Indian Economic Review (2020) 55:253–282
https://doi.org/10.1007/s41775-020-00092-4

ARTICLE

Determinants of various modes of rural non-farm sector (RNFS) employment in SAT (semi-arid tropics) and Eastern regions of India: an empirical analysis

Anviksha Drall[1] · Sabuj Kumar Mandal[1]

Published online: 27 August 2020

Abstract
The objective of the present study is to identify the major motivating factors and thereby tracing out the existence of any entry barrier for several categories of rural non-farm sector (RNFS) employment in India. We conduct our analysis using household-level data from semi-arid tropics (SAT) and Eastern regions of India for the period 2010–2014. We disaggregate the RNFS activities into various categories—wage employment, self-employment, and others—and use a multinomial logit model as the baseline model to determine the factors driving participation in the various types of non-farm employment. Furthermore, Heckman Selection Model to account for selection bias in our sample and a multinomial fractional logit model to account for the intensity of RNFS income are used. The empirical results, based on a multinomial logit model, reveal that education in general and technical education, in particular, access to credit and endowment of social capital, are the major determinants of RNFS employment in India. However, these determinants are not same across the various RNFS sub-sectors. It is found that while education affects participation in wage employment and self-employment, technical education affects participation in wage employment and others only. Also, social capital determines employment in self-employment and wage employment, but does not determine employment under the 'others' category. Other factors that determine RNFS diversification are land and non-land assets, age, and gender of the household head, household size and distance from market. Policy implications of our empirical results are also discussed.

Keywords Diversification · Rural non-farm sector · Entry barriers · Multinomial logit model · Multinomial fractional logit model · Heckman

JEL Classification Q10 · Q12

✉ Anviksha Drall
anviksha20@gmail.com

[1] Department of Humanities and Social Sciences, Indian Institute of Technology, Madras, India

1 Introduction

The farm sector has always been perceived as the major source of income for rural households in developing and transition economies. However, of late, diversification into the rural non-farm sector (hereinafter, RNFS) has become a norm across these economies. For instance, the RNFS accounted for a significant portion of the total income in Asia (32%), Africa (42%), and Latin America (40%) (Reardon et al. 1998). This could be attributed to the abysmal conditions of the farm sector which pushes the farmers (especially the small and marginal ones) into the RNFS, or to the relatively higher return in the RNFS which pulls the farmers towards the RNFS. As noted by Ellis (2007), farm households grapple with the issues of income instability and consumption smoothening, and in such a scenario, RNFS diversification becomes a vital tool. Reardon et al. (1998) discuss the growing importance of the RNFS in terms of providing purchasing power by relaxing the credit constraints in farm production and enhancing food security. Thus, the RNFS, once considered as a low productive sector, expected to wane with the rise in the development of a nation, is now discerned as a sector capable of absorbing surplus labour of the economy (Lanjouw and Lanjouw 2001). Given the fact that rural income is now not just limited to farm income, there has been an increasing focus on studying and understanding the drivers of employment in the RNFS.

Development economics literature (Reardon et al. 1998; Haggblade et al. 2007; Kassie et al. 2017) classifies the reasons behind RNFS diversification broadly into push and pull factors. The conventional pull factors that impel households to diversify into the RNFS are the relatively remunerative return to labour and/or capital and the less risky nature of investments in the RNFS (Kilic et al. 2009). On the other hand, the push factors which push households into the RNFS consist of small landholdings, large household size, and inadequate production of farm output. The farm output could be low and/or insufficient due to the households' inability to undertake adequate farm expenses resulting from their limited risk-bearing capacity, owing to imperfect insurance and credit markets (Reardon 1997; Barrett et al. 2001; Ruben 2001; Kilic et al. 2009). Thus, the households are compelled to diversify their income source to effectively mitigate the resulting risk as well as to finance farm investments. Moreover, uncertain climatic conditions coupled with market failures also act as a push factor for RNFS diversification. Finally, technological advancements in the farm sector, which leads to a decline in the demand for farm labour, also push households to look for better livelihood strategies like RNFS diversification (Lien et al. 2010).

Several empirical studies (Canagarajah et al. 2001; Corral and Reardon 2001; Escobal 2001; Woldenhanna and Oskam 2001) have been conducted to investigate the determinants of RNFS diversification. Apart from the push and the pull factors, these studies identify the individual, household, economic, and community-level characteristics which determine the household's participation in the RNFS. For instance, individual characteristics like education, technical education, age, and gender of the household members have been considered as vital determinants of

RNFS diversification in the studies undertaken by Abdulai and CroleRees (2001), Canagarajah et al.(2001), Corral and Reardon (2001), Escobal (2001), Woldenhanna and Oskam (2001), Launjow and Shariff (2004), and Jatav and Sen (2013). Household characteristics, like the household's size, have also been paid attention to in the literature (Woldenhanna and Oskam 2001; Dimova and Sen 2010; Teshome and Edriss 2013). The determinants of RNFS participation also include the household's economic characteristics, like endowment (e.g., landholdings, livestock animals, physical assets etc.) and access to loans/credit, which determine its capacity to diversify (Corral and Reardon 2001; Escobal 2001; Woldenhanna and Oskam 2001). Finally, access to public goods such as roads and electricity has also been emphasized as a determinant of RNFS diversification by many researchers (Reardon et al. 1998; Elbers and Lanjouw 2001). Most of these studies have been conducted in the context of Asian (Lanjouw and Lanjouw 2001; Schwarze and Zeller 2005; Jatav and Sen 2013) and African countries (Abdulai and CroleRees 2001; Canagarajah et al. 2001; Kassie et al 2017). Research on RNFS diversification and its determinants has drawn a lot of focus and interest in India too (Lanjouw and Shariff 2004; Lanjouw and Murgai 2009; Himanshu et al. 2013). India witnesses RNFS diversification both in terms of income and employment, with 41% of the workforce employed in the RNFS in the year 2016 (The state of food and agriculture by Food and Agriculture Organization of the United States), more so in the debilitated regions of the SAT (semi-arid tropics) where households find it difficult to sustain their livelihoods due to extreme poverty. This is so because the farm activities in these regions are mostly rainfed in nature. Consequently, frequent droughts occurring in these regions adversely affect the livelihoods and food security of the households (Singh et al. 2014; Kumar et al. 2018). Apart from this, the SAT regions experience erratic rainfall, late monsoon onset, and extreme temperatures (Diatta 2017). Moreover, the SAT households also face farm production constraints in the form of soil infertility, lack of irrigation resources, and poor rural infrastructure (Singh et al. 2014; Kumar et al. 2018). Hence, with their limited earning opportunities, the households in the SAT region are more likely to diversify into the RNFS as compared to the other regions of India. However, there is limited empirical evidence on the factors affecting RNFS diversification in the SAT regions of India (Reddy et al. 2014).

Nevertheless, most studies pertaining to RNFS diversification have considered non-farm employment in its aggregate form, without classifying it further into various sub-categories. The determinants of RNFS diversification, however, may not always be homogenous across the several RNFS sub-categories. It might be plausible that entry into some RNFS sub-sectors is more restrictive as compared to others. For instance, Unni (1991) notes that business activities and salaried jobs have higher entry barriers than agriculture or any other RNFS activity. Considering human capital as an entry barrier, Kilby and Liedholm (1988) view training as a more significant entry barrier for manufacturing than petty trade. Given the possibility that the entry barriers may vary across various RNFS subsectors, we add to the existing literature by examining whether the entry barriers to RNFS employment are homogenous across the various RNFS categories in the SAT (Andhra Pradesh, Gujarat, Karnataka, Madhya Pradesh, and Maharashtra) and Eastern states (Bihar, Jharkhand, and Odisha) of India. To our knowledge, ours is the first study that classifies RNFS employment into various categories and investigates whether the entry barriers are homogenous across those categories in the SAT and Eastern states of India. Such a study would help policy makers to identify the entry barriers across different categories of RNFS employment in these regions and render vital policy implications for enhancing the welfare of the poor, who are neither able to earn subsistent income from the farm sector nor are able to enter the RNFS.

The data for the empirical analysis come from the VDSA (Village dynamics in South Asia) micro-level data, where 1415 rural households are studied for a panel of 5 years (2010–2014). The dataset has several notable features. First, as mentioned before, the SAT regions are drought prone in nature; the vulnerability and resilience of the households to drought and other climatic uncertainties not only varies at the household level but also at the temporal level (Kumar et al. 2018). Thus, the RNFS diversification by these households may vary from time to time, which could be better captured by a dataset like the VDSA which is panel in nature. Second, the dataset befits the requirements of the study as it considers the various sources of diversification, like non-farm self-employment and wage employment, in detail for the SAT and Eastern regions where RNFS diversification is of great significance. Third, the dataset being panel in nature also accounts for the impact of unobserved household-specific omitted variables, like risk averseness and managerial ability, which may also determine RNFS diversification. This is in contrast with the prior studies in the Indian context (Lanjouw and Shariff 2004; Lanjouw and Murgai 2009; Himanshu et al. 2013; Jatav and Sen 2013) on the determinants of diversification which have relied on cross-sectional data, and did not account for the temporal dimension of diversification and unobserved household-specific factors which motivate the households to diversify.

Methodologically, this analysis is supported by various econometric models like the multinomial logit model and the fractional multinomial logit model, which account for the interdependence between various categories of RNFS employment. To the best of our knowledge, ours is the first study that employs a multinomial logit model and a multinomial fractional logit model in a panel setup to identify the homogeneity of the entry barriers across various RNFS sub-categories with an objective to bring out the differences, if any, in those barriers across categories, in the SAT and Eastern regions of India.

The empirical results confirm the existence of entry barriers across the RNFS categories—wage employment, self-employment and others—and find that the barriers are heterogeneous across these categories. It is found that while the education of the household head affects the participation in wage employment and self-employment, technical education affects participation in wage employment and others. Also, while social capital determines self-employment and wage employment, it does not determine employment under the other category.

The remainder of the paper is organized as follows. Section two reviews the studies in the context of RNFS diversification. Section three describes the data source and the patterns of income diversification in the study areas. In Section four, model specifications and the estimation procedures are discussed. Section five comprises of a brief discussion on the estimation results, followed by the concluding remarks and policy implications in section six.

2 Review of literature

An empirical estimation of the determinants of RNFS participation is generally underpinned by two steps. The first step is to unravel why the households diversify into RNFS, the answer to which is rooted in a body of theoretical literature. The second step entails the measurement of the diversification in a proper way, which has been carried out in the empirical literature on RNFS diversification. In this section, we briefly discuss the theoretical literature (Sect. 2.1) and the empirical literature (Sect. 2.2) on RNFS diversification.

2.1 Theoretical literature

To find out the reasons behind RNFS diversification, standard rural household models have been developed in the literature (Huffman 1980; Abduali and CroleRess 2001; Goodwin and Mishra 2004; Olale and Henson 2012; Demie and Zeray 2016). These models aim to maximize a non-separable utility function subject to several constraints: (a) time constraint, (b) budget constraint, (c) farm production constraint, and (d) a non-negativity constraint on the time allocated to the RNFS.

The first-order conditions of such a theoretical model provide the optimal time allocation solution between farm work, various types of non-farm work, and leisure activities. The reduced form equation for the optimal time allocated to the jth RNFS activity by the ith household is of the following form:

$$T_{ij}^{*} = T_{ij}\left(T, P, P_{y,} w_2, w_1, P_x, O, \Delta^c, \Delta^z, B, \Omega, M\right), \tag{1}$$

where T denotes the time endowment of the household members; P denotes the price vector corresponding to the vector of consumption goods and services; P_y is the price of the farm output;w_1 denotes the wage rate of hired labour, while w_2 denotes the RNFS wage rate; P_x represents the vector of price of the non-labour inputs required for farm production; O denotes the household income from other sources; Δ^c represents the household characteristics affecting the expected utility of the household, where Δ^c includes variables like the age of the household head, gender of the household head, education of the household members and the household size, among others; Δ^z represents the household characteristics affecting household production decisions. Again, Δ^z includes the age of the household head, gender of the household head, education of the household members, etc.; B represents the barriers to farm production, such as access to credit, while Ω represents the locational characteristics affecting production decisions—for example, distance from the market, infrastructure (availability of electricity) etc.; finally, M denotes the fixed factors of farm production like land or cattle.

The theoretical models can be further classified as one period (Huffman 1980; Goodwin and Mishra 2004; Olale and Henson 2012) or T period (Abduali and

CroleRees 2001; Bongole 2016), based on the number of time periods over which a household maximizes its utility.

2.2 Empirical literature

Households earn ‘income’ by engaging themselves in ‘activities’ using their ‘assets’ (Barrett et al. 2001). The literature, thus, differentiates between the three parameters, namely income, activities and assets, for measuring diversification. The most used parameters to indicate/measure diversification are ‘activity’ and ‘income’. Activity is generally measured by asking the households if they have participated in an RNFS activity or not. To measure income, either the gross value of the non-farm income or the share of non-farm income in the total income is considered. Some studies (Shehu and Abubakar 2015) also use ‘assets’, such as the number of household members (human capital) that engage in the RNFS, for the measurement of diversification.

The empirical studies measure the diversification decision based on these parameters by either considering the overall RNFS diversification or by focussing on the specific categories/modes of the RNFS. Consequently, the diversification decisions are estimated empirically using various econometric techniques. Section 2.2.1 covers the empirical studies which estimate the determinants of the overall RNFS diversification. Following this, Sect. 2.2.2 describes the empirical studies which estimate the factors affecting the modes of diversification. As mentioned in Sect. 1, research on RNFS diversification and its determinants has drawn a lot of focus and interest in India too. Hence, Sect. 2.2.1 critically reviews the empirical studies on RNFS diversification in the Indian context.

2.2.1 Empirical literature on overall diversification

The overall diversification has been measured by Abdul-Hakim and Hadijah Che-Mat (2011), Akaakohol and Aye (2014), Ghimire et al. (2014), Asfaw et al. (2017), and Kassie et al. (2017) based on the ‘activity’/participation decision. The participation decision is inferred by asking whether the household has diversified or not using the dummy variable technique. Hence, the logit/probit models have been employed for the empirical estimation of the diversification decision in this case. However, the dummy variable technique does not differentiate between the various levels of RNFS income. Hence, using ‘income’ as a parameter of diversification is preferred over ‘activity’. Several studies (Canagarajah et al. 2001; Teshome and Edriss 2013; Weldegebriel et al. 2015) use ‘non-farm income’ as the measure of diversification. For empirical estimation, Canagarajah et al. (2001) and Teshome and Edriss (2013) use the ordinary least square technique and the Tobit model, respectively. To account for the unobserved fixed effects in a panel data setup, Weldegebriel et al. (2015) apply a fixed-effects model. Nevertheless, the absolute value of non-farm income cannot capture the intensity of diversification. Thus, the ‘share of non-farm income’ in the total income has been considered by some researchers like Schwarze and Zeller (2005), Tran Quang (2014), Bongole (2016), and Demie and Zeray (2016). In this case, Schwarze and Zeller (2005), Bongole (2016), and Demie and Zeray (2016) carry out the empirical estimation using a Tobit model. Tran Quang (2014), on the other hand, uses a fractional logit model for the empirical estimation. Schwarze and Zeller (2015), based on 'activity', also use the Shannon equitability index to measure the overall diversification and estimate the same using a Tobit Model. Dimova and Sen (2010) use the Herfindahl Index to capture the concentration of diversification (based on 'activity'). To embrace the unobserved features of a household, like the attitude towards risk, they apply fixed and random effect models to estimate the diversification decisions. Besides, focussing only on the non-farm enterprise activities, Shehu and Abubakar (2015) measure diversification using the share of household members who engage in self-employment activities of the RNFS (based on 'asset'). They estimate the same using a Tobit model.

2.2.2 Empirical literature on modes of diversification

Abdulai and CroleRees (2001), Corral and Reardon (2001), Woldenhanna and Oskam (2001), and Demie and Zeray (2016) categorize diversification into various categories/modes based on the 'activity'/participation decision (whether the household participated in Category One, whether the household participated in Category Two and so on). Corral and Reardon (2001) use Probit regressions—separate regression on each category—for their empirical estimation. However, the separate regression technique does not consider the simultaneous decision-making of the households' time allocation problem. To overcome this limitation, the conditional fixed-effects logit model is used by Abdulai and CroleRees (2001). Similarly, Woldenhanna and Oskam (2001) and Demie and Zeray (2016) deploy a multinomial logit model to empirically estimate the factors affecting the modes of diversification. Since the dummy variable technique does not distinguish between the various levels of income, Corral and Reardon (2001) also classify the income from diversification into various categories (Income from Category One, Income from Category Two, and so on). In this context, they utilize a censored least absolute deviation regression on each category. Akin to Corral and Reardon (2001), Kassie et al. (2017) also classify the income from diversification into various groups. However, in contrast to Corral and Reardon (2001), they use a SUR model to account for the correlation in the error terms of several categories. Escobal (2001), Malek and Usami (2009) and Sendaza (2012), to capture the intensity of diversification, categorize the modes of diversification based on the share of RNFS income in the total income (Share of RNFS income from Category One in the total income, Share of RNFS income from Category Two in the total income, and so on). They estimate the determinants of the various modes of diversification using separate Tobit regression on each category. Besides, Woldenhanna and Oskam (2001) classify the categories based on the hours allocated to each RNFS category and estimate the same using separate Tobit regression on each category.

2.2.3 Empirical studies in the Indian context

Table 1 presents an overview of the limited empirical studies on RNFS diversification undertaken in the Indian context. Although conducted in the context of

Table 1 A brief overview of the studies on RNFS diversification in the Indian context

Author	Year	Data	Research question	Econometric methodology used
Jean O. Lanjouw and Peter Lanjouw	2001	Multiple sources; Census of India (1991) for Indian context	To analyze the role of RNFS in economic growth of the developing countries through literature survey	NA
Peter Lanjouw and Abusaleh Shariff	2004	A household survey by NCAER on 16 Indian states for the year 1993–1994	To find out the determinants of overall diversification and modes of diversification	CLAD (censored least absolute deviation) method for overall diversification; Multinomial logistic regression for modes of diversification
Peter Lanjouw and Rinku Murgai	2009	Cross-sectional data of five thick NSSO rounds (1983, 1987/1988, 1993/1994, 1999/2000 and 2004/2005)	To find out the determinants of overall diversification (for 1983 and 2004) and modes of diversification for each NSSO round	Multivariate analysis for overall diversification; multinomial logit model for modes of diversification
Himanshu et al. (2011, 2013)	2011, 2013	Cross-sectional data of five thick NSSO rounds (1983, 1987/1988, 1993/1994, 1999/2000 and 2004/2005) supplemented by a survey data on Palanpur from May 2008 to April 2010	To track the changes in rural non-farm sector at the Indian level since 1980s (overall growth of the sector; changes in employment in various RNFS sub-sectors)	Estimates based on employment and Unemployment survey for various NSS rounds
Manoj Jatav and Sucharita Sen	2013	Individual-level NSSO data from 2009–2010 employment and unemployment round (Cross-sectional data)	To find out the determinants of overall diversification	Logistic regression

developing countries, the study by Lanjouw and Lanjouw (2001) is one of the former studies in this regard. The study presents an evidence of the shift towards the RNFS, which was once expected to shrink away with the stages of development of a nation, and its contribution to economic growth, poverty alleviation, and rural employment. The rest of the studies in the table are a by-product of the literature on rural poverty and/or inequality or on the restructuring of the RNFS. For instance, Lanjouw and Shariff (2004) and Lanjouw and Murgai (2009) focus on the contribution of the RNFS in rural poverty alleviation. Hence, they find the impact of a growing RNFS on agricultural wages in India. Besides, their studies render attention to the determinants of overall RNFS diversification and the modes of diversification. Similarly, Himanshu et al. (2011) empirically estimate the impact of RNFS employment on poverty rates and agricultural wage rates. Their analysis is supplemented by various trends in the growth of the RNFS and the changes in employment in the various RNFS sub-sectors (manufacturing, services and trade & transport; salaried jobs, casual jobs, and self-employment) over the years. Additionally, Himanshu et al. (2013) assess the impact of RNFS diversification on village-level inequality. Finally, the study by Jatav and Sen (2013) attempts to understand the growth of RNFS based on the NSS data from 2009–2010 and analyze the restructuring of the RNFS towards casual labour. They also find significant entry barriers to RNFS in terms of education, age, and gender.

Our study contributes to the empirical literature on RNFS diversification in the following aspects: First, the discourse on RNFS diversification in the Indian context (Table1) attempts to analyze RNFS diversification at the all-India level. However, as discussed earlier, RNFS diversification strategies are relatively more important for the SAT regions than the other regions of India. None of the studies so far have rendered attention to these regions. Our study, on the other hand, analyzes the RNFS diversification in the SAT and the Eastern regions of India. Second, all the studies in the Indian context are based on the cross-sectional datasets which are unable to capture the unobservable features of the households that lead to RNFS diversification. Also, the temporal variations in RNFS diversification in response to climatic changes remain unexplained by a cross-sectional dataset. We overcome this limitation using the VDSA dataset which covers 1415 households for a period of 5 years (2010–2014). Third, though some of the studies so far have used an SUR model or a multinomial logit model to account for the interdependence between various RNFS employment choices, there is a plausibility of sample selection bias, since all the households are not involved in RNFS diversification. None of the studies so far address the issue of selection bias. Thus, apart from utilizing a multinomial logit model as used in the literature, we also employ a multinomial logit model that accounts for sample selection. Fourth, the multinomial logit model used in the literature treats RNFS employment merely as a polychotomous variable, thereby ignoring the intensity of participation of the households in various RNFS activities. To overcome this limitation, we utilize a multinomial fractional logit model, indeed, accounting for the sample selection bias. Fifth, the existing empirical studies so far have used only the value of livestock to account for the endowment/fixed factors of farm production while neglecting the importance of other assets. We make an improvement over the previous studies using non-land asset as a variable, which not only considers the value of livestock but also adds the value of durables, stock inventory, and farm implement to it.

Recognizing the gaps in the literature, the objective of this study is to empirically estimate the determinants, focussing on the entry barriers, of the modes of diversification in the SAT, and the Eastern regions of India in a panel setup using the VDSA dataset. We, thereby, aim to ascertain whether the entry barriers across the various RNFS sub-sectors are homogenous or sector-specific.

The next section provides a brief overview of the data source utilized for the study along with the patterns of income diversification in the study areas.

3 Data and patterns of income diversification

3.1 Data source

To fulfil the objective of this study, the dataset from the VDSA (Village Dynamics in South Asia) project was utilized. The VDSA project was executed in 2009 by ICRISAT (International Crops Research Institute for the semi-arid tropics) in collaboration with BMGF (Bill and Melindia Gates Foundation). The aim of the project was to analyze the poverty levels in the rural households of South Asia's semi-arid and humid tropic areas. The focus countries of the project were India and Bangladesh. The objective of the project was to reduce the incidence of poverty in these countries by providing data at the individual and household level (micro-data) and the district level (meso data). Since the data include information on labour allocation decisions, sources of income, education levels, access to credit, and other socio-economic characteristics of the households, it is in tandem with our data requirements. Hence, the VDSA dataset (VDSA 2013) was considered appropriate for our study.

Our study is based on the household-level panel data that is obtained from the VDSA micro-level data. In this study, we limit our analysis to the Indian subcontinent. The study is based on a sample of 1415 households for a panel of five years (2010–2014) covering eight Indian states (Andhra Pradesh, Bihar, Gujarat, Jharkhand, Karnataka, Madhya Pradesh, Maharashtra, and Odisha).

The various sources of employment which provide income to the households are broadly classified into the non-farm and farm categories. Furthermore, three types of non-farm income are distinguished—wage employment, self-employment, and others. 'Wage employment' consists of RNFS wage employment[1] and salaried jobs. 'Self-employment' constitutes tailors, all types of businesses, contractors, renting out one's own machine, running transport vehicles, and caste occupations like that of a carpenter, barber, basket maker, stone cutter, etc. Under the 'others' category comes unskilled labour, gifts and remittances, and savings and interest on deposits. Farm employment, on the other hand, comprises of three activities—crop cultivation, livestock rearing, and agricultural wage work.

[1] Wage employment can be undertaken in the farm sector too. However, we consider farm sector wage employment as a part of farm income in this study.

Table 2 Share of income (in the total income) derived from various sources of employment (2010 and 2014)

Sources of income	2010	2014
Rural non-farm sector income	28.78	42.77
Income from wage employment	9.76	17.77
Income from self-employment	10.27	15.25
Income from others	8.75	9.69
Farm income	71.19	57.23
Crop income	15.76	11.26
Farm wage income	4.85	4.94
Livestock income	50.51	41.19

Source: Authors' own calculation based on VDSA dataset

3.2 Patterns of income diversification

The patterns of diversification in the study areas are encompassed in the tables below, where diversification is measured by the share of income (derived from a particular source) in the total income of the households.

Table 2 depicts the diversification of income sources for the years 2010 and 2014. In 2010, the farm sector is the dominant source of income with almost 71% share in the total income of the households. The corresponding share for RNFS is only 29%, plausibly due to the prevalence of entry barriers in the RNFS. However, an increase in the share of RNFS income in the total income, from 29% in 2010 to 43% in 2014, is observed. Amongst the sub-sectors of the RNFS, the contribution of wage employment to the total income witnessed the largest spike from almost 10% (in 2010) to 18% (in 2014). Since a prerequisite for salaried jobs (a part of wage employment) is the possession of education in general and/or technical education in particular, this spike could be attributed to an increase in the average education of the household members (in years) from 4.9 (in 2010) to 5.6 (in 2014) and a rise in the proportion of household members having technical education from 5% (in 2010) to 7% (in 2014). Strikingly, the share of farm income in the total income decreased from 71% (in 2010) to 57% (in 2014). This is possibly due to the shift of labourers from the farm sector to the RNFS, again because of the relaxation of entry barriers in the RNFS making it attractive. Also, the Indian farm sector is crisis-ridden (Sainath 2018), and is growing at an abysmally low rate, pushing the farmers out of the farm sector, and making it unattractive. Nevertheless, the farm sector remains the dominant source of income both in 2010 and 2014.

Table 3 shows how income diversification varies with the variation in the income strata of the households over the period 2010–2014. The strata have been constructed by ranking the households according to their total income and then dividing them into quartiles (containing equal number of households). For the lowest and the highest income strata, the dominant source of income is the farm sector. The households belonging to the lowest income strata tend to enter the agricultural sector because of various entry barriers in the RNFS like the dearth of education/skill along with cash/credit constraint. This is in line with the hypothesis that the poorer

Table 3 Share of income (in the total income) derived from various sources of employment by income quartiles (2010–2014)

Sources of income	Lowest	Second	Third	Highest
Rural non-farm sector income	43.05	51.89	51.67	27.93
Income from wage employment	5.35	10.06	15.7	13.33
Income from self-employment	13.26	18.62	19.03	11.08
Income from others	24.51	23.27	16.97	3.54
Farm income	56.8	48.11	48.33	71.71
Crop income	23.77	14.09	12.9	8.73
Farm wage income	18.68	14.97	8.07	1.6
Livestock income	14.51	18.99	27.23	61.35

Source: Authors' own calculation based on VDSA dataset

Table 4 Share of income (in the total income) derived from various sources of employment by land size in hectares (2010–2014)

Sources of Income	Landless	< 2	2.01–4	4.01–10	> 10
Rural non-farm sector income	69.06	41.49	28.82	21.12	24.31
Income from wage employment	32.56	15.35	13.28	8.42	7.46
Income from self-employment	21.93	14.25	11.16	9.96	11.06
Income from others	14.57	11.99	4.26	2.75	5.84
Farm income	30.94	58.37	71.43	78.78	75.93
Crop income	0.75	13.26	10.94	9.82	4.67
Farm wage income	17.37	5.7	3.52	1.32	0.47
Livestock income	12.82	39.42	56.9	67.63	70.81

Source: Authors' own calculation based on VDSA dataset

households tend to have lesser access to the RNFS than the better-off households (Reardon et al. 1998). Those belonging to the highest income strata do not face any financial constraint. Thus, it becomes easier for them to specialize on the farm. This concords with the fact that the better-off households own productive assets and have better access to markets, especially financial markets. With the rise in wealth, households in the second and third-income strata are expected to have more education. Thus, they face lesser entry barriers in the RNFS, exhibiting high amounts of diversification. Also, the size of landholdings decreases as we move up the income strata, with 1.63, 1.58, and 1.60 ha as the average landholding size for the first-, second-, and third-income strata, respectively. Thus, with a fall in the average landholding size, the households belonging to the second- and third-income strata are more likely to diversify as compared to the households in the lowest income strata.

Table 4 shows the variation in income diversification in relation to the variation in the land size owned by the households. The table evinces that the dominant source of income for the households in all the land strata, except for the households which are landless, is the farm sector. The landless farmers earn 69%

Table 5 Share of income (in the total income) derived from various sources of employment by gender of the household head (2010–2014)

Sources of income	Male-headed households	Female-headed households
Rural non-farm sector income	34.52	56.06
Income from wage employment	12.87	20.68
Income from self-employment	13.18	18.74
Income from others	8.47	16.48
Farm income	65.61	43.94
Crop income	10.51	13.66
Farm wage income	4.6	9.85
Livestock income	50.32	20.36

Source: Authors' own calculation based on VDSA dataset

of their income from the RNFS, showing the highest amount of diversification. The small and marginal farmers earn almost 41% of their income from the RNFS, witnessing the highest amount of diversification among those who own land. The amount of diversification decreases with the increase in the land size (until 10 ha). This is because the households with large landholdings tend to specialize on the farm, depicting a negative relationship between diversification and land size. However, a 'U'-shaped relationship between land size and diversification is depicted in the table. As is evident, diversification first decreases and then rises with the land size (for households with land size > 10 ha). This concords with the hypothesis that land size and the share of non-farm income in the total income are not always inversely correlated (Reardon 2000). This inverse relationship is possible, because the households with large landholdings are sometimes compelled to diversify due to the presence of linkages farm to the RNFS. Also, the land can be used as a collateral by the households, making it easier for them to diversify into the RNFS.

Table 5 represents how income diversification varies with differences in the gender of the household head over the period 2010–2014. The male-headed households derive most (65%) of their income from the farm sector. The female-headed households, on the other hand, earn most of their income from the RNFS (56%), implying a higher level of diversification as compared to their male counterparts. This is in line with the literature (Hartog et al. 2002; Eckel and Grossman 2008; Croson and Gneezy 2009) which asserts that the degree of risk aversion varies between men and women. These studies summarize that women are more risk averse than men and are 'pushed' towards RNFS diversification (Arslan et al. 2018). Also, as noted by Weldegebriel et al. (2015), to mitigate the risk from farm activities, the female-headed households are sometimes more likely to diversify due to the rigid and patriarchal division of labour that hinders females' employment in the farm sector. For the male-headed households, the dominant income source under the RNFS is self-employment, followed by wage employment and others. However, for the female-headed households, the dominant income source under the RNFS is wage employment, followed by self-employment and others.

Table 6 Share of income (in the total income) derived from various sources of employment by region (2010–2014)

Sources of income	EAST	SAT
Rural non-farm sector income	66.22	26.27
Income from wage employment	26.92	9.13
Income from self-employment	17.22	12.3
Income from others	22.19	4.88
Farm income	33.78	73.81
Crop income	10.59	10.71
Farm wage income	2.14	5.59
Livestock income	21.08	57.54

Source: Authors' own calculation based on VDSA dataset

Table 6 depicts the variation in income diversification across the regions of the study area over the period 2010–2014. The eight states in the study are classified into two categories—EAST and SAT. The EAST category consists of the states having a humid sub-tropical weather with the annual rainfall between 1000 and 2500 mm. The EAST constitutes of the three states—Jharkhand, Bihar, and Odisha. SAT, on the other hand, consists of the semi-arid tropic states which experience sporadic climate with the features of both tropical and dry regions. The SAT regions experience lesser annual precipitation (400–750 mm) as compared to the EAST states. The SAT comprises of the five states—Andhra Pradesh, Gujarat, Karnataka, Maharashtra, and Madhya Pradesh. It is evident from the table that the states belonging to the EAST show more diversification (66%) than those belonging to the SAT, with only 26% share of RNFS income in the total income. For the EAST states, the dominant source of income under the RNFS is wage employment, followed by others and self-employment. For the SAT states, self-employment is the dominant source of income under the RNFS, followed by wage employment and others.

4 Econometric model specification

This section provides an overview of the various econometric models to be utilized to fulfil the objective of this study, along with the descriptions and summary statistics of the variables used for the estimation.

To estimate the determinants of RNFS diversification choices, we specify the empirical model[2] as follows:

[2] The variables in Eq. (1) form the vector of explanatory variables in the empirical model. However, not all the variables from Eq. (1) are used:
(i) Data on T, P and w_2 is unavailable. Hence, they were dropped in Eq. (2).

(ii) P_y is assumed to be constant across the households because of the assumption of a perfectly competitive farm output market. Thus, P_y is dropped from the empirical model.

(iii) Similarly, w_1, P_x are assumed to be constant across the households because of the assumption of perfectly competitive farm input markets. Therefore, w_1, P_x are also dropped from the empirical model.

$$
\begin{aligned}
O_{ijt} = \alpha_i + \beta_1 \mathrm{EDU}_{it} + \beta_2 \mathrm{TEDU}_{it} + \beta_3 \mathrm{CREDIT}_{it} + \beta_4 \mathrm{SOCIAL}_{it} + \beta_5 \mathrm{LANDASSET}_{it} \\
+ \beta_6 \mathrm{LANDSQ}_{it} + \beta_7 \mathrm{NONLANDASSET}_{it} + \beta_8 \mathrm{FARMINCOME}_{it} + \beta_9 \mathrm{GENDER}_{it} \\
+ \beta_{10} \mathrm{AGE}_{it} + \beta_{11} \mathrm{HHSIZE}_{it} + \beta_{12} \mathrm{DISTANCE}_{it} + \beta_{13} \mathrm{ELEC}_{it} + \beta_{14} \mathrm{EAST}_{it} + \varepsilon_{it,}
\end{aligned} \tag{2}
$$

where O_{ijt} is a variable depicting the participation of ith household in activity type 'j' at time 't'; α_i is the time-invariant individual-specific effect (unobserved heterogeneity) and ε_{it} is the composite error term. In addition to the variables from Eq. (1), variables TEDU, SOCIAL, FARMINCOME and EAST have also been used in the empirical estimation. To account for the impact of technical education/skills on RNFS diversification, TEDU has been added as a covariate in Eq. (2). In the literature (Kassie et al. 2017), social capital has been considered as one of the important covariates affecting RNFS diversification. Consequently, SOCIAL has been added as a covariate in Eq. (2). As per the literature (Huffman 1980), the size of farm output is considered as an important variable which determines the diversification of the households. Hence, FARMINCOME is also used as one of the covariates. Finally, to account for the differences in agroclimatic conditions and the state of the local economy across the households, a dummy variable EAST has been added as an explanatory variable in Eq. (2).

4.1 Multinomial logit model

We first consider the participation decision (whether household i participates in an activity type j), as measured by a dummy variable, as the dependent variable. Consequently, a multinomial logit model that accounts for the interdependence between the various employment choices, as introduced by McFadden (1974), has been utilized to assess the factors underlying the diversification choices. Here, O_{ijt} is a polychotomous variable which equals 1, 2, 3 or 4 if the household's major source of income is farm sector employment, non-farm wage employment, non-farm self-employment, or others, respectively.

4.2 Multinomial logit model with sample selection

Since all households do not derive their income from the RNFS, there is a possibility of sample selection bias that might arise while estimating Eq. (2) with a multinomial logit model. Hence, we use Heckman's two-step sample selection model which combats the problem of sample selection bias. The model consists of two equations: (i) A *selection equation* which considers the whole sample and determines the selection process (here, RNFS participation) and (ii) A *response equation* which considers only those households which derive their income from the RNFS and determines the factors affecting the outcome variable—the choice between the various modes of RNFS diversification.

The selection equation is estimated using a probit model, where the explanatory variables are the same as that in Eq. (2). To properly identify the model, at least one explanatory variable from the first stage selection equation should not be included in the second-stage response equation (Maddala 1986). A requirement for the exclusion restriction is that it must directly impact the decision to participate in the RNFS, but not the outcome of interest (here, the choice between the various RNFS sub-sectors). The response equation is as estimated in Sect. 4.1, where FARMINCOME, NONLANDASSET, and HHSIZE are excluded, and inverse mills ratio (from the selection equation) is used as an additional explanatory variable. These three variables are used as the exclusion restrictions and are used to estimate the households' participation in the RNFS by employing a probit model. Also, O_{ijt} is a trichotomous variable which equals 1, 2, or 3 if the household's major source of income is non-farm wage employment, non-farm self-employment, or others, respectively.

4.3 Fractional multinomial logit model with sample selection

Nevertheless, a multinomial logit model considers participation in an RNFS activity as a polychotomous variable and does not capture the differences in the intensity of participation in a particular category of the RNFS across the households. To overcome this limitation, we employ a fractional multinomial logit model (an extension of the fractional logit model by Papke and Wooldridge 1996) where the dependent variable is a continuous variable which represents the vector of the share of income, derived from the various sub-sectors of the RNFS, to measure RNFS participation. The model, thus, measures the variation in the share of income from an RNFS sub-sector because of changes in the explanatory variables. Therefore, the model not only captures the intensity of RNFS participation, but also better informs about the impact of push factors, such as farm income and land asset, which otherwise remain confounded in the models used in Sects. 4.1 and 4.2. As in Sect. 4.2, we again account for the possibility of sample selection bias and, therefore, Heckman's two-step sample selection method is used. Equation (2) represents the response equation, where O_{ijt} represents ith household's vector of the share of income derived from non-farm wage employment ($j=1$), non-farm self-employment ($j=2$), and others ($j=3$) at time t, where $0 \leq O_{ijt} \leq 1$ and $\sum_{j=1}^{3} O_{ijt} = 1$. As in Sect. 4.2, the household's decision to participate in RNFS is estimated using a probit model with FARMINCOME, NONLANDASSET, and HHSIZE as the identifying restrictions.

Table 7 provides an overview of the definitions of the explanatory variables used in the regression and Table 8 provides the summary statistics of these variables.

5 Estimation results and discussion

This section presents the empirical results of the various econometric methods used to estimate the homogeneity of the entry barriers across various RNFS sub-categories.

Table 9 presents the estimated results of the multinomial logit model (farm employment is the base category). The human capital variable, EDU, has a positive and significant impact only on non-farm wage employment and non-farm self-employment. This concurs with the hypothesis that education acts as an entry barrier

Table 7 Description of the explanatory variables used in the estimation

Variables	Definition
EDU	Number of years of education of the household head
TEDU	Technical education of the household members measured by the proportion of members in a household having technical education
CREDIT	Log of total credit (in Rupees), where credit from both formal and informal sources is accounted for
SOCIAL	Social capital measured by proportion of members registered with various forms of associations
LANDASSET	Land owned by the household (in ha)
LANDSQ	Square of the land owned by the household (in ha^2)
NONLANDASSET	Total value of livestock, farm implements, stock inventory, and durables (in Rupees)
FARMINCOME[a]	Log of one plus gross farm income (in Rupees)
GENDER	Gender of the household head measured by a dummy variable (GENDER = 1, if male; 0 else)
AGE	Age of the household head (in years)
HHSIZE	Household size measured by number of members present in a household
DISTANCE	Distance from the market/nearest town (in km)
ELEC	Availability of electricity in a household measured by a dummy variable (ELEC = 1, if available; 0 else)
EAST	Dummy Variable indicating whether a household belongs to an Eastern state (Bihar, Jharkhand or Odisha) (EAST = 1, if yes; 0 else)

[a]Since the decisions on farm and RNFS are made simultaneously by the households, the variable farm income grapples with the issue of endogeneity. To overcome the issue, predicted values of farm income in place of its observed values have been used, which are estimated using a Cobb–Douglas production function. For the empirical estimation, labour (family and hired), land, material inputs (fertilizers, weed, irrigation etc.), other inputs (machinery, marketing and transport), soil depth, and an year dummy were used as the explanatory variables. All the variables are found to have a significant and positive impact on farm income, except the other inputs used. Apart from farm income, the variable land asset could also be endogenous since with RNFS income, the size of land asset may go up or down depending on what the households do with the extra RNFS income. However, we could not find a suitable instrument to be used for land asset from our data set and, therefore, could not account for endogeneity arising from this variable. In the existing literature also, it is only farm income that is considered as endogenous

and those with better education are thought to perceive, interpret, and respond to information on jobs faster than their lesser educated counterparts (Huffman 1980; Barrett et al. 2001). Also, the positive relationship between EDU and non-farm self-employment converges with the fact that better-educated households are more likely to establish their own businesses (Corral and Reardon 2001). TEDU is found to have a positive and significant impact on participation in non-farm wage employment and others. The positive association between TEDU and non-farm wage employment reflects the entry barriers which hinder the households from taking up wage employment in the RNFS as compared to jobs in the farm sector. Also, the households with more technical education are more likely to go to other RNFS jobs as compared to farm employment. At first, this might seem contrived, because the other category constitutes of unskilled labour. However, with a rise in TEDU, the households are

Table 8 Summary statistics of the explanatory variables used in the estimation

Variables	Mean	SD	Min	Max
EDU	5.15	4.75	0	19
TEDU	0.07	0.14	0	1
CREDIT	10.17	1.59	4.32	15.57
SOCIAL	0.1	0.19	0	1
LANDASSET	2.05	2.75	0	35.61
LANDSQ	11.76	48.68	0	1268.24
NONLANDASSET	251,293	571,658.7	70	1.94e+07
FARMINCOME	10.78	1.66	0	15.4
GENDER	0.92	0.26	0	1
AGE	49.57	12.81	16	90
HHSIZE	5.31	2.58	1	26
DISTANCE	10.95	6	0.5	52
ELEC	0.82	0.38	0	1
EAST	0.36	0.48	0	1

Table 9 Multinomial logit estimates

	Non-farm wage employments v/s farm employment	Non-farm self-employment v/s farm employment	Others v/s farm employment
	Coeff (S.E.)	Coeff (S.E.)	Coeff (S.E.)
EDU	0.11*** (0.02)	0.04** (0.02)	−0.03 (0.02)
TEDU	3.58*** (0.51)	0.43 (0.55)	1.12* (0.65)
CREDIT	0.32*** (0.05)	0.24*** (0.05)	0.17*** (0.05)
SOCIAL	0.66* (0.34)	1.14*** (0.31)	0.14 (0.39)
LANDASSET	−0.25*** (0.07)	−0.05 (0.06)	−0.22*** (0.07)
LANDSQ	0.001 (0.01)	−0.002 (0.003)	0.01*** (0.002)
NONLANDASSET	4.85E−07*** (1.77E−07)	7.27E−07*** (1.67E−07)	3.29E−07 (2.37E−07)
FARMINCOME	−1.94*** (0.09)	−1.99*** (0.09)	−2.11*** (0.09)
GENDER	−1.01*** (0.23)	0.37 (0.26)	0.26 (0.25)
AGE	0.03*** (0.01)	0.01** (0.01)	−0.01** (0.01)
HHSIZE	0.25*** (0.03)	0.27*** (0.03)	0.24*** (0.03)
DISTANCE	0.01 (0.01)	0.04*** (0.01)	0.04*** (0.01)
ELEC	0.58** (0.23)	0.85*** (0.23)	0.14 (0.2)
EAST	0.02 (0.19)	0.11 (0.18)	0.58*** (0.19)
Year indicators (2010 = base year)			
2011	0.27 (0.21)	0.40* (0.2)	−0.003 (0.2)
2012	0.34 (0.22)	0.64*** (0.2)	−0.06 (0.22)
2013	0.89*** (0.22)	1.01*** (0.21)	0.59*** (0.22)
2014	0.92*** (0.22)	0.95*** (0.21)	0.73*** (0.22)
Log-likelihood: −2494.95			
No. of observations: 3102			

Standard errors (S.E.) in the parentheses. ***, **, and * denote statistical significance at 1, 5, and 10%, respectively more likely to participate actively in the banking sector and financial markets, which could boost their savings and interests on deposits. Our results are similar to the study by Jatav and Sen (2013) in the Indian context, which emphasizes that entry into the RNFS seems to be restrictive for those without educational achievements and technical education. Also, our findings converge with those of Lanjouw and Shariff (2004) and Lanjouw and Murgai (2009) who find that remunerative RNFS activities and regular non-farm employment in India are largely associated with the households' education levels.

Regarding the access to credit, with more credit, the households are more likely to diversify into non-farm wage employment, non-farm self-employment, and others as compared to farm employment. This is because the access to credit helps the households in solving the liquidity problems associated with the RNFS in two ways. First, the cash obtained from the credit can be used directly by the households to start a new enterprise. Second, the cash can be used to buy agricultural inputs and machinery which would enhance farm productivity. The income from the farm sector, then, can be invested into the RNFS (Demie and Zeray 2016). These findings are also consistent with the literature (Escobal 2001; Senadza 2012; Akaakohol and Aye 2014).

It is hypothesised that with the rise in the social capital of the households, the financial constraints faced by them get minimized, making them more likely to diversify into the RNFS. Moreover, with increased social capital, the entrepreneurial skills, and bargaining power of the household members also get a boost, which makes the trading of goods an easier task (Kassie et al. 2017). Our study unearths a positive relationship between social capital and all the categories of RNFS employment, except for the others category. The result is similar to that of Schwarze and Zeller (2005), Shehu and Abubakar (2015), and Kassie et al. (2017).

As noted by Woldenhanna and Oskam (2001), households with a smaller farm size diversify into the RNFS to supplement their farm income and cope with poverty. For these households, the small farm size acts as a push factor for diversification (survival-led diversification). The households with large landholdings tend to specialize on the farm and, hence, they are not compelled to diversify. This implies a negative relationship between land size and diversification. Our study shows a negative relationship between land size and non-farm wage employment. This concurs with the results of Corral and Reardon (2001), Escobal (2001), and Demie and Zeray (2016). However, the results indicate a 'U'-shaped relationship between land size and others. This accords with the discussion in Section three, which asserts that a negative relation between land size and diversification is not always true. We argue that a positive relation between land size and diversification for households with large landholding is plausible, implying an opportunity-led diversification.

The non-land asset, depicting the wealth of a household, is said to have an ambiguous effect on diversification (Reardon et al. 1992). On one hand, the more non-land asset a household possesses, the less risk averse it is and, hence, more willing to undertake investments in the RNFS. On the other hand, the portfolio theory asserts that the households with a few non-land assets are more risk averse in nature and tend to diversify more to lower the overall instability of their returns. Our results signify the importance of non-land asset in determining the households' participation in non-farm wage employment and non-farm self-employment. Interestingly, non-land asset does not have a significant impact on the participation in others.

It is hypothesised that with the rise in the level of farm income, the likelihood to participate in the RNFS diminishes because of the presence of ample subsistence income and support for the family. However, the farm income also enables the households to overcome the liquidity and credit constraints of RNFS activities, making them more likely to diversify (Woldenhanna and Oskam 2001). Our study unveils a negative and significant impact of FARMINCOME on all the categories of RNFS employment. The findings concur with that of Woldenhanna and Oskam (2001).

It is generally argued that the male-headed households are more likely to diversify as compared to their female counterparts (Dercon and Krishnan 1996; Sendaza 2012; Shehu and Abubakar 2015). However, our results indicate that female-headed households gravitate towards non-farm wage employment as compared to farm employment. This is in line with the discussion in Section three, where it is argued that females are more risk averse than males and are more likely to diversify. Moreover, the farm work (sowing, ploughing, and harvesting) is laborious in nature and is meant for males. Thus, even in the presence of land, the female heads hire male-workers for the farm work, while, in the meantime, they engage themselves in RNFS employment (Teshome and Edriss 2013). This concurs with the results of Cangarajah et al. (2001), Sendaza (2012), Shehu and Abubakar (2015) and Asfaw et al. (2017). Older household heads tend towards non-farm self-employment. This is because of the personal capital and experience gained by them with their age, making the household head more likely to invest in new enterprises (Ghimire et al. 2014). Also, AGE is found to have a positive impact on non-farm wage employment. This is because non-farm wage employment requires information flow that comes from informal social networks. These networks increase with the age of the household head, making them more likely to diversify (Abraham 2011). These results accord with that of Barrett et al. (2001) and Dimova and Sen (2010). On the contrary, the aged household heads are less likely to go for other jobs under the RNFS as compared to farm employment. This result is akin to the studies by Teshome and Edriss (2013), Ghimire et al. (2014), Shehu and Abubakar (2015), and Kassie et al. (2017), where the impact of age on diversification is found to be negative. The impact of household size on all the three types of RNFS employment is found to be positive and significant. This is in line with the hypothesis that households with a large family size tend to have a larger labour force which cannot be accommodated in the farm sector, and this acts as a push factor that makes the households more likely to diversify into the RNFS (Woldenhanna and Oskam 2001). Moreover, households with a larger family size tend to incur higher expenditures, making them more likely to distribute their work between the RNFS and farm work (Reardon et al. 1992). Our findings are consistent with that of Woldenhanna and Oskam (2001), Dimova and Sen (2010), Teshome and Edriss (2013), and Shehu and Abubakar (2015).

Reardon and Taylor (1996) emphasize the importance of access to markets (measured by DISTANCE here) and the general infrastructure (measured by ELEC here) as important variables in rural diversification. Interestingly, the impact of

DISTANCE is positive and significant on non-farm self-employment and others. This accords with the hypothesis that more the distance from the market, more is the difficulty in marketing and trading agricultural goods. This makes the households tend towards self-employment and other unskilled jobs in the rural areas, as compared to farm employment. Our result contradicts with the literature which finds a negative impact of the distance from the nearest market on RNFS diversification (Canagarajah et al. 2001; Escobal 2001; Schwarze and Zeller 2005; Teshome and Edriss 2013; Demie and Zeray 2016). ELEC has a positive and significant impact on all the categories of RNFS employment, except for others. This is analogous to the findings by Corral and Reardon (2001), Escobal (2001), Senadza (2012), and Shehu and Abubakar (2015).

It is evident from Sect. 3 that the Eastern states are more likely to engage in the other category as compared to the SAT states. Our result is supported by the fact (Table 6) that the EAST households derive 22% of their income from the others category as compared to the SAT households, which derive only 5% of their total income from other non-farm sources. This might be the result of various reasons. First, on an average, the EAST households earn only 55,225 Rupees from the farm sector, which is relatively lesser than the average farm sector earnings of the SAT households (Rs. 2,29,458). Second, the average landholding size in the EAST regions is only 1.17 ha, which is again lesser than that of the SAT households (2.56 ha). Finally, the average number of members that the EAST households have is six. Whereas, the average household size of the SAT households is only five members. The low farm income, lesser landholdings, and a larger household size might be pushing the EAST Households to diversify more, particularly into other jobs in the RNFS. Year dummies have been included in the regression to account for the year-specific effects. Compared to 2010, the households are found to be more likely to engage in non-farm wage employment in the years 2013 and 2014. Also, it is found that the likelihood of getting employed in the non-farm self-employment category is more in 2011–2014 as compared to 2010. Finally, as compared to 2010, RNFS diversification into the other category is more in the years 2013 and 2014.

Table 10 indicates the estimation results of the multinomial logit model, accounting for the possibility of sample selection bias that might occur while participating in the RNFS.

The first column shows the results of the selection equation, which estimates the determinants of RNFS participation. It is evident that the households with a larger proportion of members with technical education and those receiving a higher amount of credit are more likely to participate in the RNFS. Also, social capital and non-land asset are found to assert a positive impact on the likelihood of participation in the RNFS. As expected, farm income negatively influences RNFS participation. Female-headed households are more likely to participate in the RNFS. The households with aged heads and a larger number of household members are also more likely to go for RNFS participation. Distance from the market asserts a negative influence on RNFS participation. Households belonging to the EAST regions are found to be more likely to participate in the RNFS. Finally, more RNFS participation is observed for the years 2013 and 2014, as compared to the base year 2010.

Table 10 Multinomial logit estimates with sample selection

	Stage1: selection equation	Stage2: response equation	
	Probit estimates of RNFS participation	Non-farm wage employment v/s others	Non-farm self-employment v/s others
	Coeff (S.E.)	Coeff (S.E.)	Coeff (S.E.)
EDU	−0.001 (0.01)	0.11*** (0.01)	0.08*** (0.01)
TEDU	1.06*** (0.29)	2.46*** (0.43)	−0.44 (0.44)
CREDIT	0.07** (0.03)	0.24*** (0.04)	0.18*** (0.04)
SOCIAL	1.27*** (0.21)	0.9*** (0.27)	0.84*** (0.26)
LANDASSET	−0.02 (0.01)	−0.15*** (0.03)	0.03 (0.02)
NONLANDASSET	3.09E−07*** (1.11E−07)	–	–
FARMINCOME	−0.72*** (0.03)	–	–
GENDER	−0.78***(0.26)	−1.08*** (0.19)	−0.21 (0.2)
AGE	0.01** (0.003)	0.03*** (0.004)	0.02*** (0.004)
HHSIZE	0.1*** (0.02)	–	–
DISTANCE	−0.01** (0.01)	−0.01 (0.01)	−0.01 (0.01)
ELEC	0.08 (0.13)	0.4** (0.17)	0.53*** (0.16)
EAST	0.42*** (0.13)	−0.28* (0.15)	−0.22* (0.13)
Inverse mills ratio		0.79*** (0.27)	0.56** (0.24)
Year indicators (2010 = base year)			
2011	0.15 (0.1)	0.4** (0.16)	0.4*** (0.14)
2012	0.14 (0.11)	0.54*** (0.17)	0.65*** (0.15)
2013	0.31*** (0.12)	0.71*** (0.17)	0.64*** (0.16)
2014	0.37*** (0.13)	0.81*** (0.17)	0.68*** (0.16)
	Pseudo R^2: 0.37 No. of observations: 3102	Log-likelihood: −2661.04 No. of observations: 2662	

Standard errors (S.E.) in the parentheses. ***, ** and * denote statistical significance at 1, 5and 10%, respectively

Columns two and three present the results of the response equation estimated by a multinomial logit model (employment in others is the base category). As discussed earlier, FARMINCOME, NONLANDASSET and HHSIZE are used as the exclusion restrictions. All the three identifying restrictions are found to be significant at the 1% level. The Wald test rejected the null hypothesis that the coefficients of FARMINCOME, NONLANDASSET, and HHSIZE are simultaneously equal to zero, indicating that the exclusion restrictions are valid. Also, in the stage two-response equation, the inverse mills ratio is found to be significant at the 1 and 5% level for non-farm wage employment and non-farm self-employment, respectively, signifying that sample selection is a potential problem of the sample. As expected, EDU has a positive and significant impact on the likelihood of engaging in non-farm wage employment and non-farm self-employment, as compared to others. This is because better-educated household heads have better access to any RNFS employment type and are more likely to establish their own enterprises/businesses (Corral and Reardon 2001). As anticipated, TEDU asserts a positive influence on employment in non-farm wage employment, relative to others. However, no impact of TEDU is observed on RNFS self-employment. This is in line with the hypothesis that RNFS employment, especially in the formal sector (here, salaried jobs), requires higher skills and knowledge (Abdul-Hakim and Hadijah Che-Mat 2011). As expected, CREDIT received by the households affects non-farm self-employment positively when compared to others. This is because unskilled labour (a component of others) requires less/no amount of credit. Also, access to credit enables the households to re-allocate their physical stock of capital in the short run (and solve their liquidity problems) to take opportunities to diversify outside the farm sector (Demie and Zeray 2016). SOCIAL influences the likelihood of participation in non-farm wage employment and non-farm self-employment in comparison to others. This is because membership in co-operatives and other organizations helps the households to minimize their financial constraints, enhance their information flow on job opportunities outside the RNFS, and boost their entrepreneurial skills (Kassie et al. 2017). More the LANDASSET owned by the households, less likely are they to go for non-farm wage employment as compared to others. This might be plausible, because non-farm wage employment does not require land. Female-headed households are more likely to go for non-farm wage employment as compared to their male counterparts. This, again, concurs with the literature which asserts that female-headed households are risk averse in nature and are more likely to diversify. The result is also in line with the discussion in Section three, where female-headed households earn almost 21% of their income from non-farm wage employment, as compared to the corresponding share of the male-headed households which is only 13%. Aged household heads are more likely to gravitate towards both non-farm wage employment and non-farm self-employment, relative to others. This is possibly because aged heads have a larger household size, ergo larger unemployed labour, compelling them to diversify into the RNFS (Barrett et al. 2001). Also, as mentioned previously, aged heads tend to get information (from informal social networks) on job opportunities outside the farm sector, making them more likely to diversify (Abraham 2011). The households with access to electricity are more likely to go for non-farm self-employment. This might be plausible, since self-employment activities require more electricity than others.

ELEC also had a positive impact on non-farm wage employment. It is found that the households belonging to the EAST region are less likely to go for both non-farm wage employment and non-farm self-employment. This concurs with the fact that the EAST regions, on an average, receive lesser credit (52,756 Rupees) than the SAT regions where the average credit received by the households is around 88,601 Rupees. The low availability of credit makes the EAST regions less likely to go for self-employment. Also, the average education (in years) of the household members is 4.96 (5.46) for the EAST (SAT) households making them less likely to go for both non-farm self-employment and non-farm wage employment relative to others. Finally, the year dummies indicate a higher amount of diversification in non-farm wage employment as well as non-farm self-employment for the years 2011–2014 when compared to the year 2010.

Table 11 presents the estimated results of the multinomial fractional logit model accounting for the intensity of RNFS participation as well as the possibility of sample selection bias that might occur while participating in the RNFS. The first column shows the results of the selection equation which estimates the determinants of participation in the RNFS. Columns two and three present the results of the response equation estimated by a multinomial fractional logit model (employment in others is the base category). As discussed earlier, FARMINCOME, NONLANDASSET, and HHSIZE are considered as the exclusion restrictions. The results of this model converge with the results of the multinomial logit model with sample selection (Table 10). In the stage two-response equation, the inverse mills ratio is found to be significant at 1% for both non-farm wage employment and non-farm self-employment, signifying that sample selection is a potential problem of the sample.

6 Conclusion

Amidst the deplorable condition of the farm sector, which was once considered as the cornerstone of the rural areas in developing countries, the households' labour allocation has been shifting away from farm towards the RNFS. India is a case in point. The shift to the RNFS is compelled by various push (small landholdings, uncertain climatic conditions, rising population, inadequate farm output, price fluctuations, etc.) and pull (relatively higher returns in the RNFS) factors. Nevertheless, due to the presence of various entry barriers like access to education, lack of technical education/skills, inability to access credit, unavailability of infrastructural facilities, and lack of access to social capital, among others, the households' participation in the RNFS remains restrictive in nature. These entry barriers, however, vary across the various modes of RNFS employment (Kilby and Liedholm 1986; Unni 1991).

In this regard, our study attempts to test the existence of entry barriers in RNFS employment and investigates whether these barriers are homogenous across the various RNFS sub-sectors in the SAT and the Eastern regions of India. We have selected these regions for our study, since RNFS diversification is of greater importance to the livelihoods of the rural households in these regions as compared to the other regions of India. We add to the existing literature by laying emphasis on classifying RNFS employment into three categories—wage employment, self-employment, and others—to examine the presence of entry barriers in these different modes of RNFS diversification. We use a multinomial logit model as the baseline model to determine the factors driving participation in the various types of non-farm employment, and thereby examine the presence of entry barriers across the several categories of RNFS employment. Furthermore, the Heckman Selection Model is employed to account for the selection bias in our sample, while a multinomial fractional logit model is used to account for the intensity of the RNFS income.

Our empirical results confirm that education, in general, and technical education, in particular, as well as access to credit and endowment of social capital are the main motivating factors behind RNFS employment. Therefore, the absence of education, lack of technical education, inability to access credit, and the lower endowment of social capital may act as entry barriers to RNFS employment. Moreover, we found that these entry barriers are not homogenous across the various RNFS sub-categories. For instance, it is found that education acts as a major determinant for participation in non-farm wage employment and non-farm self-employment, but it does not affect the other categories of jobs under the RNFS. Similarly, low levels of technical education hinder participation only in non-farm wage employment and 'others'. Finally, social capital asserts a positive influence on non-farm wage employment and non-farm self-employment, but it is not found to be as influential for the other categories of RNFS jobs. The results of the regression analysis also indicate that push factors are also at play. Low farm income and large household size push the households into all the three types of diversification. Moreover, the results also indicate that land asset plays an important role as a push factor in this respect. A negative relation between the land size and participation in non-farm wage employment is indicative of survival-led diversification, wherein the lack of land asset pushes the households to diversify into the RNFS for their survival needs. On the other hand, the 'U'-shaped relationship between land size and the others category of RNFS employment is indicative of opportunity-led diversification, wherein the households increase diversification with the increase in their landholdings. We further found that female-headed households are more likely to opt for non-farm wage employment as compared to their male counterparts. Thus, we conclude that although farm households are being pushed out of the farm sector, their entry into the RNFS sub-sectors seems to be obstructed by the presence of various entry barriers.

Our empirical findings suggest that enhancing the access to education, technical education, and credit would facilitate the households to overcome the challenges of the subsistence farm sector and, thereby, would help these households earn a better livelihood by engaging them in various categories of RNFS employment. Policy initiatives like the Sarva Shiksha Abhiyan, which are dedicated to the spread of universal basic education, and providing access to educational loans to finance higher education are welcome. It is observed, by empirical studies, that formal training in

Table 11 Multinomial fractional logit estimates with sample selection

	Stage 1: selection equation	Stage 2: response equation	
	Probit estimates of RNFS participation	Non-farm wage employment v/s others	Non-farm self-employment v/s others
	Coeff (S.E.)	Coeff (S.E.)	Coeff (S.E.)
EDU	−0.001 (0.01)	0.1*** (0.01)	0.08*** (0.01)
TEDU	1.06*** (0.29)	2.46*** (0.36)	−0.38 (0.39)
CREDIT	0.07** (0.03)	0.21*** (0.03)	0.19*** (0.03)
SOCIAL	1.27*** (0.21)	0.84*** (0.23)	0.86*** (0.22)
LANDASSET	−0.02 (0.01)	−0.12*** (0.02)	0.03 (0.02)
NONLANDASSET	3.09E−07*** (1.11E−07)	–	–
FARMINCOME	−0.72*** (0.03)	–	–
GENDER	−0.78 ***(0.26)	−0.95*** (0.16)	−0.23 (0.16)
AGE	0.01** (0.003)	0.03*** (0.004)	0.01*** (0.003)
HHSIZE	0.1*** (0.02)	–	–
DISTANCE	−0.01** (0.01)	−0.004 (0.01)	−0.01 (0.01)
ELEC	0.08 (0.13)	0.42*** (0.15)	0.44*** (0.14)
EAST	0.42*** (0.13)	−0.31** (0.12)	−0.18 (0.12)
inverse mills ratio	–	0.74*** (0.23)	0.55*** (0.21)
Year indicators (2010 = base year)			
2011	0.15 (0.1)	0.29** (0.14)	0.28** (0.12)
2012	0.14 (0.11)	0.57*** (0.14)	0.57*** (0.13)
2013	0.31*** (0.12)	0.71*** (0.15)	0.55*** (0.13)
2014	0.37*** (0.13)	0.75*** (0.15)	0.59*** (0.13)
	Pseudo R^2: 0.37 No. of observations: 3102	Log-likelihood: −2673.50 No. of observations: 2662	

Standard errors (S.E.) in the parentheses. ***, **, and * denote statistical significance at 1, 5, and 10%, respectively vocational and technical education increases the wage in rural informal sectors significantly (Kumar et al. 2019). Therefore, resources should be invested to provide these types of training. Due to insufficient collateral in their possession, most of the poor households in rural areas do not have access to formal credit. In such a situation, to create informal-formal credit linkages, programmes like the SHG-Bank linkage programme should also be undertaken for providing easy access to credit as it enhances the earnings of self-employed businesses (Bairagya et al. 2020).

Apart from the standard entry barriers to RNFS employment, there are behavioral factors like inertia and status quo which could prevent the households to move out of the farm sector, even when the RNFS provides higher returns than the subsistence farm sector. Therefore, further research on RNFS employment should consider these behavioral factors and conduct a primary survey-based study to capture them.

Acknowledgement The authors thank the anonymous reviewer and editor of the journal for constructive and detailed comments on an earlier draft of this manuscript. Also, the authors are highly grateful to the members of International Crops Research Institute for the Semi-arid Tropics (ICRISAT), Patancheru for their immense help provided in extraction of the VDSA database. The contribution provided by Dr. Shalander Kumar (Principal Scientist, ICRISAT) for providing fruitful discussion regarding the subject of the manuscript is highly acknowledged by the authors. The authors would also like to extend their thanks towards Soumitra Pramanik (Senior Scientific Officer, ICRISAT) and Elias Khan (Senior Scientific Officer, ICRISAT) for providing guidance in data extraction.

Funding The authors did not receive any financial assistance from any source.

Availability of data and material The authors do not have permission to share the data used.

Code availability Codes will be made available on request.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Abdul-Hakim, R., & Hadijah Che-Mat, S. (2011). Determinants of farmer's participation in off-farm employment: A case study in Kedah Darul Aman, Malaysia. *Asian Journal of Agriculture and Rural Development, 1*, 27–37.

Abdulai, A., & CroleRees, A. (2001). Determinants of income diversification amongst rural households in Southern Mali. *Food Policy, 26*(4), 437–452.

Abraham, V. (2011). Agrarian distress and rural non-farm sector employment in India. *MPRA paper 35275*, Accessed at https://mpra.ub.uni-muenchen.de/35275.

Akaakohol, M. A., & Aye, G. C. (2014). Diversification and farm household welfare in Makurdi, Benue State, Nigeria. *Development Studies Research An Open Access Journal, 1*(1), 168–175.

Arslan, A., Asfaw, S., Cavatassi, R., Lipper, L., McCarthy, N., Kokwe, M., et al. (2018). Diversification as part of a CSA strategy: The cases of Zambia and Malawi. *Climate smart agriculture* (pp. 527–562). Cham: Springer.

Asfaw, A., Simane, B., Hassen, A., & Bantider, A. (2017). Determinants of non-farm livelihood diversification: Evidence from rainfed-dependent smallholder farmers in northcentral Ethiopia (Woleka sub-basin). *Development Studies Research, 4*(1), 22–36.

Barrett, C. B., Bezuneh, M., & Aboud, A. (2001). Income diversification, poverty traps and policy shocks in Côte d'Ivoire and Kenya. *Food Policy, 26*, 367–384.

Bairagya, I., Bhattacharya, T., & Bhattacharjee, M. (2020). Impact of credit accessibility on the earnings of self-employed businesses in India. *Journal of Asian Economics*. https://doi.org/10.1016/j.asieco.2020.101202.

Bongole, A. J. (2016). Determinants of farm and non-farm activities as sources of income amongst rural households: Evidence from Kahama District in Tanzania. *Journal of Economics and Sustainable Development, 7*,132–139.

Canagarajah, S., Newman, C., & Bhattamishra, R. (2001). Non-farm income, gender, and inequality: Evidence from rural Ghana and Uganda. *Food Policy, 26*(4), 405–420.

Corral, L., & Reardon, T. (2001). Rural nonfarm incomes in Nicaragua. *World Development, 29*(3), 427–442.

Croson and Gneezy. (2009). Gender differences in preferences. *Journal of Economic Literature, 47*(2), 448–474.

Demie, A., & Zeray, N. (2016). Determinants of participation and earnings in the rural nonfarm economy in Eastern Ethiopia. *African Journal of Rural Development (AFJRD), 1*, 61–74.

Dercon, S., & Krishnan, P. (1996). Income portfolios in rural Ethiopia and Tanzania: Choices and constraints. *The Journal of Development Studies, 32*(6), 850–875.

Diatta, D. S. F. E. S. (2017). Shock metrics for resilience analysis: An investigation of the relationship between subjective and objective indicators.

Dimova, R. D., & Sen, K. (2010). Is household income diversification a means of survival or a means of accumulation? Panel data evidence from tanzania. Available at SSRN: https://ssrn.com/abstract=1688433.

Eckel, C. C., & Grossman, P. J. (2008). Men, women and risk aversion: Experimental evidence. *Handbook of Experimental Evidence, 1*(113), 1061–1073.

Ellis, F. (2007). Household strategies and rural livelihood diversification household strategies and rural livelihood diversification. *Journal of Development Studies, 35*(1), 1–38.

Escobal, J. (2001). The determinants of nonfarm income diversification in rural Peru. *World Development, 29*(3), 497–508.

Elbers, C., & Lanjouw, P. (2001). Intersectoral transfer, growth, and inequality in rural Ecuador. *World Development, 29*(3), 481–496.

Ghimire, R., Huang, W. C., & Shrestha, R. B. (2014). Factors affecting nonfarm income diversification among rural farm households in central Nepal. *International Journal of Agricultural Management and Development (IJAMAD), 4*, 123–132.

Goodwin, B. K., & Mishra, A. K. (2004). Farming efficiency and the determinants of multiple job holding by farm operators. *American Journal of Agricultural Economics, 86*(3), 722–729.

Haggblade, S., Hazell, P. B., & Reardon, T. (Eds.). (2007). Transforming the rural nonfarm economy: Opportunities and threats in the developing world. Baltimore, MD: International Food Policy Research Institute (IFPRI), Johns Hopkins University Press.

Hartog, J., Ferrer-i-Carbonell, A., & Jonker, N. (2002). Linking measured risk aversion to individual characteristics. *Kyklos, 55*(1), 3–26.

Huffman, W. E. (1980). Farm and off-farm work decisions: The role of human capital. *The Review of Economics and Statistics*, *62*(1), 14–23.

Himanshu, H., Lanjouw, P., Mukhopadhyay, A., & Murgai, R. (2011). Non-farm diversification and rural poverty decline: a perspective from Indian sample survey and village study data, LSE Research Online Documents on Economics 38371, London School of Economics and Political Science, LSE Library.

Himanshu, H., Lanjouw, P., Murgai, R., & Stern, N. (2013). Nonfarm diversification, poverty, economic mobility, and income inequality: A case study in village India. *Agricultural Economics, 44*(4–5), 461–473.

Kassie, G. W., Kim, S., & Fellizar, F. P., Jr. (2017). Determinant factors of livelihood diversification: Evidence from Ethiopia. *Cogent Social Sciences, 3*(1), 1369490.

Jatav, M., & Sen, S. (2013). Drivers of non-farm employment in rural India: Evidence from the 2009–10 NSSO round. *Economic and Political Weekly*, *48*, 14–21.

Kilic, T., Carletto, C., Miluka, J., & Savastano, S. (2009). Rural nonfarm income and its impact on agriculture: Evidence from Albania. *Agricultural Economics, 40*(2), 139–160.

Kumar, S., Pramanik, S., Mamidanna, S., & Whitbread, A. (2018). Dichotomous outcomes and drivers of drought resilience mechanisms of smallholder farmers in semiarid India (No. 2058-2018-5334). https://doi.org/10.22004/ag.econ.276026.

Kumar, R., Mandava, S., Gopanapalli, V. S., (2019). Vocational training in India: Determinants of participation and effect on wages. *Empirical Research in Vocational Education and Training*, 11(3).
Kilby, P., & Liedholm, C. (1988). The role of non-farm activities in the rural economy. Department of Agricultural Economics and Extension; UZ/MSU Food Security Research in Southern Africa Project.
Lanjouw, J. O., & Lanjouw, P. (2001). The rural non-farm sector: Issues and evidence from developing countries. *Agricultural Economics, 26*(1), 1–23.
Lanjouw, P., & Murgai, R. (2009). Poverty decline, agricultural wages, and nonfarm employment in rural India: 1983–2004. *Agricultural Economics, 40*(2), 243–263.
Lanjouw, P., & Shariff, A. (2004). Rural non-farm employment in India: Access, incomes and poverty impact. *Economic and Political Weekly*, *39*, 4429–4446.
Lien, G., Kumbhakar, S. C., & Hardaker, J. B. (2010). Determinants of off-farm work and its effects on farm performance: The case of Norwegian grain farmers. *Agricultural Economics, 41*(6), 577–586.
Maddala, G. S. (1986). *Limited-dependent and qualitative variables in econometrics* (Vol. 3). Cambridge: Cambridge University Press.
Malek, M. A., & Usami, K. (2009). Determinants of non-farm income diversification in developed villages of Bangladesh. *American Journal of Economics and Business Administration, 1*(2), 141–149.
McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (pp. 105–142). Academic Press.
Olale, E., & Henson, S. (2012). Determinants of income diversification among fishing communities in Western Kenya. *Fisheries Research, 125*, 235–242.
Papke, L. E., & Wooldridge, J. M. (1996). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. *Journal of Applied Econometrics, 11*(6), 619–632.
Reardon, T. (1997). Using evidence of household income diversification to inform study of the rural nonfarm labor market in Africa. *World Development, 25*, 735–747.
Reardon, T. (2000). Challenges in fighting rural poverty in the globalizing economy of Latin America: Focus on institutions, markets and projects. *documento presentado en el Seminario Experiencias exitosas de combate a la pobreza rural: lecciones para una reorientación de las políticas (Valle Nevado, 27 y 28 de enero), Santiago de Chile*.
Reardon, T., & Taylor, J. E. (1996). Agroclimatic shock, income inequality, and poverty: Evidence from Burkina Faso. *World Development, 24*(5), 901–914.
Reardon, T., Delgado, C. L., & Matlon, P. (1992). Determinants and effects of income diversification amongst farm households in Burkina Faso. *Journal of Development Studies, 28*(2), 264–296.
Reardon, T., Stamoulis, K., Balisacan, A., Cruz, M. E., Berdegué, J., & Banks, B. (1998). Rural non-farm income in developing countries. *The State of Food and Agriculture, 1998*, 283–356.
Reddy, D. N., Reddy, A. A., Nagaraj, N., & Bantilan, M. C. S. (2014). Rural non-farm employment and rural transformation in India. Working Paper Series No. 57.
Ruben, R. (2001). Nonfarm employment and poverty alleviation of rural farm households in Honduras. *World Development, 29*(3), 549–560.
Sainath, P. (2018). In India, farmer's face a terrifying crisis. *The New York Times*.
Schwarze, S., & Zeller, M. (2005). Income diversification of rural households in Central Sulawesi, Indonesia. *Quarterly Journal of International Agriculture, 44*(1), 61–74.
Senadza, B. (2012). Non-farm income diversification in rural Ghana: Patterns and determinants. *African Development Review, 24*(3), 233–244.
Shehu, A., & Abubakar, N. (2015). Determinants of participation of farm households in non-farm enterprise activities in rural Nigeria. *International Journal of Economics, Commerce and Management, 3*(6), 57–71.
Singh, N. P., Bantilan, C., & Byjesh, K. (2014). Vulnerability and policy relevance to drought in the semi-arid tropics of Asia—A retrospective analysis. *Weather and Climate Extremes, 3*, 54–61.
The state of food and agriculture by Food and Agriculture Organization of the United States (2016)
Teshome, B., & Edriss, A. K. (2013). Determinants and patterns of income diversiication among smallholder farmers in Akaki district, Ethiopia. *Journal of Research in Economics and International Finance*, *2*(4), 68–78.
Tran Quang, T. (2014). Determinants of nonfarm participation among ethnic minorities in the Northwest Mountains, Vietnam. *MPRA paper 59158*, accessed at https://mpra.ub.uni-muenchen.de/59158/.
Unni, J. (1991). Regional variations in rural non-agricultural employment: An exploratory analysis. *Economic and Political Weekly*, *26*, 109–122.

VDSA. (2013). Village Dynamics in South Asia (VDSA) database, generated by ICRISAT/IRRI/NCAP in partnership with national institutes in India and Bangladesh. https://vdsa.icrisat.ac.in.

Weldegebriel, Z. B., Folloni, G., & Prowse, M. (2015). The determinants of non-farm income diversification in rural Ethiopia. *Journal of Poverty Alleviation and International Development*, *6*(1).

Woldenhanna, T., & Oskam, A. (2001). Income diversification and entry barriers: Evidence from the Tigray region of northern Ethiopia. *Food Policy, 26*(4), 351–365.