sustainability

MDPI

Article

Fertilizer Use, Value, and Knowledge Capital: A Case of Indian Farming

Bino Paul [1,*], Unmesh Patnaik [2], Subash Sasidharan [3], Kamal Kumar Murari [2] and Chandra Sekhar Bahinipati [4]

1. Centre for Human Resources Management and Labour Relations, School of Management and Labour Studies, Tata Institute of Social Sciences, Mumbai 400088, India
2. Centre for Climate Change and Sustainability Studies, School of Habitat Studies, Tata Institute of Social Sciences, Mumbai 400088, India
3. Department of Humanities and Social Sciences, Indian Institute of Technology Madras, Chennai 600036, India
4. Department of Humanities and Social Sciences, Indian Institute of Technology Tirupati, Yerpedu 517619, India

* Correspondence: bino@tiss.edu

Abstract: Using the recently released microdata covering input use in Indian agriculture, this study analyzes the relation between value and fertilizer consumption along with four layers of explanation. These layers include factors of production, knowledge capital, social identity, and human capital for both agricultural seasons. Subsequently, the study also examines the propensity to use diverse channels of information. This study uses both regression and machine learning methods for analysis. The main finding of the study is that fertilizer use is directly associated with the value of production. However, the propensity to use fertilizer is the highest for the lowest quantile. Moreover, fertilizer use is a positive covariant of select information sources. Further, similar to tangible resources, the study observes that information plays a crucial role in fertilizer use. Information channels such as extension services have a pivotal role in promoting sustainable farming, especially among marginal farms.

Keywords: fertilizer; value; information; social identity; human capital; sustainable farming; India

check for updates

Citation: Paul, B.; Patnaik, U.; Sasidharan, S.; Murari, K.K.; Bahinipati, C.S. Fertilizer Use, Value, and Knowledge Capital: A Case of Indian Farming. *Sustainability* **2022**, *14*, 12491. https://doi.org/10.3390/su141912491

Academic Editors: Margarita Maria Brugarolas Molla-Bauza and Laura Martinez-Carrasco

Received: 26 August 2022
Accepted: 28 September 2022
Published: 30 September 2022

1. Introduction

In this study, we explore the relationship between fertilizer use and value creation in India without isolating other dimensions. In vast and complex economies such as India, farming remains a vital activity. Although it generates about half of the employment, its share in the gross domestic product is less than one-sixth [1] (Bahinipati et al., 2021). Unlike in developed countries, where farms are consolidated units, farmlands in India are fragmented. Considering these contexts, combining marginal farms with high productivity is challenging in India. Agarwal & Agrawal [2] report that around half of the farmers are not interested in farming. The feasibility of farming activity depends on appropriate and sustainable use of inputs, factors of production, extension services, reducing inequalities due to social identities, and the use of human capital. This multidimensionality also raises analytical issues. Applying the conventional frequentist statistical approaches requires explanatory variables to be independent of each other. However, the reality is a picture of combining these factors by decision-makers to get outcomes. To a greater extent, machine learning approaches such as classification trees address this issue. Therefore, we combine these approaches to analyze the role of fertilizer in value creation and how the interdependence of multiple factors drives fertilizer use. This analysis is complementary to the current policy measure of scaling up the soil health card (the soil health card scheme is designed to convey quality of soil and requirement of fertilizer dosage to the farmers, with the aim of helping them use fertilizer efficiently [3]) across the states in India since 2016.

Farm productivity is crucial to ensure food security [4,5]. The usage of fertilizer is a significant contributor to farm production, apart from cropping intensity and irrigation potential [3,5]. Even though the relationship between fertilizer use and production is a stylized fact, the usage behavior is sensitive to heterogeneities that emanate from agro-climatic and socioeconomic factors. Farmers rely on diverse information sources when deciding on fertilizer usage. Available research points to overuse of chemical fertilizers that leads to soil degradation and desertification, and in turn, impacted on yield, nutrition and health [6,7], which impact agricultural sustainability in the long run (see [8]). From a policy vantage, sustainable fertilizer usage requires insights into the above-mentioned factors [7]. In this study, we develop a conceptual framework of different forms of capital or resources. It consists of five layers. The first layer consists of two factors of production: labor and capital. The second layer is the most critical input to farming, i.e., fertilizer usage. The third layer is knowledge capital [9], which is the way farmers acquire information. Quite plausibly, the acquisition of information is not a one-shot process. It involves learning and deciding whether to adopt new farming methods. To empirically capture this, our study considers five layers. The first two are proxied by using formal and private extension services, whereas the third one explores the role of media. The fourth layer is identities, which are potential sources of social capital. Identities such as gender and social group may generate advantages for some and disadvantages for others. The fifth layer is human capital, proxied by educational attainment. We do not have an a priori hypothesis regarding which dimension will be paramount. However, the importance of the most critical factor will emerge from the empirical analysis.

The research has the following objectives: first, we examine the determinants of value from farming using a multidimensional framework that includes factors of production, fertilizer use, knowledge capital, human capital, and social capital; second, we gauge the linkage between knowledge capital and above-mentioned dimensions; and third, we examine the principal drivers of fertilizer use. This paper contributes to the literature in terms of empirical framework and methodology. First, this paper uses a nationally representative dataset of farming units in India to carry out a detailed analysis of both monsoon and post-monsoon seasons. Second, our comprehensive data enables us to include the role of agroclimatic regions (India holds the second-largest agricultural land in the world, with 20 agroclimatic regions and 157.35 million hectares of land under cultivation (https://eands.dacnet.nic.in/PDF/Glance-2016.pdf, accessed on 20 July 2022)). Third, our analysis captures the influence of multiple factors on acquiring information from formal and informal sources. Finally, from a methodological point of view, we use the recently developed machine learning techniques to investigate the determinants of fertilizer use.

Agriculture is the primary source of income for many rural households in India. A variety of agricultural technologies (i.e., fertilizer, HYVs, land preparation practices, SRI), information (climate and soil), and capacity-building programs (Krushi Mela, Krushi Vigyan Kendra, etc.) have been promoted over the years (see [10]). During the last decade, several policies related to agriculture have been implemented at the national level with the aim of doubling farmers' income (e.g., National Mission on Sustainable Agriculture, Pradhan Mantri Krishi Sinchayee Yojana, Soil Health Card Scheme, Pradhan Mantri Fasal Bima Yojana, etc. [1]). Although agricultural growth is impossible without mechanization, low adoption has been reported in India [1,11–13]. Numerous studies, therefore, identified the confounding factors of adoption, and those are related to economics, farm organization, demographics, extension agents, risk aversion behavior, social learning, and environmental conditions [11,12,14]. On the other hand, several impact evaluation studies have estimated the impact of these technologies on reducing poverty and enhancing yield across African and Asian countries [10,15]. Adopting both experimental (RCT) and quasi-experimental methods (Propensity Score matching, Difference-in-Difference, Regression Discontinuity, Endogenous switching regression), most studies have observed that interventions have a positive impact on agricultural outcome indicators [10,16]. Previous studies addressed interventions such as land tenancy and tilling, extension services, irrigation, natural resource management, input technology, climate information, marketing arrangements, micro-irrigation, microfinance, and crop insurance [10,16–19].

A significant increase in crop yield during the Green Revolution (the Green Revolution was initiated in India in 1960s to increase food production and alleviate extreme poverty by introducing high-yielding varieties of seeds) in India is attributed to the adoption of exploitative agriculture [20] or intensive inorganic farming systems [21]. However, the crop yield growth rate began to taper off in a few decades [22]. A serious re-examination of the flattening of the yield pointed to an imbalanced usage of chemical fertilizer and a massive fertilizer subsidy leading to improper application [23]. Although there was some concern about deteriorating soil health, the academic research focus was mainly around crop yield, imbalanced use of fertilizer, and excessive fertilizer subsidies [23–25]. In contrast, in soil science, scholars [26–28] seemed to be taking a more holistic view of the problem by taking into account soil health, crop yield, and even economic return in their studies. When investigating the issue of disproportionate fertilizer use, there are other related issues such as dominance of urea as the main source of fertilizer and the high rate of fertilizer usage among small farmers [29]. This raises the main question that the paper seeks to explore about determinants of the use of fertilizer in the Indian context.

According to the National Sample Survey (NSS) 70th round (NSS Report No. 576: Income, Expenditure, Productive Assets and Indebtedness of Agricultural Households in India, [30]), the more minor the farm is, the more the productivity tends to decrease, whereas indebtedness directly varies with the size of the farm. These patterns seem to convey that vulnerability from farming activity seems to exist across scales. It entails innovations that sustain technically and economically feasible farming practices through process innovation, such as improvement in soil health. From the above discussion, it is evident that previous research was predominantly focused on the economics of farming, mainly measuring efficiencies; there seems to be a significant lacuna in unravelling process innovations such as improvement in soil health that are irrespective of scales. Therefore, it is essential to understand significant policy measures such as distribution of soil health cards in semi-arid regions of India and explain the factors that make this initiative sustainable.

2. Data and Methods

We used the Situation Assessment Survey (SAS) of Agricultural Households, which was part of the 70th Round conducted by the National Sample Survey Organization (NSSO), which is exceptionally detailed data of nationally representative farm households in India. We structured the microdata from the 70th round of the NSS “Situation of Agricultural Households in India” into four streams. First, the study extracts the full sample for both seasons: season 1 (covering crops from July to December 2012) and the following season 2 (covering crops from January to June 2013). The survey was carried out during 1 January–31 December 2013. The data collection was carried from the same household twice during the survey period. The first visit was from January to July 2013, while the second was from August to December 2013, covering 4529 villages across India. The first round of surveys covered 35,200 households, while the second visit covered 34,907 households. This survey used a stratified multistage design. While the census village is the first stage unit, the household is the ultimate stage unit. The stratum refers to the district level. For non-hilly states, except the state of Kerala, the substrata comprised a group of homogeneous villages in cultivated areas. The survey schedule contains 15 blocks, capturing identification of the sample, field operation, household characteristics, demographic aspects, output, inputs, value of output, expenses, assets, liability, expenditure, awareness about minimum support price, access to technical advice, and other aspects. For the analysis, this study uses household characteristics, demographic elements, the value of output and input, assets, access to technical advice, and consumption expenditure.

The study examines the explanations for three outcomes: SALES, FERTZ, and EXTN (see Table 1 for the definition of abbreviations used in the empirical model). While some explanations are common, some are specific to the outcome. There are five layers of explanatory variables. In the first layer, there are two variables: LAB and ASSETS. This layer is called factors of production (see Table 1). Substantive literature on farm economics suggests that labor and capital (assets) impact production volume [31–33]. The second layer captures fertilizer consumption. While fertilizer is one among many inputs, such as pesticides and seeds, it is the principal one in terms of value. Hence, the analysis was restricted to fertilizer. For the third layer, we included three categories of extension: FORMEXTN, PVTEXTN, and MEDIA. This layer is a proxy for knowledge capital [34,35]. The next layer is social identity. Further, two identities were also covered: social group (SOCGRP) and GENDER [36]. It is crucial to capture the advantage or disadvantages due to social identity. Finally, the fifth layer is the proxy for human capital, based on educational attainment [37]. Table 1 provides variables and definitions (see Appendix A for the descriptive statistics).

Table 1. Variable Description.

Layer	Variable	Definition
Factors of Production	LAB	Average labor cost per hectare of land (in natural logs)
	ASSETS	Agricultural assets of the household normalized by members of the household (in natural logs)
Fertilizer Consumption	FRITZ	Average consumption of fertilizer per hectare of land (in natural logs)
Knowledge Capital (Extension)	EXTN	Access to technical advice for crops from Extension Agent
	KVK	Access to technical advice for crops from Krishi Vigyan Kendra
	UNIV	Access to technical advice for crops from Agricultural University
	PRGFRM	Access to technical advice for crops from Progressive Farmer
	PVT	Access to technical advice for crops from Private Commercial Agents
	NGO	Access to technical advice for crops from Non-Governmental Organizations
	MEDIA	Access to technical advice for crops from Radio/Newspaper / Television / Internet
	FORMEX	EXTN + KVK + UNIV
	PVTEX	PRGFRM + PVT + NGO
Identity	ST	Households belonging to the social category of Scheduled Tribes (Reference Category)
	SC	Households belonging to the social category of Scheduled Castes
	OBC	Households belonging to the social category of Other Backward Classes
	OTH	Household belonging to the social category Others
	GEND	Whether the head of the household is female or male
Human Capital	IT	No general education (Reference Category)
	PRIM	The primary level of general education
	SEC	Secondary level of general education
	HSDIP	The level of general education is either Higher Secondary or Diploma
	GRAD	The level of general education is Graduate and above
Others	MPCE	Monthly per capita consumption expenditure (in natural logs)
	SURPLUS	Value of Output minus Value of Input
	FE	NSS State Region (Proxy for Agro-Climatic Conditions)
Outcome	SALES	Total output sold per hectare of land (in natural logs)
	FRITZ	Expenditure on fertilizer per hectare categorized into below median and median and above
	EXTN	MEDIA, FORMEX, PVTEX

Source: Authors' Table.

This study uses five methods for multivariate analysis. These methods include frequentist statistical and machine learning (non-parametric) approaches. We estimate Ordinary Least Square (OLS) and Simultaneous Quantile Regression (SQREG). The OLS model examines the impact of explanatory variables on the outcome of sales per hectare, subject to fixed effects emanating from the proxy of agroclimatic regions. Considering the likelihood of a fat tail of the distribution, we let the central tendency move along the axis. However, in order to ensure that the OLS estimations are robust (in terms of signature) for the entire distribution, other than the central tendency, we estimate quantile regression [38]. Quantile regression allows the analysis to be sensitive to the fat tail problem. Accordingly, quantiles were used to describe the distribution, i.e., 20, 40, 60, and 80 quantiles [39]. Equation (1) presents the econometric specification for the OLS and SQREG.

$$y = x\beta + u \tag{1}$$

In Equation (1), y represents the outcome variable, and xs' are sets of vectors of independent variables and control factors. u is the well-behaved error term. This study also used a logistic regression model for the second research question (determinants of knowledge capital) since the dependent variable is dichotomous. We estimate the following Equation (2):

$$y_i^* = x\beta + e;\ where\ y_i^* = 1\ if\ y_i = 1,\ otherwise\ 0 \tag{2}$$

In Equation (2), y_i depicts the access to extension services and the rest of the terms are as defined in Equation (1). The odds ratios (e^{β}) were used to communicate the inferences of the logistic regression. According to the framework of the logistic regression, if this ratio is more than one, for a particular variable, the odds in favor of the outcome exceed the odds against it. The reverse holds if the ratio is less than one.

In order to ensure the robustness of the results based on parametric methods (both OLS and quantile regressions), we estimate the conditional inference tree to assess the explanatory variables' interdependence in influencing the outcome. This analysis absorbs all independent variables. The rationale for using the classification trees as an analytical strategy is that the frequentist methods miss possibilities raised by the data. It is essential to understand how the combination of independent variables accounts for outcome variation. Otherwise, there is a chance of having numerous dichotomous variables; this causes a reduction in the degrees of freedom. On the other hand, a classification tree starts with a bifurcation: one branch without any subnodes and another with two subnodes and many terminal nodes.

The random forest algorithm [40] aggregates the predictions of multiple decision trees. Every decision tree is an outcome of the process of training and bootstrapping the data. It culminates in a hierarchy of ordered variables based on their importance. In random forest, recognition of importance is based on the Gini Index. This index assesses the impurity of the data to a node based on a split. The Gini Index is defined as

$$G = 2p\ (1 - p) \tag{3}$$

where p equals the proportion of positive cases assigned to a particular node. $(1 - p)$ if the fraction of negative cases. More purity of a node implies smaller Gini coefficients. In the forest, the overall importance is the average of its importance value among all trees [41]. The misclassification proportion is an important measure that matches observed and predicted categories, known as the confusion matrix. It gives the count of correct classifications. Table 2 presents the matrix, and Equation (4) presents the correct classification ratio.

$$The\ ratio\ of\ Correct\ Classification = \frac{TP + TN}{TN + TP + FP + FN} \tag{4}$$

Table 2. Definition for types of classification.

Predicted / Actual	False	True
False	True Negative (*TN*)	False Positive (*FP*)
True	False Negative (*FN*)	True Positive (*TP*)

Table 3 outlines the research questions, model, method, listing outcome, and explanatory variables. Moreover, the table provides a method for a particular model.

Table 3. Summary of research questions, model, and method.

Question	Variables (Model)		Method
	Outcome	Explanatory	
Determinants of value from farming	Sales per hectare	FP, FU, KC, ID, HC	OLS, SQREG
Linkage with knowledge capital	KC	FP, FU, ID, HC, MPCE	LOGIT
Principal drivers of fertilizer use	FU (above median & below median)	FP, ID, HC, KC, SURP	CTRE, Forest

Note: FP = Factors of Production, FU = Expenditure on Fertilizer per hectare, KC = Knowledge Capital, ID = Social Identity, HC = Human Capital, MPCE = Monthly per capital consumption expenditure, SURP = Surplus per hectare, OLS = Ordinary least square, SQREG: Simultaneous quantile regression, LOGIT = Logistic regression, CTRE = Conditional inference tree, Forest = Random Forest. Source: Authors' Table.

3. Results and Discussion

3.1. Determinants of Value from Farm

As mentioned above, there are five layers of explanation accounting for variation in average sales of farming units. In the first layer, labor cost shows the highest magnitude of impact (0.26), whereas the effect of assets is negligible (0.09) (Table 4). Consumption of fertilizer influences the outcome variable, showing a coefficient of 0.21. From the third layer, all three explanatory variables are significant. It implies that average sales tend to increase if the farmer seeks information from these sources. The coefficient varies in the range 0.11 (private extension) to 0.19 (media) (Table 4). These findings imply that seeking information is associated with a positive payoff in sales. In the fourth layer, compared to ST (reference group), OBC and others show higher coefficients (0.13 and 0.26, respectively). However, the coefficient on gender is statistically insignificant. For the last layer, represented by education, compared to the illiterates (reference group), coefficients progressively increase with higher levels of educational attainment (0.11 for primary education, 0.39 for university education) (see Table 4). These inferences have exciting implications. The first and second layers convey the significance of economic capital. The social group (fourth layer) is a proxy for how identities translate to advantages or certain forms of social capital. The third layer conveys the propensity to acquire knowledge through diverse channels. Knowledge about these channels may translate to capitalizing knowledge. Educational attainment, forming the fourth layer, represents human capital. It also leads to payoffs for farmers. A possible limitation of this inference is that it may alter the estimates if the central tendency varies along the axis, enveloping the lower and higher tail and the median. Given this, we estimate quantile regressions to the same model except for fixed effects.

Table 4. Estimated results for Season 1.

Variables	OLS	Quantile			
		20th	40th	60th	80th
FRITZ	0.211 *** (0.199)	0.369 *** (0.042)	0.347 *** (0.027)	0.288 *** (0.020)	0.268 *** (0.016)
ASSETS	0.093 *** (0.008)	0.147 *** (0.011)	0.149 *** (0.009)	0.153 *** (0.009)	0.172 *** (0.009)
LAB	0.261 *** (0.019)	0.027 (0.028)	0.093 *** (0.022)	0.146 *** (0.022)	0.205 *** (0.020)
FORMEX	0.149 *** (0.041)	0.231 *** (0.056)	0.164 *** (0.057)	0.141 *** (0.048)	0.101 *** (0.039)
PVTEX	0.111 *** (0.131)	0.090 *** (0.037)	0.043 (0.031)	0.056 (0.035)	0.075 * (0.042)
MEDIA	0.192 *** (0.033)	0.144 ** (0.065)	0.185 *** (0.054)	0.167 *** (0.044)	0.132 *** (0.053)
SC	−0.106 (0.069)	−0.190 ** (0.095)	−0.149 ** (0.073)	−0.215 *** (0.077)	−0.208 ** (0.088)
OBC	0.134 ** (0.054)	0.199 *** (0.063)	0.134 ** (0.059)	0.119 ** (0.063)	0.083 (0.056)
OTH	0.265 *** (0.057)	0.299 *** (0.058)	0.266 *** (0.059)	0.315 *** (0.053)	0.329 *** (0.046)
PRIM	0.111 *** (0.039)	0.135 ** (0.063)	0.028 (0.055)	0.055 (0.048)	0.071 (0.064)
SEC	0.151 *** (0.039)	0.053 (0.059)	0.040 (0.045)	0.039 (0.045)	0.073 (0.057)
HSDIP	0.194 *** (0.059)	0.052 (0.089)	−0.023 (0.085)	0.012 (0.068)	0.139 * (0.079)
GRAD	0.392 *** (0.058)	0.269 *** (0.079)	0.116 (0.104)	0.126 (0.109)	0.171 ** (0.083)
GEND	−0.045 (0.066)	−0.033 (0.123)	−0.019 (0.108)	−0.053 (0.092)	−0.154 ** (0.073)
CONST	5.683 *** (0.219)	5.068 *** (0.343)	5.566 *** (0.203)	6.169 *** (0.174)	6.683 *** (0.182)
OUTCOME			SALES		
FE	YES	NO	NO	NO	NO
R^2/Pseudo R^2	0.375	0.079	0.101	0.116	0.129
N	6568	6568	6568	6568	6568

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$; Robust Standard Errors in parentheses for OLS and Bootstrapped Standard Errors in parentheses for Quantile Regression; Mean Variance Inflation Factor (VIF) = 3.26.

An interesting pattern is that average labor cost shows either insignificant or no impact for the first two segments of the lower tail. In contrast, it shows a visible association with the upper quantiles (0.15 and 0.2 for the upper two quantiles, respectively; see Table 4). The coefficient for fertilizer consumption consistently declines from the lower quantile to the upper quantile. It varies in the range of 0.37 to 0.27. Compared to other variables, it emerges as the most convincing explanation for the outcome across quantiles. Assets remains a significant explanatory variable across quantiles, varying in the range of 0.15 to 0.17 (see Table 4). The coefficient of formal extension consistently declines with quantiles. For the lowest quantile, the estimated coefficient is 0.23, whereas for the highest quantile, it is 0.1. It implies that, for marginal farmers, seeking information from formal sources translates to higher average sales. However, this is not valid for private extension, which is not statistically significant (at five percent levels) across quantiles. However, media is impactful across quantiles, showing a nonlinear relationship. It increases from the lowest quantile to the center and subsequently reduces. For the social group, intergroup differences across categories are more visible at higher quantiles. The differences between the socially advantaged and disadvantaged groups at higher quantiles widen. In the case of gender, at the highest quantile, the identity of being a woman adversely impacts the outcome variable, showing a coefficient of −0.15 (15 percent decline in average sales) (see Table 4). For education, there is no impact in the middle quantiles, whereas higher education directly impacts the outcome in the lowest and highest quantile.

Regarding Season 2, fertilizer use emerges as an impactful variable accounting for variation in sales (0.18) (see Table 5). From the pool of factors, labor and assets show significant coefficients (0.22 and 0.17, respectively). In the layer of information, formal channels of extension are the strongest (0.17), followed by media (0.11) and private sources (0.07). Concerning identity, the remaining categories are insignificant for social groups except for the category of others (0.18). Moreover, gender differences turn out to be insignificant. For human capital, coefficients for all categories are significant. Quite importantly, in the case of this variable, the coefficient progressively increases as the level of education increases. It varies from 0.11 to 0.23 (see Table 5).

Table 5. Estimated results for Season 2.

Variables	OLS	Quantile			
		20th	40th	60th	80th
FRITZ	0.177 *** (0.020)	0.321 *** (0.017)	0.275 *** (0.018)	0.259 *** (0.020)	0.244 *** (0.015)
ASSETS	0.112 *** (0.008)	0.160 *** (0.012)	0.166 *** (0.009)	0.159 *** (0.006)	0.171 *** (0.007)
LAB	0.215 *** (0.019)	0.078 ** (0.032)	0.159 *** (0.024)	0.171 *** (0.021)	0.163 *** (0.016)
FORMEX	0.171 *** (0.039)	0.195 *** (0.065)	0.176 *** (0.045)	0.146 *** (0.044)	0.205 *** (0.043)
PVTEX	0.068 ** (0.029)	0.084 (0.054)	0.025 (0.043)	0.023 (0.034)	0.001 (0.043)
MEDIA	0.110 *** (0.031)	0.058 (0.048)	0.018 (0.038)	0.016 (0.038)	0.042 (0.044)
SC	−0.049 (0.064)	−0.254 *** (0.065)	−0.258 *** (0.083)	−0.166 ** (0.065)	−0.215 *** (0.065)
OBC	0.082 (0.053)	0.022 (0.064)	0.006 (0.069)	0.032 (0.052)	0.016 (0.068)
OTH	0.178 *** (0.055)	−0.030 (0.071)	0.057 (0.076)	0.15 *** (0.056)	0.213 *** (0.061)
PRIM	0.105 *** (0.037)	0.071 (0.052)	0.007 (0.044)	0.062 (0.054)	0.033 (0.034)
SEC	0.157 *** (0.035)	0.095 * (0.045)	−0.017 (0.058)	0.016 (0.034)	0.012 (0.034)
HSDIP	0.172 *** (0.052)	0.185 *** (0.067)	0.048 (0.072)	0.035 (0.063)	0.016 (0.047)
GRAD	0.234 *** (0.055)	0.129 (0.109)	−0.039 (0.068)	−0.012 (0.047)	−0.017 (0.079)
GEND	−0.084 (0.059)	0.036 (0.099)	−0.131 (0.096)	−0.067 (0.075)	−0.183 ** (0.082)
CONST	5.926 *** (0.185)	5.285 *** (0.192)	5.793 *** (0.171)	6.342 *** (0.168)	7.076 *** (0.166)
OUTCOME			SALES		
FE	YES	NO	NO	NO	NO
R^2/Pseudo R^2	0.358	0.085	0.10	0.114	0.122
N	6510	6510	6510	6510	6510

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$; Robust Standard Errors in parentheses for OLS and Bootstrapped Standard Errors in parentheses for Quantile Regression; Mean Variance Inflation Factor (VIF) = 2.68.

From lowest to highest quantile, the coefficient of fertilizer use declines consistently; the value varies from 0.24 to 0.32. However, there is a mixed pattern for assets. The coefficient shows an inverted U pattern from the 20th to the 60th quantiles. It rises from 0.16 to 0.17 and then falls to 0.16. Then it increases again to 0.17 for the 80th quantile. For labor cost, coefficients resemble an inverted U pattern. From the lowest quantile, the coefficient increases from 0.08 to 0.17 (60th quantile) and declines to 0.16 (80th quantile). Unlike the OLS model, only formal extension emerges significantly across quantiles. Coefficients show a U pattern. From the lowest quantile, it declines from 0.2 to 0.15 (60th quantile) and rises to 0.21 (80th quantile). The rest of the information channels are not statistically significant. Compared to the reference category ST, for the 60th and 80th quantile, the category others reports statistically significant coefficients (0.15 and 0.21 for 60th and 80th quantile, respectively), showing a premium. Except for the 20th quantile, education categories do not turn out to be significant. Further, gender is not statistically significant across quantiles except for the 80th one (see Table 5).

3.2. Determinants of Information Acquisition

In order to identify the determinant of information acquisition, we estimate logistic regression. From Table 6, it is evident that human capital shows a clear pattern of choice of media as an information channel with the odds ratio increasing from 1.5 (lowest education) to 2.4 (university education). The household consumption level impacts media choice, showing an odds ratio of 1.3. We observe a similar result in the case of fertilizer use, reporting an odds ratio of 1.2. However, the odds ratio for assets is just above one. Compared to the reference group, except for SCs, the rest of the categories report higher odds ratios. While OBC reports an odds ratio of 1.3, the odds ratio for others is 1.7 (see Table 6). In the case of gender of the farmer, women farmers are less inclined towards media use. The previous pattern of human capital that emerged in the media case is also valid for the formal extension. The coefficient increases from 1.3 (primary level of education) to 2.6 (university level). However, fertilizer shows a diminished impact, reporting a feeble odds ratio of 1.1 (see Table 6).

Table 6. Logistic Regression Odds Ratios.

Variables		Round 1			Round 2	
ASSETS	1.078 *** (0.013)	1.108 *** (0.018)	1.028 ** (0.012)	1.141 *** (0.015)	1.209 *** (0.021)	1.087 *** (0.014)
MPCE	1.307 *** (0.066)	1.243 *** (0.079)	1.299 *** (0.064)	1.207 *** (0.063)	1.138 * (0.082)	1.190 *** (0.062)
FRITZ	1.190 *** (0.029)	1.099 *** (0.035)	1.161 *** (0.027)	1.056 *** (0.024)	1.066 ** (0.034)	1.157 *** (0.026)
SC	1.111 (0.122)	1.383 ** (0.197)	1.109 (0.110)	0.873 (0.095)	0.983 (0.147)	1.344 *** (0.141)
OBC	1.298 *** (0.115)	1.413 *** (0.152)	1.253 *** (0.099)	1.101 (0.098)	1.059 (0.126)	1.567 *** (0.141)
OTH	1.725 *** (0.162)	1.674 *** (0.194)	1.234 ** (0.106)	1.297 *** (0.119)	1.042 (0.129)	1.465 *** (0.141)
PRIM	1.539 *** (0.099)	1.345 *** (0.111)	1.133 ** (0.067)	1.331 *** (0.085)	1.144 (0.103)	1.031 (0.064)
SEC	1.929 *** (0.123)	1.901 *** (0.154)	1.276 *** (0.074)	1.704 *** (0.104)	1.598 *** (0.137)	1.098 (0.064)
HSDIP	2.571 *** (0.255)	2.144 *** (0.276)	1.345 *** (0.127)	1.786 *** (0.166)	1.609 *** (0.215)	1.004 (0.092)
GRAD	2.391 *** (0.237)	2.569 *** (0.318)	1.101 (0.109)	2.054 *** (0.198)	2.173 *** (0.278)	0.953 (0.091)
GEND	0.826 * (0.082)	0.984 (0.125)	0.997 (0.107)	0.839 * (0.084)	0.887 (0.124)	0.815 ** (0.079)
CONST	0.005 *** (0.002)	0.115 *** (0.006)	0.003 *** (0.001)	0.031 *** (0.015)	0.02 *** (0.013)	0.004 *** (0.002)
Outcome	MEDIA	FORMEX	PVT	MEDIA	FORMEX	PVT

Table 6. *Cont.*

Variables	Round 1			Round 2		
FE			YES			
Wald Chi^2	1663.17 ***	1368.07 ***	1248.20 ***	1400.66 ***	1317.46 ***	1540.84 ***
Pseudo R^2	0.174	0.178	0.116	0.145	0.185	0.151
N	11884	11875	11821	11538	11540	11467

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$; Robust Standard Errors in parentheses.

In contrast, consumption retains the same impact, with an odds ratio of 1.24. For assets, the effect remains feeble. From the identity layer, gender is not statistically significant, whereas all categories in a social group are statistically significant. The odds ratio varies from 1.4 (SC) to 1.7 (others) (see Table 6). For private extension, human capital shows a different pattern. The category of graduates is not statistically significant, whereas others are. It varies from 1.1 (primary) to 1.3 (higher secondary). For this information channel, fertilizer is more impactful than the previous category (odds ratio of 1.16). Consumption retains its impact, showing an odds ratio of 1.3. However, the effect of assets is feeble in using this channel as a source of information. Gender is statistically insignificant, except for SCs coefficients of other categories, which are significant, hovering around 1.2 (see Table 6).

In season 2, human capital's impact on adopting media as an extension source is quite similar to the behavior during the Kharif season. The odds ratio consistently increases with the level of education; it varies in the range of 1.3 to 2.1 (see Table 6). Fertilizer usage reports a feeble impact, whereas consumption level shows an odds ratio of 1.2. Further, assets report an odds ratio of 1.4. Gender shows less than one odds ratio, whereas only the category others in the social group becomes significant (1.3). The human capital story repeats in the case of formal extension except for primary education. Whereas the impact of fertilizer use is relatively weak, consumption and assets show odds ratios of 1.1 and 1.2, respectively (see Table 6). Neither gender nor social group is significant. However, previous patterns observed in human capital do not hold for the choice of a private extension, with none of the categories having any significance. A meaningful change is that fertilizer use becomes more impactful, with an odds ratio of 1.16 (see Table 6). The consumption level reports an odds ratio of 1.2; however, the impact on assets is relatively weak. The odds ratio for gender is less than one. All categories show significant odds ratios for the social group, varying between 1.3 (SC) and 1.6 (OBC) (see Table 6).

3.3. Determinants of Fertilizer Use: A Machine Learning Approach

The determinants of fertilizer use are presented using machine learning techniques and are presented in Figure 1. This tree has 21 nodes, of which 11 are terminating ones. Labor accounts for the highest impurity reduction in the tree. Chances of above-median fertilizer use vary from one-fourth to three-fifths. The lowest probability of fertilizer use is for the terminal node, which is a combination of high labor cost, OBC, SC, and others in the social group and access to private information as a source of extension. In contrast, the combination of low labor cost and the social group others shows the highest probability, followed by the combination of low labor cost and the rest of the social group categories. Only these two combinations show the likelihood of fertilizer use to exceed half. For two combinations, probability hovers around two-fifths. The combination is of high labor cost, ST in the social group and illiterate, diploma, higher secondary in the educational attainment and male in the gender category. The second combination covers high labor cost, ST in the social group, graduate and primary in educational attainment, and low surplus. The rest of the nodes vary from one-fourth to slightly less than one-fifth.

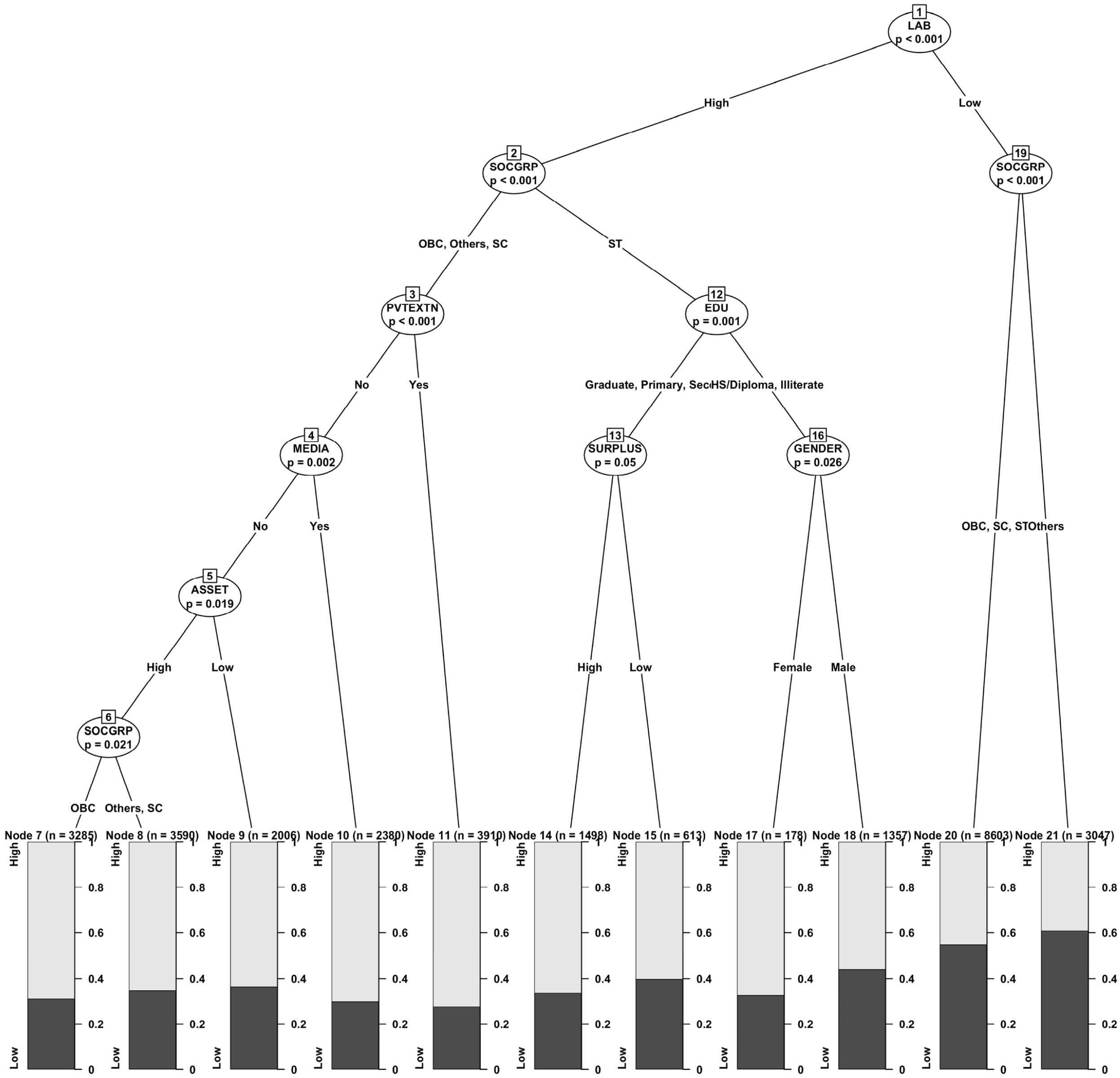

Figure 1. Output of Conditional Inferential Tree for Season 1.

For season 2 (Figure 2), the tree consists of 19 nodes. Ten are terminating ones. The probability varies in the range of one-fourth to four-fifths. The lowest probability is for the combination of high labor cost, use of private extension, social category OBC, SC, and others, and no media usage. On the other hand, the combination of low labor cost, no use of media, formal extension, and ST in the social category shows the highest probability. The next highest probability is the combination of low labor cost and media usage (a bit above three-fifths). Two more combinations show probabilities closer to three-fifths. First is the combination of low labor cost, no use of media, and others in the social group. The second consists of low labor cost, no media, OBC and SC in social groups, and usage of a private extension. The combination of low labor cost, no media, OBC, SC, ST, and no formal extension reports a probability of close to half. Among the rest of the terminating nodes, only one shows a probability of two-fifths. It consists of high labor costs, the use of private extensions, and groups belonging to ST.

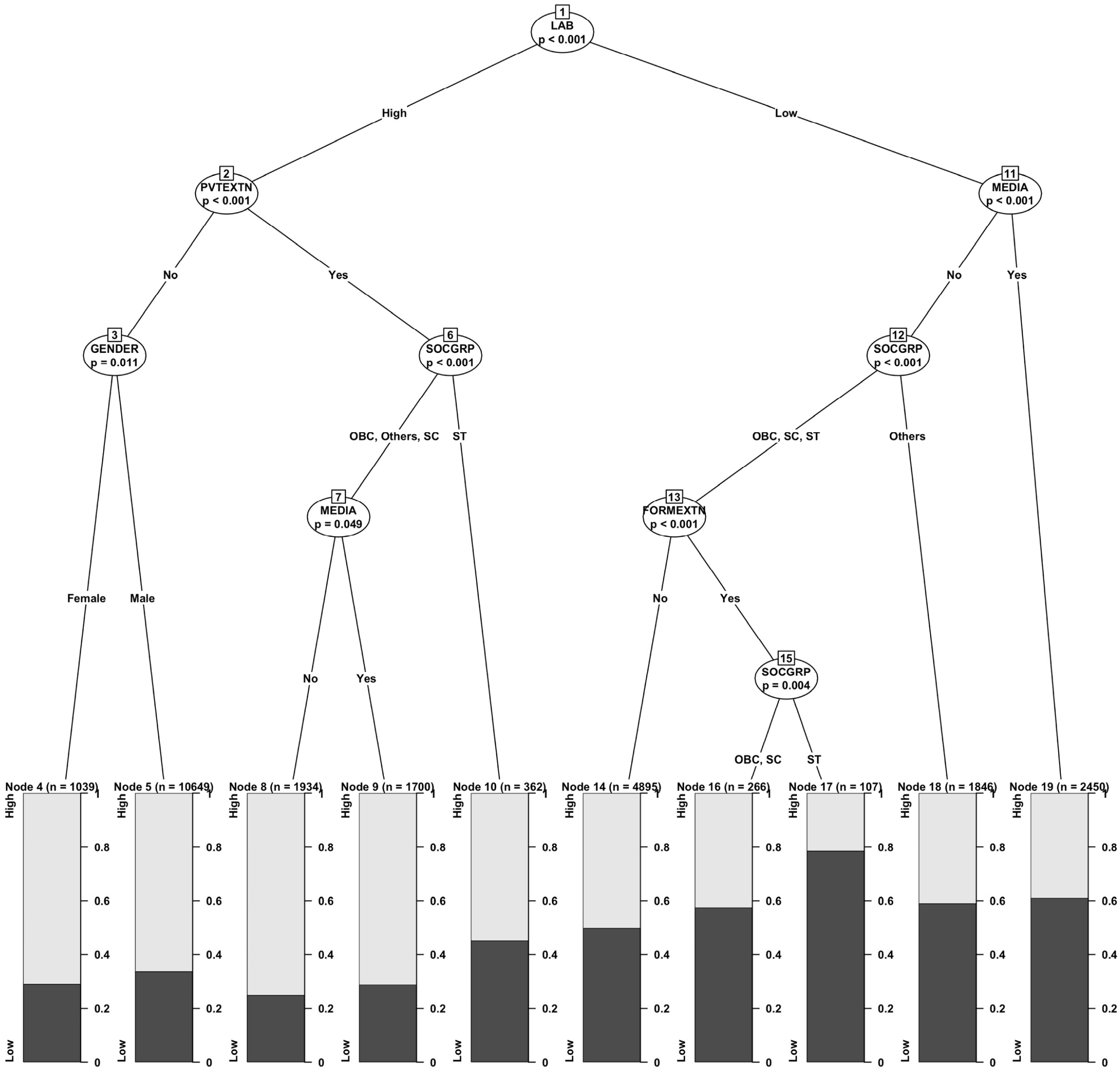

Figure 2. Output of Conditional Inferential Tree for Season 2.

To gauge the importance of the most crucial determinant of fertilizer use, we use a random forest approach (see Figure 3). The random forest draws cues from several random conditional trees from the sample; this method generates an ordered plotting of factors—the criterion for variable importance is the mean decrease in entropy measured by the Gini Index. For season 1, surplus emerges as the most critical variable (see Figure 3A). The other two variables which stand out are formal extension and private extension. However, the rest of the variables are homogeneous regarding variable importance. This pattern does not apply for season 2 (see Figure 3B). In this case, formal extension is most important. Assets and surplus also stand out in terms of importance. The rest are more or less similar to each other. In a nutshell, formal extension emerges as the most important common factor. Compared to the other forms of capital, knowledge capital is more crucial in explaining fertilizer usage by farmers in India. It is entirely plausible that extension services play a significant role in decision-making by the farmers.

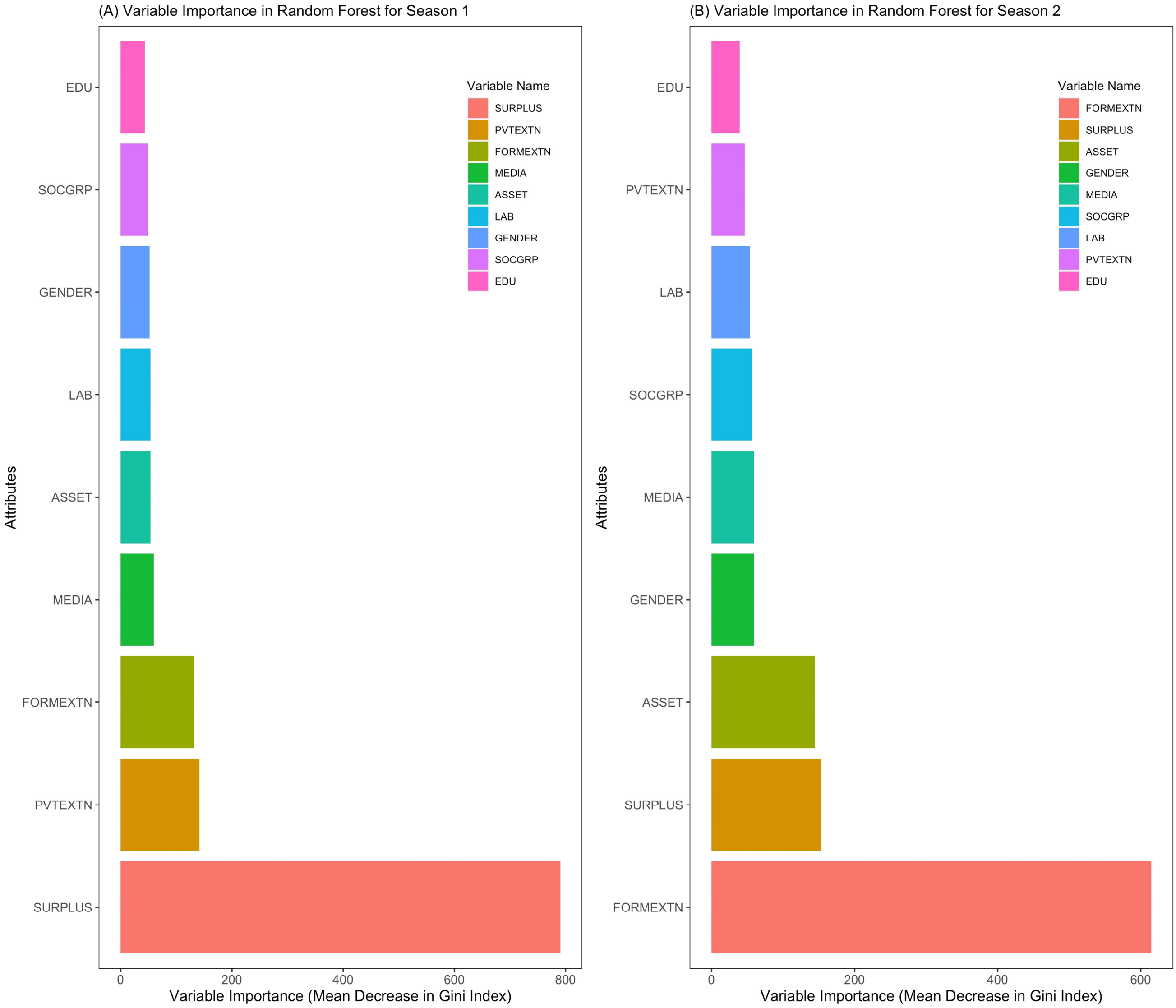

Figure 3. Output of Random Forest Analysis.

4. Conclusions

In this study, we attempted to analyze the determinants of fertilizer usage of farm households in India using a comprehensive nationally representative dataset. The analysis of farm households data reveals that: (a) fertilizer expenditure is an impactful determinant of a farm's performance across agroclimatic regions and seasons, (b) even though extension services also influence a farm's performance, the propensity to seek these services depends on human capital, and (c) access to an extension service makes a huge positive impact on consumption of fertilizer. From the food security angle, these findings suggest a need for developing an in-depth understanding of how systems such as extension services can contribute to the sustainable use of fertilizer. Quite importantly, the strong linkage between the lower tail (marginal farmers) and higher propensity to use fertilizer raises questions if different sources of information convey the context and the meaning of the sustainable use of fertilizer. Extension services or media seem to be impactful gatekeepers of the system that determines the intensity of input usage or the adoption of new technology. Here, the policy challenge is to create a common ground between sustainable use of input, value creation by farms, and innovative dissemination of information. In the future, the policy that does not integrate these three aspects may pave the way for disruptive changes in food security. If the information dissemination is not mindful of the heterogeneity of farming units, it results in unbalanced input use. This behavior may culminate in the viciousness of low farm productivity that causes shocks to food security.

Author Contributions: Conceptualization, B.P., U.P. and K.K.M.; Formal analysis, B.P. and U.P.; Methodology, B.P., U.P., S.S., K.K.M. and C.S.B.; Writing—original draft, U.P., K.K.M. and C.S.B.; Writing—review & editing, S.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Indian Council of Social Science Research (ICSSR), New Delhi, under the Impactful Policy Research in Social Science scheme (IMPRESS/P2504/312/18-19/ICSSR).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data is curated from public sources. The dataset can be made available on request from the authors.

Acknowledgments: We are grateful to two anonymous reviewers for comments and suggestions. Views expressed in the paper are those of the authors and not those of the supporting or collaborating institutions. The usual disclaimers apply.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Descriptive Statistics for Continuous Variables.

Variables	Round 1				Round 2			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
FRITZ	7.67	1.08	1.93	15.02	7.64	1.21	2.04	14.99
ASSETS	5.55	2.03	−1.10	15.89	5.32	1.94	−1.10	13.92
LAB	8.01	1.11	1.43	15.71	7.83	1.18	1.97	15.05
SALES	10.05	1.48	2.20	16.22	9.92	1.46	3.91	15.24
MPCE	7.12	0.56	−1.95	13.33	7.19	0.54	−2.30	11.33
SURPLUS	9.36	1.46	1.73	19.29	9.11	1.48	2.23	17.03

Table A2. Descriptive Statistics for Categorical Variables.

Variables	Categories (%)	
	Round 1	Round 2
EXTN	Yes (7.91); No (92.09)	Yes (7.65); No (92.35)
KVK	Yes (4.3); No (95.7)	Yes (4.61); No (95.39)
UNIV	Yes (1.85); No (98.15)	Yes (1.89); No (98.11)
PRGFRM	Yes (19.54); No (80.46)	Yes (20.9); No (79.1)
PVT	Yes (6.12); No (93.88)	Yes (7.25); No (92.75)
NGO	Yes (1.16); No (98.84)	Yes (1.46); No (98.54)
MEDIA	Yes (23.74); No (76.26)	Yes (26.3); No (73.7)
FORMEX	Yes (12.53); No (87.47)	Yes (12.61); No (87.39)
PVT	Yes (23.54); No (76.46)	Yes (25.68); No (74.32)
ST	18.96	19.01
SC	13.24	13.25
OBC	40.32	40.28
OTH	27.48	27.46
GEND	Male (91.58); Female (8.42)	Male (91.61); Female (8.39)

Table A2. *Cont.*

Variables	Categories (%)	
	Round 1	Round 2
ILT	ILT (34.41)	ILT (34.41)
PRIM	26.53	26.53
SEC	27.64	27.64
HSDIP	6.13	6.13
GRAD	5.29	5.29

Table A3. Confusion Matrix for Conditional Inference Trees.

Sample Details	Predicted / Actual	Number of Cases		Correct Classification Ratio
		Yes	No	
Season 1	Yes	12,627	6190	0.631
	No	5067	6583	
Season 2	Yes	13,103	7477	0.631
	No	1846	2823	

Table A4. Confusion Matrix for Random Forest.

Sample Details	Predicted / Actual	Number of Cases		Correct Classification Ratio
		Yes	No	
Season 1	Yes	13,101	4593	0.631
	No	6653	6120	
Season 2	Yes	11,919	3029	0.639
	No	6098	4202	

References

1. Bahinipati, C.S.; Kumar, V.; Viswanathan, P.K. An Evidence based systematic review on farmers' adaptation strategies in India. *Food Secur.* **2021**, *13*, 399–418. [CrossRef]
2. Agarwal, B.; Agrawal, A. Do farmers really like farming? Indian farmers in transition. *Oxf. Dev. Stud.* **2017**, *45*, 460–478. [CrossRef]
3. Fitzpatrick, I.A.; Millner, N.; Ginn, F. Governing the soil: Natural farming and bio-nationalism in India. *Agric. Hum. Values* **2022**, 1–16. [CrossRef]
4. Akbari, M.; Foroudi, P.; Shahmoradi, M.; Padash, H.; Parizi, Z.S.; Khosravani, A.; Ataei, P.; Cuomo, M.T. The Evolution of Food Security: Where Are We Now, Where Should We Go Next? *Sustainability* **2022**, *14*, 3634. [CrossRef]
5. Yousaf, M.; Li, J.; Lu, J.; Ren, T.; Cong, R.; Fahad, S.; Li, X. Effects of fertilization on crop production and nutrient-supplying capacity under rice-oilseed rape rotation system. *Sci. Rep.* **2017**, *7*, 1270. [CrossRef] [PubMed]
6. Lal, R. Soil degradation as a reason inadequate human nutrition. *Food Secur.* **2009**, *1*, 45–57. [CrossRef]
7. Lal, R. Restoring soil quality to mitigate soil degradation. *Sustainability* **2015**, *7*, 5875–5895. [CrossRef]
8. Lampridi, M.G.; Sorensen, C.G.; Bochtis, D. Agricultural sustainability: A review of concepts and methods. *Sustainability* **2019**, *11*, 5120. [CrossRef]
9. Rahman, M.M.; Connor, J.D. Impact of Agricultural Extension Services on Fertilizer Use and Farmers' Welfare: Evidence from Bangladesh. *Sustainability* **2022**, *14*, 9385. [CrossRef]
10. IEG (Independent Evaluation Group). *Impact Evaluations in Agriculture: An Assessment of the Evidence*; World Bank: Washington, DC, USA, 2011.
11. Bahinipati, C.S.; Patnaik, U. What motivates farm-level adaptation in India? A systematic review. In *Climate Change and Community Resilience: Insights from South Asia*; Springer: Berlin/Heidelberg, Germany, 2022; Chapter 4, pp. 49–68.
12. Foster, A.D.; Rosenzweig, M.R. Microeconomics of technology adoption. *Annu. Rev. Econ.* **2010**, *2*, 395–424. [CrossRef]
13. Viswanathan, P.K.; Kavya, K.; Bahinipati, C. Global patterns of climate resilient agriculture: A review of studies and imperatives for empirical research framework for India. *Rev. Dev. Change* **2020**, *25*, 169–192. [CrossRef]
14. Feder, G.; Just, R.E.; Zilberman, D. Adoption of agricultural innovations in developing countries: A survey. *Econ. Dev. Cult. Chang.* **1985**, *33*, 255–298. [CrossRef]
15. Boudot, C.; Butler, A.; Dugal, N. Evaluating technologies for agricultural development: How to capture the true impact? *Sci. Changements Planét. Sécheresse* **2013**, *24*, 374–384. [CrossRef]

16. Patnaik, U.; Das, P.K.; Bahinipati, C.S. Developmental Interventions, Adaptation Decision and Farmers' Well-being: Evidence from Drought-prone Households in rural India. *Clim. Dev.* **2019**, *11*, 302–318. [CrossRef]
17. Bahinipati, C.S.; Viswanathan, P.K. Can Micro-Irrigation Technologies Resolve India's Groundwater Crisis? Reflections from Dark-Regions in Gujarat. *Int. J. Commons* **2019**, *13*, 848–858. [CrossRef]
18. Biswal, D.; Bahinipati, C.S. Why are farmers not insuring crops against risks in India? A Review. *Prog. Disaster Sci.* **2022**, *15*, 100241. [CrossRef]
19. Patt, A.; Suarez, P.; Gwata, C. Effects of seasonal climate forecasts and participatory workshops among subsistence farmers in Zimbabwe. *Proc. Nat. Acad. Sci. USA* **2005**, *102*, 12623–12628. [CrossRef] [PubMed]
20. Swaminathan, M.S. An evergreen revolution. *Crop Sci.* **2006**, *46*, 2293–2303. [CrossRef]
21. Thakur, D.S.; Sharma, K.D. Organic farming sustainable agriculture and meeting the challenges of food security in 21st century: An economic analysis. *Indian J. Agric. Econ.* **2005**, *60*, 1–15.
22. Prasad, J. Soil health management—A key for sustainable production. *J. Indian Soc. Soil Sci.* **2015**, *63*, 6–13.
23. Shamrao, T.S. A study of fertilizer policy in India. *Int. J. Agric. Sci.* **2011**, *3*, 145. [CrossRef]
24. Gulati, A.; Narayana, S. Demystifying fertilizer and power subsidies in India. *Econ. Polit. Wkly.* **2000**, *35*, 784–794.
25. Singh, B. Are nitrogen fertilizers deleterious to soil health? *Agronomy* **2018**, *8*, 48. [CrossRef]
26. Chander, G.; Wani, S.P.; Krishnappa, K.; Sahrawat, K.L.; Parthasaradhi, G.; Jangawad, L.S. Soil mapping and variety-based entry-point interventions strengthening agriculture-based livelihoods-exemplar case of 'Bhoochetana' in India. *Curr. Sci.* **2016**, *110*, 1683–1691. [CrossRef]
27. Chander, G.; Wani, S.P.; Sahrawat, K.L.; Dixit, S.; Venkateswarlu, B.; Rajesh, C.; Pardhasaradhi, G. Soil test-based nutrient balancing improved crop productivity and rural livelihoods: Case study from rainfed semi-arid tropics in Andhra Pradesh, India. *Arch. Agron. Soil Sci.* **2014**, *60*, 1051–1066. [CrossRef]
28. Chander, G.; Wani, S.P.; Sahrawat, K.L.; Kamdi, P.J.; Pal, C.K.; Pal, D.K.; Mathur, T.P. Balanced and integrated nutrient management enhanced and economic food production: Case study from rainfed semi-arid tropics in India. *Arch. Agron. Soil Sci.* **2013**, *59*, 1643–1658. [CrossRef]
29. Murari, K.K.; Jayaraman, T.; Chakraborty, S. *Fertilizer Use and the Small Scale Farms. How Do Small Farmer Fare? Evidence from Village Studies in India*; Tulika Books: Chennai, India, 2017; pp. 201–229.
30. National Sample Survey Organization (NSSO). *Government of India, Income, Expenditure, Productive Assets and Indebtedness of Agricultural Households in India*; NSS Report No.576; National Sample Survey Organization: New Delhi, India, 2013.
31. Chavas, J.P.; Chambers, R.G.; Pope, R.D. Production economics and farm management: A century of contributions. *Am. J. Agric. Econ.* **2010**, *92*, 356–375. [CrossRef]
32. Cornia, G.A. Farm size, land yields and the agricultural production function: An analysis for fifteen developing countries. *World Dev.* **1985**, *13*, 513–534. [CrossRef]
33. Joshi, P.K.; Pal, S.; Birthal, P.S.; Bantilan, M.C.S. *Impact of Agricultural Research: Post-Green Revolution Evidence from India*; National Centre for Agricultural Economics and Policy, Research, New Delhi International Crops Research Institute for Semi-Arid Tropics: Patancheru, India, 2005.
34. Evenson, R.E.; Jha, D. The contribution of agricultural research system to agricultural production in India. *Indian J. Agric. Econ.* **1973**, *28*, 212–230.
35. Garforth, G.; Jones, C.; Jones, G.; Garforth, C. The history, development, and future of agricultural extension. In *Improving Agricultural Extension: A Reference Manual*; Swanson, B.E., Bentz, R.P., Sofranko, A.J., Eds.; FAO: Rome, Italy, 1997.
36. Omvedt, G. Capitalist Agriculture and rural classes in India. *Econ. Polit. Wkly.* **1981**, *15*, A140–A159. [CrossRef]
37. Huffman, W.E. Human capital: Education and agriculture. *Handb. Agric. Econ.* **2001**, *1*, 333–381.
38. Murari, K.K.; Mahato, S.; Jayaraman, T.; Swaminathan, M. Extreme temperatures and crop yield in Karnataka, India. *Rev. Agrar. Stud.* **2019**, *8*. Available online: http://ras.org.in/91036314fb4c1338277b3da6a01a6a36 (accessed on 20 June 2022).
39. Koenker, R.; Hallock, K.F. Quantile regression. *J. Econ. Perspect.* **2001**, *15*, 143–156. [CrossRef]
40. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
41. Qi, Y. Random Forest for Bioinformatics. 2012, pp. 307–323. Available online: http://www.cs.cmu.edu/~{}qyj/papersA08/11-rfbook.pdf (accessed on 20 May 2022).