The Impact of Adopting Improved Wheat Technology on the Productivity and Income of Houeholds in Misha District, Southern Ethiopia

Melkamu Tilaye
Ethiopian Institute of Agricultural Research, Wondogenet Agricultural Research Center, P.O. Box. 198, Shashemene, Ethiopia

Abstract
Wheat rust disease is a major constraint of wheat production in Ethiopia. Thus, this study is carried out to examine the impact of rust resistant wheat varieties on the productivity and income of households in Misha district in Southern Ethiopia. Using a household survey, cross-sectional data were collected from 387 randomly selected households. Descriptive statics and propensity score matching methods were used for data analysis to achieve the objectives of the study. Propensity score matching method was used to measure the impact of adopting rust-resistant improved wheat varieties on productivity and income of households. The study has found that age and education level of household head, land size, livestock holding, frequency of extension contact, and access to credit services were factors that significantly affected adoption of rust-resistant improved wheat varieties. Using the propensity score matching method, the study found that the adoption of rust-resistant improved wheat had a positive impact on average wheat productivity and income of households. Moreover, the results were insensitive to unobserved heterogeneity bias. This indicates that adoption of the technology has a positive contribution to households' wellbeing. Therefore, government and concerned bodies should better give due attention to the development, dissemination, and scaling up of rust-resistant improved wheat varieties.
Keywords: Adoption, Impact, Propensity Score Matching, Wheat.
DOI: 10.7176/EJBM/14-19-03
Publication date: October 31st 2022

Introduction

Wheat (Triticum aestivum) is the world's leading cereal grain which is used by more than one-third of the population of the world as a staple food (FAO, 2018). Wheat is one of the most important food grain crops both in production and nutrition and thought that it is the first crop ever to be cultivated and plays a major role in human's economic and social development worldwide (Thabet and Najeeb, 2017; USDA, 2018). It is the major cereal crop that plays a significant role in feeding a hungry world and improving global food security (Mengistu and Belay, 2016).

Wheat is a vital staple food crop in Ethiopia and since 2005 the country has been the largest producer of wheat in sub-Saharan Africa (Hodson *et al.*, 2020). Wheat is grown on 1.6-1.8 million ha, annually, with an estimated 5 million farming households dependent on the crop (CSA, 2018). There are two wheat species, which are dominantly grown in Ethiopia. These two economically important wheat species are bread wheat and durum wheat (Hodson *et al*., 2020). Durum wheat is native to Ethiopia and mainly grown in the Central and Northern highlands and bread wheat is a recent introduction to Ethiopia. Like other cereal crops produced in the country, wheat has various uses. In Ethiopia, wheat grain is used in the preparation of different traditional and modern processed food products such as injera and other industrial processed products like pasta and macaroni. Moreover, wheat straw is commonly used as a roof tacking material and as a feed for animals. So, wheat is an important cereal crop that should get the emphasis on both its production and its marketing (Bekele *et al*., 2014; Aklilu *et al*., 2015).

Demand for wheat is growing rapidly in Ethiopia, reflecting population growth and shifting dietary patterns linked to urbanization that are mirrored across other eastern and southern African countries (Mason *et al.*, 2015). On average majority of farmers produces 2.7 ton/ha which is less than the yield attained at the research stations and on-farm, that is 7 ton/ha and 6 ton/ha, respectively (Fisseha *et al.*, 2020). Despite this low productivity, the demand for wheat has been increasing in both urban and rural areas of the country (Bekele *et al.*, 2014). Although there exist recent productivity gains, shortfalls remain and drastically narrowing the gap between supply and demand; self-sufficiency in wheat production is a high national priority. Food security problems and the need to decrease spending of scarce foreign currency reserves on costly wheat imports have paramount importance to the Government of Ethiopia (Hodson *et al*., 2020).

During the 2019 production season, the national average wheat productivity of Ethiopia was 2.97 tons per hectare (t/ha), which was lower than the average productivity of Zambia and Egypt whose productivity was 6.68 t/ha and 6.38t/ha, respectively (FAOSTAT, 2019). The low productivity is attributed to several factors including biotic (diseases, insects, weeds, and others) and abiotic (low and high rainfall, temperature, low adoption of new agricultural technologies). Among the biotic factors, wheat rust has been the most devastating disease in Hadiya zone and Ethiopia causing up to 100% yield losses on susceptible varieties during the epidemic year (Belayneh *et al*., 2012; Alemayehu *et al*., 2020).

Wheat production in Ethiopia faces various climate-related constraints climate-related challenges such as frost, drought, and rust. Of which, wheat rust disease is one of the major wheat production problems which has been imposing a negative impact on wheat production and productivity in Ethiopia including Misha district of Hadiya zone. Since all improved wheat varieties are not rust disease resistant, these susceptible varieties have been attacked during the occurrences of rust diseases. According to Fetch and Callum (2014), Lidwell-Durnin and Lapthorn (2020), wheat rust diseases are the most damaging disease worldwide and are widely distributed across wheat-growing regions. According to Messay *et al*. (2013), Getnet *et al*. (2020), and Hodsosn *et al*. (2020), stem rust and yellow rust are the most biotic constraints to wheat production in Ethiopia. Recurrent rust epidemics have caused large-scale production losses and low wheat productivity in Ethiopia. Alemayehu *et al*. (2020) indicated that wheat rust is one of the major constraints of wheat production in Hadiya zone and found that Misha district had a high prevalence of wheat rust diseases. To improve wheat productivity and reduce the effect of wheat rust diseases, the government and many stakeholders were engaged in promoting and popularization of newly released improved wheat varieties. However, the studies on the impacts of these technologies on farmers' livelihoods were limited (Bekele *et al.,* 2014; Tesfaye *et al*., 2016). Although Misha district is one of the most potential wheat-producing districts of Hadiya zone and is familiar with the adoption of improved wheat technologies; Studies on the impact of adopting improved wheat technologies in Misha district as well as Ethiopia were limited.

Furthermore, there were some empirical studies on the impacts of improved crop varieties, but most studies have analyzed the impact of improved wheat on productivity or other outcome variables on households and did not clearly state whether these improved wheat varieties were rust-resistant or not since all improved wheat is not rust-resistant. According to Zewdu *et al*. (2017), While significant researches on the adoption and impacts of improved crop varieties exist, most studies have analyzed yield effects in general without distinguishing between different varietal traits and characteristics such as improved rust resistance with improved susceptible wheat varieties. For example studies by Tesfaye *et al*.(2016) studied the impact of bread wheat varieties on productivity and income in Arsi zone, Tesfaye *et al.* (2018) examined the impact of improved wheat variety on productivity in Oromia Regional State, Fitsum (2018) studied the impact of improved wheat varieties adoption on productivity in different Agroecological Zones of Ethiopia, Hiwot (2018), studied the impact of the adoption of improved wheat varieties on productivity and food security in Girar Jarso woreda, Northen Shewa zone of Oromiya region. All of these above-mentioned studies examined the impacts of improved wheat varieties in general, without distinguishing between different varietal characteristics like whether these varieties were improved varieties that are rust-resistant or not. Because all improved varieties are not rust-resistant and as a result, most improved varieties have been devastated due to the occurrence of rust disease as reported in different literature. To fill this research gap, this study examined the impact of rust-resistant improved wheat varieties on the productivity and income of households in Misha district.

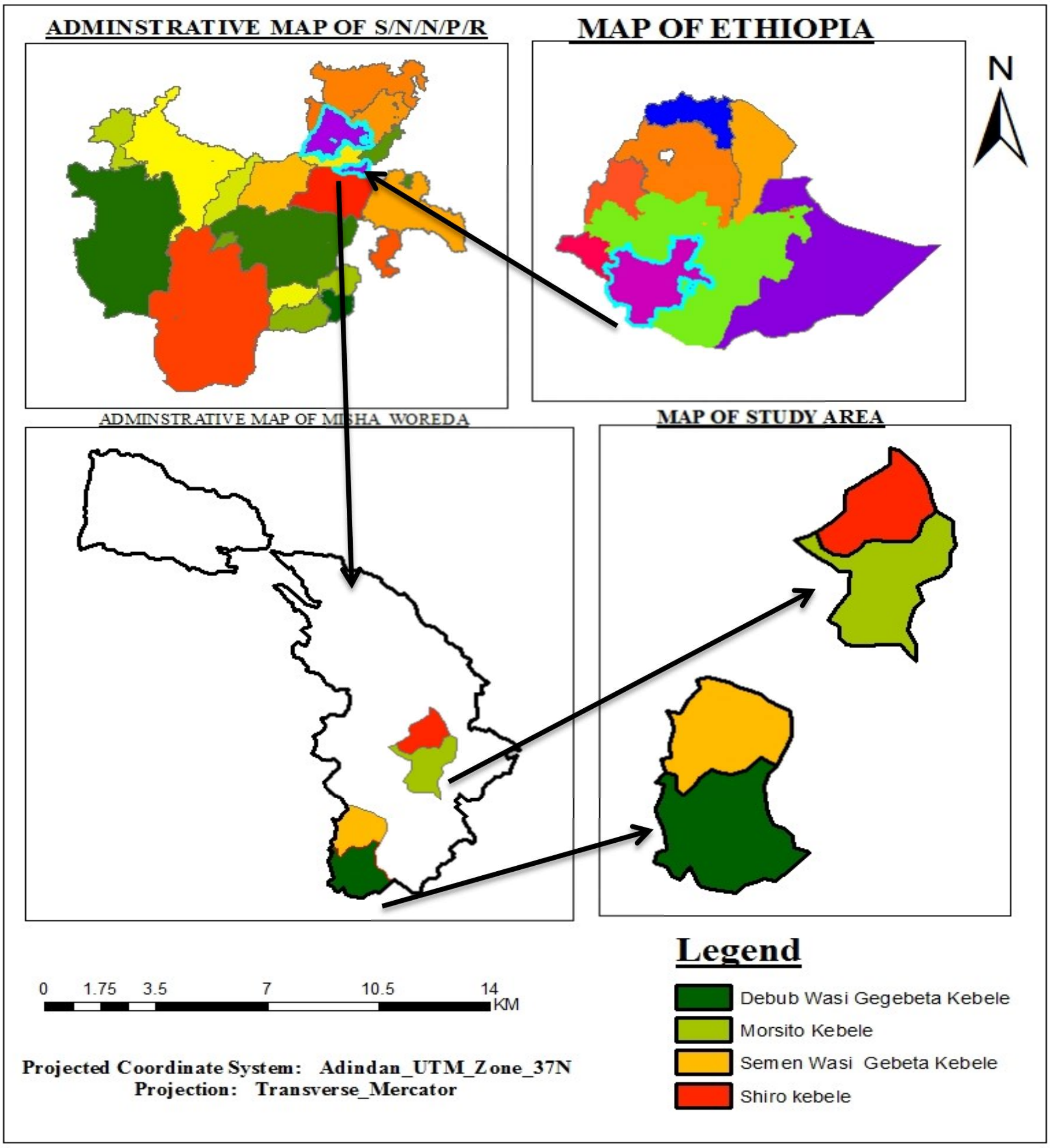
Moreover, most studies fail to accurately define the region of common support during the implementation of the propensity score matching method of impact evaluation. According to Heckman *et al*. (1997a), a violation of the common support condition is a major source of evaluation bias. In addition to this, most studies measured the impacts of improved wheat technology packages using propensity score matching, did not test the estimated average treatment effects were free from unobserved bias (example studies by: Tsegaye and Bekele, 2012; Tesfaye *et al*., 2016; Hiwot, 2018; Tesfaye *et al*., 2018; Fitsum, 2018; Baye *et al.,*2019), the afro-mentioned empirical studies used propensity score matching for impact evaluation but those studies did not accurately define the region common support and did not conduct sensitivity test for unobserved factors that affect treatment and outcome variables. Unlike previous studies, this study employed propensity score matching and filled these research gaps. Thus, to fill these gaps, this study had correctly defined common support region and test the existence of unobserved characteristics or variables which affect assignment into treatment (in this case, adoption of rust-resistant improved wheat varieties) and the outcome variable that is, wheat productivity and income simultaneously using sensitivity analysis. Therefore, this study was designed to assess the impact of adopting rust-resistant improved wheat varieties on the productivity and income of households in Misha district.

2. Materials and Methods

2.1. Description of the Study Area

This study was conducted in Misha district, which is found in Hadiya administrative zone of the Southern Nations Nationalities and Peoples Regional State of Ethiopia. The district is located at a distance of 253 km away from Addis Ababa, 207 km from Hawassa, and 18 km from Hossana. The geographic location of the district is at 7°08' N latitude and 37°81' E longitude. It is bounded by Silte zone in East, Guraghe Zone in North, Gombora Woreda in South, and Gibe Woreda in West direction. The total area of the district covers approximately 304.07 km^2 with an average population density of about 252 households per kilometer square. The land feature of Misha district is characterized by sloppy and flat, in addition it is characterized by a humid tropical climate. The mean annual maximum and minimum temperatures of the district are 24^0C and 13^0C respectively. About 70 percent of the land of Misha district lies in the weyena dega (mid-altitude between 1500 and 2500 m.a.s.l), 20 percent of the district lies in the dega (high altitude above 2500 m.a.s.l.) and 10 percent lies in kola (low altitude below 1500 m.a.s.l) agro-climatic zones. Agricultural activity is the main means of livelihood for the majority of Misha district population. In terms of economic activities, the Woreda community fully experienced animal rearing and crop production (mixed farming system). The most dominant cereal crops produced in this district are wheat, teff, maize, sorghum, bean, pea, and other cash crops like chat, coffee, and vegetables (Shigute and Anja 2018; Girma *et al*., 2019).

Map of the Study Area



Source: Ethio_map of Shapefile

2.2. Sample Size and Sampling procedure

The required sample size was determined by using Yamane (1967) sample size determination formula. A simplified formula by Yamane (1967) was used to determine the required sample size at 95% confidence level, and 5% (0.05) level of precision.

$$n = \frac{N}{1 + N(e)^2}$$

Where:
n = the required sample size
N = population size
e = is the level of precision

In Misha district, there were about 11,683 wheat producer households of which 10,318 were male and 1,365 were female-headed (Misha district agricultural development office report, 2021). Therefore, the sample size was determined as follows:

$$\boldsymbol{n = \frac{11,683}{1 + 11,683(0.05)^2}}$$

= **387** households

The sampling method used for this study was a mixed method of purposive and simple random sampling, which involves three stages. First, purposive selection of potential wheat production kebeles of the woreda was conducted based on the data on production potential of each kebele from the woreda agricultural and rural development office to get sufficient data for impact evaluation. In the meantime, four wheat potential production kebeles were selected. Then at the second stage on these four selected kebeles: households were stratified into two strata. That is households who cultivate rust-resistant improved wheat varieties and non-rust resistant wheat varieties, which was conducted in collaboration with development agents of the respective kebeles. Finally, a sample of households from each stratum was selected through a simple random sampling technique based on probability proportional to the size of the population for each kebeles.

The sample size for each kebele was determined as:

$$\text{ni} = n(\frac{Ni}{N})$$

Where: **ni** is the sample size from each selected kebele, **n** is the total sample size of the study, which is the sum of the sample size of the four kebeles, and **Ni** is total wheat farm households in respective kebeles, and N is the total population or wheat farm households of the four kebeles combined. Table 1 below shows the proportional sample size of four kebeles.

Table 1. Distribution of sample households in each kebele

Sample Kebeles	Household size	Adopters		Non-adopters		Total Sample
		Total	Sample	Total	Sample	
Debub Wasgebeta	576	230	43	346	61	104
Semen Wasgebeta	568	227	41	341	61	102
Morsito	519	208	38	311	56	94
Shiro	485	194	35	291	52	87
Total	2148	859	157	1289	230	387

Source: Misha woreda agriculture office and own computation (2021).

2.3. Method of Data Collection

For this study household survey technique used to collect the primary data from all sample respondents. On the other hand, secondary data were collected, from woreda and kebele agricultural and development offices reports, review of different documents such as research and reports of different organizations, published journals articles, books, proceedings related to this study. A household survey was conducted by using a structured questionnaire. Before commencing the data collection process training and orientation about data collection were given to enumerators. Then data collection was conducted with structured questionnaires using trained enumerators from a sample of 387 households using face to face interview method. During the time of data collection monitoring and controlling of data collection activities of all enumerators was done by the supervisor.

2.4. Method of Data Analysis

The study employed descriptive, and econometrics methods of data analysis.

2.4.1. Descriptive analysis

Descriptive statistics were used to provide a summary statistic of the demographic, socio-economic, and institutional variables. In addition, to those summary statics, the chi-square test was used to identify the proportional difference in categorical variables between adopters and non-adopters, and a t-test was used to test mean differences of the continuous variables between groups of rust-resistant improved wheat variety adopters and non-adopters.

2.4.2. Econometric model

Propensity score method

Propensity score matching (PSM) method is a quasi-experimental method to estimate causal treatment effects.

PSM is a method to match program participants with non-participants typically using individual observable characteristics. Each program participant is paired with a small group of non-participants in the comparison group that is most similar in the probability of participating in the program (Becker and Ichino, 2002). It matches control groups to treatment groups based on observed characteristics or by propensity scores; the closer this score, the better the match. Unlike econometric regression methods, PSM compares only comparable observations and does not rely on parametric assumptions to identify the impacts of programs and it does not impose a functional form of the outcome, thereby avoiding assumptions on functional form and error term distributions, e.g., linearity imposition, multicollinearity, and heteroskedasticity issues. In addition, the matching method emphasizes the problem of common support, thereby avoiding the bias due to extrapolation to the non-data region (Becker and Ichino, 2002; Caliendo and Kopeining, 2008).

Participants and non-participants of the technology adoption may not be directly comparable, since participants and nonparticipants usually differ even in the absence of treatment which means outcomes of the two groups differ even in the absence of the treatment. This problem is known as selection bias. Therefore, before proceedings to future counterfactuals, first need for comparability establishment to avoid initial difference (Caliendo and Kopeinig, 2008). The PSM approach tries to capture the effects of different observed covariates X on adoption, in a single propensity score or index. Then, outcomes of adopters and non-adopters with similar propensity scores are compared to obtain the adoption effect (Rosenbaum and Rubin, 1983). For the aforementioned reasons, the study uses the propensity score matching method to measure the impact of rust-resistant improved wheat varieties on the wheat productivity and income of households.

I. Propensity Score Estimation

The first stage in the propensity score matching method is to estimate propensity scores. When estimating the propensity score, two choices have to be made. The first one concerns the model to be used for the estimation and the second one the variables to be included in this model. For the binary treatment case, where we estimate the probability of participation versus nonparticipation, logit and probit models usually yield similar results (Caliendo and Kopeinig, 2008). The matching strategy builds on the CIA, requiring that the outcome variable(s) must be independent of treatment conditional on the propensity score. Hence, implementing matching requires choosing a set of variables *X,* that credibly satisfies this condition. Heckman *et al.* (1997a) and Dehejia and Wahba (1999) show that omitting important variables can seriously increase bias in resulting estimates. Only variables that influence simultaneously the participation decision and the outcome variable should be included (Sianesi, 2004; Smith and Todd, 2005). Thus, the first step in propensity score matching is to predict the propensity score using a logit/probit model. For this study logit model was selected to estimate propensity scores, since this model is an extremely flexible and easily used model from a mathematical point of view and results in a meaningful interpretation (Hosmer and Lemeshew, 2000). The mathematical formulation of the logit model is as follows:

$$Li = \text{Ln}\left(\frac{pi}{1-pi}\right) = Zi = \beta o + \beta 1X1 + \cdots + \beta nXn + Ui \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Where: Li is the log of the odds ratio, L is the logit, Zi: is a function of n-explanatory variables, i.e., $Z_i = \beta_0 + \beta_1X_1 + \cdots + \beta_nX_n$, Pi probability of adoption which, ranges between 0 and 1.

II. Defining Overlap and Common Support

Heckman *et al.* (1997a) point out that a violation of the common support condition is a major source of evaluation bias. Only the subset of the comparison group that is comparable to the treatment group should be used in the analysis (Dehejia and Wahba, 1999). According to Caliendo and Kopeinig (2008), the region of common support can be determined by comparing the minima and maxima of the propensity score in treated and control groups. The common support region is the region within the minimum and maximum propensity scores of treated (adopters of rust-resistant improved wheat varieties) and comparison groups (non-adopters), respectively, and it will be demarcated by cutting off those observations whose propensity scores are smaller than the minimum of the treated group and greater than the maximum of the comparison groups.

III. Choosing a Matching Algorithm

The most commonly used matching algorithms are:

Nearest Neighbor Matching (NN): the individual from the comparison group is chosen as a matching partner for a treated individual that is closest in terms of the propensity score. Several variants of NN matching are proposed, e.g., NN matching with replacement and without replacement. In the former case, an untreated individual can be used more than once as a match, whereas in the latter case it is considered only once. Matching with replacement involves a trade-off between bias and variance. If we allow replacement, the average quality of matching will increase and the bias will decrease (Caliendo and Kopeinig, 2008).

Caliper Matching: Nearest neighbor matching faces the risk of bad matches if the closest neighbor is far away. This can be avoided by imposing a tolerance level on the maximum propensity score distance (caliper). Hence, caliper matching is one form of imposing a common support condition. When we use caliper matching, bad matches are avoided and the matching quality rises. Applying caliper matching means that an individual from the comparison group is chosen as a matching partner for a treated individual that lies within the caliper (propensity range) and is closest in terms of propensity score (Caliendo and Kopeinig, 2008).

Radius Matching: a variant of caliper matching which is called radius matching. The basic idea of this variant is to use not only the nearest neighbor within each caliper but all of the comparison members within the caliper. A benefit of this approach is that it uses only as many comparison units as are available within the caliper and therefore allows for usage of extra (fewer) units when good matches are (not) available. Hence, it shares the attractive feature of oversampling but avoids the risk of bad matches (Dehejia and Wahba 2002).

Kernel Matching: Kernel matching (KM) is a nonparametric matching estimator that uses weighted averages of all individuals in the control group, on the choice of the kernel function to construct the counterfactual outcome. Thus, one major advantage of this approach is the lower variance which is achieved because more information is used. A drawback of these methods is that possibly observations are used that are bad matches. Hence, the proper imposition of the common support condition is of major importance for kernel matching (Caliendo and Kopeinig, 2008). As Smith and Todd (2005) note, kernel matching can be seen as a weighted regression of the counterfactual outcome on an intercept with weights given by the kernel weights. Weights depend on the distance between each individual from the control group and the participant observation for which the counterfactual is estimated. The average places higher weight on persons close in terms of the propensity score of a treated individual and lower weight on more distant observations. When applying kernel matching, one has to choose the kernel function and the bandwidth parameter (Caliendo and Kopeinig, 2008).

The performance of different matching estimators varies case-by-case and depends largely on the data structure at hand (Zhao, 2003). For this study, the choice of matching algorithm is done by using criteria such as: the amount of matched sample lies on-support, Psedou-$R^{2,}$ and covariate balance after matching.

IV. Testing the balance of propensity scores and covariates

At this stage, whether the matching procedure is able to balance the distribution of the relevant variables in both the control and treatment groups is checked. This approach is to compare the situation before and after matching and check if there remain any differences after conditioning on the propensity score. If there are differences, matching on the score was not completely successful and remedial measures have to be done, e.g., by including interaction terms in the estimation of the propensity score. Rosenbaum and Rubin (1983) state that:

X Ц $D|P(D=1|X)$……………………………………………………… (2)

This means that after conditioning on P (D = 1|X), additional conditioning on X should not provide new information about the treatment decision. Hence, if after conditioning on the propensity score there is still a dependence on X, this suggests either misspecification in the model used to estimate P (D = 1|X) (Smith and Todd, 2005) or a fundamental lack of comparability between the two groups (Blundell *et al.*, 2005).

Standardized Bias

One suitable indicator to assess the distance in marginal distributions of the X variables is the standardized bias (SB) suggested by Rosenbaum and Rubin (1985). For each covariate X, it is defined as the difference of sample means in the treated and matched control subsamples as a percentage of the square root of the average of sample variances in both groups. The SB before matching is given by:

$$\mathrm{SB}_{\text{before}} = 100\ \frac{\bar{X}_1-\bar{X}_0}{\sqrt{0.5\cdot(V_1(X)+V_0(X))}}\text{……………………………………………… (3a)}$$

The SB after matching is given by:

$$\mathrm{SB}_{\text{after}} = 100\ \frac{\bar{X}_{1M}-\bar{X}_{0M}}{\sqrt{0.5\cdot(V_{1M}(X)+V_{0M}(X))}}\text{…………………………………………… (3b)}$$

Where: X 1 (V 1) is the mean (variance) in the treatment group before matching and X 0 (V 0) the analogue for the control group. X 1M (V 1M) and X 0M (V 0M) are the corresponding values for the matched samples. Rosebaum and Rubin (1985) suggested that standardized bias less than 20% after matching indicates covariates are balanced; thereby there is no more mean difference exist between adopters and non-adopters.

t-Test

A similar approach uses a two-sample *t*-test to check if there are significant differences in covariate means for both groups (Rosenbaum and Rubin, 1985). Before matching, differences are expected, but after matching the covariates should be balanced in both groups and hence no significant differences should be found. The *t*-test might be preferred if the evaluator is concerned with the statistical significance of the results.

Joint Significance and Pseudo-R^2

Sianesi (2004) suggested comparing participants and matched nonparticipants, using pseudo-R^2s before and after matching. The pseudo-R^2 indicates how well the regressors X explain the participation probability. After matching there should be no systematic differences in the distribution of covariates between both groups and therefore the pseudo-R^2 should be fairly low. Moreover, also suggested a likelihood ratio test on the joint significance of all regressors in the probit or logit model should not be rejected before and should be rejected after, matching.

V. The Average Treatment Effect

The average treatment effect on the treated (ATT) is given by the difference in mean outcome (grain yield and farm income) of matched adopters and non-adopters that have common support conditional on the propensity score. The mean effect of adopting rust-resistant improved wheat varieties, therefore, is given by:

$Ti = E(Y_1|Di = 1) - E(Y_0|Di = 0)$…………………………………………… (4)

Where: Ti, is a treatment effect, Y is the outcome (grain yield and farm income) and Di is a dummy whether household *i,* has the treatment or not. However, one should note that Y (Di = 1) and Y (Di = 0) cannot be observed for the same household at the same time. Due to this fact, estimating the individual treatment effect was not possible and one has to shift to estimating the average treatment effects of the population rather than the individual one. Therefore, following Takahashi and Barrett (2013), the average treatment effect on treated (ATT) can be defined as:

$ATT = E\{Y_1 - Y_0|D = 1\} = E(Y_1|D = 1) - E(Y_0|D = 1)$……………………… (5)

Where: Y_1 = the outcome in the treated condition, Y_0 = the outcome in the control condition; and D = Dummy, indicator variable denoting adoption of rust-resistant improved wheat varieties.

We can observe the outcome variable of adopters E (Y_1| D = 1), but we cannot observe the outcome of those adopters had they not adopted E (Y_0 |D = 1), and estimating the ATT using equation (5) may therefore lead to biased estimates (Takahashi and Barrett, 2013). Propensity score matching relies on an assumption of conditional independence where conditional on the probability of adoption, given observable covariates, an outcome of interest in the absence of treatment Y_0 and adoption status, D are statistically independent (Takahashi and Barrett, 2013). Rosenbaum and Rubin (1983) define the propensity score or probability of receiving treatment as:

$p(X) = \text{pr }(D = 1)|X$…………………….……….……….……..………… (6)

Another important assumption of PSM is the common support condition, which requires substantial overlap in covariates between adopters and non-adopters, so that households being compared have a common probability of being both an adopter and a non-adopter, such that 0 < $p(X)$ < 1 (Takahashi and Barrett, 2013). If the two assumptions are met, then the PSM estimator for ATT can be specified as the mean difference outcomes of the adopters matched with non-adopters who are balanced on the propensity scores and fall within the region of common support, expressed as:

$$E[(Y_1|D = 1) - E(Y_0|D = 0)] = \tau_{ATT} + E[(Y_0|D = 1) - E(Y_0|D = 0)]$$

The difference between the left-hand side of the equation and τATT is the so-called 'selection bias'. The true parameter ATT is only identified if there is no selection bias:

$E[(Y_0|D = 1) - E(Y_0|D = 0)] = 0$ thereby,

$ATT = E(Y_1|D = 1, p(X)) - E(Y_0|D = 0, p(X))$………………..…………..… (7)

VI. Sensitivity Analysis

The estimation of treatment effects with matching estimators is based on selection on observables assumption. However, if there are unobserved variables that affect assignment into treatment and outcome variables simultaneously, a hidden bias might arise (Rosenbaum, 2002). Sensitivity is to check how strongly an unmeasured variable influences the selection process to undermine the implications of matching analysis (Caliendo and Kopeinig, 2008).

The main question to be answered in sensitivity analysis is whether inference about treatment effects may be altered by unobserved factors. In other words, one wants to determine how strongly an unobserved variable influence the selection process in order to undermine the implications of matching analysis.

The bounding approach of sensitivity analysis proposed by Rosenbaum (2002) is used to check the sensitivity of the estimated average treatment effect on treated with respect to deviation from the Constant Independence Assumption. The bounding approach does not test the confoundedness assumption itself. However, it provides evidence on the degree to which the significance of results hinges on this assumption. If the results turn out to be sensitive, the evaluator might have to think about the validity of his identifying assumption and consider other estimation strategies.

Diagnostics tests

1. Testing the goodness of fit of logistic regression

The goodness-of-fit tests, which can help to decide whether the model used in the analysis was correctly specified or not. According to Lomax and Hahs-Vaughn (2012) testing the significance of the overall logistic regression model enables to determine the overall model fit and provides evidence of the extent to which the predicted values accurately represent the observed values. This study used goodness of fit tests like Hosmer–Lemeshow goodness-of-fit test to test the specified logistic regression model that was used for propensity score estimation have a good fit.

Hosmer–Lemeshow goodness-of-fit test: Hosmer-Lemeshow goodness-of-fit test is one tool used to examine the overall model fit. The Hosmer-Lemeshow statistic is computed by dividing cases into deciles (i.e., 10 groups) based on their predicted probabilities. Then a chi-square value is computed based on the observed and expected frequencies. Statistically non-significant results for the Hosmer–Lemeshow test indicate the model has an acceptable fit (Lomax and Hahs-Vaughn, 2012). The test statics is based on the P-value of the Hosmer-Lemeshow goodness-of-fit test. If test statics have a low p-value (below 0 .05), then the model is not fit. If it's high (greater than 0.05), then the model passes the test, that is the model has a good fit.

2. Multicollinearity test

The term multicollinearity originally meant the existence of a perfect or exact linear relationship among some or all explanatory variables of a regression model. Multicollinearity is perfect if the regression coefficients of the explanatory variables are indeterminate and their standard errors are infinite. If multicollinearity is less than perfect, the regression coefficients, although determinate, possess large standard error, which means the coefficients cannot be estimated with great precision or accuracy. Multicollinearity is a question of degree and not of kind. The meaningful distinction is not between the presence and the absence of multicollinearity, but between its various degrees (Gujarati, 2004). The no-multicollinearity problem applies to logistic regression models with multiple predictors just as it was in multiple regressions (Lomax and Hahs-Vaughn, 2012).

This study used two methods or rules to detect the existence of multicollinearity in the logistic regression model. For continuous explanatory variables existence of multicollinearity was tested using variance inflated factor and tolerance. For dummy or discrete variables explanatory variables, Contingency coefficient is used to detect the existence of multicollinearity problems.

For continuous explanatory variables, multicollinearity was detected with variance inflation factor (VIF). The speed with which variances and covariance increase can be seen with the variance-inflating factor (VIF), which is defined as:

$$\mathrm{VIF} = \frac{1}{\left(1 - R_i^2\right)}$$

Where, Ri^2 represents the coefficient of determination for regressing, the i^{th}, independent variable on the remaining ones. VIF shows how the variance of an estimator is inflated by the presence of multicollinearity.

The larger the value of VIF, the more collinear the predictor variable. As a **rule of thumb**, if the VIF of a variable exceeds 10, which will happen if its R^2 exceeds 0.90, that variable is said to be highly collinear (Gujarati, 2004).

2.2. Multicollinearity test for dummy variables

Similarly, for dummy variables contingency coefficients test were employed using the following formula.

$$\boldsymbol{C} = \sqrt{\frac{\boldsymbol{x^2}}{\boldsymbol{n + x^2}}}$$

Where: C is contingency coefficient, is the chi-square value and n=total sample size. For dummy variables if the value of contingency coefficients is greater than 0.75 the variable is said to be collinear.

Definition of variables and hypothesis

Outcome variables: wheat productivity and income were used to estimate the impact of rust-resistant improved wheat varieties. Wheat productivity is the wheat production obtained from one hectare. While income is the annual total income from wheat production of households measured in Ethiopian birr.

Dependent variable: the dependent variable is the adoption decision of rust-resistant improved wheat varieties. The variable takes the value of 1 for the household that cultivated rust resistant improved wheat varieties during the 2020/2021 production year and 0 for a household that did not cultivate rust-resistant improved wheat varieties.

Independent variable: for this study independent variables were selected based on the literature of past research findings on the adoption and impact of agricultural technology. Major variables expected to influence adoption and outcome variables of improved wheat varieties were selected.

Sex: It is a dummy variable that takes a value of 1 if the household head is male and 0, otherwise. Tesfaye *et al.* (2016), found that male-headed households were more likely to adopt improved wheat varieties. Being male was expected to have a positive influence on the adoption of rust-resistant improved wheat varieties.

Age of household head: It is a continuous variable measured in the number of years of the household head. It has an important role in the production process. As the age of the household head increases, the probability of adopting a technology likely to decrease. Milkias (2020) indicated that the age of the household head was a determinant factor for adopting high-yielding wheat varieties. Fear of risk for technologies is more observed in older farmers than younger farmers. Therefore, age of the household head was hypothesized to have a negative influence on adoption.

Education level of household head: level of education is assumed to increase a farmer's ability to obtain process, and use information relevant to the adoption of rust-resistant improved wheat varieties. an increase in the level of education of a household increases the probability of adopting rust-resistant improved wheat varieties. Hiwot (2018), indicated that the education level of farmers had a positive and significant influence on the adoption of improved wheat varieties. On the contrary, Tesfaye *et al.* (2016) indicated that the level of education of a household head decreases the probability adopting of improved wheat varieties. This study was hypothesized that education level has positively affected the decision to use rust-resistant improved varieties.

Family size: Adopting improved varieties requires labor for preparation of land, management practices, and proper harvesting so a large number of family sizes has a positive relationship for the decision to use rust-resistant improved varieties. Leake and Adam (2015) indicated that family size had positively affected the adoption decision of improved wheat varieties. This study was hypothesized a large number of family sizes have positively influenced the adoption decision of rust-resistant improved wheat varieties.

Farm size: It is a continuous variable. It refers to the area of cultivated land possessed by the respondents or their families. The study assumes that the larger the farm size the farmer has the opportunity to increase income and productivity since farmers who have large farm has a better opportunity to either renting part of their land for an income-generating activity or gain better farm income from production different crops in his farmland as compared to farmers with low landholding. Therefore, large landholding enables farmers to have better farm income to purchase improved farm technologies like rust-resistant improved wheat varieties. The study was hypothesized land size expected to have positively influenced the adoption of rust-resistant improved wheat varieties.

Farming experience: previous experience of farmers can be expected to either enhance or diminish their level of confidence. It is anticipated that with more experience, farmers could become risk-aversive regarding the adoption of specific wheat varieties (Tesfaye, 2001). On the other hand, it is assumed with increased years of experience in farming, farmers are generally better able to assess the relevance of new technologies. This often comes from their interactions with their neighbors and the outside people. Because of their experience, they also tend to be better placed to acquire the needed skills to use the technologies compared with younger ones (Tolesa, 2014). This study was hypothesized that having more farming experience has positively influenced the decision to use rust-resistant improved wheat varieties.

Livestock holding: It is the total number of livestock holding of the farmer and it is measured in the Tropical Livestock Unit (TLU). Livestock holding is taken as an asset; farmers who hold large livestock implies they can easily generate an income that enables them to cover the required cost of improved rust-resistant seed and other farm technologies. This study was hypothesized having a large livestock number has positively influenced the adoption decision of rust-resistant improved wheat varieties.

Frequency of extension service: The efficiency of extension service depends on the frequency of those extension agent visits of a given household in a specific crop year. According to Kidane (2001), a high frequency of extension contact accelerates the effective dissemination of information that enhances the adoption of new agricultural technologies. Therefore, this study expected that more access to extension services has positively influenced the decision to use rust-resistant improved varieties.

Access to credit: access to credit measured as a dummy variable takes a value of 1 if the farmer has access to credit and 0 otherwise. Access to credit can solve farmers' financial constraints (Tesfaye, 2001). Therefore, this study was hypothesized that access to credit services has a positive influence on farm households' adoption decision of rust-resistant improved wheat varieties.

Membership of farmers' cooperative: It is a dummy independent variable represented by 1 if the household head participates in membership in the farmer cooperative and 0 otherwise. This study was hypothesized that being a participant in a farmer cooperative has positively influenced the adoption of rust-resistant improved wheat varieties.

Distance to nearest markets: It is a continuous variable measured in walking minutes that the household travel to reach the nearby market. It is expected that a short distance to the nearest market has a positive contribution to the adoption of improved varieties and vice versa. This study was hypothesized as distance to the nearest market increases it would have negatively influenced the adoption decision of rust-resistant improved wheat varieties.

Dependency ratio: It is a measure of the number of dependents aged zero to 14 years and over the age of 65 years, compared with the total population aged 15 to 64 years. This demographic indicator gives insight into the number of people of non-working age, compared with the number of those of working age. It is also used to understand the relative economic burden of the workforce. The number of dependents in the household may reduce the household income available for investments, thus discouraging adoption. Adeoti (2008) indicated that the dependency ratio has a negative effect on the probability of the adoption of irrigation technology. This study was hypothesized dependency ratio has negatively affected the adoption of rust-resistant wheat verities.

CHAPTER THREE

3. RESULTS AND DISCUSSION

3.1. Wheat Varieties used by Sample Households

The major improved wheat varieties used by sample households on the cropping season 2020/21 were Danda'a (also known as Danphe), Digalu, Huluka, Ogelcho, and Shorima. From these wheat varieties, Denda'a is among the improved wheat varieties that were released as rust disease-resistant improved wheat in 2010. Using observation, key informant survey, and focus group discussion with farmers from all these varieties Danda'a wheat variety was the most rust-resistant improved wheat variety but other improved wheat was seriously attached by the occurrence of wheat rust diseases. Most farmers have planted Oglecho improved wheat variety in anticipation to get a high wheat yield. But in this year, due to the occurrence of severe wheat rust diseases, this variety was attached by rust disease and as a result, producers of wheat were experienced a huge loss in wheat production.

3.2. Descriptive Statistics of Variables

3.2.1. Descriptive statistics results of categorical variables

Sex of the household head: as shown in table 2 below, from the total sample household's female share 27.39% and the rest 72.61% were male-headed households. 23.57% of rust-resistant improved wheat variety (RURWV) adopters were female-headed and the rest 76.43% were male-headed households. On the other hand, about 30% of non-adopters of RURWV were female-headed and the rest 70% were male-headed households. The Pearson chi-square test analysis results of RURWV adopters and non-adopters regarding the sex of head shows that there was no statistically significant difference in the proportion of sex between rust-resistant improved wheat variety adopters and non-adopters.

Access to credit: households were asked to answer the question of whether they had access and availability to credit services in their locality. As indicated in table 2 below, from the total sample household's 64.86% of households had access to credit and the rest 35.14% had no access to credit services. 76.43% of RURWV adopters had access to credit services and 23.57 % had no access to credit services. On the other hand, about 56.96% of non-adopters of RURWV had access to credit and the rest 43.04% of households did not have access to credit services. The Pearson chi-square test results revealed that there was a statistically significant difference (pr = 0.000, which is less than 0.01) in access to credit services between RURWV adopters and non-adopters.

Membership of farmers Cooperatives: as shown in table 2 below from the total sample household's 55.3% of households were members of farmer's cooperatives and the rest 44.7% were not members of farmer's cooperatives. 63.69% of RURWV adopters were members of farmers' cooperatives and the rest 36.31% were not members of farmer's cooperatives. On the other hand, about 49.57% of non-adopters of RURWV were members of farmer's cooperatives and the rest 50.43% of households were not members of farmer's cooperatives. The chi-square test result revealed that there was a statistically significant difference (p-value less than 0.01) in membership of farmers cooperatives between adopters of RURWV and non-adopters.

Table 2. Summary of frequency of categorical variables

Variables		RUWV Adopters		RURWV Non-adopters		Total		
		Frequency	Percent	Frequency	Percent	Frequency	Percent	χ^2 (chi2)
Sex_hh	Female	37	23.57	69	30.00	106	27.39	
	Male	120	76.43	161	70.00	281	72.61	1.942
ACRD	Yes	120	76.43	131	56.96	251	64.86	
	No	37	23.57	99	43.04	136	35.14	15.529***
MCOP	Yes	100	63.69	114	49.57	214	55.30	
	No	57	36.31	116	50.43	173	44.70	7.535***

Source: Own computation using survey data (2021).

Note: *, ** and *** represents significant at 10%, 5%, and 1% level of significance respectively.

3.2.2. Descriptive statistics results of continuous variables

Age of household head: as shown in table 3 below, the average age of household head for whole observation was 42.79 years. For adopters of RURWV, the average age of household head was 41.35 and for non-adopters of RURWV, the average age of household head was 43.78 years. The result shows that both adopters and non-adopters on average lied in the economically productive age. The mean difference in the average age between adopters and non-adopters was 2.43 years in absolute terms and it was statistically significant at 5% level of significance. Moreover, the result revealed that those adopters of rust-resistant improved wheat varieties were on average younger than non-adopters, which implies that they are economically active labor than non-adopters and it may enable them to easily grasp new information and not stick to old technologies.

Education level: as shown in table 3 below, the average education enrollment level of household head of all sample observation was 6.89 years of schooling, approximately grade 7. Moreover, the result revealed that the mean education enrollment of adopter household heads was 7.78 schooling years which is proximately graded 8, whereas for those non-adopter rust-resistant improved wheat varieties their average education level was grade 6. The result revealed that sampled households on average attained their primary education (grade 1-8). The mean difference in education level between the group of adopters and non-adopters was 1.51 years of schooling in absolute terms and this mean difference was significant at 1% level of significance. The result shows that on average adopters have a higher level of educational attainment than non-adopters. This implies that as the education level increases household's capacity and skills of gathering more information from different sources increases and this assist them in their decision to adopt rust-resistant improved wheat varieties.

Farm experience: as shown in table 3 below, the average farming experience of the household head for all sample observation was 17.48 years. For adopters of RURWV, the average farming experience of household head was 17.22 and for non-adopters of RURWV, the average farming experience of household head was 17.65 years. The mean difference in average farming experience between adopters and non-adopters was 0.43 years in absolute terms and using t-statistics this mean difference revealed that there was no statistically significant difference in the average farming experience of households between adopters and non-adopters.

Distance to nearest market: as shown in table 3 below, the average walking distance to the nearest market in walking minutes for combined sample observation was 35.29 minutes. Moreover, the result revealed that on average adopters of RURWV household heads took 33.43 walking minutes to their nearest market, whereas for those non-adopters of RURWV on average it took 36.57 minutes. The mean difference in walking minutes between adopters and non-adopters was 3.14 minutes in absolute terms, and it was statistically significant at 5% level of significance. This indicates that those who were non-adopters of rust-resistant improved wheat varieties took much time to reach their nearest market and it may have a negative consequence in the adoption decision of rust-resistant improved wheat varieties.

Family size of household: as shown in table 3 below, on average adopters RURWV and non-adopters RURWV had family sizes 7.42 and 7.10 respectively, which was approximately 7 persons per household for both adopters and non-adopters. But the mean difference in average family size between groups adopters and nonadopters was 0.32 in absolute terms and it was statistically significant at 10% percent level of significance. The result shows that on average adopters have a large family size than non-adopter households. As it is indicated in table 4, these adopters had a low dependency ratio as compared to non-adopters this revealed that they have a more economically active labor force that enables them to gain more income and this in turn has a positive contribution for adoption.

Land size owned: the area of land cultivated measures the availability of land for agricultural production, a household with more landholding has the opportunity to produce more crops and thereby generate more income. From table 3 below, the average cultivated landholding of all sample households was 0.73 hectare of land. On the other hand, the average landholding of RURWV adopters were 0.81 hectare whereas the average landholding of respondents who did not adopt rust-resistant improved wheat varieties was 0.68 hectare. The mean difference between adopters and non-adopters in average land size was 0.13 hectare in absolute terms and it was significant at 1% level of significance. Adopters of rust-resistant improved wheat varieties had more land size than non-adopters. This indicates adopters had more possibility to diversify crop production and adopt rust-resistant improved wheat varieties.

Livestock holding (TLU): as shown in table 3 below, on average the livestock holding of all sample households in tropical livestock units was 6.08 tropical livestock units (TLU). The average livestock holding of adopters of RURWV was 6.74 tropical livestock units whereas the average livestock holding of respondents who did not adopt rust-resistant improved wheat varieties was 5.62 tropical livestock units. The mean difference in livestock holding in terms of tropical livestock holding unit between adopters and non-adopters was 1.12 TLU in absolute terms and it was statistically significant at 1% level of significance. Adopters of RURWV had more livestock holding as compared to nonadopters of RURWV. This implies large livestock holding increases the probability of gaining income from sales of livestock and positively contributes to the adoption decision of households.

Frequency of extension service: as indicated in table 3 below, for all sample households the average number of frequencies farmers get extension services from development agents during cropping season was (3.88) approximately 4 times. The average number of frequencies of extension contact of households with development agents during wheat cropping season for those who adopt RURWV was 4.29 and for non-adopter households, the average number of extension agent visits during wheat cropping season was 3.6 times. The mean difference in extension contact frequencies between adopters and non-adopters was 0.69 times in absolute terms and this mean difference was statistically significant at 1% level of significance. This revealed that adopters of rust-resistant improved wheat varieties had more extension visits than non-adopters. This implies that adopters have better extension services and information on agricultural technologies than non-adopter and this in turn positively contributes to RURWV adoption.

Dependency ratio: as shown in table 3 below, the average dependency ratio of all sample households was 78.99%. On the other hand, the mean dependency ratio of adopters of RURWV was 76.23% whereas the average dependency ratio of respondents who did not adopt RURWV was 80.87%. The mean difference in the dependency ratio between adopters and non-adopters was 4.64% in absolute terms, and this mean difference was not statistically significant. Even though their difference is not statistically significant relatively non-adopters have a high dependency ratio it implies there is a relatively lower economically active labor force, therefore, it may have an indirect negative contribution to adoption.

Table 3. Summary and mean comparison of continuous variables

Variables	RURWV Adopters (n=157)	RURWV Non-adopters (n=230)	Mean difference	Combined sample (n=387)	T- value
Age of household head	41.35	43.78	-2.43	42.79	2.524**
Education level	7.78	6.27	1.51	6.89	6.713***
Farm experience	17.22	17.65	-0.43	17.48	0.496
Distance to market (minutes)	33.43	36.57	-3.14	35.29	2.230**
Family size	7.42	7.1	0.32	7.23	1.668*
Land size (ha)	0.81	0.68	0.13	0.73	5.419***
Livestock holding (TLU)	6.74	5.62	1.12	6.08	4.068***
Frequency of extension service	4.29	3.60	0.69	3.88	5.871***
Dependency ratio	76.23	80.87	-4.64	78.99	0.675

Source: Own computation using survey data (2021).

Note: *, ** and ***, indicate significance at 10% level of significance, 5% level of significance, and 1% level of significance respectively.

3.3. Econometric Models

3.3.1. Diagnostic test of logistic regression model

This study used logistic regression to estimate propensity scores for impact evaluation. For the analysis to be valid, the model has to satisfy the assumptions of logistic regression. Therefore, before using the model to make any statistical inference, the study checked that the logistic regression model used fits sufficiently well using major diagnostic tests of the logistic regression model. The details of model diagnostic tests of the logistic regression model used in the study are presented as follows.

1. Goodness of Fit Measures for Logistic Regression

The goodness-of-fit (GOF) tests can help us to decide whether the model is correctly specified. The goodness of fit is based on hypothesis testing of the following type:

H0: model is exactly correct and *HA*: model is not exactly correct

If the null hypothesis is rejected, the model is not exactly correct or not correctly specified. On the other hand, if we do not reject null hypnosis the model is correct or correctly specified. GOF tests produce p-value helps to test the fitness of the logistic regression model. If it is low (below 0 .05), the model does not fit or reject the null hypothesis which says the model is fit. If it's high (greater than 0.05), then the model passes the test to accept the null hypothesis which says the model is fit (Lomax and Hahs-Vaughn, 2012).

The study used the Hosmer-Lemeshow goodness-of-fit test is one of the tools used to examine the overall model fit. The Hosmer & Lemeshow test provides a global fit test, testing the 'estimated model to one that has a perfect fit. If this test is not significant, then you have evidence of a correctly specified model. If it is significant, then you have evidence that the model is misspecified (Pituch and Stevens, 2016). Table 4 shows that (Hosmer-Lemeshow chi2 (8) = 6.52, Prob > chi2 = 0.5890), prob chi2 is greater than critical value 0.05 which was insignificant, this result revealed that the model had an acceptable fit or correctly specified, in other words, it means that we accept the null hypothesis which indicates the model was correctly specified.

Table 4. Hosmer-Lemeshow chi-square model specification test

Logistic model for RURWV, goodness-of-fit test
(Table collapsed on quantiles of estimated probabilities)
number of observations = 387
number of groups = 10
Hosmer-Lemeshow chi2(8) = 6.52
Prob > chi2 = 0.5890

Source: Own computation using survey data (2021)

2. Multicollinearity test

2.1. Multicollinearity test of continuous explanatory variables

For continuous explanatory variables, multicollinearity was detected with the help of tolerance and its reciprocal, called variance inflation factor (VIF). According to Lomax and Hahs-Vaughn (2012), tolerance is the percentage of the variance in a given predictor that cannot be explained by the other predictors. Tolerance close to 1 indicates that there is no multicollinearity problem, whereas a value close to zero suggests that multicollinearity may be a threat. The VIF is defined as the reciprocal of tolerance. VIF shows that how much the variance of the coefficient estimate is being inflated by multicollinearity. Values of VIF exceeding 10 are often regarded as indicating the existence of multicollinearity. From table 5 below, all continuous explanatory variables had tolerance values closer to one, and variance inflating factors of all explanatory variables were below 2, which indicates VIF all these explanatory variables were less than critical VIF value 10. So, by using the rule of thumb (that is if the VIF of a variable exceeds 10, that variable is said to be highly collinear) there was no multicollinearity or collinearity problem between explanatory variables in the specified logistic regression model.

Table 5. Multicollinearity test for continuous variables

Variable	VIF	Tolerance	R-Squared
Age_hh	1.29	0.7760	0.2240
Educ_level	1.11	0.9043	0.0957
Farm_Exp	1.37	0.7307	0.2693
Mrk_Dist	1.19	0.8377	0.1623
TFAMSIZE	1.05	0.9502	0.0498
LDOW	1.15	0.8726	0.1274
LHTLU	1.19	0.8406	0.1594
FRQEXN	1.16	0.8655	0.1345
DPR	1.14	0.8775	0.1225
Mean VIF	1.18		

Source: Own computation using survey data (2021)

2.2. Multicollinearity test for desecrate variables

This study used contingency coefficient to detect the existence of multicollinearity between desecrate variables. As shown in table 6 below, the contingency coefficients between of explanatory variables where is less than 0.75. So, using this rule of thumb method of detecting multicollinearity, there is no multicollinearity problem between these desecrate variables.

Table 6. Contingency coefficient for discrete variables

	Sex_hh	ACRD	MCOP
Sex_hh	0.707		
ACRD	0.112	0.707	
MCOP	0.039	0.121	0.707

Source: Own computation using survey data (2021)

3.3.2.The impact of adopting rust-resistant wheat varieties

3.3.2.1. Estimating propensity scores

Estimating the propensity score is the first and crucial step in using propensity score matching as an evaluation strategy to predict the probability of adoption of rust-resistant improved wheat varieties. The model used for propensity score estimation was the logistic regression model and this model consists of a range of predictor variables that are most likely to influence both adoption of rust-resistant improved wheat varieties and the outcome variables. In propensity score estimation after estimation of propensity score, first, there is an identification of an optimal number of blocks in which mean propensity for treated and controls in each block close to each other. Then after there should be a balance between the mean propensity score of treated and untreated, if the balancing property is not satisfied, the corrective measure should be taken. For this study, the number of blocks were five and the balancing property was satisfied.

The propensity score estimation in table 7, revealed that, likelihood ratio chi-square test with 12 degrees of freedom (LR chi2 (12) = 121.14, prob > chi2 = 0.0000), this implies that the null hypothesis which indicates all coefficients are simultaneously equal to zero is rejected at 1% level of significance (prob > chi2 = 0.0000; which is has p-value < 0.01). In other words, it means coefficients of explanatory variables were different from zero. The pseudo-R^2 value was 0.2318 which is low. This indicates that there were no symmetric differences in the distribution of covariates between adopters and non-adopters of RURWV, it indicates the adoption of rust-resistant improved wheat varieties was fairly random. Age of household head, education level of household head, land holding, livestock holding, frequency of extension contact and access to credit were factors significantly affected the adoption of rust-resistant improved wheat varieties.

Table 7. Logistic regression result of propensity scores

RURWV	Coef.	Std. Err.	Z	P>z
Age_hh	-0.054	0.017	-3.24***	0.001
Sex_hh	0.221	0.296	0.75	0.455
Educ_level	0.364	0.069	5.30***	0.000
Farm_Exp	0.015	0.017	0.89	0.373
Mrk_Dist	-0.011	0.010	-1.10	0.272
TFAMSIZE	0.100	0.070	1.44	0.151
LDOW	1.878	0.581	3.24***	0.001
LHTLU	0.098	0.052	1.87*	0.061
FRQEXN	0.368	0.124	2.97***	0.003
ACRD	0.825	0.282	2.93***	0.003
MCOP	0.201	0.262	0.77	0.443
DPR	-0.003	0.002	-1.28	0.200
_cons	-5.361	1.140	-4.70	0.000

Logistic regression	Number of obs	=	387
	LR chi2(12)	=	121.14***
	Prob > chi2	=	0.0000
Log likelihood = -200.88985	Pseudo R2	=	0.2318

Source: own computation using survey data (2021)

Figure 1 describes the distribution of the household with respect to the estimated propensity scores. In case of adopter households, most of them are found in partly the middle and partly in the right side of the distribution.

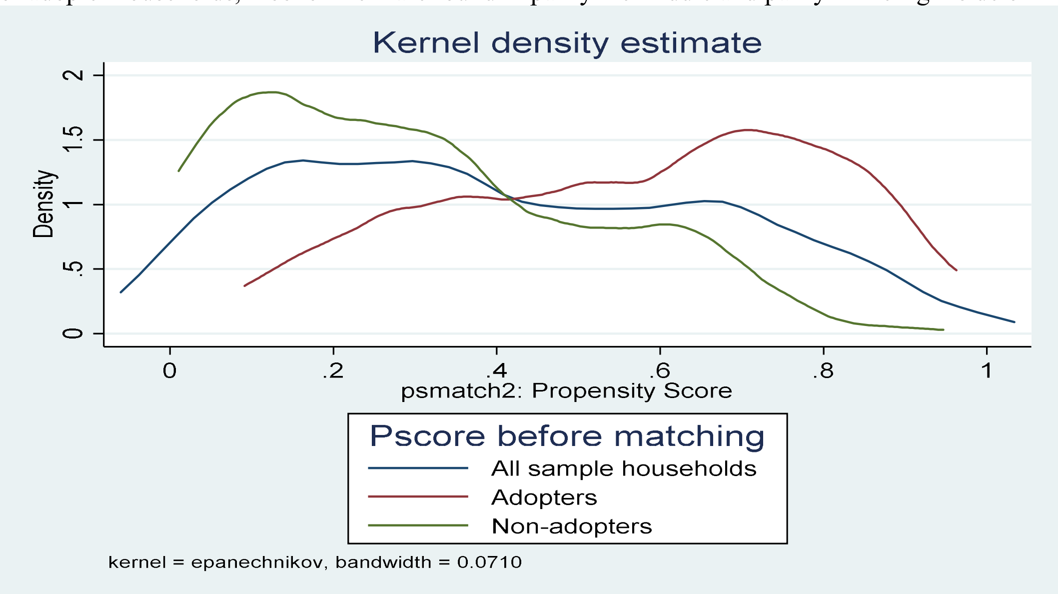


Figure 1. Kernel density of propensity scores of distributions

3.3.2.2. Overlap and Common Support

After estimation of propensity scores the next crucial step is to ensure the existence of a region of common support between adopters and non-adopters of rust-resistant improved wheat variety (RURWV). Sufficient overlap in propensity scores for treated and control is required to ensure the estimation of treatment effect is not biased. Based on the results shown in table 8 below, the predicted propensity score for RURWV adopters had a minimum of 0.0913378 and a maximum of 0.9626534 with a mean propensity score of 0.5699304 which is approximately 0.6. On the other hand, the predicted propensity score for non-adopters of RURWV had a minimum of 0.0106588 and a maximum of 0.9466604 with a mean propensity score of 0.2935692 which is approximately 0.3.

Thus, using the method of comparing the minima and maxima of the propensity score in both groups (adopters and non-adopters), the minimum propensity score of adopters was 0.0913378 and the maximum propensity score of non-adopters was 0.9466604, therefore the common support region is [0.0913378 to 0.9466604], this means households with propensity score less than the minimum (0.0913378) and larger than maximum (0.9466604) are off-support and not considered for matching and estimation of average treatment effect. Based on this, a total of 53 households of which 50 were from non-adopter households and 3 from adopters' households which accounts for 13.7% of households of total sample households were out of the support region. This also implies the study has enough support region and satisfies the requirement of sufficient overlap and support.

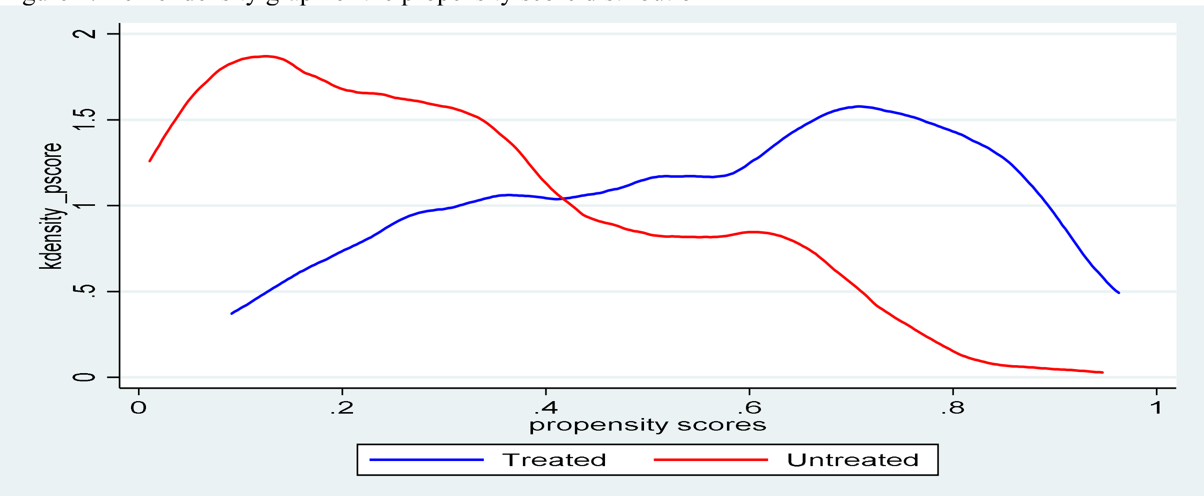
Table 8. Summary of estimated propensity score of households

Propensity score	Mean	Std. Dev.	Minimum	Maximum
RURWV Adopters	0.5699304	0.235332	0.0913378	0.9626534
RURWV non-adopters	0.2935692	0.2120474	0.0106588	0.9466604
Full sample	0.4056848	0.2598457	0.0106588	0.9626534

Source: Own computation using survey data (2021)

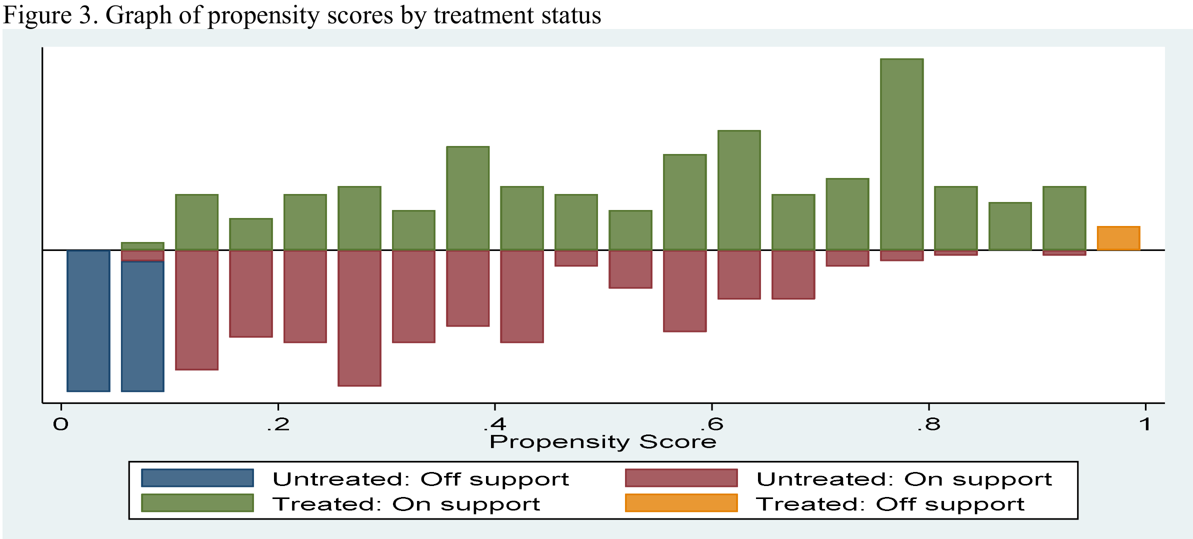
The common support region was also shown clearly in the following figure 3 below. The bottom half of the graph shows the propensity score distribution of non-adopters of RURWV and the upper half of the graph represents the propensity score distribution of adopters of RURWV. The density of the scores is shown in figure 2 below. The densities of the propensity scores are on the y-axis and propensity scores on the x-axis. As it is shown in the kernel density graph, the density distribution of estimated propensity scores for adopters(treated) and non-adopters of RURWV (untreated), the assumption of common support condition was satisfied, and there exists enough overlap in the distribution of propensity scores of the two groups.

Figure 2. Kernel density graph of the propensity score distribution



Source: Own computation using survey data (2021)

Figure 3. Graph of propensity scores by treatment status



Source: Own computation using survey data (2021)

Note: Treated indicates adopters of RURWV and Untreated indicates non-adopters of RURWV.

Treated on support indicates these households in the group of adopters of RUWV who found a suitable match, whereas treated off support indicates that households in the group of adopters of RURWV who did not find a match from non-adopters of RURWV.

3.3.2.3. Choosing Matching Algorithm

Another important step is choosing the best matching algorithm. To choose a matching algorithm Sianesi (2004); Dehejia and Wahba (2002) proposed criteria such as: large matched sample size, low pseudo-R^2, a large number of insignificant variables after matching (covariance balance test), and joint insignificant of all regressors of logit or probit analysis after matching (in this case logistic regression analysis).

Table 9, shows results obtained from major matching algorithms namely nearest neighbor matching, caliper matching, radius matching, and kernel matching. Based on the above-mentioned matching algorithm selection criteria best fit matching estimator who had large matched sample size, low pseudo-R^2, a large number of insignificant explanatory variables after matching (insignificant t-test of explanatory variables after matching and mean bias less than 20%), was radius matching with radius caliper (0.1); therefore, for this study, this matching estimator used to estimate the average treatment effect.

Table 9. Performance of matching estimators

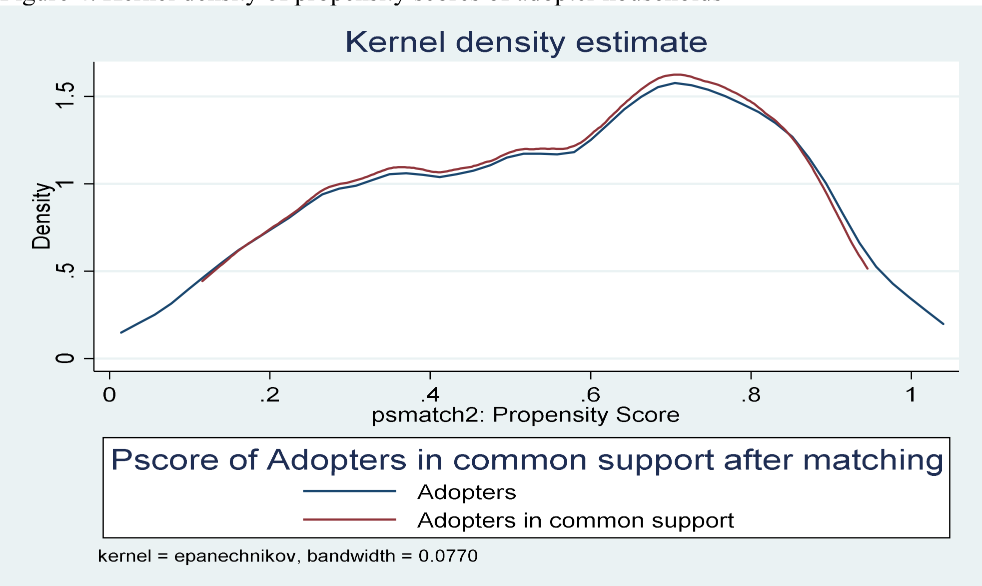
Matching estimators	Performance Criteria		
	Pseudo-R^2	Balancing test*	Matched sample size
Nearest neighbor matching			
Nearest neighbor 1	0.022	11	334
Nearest neighbor 2	0.021	11	334
Nearest neighbor 3	0.018	12	334
Nearest neighbor 4	0.016	12	334
Nearest neighbor 5	0.015	12	334
Caliper matching			
Caliper 0.01	0.013	12	293
Caliper 0.1	0.022	11	334
Caliper 0.25	0.022	11	334
Caliper 0.5	0.022	11	334
Radius matching			
Radius 0.01	0.018	12	293
Radius 0.1	**0.013**	**12**	**334**
Radius 0.25	0.028	11	334
Radius 0.5	0.087	7	334
Kernel Matching			
Bandwidth 0.01	0.019	12	293
Bandwidth 0.1	0.015	12	334
Bandwidth 0.25	0.018	12	334
Bandwidth 0.5	0.056	11	334

Source: Own computation using survey data (2021)

Note: radius matching estimator with radius caliper 0.1 has low pseudo-R^2 value =0.013, large balanced covariates =12, and large amount of matched sample size =334.

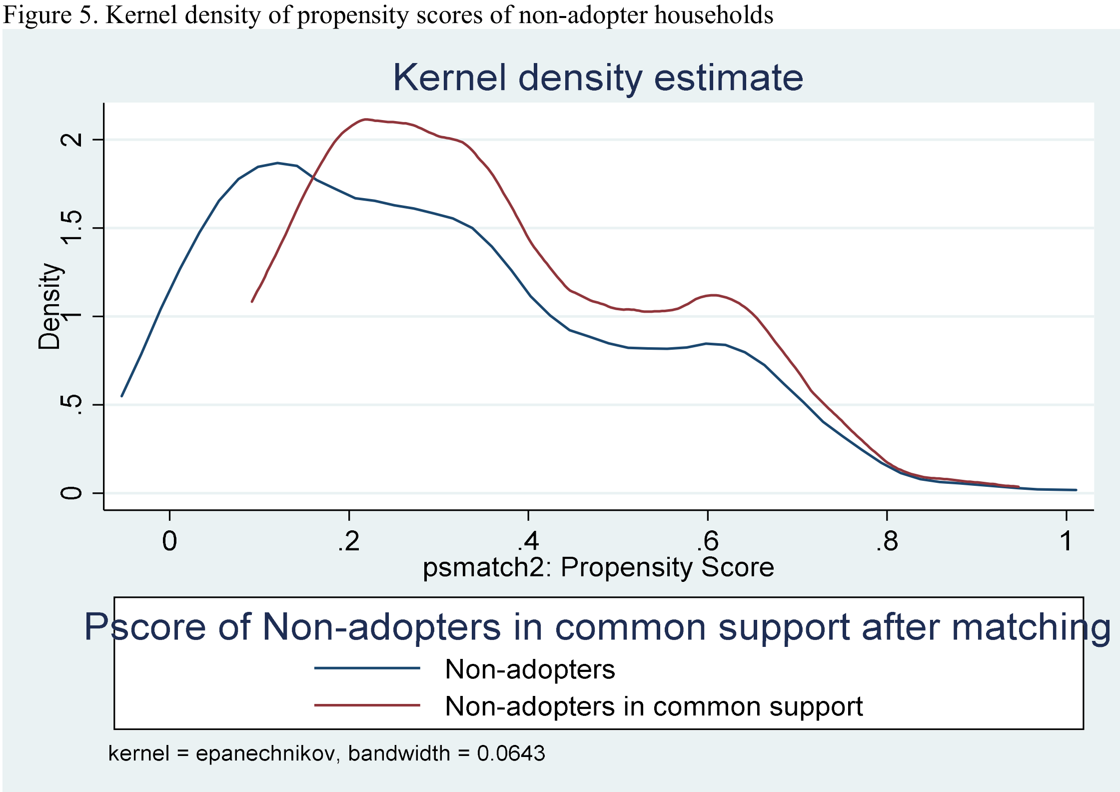
The quality of matching can also be assessed by visual inspection using graphs. To do so, we plotted the graphs of estimated propensity scores for adopter and nonadopter households after matching (Figure 4 and 5). Obviously, the distributions of the estimated propensity scores were somehow skewed to the right for adopter households and to the left for non-adopter households. However, the region of common support was sizable and the distribution of the graph appeared even more similar after matching.

Figure 4. Kernel density of propensity scores of adopter households



Source: Own computation using survey data (2021)

Figure 5. Kernel density of propensity scores of non-adopter households



Source: Own computation using survey data (2021)

3.3.2.4. Testing the balance of propensity scores and covariates

In addition to propensity score, the matching procedure has to be checked if it is able to balance the distribution of relevant variables in both the control and treatment groups. The basic idea of this approach is to compare the situation before and after matching and check if there remain any differences after conditioning on the propensity score. If there are differences, matching on the score was not completely successful, and remedial measures have to be done (Caliendo and Kopeinig, 2008). According to Khandker *et al*. (2010) balancing tests can also be conducted to check whether, the average propensity score and mean of covariates between treated and control groups are the same.

The balancing powers of the estimations are ascertained by considering different testing methods such as the reduction in the mean standardized bias between the matched and unmatched households, equality of means using t-test, values of Pseudo-R^2, and chi-square test for joint significance of the variables used. Table 10 presents propensity score and covariance balance test before and after matching by using the selected matching algorithm (in this case radius matching with radius caliper 0.1).

The standardized bias before matching and after matching, total bias reductions obtained by the matching procedure as shown in columns five and six in table 10, the standardized difference in propensity score and covariates before matching was in between of 5.1% and 123.4% in absolute value and 9 variables had absolute mean bias greater than 20%, which shows RURWV adopters and non-adopters were not balanced in these 9 variables including mean propensity scores. Whereas standardized difference of variables for all covariates and propensity score lies between 0.2 % and 14.4% after matching. This is fairly below the critical level of 20% suggested by Rosenbaum and Rubin (1985). Therefore, the process of matching creates a high degree of covariate and propensity score balance between the adopter and non-adopter samples that are ready to use in the estimation procedure.

Similarly, the t-test in table 10 revealed that before matching 8 variables and the propensity score in total 9 variables was statistically significant at 5% level of significance, which means households who were adopters of RURWV and non-adopters of RURWV had significant differences in means of these variables. But after matching all variables (12 variables) and propensity score were insignificant t-value ($p > 0.05$), which means mean differences of all these variables between RURWV adopters and non-adopters have no statistical significance. This implies the covariates were balanced between the two groups. Based on the above evidence, the overall variables were balanced and the assumption of no selection bias was satisfied.

Table 10. Summary of propensity score and covariate balance test

		Mean			%reduction	t-test	
Variables	Sample	Treated	Control	%bias	|bias|	T	p>|t|
_pscore	Unmatched	0.56993	0.29357	123.4		12.04	0.000
	Matched	0.56235	0.54655	7.1	94.3	0.61	0.542
Age_hh	Unmatched	41.35	43.778	-26.2		-2.52	0.012
	Matched	41.519	41.84	-3.5	86.8	-0.34	0.732
Sex_hh	Unmatched	0.76433	0.7	14.5		1.39	0.164
	Matched	0.75974	0.82356	-14.4	0.8	-1.38	0.169
Educ_level	Unmatched	7.7834	6.2739	70.3		6.71	0.000
	Matched	7.7273	7.6774	2.3	96.7	0.22	0.830
Farm_Exp	Unmatched	17.223	17.652	-5.1		-0.50	0.620
	Matched	17.357	17.312	0.5	89.5	0.05	0.962
Mrk_Dist	Unmatched	33.433	36.565	-23.2		-2.23	0.026
	Matched	33.5	35.424	-14.2	38.6	-1.23	0.220
TFAMSIZE	Unmatched	7.4204	7.1	17.4		1.67	0.096
	Matched	7.4221	7.2946	6.9	60.2	0.63	0.531
LDOW	Unmatched	0.80732	0.67609	55.3		5.42	0.000
	Matched	0.80195	0.80232	-0.2	99.7	-0.01	0.988
LHTLU	Unmatched	6.7377	5.6237	41.1		4.07	0.000
	Matched	6.6816	6.5123	6.2	84.8	0.59	0.553
FRQEXN	Unmatched	4.2866	3.6043	63.6		5.87	0.000
	Matched	4.2727	4.257	1.5	97.7	0.15	0.884
ACRD	Unmatched	0.76433	0.56957	42.1		4.01	0.000
	Matched	0.75974	0.76046	-0.2	99.6	-0.01	0.988
MCOP	Unmatched	0.63694	0.49565	28.7		2.77	0.006
	Matched	0.63636	0.58828	9.8	66.0	0.86	0.388
DPR	Unmatched	76.226	80.869	-7.1		-0.68	0.500
	Matched	76.575	70.09	9.9	-39.7	0.92	0.357

Source: Own computation using survey data (2021)

In addition, as it is presented in table 11 below, before matching, the standardized mean bias for overall covariates used in propensity score estimation was 39.8%; after matching this standardized mean bias reduced to 5.9% which is below the critical value suggested by Resenbaum and Rubin (1985).

According to Sianesi (2004), after matching there should be no systematic differences in the distribution of covariates between both groups and therefore the pseudo-R^2 should be fairly low. Also suggested that the likelihood ratio test on joint significance of all regressors in the probit or logit model should not be rejected before matching, and should be rejected after matching. Based on this mentioned criteria, table 11 below shows that pseudo-R^2 was 0.234 and significant p-value of likelihood ratio test which was 0.000 before matching, but pseudo-R2 was 0.014 is fairly low and insignificant p-value of likelihood ratio test which was 0.949 after matching and low standard bias revealed that the selected matching estimator in this case radius matching with radius caliper 0.1, was successfully balanced the distribution of covariates between groups of adopters and non-adopters of rust-resistant improved varieties.

Table 11. Matching quality indicators

Sample	Ps R2	LR chi2	p>chi2	Mean Bias
Unmatched	0.234	122.30	0.000	39.8
Matched	0.014	5.93	0.949	5.9

Source: Own computation using survey data (2021)

3.3.2.5. Estimating the average treatment effect

This section presents evidence of whether the adoption of rust-resistant improved wheat varieties has an impact on the household's productivity and income. The average yield of wheat production per hectare was used as an indicator of farmers' productivity for both adopters and non-adopters of rust-resistant improved wheat varieties. Annual total income from wheat production in terms of Ethiopian birr was used as an indicator of income for households.

Table 12 below, shows that the mean impact of adopting RURWV on households' wheat productivity in quintal per hectare and households' income in Ethiopian birr. As shown in table, the average treatment effect on treated (ATT) revealed that the average wheat productivity of adopters of RURWV increases by 16.62 quintals per hectare (1.662 ton/ha) as compared to non-adopters of RURWV and the t-test result shows that, this impact was statistically significant at 1% level of significance. The result agrees with the study by Tesfaye *et al*. (2016) reported that improved wheat verities had a positive and significant effect on the productivity of farmers. Similarly, Tesfaye *et al*. (2018) found that improved wheat variety adoption significantly increased wheat productivity, and Adane *et al*. (2019) also found that adoption of improved soybean varieties had a significant and positive effect on productivity. This implies the adoption of rust-resistant wheat technologies has a significant and positive contribution to the productivity of households.

The result from table 12, indicates that on average income from wheat production for adopters of RURWV increased by Ethiopian birr 10,460.63 as compared to being non-adopters of RURWV and this mean impact was statistically significant at 1% level of significance. This result agrees with studies by Khonje *et al.* (2015) reported that the adoption of improved wheat leads to a significant and positive effect on crop income; Tesfaye *et al*. (2016); Regasa and Degye (2019) found adoption of improved wheat varieties had a significant and positive impact on the income of households. This result implies that the adoption of improved wheat varieties has a significant role in increasing productivity and in turn increases income gained from wheat production.

Table 12. Average treatment effect on treated (ATT) on productivity and Income

Variables	Treated	Controls	Difference	S.E.	T-stat
Wheat yield (qt/ha)	34.669	18.051	16.618	0.931	17.85***
Income (in birr)	29,785.71	19,325.08	10,460.63	2018.38	5.18***

Source: Own computation using survey data (2021)

Note: *** represents significant at 1% level of significance.

3.3.2.6. Sensitivity analysis of average treatment effects

If there are unobserved variables that simultaneously affect assignment into treatment and the outcome variable hidden bias might arise to which matching estimators are not robust (Rosenbaum, 2002). Sensitivity analysis of average treatment effect is to check whether or not inference about treatment effects may be altered by unobserved factors. For this study, the bounding approach sensitivity analysis proposed by Rosenbaum (2002) was used to assess the sensitivity of the average treatment effects for unobserved covariates. The rbounds calculates Rosenbaum bounds for average treatment effects on the treated in the presence of unobserved heterogeneity (hidden bias) between treatment and control cases. The procedure then calculates Wilcoxon sign rank tests that give upper and lower bound estimates of significance levels at given levels of hidden bias (DiPrete and Gangl, 2004). Given the positive average treatment effect on productivity and income of households, the lower bounds under the assumption that we have under-estimated the true treatment effect are somewhat less interesting (Becker and Caliendo, 2007). Thus, to test the sensitivity of the average treatment effect of these outcome variables are using upper bound Wilcoxon positive significance level (sig+) at different critical value gamma (***e^γ*)**.

As shown in table 13 below, sensitivity analysis of outcome variables namely: wheat yield in quintal per hectare and income of households. Gamma (***e^γ***) measures a degree of departure from the study that is free from hidden bias. The critical level of '***e^γ*** 'in the table represents the levels at which causal inference of the significant impact of adopting RURWV is questionable. The results in columns 2 to 3 show that, upper bounds of Wilcoxon significance level (Sig+) at a different level of gamma (*e^γ*). The result revealed that, if the matched pairs of adopters and non-adopters of RURWV allowed to differ in odds adopting of RURWV by a factor of 2.5 (150%) in unobserved characteristics, the impact of adopting RURWV on wheat productivity and income of households is still significant at 10% level of significance. This indicates the estimated impacts are insensitive to hidden bias. According to Hui *et al*. (2020), if the results of sensitivity analysis are significant until the value of gamma (Γ) is close to 2, it can be considered that the empirical result of propensity score matching is insensitive to unobserved bias. Based on this, the results have a good justification to insure the average treatment effects on treated using propensity score matching were insensitive to unobserved bias. This revealed that the study includes important covariates that affected both adoptions of RURWV and outcome variables and the estimated average treatment effect on treated (ATT) for both wheat productivity and income of households from wheat production were insensitive to hidden biases or unobserved covariates.

Table 13. Rosenbaum bounds sensitivity analysis of hidden bias

	Scores on productivity	**Scores on income**
Gamma(e^{γ})	**sig+**	**sig+**
1	0	1.6e-09
1.1	0	3.1e-08
1.2	0	3.5e-07
1.3	0	2.7e-06
1.4	0	0.000015
1.5	0	0.000061
1.6	0	0.00021
1.7	2.2e-16	0.000605
1.8	1.3e-15	0.001515
1.9	7.1e-15	0.003366
2	3.3e-14	0.00676
2.1	1.3e-13	0.012449
2.2	4.7e-13	0.021281
2.3	1.5e-12	0.034108
2.4	4.3e-12	0.051687
2.5	1.1e-11	0.074581
2.6	2.8e-11	0.103087

Source: own computation using survey data (2021)
Gamma(e^{γ}) = log odds of differential assignment due to unobserved factors
sig+ = upper bound significance level, and sig- = lower bound significance level

CHAPTER FOUR

4. SUMMARY, CONCLUSION, AND RECOMMENDATION

4.1. SUMMARY AND CONCLUSION

This study analyzed the impacts of adopting rust resistant improved wheat technology on household productivity and income in Misha district, Hadiya zone, Southern nation nationalities and people's region Ethiopia. The objectives of this study were to analyze the impact these rust-resistant improved wheat technologies or varieties on the productivity and income of households in Misha district. Descriptive statistics, propensity score matching method, and cross-sectional survey data were used to achieve the objectives of the study. Data on farm households' demographic, socio-economic characteristics, and institutional factors were collected from randomly selected 387 sample households.

Adopting rust-resistant improved wheat technology is one way of improving farmers' wheat production and decreasing yield loss due to currently occurring wheat rust diseases. For this study analysis of factors influencing adoption of rust-resistant improved wheat varieties were conducted using a binary logistic regression model. The analysis of data shows that education level of household head, land size, livestock holding in tropical livestock unit, frequency of extension contacts, and access to credit services were factors positively and significantly affected the probability of adopting rust-resistant improved wheat varieties. On the other hand, age of household head was negatively and significantly affected the adoption of rust-resistant improved wheat varieties.

The estimation of the average treatment effect of adoption of rust-resistant improved wheat varieties was done using radius matching with radius caliper of 0.1, which satisfies the criteria of best matching estimator such as: low pseudo R^2(0.13), satisfying balancing test, and a large amount of matched sample size. Using the method of minima and maxima comparison of propensity scores the common support region was [0.0913378 to 0.9466604]. Thus, 3 households from adopters and 50 households from non-adopters in total 53 (13.7%) households became out of support.

The findings on the impact of adopting rust-resistant improved wheat varieties revealed that adoption of rust-resistant improved wheat varieties had a positive and significant effect on wheat productivity and income of households. On average the productivity of adopters of RURWV increased by 16.62 quintal per hectare (1.662 ton/ha) wheat as compared to matched non-adopters and it was statically significant at 1% level of significance. In addition to this, the income of adopters of RURWV increased birr 10,460.63 as compared to those non-adopters and this result was statistically significant at 1% level of significance.

The study concludes wheat rust is one of the major limiting factors for wheat production that has been affecting wheat production and productivity. Therefore, adopting rust-resistant improved wheat varieties have a crucial role for farmers to increase productivity and income gained from wheat production, this, in turn, improves the livelihood of farmers engaged in wheat production. This also implies it will contribute to the reduction of poverty in rural areas.

4.2. RECOMMENDATION

Based on the findings of this research the following recommendations are forwarded:

Adoption of rust-resistant improved wheat varieties was found to have a positive and significant effect on wheat productivity and income of households. As a result, further scaling up of rust-resistant improved wheat technologies through awareness creation and training for farmers by extension agents and concerned other government officials will help to sustainably enhance productivity and improve the wellbeing of households. In addition to this, using focus discussion the study found that there was a shortage of supply or access to rust-resistant improved wheat seeds and other farm inputs like chemicals in the study area. Therefore, this study recommends the government should encourage and support seed suppliers to supply adequate improved wheat seed in a way that will be accessible for all farmers thereby these farm households will be benefited from using these rust-resistant improved wheat seeds.

Wheat rust was a critical problem in Misha district and Ethiopia as a whole, therefore agricultural research and extension activities need to give due attention to the development and dissemination of rust-resistant improved wheat varieties. Generation of rust disease-resistant wheat varieties from time to time, promotion and scaling-up of those rust-resistant improved wheat needs due considerations and concern to increase productivity and improve the livelihood of farm households.

REFERENCES

Adane Tufa Hirpa, Arega Alene, Manda, J., Akinwale, M.G., Chikoye,D., Shiferaw Feleke, Tesfamicheal Wossen, and Manyong, V. (2019). The productivity and income effects of adoption of improved soybean varieties and agronomic practices in Malawi. *World development* 124:104631.

Adeoti, A. I. (2008). Factors influencing irrigation technology adoption and its impact on household poverty in Ghana. *Journal of Agriculture and Rural Development in the Tropics and Subtropics (JARTS)*, *109*(1): 51-63.

Aklilu Nigussie, Adem Kedir, Abiy Adisu, Getnet Belay, Desta Gebrie and Kidane Desalegn. (2015). Bread wheat production in small scale irrigation users agro-pastoral households in Ethiopia: Case of Afar and Oromia regional state. *Journal of Development and Agricultural Economics*, 7 (4): 123-130.

Alemayehu Ayele, Yilma Dessalegn, and Selamawit Tesfaye. (2020). Status of wheat rust diseases in Hadiya Zone, Ethiopia. *Journal of Biology, Agriculture and Healthcare,*10(5): 33-37.

Baye Belay, Fitsum Daniel and Eyob Bezabeh. (2019). Impact of Improved Wheat Technology Package Adoption on Productivity in Ethiopia. *Greener Journal of Agricultural Sciences*, 9(1): 76–85.

Becker, S. O., and Caliendo, M. (2007). Sensitivity analysis for average treatment effects. *The Stata Journal,* 7(1): 71-83.

Becker, S. O., and Ichino, A. (2002). Estimation of average treatment effects based on propensity scores. *The stata journal*, 2(4):358-377.

Bekele Shiferaw, Menale Kassie, Moti Jaleta and Chilot Yirga. (2014). Adoption of improved wheat varieties and impacts on household food security in Ethiopia. *Food Policy*, 44:272-284.

Belayneh Admassu, Friedt, W., and Ordon, F. (2012). Stem rust seedling resistance genes in Ethiopian wheat cultivars and breeding lines. *African Crop Science Journal*, 20(3):149-162.

Blundell, R., Dearden, L., and Sianesi, B. (2005). Evaluating the impact of education on earnings in the UK: models, methods and results from the NCDS. *Journal of the Royal Statistical Society, Series A*, 168(3): 473-512.

Caliendo, M., and Kopeinig, S. (2008). Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys*, 22 (1):31-72.

CSA (Central Statistics Agency). (2018). Agricultural Sample Survey 2017/2018 (2010 E.C.). Report on Area and Production of Major Crops (Private Peasant Holdings, Meher Season). Volume I, Addis Ababa, Ethiopia.

Degefu Kebede, Mengistu Ketema, Nigussie Dechassa, and Feyisa Hundessa. (2017). Determinants of adoption of wheat production technology package by smallholder farmers: Evidences from eastern Ethiopia. *Turkish Journal of Agriculture-Food Science and Technology*, *5*(3): 267-274.

Dehejia, R. and Wahba, S. (1999). Causal effects in nonexperimental studies: reevaluating the evaluation of training programs. *Journal of the American Statistical Association,* 94(448):1053-1062.

Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies.

Review of Economics and statistics, 84(1):151-161.
DiPrete, T. A., and Gangl, M. (2004). Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments. *Sociological methodology*, 34(1): 271-310.
FAO. (2018). The state of food security and nutrition in the world. (Building climate resilience for food security and nutrition) (pp.202). FAO. License: CC BY-NC-SA 3.0 IGO. ISBN 978-92-5-130571-3. https://creativecommons.org/licenses/by-nc-sa/3.0/igo.
Fetch, T., & McCallum, B. (2014). Increased virulence of wheat rusts and the threat to global crop production. In *Future challenges in crop protection against fungal pathogens* (pp. 249-266). Springer, New York, NY.
Fisseha Zegeye, Bamlaku Alamirew, Degefa Tolossa. (2020). Analysis of Wheat Yield Gap and Variability in Ethiopia. International Journal of Agricultural Economics, 5(4): 89-98.
Fitsum Daniel. (2018). Impact of improved wheat varieties adoption on productivity: Disparity among Agro-ecological Zones of Ethiopia. *Greener Journal of Agricultural Sciences*, 8(9): 179-196.
FOSTAT. (2019). Crops and livestock products. available at: https://www.fao.org/faostat/en/#home, accessed on November, 2021.
Getnet Muche Abebele, Merkuz Abera Admasu and Bekele Hundie Agdu. (2020). Field Evaluation of Bread Wheat (Triticum aestivum L.) Genotypes for Stripe Rust (Puccinia striiformis W.) Resistance in Arsi Highlands of Oromia Region, South-Eastern-Ethiopia. *Journal Plant Pathol Microbiol*, 11 (521):1-9.
Girma Woldemichael, Meseret Chimdessa and Anteneh Abebe. (2019). Species Diversity and Use of Home gardens in Misha Woreda, Hadiya Zone of the Southern Nations, Nationalities and Peoples Regional State, Ethiopia. *International Journal of Food Science and Agriculture,* 2(7):118-129.
Gujarati, D.N. (2004). ***Basic econometrics*** **(**4th ed.). Tata McGraw-Hill publishing company: New Delhi, India.
Heckman, J., Ichimura, H., and Todd, P. (1997a). Matching as an econometric evaluation estimator: evidence from evaluating a job training programme. *Review of Economic Studies,* 64(4): 605-654.
Hiwot Hailu. (2018). Impacts of adopting improved wheat varieties on food security in Girar Jarso woreda, north Shewa zone, Oromia region, Ethiopia. Master's thesis. Addis Ababa University, Addis Ababa, Ethiopia.
Hodson, D. P., Moti Jaleta, Kindie Tesfaye, Chilot Yirga, Habtamu Beyene, Kilian, A., Carling, J.,Tesfaye Disasa, Sisay Kidane Alemu, Teshale Daba and Yared Alemayehu. (2020). Ethiopia's transforming wheat landscape: tracking variety use through DNA fingerprinting. *Scientific reports*, 10(1): 1-13.
Hosmer, D. W., and Lemeshow, (2000). *Applied logistic regression* (2nd ed.). John Wiley & Sons Inc., New York.
Hui, J., & Lili, Li, Xinyi, Qian and Yiwu, Z. (2020). Can rural e-commerce service centers improve farmers' subject well-being? A new practice of 'internet plus rural public services' from China. *International Food and Agribusiness Management Review*, 23(1030-2021-195): 681-695.
Khandker, S.R., Koolwal, G.B., and Samad, H.A. (2010). *Handbook on impact evaluation: quantitative methods and practices.*World Bank Publications: Washington DC.
Khonje, M., Manda, J., Arega Alene., and Minale Kassie. (2015). Analysis of adoption and impacts of improved maize varieties in eastern Zambia. *World development*, 66:695-706.
Kidane, Geberemariyam. (2001). Factors Affecting adoption of new wheat varieties in Tigray, Ethiopia: the case of Hawzien district. MSc thesis. Haramaya University, Haramaya.
Leake Gebresilassie., and Adam Bekele. (2015). Factors determining allocation of land for improved wheat variety by smallholder farmers of northern Ethiopia. *Journal of Development and Agricultural Economics*, *7*(3):105-112.
Lidwell-Durnin, J., and Lapthorn, A. (2020). The threat to global food security from wheat rust: Ethical and historical issues in fighting crop diseases and preserving genetic diversity. *Global Food Security*, 26:100446.
Lomax, R.G., and Hahs-Vaughn, D.L. (2012). *An introduction to statistical concepts* (3rd ed.). Routledge Taylor & Francis Group: New York.
Mason, N.M., Jayne, T.S. and Bekele Shiferaw. (2015). Africa's rising demand for wheat: trends, drivers, and policy implications. *Development Policy Review*, 33(5): 581-613.
Mengistu Ketema and Belay Kassa. (2016). Impact of Technology on Smallholder Wheat Production in Bale Highlands of Ethiopia: Application of Output Decomposition Model. *Turkish Journal of Agriculture-Food Science and Technology*, 4(6), 446-454.
Mesay Yami, Tesafye Solomon, Bedada Begna, Fekadu Fufa, Tolesa Alemu and Dawit Alemu. (2013). Source of technical inefficiency of smallholder wheat farmers in selected waterlogged areas of Ethiopia: a translog production function approach. *African Journal of Agricultural Research,* 8(29): 3930-3940.
Milkias Dawiet. (2020). Analysis on Determinants of Adoption of Improved Wheat Technology in Liben Jewi District, Oromia Region, Ethiopia. *International Journal of Applied Agricultural Sciences*, 6(3): 36-43.
Pituch, K.A., and Stevens, J.P. (2016). *Applied multivariate statistics for the social sciences*. Routledge: New York.

Regassa Dibaba, and Degye Goshu. (2019). Determinants of High Yielding Wheat Varieties Adoption by Small-Holder Farmers in Ethiopia. *Journal of Natural Sciences Research,* 9(12):14-23.

Rosenbaum, P. and Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika,* 70(1): 41-55.

Rosenbaum, P. and Rubin, D. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician* 39(1): 33-38.

Rosenbaum, P.R. (2002). *Observational Studies*. Springer: New York.

Shigute Etalema and Anja Abera. (2018). Small ruminant production and constraints in Misha Woreda, Hadiya Zone, Southern Ethiopia. *International Journal of Livestock Production*, 9(8):192-197.

Sianesi, B. (2004). An evaluation of the Swedish system of active labor market programmes in the 1990s. *Review of Economics and Statistics,* 86(1):133-155.

Smith, J. A., and Todd, P. E. (2005). Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of econometrics*, 125(1-2):305-353.

Takahashi, K., and Barrett, C. B. (2013). The system of rice intensification and its impacts on household income and child schooling: Evidence from rural Indonesia. *American Journal of Agricultural Economics*, 96(1): 269-289.

Tesfaye Haregewoin, Baye Belay, Eyob Bezabeh, Kaleb Kelemu, Daniel Hailu and Fitsum Daniel. (2018). Impact of Improved Wheat Variety on Productivity in Oromia Regional State, Ethiopia. *Greener Journal of Agricultural Sciences*, 8(4): 074–081.

Tesfaye Solomon, Bedada Begna and Mesay Yami. (2016). Impact of improved wheat technology adoption on productivity and income in Ethiopia. *African Crop Science Journal*, 24(1):127-135.

Tesfaye Zegeye, Girma Taye, Tanner, D., Verkuijl, H., Aklilu Agidie and Mwangi,W. (2001). Adoption of improved bread wheat varieties and inorganic fertilizer by small-scale farmers in Yelmana Densa, and Farta districts of Northwestern Ethiopia, CIMMYT.

Thabet, M., and Najeeb, K. M. (2017). Impact of wheat leaf rust severity on grain yield losses in relation to host resistance for some Egyptian wheat cultivars. Middle East J, 6(4):1501-1509.

Tsegaye Mulugeta, and Bekele Hundie. (2012). Impacts of adoption of improved wheat technologies on households' food consumption in southeastern Ethiopia (No. 1007-2016-79620).

USDA (United States Department of Agriculture). (2018). World Agricultural Production. Circular Series Report.

Yamane, T. (1967). *Statistics: an introductory analysis* (2nd ed.). Harper and Row: New York.

Zewdu Ayalew, Moti Jaleta and Qaim, M. (2017). Yield effects of rust-resistant wheat varieties in Ethiopia. *Food security*, 9(6): 1343-1357.

Zhao, Z. (2003). Data issues of using matching methods to estimate treatment effects: an illustration with NSW data set. *China Center for Economic Research*, 1-37.