The data for this study are taken from the Cambodia Inter-Censal Agriculture Survey in 2019 (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 (FAO), 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 the CAC13 and the reference period is between July 2018 and June 2019. It collects data on households' characteristics, crop cultivation, livestock and poultry, and aquaculture and capture fishing operations in 25 provinces including Banteay Meanchey, Battambang, Kampong Cham, Kampong Chhnang, Kampong Speu, Kampong Thom, Koh Kong, Kampot, Kandal, Kep, Kratie, Mondul Kiri, Otdar Meanchey, Phnom Penh, Preah Sihanouk, Preah Vihear. Prey Veng, Pailin, Pursat, Ratanak Kiri, Svay Rieng, Steng Treng, Siemreap, Thbang Khmum, and Takeo.[2]

The analyses use three datasets from the survey: *Main*, *Parcels_Homelots*, and *Members*. (1) the *Main* dataset includes household-level data of various agricultural activities such as crop cultivation, raising livestock and poultry, and aquaculture and capture fishing operations. There are 15,985 observations in the data. (2) the *Parcels_Homelots* dataset covers paddies and other crop cultivation activities at the parcel or home lot level. In the survey, a household cultivates up to 16 parcels in addition to home lots, resulting in 30,221 observations.[3] (3) the *Members* dataset contains data on the characteristics and agricultural activities of each member in the household. Because there are many members in a household, the dataset identifies 63,029 individuals. Table A in the

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

[3] In the *Parcels_homelots* dataset, 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 data cleaning process.

Appendix provides information on relevant survey questions used in the study.

3.2. Rice productivity

In examining the factors that contribute to rice farm productivity, the analysis is carried out at the parcel level using the *Parcels_Homelots* dataset, supplemented with certain indicators extracted from the *Main* and *Members* datasets.[4] Because we focus only on paddy production, the sample is reduced to 11,419 parcels. As described earlier, previous studies have used the gross value of crop production or production per acre as a measure of productivity. In this study, rice farm productivity/yield is defined as the amount of rice output in kilograms per hectare per harvest. Thus, $Yield = q/an$ where q is the total quantity of rice harvested in kg, a is cultivated area in ha, n is the number of harvests.[5] Within our sample, most parcels constitute one harvest, approximately 86 %, while two harvests 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 1 presents descriptive statistics of the variables at the parcel level. These variables include the main variable of interest, rice farm productivity, and control variables including agricultural shocks, household characteristics, crop characteristics, and agricultural inputs. The average farm productivity (*Yield*) was 2186 kg per ha per harvest, ranging from zero to as high as 25,000 kg. The agricultural shock is a binary variable which identifies the farmland that experienced disasters including flood, drought, insects, crop disease, etc. Approximately 37 % of the parcels experienced agricultural shocks.

The variables representing household characteristics are appended mostly from the *Members* dataset. They vary over households but are constant across parcels within the household. The average age of a household head is about 49. About 77 % and 87 % of them are male and married, respectively. The education level of a household head is categorized into four binary variables: illiterate, primary, secondary, and high school+ . Approximately 21 % of the household heads are illiterate or never receive formal education; around half of them have primary education; 20 % have secondary education; and 9 % have attained high school or higher education. Agricultural training is uncommon among Cambodian farmers; only about 26 % of the farmers have received such training. The average number of kids who are 14 or under is one and the average number of seniors who are 65 or above is 0.27. We define household labor as the total number of hours all household members have worked on the farms during the wet season. On average, they spend about 895 h on the farms. Agricultural loans and household wealth are appended from the *Main* dataset. The access to agricultural loans in Cambodia remains limited; only 23 % of the households reported having a loan used for agricultural purposes. The household wealth is proxied by residential wall materials, which is an ordinal categorical variable of six, ascendingly in order from earth, bamboo, aluminum, cement sheets, wood, and concrete.

The variables for crop characteristics appended from the *Parcels_Homelots* dataset. The rice farms producing three different varieties of paddies have been surveyed and thus included in the study: non-aromatic, aromatic, and sticky paddies, which account for about 81 %, 17 %, and 1.6 % of the sample, respectively. Farmers may grow more than one variety of paddy on a parcel. The indicator for the mixed variety shows merely 10 % of the parcels fall in this category. Some households employed modern varieties on their farms. We measure this variable as the percentage of certified modern varieties of crop seeds grown on a parcel. The data show that only 10 % of the parcels adopted modern varieties.

The variables for agricultural inputs are appended from the *Parcels_Homelots* and *Main* datasets. Approximately 88 %, 44 % and 64 % of the paddy parcels surveyed are reported to have been fertilized, irrigated and applied pesticide, respectively. Parcel is a binary variable which indicates that about 94 % of the farmlands are parcels and only 6 % are home lots.[6] We also control for the number of parcels used by the household for agricultural production, which ranges from zero to 6. It represents the number of fragmented lands intended for rice cultivation. On average, a household used about 2.3 parcels for rice production. The last input is a dummy variable which indicates whether there are hired workers on the farm. Approximately 49 % of the parcels reported having hired labor.

3.3. Market participation

For the household market participation, the current study focuses on the first-stage participation decision. The CIAS19 dataset 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*. We define a binary dependent variable, *Market*, taking a value of one for households whose rice production is mainly for sale (commercialization) and zero for households whose production is intended mainly for home

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

[5] 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: wet and dry seasons. Additionally, farmers may plant early wet season rice to supplement existing income or stocks of rice for household consumption (Chea 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 yield is at around 2186 kg, significantly lower than 3250 kg reported by the World Bank during 2018–2019 harvesting season.

[6] Home lot is a farm plot where the farmer lives and grows crops on while parcel is a plot only used for farming.

Table 2
Descriptive statistics for market participation at the household level.

	Subsistence			Commercialization		
	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.
Yield***	2634	1999	1168	1967	2293	1415
Total harvest***	2634	3457	11,051	1967	9716	21,546
Area x No. of harvests***	2638	1.862	6.917	1969	4.266	8.038
Agricultural shocks***	2638	0.359	0.480	1969	0.413	0.493
Household 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
Agricultural training***	2638	0.221	0.548	1969	0.382	0.751
Number of children	2638	1.019	1.048	1969	0.969	1.060
Number of seniors**	2638	0.266	0.566	1969	0.220	0.511
Household labor***	2638	900.0	907.9	1969	1270	1119
Agricultural 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
Agricultural 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
Proportion of commercial farmers***	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.

consumption (subsistence). We use household-level data from the *Main* dataset for the households who cultivated mainly during the wet/main season regardless of their activities during the dry/low season.[7] The sample consists of 4607 households. Approximately 43 % of the households grow rice for sale. The variables for crop characteristics and agricultural inputs are re-defined and aggregated over parcels within a household. Table A in the Appendix provides more information on the definition and calculation of those variables.

Table 2 presents the descriptive statistics of the variables by either subsistence or commercialization. The independent sample *t*-test and *Chi-squared* test are conducted to check the differences between subsistence and commercial farmers. Except for the variables for primary, high school+ , number of children, and mixed paddy,[8] the results indicate that there are statistically significant differences between subsistence and commercial households. The average rice yield for subsistence households is about 2000 kg per ha per harvest compared to about 2300 kg per ha per harvest for commercial households.

There are some variables worth pointing out. While only about 22 % of the subsistence households obtained agricultural loans, 37 % of the households who participated in the market did so. Non-aromatic paddy is a primary variety grown in Cambodia and it is more prevalent among subsistence households. About 79 % of subsistence farmers grew non-aromatic paddy compared to 69 % for commercial households. However, for fragrant rice, the share of commercial households that grew aromatic paddies is twice as large as that of subsistence households. The same is true for the adoption of modern variety.

3.4. Empirical method

In examining the factors determining to rice farm productivity, we estimate the following model:

$$Yield_{ij} = \mu Shock_{ij} + \pi X_{ij} + \Omega_p + v_{ij} \quad (1)$$

where $Yield_{ij}$ is the productivity of parcel j belonging to household i. X is the vector of variables including household characteristics,

[7] There are two rice-growing seasons in Cambodia. The wet or rainy season is the main period of rice cultivation. In the dry or low season, while some farmers also grow rice, others grow other crops or seek other work.

[8] Mixed paddy refers to household that grows both aromatic and non-aromatic paddy. Sticky paddy is dropped due to small number of observations.

Table 3
Rice farm productivity.

Dependent variable: Yield	(1)	(2)	(3)
Agricultural shocks	-281.4***	-293.7***	-247.3***
	(27.63)	(29.02)	(29.24)
Household characteristics			
Age	1.023	0.325	-0.147
	(1.344)	(1.399)	(1.391)
Male	99.79**	98.79**	64.99
	(40.92)	(43.41)	(43.38)
Married	-56.73	-48.47	-31.80
	(49.36)	(51.43)	(51.59)
Primary	180.8***	191.1***	180.3***
	(32.20)	(33.43)	(33.49)
Secondary	272.3***	285.5***	276.9***
	(40.11)	(41.71)	(41.96)
High school+	216.6***	202.6***	203.8***
	(51.87)	(54.20)	(54.29)
Agricultural training	50.22**	47.12**	28.20
	(20.63)	(21.83)	(22.21)
Number of children	-49.99***	-53.24***	-48.32***
	(12.65)	(13.22)	(13.15)
Number of seniors	8.854	-1.120	9.367
	(23.81)	(24.63)	(24.51)
Household labor	0.007	0.006	0.027*
	(0.014)	(0.015)	(0.015)
Agricultural loan	95.14***	95.30***	92.69***
	(30.91)	(32.44)	(32.50)
Wealth	-10.78	-11.98	2.096
	(12.21)	(12.69)	(12.59)
Crop Characteristics			
Aromatic paddy		88.97**	114.9***
		(36.89)	(36.89)
Sticky paddy		426.4**	520.8***
		(191.7)	(202.7)
Mixed variety		-174.6***	-144.9***
		(47.37)	(49.16)
Modern variety		2.111***	1.502***
		(0.522)	(0.506)
Agricultural 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)
Total area			-10.23***
			(3.828)
Constant	1906***	2008***	1614***
	(11.40)	(120.3)	(140.4)
Province dummies	yes	yes	yes
Adjusted R-squared	0.120	0.128	0.156
Observations	11,172	10,194	10,053

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

crop characteristics, and inputs. Note again that the yield analysis is conducted at the parcel level. Ω_p is the province dummies and v_{ij} is the error term.

In estimating the effect of farm productivity on the likelihood of the households participating in the market, we employ the following model:

$$Market_i = \alpha Yield_i + \beta X_i + \Omega_p + u_i \quad (2)$$

where $Market_i$ is, as discussed in the data section, a binary variable which indicates if households grew the crop for consumption

(subsistence) or sale (commercialization). Due to the natural of the dependent variable, we use Probit regressions.

In estimating Eq. (2), we encounter an endogeneity problem. The endogeneity problem can be due to the reverse causation between *Yield* and *Market*. For instance, while an increase in productivity should give farmers the surplus harvest for sale, the ability to sell the crop also allows farmers to afford better tools and technology to improve farm productivity. In addition, it is possible that an unobserved factor that is not controlled in the equation may simultaneously affect both *Yield* and *Market*, resulting in the endogeneity problem. For instance, it is also possible that a good local infrastructure can be conducive to both yield improvement and commercialization. Other factors such as the rising cost of rural labor or an incentive to work off-farm may put pressure on farm productivity and at the same time affect the farmers' likelihood to commercialize.

With the binary nature of the *Market* variable, to resolve the endogeneity of *Yield* we utilize a control function approach (for the application of this approach, see Blundell & Smith, 1989; Khun et al., 2020; Petrin & Train, 2010; for the theory, see Rivers & Vuong, 1988; Train, 2003). This approach is applicable when the estimated equation, which is a Probit equation, is a non-linear model. Because *Shock* is a force of nature, which can be assumed to be exogenous, it can serve as an instrumental variable in the control function. The shock variable is taken from the question '*Did any severe shocks hit the holding or household during the reference period?*' It takes the value of 1 for 'Yes' and zero for 'No'. The error term in Eq. (1) is used in place of *Yield* in Eq. (2). Since the analysis of market participation of households is conducted at the household level, all variables are aggregated to the household level (see Table A in the Appendix). The error term from Eq. (1) is aggregated by taking the average for the households with multiple plots for each crop.

4. Results

4.1. Rice productivity

First, we estimate the productivity regressions of Eq. (1). To ensure consistency of the results, we include household characteristics, crop attributes, and inputs one at a time. Table 3 reports the estimation results, which are quite consistent. The effect of agricultural shocks on farm productivity is negative and statistically significant at the 99 % confidence level. That is, farms that experienced flood, drought, or other natural disasters, on average, produce about 247 kg lower in yields per ha per harvest. This represents more than 10 % loss of yields considering the fact that the average rice yields in the sample is about 2186 kg. The negative effects of severe climate events on agricultural production are well-documented and this result further echoes the vulnerability of farmers' livelihood as severe climate events become more frequent.

Some of the household characteristics are of statistical and economic significance. Specifically, education, be it primary, secondary or high school/higher education of the household head, is an important factor, relative to the head with no formal education. Rice parcels headed by a farmer with primary education, on average, experience approximately 180 kg higher in yields than those headed by a non-educated one. The difference is even greater for farmers with secondary schooling, 277 kg 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 find that educated individuals can better manage new technologies than their less educated counterparts. In addition, the results show that households with more children have lower farm productivity; an additional minor lowers the yields by about 48 kg. The households with agricultural loans have higher farm productivity. This finding further reaffirms the favorable effect of credit on agricultural productivity (see Akudugu et al., 2012; Ali et al., 2014; Foltz, 2004; Lawal et al., 2009).

All variables for crop attributes including aromatic paddy, sticky paddy, mixed variety, and modern variety are all statistically significant. That is, aromatic and sticky paddies produce higher yields than non-aromatic ones. On average, they offer about 115 kg and 521 kg higher in yield, respectively. This is in line with the finding by Bunthan et al. (2018) who contend that the aromatic paddy gives higher yields than non-aromatic paddy albeit with higher production cost. The results also indicate that certified modern varieties are more productive. A 10-percentage-point increase in the share of certified modern varieties raises rice yields by about 15 kg per ha per harvest. On the other hand, growing mixed varieties results in about 145 kg lower yields.

Finally, the results show that agriculture inputs are the essential elements in determining farm productivity. Irrigation, pesticides, and hired laborers improve rice yields by about 394 kg, 210 kg, and 90 kg, respectively. The coefficient for fertilizer is not significant. However, fertilizer and pesticide are highly correlated. It turns significant when pesticide is excluded from the regression. This result, to a certain extent, further confirms the role of fertilizer, irrigation and pesticide as suggested by Yu and Fan (2011) and Chun (2014) in the Cambodian context. The positive coefficient of parcel indicates that it is more efficient to cultivate on parcels than on home lots, the difference being about 171 kg. In addition, the coefficient for the number of parcels is negative and significant at the 99 % confidence level. That is, an additional parcel reduces farm productivity by 60 kg. This suggests that fragmented land is linked to lower productivity. Finally, the result for the area variable confirms the inverse productivity–size relationship. A one-hectare increase in the cultivated area lowers farm productivity by 10 kg. 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).

4.2. Market participation

Table 4 reports the results from the estimation of Eq. (2). The specifications are analogous to those in Table 3 for consistency reasons. The marginal effect for the full specification is provided for interpretation purposes. The results show that the impact of farm productivity on market participation is positive and statistically significant at the 99 % confidence level. That is, the probability that a household participates in the market goes up by 20 % when farm productivity increases by 1000 kg. This is an economically sizable

Table 4
Household market participation.

Dependent var.: Market	(1)	(2)	(3)	Marginal effects
Yield	0.0004***	0.0005***	0.001***	0.0002***
	(0.000)	(0.000)	(0.000)	(0.000)
Household characteristics				
Age	-0.003	-0.003	-0.003	-0.001
	(0.002)	(0.002)	(0.002)	(0.001)
Male	0.177**	0.191**	0.145*	0.041*
	(0.075)	(0.080)	(0.081)	(0.024)
Married	0.153*	0.131	0.107	0.030
	(0.086)	(0.087)	(0.089)	(0.025)
Primary	0.024	0.016	-0.041	-0.012
	(0.066)	(0.069)	(0.070)	(0.020)
Secondary	0.129	0.090	0.007	0.002
	(0.087)	(0.091)	(0.095)	(0.027)
High school+	-0.009	-0.030	-0.145	-0.041
	(0.099)	(0.101)	(0.100)	(0.028)
Agricultural training	0.156***	0.127***	0.112***	0.032***
	(0.036)	(0.038)	(0.039)	(0.011)
Number of children	-0.040	-0.032	-0.025	-0.007
	(0.025)	(0.027)	(0.028)	(0.008)
Number of seniors	-0.046	-0.048	-0.049	-0.014
	(0.039)	(0.041)	(0.041)	(0.012)
Household labor	0.0001***	0.0001***	0.0001***	0.00003***
	(0.000)	(0.000)	(0.000)	(0.000)
Agricultural loan	0.264***	0.248***	0.183***	0.052***
	(0.049)	(0.056)	(0.061)	(0.018)
Wealth	0.084***	0.100***	0.086***	0.025***
	(0.020)	(0.021)	(0.023)	(0.007)
Crop attributes				
Aromatic paddy		-0.007	-0.060	-0.017
		(0.091)	(0.098)	(0.028)
Mixed paddy		0.088	-0.042	-0.012
		(0.076)	(0.088)	(0.025)
Modern variety		0.004***	0.004***	0.001***
		(0.001)	(0.001)	(0.000)
Agricultural inputs				
Fertilizer			-0.032	-0.009
			(0.049)	(0.014)
Irrigation			0.046	0.013
			(0.064)	(0.018)
Pesticide			0.127***	0.036***
			(0.048)	(0.014)
Hired workers			0.131**	0.037**
			(0.055)	(0.016)
Number of parcels			0.111***	0.032***
			(0.017)	(0.005)
Total area			0.026**	0.007**
			(0.012)	(0.003)
Proportion of commercial farmers			0.931	0.265
			(0.184)	(0.527)
Constant	-1.372***	-1.751***	-2.790***	
	(0.312)	(0.305)	(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.

magnitude and an encouraging result. In other words, if the average farm productivity of the subsistence households rises to the average level of farm productivity of the commercial farmers (from 1999 kg to 2293 kg), that implies that 155 more subsistence households would participate in the market. Thus, their income would increase as well. The Wald test of the null hypothesis that *Yield* is not endogenous is rejected, indicating that *Yield* is endogenous.

Some control variables deserve some explanation. A male headed household is about 4 % more likely to participate in the market than a household led by a female counterpart. Education does not matter in farmers' intended commercialization, but as Table 3 indicates it does for productivity. This indicates education does not affect market participation *directly*, but through its effect on productivity. The results for agricultural training, agricultural loan, and wealth are also interesting. Farmers who intend to commercialize their production are more likely to seek agricultural training and rich households or those with agricultural loans are

Table 5
Household market participation for wet and dry season farmers.

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

also more likely to commercialize their paddies. These findings are in line with the conclusion in Fafchamps and Hill (2005) and Kim et al. (2016).

The results for the crop characteristics show that there are no significant differences in the likelihood to commercialization among households that grow aromatic, non-aromatic, or mixed paddies while farmers that grow modern variety are more likely to sell their crops. A 10-percentage-point increase in the share of certified modern variety leads to about one percent increase in the likelihood of being commercial farmers.

Most agriculture input variables are shown to be the essential determinants of commercialization, except for fertilizer, and irrigation. 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

Table 6
Farm productivity from the Newey's two-step estimations.

Dependent variable: Yield	(1)	(2)	(3)	(4)
Agricultural shocks	-274.2***	-258.3***	-211.5***	-206.6***
	(42.16)	(44.15)	(44.18)	(44.29)
Household characteristics				
Age	-0.429	-1.040	-1.168	-0.510
	(1.885)	(1.945)	(1.943)	(1.925)
Male	166.4***	146.0***	112.9**	96.92*
	(57.07)	(58.60)	(58.71)	(58.16)
Married	-181.7***	-154.0**	-131.3*	-95.36
	(74.13)	(75.36)	(75.07)	(74.06)
Primary	187.5***	184.3***	178.8***	188.2***
	(49.96)	(51.34)	(51.53)	(51.11)
Secondary	249.5***	246.8***	242.7***	249.3***
	(61.10)	(62.83)	(63.11)	(62.46)
High school+	244.6***	226.6***	226.6***	231.3***
	(79.51)	(81.88)	(82.44)	(81.64)
Agricultural training	31.94	35.09	20.78	38.15
	(28.72)	(30.34)	(30.44)	(30.31)
Number of children	-54.05***	-55.40***	-49.43***	-43.89**
	(18.84)	(19.35)	(19.36)	(30.31)
Number of seniors	7.827	3.478	16.55	28.47
	(35.64)	(36.46)	(36.29)	(36.10)
Household labor	0.031	0.029	0.038*	0.035*
	(0.020)	(0.020)	(0.020)	(0.020)
Agricultural loan	20.09	46.38	51.76	57.50
	(41.69)	(43.44)	(43.73)	(43.39)
Wealth	-17.52	-25.98	-9.089	-16.32
	(18.05)	(18.66)	(18.64)	(18.48)
Crop characteristics				
Aromatic paddy		357.9***	353.2***	406.6***
		(58.08)	(58.05)	(60.94)
Mixed paddy		163.5***	260.6***	299.0***
		(62.21)	(63.45)	(63.87)
Modern variety		-1.188*	-1.607**	-1.508**
		(0.717)	(0.722)	(0.713)
Agricultural inputs				
Fertilizer			-85.47**	-78.52**
			(36.09)	(36.00)
Irrigation			204.7***	191.5***
			(24.27)	(24.08)
Pesticide			71.67***	75.27***
			(28.85)	(35.99)
Hired workers			52.81	36.74
			(42.07)	(41.69)
Number of parcels			-68.54***	-74.79***
			(13.82)	(13.63)
Total area			-5.573	-5.481
			(3.533)	(3.473)
Proportion of commercial farmers			944.9	486.0
			(1679)	(1669)
Price				-0.352***
				(0.104)
Constant	2087***	2123***	1586*	2174**
	(159.2)	(168.3)	(921.6)	(919.4)
Province dummies	yes	yes	yes	yes
Adjusted R-squared	0.136	0.146	0.170	0.173
Observations	4499	4175	4100	4006

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

inputs in rice commercialization rather than their impact through productivity.

4.3. Robustness check

To test whether or not the results are robust, we carry out the following tests.

Table 7
Household market participation from the Newey's two-step estimations.

Dependent variable.: Market	(1)	(2)	(3)	(4)
Yield	0.0004**	0.001***	0.001**	0.001**
	(0.000)	(0.000)	(0.000)	(0.000)
Household characteristics				
Age	-0.003	-0.003	-0.004	-0.004
	(0.002)	(0.002)	(0.003)	(0.003)
Male	0.186***	0.209***	0.165**	0.154*
	(0.074)	(0.079)	(0.083)	(0.083)
Married	0.161*	0.143	0.122	0.122
	(0.097)	(0.101)	(0.107)	(0.105)
Primary	0.026	0.017	-0.046	-0.058
	(0.069)	(0.076)	(0.084)	(0.088)
Secondary	0.136	0.099	0.008	-0.011
	(0.086)	(0.094)	(0.106)	(0.110)
High school+	-0.010	-0.033	-0.165	-0.157
	(0.104)	(0.112)	(0.124)	(0.128)
Agricultural training	0.164***	0.139***	0.128***	0.122***
	(0.034)	(0.038)	(0.039)	(0.041)
Number of children	-0.043*	-0.035	-0.029	-0.030
	(0.024)	(0.027)	(0.029)	(0.029)
Number of seniors	-0.049	-0.053	-0.056	-0.071
	(0.042)	(0.045)	(0.047)	(0.049)
Household labor	0.0001***	0.0001***	0.0001***	0.0001***
	(0.000)	(0.000)	(0.000)	(0.000)
Agricultural loan	0.279***	0.272***	0.209***	0.194***
	(0.048)	(0.052)	(0.056)	(0.058)
Wealth	0.089***	0.110***	0.099***	0.102***
	(0.022)	(0.024)	(0.025)	(0.026)
Crop Characteristics				
Aromatic paddy		-0.007	-0.068	-0.131
		(0.098)	(0.116)	(0.134)
Mixed paddy		0.096	-0.048	-0.106
		(0.080)	(0.103)	(0.114)
Modern variety		0.005***	0.005***	0.005***
		(0.001)	(0.001)	(0.001)
Agricultural inputs				
Fertilizer			-0.037	-0.051
			(0.052)	(0.053)
Irrigation			0.053	0.049
			(0.066)	(0.066)
Pesticides			0.145***	0.146***
			(0.042)	(0.044)
Hired workers			0.150***	0.153***
			(0.055)	(0.056)
Number of parcels			0.127***	0.129***
			(0.026)	(0.029)
Total area			0.030***	0.029***
			(0.005)	(0.005)
Proportion of commercial farmers			1.063	0.757
			(2.092)	(2.125)
Price				0.004
				(0.000)
Constant	-1.448***	-1.920***	-3.186***	-3.452***
	(0.411)	(0.485)	(1.196)	(1.294)
Province dummies	yes	yes	yes	yes
Wald Chi-squared	836.0	792.6	851.3	809.9
Observations	4499	4175	4100	4006

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.

4.4. Wet and dry season farmers

Paddy farming in Cambodia can be done in both wet and dry seasons and farmers whose livelihood primarily depends on crop cultivation do both. So, to check if the earlier results still hold among these farmer households, we restrict the sample to the households who cultivate in both wet and dry seasons. As a result, the number of observations reduces considerably, from 4100 to only 1638. Table 5 reports the results. The finding for the impact of farm productivity on market participation is consistent. The likelihood of commercialization increases by 20 % when farm productivity rises by 1000 kg.

While the coefficients for wealth, modern variety, number of parcels, and total area are largely in line with those in Table 4, those

Table A
Survey questions and variables.

Defined Variables					
Yield at parcel level	Market participation at holding level	Dataset	Variable ID	Questions	Response
$\text{Yield} = \frac{\text{Harvests}}{\text{Area x No of Harvest}}$	$\text{Yield} = \frac{\sum_i \text{Harvests}}{\sum_i \text{Area x No of Harvest}}$	Parcels_homelots	s03q07a1	What was the total quantity harvested during the last 12 months?	0–2,000,000 kg
			s03q05a1	What area was planted? (in hectares)	0.0001–300
			s03q06a	How many harvests did you have for the crop in Ref. Period?	No harvest, One harvests, Two harvests, Three harvest, Four harvests, Continuous harvest
Proportion of commercial farmers by province = Number of commercial farmers / Total farmers in the province	Market: Mainly for sale = 1 Mainly for home consumption = 0 Proportion of commercial farmers by province = Number of commercial farmers / Total farmers in the province	Main	s02q23	What is the main intended destination of your agricultural production?	Mainly for home consumption, Mainly for sale
Agriculture shocks: Yes = 1; No = 0	Agriculture shocks: Yes = 1; No = 0	Main	s05q17	Did any severe shocks hit the holding or household during Ref. Period?	Yes, No
Age of categories: 20, 35, 55, 65. Number of children = Sum of those aged 0–14 Number of seniors = Sum of those aged 65 years and older	Age of categories: 20, 35, 55, 65. Number of children = Sum of those aged 0–14 Number of seniors = Sum of those aged 65 years and older	Members	s06q03f	Age, in completed years	0–14 years, 15–24 years, 25–44 years, 45–64 years, 65 years and older
Male = 1; Female = 0	Male = 1; Female = 0	Members	s06q03a	Sex	Male, female
Married = 1; zero, otherwise.	Married = 1; zero, otherwise.	Members	s06q03g	Marital status	Married, Widowed, Single, Seperated or divorced
Illiterate: None = 1; zero, otherwise. Primary = 1; zero, otherwise. Secondary = 1; zero, otherwise. High school+ = 1 for high schools and above; zero, otherwise.	Illiterate: None = 1; zero, otherwise. Primary: Primary = 1; zero, otherwise. Secondary: Secondary = 1; zero, otherwise. High school+ = 1 for high schools and above; zero, otherwise.	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
Agricultural training = Sum of members who are agriculturally trained	Agricultural training = Sum of members who are agriculturally trained	Members	s06q03j	Has ever received any formal training on agriculture?	Yes, No
Agricultural loan: Yes, Yes = 1; zero, otherwise.	Agricultural loan: Yes, Yes = 1; zero, otherwise.	Main	s08q04	Did you or any member of your household have a loan?	Yes, No
		Main	s08q04b	Was any part of the loan used for agricultural purposes?	Yes, No
Household labor = Number of months worked x Average number of days worked per month x Average number of hours worked per day	Household labor = Number of months worked x Average number of days worked per month x Average number of hours worked per day	Members	s07q01a	Number of months worked on the holding during the wet season	From 0–6
		Members	s07q01b	Average number of days worked per month	From 1–31

(*continued on next page*)

Table A (*continued*)

Defined Variables					
Yield at parcel level	Market participation at holding level	Dataset	Variable ID	Questions	Response
				during the wet season	
		Members	s07q01c	Average number of hours worked per day during the wet season	From 1–12
Wealth: Earth = 1; Bamboo/ thatch/grass/reeds = 2; Galvanised iron/ aluminum/other metal = 3; Asbestos cement sheets = 4; Wood/ plywood = 5; Concrete/ brick/stone = 6	Wealth: Earth = 1; Bamboo/thatch/grass/ reeds = 2; Galvanised iron/ aluminum/other metal = 3; Asbestos cement sheets = 4; Wood/ plywood = 5; Concrete/brick/stone = 6	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)
Non-aromatic paddy = 1; zero, otherwise. Aromatic paddy = 1; zero, otherwise. Sticky paddy = 1; zero, otherwise.	Non-aromatic paddy = 1; zero, otherwise. Aromatic paddy = 1; zero, otherwise. Mixed Paddy = 1 for a holding having parcels with a combination of any two types of paddy; zero, otherwise.	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
Mixed variety = 1; zero, otherwise.	NA	Parcels_homelots	s03q04d	How many varieties of the crop were used?	More than one variety, NA, One variety
Modern variety = Share of certified modern variety	Modern variety = Average share of certified modern variety by the household	Parcels_homelots	s03q04e	What share of the crop seed consisted in certified modern varieties?	From 0% to 100%
Fertilizer: Yes = 1; No = 0	Fertilizer = 0 for no parcels fertilized, 1 for some parcels fertilized, or 2 for all parcels fertilized	Parcels_homelots	s03q04a	Were fertilizers used on the crop?	Yes, No
Irrigation: Yes = 1; No = 0	Irrigation = 0 for no parcels irrigated, 1 for some parcels irrigated, or 2 for all parcels irrigated	Parcels_homelots	s03q04c	Was the crop irrigated during Ref.Period?	Yes, No
Pesticide: Yes = 1; No = 0	Pesticide = 0 for no parcels applied pesticide, 1 for some parcels applied, or 2 for all parcels applied	Parcels_homelots	s03q04b	Were pesticides used on the crop?	Yes, No
Hired workers: Yes = 1; No = 0	Hired workers: Yes = 1; No = 0	Main	s07q04	Holding has any paid/unpaid workers who were not part of the hh(s) of holder(s)	Yes, No
Parcel: Parcel = 1; Home lot = 0	NA	Parcels_homelots	homelot_parcel	Is this parcel a home lot or other parcel?	Home lot, Parcel
Number of parcels	Number of parcels	Main	s03q01	How many parcels did you use for agricultural production?	From 0–6 and more parcel
Province dummies	Province dummies	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

(*continued on next page*)

Table A (*continued*)

Defined Variables					
Yield at parcel level	Market participation at holding level	Dataset	Variable ID	Questions	Response
					Sihanouk, Kep, Mondul Kiri, Preah Vihear, Ratanak Kiri, Stung Treng, Kratie, Kampong Speu

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

for primary, high school+ , aromatic paddy, and mixed paddy seem counter intuitive.

4.5. Newey's two-step estimations

Another test is to use Newey's (1987) minimum $\chi 2$ estimations to obtain both the productivity and market participation results. The specification requires that both estimations are carried out at the same household level. The results for the farm productivity and market participation equations are reported in Tables 6 and 7, respectively. The results in both tables are largely consistent, especially the impact of farm productivity on market participation. The main factors such as agricultural shock, levels of education, number of children, irrigation, pesticide, and number of parcels remain significant in determining farm productivity. A few variables such as agricultural loan, hired workers, and total area turn insignificant.[9]

In addition, one may suggest that the value of the crops can also drive market participation of the farmers. We include the price variable in the estimations and the main results are still quite consistent. The coefficient of the price variable in the market participation regression (column 4 of Table 7) is positive and statistically significant, suggesting that an increase in the price of crops encourages farmers' market participation. However, the negative sign of the price variable in the yield regression (column 4 of Table 6) is counter-intuitive. In fact, the changes of the crop price should affect farm productivity and if it does, it should be through its effects on farmers' investment in agricultural inputs.

5. 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: farm productivity and household market participation. Using a fairly large sample of rice farms, it first sets out to determine key factors that influence farm productivity. Second, it investigates the effect of farm productivity on households' decision to participate in the market. Unlike previous studies in the literature, we exploit the available information on agricultural shock as an instrument variable to address the endogeneity of farm productivity in the estimation of its impact on market participation. Another important contribution is that we calculate farm productivity as the amount of yield per ha per harvest by considering multiple harvests on a parcel in a year.

We find that if farm productivity increases by a ton, the probability that a farmer household commercializes their crops goes up by 20 %. This suggests that the government should gear up the efforts to raise the productivity of rice farms in order to boost rice commercialization. The results also show that general education, adopting aromatic or sticky paddy as opposed to non-aromatic paddy, growing modern varieties, utilizing irrigation, applying pesticides, and farming on rice parcels as opposed to home lots are the significant factors enhancing farm productivity. On the other hand, agriculture shocks, growing mixed varieties on a parcel, and fragmented farmlands adversely affect productivity. Recall that the data show that only about 44% of the parcels are irrigated and only 10% of the parcels adopted modern varieties.

For Cambodia, the importance of improved rice productivity and market participation cannot be understated 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 perspectives on the issues and possibly contribute to potential policies and reforms.

Declaration of Competing Interest

Channary Khun acknowledges the financial support from the International Fund for Agricultural Development. We declare that there is no relationship that may pose conflict of interest. Opinions expressed in this paper are solely ours and do not represent any opinions of the institutions we belong to.

[9] Note that Table 7 does not present the marginal effects because they cannot be directly calculated and obtaining corresponding standard errors is a rather complex task.

Appendix

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). https://doi.org/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

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

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.

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

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), 71–76.

Chea, S., 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). Retrieved from ⟨https://doi.org/10.1007/978–981-15–0998-8_13⟩.

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). Retrieved from ⟨doi.org/10.1007/978–981-15–0998-8⟩.

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⟩.

Datt, G., & Ravallion, M. (1998). Farm productivity and rural poverty in India. *Journal of Development Studies, 34*(4), 62–85. https://doi.org/10.1080/00220389808422529

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. In *Proceedings of the 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 from 2021, ⟨https://www.fao.org/faostat/en/#home⟩. (Accessed 1 August 2021).

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

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/10.3390/economies4040022

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

Khun, C., Lahiri, S., & Lim, S. (2020). Why do U.S. parents prefer private to foster care adoptions? The role of adoption subsidies, gender, race, and special needs. *Economic Inquiry, 58*(4), 1757–1782.

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

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.

Lyons, D. C., & Thompson, R. L. (1981). The effect of distortions in relative prices on corn productivity and exports: A cross-country study. *Journal of Rural Development, 4*, 83–102.

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.

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

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

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

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. In *Proceedings of the 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. In *Proceedings of the 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

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). Retrieved from https://doi.org/10.22004/ag.econ.11463.

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

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.

World Development Indicators (2021) *World Bank's World Development Indicators* Online. https://databank.worldbank.org/source/world-development-indicators.

Yu, B., & Fan, S. (2011). Rice production response in Cambodia. *Agricultural Economics, 42*(3), 437–450. https://doi.org/10.1111/j.1574-0862.2010.00522.x