



FACTORS AFFECTING WHEAT YIELD VARIATION OF SMALLHOLDER FARMERS IN PUNJAB PROVINCE: APPLICATION OF ORDERED LOGISTIC MODEL

R. Saeed^{1*}, M. Qasim², I. Mahmood¹ and W. Akhtar²

¹Social Sciences Research Institute (PARC), AARI, Jhang Road, Faisalabad, Pakistan

²Social Sciences Research Institute, NARC, Chak Shahzad, Islamabad, Pakistan

ABSTRACT

Many socioeconomic, institutional and biophysical factors are causing high wheat yield variation among wheat growers in the country in general and in the Punjab province in specific. Ultimate purpose of present study was to determine factors affecting probability of wheat yield being in low, medium or high ordinals given the set of yield changing inputs. Cross-sectional data collected from randomly selected 320 wheat growers with 80 respondents from each of four agro-ecological regions of Punjab Province was analyzed through proportional odds model to obtain the study objectives. The study found out some socio-economic and agro-ecology related variables such as age, tractor ownership, role of income diversification through part-time farming, the contribution of smallholder tenants and owner-cum-tenants and chemical fertilizers that can significantly affect wheat yield categories of low, medium or high. Based on findings, it is imperative to support young innovative farmers having their own farm machinery, generate off-farm/on-farm income generating avenues for part-time farmers, and provide more facilities to smallholder farmers of both tenants as well as owner-cum-tenant class in enhancing their wheat production of higher level. Moreover, agriculture advisory services should focus more on the cotton-wheat zone accompanied by appropriate use of seed rate, chemical fertilizers, and plant protection measures to enhance wheat yield in Punjab Province.

Keywords: ordered logistic regression, smallholder farmers, Wheat, yield categories

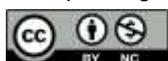
INTRODUCTION

Wheat crop has occupied a prominent position in formulation of agriculture sector policies in Pakistan (GoP, 2020). Literature suggests that wheat is one of the major crops grown by 80% percent farmers who are predominantly small and medium-sized farmers (FAO, 2013) and has high demand from consumers (household consumption of wheat is almost 9%) yet its average yield despite irrigated lands remains low compared to countries with comparable agro-climatic conditions. Even wheat yield levels are exceptionally different and low in rainfed than irrigated areas of Pakistan (Jan *et al.*, 2021). Although Pakistan claims self-sufficiency in wheat production, existing wheat yield depicts more than 50% yield gaps than the developed countries (Aslam, 2016). Wheat is the most important food crop as more than 60% people of Pakistan are totally dependent on wheat for their regular diet (Shah *et al.*, 2017). The total wheat production of 20-24 million tons in a year is mostly consumed locally; however, the country

had to import wheat in substantial quantities in the years of shortage (FAO, 2013). The self-sufficiency in food, therefore, remained a prime goal for the country as increased wheat production and yield can mitigate the effect of higher food prices. Now-a-days, wheat has gained more popularity as one of the staple and cash crops in general and in Punjab particularly due to increased demand from consumers in the wake of COVID-19 outbreak. If IIASAs (2008) possibility of Chinese diet change, based on the present trend of economic crisis, is considered then there is high a probability of Chinese reversion to old commune system in agriculture when rice, wheat and tubers were the hallmark of 1950s to 1960s. These circumstances may give favor to Pakistan in benefiting from more production of wheat and export to major COVID-19 hit countries because wheat harvesting and marketing was spared the negative effects of COVID-19 particularly in Punjab (ADB, 2020).

Based on the latest economic survey of Pakistan (2018-19), wheat accounts for 8.9 percent value added to agriculture and 1.6 percent to GDP of the country. Punjab has been

*Corresponding author: rashedkasuri@gmail.com



leading in per acre wheat production than other provinces up to 2003-04 but situation reversed after this and now Sind holds first position in per acre wheat production in the country. During study year 2017-18, Sind produced 33.39 quintals per hectare as compared to Punjab which produced 29.23 quintals per hectare (AMIS, 2018). The national average wheat production of 29 q/ha in the country is far less compared to world average of plus 40 q/ha (Alemu, 2015). Over the last decade, the yield of all important crops is low than desirable levels mainly due to stagnant cultivated area, productivity and climate changes (GoP, 2019). Resultantly, the low yield has given tough time to the country in meeting high consumer demands. Keeping in view that wheat is second most important staple crop for rural than urban masses in Pakistan (GoP, 2011), it is imperative to identify other ways of increasing wheat yield despite adoption of improved inputs, management practices and technologies (Dorosh and Rashid, 2013). Though the agricultural extension department is striving hard to promote improved technologies in the study area to enhance wheat yield for many years, the national average wheat yield could not go up from 28.5 quintals per hectare with higher yield variations among farmers (AMIS, 2018). Past research shows numerous wheat growers in the country are getting much below or above the average national wheat production. Therefore, it is imperative to identify the socio-economic and biophysical factors that determine the relatively low, medium and high wheat yield groups among farmers. Empirical knowledge is scanty on the comparative analysis of the factors that are causing wheat yield variations in the country. It is expected that current research will help in better targeting the agricultural extension services towards successful adoption and promotion of wheat crop husbandry practices. Present study is an effort to fill this knowledge gap. The study was undertaken with the major objective of i) exploring demographic information of the respondent ii) determining factors affecting low, medium and high wheat yield levels among farmers, and iii) suggesting policy recommendations for better wheat husbandry practices.

MATERIALS AND METHODS

Description of the study area

The Punjab Province is the most agricultural province of Pakistan where 80% of the country's wheat and cotton is produced (GoP, 2018). The

Punjab province is divided into 14 agro-ecological zones (AEZs) on the basis of weather and soil characteristics etc. (Ahmad *et al.*, 2019). Such types of variations make the provincial lands good for raising diverse agricultural crops. Important crops grown in the province are wheat, cotton, rice, sugarcane, maize and different types of fodders and oilseeds. However, Punjab Province alone provides almost 3/4th of wheat to the country.

Sampling methods

A simple random sampling procedure was adopted for sample selection. The study sample consisted of 4 agro-ecological zones for data collection. The survey team of Pakistan Agricultural Research Council randomly approached 80 farming households from each zone for collecting farm level data. Thus, the total sample comprised 320 farmers.

Methods of data collection

Present study mainly relies on farm-level original data; therefore, a survey was conducted in the study area of Punjab Province. Well-designed and properly tested questionnaire was used to collect both quantitative and qualitative information. Data comprised households socio-economic characteristics (household education level, time spent in farming, farming experience), land holding (agricultural land, ownership of tractor, area under wheat), farm inputs utilization (seed rate, fertilizers, pesticides, farm yard manure), farm output (wheat yield). Secondary information was collected from published and un-published documents to support research findings.

Data analysis

First of all, descriptive, inferential and analysis of variance techniques were used to know about the nature of data. Then, to meet the objectives, ordered logistic regression also known as proportional odds model was applied to identify the factors impacting wheat yield at different groups of low, medium or high. Previous research suggests that variation in yields of crops among farmers of the study area is a common feature on grounds of varying agro-ecological and socio-economic conditions of the households. Present study also proves this, for example, some farmers achieved wheat yield much below the mean yield level of the sample households and others crossed yield even 60 quintals/hectare (Table 2). In practice, to measure this yield variation of categorical

dependent variable, ordered logistic or ordered probit regression works well giving similar results. Here, wheat yield categories of low, medium or high were regressed on some explanatory variables through ordered logistic regression. The actual values taken on by the outcome variable are irrelevant in present model, except that larger values are supposed to correspond to higher outcomes. In current study, yield categories were taken as 1, 2, and 3 to represent low, medium or high yield groups. Average yield and standard deviation of sample farmers as estimated in (Table 2) were used to develop categories of dependent variable. Hence, yield levels within the range of average plus or minus one standard deviation were treated as medium, and those coming in the range of average minus more than one standard deviation were treated as low and finally those with average plus more than one standard deviation were named as high yield groups.

Table 1. Descriptions and measurement of variables

Variables	Measurements	Descriptions
Dependent Outcome	Groups (low, medium, and high)	Group 1=low, yield less than 32 q/ha, 2=medium, yield within ≥ 32 and ≤ 48, and 3=high if more than 48 q/ha
Education	Year	Education of the respondent in years
Age	Year	Present age of the farmer respondents
Farm sizes	Small, medium, large	Small farmers=land less than 12.5 acres, medium=12.5 to 25, large= above 25 acres
Agro-ecology	Rice-Wheat, Cotton-Wheat, Mixed Crop, and Rain-fed	1=Rice-Wheat, 2=Mixed Crop, 3=Cotton-Wheat, and 4= Rain-fed/low intensity zone in Punjab Province
Time for farming	Full/Part Time	The involvement of respondent in farming, 1=full time, 0 otherwise
Land tenancy	Owner, tenant, owner-cum-tenant	Tenancy 1=owner, 2=owner-cum-tenant, and 3=tenant
Tractor ownership	Own/rent	Source of tractor, 1=owned, 0 otherwise
Seed rate	Kg	Applied seed rate of wheat per acre
Farm manure (FYM)	Yes/No	Use of FYM in wheat field, 1=yes, and 0 otherwise
Fertilizers	Kg	Quantity of chemical fertilizers used per acre
Pesticides	Spray	Number of pesticide sprays on wheat crop
Weedicide	Sprays	Number of weedicide sprays for weeds management
Harvesting	Manual/Mechanical	Farmers practice of harvesting wheat, 1>manual, and 0 otherwise

In ordered logistic model, the probability of outcome value is determined as a linear function of the predictor variables associated with cut-points and chances of expecting outcome “i”

corresponds to the probability that the estimated linear function, plus random error, is within the range of the cut-points estimated for the outcome (Equation 1).

$$P_{ij} = \Pr (y_{ij} = i) \Pr (k_{i-1} < x_{ij}\beta + u_i \leq k_i) = 1 / 1 + \exp (-k + x_{ij}\beta) - 1 / 1 + \exp (-k - 1 + x_{ij}\beta) \dots \dots \dots (1)$$

Here, u_j is supposed to be logistically distributed in ordered logit and the coefficients of β ($\beta_1, \beta_2, \beta_3, \dots, \beta_k$) along with the cut-points k_1, k_2, \dots, k_{k-1} , where k means number of possible outcomes. Whereas, the value of k_0 is defined as $-\infty$, and k_k as $+\infty$ (Kebede *et al.*, 2017). Coefficients and cut-points are estimated using maximum likelihood but the cut-points are not ordinarily elaborated in the results. In Table 1, based on prior literature and theoretical justifications, following predictor variables with measurement of variables have been provided to determine the categorical dependent variable.

RESULTS AND DISCUSSION

Descriptive results

According to preliminary results, average yield was estimated as 39 quintal/ha, ranging from lowest achieved of 13 to highest yield value of 64 q/ha. However, this provincial average yield is quite reasonable than the national average yield as Badar *et al.* (2015) reported as about 28 quintal per hectare while studying time series country level data. Two things are clear now. First, wide output difference among farmers is evident from values of variance and standard deviation. Second, average yield seems to follow normality due to median yield of 39 and coefficient of skewness of 0.03. Therefore, more struggles are required to lessen a wide yield gap among the low and high wheat yield groups of respondents. Moreover, the current wheat yield of 38.85 quintals per hectare in Punjab Province is well above the national average wheat yield of 29 quintal per hectare in the country. But when converted to tons, the 38.85 q/ha yield of the province becomes almost 4.28 US tons per hectare which is far less than potential yield of 6 tons per hectare in Pakistan (FAO, 2013).

Table 2. Summary statistics of yield

Parameters	Output (q per ha)
Average	38.85
S.D	7.844
Lowest yield	12.85
Highest yield	64.25
Median	39.54
Variance	61.52
Skewness	0.036
Observations	320

Table 3 below reveals frequency, percent of farmers and average yields coming under three wheat yield levels. Large number of respondents (64%) obtained yield of middle category covering yield from 32 to 48 q/ha. From general understanding of the data, it appears that medium level category would have more probability of entering into a higher wheat yield category. Only 15% farmers could obtain high yield level while about 20% achieved low yield level. All this implies that combined probability of both low and medium farmers might be higher of entering into higher yield category than large farmers as evident from joint percentage (85%) of farmers subject to providing appropriate farm input packages and given their prevailing/constant socio-economic and agro-ecological characteristics.

Table 3. Statistics on average yield by category

Statistics	Frequency	Percent	Cumulative percent	Mean yield	Std. Dev.
Low	66	20.6	20.6	28.09	3.243
Medium	205	64.1	84.7	39.51	4.079
High	49	15.3	100	50.56	4.763
Total	320	100		38.85	-

The ANOVA technique was applied to judge the average yield variation among sample respondents as reported in Table 4. The ANOVA results show that F-statistic with value of 443.31 is highly significant on *p*-value of 0.0000, revealing significant average wheat production variation among prescribed three yield groups. Therefore, it is high time to determine the factors affecting for households' yield variation into low, medium and high. Improving wheat yield necessitates targeting the factors that make up relatively low or high wheat yield.

Table 4. ANOVA statistics of yield categories

Analysis of variance					
Source	SS	df	MS	F	Prob> F
Between groups	14456.3	2	7228.16	443.31*	0.0000
Within groups	5168.7	317	16.3		
Total	19625	319	61.52		
Bartlett's test for equal variances: $\chi^2(2) = 8.3386^*$ Prob> $\chi^2 = 0.015$					

*Significant at 1% significance level

Econometric estimation results

Table 5 presents ordered logistic regression estimates after the 4th iteration. Total 320 observations used in the analysis are displayed at the start of output Table. The model Likelihood Ratio (LR) chi-square having value of 68.21 at 17 degrees of freedom was highly significant with *p*-value less than 0.01. The significant chi-square shows overall fitness of

the model with all its variables as compared to null model with no predictors. In simple words, the predictor variables significantly affect wheat yield categories of wheat growers. The effect of individual variables is elaborated with reference to respective coefficients and associated z-score with two tail *p*-values. Value of Pseudo R² (McFadden R²) is approximately 12%. The positive coefficients for variables reveal that farmers with more units of positive inputs have more chances of obtaining higher level of wheat yield given that all other variables included in the model are kept constant and inverse for negative coefficients i.e. farmers with more units of the variables may have less probability of entering into higher yield group.

Overall 6 variables are found statistically significant but with different significance levels and signs. Positive and significant coefficient holder variables inform that remaining variables included in the model being same, farmers owning their tractor and the owner-cum-tenant (OCT) and tenant farmers, farmers of the cotton-wheat zone and applying adequate chemical fertilizers tend to have higher wheat yield levels per hectare. On the contrary, the age of the farmers and time devoted to farming are negatively but significantly affecting wheat production. Keeping other variables constant, for old farmers, the households may have less probability of obtaining higher wheat yield versus the combined medium and low yield categories. Similarly, keeping other variables constant, the farmers devoting full time to crops, have less probability of entering into higher yield versus the combined medium and low categories. For example, one unit (year in this case) increase in farmers age, we expect a 0.01 decline in the log odds of being in a higher yield level versus the joint effect of low and medium categories, given that all other variables of model are held constant. This finding seems logical as mostly old-farmers are conservative using traditional methods of crop-raising and hence, may have less agricultural output. Likewise, holding other factors constant, for full-time farmers, the log odds of being in a higher yield category versus the combined low and medium yield categories are 0.66 times less as compared to part-time farmers. This is in line with the production efficiency theory of being production dependent on off-farm work and off-farm's work dependency on production efficiency (Lien *et al.*, 2010). The finding is endorsed by Musafiri and Sjolander (2017) that off-farm work hours for income diversification of smallholder part-time

farmers can generate spill-over effects for enhancing agricultural productivity of smallholder farmers. Woldehanna (2000) also concluded that farmers having diversified sources of income generate more agricultural output. That is why development practitioners connect development of a non-farm sector with agricultural development since the non-farm sector reinforces the agricultural income (Sokcheng, 2014).

Some variables such as education, farmyard manure, pesticides and weedicides like inputs are not significantly affecting the yield categories but still these bear theoretically expected signs. For example, education, farmyard manure, pesticides, and weedicides like inputs have positive signs and the method of wheat crop harvesting has negative sign. The farmers practicing manual wheat harvesting may have less probability of shifting to a higher yield category than farmers relying on mechanized harvesting; in line with the finding that rich mechanized cotton-wheat zone of the province has more probability of going to a higher wheat yield group as evident in results of ordered logistic regression. The fertilizer consumption is a major input for higher wheat production but contrary to present study, Chandio *et al.* (2016) found a non-significant effect on total wheat production in Pakistan. Most farmers regard fertilizers alone can enhance the yield and disregard proper application of other best practices. Non-adoption of recommended best practices could affect negatively on the level of yield as evident from seed rate in present case. The ordered logistic regression is without intercept term as its effect is absorbed in cut-points which are generally not elaborated but show where the latent variable (continuous variable) is cut to make the three levels/groups that we have specified for our dependent variable. Table 6 reports the predicted probabilities for each yield category subject to setting off all predictor variables of the model to their average values. Predicted probabilities are normally obtained after performing ordered logistic regression. Predicted probabilities are usually easier to understand than coefficients or the odd ratios obtained in ordered logistic regression. Average probability of having the first wheat category was 0.312 in the sample at the average values of predictor variables. In the same way, probability of yield being in the medium yield group was lower with average value of 0.19, with lowest probability of 0.0152 to highest probability of 0.8069.

Table 5. Parameter estimates for three yield groups

Ordered logistic regression		Observations		320
		LR chi ² (17)		68.21
		Prob> chi ²		0.000*
Log likelihood	-253.32	Pseudo R ²		0.1187
Variables	Coeff.	S.E	Z	p> z
Age	-0.0169	0.009	-1.77	0.077***
Education	0.032	0.031	1.04	0.300
Part-time (ref. category)				
Full time participation	-0.667	0.351	-1.90	0.058**
Small farm (reference)				
Medium	0.067	0.339	0.20	0.842
Large	-0.345	0.381	-0.91	0.365
Tractor rented (ref)				
Tractor owned	0.668	0.307	2.17	0.030**
Rice-Wheat (ref. zone)				
Mix crop	0.644	0.422	1.53	0.127
Cotton-Wheat	1.074	0.382	2.81	0.005*
Rain-fed/low intensity	0.055	0.425	0.13	0.897
Owner (ref. tenancy)				
Owner-cum-tenant	0.595	0.297	2.00	0.045**
Tenant	1.197	0.516	2.32	0.021**
Farmyard Manure Non-user (ref)				
FYM user	0.17	0.258	0.66	0.511
Seed rate	-0.037	0.028	-1.30	0.194
Fertilizers	0.007	0.002	3.53	0.000*
Pesticides	0.122	0.385	0.32	0.751
Weedicide	0.345	0.285	1.21	0.225
Mechanical harvesting (ref)				
Manual harvesting	-0.642	0.312	-0.21	0.837
/cut1	-2.11	1.666		
/cut2	1.557	1.659		

*p < 0.01; **p < 0.05; ***p < 0.10

Table 6. Predicted probabilities for yield categories

Output	Average	S.D	Minimum	Maximum
1= Low	.3127	.1576	.0338	.6156
2= Medium	.1950	.1387	.0152	.8069
3= High	.1088	.096	.0123	.4938

CONCLUSION AND RECOMMENDATIONS

The ultimate purpose of current study was to identify the socio-economic and farm inputs that could affect entry of farmers into relatively low, medium and high wheat yield categories. Cross-sectional data collected from randomly selected respondents from agro-ecological zones was used to apply ordered logistic regression to know factors causing yield variation among farmers. The research segregated very useful socio-economic and agro-ecology related variables such as old age, tractor ownership, role of part-time farming activities as a proxy of income diversification, the contribution of tenants and owner-cum-tenants and fertilizers that can enhance wheat yield of higher levels. Based on findings, it is imperative to support young innovative farmers having their own farm machinery, strengthen off-farm/on-farm income generating avenues for part-time farmers, and provide more facilities to smallholder farmers of both tenants as well as owner-cum tenant class in enhancing their wheat production of higher level. Hence, agriculture advisory services

should focus more on the cotton-wheat zone accompanied by appropriate use of seed rate, chemical fertilizers, and plant protection measures to enhance wheat yield in the study area.

AUTHOR'S CONTRIBUTION

R. Saeed: Study plan, data analysis and write-up of results and discussion

M. Qasim: Helped in writing introduction and abstract

I. Mahmood: Helped in data collection and references

W. Akhtar: Participated in data collection and review of literature

REFERENCES

- Ahmad, A., I. Ahmad, M. R. Khan, S. H. H. Shah, M. A. Kamran, S. A., Wajid, M. Amin, A. Khan, M. N. Arshad, M. J. M. Cheema, Z. A. Saqib, R. Ullah, K. Ziaf, A. UIHuq, S. Ahmad, M. Fahad, M. M. Waqas, A. Abbas and A. Iqbal. 2019. Agro-ecological zones of Punjab, Pakistan, Rome, FAO.
- Alemu, T. 2015. Determinants of wheat yield variation of smallholders in South Eastern Ethiopia: Application of ordered logistic regression. *Science, Technology and Arts Research Journal*, 4 (3): 61-66.
- AMIS, 2018. Crop yield data 1947-48 to 2017-18, Province wise average yield of wheat from 1947-48 to 2017-18. Agriculture Marketing Information Service, Directorate of Agriculture (Economics and Marketing) Punjab, Lahore.
- Asian Development Bank. 2020. ADB briefs 149, August, 2020.
- Aslam, M. 2016. Agricultural productivity current scenario, constraints and future prospects in Pakistan. *Sarhad Journal of Agriculture*, 32 (4): 289-303.
- Badar, N., H. Fatima, A. Jabbar and M. Asif. 2015. Major food crops production and yield forecast in Pakistan. *Pakistan Journal of Agricultural Research*, 28 (3): 295-302.
- Chandio, A. A., Y. Jiang, M. A. Joyo and A. Rehman. 2016. Impact of area under cultivation, water availability, credit disbursement, and fertilizer off-take on wheat production in Pakistan. *Journal of Applied Environmental and Biological Sciences*, 6 (10):10-18.
- Dorosh, P. and S. Rashid. 2013. *Food and Agriculture in Ethiopia: Progress and Policy Challenges*. University of Pennsylvania Press Philadelphia, USA.
- FAO, 2013. Pakistan review of the wheat sector and grain storage issues. FAO investment center, country highlights.
- GoP, 2019. Economic Survey of Pakistan, Ministry of Finance, Government of Pakistan, Islamabad.
- GoP, 2020. Economic Survey of Pakistan, Ministry of Finance, Government of Pakistan, Islamabad.
- GoP, 2018. Punjab agriculture policy 2018. Lahore.
- IIASA.ac. 2008. IIASA Arguments- Prediction Error: Diet.
- Jan, I., S. Khan and T. Mahmood. 2021. Effects of rainwater harvesting on yield of wheat and maize crops in Pakistan. *Pakistan Journal of Agriculture, Agricultural Engineering and Veterinary Sciences*, 37 (1): 29-35.
- Kebede, D., M. Ketema and N. Dechassa. 2017. Disparity in adoption of wheat production technology packages in Eastern Ethiopia, *Review of Agricultural and Applied Economics*, 20 (2): 22-29.
- Lien, G., S. C. Kumbhakar and J. B. Hardaker. 2010. Determinants of off-farm work and its effects on farm performance: the case of Norwegian grain farmers. *Agricultural Economics*, 41 (6): 577-586.
- Musafiri, I. and P. Sjolander. 2018. The importance of off-farm employment for smallholder farmers in Rwanda. *Journal of Economic Studies*, 45 (1): 14-26.
- Shah, T., A. Z. Khan, M. Numan, W. Ahmad, M. Zahoor, M. Ullah. and A. Jalal. 2017. Nutrient uptake and yield of wheat varieties as influenced by foliar potassium under drought condition. *Cercetari Agronomic in Moldova*, 50 (2): 5-20.
- Woldehanna, T. 2000. Economic analysis and policy implications of farm and off-farm employment: a case study in the Tigray region of Northern Ethiopia. Ph.D. Thesis, Wageningen University, Netherland.

(Received: September 21, 2020; Accepted: August 12, 2021)