



AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.

Farm machinery use and maize yields in China: an analysis accounting for selection bias and heterogeneity

Xiaoshi Zhou , Wanglin Ma , Gucheng Li and
Huanguang Qiu[†]

Crop production in developing and emerging countries is increasingly dependent on the usage of farm machinery. However, it remains unclear whether low-productive and high-productive farmers benefit equally from farm machinery use. To address the research gap, this study examines the potential heterogeneous effects of farm machinery use on maize yields, using an unconditional quantile regression model and survey data from China. We employ a control function approach to address the selection bias issue associated with farm machinery use. The empirical results show that the use of farm machinery significantly increases maize yields for all the selected quantiles (except for the 80th quantile); the low-productive farmers tend to benefit more from farm machinery use relative to their high-productive counterparts; and farm machinery use reduces the inequality and variability of maize yields.

Key words: farm machinery use, maize yield, UQR model, control function approach, China.

1. Introduction

Farm machinery has long been playing an important role in enhancing land productivity and promoting sustainable agricultural practices such as conservation agriculture and intensive production practices (Benin, 2015; Sims *et al.* 2016; Qiao 2017; Adu-Baffour *et al.* 2019; Takeshima *et al.* 2020; Van Loon *et al.* 2020). Farm machinery enables to substitute manual labour and traditional tools such as draft animals and hand holes, and it has the

* Xiaoshi Zhou acknowledges the research fund support for Young Teachers of School of Economics and Management, Nanjing University of Science and Technology (No. JGQN2010). Gucheng Li acknowledges the financial support from the National Natural Science Foundation of China (No. 71873050). Huanguang Qiu acknowledges the financial support from the National Natural Science Foundation of China (No. 51711520318 and No. 71673290). The authors would like to thank the two anonymous reviewers for their comments and suggestions.

[†]Xiaoshi Zhou is a Lecturer at the School of Economics & Management, Nanjing University of Science and Technology, Xiaolingwei Street, Nanjing, 210094, China. Wanglin Ma is a Senior Lecturer at the Department of Global Value Chains and Trade, Faculty of Agribusiness and Commerce, Lincoln University, Christchurch, Lincoln, 7647, New Zealand. Gucheng Li is a Professor (email: lgcabc@mail.hzau.edu.cn) at the College of Economics and Management, Huazhong Agricultural University, Wuhan, 430070, China. Huanguang Qiu is a Professor at the School of Agricultural Economics and Rural Development, Renmin University of China, Beijing, 100872, China.

potential to save production costs and reduce drudgery. Thus, understanding the association between farm machinery use and agricultural performance has been of great interest of researchers and policymakers.

Globally, farm machinery is a crucial input for agricultural production. Mechanisation is regarded as a motor for agricultural transformation. Most of farm operations in developed countries are accomplished by farm machinery, while farm machinery in other parts such as in sub-Saharan African, Latin America and South Asia is almost negligible given total cultivated land area (Van den Berg *et al.* 2007; Baudron *et al.* 2015; Adekunle *et al.* 2016; Aryal *et al.* 2019; Müller 2020; Paudel *et al.* 2020; Van Loon *et al.* 2020). For example, the proportion of farmland cultivated by tractors in Africa south of the Sahara is only 10 per cent, compared to 35 per cent in South Asia and 50 per cent in Latin America (Benin 2015). A study by Takeshima *et al.* (2017) also reveals that <8 per cent of the farms in Nepal have used mechanised tillage in hills and mountains. Kienzle *et al.* (2013) reviewed patterns and progress of agricultural mechanisation for rural development around the world and concluded that without farm machines, farmers would struggle to emerge from subsistence production. This is true as agricultural mechanisation enables farmers to increase farm productivity through production intensification and land expansion.

In China, the government has made efforts to promote the usage of farm machinery among smallholder farmers through both policy and financial instruments. The promulgation of the 'Agricultural Mechanization Promotion Law' in 2004 is one of such efforts. Like some other developing countries such as Ghana (Benin 2015), the Chinese government has also provided subsidies to increase the adoption rates of farm machines. The amount of agricultural machinery subsidies offered by the Chinese government has experienced a significant increase from 70 million yuan in 2004 to 17.4 billion yuan in 2018 (NBSC 2019).¹ As a result, the total power of farm machinery has experienced a dramatic increase, shifting from 640.3 million kilowatts in 2004 to 1,004 million kilowatts in 2018. Buying machinery services, purchasing and renting machinery are the three main options to obtain farm machines in China. The well-developed agricultural machinery service market has enabled smallholder farmers to access farm machines in agricultural production easily (Ji *et al.* 2012; Yang *et al.* 2013; Ma *et al.* 2018). Different types of farm machines (e.g. rotary cultivator, film applicator, seeder, power sprayer, fertiliser distributor, harvester, tractor, thresher and dryer) have been used by farmers to undertake various farm activities (e.g. ploughing land, applying the agricultural film, sowing seeds, spraying pesticides, applying fertiliser, harvesting, transporting, threshing and drying crops).

Several studies have investigated the association between farm machinery use and agricultural production, and they assume that both low-productive and high-productive farmers benefit equally from farm machinery use (e.g.

¹ Yuan is the Chinese currency, which 1 USD = 7.07 Yuan in April 2020.

Benin 2015; Liu *et al.* 2016; Wang *et al.* 2016; Ma *et al.* 2018; Paudel *et al.* 2019; Takeshima *et al.* 2020). For example, Benin (2015) found that farm machinery use significantly increases crop yields in Ghana. A study on Nepal by Paudel *et al.* (2019) finds that adoption of scale-appropriate mechanisation increases rice productivity by 1,110 kg/ha. In their investigation on China, Ma *et al.* (2018) also found a positive association between farm machinery use and maize yields. However, the effects of farm machinery use may be uneven among different types of farmers. Wang *et al.* (2016) showed that large-scale farming increases the applicability to mechanical production, and households cultivating large farms appear to benefit more from farm machinery use than their counterparts who cultivate small farms.

Most of the studies mentioned above have focused on the homogenous production effects of farm machinery use. However, farm machinery use may affect crop yields of low-productive and high-productive farmers differently due to the differences in household and farm-level characteristics and social-economic conditions (FAO 2013; Adekunle *et al.* 2016; Foster and Rosenzweig 2017; Adu-Baffour *et al.* 2019; Yi *et al.* 2019; Takeshima *et al.* 2020). The uneven impacts of farm machinery use stemming from the heterogeneous rural households would result in crop yield inequality among farmers. If farm machinery use has a positive impact on crop yields exclusively for high-productive farmers, the programs merely promoting agricultural mechanisation would not be a balanced and fair policy option to improve farm productivity and rural household welfare. Inappropriate mechanisation promotion policy would result in serious inequity consequences (FAO 2018). Thus, understanding the heterogeneous effects of farm machinery use on crop yields has important policy implications. However, it remains unclear whether and to what extent farm machinery use affects crop yields differently for farm households having different levels of productivity, especially in the context where farmers decide themselves (self-selection) whether they should use the machines on farms.

The primary objective of this study is, therefore, to investigate the heterogeneous effects of farm machinery use on crop yields. We attempt to make three contributions to the literature. First, we employ an unconditional quantile regression model to shed light on the heterogeneous effects of farm machinery use on maize yields. This is different from previous studies that have assumed a homogenous relationship between farm machinery use and crop production (Benin 2015; Wang *et al.* 2016; Qiao 2017; Ma *et al.* 2018; Paudel *et al.* 2019). Second, we take into account the selection bias issue of farm machinery use and address it using the control function approach proposed by Wooldridge (2015). Farmers decide themselves whether to use machinery on their farms (self-selection). Their farm machinery use decision may be affected by both observed factors (e.g. age, education and farm size) and unobserved factors (e.g. innate abilities, motivations and risk preference) (Ma *et al.* 2018; Zhang *et al.* 2019). The fact leads to the sample selection bias, an issue that should not be ignored.

Third, we add to the literature by investigating whether farm machinery use is associated with yield variability and yield inequality among maize farms, measured by sampling variance and Gini coefficient, respectively. Previous studies have examined the variability and inequality of income (Chang *et al.* 2012; Evans *et al.* 2019), consumption (Tamkoç and Torul, 2020), economic development (Gibson 2018) and health outcomes (Lai *et al.* 2008; Hoque *et al.* 2019). For example, using sampling variance and Gini coefficient, Chang *et al.* (2012) examined the impact of eco-label use on income variability and income inequality of aquaculture producers in Taiwan, and Hoque *et al.* (2019) examined the health inequality among Bangladeshi children of age 6–59 months. However, to date, no previous studies have taken into account the association between agricultural technology adoption and crop yield variability and inequality.

Our econometric analysis utilises household survey data collected from smallholder maize farmers in China. Globally, China is the second-largest maize-producing country, and its maize output is only behind the United States (FAOSTAT). Over the past three decades, the total maize output in China has dramatically increased from 62.6 million tonnes in 1980 to 257.17 million tonnes in 2018 (NBSC 2019). Despite the significant increase in the total output of maize production, the maize market in China is shifting from an oversupply to a short supply due to the rapid economic development and the growing demand for maize. In 2018, the imported maize by China reached 3.52 million tons, which is mainly used to supplement the shortage of the domestic maize supply for human food, animal feed and industrial production. The shortage of maize supply in China can be attributed partly to the low productivity of maize production. The maize productivity in China is significantly lower than that in other major maize-producing countries such as the United States and Canada over the years (see Figure S1). For example, in 2018, the outputs of maize in the United States and Canada are 11,864 kg/ha and 9,705 kg/ha, respectively, while that in China is only 6,104.2 kg/ha.

The rest of the paper is organised as follows. Section 2 presents the theoretical framework and econometric approach, and it is followed by Section 3 that presents the data and descriptive statistics. Section 4 presents and discusses the empirical results, while Section 5 concludes with policy implications.

2. Theoretical framework and econometric approach

2.1 Theoretical framework

Following previous studies on agricultural technology adoption (e.g. Song *et al.* 2018; Ma and Abdulai 2019; Tufa *et al.* 2019; Amadu *et al.* 2020), this study approaches farm machinery use as a choice problem within a random utility maximisation framework. The random utility theory assumes that a farmer who seeks to maximise utility will compare the utility from using the

machinery technology (U_U) with that from not using the technology (U_N). The farmer will choose to use the technology if the utility difference (M^*) is greater than zero, that is $M^* = U_U - U_N > 0$. The unobservable net utility M^* can be expressed as function of observable elements in the following latent variable model:

$$M_i^* = \alpha_i L_i + \tau_i \text{ with } M_i = 1(M_i^* > 0) \quad (1)$$

where M_i is a binary indicator that equals 1 for farmer i who uses farm machinery and 0 otherwise; α_i is a vector of parameters to be estimated; L_i is a vector of household and farm characteristics; and τ_i is an error term.

Previous studies have shown that farm machinery use can affect farm outputs (Benin 2015; Abass *et al.* 2017; Ma *et al.* 2018; Paudel *et al.* 2020). To link the farm machinery use decision with farm outputs such as maize yields, we assume that a risk-neutral farmer i maximise net returns, π , subject to competitive inputs and output markets and a single-output technology. This may be expressed as follows:

$$\text{Max } \pi = P_Q Q(P_I, L) - P_I I \quad (2)$$

where P_Q is the market price for maize and Q is the expected maize output level; P_I is a vector of input prices and I represents a vector of input quantities. The net returns of maize production can be expressed as a function of farm machinery use M , input and output prices, and farm and household characteristics as follows:

$$\pi = \pi(M, P_I, P_Q, L_i) \quad (3)$$

Application of Hotelling's Lemma with respect to the market price for maize to Equation (3) yields the reduced form equation for maize output supply:

$$\frac{d\pi}{dP_Q} = Q = Q(M, P_I, P_Q, L_i) \quad (4)$$

The specification in Equation (4) shows that the level of maize output is affected by farm machinery use, inputs and output prices, and farm and household characteristics. In the following, we apply a reduced form approach to relate farm machinery use to maize production for estimation purposes.

2.2 Selection bias issue of farm machinery use

The empirical model for examining the relationship between farm machinery use and maize yields can be specified as a general maize production function:

$$\ln Y_i = \alpha_0 + \alpha_m M_i + \sum_{j=1}^4 \alpha_j \ln X_{ij} + \sum_{n=1}^{11} \alpha_n A_{in} + \varepsilon_i \quad (5)$$

where \ln denotes natural logarithm; Y_i refers to the outcome variables such as maize yields per mu (1 mu = 1/15 hectare) for household i ; M_i refers to the farm machinery use status; X_{ij} represents j ($j=4$) types of production inputs including farm size, pesticide, fertiliser and seed; A_{in} refers to n ($n=11$) explanatory variables that represent household and farm-level characteristics (e.g. age, gender, education, household size, extension contact and soil quality) that are expected to affect maize yields; α_0 is a constant; α_m , α_j and α_n are the corresponding parameters to be estimated; and ε_i is a classical random error term. In particular, α_m is used to capture the effect of farm machinery use on maize yields. If $\alpha_m > 0$ and is statistically significant, this would suggest that farm machinery use increases the maize yields and vice versa.

In Equation (5), if farm machinery use, M_i , is exogenous, we can use a simple ordinary least squares (OLS) regression model to analyse the impact of farm machinery use on maize yields. However, as discussed earlier, both observed factors and unobserved factors may influence farmers' decision to use machines on their farms. For example, farmers with better managerial skills may prefer to use innovative agricultural technologies such as farm machinery. Farmers' managerial capabilities cannot be observed by analysts directly, but they still affect farm machinery use and maize yields. The fact leads to potential selection bias issue of variable M_i in Equation (5). Failing to account for the selection bias issue that is associated with farm machinery use variable would produce biased and inconsistent estimates.

To address the selection bias issue and to achieve more consistent results, we follow previous studies (Wooldridge, 2015) and apply a control function approach that combines farm machinery use equation with maize production function equation in the present study.

2.3 Control function approach

The control function approach estimates a farm machinery use equation at the first stage and a maize production equation at the second stage. In particular, the inverse Mills ratio is predicted from the first-stage estimation and then added as an extra regressor in the second-stage regression to account for the self-selection bias issue of farm machinery use.

2.3.1 First-stage analysis: modelling the decision of farm machinery use

Farmers' decision to use farm machinery can be estimated by Equation (1). We assume that $L_i = Z_i + S_i$, with Z_i representing a vector of exogenous determinants of farm machinery use and S_i representing an instrumental

variable. Therefore, Equation (1) can be rewritten as follows:

$$M_i^* = \beta_z Z_i + \beta_s S_i + \mu_i \text{ with } M_i = 1 (M_i^* > 0) \quad (6)$$

where β_z and β_s are corresponding parameters to be estimated; and μ_i is a classical random error term.

The probability of farm machinery use can be expressed as:

$$\Pr(M_i = 1) = \Pr(\beta_z Z_i + \beta_s S_i + \mu_i > 0) = 1 - F(-\beta_z Z_i - \beta_s S_i) \quad (7)$$

where $F(\cdot)$ is the cumulative distribution function which depends on the error term μ_i . The parameters, β_z and β_s , can be obtained by estimating a dichotomous model (i.e. probit model in this study) to maximise the following log-likelihood function (Greene 2018):

$$\ln L = \sum_{i=1}^n M_i \times \Phi(\beta_z Z_i + \beta_s S_i) + (1 - M_i)[1 - \Phi(\beta_z Z_i - \beta_s S_i)] \quad (8)$$

where $\Phi(\cdot)$ is the cumulative density function.

In the present study, we employ a variable representing the smartphone use status of the household head as an instrumental variable as defined by S_i in Equations (6–8). The purpose of using an instrumental variable in the farm machinery use equation is to account for unobserved factors that may bias the impact of farm machinery use on maize yields in the second-stage estimation of the control function approach (Wooldridge 2015). It is rational to use the household head's smartphone use status as an instrumental variable. The existing literature has shown that the use of information and communication technologies such as mobile phones/smartphones facilitates agricultural technology adoption (Baumüller 2012; Aker and Ksoll 2016; Maredia *et al.* 2018; Michels *et al.* 2020). Baumüller (2012) pointed out that the use of mobile phones can facilitate agricultural technology adoption among farmers in developing countries because it helps overcome some of the obstacles to technology adoption by facilitating access to information and learning, financial services, and input and output markets. Thus, smartphone use is also expected to have a direct impact on farmers' decision on farm machinery use. However, smartphone use is not expected to affect maize yields directly. The Pearson correlation analysis results support the assumption. In particular, the Pearson correlation analysis shows that smartphone use is significantly correlated with farm machinery use (correlation coefficient = 0.187, $P < 0.01$), while it has an insignificant correlation with maize yields (correlation coefficient = 0.048, $P > 0.10$).

2.3.2 Second-stage analysis: modelling the impacts of farm machinery use on maize yields

The second-stage analysis of the control function approach estimates the maize production function, focusing on the analysis of the effects of farm machinery use on maize yields. Within the framework of the control function approach, the inverse Mills ratio (*IMR*) is predicted after estimating farm machinery use equation in the first stage and then added as an extra covariate in the second-stage estimation. Therefore, the maize production function Equation (5) can be rewritten as follows:

$$\ln Y_i = \gamma_0 + \gamma_m M_i + \sum_{j=1}^4 \gamma_x \ln X_i + \sum_{n=1}^{11} \gamma_a A_n + \gamma_{imr} IMR + \nu_i \quad (9)$$

where M_i , X_i and A_n are variables defined previously; γ_m , γ_x and γ_a are corresponding parameters to be estimated; γ_{imr} represents the estimated parameter corresponding with the *IMR* term; and ν_i refers to the random error term.

Both the conditional quantile regression (CQR) model and the unconditional quantile regression (UQR) model can be potentially applied to estimate the heterogeneous effects of farm machinery use on maize yields. Compared with the UQR model, the CQR model provides a narrower interpretation of the effect of farm machinery use on maize yields because the resulting distribution of maize yields will be mostly conditional on the selection of other control variables used in the UQR model (Firpo *et al.* 2009; Mishra *et al.* 2015; Ma *et al.* 2019, 2020). Given that the UQR model enables to overcome the weakness of the CQR model, and it provides a better estimate, the UQR model is used in the present study. The UQR model has been applied recently (Ferraro *et al.* 2018; Fernandez and Bucaram 2019; Ma *et al.* 2020). For example, Fernandez and Bucaram (2019) employed the UQR model to examine the heterogeneous effects of environmental amenities on the housing prices in New Zealand. Ferraro *et al.* (2018) also applied the UQR model to investigate the impact of the statutory minimum wage on the wage distribution in Estonia, and they found that the minimum wage contributes to lower wage inequality.

The UQR model is estimated based on implementing the re-centred influence function (RIF) that is designed to measure how a change in the underlying distribution of an outcome affects distributional statistics such as median, different quantiles, variance and Gini coefficient. The RIF function in the UQR model can be estimated using the STATA commands of 'rifreg' (Firpo *et al.* 2009). In particular, the marginal effect for a given distribution statistic can be calculated by averaging the RIF regression regarding the change in the distribution of covariates. Following Firpo *et al.* (2009), the linear RIF regression for maize production can be specified as follows:

$$E[RIF(\ln Y_i; q_\tau) | M_i, \ln X_i, A_n, IMR] = \lambda_o + \lambda_m M_i + \sum_{j=1}^4 \lambda_x \ln X_i + \sum_{n=1}^{11} \lambda_a A_n + \lambda_{imr} IMR + \omega_i \tag{10}$$

where λ_m represents the marginal effects of farm machinery use at the maize yield distribution quantile q_τ ; λ_o is a constant; λ_x and λ_a are parameters to be estimated; and ω_i refers to an error term.

It is important to note that because some farmers did not use the pesticides, the log-transformation process would produce missing values. Following Bellemare *et al.* (2013), we employ a logarithmic-like transformation (i.e. inverse hyperbolic sine transformation, IHS) to address the zero-value issue of pesticide variables.²

2.4 Yield variance and Gini coefficient

In addition to estimating the yield impacts of farm machinery uses in this study, we are also interested in estimating the impacts of farm machinery use on maize yield variance and Gini coefficient. Following previous studies (Lai *et al.* 2008; Firpo *et al.* 2018; Hoque *et al.* 2019), the distributional variance of maize yields can be calculated as follows:

$$Variance = \sigma_Y^2 = \int (y - \mu_Y)^2 f(y) dy \tag{11}$$

where Y is the maize yield variable given by $Y = [y_1, y_2, \dots, y_n]$, and n is the sample size. μ_Y refers to the average maize yields; $f(y)$ is the probability density function.

The Gini coefficient is measured as follows:

$$Gini_Y = 1 - \frac{2}{\mu_Y} R_Y \tag{12}$$

where $R_Y = \int_0^1 GL_Y(p) dp$ with $p = F_Y(y)$ and where $GL_Y(p) = \int_{-\infty}^{F_Y^{-1}(p)} y dF_Y(y)$ is the generalised Lorenz ordinate of F_Y which refers to the cumulative distribution function.

² The IHS specified function is expressed as $IHS(x) = \ln(x + \sqrt{x^2 + 1})$, where x refers to an arbitrary random variable.

3. Data and descriptive statistics

3.1 Data

In this study, we use data drawn upon a randomised rural household survey of Chinese smallholder farmers in January 2017, and the information refers to the production year 2016. A multistage sampling strategy was adopted for selecting the samples. In the first stage, three provinces, including Gansu, Shandong and Henan, were selected based on the geographic characteristics and economic development levels. Shandong is located in the eastern region of China, and it is one of the most developed areas. Henan and Gansu are located in central plain and western hilly regions, respectively. In 2016, the maize outputs in Shandong, Henan and Gansu were 20.65, 17.46 and 4.61 million tons, respectively, together accounting for about 16.58 per cent of the country's total maize outputs. In the second stage, one city within each selected province was randomly chosen, and this includes Dingxi City in Gansu, Heze City in Shandong and Sanmenxia City in Henan. In the third stage, three villages were randomly selected from each city. Finally, around 45–55 households were randomly selected and interviewed in each village, resulting in a total of 493 samples.

Face-to-face interviews were conducted by well-trained enumerators who spoke both Mandarin and local dialects, using a detailed structured questionnaire. The enumerators were hired from local universities. The survey gathered information covering household and farm-level characteristics (e.g. age, gender, education, farm size and household size), farm machinery use status in maize production, maize yields, the use of production inputs (e.g. fertiliser, pesticide and seed) and distance to credit institutions.

We computed the sample size using Cochran's sample size determination formula due to lack of information on the population of smallholder maize households in the sampled regions. Cochran's formula is given as $n_0 = pqZ^2/e^2$, where we assume a margin of error e of 5 per cent, a probability or p -value of 0.5, a confidence level of 95 per cent with the corresponding Z -value of 1.96, thus giving a minimum sample size, $n_0 = [(0.5)(0.5)(1.96)^2]/(0.05)^2 = 385$. Thus, collecting a random sample of at least 385 households would be enough to give us the confidence level we need. As detailed in Table 1, the present paper relied on a sample size of 493 respondents to ensure precision.

The outcome variable used in this study is maize yields, which is measured by kg/mu (1 mu = 1/15 hectare). Farm machinery use is the key explanatory variable, which refers to a dummy variable identifying whether a household used a rotary cultivator for land preparation, that is land ploughing and levelling (1 = users and 0 = nonusers). We focus on farm machinery use in the land preparation stage rather than other production stages because land preparation has been identified as the most power-intensive stage in agricultural

production (Kienzle *et al.* 2013; Zhou *et al.* 2018), and it directly affects the use of inputs (e.g. seed, fertiliser and pesticide) and finally affects maize outputs. With respect to the selection of other explanatory variables that are expected to affect maize yields, we draw on the existing literature on farm machinery use and agricultural production (e.g. Benin 2015; Ma *et al.* 2018; Paudel *et al.* 2019; Justice and Biggs 2020; Takeshima *et al.* 2020; Van Loon *et al.* 2020).

3.2 Descriptive statistics

The definition and descriptive statistics of the variables used in the analysis are presented in Table 1. The table reveals that the average maize yields are 483.21 kg/mu (equivalent to 7,248 kg/ha), which are quite similar with the maize yields in the northern spring maize zone (500 kg/mu) but higher than the yields in the Yellow-Huai Valley summer maize zone (417 kg/mu), as reported by Ma and Maystadt (2017) for China. In their investigation of Zambia, Manda *et al.* (2016) reported that the average maize yields are 2,686 kg/ha for rural farmers. Averagely 85.6 per cent of households have used farm machinery in the land preparation stage. In our sample, the average farm size for maize production is 3.53 mu (equivalent to about 0.24 hectare), while the largest farm size is approximately 30 mu (equivalent to 2 hectares). The finding is consistent with the situation in China, where smallholder is still the mainstream in agricultural production.

The differences in the means of household demographic and farm-level characteristics variables between farm machinery users and nonusers are presented in Table 2. The last two columns in Table 2 report the mean differences and the corresponding *t*-value of the *t*-test. The table shows that the household heads in the farm machinery users tend to have low education level and a smaller household size than their nonuser counterparts. Farm machinery users are more likely to contact extension agents and are more likely to have access to better transportation condition than nonusers. The mean farm size for farm machinery users is around 1.6 mu larger than that for nonuser. The inputs of maize production for the farm machinery users appear to differ from the nonusers. Compared with nonusers, farm machinery users spend more on fertiliser but less on pesticide and seed. Our findings in the upper part of Table 2 also show that farm machinery users obtain about 34 kg/mu more maize yields than nonusers. The finding seems to indicate that farm machinery enables to increase maize yields. However, the simple mean comparison does not provide consistent results because it fails to account for the observable and unobservable confounding factors, which may generate misleading conclusions. Therefore, this study employs a control function method to address the issue of sample selection bias and estimate unbiased effects of farm machinery use on maize yields.

Table 3 presents the statistic information of household demographic and farm-level characteristics variables by maize yield quantiles at the 20th, 50th and 80th levels. The last column of Table 3 reports the *F*-statistic of the *F*-test

Table 1 Definition and descriptive statistics

Variables	Definition	Mean	SD
Maize yields	Maize yields (kg/mu)†	483.211	129.815
Farm machinery use	1 if household use farm machinery in land ploughing and levelling, 0 otherwise	0.856	0.351
Age	Age of household head (years)	46.787	10.323
Gender	1 if household head is male, 0 otherwise	0.836	0.371
Education	Schooling years of household head (years)	6.779	02.76
Household size	Number of members residency in a household	4.552	1.447
Extension contact	1 if household receives extension service, 0 otherwise	0.203	0.403
Transportation condition	1 if transportation from the village to the train/ bus station is convenient, 0 otherwise	0.753	0.432
Soil fertility	1 if soil is fertile, 0 otherwise	0.28	0.449
Access to credit	1 if household has access to credit, 0 otherwise	0.428	0.495
Irrigation	1 if farmland could be irrigated, 0 otherwise	0.998	0.026
Farm size	Total farm size used to cultivate maize (mu)	3.514	2.956
Pesticide costs	Pesticide costs (yuan/mu)‡	25.752	28.698
Fertiliser costs	Fertiliser costs (yuan/mu)	151.887	66.611
Seed costs	Seed costs (yuan/mu)	70.189	57.131
Smartphone	1 household head uses a smartphone, 0 otherwise	0.645	0.479
Gansu	1 if household residents in Gansu, 0 otherwise	0.327	0.469
Henan	1 if household residents in Henan, 0 otherwise	0.345	0.476
Shandong	1 if household residents in Shandong, 0 otherwise	0.329	0.47

Note: †1 mu = 1/15 hectare;

‡Yuan is Chinese currency: 1 USD = 6.70 yuan in 2017. SD, standard deviation.

that examines whether group means of quantiles are equal. The results in the table demonstrate the presence of discrepancy of demographic and farm-level attributes among smallholder farms. In particular, we show that farm machinery use, transportation condition, access to credit, farm size, fertiliser costs and the location regions are statistically significant across the selected quantiles of maize yields among farm households. The findings suggest that farm machinery use may have heterogeneous effects across the distribution of maize yields of rural households.

Figure 1 plots the maize yields distributions and the visible disparity between farm machinery users and nonusers. The figure shows that the maize yields distribution is unevenly distributed for both farm machinery users and nonusers, and the maize distribution of farm machinery users is less unequal than that of nonusers across the board. The findings suggest that farm machinery use may have the potential to reduce the inequality of maize yields among rural households.

4. Empirical results

4.1 The determinants of farm machinery use

The estimates for the factors that affect the use of farm machinery are presented in Table 4. As indicated previously, the probit regression model is

Table 2 Mean differences in households and farm-level characteristics between farm machinery users and nonusers

Variables	Machinery user (<i>N</i> = 422)	Nonuser (<i>N</i> = 71)	Mean difference	<i>t</i> -value
Maize yields	488.132 (125.100)	453.96 (152.668)	34.172**	2.059
Age	46.775 (10.250)	46.859 (10.82)	-0.084	-0.064
Gender	0.829 (0.377)	0.873 (0.335)	-0.044	-0.922
Education	6.68 (2.632)	7.366 (3.39)	-0.686*	-1.943
Household size	4.486 (1.439)	4.944 (1.443)	-0.458**	-2.480
Extension contact	0.227 (0.420)	0.056 (0.232)	0.171***	3.349
Transportation condition	0.801 (0.400)	0.465 (0.502)	0.336***	6.301
Soil fertility	0.268 (0.443)	0.352 (0.481)	-0.084	-1.465
Access to credit	0.377 (0.485)	0.732 (0.446)	-0.356***	-4.023
Irrigation	0.999 (0.014)	0.993 (0.059)	0.006*	1.844
Farm size	3.744 (3.091)	2.145 (1.317)	1.599***	4.291
Pesticide costs	24.164 (27.662)	35.192 (32.875)	-11.028***	-3.020
Fertiliser costs	156.902 (66.834)	122.077 (57.168)	34.825***	4.142
Seed costs	66.222 (58.953)	93.768 (37.147)	-27.546***	-3.810
Smartphone	0.68 (0.467)	0.437 (0.499)	0.243***	4.023
Gansu	0.251 (0.434)	0.775 (0.421)	-0.523***	-9.439
Henan	0.365 (0.482)	0.225 (0.421)	0.140**	2.297
Shandong	0.384 (0.487)	0.000 (0.000)	0.384***	6.638

Note: **P* < 0.1;

***P* < 0.05;

****P* < 0.01. Standard errors in parentheses.

used to estimate the Equation (6). Because the estimated coefficients of the variables cannot be interpreted directly, we calculate and present in the last column of Table 4 the estimates for the marginal effects of explanatory variables to facilitate the interpretation. The marginal effects are calculated by multiplying the coefficient estimates β by $\Phi(\beta X_i)$ at the mean values of X_i . The *chi-square* statistic presented at the bottom of Table 4 is significant at the 1 per cent level. The finding suggests that the independent variables that represent household and farm-level characteristics affect farm machinery use decision jointly. The McFadden R^2 , an indication of goodness of fit, is 0.45, suggesting a reasonable fit and excellent predictive power of our model (Maddala 1986).

The results show that the marginal effect of the education variable is negative and statistically significant, suggesting that an additional year of education decreases the probability of farm machinery use. A possible reason is that better-educated farmers are more likely to exit the farm but choose to participate in off-farm work activities that are usually better paid (Takeshima *et al.* 2018). The estimated marginal effect of the household size variable is negative and statistically significant from zero, suggesting that large households are 2.3 per cent less likely to use farm machines. The findings are in accordance with the result of Ma *et al.* (2018) who found that households with large members tend to have a greater supply of family farm

Table 3 Statistics of variables by selected maize yield quantiles

Variables	20th	50th	80th	F-value
Farm machinery use	0.737 (0.442)	0.864 (0.344)	0.932 (0.252)	12.09***
Age	46.74 (9.37)	46.97 (10.77)	46.72 (10.12)	0.55
Gender	0.788 (0.41)	0.84 (0.37)	0.85 (0.36)	0.30
Education	7.071 (2.71)	6.83 (2.75)	6.52 (2.63)	1.26
Household size	4.39 (1.33)	4.58 (1.60)	4.55 (1.41)	0.85
Extension contact	0.25 (0.44)	0.25 (0.43)	0.20 (0.40)	2.27
Transportation condition	0.64 (0.48)	0.65 (0.48)	0.85 (0.36)	15.87***
Soil fertility	0.22 (0.43)	0.16 (0.37)	0.32 (0.47)	1.20
Access to credit	0.455 (0.5)	0.442 (0.498)	0.304 (0.462)	7.22***
Irrigation	0.993 (0.054)	1.000 (0.000)	0.999 (0.016)	1.91
Farm size	2.77 (2.28)	3.79 (2.56)	4.19 (3.99)	8.66***
Pesticide costs	22.19(28.04)	17.96(19.44)	26.53 (32.33)	1.41
Fertiliser costs	138.54 (53.87)	133.53 (41.06)	160.84 (61.47)	14.00***
Seed costs	60.02(30.21)	67.23(87.50)	63.61(34.47)	0.47
Smartphone use	0.66 (0.48)	0.56 (0.50)	0.68 (0.47)	1.71
Gansu	0.26 (0.44)	0.24 (0.43)	0.24 (0.43)	0.48
Henan	0.68 (0.47)	0.42 (0.50)	0.24 (0.43)	34.05***
Shandong	0.06 (0.24)	0.34 (0.48)	0.52 (0.50)	47.87***

Note: * $P < 0.1$;

** $P < 0.05$;

*** $P < 0.01$. Standard deviation in parentheses.

labour, which leads to less demand for the labour-saving technologies such as farm machines. The extension contact variable has a positive and statistically significant marginal effect, suggesting that farmers contacting extension agents are 12 per cent more likely to use farm machinery. The finding highlights the importance of production extension and information in promoting machinery service in rural areas. Our finding is consistent with the result of Pan *et al.* (2018) who found that the agricultural extension program facilitates smallholder women farmers to use better basis cultivation methods and achieve improved food security in Uganda.

The transportation condition variable appears to have a positive and statistically significant effect on farm machinery use. This finding suggests that farmers who have access to convenient transportation are 9.3 per cent more likely to use farm machines. Convenient transportation would help facilitate smallholders to access machinery services. The positive and significant marginal effect of the access to credit variable suggests that access to credit increases the probability of using farm machinery among farmers. This finding is consistent with Mottaleb *et al.* (2017), who also found credit access significantly increases the likelihood of purchasing power tillers in Bangladesh. Access to credit helps alleviate capital constraints and thus enables farmers to purchase machinery or machinery services. The marginal effect of irrigation variable is positive and statistically significant, suggesting access to irrigation increases the probability of farm machinery use.

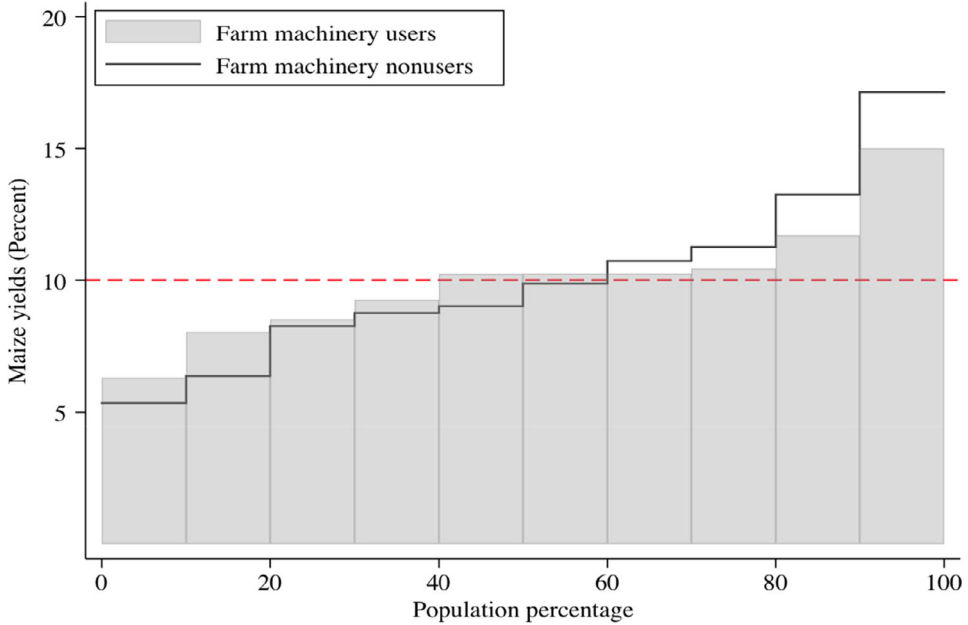


Figure 1 The maize yields distribution of farm machinery users and nonusers.
Note: The red line refers to the share of 10% of maize yields when each decile group would receive an equal distribution. [Colour figure can be viewed at wileyonlinelibrary.com]

It appears that the inputs of maize production are positively and significantly correlated with farm machinery use. The significant marginal effect of farm size variable suggests that the larger farmers are more likely to use farm machines. The finding is in line with Lai *et al.* (2015), who found that farm machinery use is determined by the total operating area of wheat and corn in Henan and Shandong provinces of China. Farm machinery use is positively correlated with the usage of productivity-increasing inputs such as fertilisers and pesticides. Seed costs decrease the probability of farm machinery use. Finally, our results show that relative to farmers in Shandong (reference group), those in Gansu are less likely to use farm machinery while those in Henan are more likely to use it. The findings suggest the presence of location-fixed effects affecting farm machinery use.

4.2 Impacts of farm machinery use on maize yields

The results for the UQR estimates on the impacts of machinery use and other explanatory variables on maize yields, which are estimated using the Equation (10), are presented in Table 5. We only report and discuss the UQR results for the 20th, 50th and 80th quantiles for the sake of brevity. Nevertheless, In Figure S2, we plot the effects of farm machinery use on maize yields distributions for a better understanding. For comparison, we

Table 4 Probit model estimation of farm machinery use

Variables	Coefficients	<i>t</i> -value	Marginal effects
Age	-0.008 (0.010)	-0.752	-0.001 (0.001)
Gender	-0.072 (0.278)	-0.260	-0.009 (0.035)
Education	-0.079 (0.034)**	-2.288	-0.010 (0.004)**
Household size	-0.178 (0.083)**	-2.154	-0.023 (0.010)**
Extension contact	0.944 (0.300)***	3.148	0.120 (0.037)***
Transportation condition	0.731 (0.218)***	3.349	0.093 (0.026)***
Soil fertility	-0.088 (0.218)	-0.404	-0.011 (0.028)
Access to credit	5.273 (1.737)***	3.036	0.671 (0.210)***
Irrigation	0.853 (0.220)***	3.881	0.109 (0.027)***
Farm size (<i>ln</i>)	0.242 (0.085)***	2.854	0.031 (0.011)***
Pesticide costs (<i>ln</i>)	0.850 (0.245)***	3.470	0.108 (0.029)***
Fertiliser costs (<i>ln</i>)	1.108 (0.301)***	3.678	0.141 (0.037)***
Seed costs (<i>ln</i>)	-6.520 (0.630)***	-10.354	-0.830 (0.095)***
Gansu	-4.571 (0.479)***	-9.536	-0.582 (0.079)***
Henan	0.776 (0.228)***	3.397	0.099 (0.028)***
Smartphone use	-0.089 (0.224)	-0.396	-0.011 (0.029)
Constant	-10.270 (3.019)***	-3.402	
<i>Chi</i> ²	960.117		
<i>Pob</i> > <i>Chi</i> ²	0.000		
McFadden <i>R</i> ²	0.450		
Observations	493		

Note: **P* < 0.1;

***P* < 0.05;

****P* < 0.01. The reference province is Shandong. Standard errors in parentheses.

also estimate the effects of farm machinery use on maize yields using the OLS model and present the results in the last column of Table 5.

The results presented in Table 5 show that the effects of farm machinery use are uniformly positive across the UQRs. The findings suggest that farm machinery use is associated with an increase in maize yields, which are consistent with the results of Benin (2015) on Ghana. Following Mishra *et al.* (2015) and Ma *et al.* (2020), the proportional impact of the discrete farm machinery use on maize yields is measured as $p_i = [\exp(\lambda_i) - 1]$, where λ_i is the coefficient of the farm machinery use variable. We show that farm machinery use increases maize yields by 12% [$\exp(0.111) - 1$] at the 20th quantile and 4% [$\exp(0.041) - 1$] at the 50th quantile. Such effects could not be observed if we only estimate the OLS model, which shows that farm machinery use increases maize yields by 13 per cent on average. At the higher quantiles such as the 80th quantile, farm machinery use does not have a significant impact on maize yields.

In terms of other explanatory variables that affect maize yields, our UQR estimates show that an additional year of education significantly increases maize yields by 1.3 per cent only at the 80th quantile, while an additional increase in household member increases maize yields by 0.8 per cent at the 50th quantile. In comparison, the OLS estimates show that the educational level of household heads significantly increases maize yields, but household

Table 5 The impact of farm machinery use on maize yields distribution at the selected quantiles: UQR model estimation

Variables	Dependent variable = <i>ln</i> (maize yields)		
	20th	50th	80th
Farm machinery use	0.111 (0.044)**	0.041 (0.020)**	0.040 (0.064)
Age	-0.000 (0.001)	0.000 (0.001)	0.004 (0.002)**
Gender	0.033 (0.026)	0.019 (0.016)	0.009 (0.039)
Education	0.002 (0.004)	-0.000 (0.002)	0.013 (0.007)*
Household size	0.006 (0.007)	0.007 (0.004)**	0.004 (0.010)
Extension contact	0.033 (0.033)	0.003 (0.019)	-0.070 (0.044)
Transportation condition	-0.023 (0.036)	0.046 (0.019)**	0.080 (0.042)*
Soil fertility	0.094 (0.031)**	0.092 (0.015)***	0.074 (0.042)*
Access to credit	0.005 (0.022)	-0.019 (0.012)	0.029 (0.034)
Irrigation	0.427 (0.304)	-0.076 (0.165)	-0.075 (0.194)
Farm size (<i>ln</i>)	0.091 (0.018)**	0.044 (0.010)***	0.058 (0.025)**
Pesticide costs (<i>ln</i>)	0.013 (0.013)	0.020 (0.005)***	0.012 (0.014)
Fertiliser costs (<i>ln</i>)	0.035 (0.028)	0.039 (0.015)***	0.129 (0.050)**
Seed costs (<i>ln</i>)	0.061 (0.032)*	0.088 (0.025)***	0.092 (0.064)
Gansu	-0.095 (0.051)*	-0.114 (0.034)***	0.234 (0.098)**
Henan	-0.236 (0.034)***	-0.148 (0.017)***	-0.099 (0.038)***
IMR	0.017 (0.053)	0.009 (0.027)	-0.107 (0.075)
Constant	5.527 (0.411)***	6.150 (0.237)***	5.335 (0.535)***
Adjusted <i>R</i> ²	0.223	0.389	0.221
Observations	493	493	493
			0.396
			4.739 (0.385)***
			0.002 (0.044)***
			0.002 (0.001)**
			0.033 (0.024)
			0.010 (0.005)**
			0.010 (0.006)
			0.007 (0.029)
			0.002 (0.030)
			0.147 (0.028)***
			0.016 (0.020)
			0.813 (0.280)***
			0.062 (0.018)***
			0.016 (0.012)
			0.033 (0.026)
			0.134 (0.043)***
			-0.061 (0.054)
			-0.265 (0.029)***
			-0.028 (0.049)
			4.739 (0.385)***
			0.396
			493

Note: * $P < 0.1$;

** $P < 0.05$;

*** $P < 0.01$. The reference province is Shandong. Standard errors in parentheses.

size does not have a significant impact on maize yields on average. Therefore, the UQR model provides more information by examining the impact of education and household size across the quantiles of maize yields.

Transportation condition plays a vital role in determining high maize yields. Our results show that access to better transportation condition enables to increase maize yields by 4.7 per cent at the 50th quantile and 8.3 per cent at the 80th quantile. The findings are consistent with the findings in previous studies which have shown that better infrastructure in rural areas enables to help increase agricultural productivity (Qin and Zhang, 2016). Soil fertility variable has a positive and statistically significant impact on maize yields. The UQR results show that cultivating on the land with fertile soil increases maize yields by around 10 per cent at the 20th and 50th quantiles and 8 per cent at the 80th quantile.

Since the explanatory variables, including farm size, pesticide costs, fertiliser costs and seed costs, and the dependent variable are expressed in logarithmic forms, the estimated coefficients of these variables reflect the elasticities. The results show that a 1 per cent increase in farm size increases maize yields by 0.09 per cent at the 20th quantile, 0.04 per cent at the 50th quantile and 0.06 per cent at the 80th quantile. Pesticide costs only have a significant impact on maize yields at the 50th quantile, and our results show that a 1 per cent increase in pesticide costs increases maize yields by 0.02 per cent. Regarding the impact of fertiliser costs on maize yields, our results show that a 1 per cent increase in fertiliser costs increases maize yields by 0.04 per cent at the 50th quantile and 0.13 per cent at the 80th quantile. The UQR results also show that a 1 per cent increase in seed costs significantly increases maize yields by 0.06 per cent at the 20th quantile and 0.09 per cent at the 50th quantile, while seed costs have no statistically significant impact on maize yields at the 80th quantile. The positive effects of agricultural inputs such as fertiliser, pesticide and seed on crop yields have also been reported in previous studies (e.g. Prishchepov *et al.* 2019).

The differences in maize yields exist among survey regions. Our results show that at the 20th and 50th quantiles, the maize yields in Gansu and Henan are lower than that in Shandong (reference region). At the 80th quantile, compared with maize yields in Shandong, the maize yields in Gansu are higher while that in Henan are lower. The findings suggest the presence of location-fixed effects (e.g. the differences in terms of climate, institutional arrangements and topography) that may also affect maize yields. Finally, the coefficients of the *IMR* term are insignificant across the quantiles, suggesting that the selection bias issue associated with farm machinery use variable is consistently corrected (Wooldridge 2015).

For comparison, we also estimated the impacts of farm machinery use on maize yields using the CQR model and presented the results in Table S1. The results show for households using farm machinery, the increase in maize yields at the 20th quantile and 50th quantile is around 21% ($\exp[0.191]-1$) and 11% ($\exp[0.107]-1$), respectively, which are higher than the effects

Table 6 Estimation of the effects of farm machinery use on maize yields variance and Gini coefficient

Variables	Variance		Gini coefficient	
	Coefficients	<i>t</i> -value	Coefficients	<i>t</i> -value
Farm machinery use	-0.027 (0.022)	-1.238	-0.006 (0.003)**	-1.981
Age	0.002 (0.001)**	2.412	0.000 (0.000)***	2.807
Gender	-0.018 (0.017)	-1.060	-0.004 (0.002)	-1.548
Education	0.005 (0.003)**	1.967	0.001 (0.000)*	1.787
Household size	0.001 (0.004)	0.196	0.000 (0.001)	0.168
Extension contact	-0.032 (0.019)*	-1.661	-0.005 (0.003)*	-1.757
Transportation condition	0.012 (0.020)	0.608	0.003 (0.003)	0.936
Soil fertility	0.015 (0.017)	0.925	0.000 (0.002)	0.022
Access to credit	0.033 (0.014)**	2.383	0.005 (0.002)**	2.332
Irrigation	-0.915 (0.241)***	-3.791	-0.116 (0.034)***	-3.368
Farm size (<i>ln</i>)	-0.051 (0.012)***	-4.198	-0.008 (0.002)***	-4.566
Pesticide costs (<i>ln</i>)	-0.001 (0.006)	-0.111	-0.000 (0.001)	-0.308
Fertiliser costs (<i>ln</i>)	0.011 (0.017)	0.620	0.002 (0.002)	0.671
Seed costs (<i>ln</i>)	0.007 (0.021)	0.320	-0.002 (0.003)	-0.550
Gansu	0.046 (0.034)	1.343	0.013 (0.005)***	2.688
Henan	0.061 (0.019)***	3.234	0.014 (0.003)***	5.160
IMR	-0.041 (0.030)	-1.361	-0.006 (0.004)	-1.500
Constant	0.860 (0.298)***	2.889	0.129 (0.042)***	3.049
Adjusted R^2	0.166		0.216	
Observations	493		493	

Note: * $P < 0.1$;** $P < 0.05$;*** $P < 0.01$. The reference province is Shandong. Standard errors in parentheses.

estimated from the UQR model (i.e. 12 per cent at the 20th quantile and 4 per cent at the 50th quantile). The differences also exist for other explanatory variables. For example, the CQR results presented in Table S1 show that a 1 per cent increase in seed costs increases maize yields by 0.13 per cent at the 20th quantile, but the UQR results presented in Table 5 show that a 1 per cent increase in seed costs only increases maize yields by 0.06 per cent at the same quantile. It is not implausible that the CQR model overestimates the effects of farm machinery use and other explanatory variables on maize yields. The quantiles in the CQR model are defined conditional on the employed covariates, and it is impossible for the CQR model to freely add or delete control variables without redefining the quantiles (Firpo *et al.* 2009; Mishra *et al.*, 2015; Ma *et al.* 2020). Therefore, the UQR model provides more convincing results.

4.3 The impacts of farm machinery use on the equality of maize yields

In addition to examining the heterogeneous effects of farm machinery use on maize yields, we investigate the variability and inequality of maize yields among rural households. Rather than using the different quantiles, we estimate a production function by using maize yields variance and an

estimated Gini coefficient as the dependent variables to investigate the role of farm machinery use in determining the maize yields inequality. The results, which are presented in Table 6, show that the variable representing farm machinery use is negatively and significantly associated with the Gini coefficient. Farm machinery use has a negative and insignificant impact on the variance of maize yields. The findings suggest that farm machinery use decreases maize yields inequality and indicate that using machinery on farms is a promising way to enhance maize productivity and reduce the inequality of maize yields among rural households.

4.4 Further analyses

4.4.1 *Impact of machinery use intensity on maize yields*

In our previous analysis, farm machinery use has been measured as a binary variable. To enrich our understanding regarding the association between farm machinery use and maize yields, we have conducted three additional analyses by measuring machinery use in other ways. In the first analysis, we estimated the impact of farm machinery expense per unit of land on maize yields, using a control function approach to address the endogeneity issue of farm machinery expense. The results (Table S2) show that farm machinery expense has a positive but insignificant impact on maize yields.

In the second analysis, we have analysed the impact of farm machinery use at different production and postharvest stages on maize yields. These include seven production stages and five postharvest stages (see definitions of the 12 variables in Table S3). The empirical results, which are presented in Table S4, show that adoption of farm machinery for land ploughing and threshing significantly increases maize yields, but farm machinery use for weeding and drying decreases maize yields significantly. It should be noted here it is not possible to address the endogeneity issues of the 12 variables that represent different production and postharvest stages of maize production simultaneously, due to the lack of efficient instrumental variables and methodologies. Nevertheless, we have provided some useful insights regarding the farm machinery use at different production and postharvest stages and maize yields.

In the third analysis, we examined the impact of machinery use intensity on maize yields, using the control function approach to address the endogeneity issue of machinery use intensity. In particular, machinery use intensity is measured as the accumulative values (1 = yes) of farm machinery use at the 12 production and postharvest stages of maize production, as illustrated in Table S3. The empirical results (Table S5) show that farm machinery use intensity has a positive and statistically significant impact on maize yields.

4.4.2 *Impact of farm machinery use on inputs costs*

In addition to influencing farm output, farm machinery use may have a direct impact on production inputs. To explore this, we analyse the impact of farm

machinery use on pesticide costs, fertiliser costs and seed costs. After addressing the endogeneity issue of the farm machinery use using the control function approach, we show (Table S6) that farm machinery use significantly increases pesticide costs, fertiliser costs and seed costs. Despite the positive relationship between farm machinery use and inputs costs, it is not clear whether machinery use increases maize yields directly or indirectly through increased production inputs. We addressed this research gap by estimating a mediation model. Taking fertiliser use as an example, we identify whether farm machinery use affects maize yields directly or indirectly through affecting fertiliser use. The estimated results (Table S7) show that farm machinery use has contributed 86 per cent direct effects on maize yields and 14 per cent indirect effects through affecting fertiliser use.

5. Conclusions and policy implications

In this paper we analysed the heterogeneous effects of farm machinery use on maize yields by employing an unconditional quantile regression model and household survey data collected from 493 households in rural China. To address the selection bias issue associated with the self-selection of farm machinery use, a control function was applied. We also estimated the impact of farm machinery use on maize yields variability and inequality using both the Gini coefficient and variance of maize yields.

The empirical results estimated from the first-stage estimation of the control function approach showed that extension contact, transportation condition, access to credit and irrigation, farm size, and the costs of production inputs including pesticide, fertilisers and seeds are main factors that drive smallholder maize farmers' decisions to use farm machinery.

The UQR results revealed that farm machinery use is associated with higher maize yields, but the effects of farm machinery use are not homogenous. In particular, we showed that farm machinery use increases maize yields by 12 per cent at the 20th quantile and 4 per cent at the 50th quantile, and the low-productive farmers tend to benefit more from farm machinery use relative to their high-productive counterparts. We found that maize yields are positively and significantly affected by soil fertility, farm size, pesticide costs, fertiliser costs and seed costs. The estimates for the effects of farm machinery use on maize yields Gini coefficient confirmed that farm machinery could be an equality-friendly technology that balances the maize productivity and yields equality. The additional analyses revealed that adoption of farm machinery for land ploughing and threshing significantly increases maize yields, but farm machinery use for weeding and drying decreases maize yields significantly. Pesticide costs, fertiliser costs and seed costs were also positively affected by farm machinery use.

Our findings of this study have important policy implications for sustainable agricultural production and rural development. The finding of the positive and significant yield-enhancing effects of farm machinery use

highlights the necessity of government policies that support the formation of farm machinery service market and provide favourable and affordable machinery extension services for smallholder farmers with low crop yields. The finding of the positive association between access to credit and farm machinery use underscores the importance of government financial programs in helping relax smallholder rural farmers' credit constraints. Both extension contact and transportation variables have positive and statistically significant impacts on farm machinery use, and the findings suggest that policies aimed at facilitating convenient transportation access and extension agent access of smallholder farms would help increase the likelihood of farm machinery use. High maize yields significantly matter with soil fertility. Thus, the government should consider to develop agricultural programs to help promote sustainable soil management programs in rural regions and increase farm productivity. Because small and fragmented land has been perceived as an obstacle, promoting land consolidation should be enhanced to support agricultural mechanisation.

A limitation of this study is that we have only considered the heterogeneous effects of farm machinery use in the land preparation stage on maize yields. Such effects may also exist due to the use of farm machinery in other production stages such as pesticide and fertiliser application, which is a promising area for future studies to have an investigation. In addition, the analysis of the present study only focuses on maize farmers and data collected from three provinces in China, and the studies focusing on other crops and other regions/countries are necessary to better understand the heterogeneous relationship between farm machinery use and agricultural performance in a broad context. When the required data are available, future studies may also investigate how crop productivity is affected by the intensity of farm machinery use in the land preparation stage (e.g. the costs of farm machinery use, hours of machinery use or the ratio of land using farm machinery to the total land).

Data availability statement

The data that support the findings of this study are available from the leading author, Xiaoshi Zhou, upon reasonable request.

References

- Abass, A., Amaza, P., Bachwenkizi, B., Wanda, K., Agona, A. and Cromme, N. (2017). The impact of mechanized processing of cassava on farmers' production efficiency in Uganda, *Applied Economics Letters* 24, 102–106.
- Adekunle, A., Osazuwa, P. and Raghavan, V. (2016). Socio-economic determinants of agricultural mechanisation in Africa: A research note based on cassava cultivation mechanisation, *Technological Forecasting and Social Change* 112, 313–319.

- Adu-Baffour, F., Daum, T. and Birner, R. (2019). Can small farms benefit from big companies' initiatives to promote mechanization in Africa? A case study from Zambia, *Food Policy* 84, 133–145.
- Aker, J. and Ksoll, C. (2016). Can mobile phones improve agricultural outcomes? Evidence from a randomized experiment in Niger, *Food Policy* 60, 44–51.
- Amadu, F.O., McNamara, P.E. and Miller, D.C. (2020). Understanding the adoption of climate-smart agriculture: A farm-level typology with empirical evidence from southern Malawi, *World Development* 126, 104,692.
- Aryal, J.P., Rahut, D.B., Maharjan, S. and Erenstein, O. (2019). Understanding factors associated with agricultural mechanization: A Bangladesh case, *World Development Perspectives* 13, 1–9.
- Baudron, F., Sims, B., Justice, S., Kahan, D.G., Rose, R., Mkomwa, S., Kaumbutho, P., Sariah, J., Nazare, R., Moges, G. and Gérard, B. (2015). Re-examining appropriate mechanization in Eastern and Southern Africa: two-wheel tractors, conservation agriculture, and private sector involvement, *Food Security* 7, 889–904.
- Baumüller, H. (2012). Facilitating Agricultural Technology Adoption Among the Poor: The Role of Service Delivery Through Mobile Phones, ZEF Working Paper Series 93. Available from URL: <https://doi.org/10.2139/ssrn.2237987>. <https://doi.org/10.20955/r.85.67> (accessed 10 Apr 2020).
- Bellemare, M.F., Barrett, C.B. and Just, D.R. (2013). The welfare impacts of commodity price volatility: evidence from rural Ethiopia, *American Journal of Agricultural Economics* 95, 877–899.
- Benin, S. (2015). Impact of Ghana's agricultural mechanization services center program, *Agricultural Economics* 46, 103–117.
- Chang, H., Gibson, J., Lai, D., Huang, J., Risser, J.M., Kapadia, A.S. and Chang, H. (2012). Does the use of eco-labels affect income distribution and income inequality of aquaculture producers in Taiwan?, *Ecological Economics* 80, 101–108.
- Evans, M.D.R., Kelley, J., Kelley, C.G.E. and Kelley, S.M.C. (2019). Income inequality in the great recession did not harm subjective health in Europe, 2003–2012. Available from URL: *Applied Research in Quality of Life*, <https://doi.org/10.1007/s11482-019-09741-0> (accessed 10 Apr 2020).
- FAO (2013). *Mechanization for Rural Development: A review of patterns and progress*. Integrated Crop Management, Plant Production and Protection Division, 20. Food and Agriculture Organization of the United Nations, Rome.
- FAO (2018). *Sustainable Agricultural Mechanization*. Plant Production and Protection Division. Food and Agriculture Organization of the United Nations (FAO), Rome. Available from URL: <http://www.fao.org/3/a-i7473e.pdf> (accessed 10 Apr 2020).
- Fernandez, M.A. and Bucaram, S. (2019). The changing face of environmental amenities: Heterogeneity across housing submarkets and time, *Land use policy* 83, 449–460.
- Ferraro, S., Meriküll, J. and Staehr, K. (2018). Minimum wages and the wage distribution in Estonia, *Applied Economics* 50, 5,253–5,268.
- Firpo, S., Fortin, N.M. and Lemieux, T. (2009). Unconditional Quantile Regressions, *Econometrica* 77, 953–973.
- Firpo, S., Fortin, N. and Lemieux, T. (2018). Decomposing wage distributions using recentered influence function regressions, *Econometrics* 6, 28.
- Foster, A. and Rosenzweig, M. (2017). Are There Too Many Farms in the World? Labor-Market Transaction Costs, Machine Capacities and Optimal Farm Size. *NBER Work. Pap. No. 23909*.
- Gibson, J. (2018). Forest loss and economic inequality in the Solomon Islands: using small-area estimation to link environmental change to welfare outcomes, *Ecological Economics* 148, 66–76.
- Greene, W.H. (2018). *Econometric analysis*, 8th edn. Pearson, New York.

- Hoque, A., Hasan, S. and Seo, B. (2019). Health inequality in bangladeshi children of age 6–59 months using hemoglobin data: a nonparametric analysis, *The Journal of Developing Areas* 53(3), 6–59.
- Ji, Y., Yu, X. and Zhong, F. (2012). Machinery investment decision and off-farm employment in rural China, *China Economic Review* 23, 71–80.
- Justice, S. and Biggs, S. (2020). The spread of smaller engines and markets in machinery services in rural areas of South Asia, *Journal of Rural Studies* 73, 10–20.
- Kienzie, J., Ashburner, J.E. and Sims, B.G. (2013). *Mechanization for rural development: A review of patterns and progress from around the world*, *Integrated Crop Management*. Integrated Crop Management, Vol. 20. Food and Agriculture Organization of the United Nations, Rome.
- Lai, D., Huang, J., Risser, J.M. and Kapadia, A.S. (2008). Statistical properties of generalized Gini coefficient with application to health inequality measurement, *Social Indicators Research* 87, 249–258.
- Lai, W., Roe, B. and Liu, Y. (2015). Estimating the Effect of Land Fragmentation on Machinery Use and Crop Production, Agricultural & Applied Economics Association and Western Agricultural Economics Association Annual Meeting, San Francisco, CA, July 26–28.
- Liu, Y., Violette, W. and Barrett, C.B. (2016). *Structural Transformation and Intertemporal Evolution of Real Wages, Machine Use, and Farm Size–Productivity Relationships in Vietnam*, *IFPRI Discussion Paper 01525*. International Food Policy and Research Institute, Washington, DC.
- Ma, W. and Abdulai, A. (2019). IPM adoption, cooperative membership and farm economic performance, *China Agricultural Economic Review* 11, 218–236.
- Ma, J. and Maystadt, J.-F. (2017). The impact of weather variations on maize yields and household income: Income diversification as adaptation in rural China, *Global Environmental Change* 42, 93–106.
- Ma, W., Renwick, A. and Grafton, Q. (2018). Farm machinery use, off-farm employment and farm performance in China, *Australian Journal of Agricultural and Resource Economics* 62, 279–298.
- Ma, W., Renwick, A. and Greig, B. (2019). Modelling the heterogeneous effects of stocking rate on dairy production: an application of unconditional quantile regression with fixed effects, *Applied Economics* 51, 4,769–4,780.
- Ma, W., Nie, P., Zhang, P. and Renwick, A. (2020). Impact of Internet use on economic well-being of rural households: Evidence from China, *Review of Development Economics* 24, 503–523.
- Maddala, G.S. (1986). *Limited-dependent and qualitative variables in econometrics*. Cambridge University Press, Cambridge.
- Manda, J., Alene, A.D., Gardebroeck, C., Kassie, M. and Tembo, G. (2016). Adoption and impacts of sustainable agricultural practices on maize yields and incomes: evidence from rural Zambia, *Journal of Agricultural Economics* 67, 130–153.
- Maredia, M.K., Reyes, B., Ba, M.N., Dabire, C.L., Pittendrigh, B. and Bello-Bravo, J. (2018). Can mobile phone-based animated videos induce learning and technology adoption among low-literate farmers? A field experiment in Burkina Faso, *Information Technology for Development* 24, 429–460.
- Michels, M., Fecke, W., Feil, J.-H., Musshoff, O., Pigisch, J. and Krone, S. (2020). Smartphone adoption and use in agriculture: empirical evidence from Germany, *Precision Agriculture* 21, 403–425.
- Mishra, A.K., Mottaleb, K.A. and Mohanty, S. (2015). Impact of off-farm income on food expenditures in rural Bangladesh: an unconditional quantile regression approach, *Agricultural Economics* 46, 139–148.

- Mottaleb, K.A., Rahut, D.B., Ali, A., Gérard, B. and Erenstein, O. (2017). Enhancing smallholder access to agricultural machinery services: lessons from Bangladesh, *The Journal of Development Studies* 53, 1,502–1,517.
- Müller, M. (2020). Leadership in agricultural machinery circles: experimental evidence from Tajikistan, *Australian Journal of Agricultural and Resource Economics* 64, 533–554.
- NBSC (2019). *China Statistics Yearbook 2019*. National Bureau of Statistics of China, Beijing.
- Pan, Y., Smith, S.C. and Sulaiman, M. (2018). Agricultural extension and technology adoption for food security: evidence from Uganda, *American Journal of Agricultural Economics* 100, 1,012–1,031.
- Paudel, G.P., Kc, D.B., Rahut, D.B., Justice, S.E. and McDonald, A.J. (2019). Scale-appropriate mechanization impacts on productivity among smallholders: Evidence from rice systems in the mid-hills of Nepal, *Land Use Policy* 85, 104–113.
- Paudel, G.P., Gartaula, H., Rahut, D.B. and Craufurd, P. (2020). Gender differentiated small-scale farm mechanization in Nepal hills: An application of exogenous switching treatment regression, *Technology in Society* 61, 101,250.
- Prishchepov, A.V., Ponkina, E., Sun, Z. and Müller, D. (2019). Revealing the determinants of wheat yields in the Siberian breadbasket of Russia with Bayesian networks, *Land Use Policy* 80, 21–31.
- Qiao, F. (2017). Increasing wage, mechanization, and agriculture production in China, *China Economic Review* 46, 249–260.
- Qin, Y. and Zhang, X. (2016). The road to specialization in agricultural production: evidence from rural China, *World Development* 77, 1–16.
- Sims, B., Hilmi, M. and Kienzie, J. (2016). Agricultural mechanization: A key input for sub-Saharan African smallholders, *Integrated Crop Management* 23. Food and Agriculture Organization of the United Nations, Rome.
- Song, C., Liu, R., Oxley, L. and Ma, H. (2018). The adoption and impact of engineering-type measures to address climate change: evidence from the major grain-producing areas in China, *Australian Journal of Agricultural and Resource Economics* 62, 608–635.
- Takehima, H., Adhikari, R.P., Shivakoti, S., Kaphle, B.D. and Kumar, A. (2017). Heterogeneous returns to chemical fertilizer at the intensive margins: Insights from Nepal, *Food Policy* 69, 97–109.
- Takehima, H., Houssou, N. and Diao, X. (2018). Effects of tractor ownership on returns-to-scale in agriculture: Evidence from maize in Ghana, *Food Policy* 77, 33–49.
- Takehima, H., Hatzenbuehler, P.L. and Edeh, H.O. (2020). Effects of agricultural mechanization on economies of scope in crop production in Nigeria, *Agricultural Systems* 177, 102,691.
- Tamkoç, M.N. and Torul, O. (2020). Cross-sectional facts for macroeconomists: wage, income and consumption inequality in Turkey, *The Journal of Economic Inequality* 18(2), 239–259.
- Tufa, A.H., Alene, A.D., Manda, J., Akinwale, M.G., Chikoye, D., Feleke, S., Wossen, T. and Manyong, V. (2019). The productivity and income effects of adoption of improved soybean varieties and agronomic practices in Malawi, *World Development* 124, 104,631.
- Van den Berg, M.M., Hengsdijk, H., Wolf, J., Van Ittersum, M.K., Guanghuo, W. and Roetter, R.P. (2007). The impact of increasing farm size and mechanization on rural income and rice production in Zhejiang province, China, *Agricultural Systems* 94, 841–850.
- Van Loon, J., Woltering, L., Krupnik, T.J., Baudron, F., Boa, M. and Govaerts, B. (2020). Scaling agricultural mechanization services in smallholder farming systems: Case studies from sub-Saharan Africa, South Asia, and Latin America, *Agricultural Systems* 180, 102,792.
- Wang, X., Yamauchi, F., Otsuka, K. and Huang, J. (2016). Wage growth, landholding, and mechanization in Chinese agriculture, *World Development* 86, 30–45.
- Wooldridge, J.M. (2015). Control function methods in applied econometrics, *Journal of Human Resources* 50, 420–445.

- Yang, J., Huang, Z., Zhang, X. and Reardon, T. (2013). The rapid rise of cross-regional agricultural mechanization services in China, *American Journal of Agricultural Economics* 95, 1,245–1,251.
- Yi, Q., Chen, M., Sheng, Y. and Huang, J. (2019). Mechanization services, farm productivity and institutional innovation in China, *China Agricultural Economic Review* 11(3), 536–554.
- Zhang, J., Wang, J. and Zhou, X. (2019). Farm machine use and pesticide expenditure in maize production: health and environment implications, *International Journal of Environmental Research and Public Health* 16, 1,808.
- Zhou, X., Ma, W. and Li, G. (2018). Draft animals, farm machines and sustainable agricultural production: insight from China, *Sustainability* 10, 3,015.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. The trend of productivities of maize in China and America (1961–2018).

Figure S2. The effects of farm machinery use on maize yields distribution.

Table S1. The impact of farm machinery use on maize yields distribution at the selected quantiles: CQR model estimation.

Table S2. The impacts of farm machinery expense on maize yields: Control function approach estimation.

Table S3. Descriptive statistics of the 12 farm machinery use variables.

Table S4. The impact of farm machinery use on maize yields by production stage.

Table S5. The impacts of machinery use intensity on maize yields: Control function approach estimation.

Table S6. The impacts of farm machinery use on pesticide costs, fertilizer costs and seed costs: Control function approach estimation.

Table S7. The impacts of farm machinery use on maize yields: mediation analysis.