WP-2018-023

Looking Beyond the Farm and Household: Determinants of On-farm Diversification in India

Varun Kumar Das

Indira Gandhi Institute of Development Research, Mumbai
November 2018

Looking Beyond the Farm and Household: Determinants of On-farm Diversification in India

Varun Kumar Das

Email(corresponding author): varun@igidr.ac.in

Abstract

This paper analyses the factors affecting on-farm diversification decision. Notwithstanding the influence of farm and household conditions, studies have also highlighted the role of external pull factors on farm diversification. Though appraised in scholarship, this aspect has eluded an empirical scrutiny in literature. Taking India as a case, this study shows that apart from farm and household factors, there is a broader agro-ecological and structural feature which impacts on-farm diversification decision. Correcting for endogeneity in a seemingly unrelated system of ordered probit models, a "three-stage" residual inclusion model is estimated. The findings show that proximity to social infrastructure such as schools, colleges, and access to public transport matters for diversification. Results also show that though urbanization may increase demand for variety of products, it might as well impinge upon farm labor supply as non-farm opportunities also rises with urbanization. Thus, the underlying structure of the local economy also merits attention while understanding on-farm diversification process.

Keywords: Farm diversification; Pull factors; External conditions; Residual inclusion; Economies of density; Agro-ecology; Urbanization; Structural transformation

JEL Code: Q12; Q18; R23; R53

Acknowledgements:

My profound gratitude to Dr. A. Ganesh-Kumar remains.

Looking Beyond the Farm and Household: Determinants of On-farm Diversification in India

Varun Kumar Das*

Abstract

This paper analyses the factors affecting on-farm diversification decision. Notwithstanding the influence of farm and household conditions, studies have also highlighted the role of external pull factors on farm diversification. Though appraised in scholarship, this aspect has eluded an empirical scrutiny in literature. Taking India as a case, this study shows that apart from farm and household factors, there is a broader agro-ecological and structural feature which impacts on-farm diversification decision. Correcting for endogeneity in a seemingly unrelated system of ordered probit models, a "three-stage" residual inclusion model is estimated. The findings show that proximity to social infrastructure such as schools, colleges, and access to public transport matters for diversification. Results also show that though urbanization may increase demand for variety of products, it might as well impinge upon farm labor supply as non-farm opportunities also rises with urbanization. Thus, the underlying structure of the local economy also merits attention while understanding on-farm diversification process.

Keywords: Farm diversification; Pull factors; External conditions; Residual inclusion; Economies of density; Agro-ecology; Urbanization; Structural transformation
JEL classification: Q12; Q18; R23; R53

*Indira Gandhi Institute of Development Research, Mumbai. Corresponding email: varun@igidr.ac.in
My profound gratitude to Dr. A. Ganesh-Kumar remains.

1 Introduction

In a developing economy, it is usual for farmers to be engaged in more than one farm activity. Such on-farm diversification involves farmers engaging in different types of crops, animal husbandry activities or both. On-farm diversification often provides buoyancy for farm households to tide over uncertain production and consumption environment. Under the challenges of climate change and price volatility, adaptation of crop diversification and mixed farming practices help in preserving the portfolio of food production system (Auffhammer and Carleton, 2018; FAO, 2018). It also helps in ensuring food self-sufficiency and overall food security in an economy (Pingali and Rosegrant, 1995; Joshi et al., 2007). Apart from mitigating risk of farmers, at a micro level, on-farm diversification contributes towards alleviating poverty (Ellis, 2000; Meert et al., 2005; Birthal et al., 2006; Weinberger and Lumpkin, 2007; Chhatre et al., 2016; Birthal et al., 2015b; Michler and Josephson, 2017; Thapa et al., 2018).

However, there are often many structural impediments which may restrict the diversification potential of a farm household (Ellis, 1998, 2000; Rigg, 2006). Lack of proper rural infrastructure impinges upon both production and marketing avenues for agricultural producers. Add to this, the vagaries of nature and price uncertainties, farming entrepreneurship becomes even more constrained. Consequent to these reasons, sustainability and viability of a farming life itself becomes a concern (Tilman et al., 2002; Lee, 2005; Birthal et al., 2015a; Agarwal and Agrawal, 2017). As such, farm diversification becomes more of a necessity than a choice.

The drivers of on-farm diversification may be broadly categorized into two, viz. “push” and “pull” factors (Barrett et al., 2001; Reardon et al., 2007; Xiaoping et al., 2007; Haggblade et al., 2010; Nakajima et al., 2018). Push factors include farmer’s resource availability in the farm and household to undertake farm diversification. It could also imply strategies to hedge against both production and market risks endemic to individual agricultural commodities. Pull factors on the other hand, includes access to better infrastructure, urbanization, changing preferences for different types of high valued agricultural products, etc. Another important pull factor in a developing economy could be the underlying structural transformation in the economy. Structural changes results in labor moving from the farm to the non-farm sector. This could in turn affect farm diversification through the labor market channel.

Literature so far has primarily emphasized on push factors. Such push factors are basically farm and household level characteristics which impact farm diversification process. These includes access to modern technologies, irrigation facility, access to credit, family size, income level, social group, education, etc. (Joshi et al., 2004, 2007; Birthal et al., 2013, 2015b). However, confining diversification decision only to individual farm and household characteristics leaves out several other important pull factors which are external to farm and household environment. It is important to understand the larger operating environment under which farming decisions are made. Such pull factors which are beyond the farm and household condition could have a significant bearing on farm diversification process.

While the effect of such pull factors on agricultural growth has attracted considerable attention in the literature, however its parallel understanding in case of on-farm diversification is still due. Although hinted in the literature, studies have not been able to bring the exegesis of pull factors into the discussion on on-farm diversification. Such pull factors could be the local structure of the economy, access to urban centers, public transport facility, non-farm employment, etc., which could be critical for diversification process. A reason as to why these features have eluded the literature so far may be because most of such studies usually relied upon survey data which are limited to only farm and household level information.

The above discussion has highlighted the necessity of an integrated analysis of on-farm diversification, taking into account the "structural features" of the economy that go beyond the farm and household situation. In this stride, taking India as a case, this paper attempts to explain on-farm diversification process by coalescing two sets of data. One comprising of farm and household level push factors, and the other, comprising of external structural pull factors which may influence on-farm diversification. Farm and household level push factors include farm size, farm inputs, irrigation, household size, age, education, social group, etc. On the other hand, external pull factors consists of broad agro-ecological and structural features under which a farm operates. These are village neighborhood network, social infrastructure, access to public transport, urbanization and structural transformation. All these external pull factors could impinge upon a farm household's on-farm diversification decision process.

In this study on-farm diversification is measured using a count method. However, in the literature so far, crop diversification has been measured using index method such as Simpsons index, Herfindahl index, etc. But, as will be explained later, on-farm diversification which includes both crop and animal husbandry activities, is difficult to measure using an index method.

To identify the factors influencing on-farm diversification decision, an ordered probit model is estimated. As will be discussed later, there are variations in on-farm diversification over the two agricultural seasons in India – *kharif* and *rabi*. The weather conditions in the two seasons vary quite dramatically across the country. Especially rainfall, temperature, irrigation availability, entomological impacts, etc. are different in the two seasons. Due to these varying ecological conditions, and other plausible reasons pertaining to labor market conditions, both the extent and drivers of diversification could differ across these two seasons. Hence, in this study ordered probit models are estimated separately for the two seasons. With most of the external structural variables available only on an annual basis, the error terms over the two seasons to be correlated. Therefore, the model is estimated under a seemingly unrelated regression (SUR) framework. As suggested by Terza (2017) and Wooldridge (2015), potential endogeneity is corrected in a non-linear set-up using a two stage residual inclusion (2SRI) control function (CF) approach. Since, a 2SRI approach is applied in a SUR set-up, it eventually models a "three stage" residual inclusion (3SRI) estimation.

The rest of the discussion is structured as follows: the next section 2 discusses the external structural conditions which may be relevant for on-farm diversification process. A discussion on measuring on-farm diversification and data sources is presented in section 3. The empirical strategy applied in this study is mentioned in section 4. Section 5 presents the results and discusses the policy implications of this study. Section 6 concludes the paper.

2 Factors beyond the farm and household

The external pull factors affecting on-farm diversification could be at different levels. Neighborhood and local economy factors at the village level could influence farming decisions. However, villages across a particular district share common agro-ecological and infrastructure conditions. Thus, factors at the district level could also impinge upon on-farm diversification. Agricultural policies are usually formed at the state level. Farm policies of a state could also affect diversification decision. These broad external factors are discussed below.

An important way through which farm decisions may be affected is through farm labor availability. Labor markets are usually not well integrated in India, with farm households dependent on local village labor. Another way farm diversification decision might be affected is through prices. Though agricultural prices ought to equalize owing to law of one price, it may not be the case as both rural input and product markets are not integrated in India. Rural households function under village level market conditions and hence farming decisions may be affected by village level relative prices of agricultural produce. Even ranking of various farm products in terms of their relative profitability can differ from village to village as markets are not integrated.

Another important way, through which farm households may be affected, is through village neighborhood network. Holmes and Lee (2012) point out that, farmers often benefit from growing the same crop as their neighbors. They benefit from what they call "economies of density". That is, economies of scale achieved when farmers grow the same crop as their neighboring fields. Robalino and Pfaff (2012) discuss about neighborhood interaction in causing deforestation in Costa Rica. Similarly, Richards (2018) argues that neighborhood influences household land use behavior in Brazil. In case of India, Birthal et al. (2015b) note that there is a village neighborhood effect on adoption of HVC (high valued crops). Nonetheless, even if these village level neighborhood effects seems to be theoretically convincing, it has not been empirically tested. This study makes an attempt in going beyond the household conditions by taking into account village level factors in understanding on-farm diversification decisions.

Studies so far have also highlighted the contribution of physical infrastructure such as roads and transportation, in affecting agricultural growth and rural development (Fan et al., 2008; Motamed et al., 2014; Asher and Novosad, 2016; Aggarwal, 2018). Apart from bringing the input and output markets closer to the farm, access to roads and transport also raises the relative profitability of output. However, one has to note that even though provisioning of such infrastructure brings down the associated opportunity cost of diversification, it might as well affect farm labor availability and hence farm diversification process. As such, road and transport facilities might affect diversification either ways. This aspect remains to be understood in the context of diversification decision literature.

Access to educational facilities is well understood to positively impact agricultural growth and productivity (Fan et al., 2000; Chand et al., 2012). However, even though education is socially desirable, it can impact on-farm diversification through labor supply effect, especially family labor when members of the farming household pursue education. For example, high value crops (HVC) and animal husbandry, demand by and large more labor than other farming activities. As a result, farm households may face labor constraints. In such a situation, to have a positive impact of educational facilities on diversification, it is essential that viable labor saving technologies penetrate into agricultural production. Therefore, it is important to understand how provisioning of educational facilities affect farm diversification. This could be done only if local information on educational facilities is taken into account while analyzing farm diversification decision.

Palmer-Jones and Sen (2003) and Krishna Kumar et al. (2004), note that agro-ecological and environmental characteristics significantly explain differences in agricultural growth. Agricultural land availability, fallow land, source of irrigation, power availability for agriculture, forest, etc. could also determine farm diversification decisions. Apart from reflecting the local structural conditions for agriculture, these factors could also impact the local biodiversity which may have a bearing on on-farm diversification. This paper tries to bring on board these external agro-ecological and structural conditions which till now have not been well understood in the literature on farm diversification.

Rising urbanization and structural transformation in the economy naturally impacts the agricultural sector. Proximity to urban centers influences the production behavior of rural households (Rao et al., 2007; Lange et al., 2013; Chatterjee et al., 2015; Vandercasteelen et al., 2018). Increase in urbanization leads to greater demand for variety and high value products. Though urbanization may provide a bigger market for the farm households, it may also adversely impact farm diversification through lack of adequate farm labor. Urbanization not only implies a larger market, but under the Lewisian dual-sector model, it also provides greater non-farm employment opportunity (Lewis, 1954). As urban centers grow, the accompanying structural transformation may pull out labor from agriculture. It depends on the level of urbanization or the class of urban centers, as to whether urbanization helps diversification through higher demand, or it adversely affects diversification by pulling out labor from the agricultural sector. Thus, structural changes without the adoption of labor saving technologies on the farm, may restrict farm diversification process. Not limited to it, urbanization may even have a varied impact over agricultural seasons. For example, harvesting of orchards and seasonal crops, which happens in a particular season, demands more labor. In such a case, urbanization may make labor availability an impediment for diversification in such seasons. The impact of urbanization and its underlying structural change could only be understood if extra information beyond the household is brought in to the analysis of on-farm diversification process.

States in India also differ by general economic levels. A higher level of economic condition could imply that a larger proportion of population would be engaged in non-farm activities. As a result, supply of labor could be constrained and hence impact farm diversification decision. To counter this, it is important whether state agricultural expenditure favors or constraints diversification. Since public investment in agriculture impacts agricultural growth (Chand and Kumar, 2004; Gulati et al., 2017), it is also equally plausible that state policy factors could also affect agricultural diversification. This study makes an attempt to include such state policy factors in explaining farm diversification decision.

3 Data: Source, measurement, and description

The data for this study has been compiled from various sources. The farm and household data is taken from the latest available all India survey on agricultural households conducted by the National Sample Survey Office. However, this data source does not provide information on district and state factors. Hence, this data had to be complemented with information collected from the other statistical sources. We first discuss about measuring on-farm diversification, followed by discussing all the data sources used in this study.

3.1 On-farm diversification

Available literature so far has ascribed to only crop diversification in its analysis of agricultural diversification. But, it is very common in India for agricultural households to be engaged in mixed farming (Dadhich, 2015; Kishore et al., 2016). To assume on-farm diversification as only diversification among crops leaves many other non-crop farming activities out of its fold. This presents a skewed view of farming activities of an agricultural household. Therefore, this study considers on-farm diversification to consist of both crop and animal husbandry activities.

Farm diversification is usually measured in the literature using Simpsons index, Herfindhal index, etc. Such studies use output or land shares to compute diversification index. However, problem arises when animal husbandry activities is also included in the ambit of on-farm diversification. Obviously, land shares cannot be used to calculate diversification. The only option is then to use value of output shares.

However, prices of different commodities are not the same. Even with constant prices, relative prices may be different. As such, different items could have different values, which may not be commensurate with allocation of resources. For example, consider a farm household involved in mixed farming. One commodity may be of high value (say, spices, poultry etc.) than the other (say, paddy, wheat, etc.), which may be relatively of lower value. But, according to share of input use for the high valued commodity may be less compared to the other. Also, input use and requirements for the two commodities may be completely different. Therefore, it is difficult to use index method either with land share or output to measure on-farm diversification.

In order to account for on-farm diversification, this study applies a count method to measure diversification. A higher count would imply a high degree of diversification and a lower count implies a low level of diversification. In a count method, the number of crops and animal husbandry activities are counted to account for the degree of on-farm diversification by a farm household. Count method brings out the competitive claims on time as a resource input. As such, it allows for the possibility of misallocation of time, resulting in sub-optimal level of output. Nevertheless, it has to be kept in mind that count method assigns equal weights to each activity regardless of its intensity.

Data on on-farm diversification of farm households is taken from the 70^{th} round National Sample Survey (NSS), Situation Assessment Survey (Schedule 33)[1]. This is an all India survey which surveyed 34,907 agricultural households for the period July 2012 to June 2013. The 70^{th} NSS survey is canvassed in two visits. The first Visit 1 is for *kharif* season from July-December 2012, and the second one is Visit 2 canvassed for the *rabi* season January-June 2013. Among the 34,907 agricultural households, a further careful scrutiny leaves only 25,021 households reporting information on agricultural production during the two broad

[1]Though a similar survey was conducted in its 59^{th} round NSS survey (2002-2003), a change in the definition of agricultural household renders the 59^{th} and 70^{th} round incomparable. See Chapter 5, Instruction Manual, 70^{th} round NSSO (2014).

seasons. This is so because, we find that there is incomplete information either for the dependent variable of agricultural production, or for the explanatory variables. Hence, for this reason we have to restrict the analysis to 25,021 households which report the required information in either season.

During Visit 1, around 79.3% of farm households reported to have been engaged in more than one farm activity. Around 48.8% of households are reported to have been cultivating more than one type of crop, and about 29.17% reported to have been engaged in more than one type of animal husbandry activity during this season. In Visit 2, about 72.6% of farm households reported to have been engaged in more than one on-farm activity. 43.4% were engaged in more than one crop and 34.3% were engaged in more than one animal husbandry activity during this season.

The study is restricted to 20 major states for which data on non-farm variables are available. These states are Assam, Andhra Pradesh, Bihar, Chhattisgarh, Gujarat, Haryana, Himachal Pradesh, Jharkhand, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Odisha, Punjab, Rajasthan, Tamil Nadu, Telangana, Uttarakhand, Uttar Pradesh, and West Bengal. Though the study does not cover the Union Territories, Jammu and Kashmir, Arunachal Pradesh, Meghalaya, Mizoram, Manipur, Nagaland and Tripura, it however covers around 85 per cent of total districts in India.

There is variation in climatic and geographic conditions across districts in India. Depending on the availability of rainfall, irrigation, agro-climatic conditions, etc. diversification pattern changes across districts and over the two seasons. Around Figure 1 gives the variation in crop diversification across districts in the two seasons. we see that though most of the districts in central India diversify within 2 to 3 crops in *kharif* (Visit 1), diversification significantly falls during the *rabi* (Visit 2) season. However, for animal husbandry diversification there is not much change in pattern across districts or seasons (Figure 2). It is interesting to note that crop and animal husbandry diversification has a complementary pattern i.e. animal husbandry activities are higher in regions where crop activities are higher.

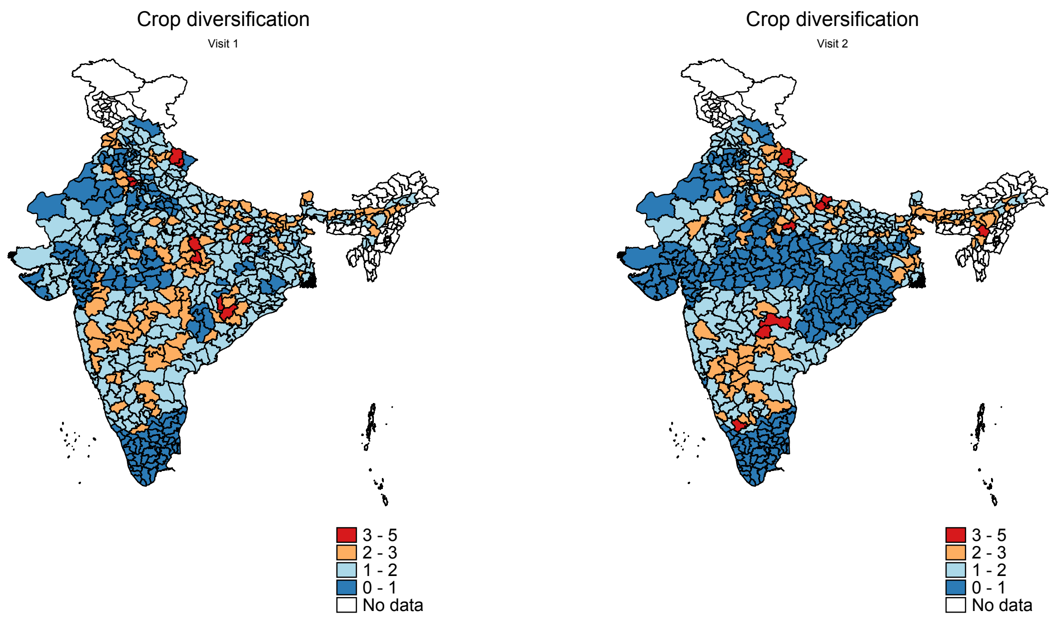


Figure 1: Crop diversification during *kharif* and *rabi* seasons in India
Source: Author's calculation based on NSS 70^{th} round.

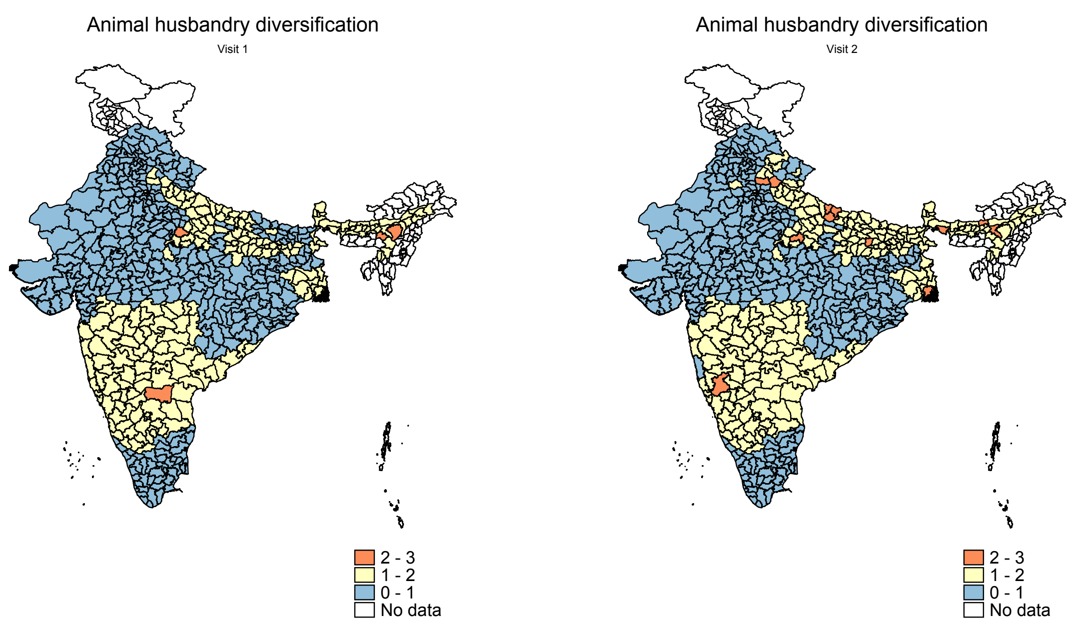
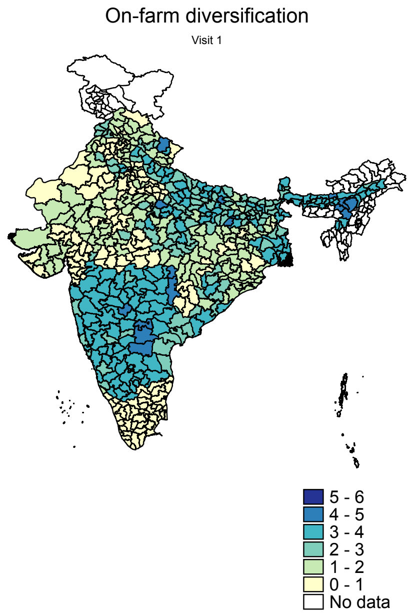


Figure 2: Animal husbandry diversification during *kharif* and *rabi* seasons in India
Source: Author's calculation based on NSS 70^{th} round.



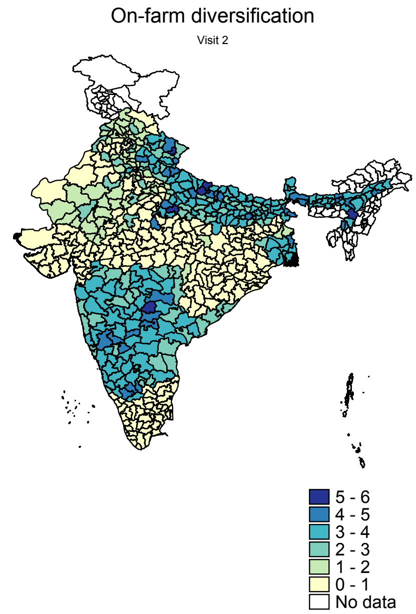


Figure 3: On-farm diversification during *kharif* and *rabi* seasons in India
Source: Author's calculation based on NSS 70^{th} round.

In both Visit 1 and Visit 2, on an average farm households are engaged in around 3 on-farm activities, with a maximum of households being engaged in 9 on-farm activities[2]. Overall on-farm diversification among crops and animal husbandry activities across districts are presented in Figure 3. There is less amount of diversification in Western and Central India. States like Kerala and Tamil Nadu also have less diversification. High levels of on-farm diversification is to be found in most of the districts of Uttar Pradesh, Maharashtra, and Karnataka. In fact, this high degree of on-farm diversification in these three regions could be entirely due to high rate of animal husbandry activities in these districts. There is some amount of fall in diversification during Visit 2. This may be due to shortage of irrigation and rainfall availability during the *rabi* season.

Thus, clearly there is variation in farm diversification across districts and over the two seasons. Hence, analysis is carried out separately for the two seasons.

[2]See Appendix 2.

3.2 Drivers of on-farm diversification

3.2.1 Farm and household factors

The NSS 70^{th} round provides information on farm and household level explanatory variables for our study. Details regarding the farm and household variables are given in Appendix 1. Under the farm head, we have data on various farm inputs, information about credit, produce sale etc. Most of the farm input variables are measured as as dummy variables because NSS records input usage in terms of value rather than actual usage. Information on market sale of crop and livestock are reported separately.

Summary statistics of these variables are presented in Appendix 2, for Visit 1 and Visit 2 separately. On an average total land of a farm household is reported to be around 1.5 hectares. Around 50% of farm household's land is under irrigation. More than half of farm households report to have used purchased seeds. On an average, around 58% of households report to have outstanding formal credit. About 78% of households employ hired farm labor in Visit 1, and this falls to 61% during Visit 2. 84% and 63% of farm households used fertilizers in Visit 1 and Visit 2, respectively. Almost all households report to have used farm machinery in Visit 1, but only 79% report farm machinery use in Visit 2. About 21% of households report to have received technical advice through newspapers, radio, TV, etc.

For controlling for household characteristics, we have information on household size, social group (SC/ST), gender of the household head, average age of the household, average education of the household, and proportion of household dependents. Non-farm economic activity is captured by the two variables on proportion of household members engaged in non-farm work and MGNREGA. Economic condition of the household is reflected by its monthly *per capita* consumption expenditure.

93% of households are headed by males. About 37% of household heads are illiterate. SC/ST households comprise of 28% of households. Average age of households is around 31 years, with a standard deviation among household members of 17 years. Household male to female ratio is around 52%. On an average, monthly *per capita* consumption is around Rs. 1340/- during Visit 1, and around Rs. 1638/- during Visit 2.

3.3 Village, district, and state factors

(i) Village level

As discussed earlier, village networks may impact farm level decisions through the labor and price channel. Using NSS data we generate for each household, the relative prices of exact crops the household is cultivating and livestock product the household is producing. Birthal et al. (2015b) only considered the proportion of households engaged in cultivating high-value crops (HVCs) in each village. But, farm households may be engaged in non-HVCs and animal husbandry as well. Therefore, this study takes into account not only information on village level relative prices of all types of crops cultivated by a particular household, but also has information on village level relative price of animal products. To capture the local labor market structure, the proportion of households engaged in non-farm activities in each village is also calculated. On an average 45% of households in a village have non-farm workers in Visit 1, and 50% during Visit 2.

(ii) District level

To bring in further information at the district level which may be crucial for farm diversification, relevant district data is compiled from Census GoI (2011) and other statistical sources. Since, village codes are not the same for NSS, Census and other sources, most of the district level variables are measured in terms of proportion.

The social demography of the district is given by proportion of SC & ST households in a district. SC & ST households are usually resource poor. This will indicate the general resource availability in the district. On an average 18% of district's population are SC, and 11% are ST population. Access to social infrastructure measured as proportion of villages with schools, proportions of villages with colleges within 5 kms, proportion of villages with access to public transport, and proportion of villages with towns within 5 kms. Accessibility of schools and colleges will inform whether availability of family labor is a binding condition for diversification. On an average there are 21% villages with middle & high schools.

Information on agricultural infrastructure is compiled from four sources viz. GoI (2011), GoI (2013a), RBI (2012), and EPWRF (2012). These provides information on net sown area in the district, proportion of fallow land, ratio of area under ground and well irrigation to NSA, per day average power availability for agricultural use, proportion of villages with veterinary hospitals, and proportion of agricultural credit. On an average, 56% of district's area is under NSA, and around 32% are fallow land. On an average, 37% of credit in a district is for agriculture.

Data on agro-climatic condition and soil characteristics is derived from different sources. District level rainfall deviation is taken from Rainfall Statistics in India (GoI, 2013b). Further, agro-climatic and soil conditions such as soil moisture availability index (MAI) and soil type data is compiled using ICRISAT (2018). Moisture availability and soil types reflect the district's natural suitability for diversification.

Level of structural transformation of a particular district is captured by average number of village manufactured products in the district, and urban centres in the district categorized into six different classes. This information is collected from GoI (2011). The urban centres are classified by population sizes. Class 1 and 2 cities are large urban centres with a population between 1,00,000 and above, and 50,000 to 99,999, respectively. Class 3, 4, 5, and 6 are relatively smaller urban centres. On an average, each district has one Class 1 and / or Class 2 cities, and ten Class 3, 4, 5 and 6 cities combined.

(iii) State level

To account for state level variation in pull factors, data is collated from RBI (2013) and RBI (2015). State agricultural policy characteristics is captured by per hectare agricultural capital and revenue expenditure. The general economic condition of the state is represented by *per capita* state gross domestic product. On an average a state spends Rs. 3,174.8 million per hectare on agriculture. Out of this only 18% is capital expenditure. The average *per capita* state GDP is around Rs. 81,061/-.

4 Estimation strategy

4.1 Ordered probit in a seemingly unrelated system

Motivated in the light of discussion above, apart from farm and household level "push" factors, on-farm diversification choice may also be influenced by external "pull" conditions. To account for this, on-farm diversification decision for the two seasons — *kharif* (Visit 1) and *rabi* (Visit 2) may be specified as follows:

$$\begin{aligned} y^*_{ivds1} = \alpha_0 + \alpha_1 FARMCH_{ivds1} + \alpha_2 HHLDCH_{ivds1} + \alpha_3 VILLCH_{vds1} + \alpha_4 DISTCH_{ds} \\ + \alpha_5 STATEPOL_s + \epsilon_{ivds1} \end{aligned} \tag{1}$$

$$\begin{aligned} y^*_{ivds2} = \beta_0 + \beta_1 FARMCH_{ivds2} + \beta_2 HHLDCH_{ivds2} + \beta_3 VILLCH_{vds2} + \beta_4 DISTCH_{ds} \\ + \beta_5 STATEPOL_s + \epsilon_{ivds2} \end{aligned} \tag{2}$$

In equations (1) and (2), the two seasons are represented by the subscripts 1 (*kharif* / Visit 1) and 2 (*rabi* / Visit 2). However, y^*_{ivds1} and y^*_{ivds2} are latent variables and are not observed. Rather, the response variable which is observed is y_{ivdst}, $t = 1, 2$:

$$y_{ivdst} = \begin{cases} 0 \text{ if } y^*_{ivdst} \leq 0 \\ 1 \text{ if } 0 < y^*_{ivdst} \leq \delta_1 \\ n \text{ if } \delta_{n-1} < y^*_{ivdst} \leq \delta_n \end{cases} \tag{3}$$

Here, $y_{ivdst} = 1, 2, \ldots n$ is the level of on-farm diversification (denoting the total number of crops and animal husbandry activities) of household i, in village v, district d, state s, in season t. $FARMCH_{ivdst}$ is the individual farm characteristics. $HHLDCH_{ivdst}$, are the household socio-economic and demographic characteristics. $VILLCH_{vdst}$ and $DISTCH_{ds}$, are the village and district characteristics, respectively. $STATEPOL_s$, are the state policy variables. And, ϵ_{ivdst} is a random error term and $\epsilon_{ivdst} \sim N(0,\sigma^2)$

Since, the dependent variable y_{ivdst} denotes level of on-farm diversification, equations (1) & (2), may be estimated using ordered probit models. However, as discussed earlier, some of the explanatory variables (external to fam and household conditions) are common to both the seasons. This may result in correlation between the error terms of the two seasons for a given v, d and s, to be correlated as follows:

$$E(\epsilon_{it}) = 0 \; \forall \; t = 1, 2$$

$$E(\epsilon_{it} \; \epsilon_{js}^{'}) = \begin{pmatrix} \sigma_{11} I_2 & \cdots & \sigma_{1N} I_2 \\ \vdots & \ddots & \vdots \\ \sigma_{N1} I_2 & \cdots & \sigma_{NN} I_2 \end{pmatrix} \tag{4}$$

Hence, to account for the correlation in the error terms between the two seasons, the ordered probit models are estimated in a seemingly unrelated regression system.

4.2 Correcting for endogeneity: A residual inclusion approach

As discussed till now, the identification strategy is to consider an ordered probit model in a SUR system. Concern remains over endogeneity of explanatory variables. Such biases are usually corrected by using instrumental variables (IV) which are correlated with the endogenous variable but do not directly impact the dependent variable of interest. A precondition is that the IV is excluded from the main structural equation under consideration. Consider a non-linear relationship (say, a logit or a probit) as follows:

$$Y = \psi(X_1, X_2, X_u) + \epsilon \tag{5}$$

Here, X_1 and X_2 to be two observed explanatory variables for the dependent variable Y. And, X_u is an unobserved explanatory variable such that $Cor(X_1, X_u) \neq 0$. This would imply that $Cor(X, \epsilon) \neq 0$, and result in bias estimates.

Suppose, Z is a possible instrument for the observed endogenous variable X_1. Then the first stage regression for estimating X_1 is given by:

$$X_1 = \omega(Z, X_2, X_u) + u \tag{6}$$

Here, $E(Z'u) = 0$ and $E(X'u) = 0$. From (6), $\hat{X}_1$ can be estimated and substituting $\hat{X}_1$ for X_1 in the second stage equation (5) corrects for endogeneity bias by clubbing the unobserved part $\hat{X}_u$ (contained in $\hat{X}_1$) with the second stage error term ϵ.

However, Terza (2017) and Wooldridge (2015), note that in a non-linear set-up, correcting for endogeneity using $\hat{X}_1$ as an explanatory variable in the structural model (second stage) results in inconsistent estimates. This is so because the $\hat{X}_u$ contained in $\hat{X}_1$ from the first stage in (6) cannot be additively assigned to the error term ϵ in the second stage (5) due to non-linearity of the function $\psi(X_1, X_2, X_u)$.

To address this problem of endogenous explanatory variable in a non-linear structural model, a "residual inclusion" (à la Terza (2017)) or a "control function" (à la Wooldridge (2015)) method may be applied. Some recent empirical studies have applied similar methods (Terza et al., 2008; Giles and Murtazashvili, 2013; Woldeyohanes et al., 2017; Michler and Josephson, 2017). This technique helps in correcting endogeneity when structural models are non-linear. Instead of estimating $\hat{X}_1$ from the first stage or reduced form equation (6), a residual estimate or a control function in the form of û is estimated as follows:

$$\hat{u} = X_1 - \hat{X}_1 \tag{7}$$

The estimated $\hat{u}$ from equation (7) is then replaced for X_1 in the second stage non-linear model (5). The inclusion of the estimated error term from the first stage would then control for the endogeneity of X_1 in the structural equation. The second stage is then estimated using maximum likelihood estimation (MLE). Hence, it leads to a two stage residual inclusion (2SRI) estimation. For a more detailed discussion on 2SRI or CF estimation approach, see

Terza (2017) and Wooldridge (2015).

As discussed earlier, since most of the explanatory variables external to the farm and household conditions are available only on an annual basis, the error terms over the two seasons may be correlated. Hence, the two ordered probit models are estimated for (Visit 1) and *rabi* (Visit 2) under seemingly unrelated regression (SUR) set-up. In this approach, the endogeneity in Visit 1 is corrected with a control function for Visit 1, and the endogeneity in Visit 2 is corrected with a control function for Visit 2. These two endogeneity corrected 2SRI models are then estimated in a seemingly unrelated regression (SUR) framework. Hence, it results in estimating a three stage residual inclusion (3SRI) model.

5 Results & policy implications

5.1 Do beyond the farm & household factors matter for diversification?

The explanatory variables are categorized into farm, household, village, district and state levels. For model specification, the analysis is first followed in a hierarchical order, from the local farm and household level factors, to the village, district and state level factors. Table 1 shows the four hierarchical models A, B, C, and D, consisting of explanatory variables from the farm and household level, to village, district and state location and policy variables. Models A and B consist of information based on NSS data. In addition to these, information from Census and other sources are added to generate models C and D.

We try to check which of the models has a better explanatory power based on Likelihood Ratio (LR) tests. When we compare model B over model A (Table 2), we find that in both two visits model B is a better specification than model A (p-value = 0.00). Similarly, when we compare model C over model B, we find model C is a better specification. Again, model D is a better specification over model C. This implies that external pull factors at the village, district and state levels, which are beyond the control of farm and household factors, influences on-farm diversification decision. Hence, for the rest of the analysis, we

Table 1: Hierarchical model specification & data description

Models	A		$B = A + Village$	$C = B + District$	$D = C + State$
Levels	Farm	Household	Village	District	State
Variables	Farm inputs	Soc. Group	Prop. in non-farm activity	Social demography	Agri. Exp.
	Credit	Age	Rel. price of major crop	Public & social infrastructure	(cap., rev.)
	Insurance	Education	Rel. price of animal product	(Rds., pub. tp., sch., dist. to town/coll., etc.)	PC GSDP
	Sale	Non-farm		Financial infrastructure	
	Tech. advice	MPCE		(financial serv., dist. agri. Credit, shg)	
				Agricultural infrastucture	
				(agri pwr., agri. mkt., manf. pdt.,)	
				Land use pattern (NSA, forest, fallow)	
				Agro-clim. (rain, mai, soil)	
				Urban agglomerations (city classifications)	
Source	NSS	NSS	NSS	NSS + Census + other sources	NSS + Census + other sources

consider model D specification with all the information on farm, household, village, district, and state variables.

Table 2: Likelihood ratio tests for models

Models	Visit 1		Visit 2	
	LR-chi2	p-value	LR-chi2	p-value
B vs A	2929.1	0.00	1511.2	0.00
C vs B	917.63	0.00	1278.06	0.00
D vs C	86.77	0.00	441.19	0.00

5.2 Testing for endogeneity

The external "pull" variables are determined at the village, district and state levels. These variables are exogenous to individual farm and household conditions. However, there might be endogeneity at the farm level explanatory variables. Information on most of the farm input usage such as irrigation use, fertilizer use, hired labor, animal labor, etc. are not differentiable by crop or animal activity. These are at the farm level as a whole. Hence, problem with endogeneity or simultaneity bias does not arise with regard to these inputs. But, there might be potential simultaneity bias of on-farm diversification and usage of crop seeds and animal seeds as farm level explanatory variables.

Being homogeneous clusters, village neighborhood decisions usually lead to complementary decisions to reap benefits of scale density (Cooper and John, 1988; Holmes and Lee, 2012).

The higher the proportion of households in the village growing the same crop or involved in the same animal husbandry activity, less will be the incentive for households to diversify and deviate. An important decision made by an agricultural household is investment in purchase of seeds. Neighborhood networks have been found to influence adoption of seeds in Tanzania through sharing of information and also even by sharing seeds (Larsen, 2018). Farmers in India often take advice and suggestions from their village folks on new seeds (Flachs and Stone, 2018). Thus, it is much likely that a farmer's adoption of seeds will be affected by the village neighborhood. It is conceivable that a household's adoption / purchase of seeds (for both crops and animal husbandry activities) will be affected by the extent to which the village is growing exactly the same crops. Thus, the variable on adoption of seeds may be endogenous to the household attributes. In order to control for this endogeneity, a control function approach is adopted. A candidate control for this would be the proportion of households in the village growing exactly the same crops. In a similar way, the same can be expected for purchase of cattle, poultry, goat, and other animal husbandry seeds.

Thus the proportion of households growing exactly the same crops in the village is a candidate control function for proportion of land under purchased seeds[3]. And, for purchase of cattle, goat, poultry and other seeds, we consider the proportion of households engaged in the same animal husbandry activities in the village as a control function for purchase of animal husbandry seed[4]. Since, the structural model is a non-linear ordered probit model the usual linear model endogeneity tests (such as Hausman tests) cannot be applied to check for endogeneity. In a residual inclusion approach, the coefficient of control function in the first stage needs be significant (Woldeyohanes et al., 2017). Hence we test the coefficients from the first stage OLS coefficients. The coefficient of proportion of crops grown by neighborhood is significantly different from zero for the two seasons (Table 3). However, the coefficient of proportion of households in the same animal husbandry activities is not significant for purchase of animal husbandry seeds. This implies that proportion of land under purchased seed

[3]This is possible because the NSS schedule records the crop codes for each of the crops grown by a household.

[4]NSS however do not record the animal husbandry seed code. It only reports whether any purchase is made for cattle, goat, poultry, and other seeds.

is endogenous and we have one control function for controlling this endogeneity. This also implies that though crop choice by a household is influenced by village neighborhood effect, animal husbandry activity choice on the other hand is not impacted by the neighborhood.

Table 3: Test for control functions

Proportion of land under purchased seeds			
Control function: Proportion of households growing the same crops in the village	Coefficient	p-value	Endogenous/ exogenous
Visit 1	-0.070	0.001	Endogenous
Visit 2	-0.040	0.099	Endogenous

Purchase of cattle, goat, poultry, other seed			
Control function: Proportion of households engaged in the same animal husbandry activities in the village	Coefficient	p-value	Endogenous/ exogenous
Visit 1	0.014	0.681	Exogenous
Visit 2	-0.021	0.475	Exogenous

As discussed earlier, application of residual inclusion in a non-linear model under SUR framework leads to a three stage residual inclusion (3SRI) model. We apply the respective residual or control functions in the two ordered probit models, and estimate a 3SRI model.

5.3 Final results

Ordered probit models are considered for model D Visit 1 (*kharif*) and Visit 2 (*rabi*) separately. Since, there could be correlation between the error terms of the two visits, the ordered probit models are estimated under a seemingly unrelated regression (SUR) framework. Further, potential endogeneity in the ordered probit SUR model are corrected using residual inclusion control function approach. Hence, a "3SRI" system estimation. The signs for the SUR and 3SRI estimation are presented in Table 4. The following discussion interprets result from the 3SRI estimation only. Regression coefficients are reported in Appendix 3.

5.3.1 Farm

As discussed above, the proportion of households growing exactly the same group of crops in the village is used as a control function for endogenous proportion of land under purchased seeds. Under the SUR model, during Visit 1 Visit 2, purchase of seeds may imply a higher degree of diversification. However, we see that controlling for the endogeneity results in negative significant effect on on-farm diversification. This shows that higher the proportion of households engaged in cultivation of same crop in the village, lower would be the probability to deviate and go for diversification. This result reinforces the argument of "economies of density" as proposed by Holmes and Lee (2012). It also confirms that farmers are purchasing seeds for crops as grown by their neighbors. As the farmers may gain from "strategic complementarity" (Cooper and John, 1988), it is better for a farmer to grow the same set of crops as his/her neighbors than to grow some other crops.

Though share of land under irrigation does not impact diversification in the first season, it however negatively impacts probability of diversification in Visit 2. So, whatever irrigation is available during the dry *rabi* season, are channelled towards a few crops. Farm size has a negative effect on the probability of diversification during Visit 1, but has positive impact during Visit 2. However, at a higher level (square term) farm size positively influences probability of diversification in both the two seasons. Hiring of human labor and usage of fertilizer negatively impacts diversification decision in both the seasons. This implies that farmers restrict the usage of these farm inputs into few crops. Though electricity has a negative impact on probability of diversification during Visit 1, it does not seem to matter in Visit 2. Hiring of machines raises the probability of diversification during the *kharif* season, but lowers its probability during *rabi* season. Purchase of poultry and other animal husbandry seeds reflects higher chances of diversification during Visit 2, though it may imply lower diversification in Visit 1. If the household has any outstanding credit, then lower will be the probability of the household to diversify. Progressive farmers often specialize in growing of particular crop or raising particular livestock. Hence, technical advice received from progressive farmer leads helps in gaining advantage of scale density (Holmes and Lee, 2012).

Table 4: Determinants of on-farm diversification in India

	SUR		3SRI	
	Visit 1	Visit 2	Visit 1	Visit 2
FARM				
Proportion of land under purchased seeds	+++	+++		
Control function			- - -	- - -
Proportion of land under irrigation	- - -	-		- - -
Dummy insurance			- - -	+++
Total land			- - -	++
Square total land			+++	+++
Dummy human labour		+++	- - -	- - -
Dummy fertilizer	- - -	- - -	- - -	- - -
Dummy electricity	+++		- - -	
Dummy machine repair		+++	- - -	+++
Dummy machine hire		+++	+++	- - -
Dummy cattle seed	++	+++	- - -	- - -
Dummy goat seed		- -	+++	- - -
Dummy poultry seed		++	- - -	+++
Dummy other seed		+++	- - -	+++
Dummy green fodder			++	
Dummy dry fodder	++	+++	- - -	- - -
Dummy veterinary		+++	+++	- - -
Dummy outstanding credit	+++	+++	- - -	- - -
Dummy tech advice (AU)	++		+++	- - -
Dummy tech advice (PF)	- - -		- - -	- - -
Dummy tech advice (RN)	+++	+	+++	+++
Crops sale under MSP (No.)		++	- - -	- - -
Square of crops sale under MSP	+++		+++	+++
Livestock market sale (No.)		++	- - -	+++
HOUSEHOLD				
Dummy head male			- -	- - -
Dummy head illiterate		+	+++	
Household size			-	- - -
Dummy SC/ST household			+++	
Avg household age			+++	- - -
SD household age	- -		- - -	+++
Proportion of dependents			- - -	
Male female ratio			++	
Avg education (adults in agri)			+++	- -

Table 4: Determinants of on-farm diversification in India (*contd*)

	SUR		3SRI	
	Visit 1	Visit 2	Visit 1	Visit 2
Proportion of principal non-farm workers				+++
Proportion of MGNREGA workers			- - -	- - -
Monthly per capita consumption			- - -	+++
Square of monthly per capita consumption				- - -
VILLAGE				
Village proportion of non-farm workers			- - -	+++
Relative price of crops			- - -	- - -
Relative price of animal husbandry		+++	- - -	- - -
DISTRICT				
(i) *Infrastructure*				
Proportion of SC			- - -	- - -
Proportion of ST	- -	- - -	- - -	- - -
Proportion of NSA to total district area	++	+++	+	+++
Proportion of fallow land to total district area			+++	+++
Proportion with middle and high schools	- - -	-	+++	+++
Proportion with colleges within 5 kms		+++		+++
Proportion with towns within 5kms	- -	- - -	- - -	- - -
Proportion with public transport			- - -	++
Public transport * towns within 5 kms	++	+++	+++	+++
Proportion with vet hospital	++		++	
Proportion with SHG	+++		- - -	+++
Avg number of village manufactured goods			+++	+++
Square avg no. of vill manufactured goods	- - -	- -	- - -	- - -
Proportion SHG * avg manufactured goods	+++		+	- - -
Proportion of agri credit to total credit	- - -		+++	+++
Proportion SHG * proportion agri credit to total credit			- - -	- - -
Avg power available for agri use (hrs./day)	++	- - -		+++
Square avg power available for agri use (hrs./day)	- - -	+++	+++	- - -
Ratio of ground and well irrigation to total NSA	- -	-	+++	+++

Table 4: Determinants of on-farm diversification in India (*contd*)

	SUR		3SRI	
	Visit 1	Visit 2	Visit 1	Visit 2
DISTRICT				
(ii) *Urban agglomeration*				
Class 1 cities	- - -			- - -
Square of class 1 cities	++		- - -	+++
Class 2 cities	-		- - -	+++
Square of class 2 cities			+++	- - -
Class 3, 4, 5, 6 cities		+++	+++	- - -
Square of class 3, 4, 5, 6 cities		- - -	- - -	+++
Ratio of class 1 and 2 cities to total cities	++	+	+++	
STATE				
Ratio of capital to total agri expenditure		+++	- - -	+++
Total agri expenditure (Rs. million/hect)	- - -	+	- - -	+++
Square total agri expenditure	+++		+++	- - -
Per capita state GDP	- - -			+++
Square per capita state GDP	++		-	- - -
District rainfall deviation	- - -	+++	+++	+++
District soil moisture availability index	*Included*	*Included*	*Included*	*Included*
Proportion of 19 soil types in the district	*Included*	*Included*	*Included*	*Included*
N	25021	25021	25021	25021

Note: The first two columns are SUR estimation results, without endogeneity correction. The last two columns are endogeneity corrected 3SRI estimation results. The *control functions* for proportion of land under purchased seeds for each seasons are the share of households in the village growing exactly the same crops. Technical advice from AU refers to agricultural university, PF refers to progressive farmers, and RN refers to radio, newspaper, TV etc. Among the district level variables, average power availability for agricultural use and district rainfall deviation are available for *kharif* and *rabi* seasons separately. 19 different soil types were considered in this study (as in ICRISAT data). The proportion of land area in each district for each of these 19 different soil types were calculated.
+++/- - - significant at 1% level, ++/- - significant at 5% level, +/- significant at 10% level.

However, technical wisdom received from mass media sources such as radio, newspaper, TV shows etc. positively impacts decision to diversify. Higher the number of crops sold at a minimum support price (MSP), lower will be the probability for the agricultural household to diversify. But, sale of livestock during the lean season of *rabi* positively influences the farmer to diversify.

5.3.2 Household

Under household characteristics, having a male headed household leads to lower probability of diversification. Higher the number of household members, lower will be the influence on diversification decision. A Scheduled Caste or Scheduled Tribe (SC/ST) household has a higher probability of diversification during Visit 1. A higher aged household will have a higher probability to diversify during Visit 1 but will have a lower probability to diversify during Visit 2. If the standard deviation of the age of household members is high, lower will be the probability of household to diversify during Visit 1 but will increase the probability of diversification in Visit 2. Higher the proportion of dependents in the household (household members with age < 15 and > 60), lower would be the probability to diversify during Visit 1. Average education of adult (age > 15) household members engaged in agriculture increases the probability to diversify during Visit 1 but lowers its probability during Visit 2. This may because during the Visit 2 (*rabi*) most of the academic examinations take place during this season in India. Hence, those adult members who are still in school or colleges would not be available to provide extra hand in farming activities. Therefore, probability of diversification may decrease during Visit 2. If a higher proportion of household members are engaged in public work programs like MGNREGA[5], lower would be the probability of diversification during both the seasons. This may be too working through the labor market channel. A higher level of household per capita consumption expenditure implies a lower level of diversification during the *kharif* season, but would increase its probability during the *rabi* season. But, at a higher level of consumption expenditure this probability goes down. This may be because during the *rabi* season, due to lack of rainfall and other resources,

[5]Mahatma Gandhi National Rural Employment Guarantee Act.

a poorer household need to diversify more in order to reduce the production risk involved during that season.

5.3.3 Village

A higher proportion of households in the village engaged in non-farm work decreases the probability of diversification during Visit 1, but raises its probability during Visit 2. This may be because during *kharif* season due to rainfall availability, there is higher demand for farm labor, which then becomes a constraint. Whereas, during the lean *rabi* season, the demand for farm labor goes down, which in turn relaxes the labor demand constraint. So, whatever labor is available is sufficient for diversification.

A higher relative price of crops grown and animal husbandry activities of the household dampen the probability for further diversification. This is obvious from the fact that higher the prices of the produce of the household, lower will be the incentive to deviate from those activities. This is another way of suggesting that economies of density is at work through the price channel (Holmes and Lee, 2012).

5.3.4 District

(a) Socioeconomic & infrastructure

A higher proportion of Scheduled Caste and Scheduled Tribe (SC & ST) population in the district decreases the probability of on-farm diversification. SC & ST households are generally resource poor households with unfavorable socio-economic conditions. A district having a higher proportion of SC & ST households will have less diversification potential.

A higher proportion of net sown area (NSA) to total geographical area of the district would lead to more chances of diversification in both the two seasons. This signifies that more the land available for sowing, higher will be the probability of growing different commodities.

Interestingly, a higher proportion of fallow land[6] in the district would also increase the probability of diversification in both the Visits. It may be because increase in fallow land helps in regaining of soil fertility which is essential for supporting a variety of cropping system.

Provision of high and secondary schools in the village, affects farm diversification through the labor market. A higher proportion of villages in the district with schools within the village raises farm diversification. Having a nearby school provides for easy availability of a helping hand in the household. This is specially true for animal husbandry activities, which are relatively labor intensive than crop. Similarly, if the proportion of of villages with colleges within a distance of less than 5 kms increases, it will provide a quick helping hand in farm activities and hence raise diversification. This brings out the importance of family labor in farming activities.

A higher proportion of villages in the district with towns within less than 5 kms, reduces the probability of on-farm diversification. However, its interaction with access to public transport, positively influences on-farm diversification. This implies that having a nearby town with access to public transport facility in villages increases the probability of diversification.

Number of village manufactured products has an inverted U-shaped relationship with probability of diversification. Initially, it helps in raising probability of diversification. This is because agricultural households benefit from "pluriactivity" in their location vicinity (Rigg, 2006). However, at a higher level of village manufactured products, the probability of diversification goes down. This may imply structural transformation a at higher level of manufactured products in the villages.

Access to credit helps in agricultural growth (Das et al., 2009; Narayanan, 2016). A higher proportion of credit towards agricultural and allied activities in the district, helps in increasing the probability of farm diversification. Its interaction with SHG, however, leads to a negative impact on diversification decision. This may be so because association with SHGs

[6] This includes culturable waste land, fallow and current fallow land.

usually lead to working in a group which concentrate on investing in a particular agricultural commodity, thereby decreasing farm diversification decision. Hence, there is need for policy changes for agricultural credit through SHGs to have a positive impact on farm diversification (Mohan, 2006).

Increase in the availability of power for agricultural use during the *rabi* season, lead to increase in diversification, but at a higher level probability of diversification falls. However, at a higher level, availability of power during *kharif*, will lead to an increase in probability of diversification. Increase in the ratio of irrigation from ground and wells to total NSA, raises probability of diversification.

(b) Urban agglomeration

Our results show that, presence of urban centres in the district may not have the same impact on diversification decisions during the two seasons. Also, different types of urban centres have different impact within a season. Class 1 cities have a U-shaped relationship with the probability of diversification during Visit 2 i.e. at an initial level presence of class 1 cities have a negative influence on diversification, but as its number increases, it has a positive impact on diversification decision. Similarly, presence of class 2 type of cities has a U-shaped relationship with probability of diversification during Visit1, but the relationship becomes inverted U-shaped during Visit 2. While Class 3, 4, 5, 6 cities have an inverted U-shaped relationship with probability of diversification during Visit 1, it has a U-shaped relationship during Visit 1. Having more proportion of class 1 2 cities in the district raises the probability of diversification decision during Visit 1 only.

(c) Soil

Apart from controlling for moisture content of soil in the district, this study also considered 19 different types of soil quality. However, data on soil types are available only at the district level. This study considered the proportion of area in each district, under each of these 19 different soil types. Based on the 3SRI results, we categorized districts with soil types favourable or unfavourable for diversification according to the following:

Favourable for on-farm diversification = 1

Not favourable for on-farm diversification = -1

No effect on on-farm diversification = 0

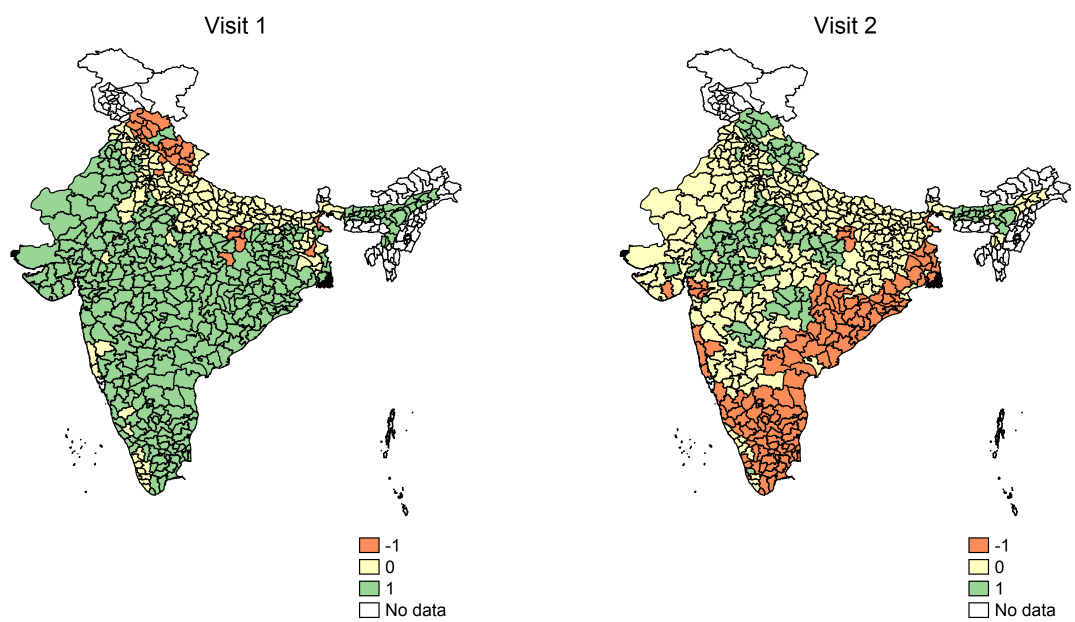


Figure 4: District soil with $+/-$ or no effect on on-farm diversification in India
Source: Author's calculation.

The soil district maps are presented in Figure 4. We find that during the *kharif* (Visit 1) season, except Central and Western Uttar Pradesh, Northern Bihar, Uttarakhand, Northern Punjab, Western Himalayan regions, almost all the rest of the districts in India have soil types which can positively influence on-farm diversification decisions. Soil characteristics in the Upper Himalayan regions of Himachal Pradesh, Uttarakhand, some districts in West Bengal, Western Maharashtra, Kerala either disfavours or has no-effect on the probability of on-farm diversification. However, during *rabi* (Visit 2) season, the situation changes drastically. Almost all the coastal districts of India have soil types which does not favour diversification decision. Compared to the Western Ghats, this phenomenon extends much more deeper into the interior districts of the Eastern Ghats. All the districts of Tamil Nadu, Andhra Pradesh, some regions of Telangana, Odisha, West Bengal, have soil types which have a negative impact on the probability of on-farm diversification. In the remainder of the districts, most of the soil types turn to have no impact on diversification decision. Soil types either provide a natural structural stimulus or barrier for on-farm diversification decision.

5.3.5 State

Ratio of capital expenditure on agriculture and allied activities to total agricultural expenditure in the state improves the probability of diversification during Visit 2, but leads to a decline in probability of diversification during Visit 1. During Visit1, total expenditure on agriculture and allied activities raises the probability of diversification but only at a higher level. Though expenditure under this heading increases the probability of diversification during Visit 2, but at a higher level it may lead to negative influence on diversification. A higher per capita state GDP signifies a higher probability of diversification during Visit 2 only. But, at a higher level, it has a negative impact on diversification decision in both the two seasons.

5.4 Discussion & policy implications

5.4.1 Agro-ecology

The above results show that agro-ecological factors, which are external to farm and household conditions, significantly affect on-farm diversification decision. Apart from rainfall and soil characteristics, we have brought in factors like NSA, fallow land, irrigation facility into our analysis. NSA and fallow land are favorably disposed towards farm diversification, whereas irrigation availability helps in diversification during the lean *rabi* season. The agro-ecological environment at large has a bearing on farm diversification process. This implies that restricting analysis of farm diversification with only farm and household conditions may limit our understanding of on-farm diversification process.

Most of the districts have soil types which have a positive impact on on-farm diversification decision during the *kharif* season. However, a large number of districts turn out to have soil types, which negatively influences on-farm diversification during the *rabi* period. This is particularly true for most of the districts in the coastal and peninsular India. Soil types provide a natural hurdle for on-farm diversification. Hence, policy makers should look for alternate means of encouraging diversification in such regions during the *rabi* season.

5.4.2 Local infrastructure

We find that provision of social infrastructure such as schools and colleges positively impacts on-farm diversification. Ease of access to schools and colleges influences farm diversification through the labor market. Having a nearby educational facility may be crucial for easy application of young family labor in the farm, especially in case of animal husbandry activities such as rearing of cattle, etc. The general understanding is that, as the demand for education increases, it will pull out labor from the agricultural sector. However, our results show that accessible educational facility may in fact positively influence farm diversification. This indicates that family labor still matters for farming activities in India. This has a policy relevance for provisioning of educational institutions within the vicinity of villages so that these are easily accessible to the farm households.

Due to higher wages in the non-farm sector, it is widely construed that labor is moving away from the farm to the non-farm sector (Chand and Srivastava, 2014). This implies that having a nearby town would aggravate this condition, and hence negatively impact diversification. However, we find that if villages have accessible public transport facility, then having a nearby town would rather lead to an increase in farm diversification. Thus, to have a positive impact of nearby towns on on-farm diversification, it is important for policy makers to make villages more accessible through better public transport facilities.

5.4.3 Urbanization & structural transformation

Urbanization leads to increase in the demand for variety of agricultural products. As urbanization rises demand for high value products such as meat, dairy, etc. also increases (Rao et al., 2007). Though urbanization might provide a bigger market for variety of products, it may as well negatively impact on-farm diversification through the labor market. A Lewisian view would be that urbanization results in better non-farm employment opportunity, which may pull out farm labor from agriculture (Lewis, 1954). Consequently, labor supply may become a constraint for on-farm diversification. However, our results suggests that, for urbanization to have a positive (through higher demand) or negative (through farm labor constraint) impact on diversification, its not only urbanization *per se* which matters, but also the level or degree of urbanization. We also find a seasonal effect of urbanization on farm diversification.

Structural transformation has resulted in the share of non-farm sector to rise in the rural economy (Chand et al., 2017). There may be positive feedback effect into the farm sector through higher agricultural investments. Our results show a positive association between number of village manufactured products and on-farm diversification. However, at a higher level, negative impact on diversification may set in by impinging on farm labor supply.

As the economy grows, there would be further increase in urbanization and structural transformation in the economy. In order to ensure that such changes make a positive impact on farm diversification, it is important for policy makers to provide viable labor replacing technologies in the agricultural sector. This would make the farm sector forbear the impact of shifting labor from the farm to the non-farm sector.

5.4.4 Neighborhood network & scale density

We find that neighborhood can impact farm diversification through the village labor market. Relative prices of crops and animal husbandry also influences farm diversification. In controlling for endogeneity in our analysis, we find that farmers are encouraged to grow what their neighbors are growing. All these implies that there is a neighborhood network effect on on-farm diversification decision. Flachs and Stone (2018) note that farmers often interact with their neighborhood for getting information on seeds. Farm households benefit by adopting a "complementary strategy" (Cooper and John, 1988) and hence gain "economies of density" (Holmes and Lee, 2012) by growing the same crop as by the rest of the villagers. This has implication for adoption of farm diversification practices. Calibrated policies are required for taking into account the village community as whole rather than targeting individual farmers in isolation.

6 Conclusion

Apart from farm and household factors, there is a larger context and non-farm environment under which a farm operates. This study contributes to the understanding of such pull factors which are external to farm and household conditions which impinges on on-farm diversification process. Although a lot of studies have looked into the contribution of farm and household factors, literature is often oblivious to the contribution of external factors towards farm diversification. This paper tries to fill that gap in the literature by bringing in that extra information beyond the farm and household which can significantly affect farm diversification.

Using latest data on agricultural households in India, and coalescing it with external pull factor information drawn from several other sources, this paper assesses the contribution of farm, household, village, district, and state variables towards on-farm diversification decision. Applying a novel residual inclusion method of controlling for endogeneity, a "three stage" residual inclusion (3SRI) model is estimated. Based on the estimation results, this study finds that village network influences on-farm diversification through cultivation of same crops as by the neighborhood. This reaffirms that farm household reap benefits from "economies of density" as noted by Holmes and Lee (2012). Results also show that there is a broader agro-ecological environment under which a farm operates and these external factors could significantly influence on-farm diversification decision. Provision of accessible public infrastructure such as schools, colleges and public transport positively impacts on-farm diversification. For urbanization and structural transformation to have a positive influence on farm diversification, viable labor replacing technologies should be made available to the agricultural sector. Capital and total expenditure on agricultural and allied activities raises the probability of diversification during the lean period of *rabi* season.

A limitation of this analysis is that it provides a static view of the determinants of on-farm diversification. Provided a panel data set is available for non-farm explanatory variables, it would be interesting to understand the spatio-temporal impact of non-farm determinants on on-farm diversification decisions.

References

Agarwal, B. and Agrawal, A. (2017). Do farmers really like farming? Indian farmers in transition. *Oxford Development Studies*, 45(4):460–478.

Aggarwal, S. (2018). Do rural roads create pathways out of poverty? Evidence from India. *Journal of Development Economics*, 133:375–395.

Asher, S. and Novosad, P. (2016). Market access and structural transformation: Evidence from rural roads in india. *Manuscript: Department of Economics, University of Oxford*.

Auffhammer, M. and Carleton, T. A. (2018). Regional crop diversity and weather shocks in India. *Asian Development Review*, 35(2):113–130.

Barrett, C. B., Reardon, T., and Webb, P. (2001). Nonfarm income diversification and household livelihood strategies in rural Africa: Concepts, dynamics, and policy implications. *Food policy*, 26(4):315–331.

Birthal, P. S., Jha, A., Joshi, P., Singh, D., et al. (2006). Agricultural diversification in north eastern region of India: Implications for growth and equity. *Indian Journal of Agricultural Economics*, 61(3):328.

Birthal, P. S., Joshi, P. K., Roy, D., and Thorat, A. (2013). Diversification in Indian agriculture toward high-value crops: The role of small farmers. *Canadian Journal of Agricultural Economics*, 61(1):61–91.

Birthal, P. S., Roy, D., Khan, M. T., and Negi, D. S. (2015a). Farmers' preference for farming: Evidence from a nationally representative farm survey in India. *The Developing Economies*, 53(2):122–134.

Birthal, P. S., Roy, D., and Negi, D. S. (2015b). Assessing the impact of crop diversification on farm poverty in India. *World Development*, 72:70–92.

Chand, R. and Kumar, P. (2004). Determinants of capital formation and agriculture growth: Some new explorations. *Economic & Political Weekly*, pages 5611–5616.

Chand, R., Kumar, P., and Kumar, S. (2012). Total factor productivity and returns to public investment on agricultural research in India. *Agricultural Economics Research Review*, 25(2).

Chand, R. and Srivastava, S. K. (2014). Changes in the rural labour market and their implications for agriculture. *Economic & Political Weekly*, 49(10):47–54.

Chand, R., Srivastava, S. K., and Singh, J. (2017). Changes in rural economy of India, 1971 to 2012. *Economic & Political Weekly*, 52(52):64–71.

Chatterjee, U., Murgai, R., and Rama, M. (2015). Employment outcomes along the rural-urban gradation. *Economic & Political Weekly*, 50(26):5–10.

Chhatre, A., Devalkar, S., and Seshadri, S. (2016). Crop diversification and risk management in Indian agriculture. *DECISION*, 43(2):167–179.

Cooper, R. and John, A. (1988). Coordinating coordination failures in Keynesian models. *The Quarterly Journal of Economics*, 103(3):441–463.

Dadhich, C. L. (2015). Role of dairying in diversification of Indian agriculture. In Ghosh, M., Sarkar, D., and Roy, B. C., editors, *Diversification of Agriculture in Eastern India*, pages 99–107. Springer.

Das, A., Senapati, M., and John, J. (2009). Impact of agricultural credit on agriculture production: An empirical analysis in India. *Reserve Bank of India Occasional Papers*, 30(2):75–107.

Ellis, F. (1998). Household strategies and rural livelihood diversification. *The Journal of*

Development Studies, 35(1):1–38.

Ellis, F. (2000). The determinants of rural livelihood diversification in developing countries. *Journal of Agricultural Economics*, 51(2):289–302.

EPWRF (2012). EPW Research Foundation India Time Series. Technical report, EPW Research Foundation Accessed online on May 4 2018.

Fan, S., Gulati, A., and Thorat, S. (2008). Investment, subsidies, and pro-poor growth in rural India. *Agricultural Economics*, 39(2):163–170.

Fan, S., Hazell, P., and Thorat, S. (2000). Government spending, growth and poverty in rural India. *American Journal of Agricultural Economics*, 82(4):1038–1051.

FAO (2018). The State of Food Security and Nutrition in the World. Technical report, FAO, IFAD, UNCIEF, WFP, WHO.

Flachs, A. and Stone, G. D. (2018). Farmer knowledge across the commodification spectrum: Rice, cotton, and vegetables in Telangana, India. *Journal of Agrarian Change*.

Giles, J. and Murtazashvili, I. (2013). A control function approach to estimating dynamic probit models with endogenous regressors. *Journal of Econometric Methods*, 2(1):69–87.

GoI (2011). Village Directory, Towns and Urban Agglomerations classified by population size. Technical report, Office of the Registrar General Census Commissioner, Ministry of Home Affairs, Government of India.

GoI (2013a). Land Use Statistics. Technical report, Department of Economics and Statistics, Ministry of Agriculture, Cooperation and Farmers Welfare.

GoI (2013b). Rainfall statistics in India. Technical report, Rainfall Statistics in India, 2012-2013, India Meteorological Department, Ministry of Earth Sciences.

Gulati, A., Rajkhowa, P., and Sharma, P. (2017). Making rapid strides-Agriculture in Madhya Pradesh: Sources, drivers, and policy lessons. *Working Paper ICRIER*.

Haggblade, S., Hazell, P., and Reardon, T. (2010). The rural non-farm economy: Prospects for growth and poverty reduction. *World development*, 38(10):1429–1441.

Holmes, T. J. and Lee, S. (2012). Economies of density versus natural advantage: Crop choice on the back forty. *Review of Economics and Statistics*, 94(1):1–19.

ICRISAT (2018). District Level Data. Technical report, International Crop Research Institute for Semi-Arid Tropics, Hyderabad.

Joshi, P. K., Gulati, A., and Birthal, P. S. (2007). Agricultural diversification in India: Status, nature and pattern. In Joshi, P. K., Gulati, A., and Cummings Jr., R., editors, *Agricultural Diversification and Smallholders in South Asia*, pages 219–242. Academic Foundation, New Delhi.

Joshi, P. K., Gulati, A., Birthal, P. S., and Tewari, L. (2004). Agriculture diversification in South Asia: Patterns, determinants and policy implications. *Economic and Political Weekly*, pages 2457–2467.

Kishore, A., Birthal, P. S., Joshi, P., Shah, T., and Saini, A. (2016). Patterns and drivers of dairy development in India: Insights from analysis of household and district-level data. *Agricultural Economics Research Review*, 29(1):1–14.

Krishna Kumar, K., Rupa Kumar, K., Ashrit, R., Deshpande, N., and Hansen, J. (2004). Climate impacts on Indian agriculture. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 24(11):1375–1393.

Lange, A., Piorr, A., Siebert, R., and Zasada, I. (2013). Spatial differentiation of farm diversification: How rural attractiveness and vicinity to cities determine farm households response to the CAP. *Land Use Policy*, 31:136–144.

Larsen, A. F. (2018). When knowledgable neighbors also share seedlings: Diffusion of banana cultivation in Tanzania. *Agricultural Economics*.

Lee, D. R. (2005). Agricultural sustainability and technology adoption: Issues and policies for developing countries. *American Journal of Agricultural Economics*, 87(5):1325–1334.

Lewis, W. A. (1954). Economic development with unlimited supplies of labour. *The Manchester School*, 22(2):139–191.

Meert, H., Van Huylenbroeck, G., Vernimmen, T., Bourgeois, M., and Van Hecke, E. (2005). Farm household survival strategies and diversification on marginal farms. *Journal of Rural Studies*, 21(1):81–97.

Michler, J. D. and Josephson, A. L. (2017). To specialize or diversify: Agricultural diversity and poverty dynamics in Ethiopia. *World Development*, 89:214–226.

Mohan, R. (2006). Agricultural credit in India: Status, issues and future agenda. *Economic and Political Weekly*, pages 1013–1023.

Motamed, M. J., Florax, R. J., and Masters, W. A. (2014). Agriculture, transportation and the timing of urbanization: Global analysis at the grid cell level. *Journal of Economic Growth*, 19(3):339–368.

Nakajima, M., Otsuka, K., and Yamano, T. (2018). Jobs off the farm: Wealth, human capital, and social group in rural eastern India. *The Journal of Development Studies*, 54(1):111–132.

Narayanan, S. (2016). The productivity of agricultural credit in India. *Agricultural Economics*, 47(4):399–409.

Palmer-Jones, R. and Sen, K. (2003). What has luck got to do with it? A regional analysis of poverty and agricultural growth in rural india. *The Journal of Development Studies*, 40(1):1–31.

Pingali, P. L. and Rosegrant, M. W. (1995). Agricultural commercialization and diversification: Processes and policies. *Food policy*, 20(3):171–185.

Rao, P. P., Birthal, P. S., Joshi, P., and Kar, D. (2007). Agricultural diversification towards High-Value Commodities and role of urbanisation in India. In Joshi, P. K., Gulati, A., and Cummings Jr., R., editors, *Agricultural Diversification and Smallholders in South Asia*, pages 243–270. Academic Foundation, New Delhi.

RBI (2012). District Banking Statistics for 2012-13. Technical report, Reserve Bank Of India.

RBI (2013). Handbook of Statistics on Indian States for 2012-13. Technical report, Reserve Bank Of India.

RBI (2015). Study of state budgets for 2012-13. Technical report, Reserve Bank of India.

Reardon, T., Stamoulis, K., and Pingali, P. (2007). Rural nonfarm employment in developing countries in an era of globalization. *Agricultural Economics*, 37:173–183.

Richards, P. (2018). Its not just where you farm; its whether your neighbor does too. how agglomeration economies are shaping new agricultural landscapes. *Journal of Economic Geography*, 18(1):87–110.

Rigg, J. (2006). Land, farming, livelihoods, and poverty: Rethinking the links in the Rural South. *World Development*, 34(1):180–202.

Robalino, J. A. and Pfaff, A. (2012). Contagious development: Neighbor interactions in deforestation. *Journal of Development Economics*, 97(2):427–436.

Terza, J. V. (2017). Two-stage residual inclusion estimation: A practitioners guide to Stata implementation. *Stata Journal*, 17(4):916–938.

Terza, J. V., Basu, A., and Rathouz, P. J. (2008). Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. *Journal of Health Economics*, 27(3):531–543.

Thapa, G., Kumar, A., Roy, D., and Joshi, P. (2018). Impact of crop diversification on rural poverty in Nepal. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, 66(3):379–413.

Tilman, D., Cassman, K. G., Matson, P. A., Naylor, R., and Polasky, S. (2002). Agricultural sustainability and intensive production practices. *Nature*, 418(6898):671.

Vandercasteelen, J., Beyene, S. T., Minten, B., and Swinnen, J. (2018). Cities and agricultural transformation in Africa: Evidence from Ethiopia. *World Development*, 105:383–399.

Weinberger, K. and Lumpkin, T. A. (2007). Diversification into horticulture and poverty reduction: A research agenda. *World Development*, 35(8):1464–1480.

Woldeyohanes, T., Heckelei, T., and Surry, Y. (2017). Effect of off-farm income on smallholder commercialization: Panel evidence from rural households in Ethiopia. *Agricultural Economics*, 48(2):207–218.

Wooldridge, J. M. (2015). Control function methods in applied econometrics. *Journal of Human Resources*, 50(2):420–445.

Xiaoping, S., Heerink, N., and Futian, Q. (2007). Choices between different off-farm employment sub-categories: An empirical analysis for Jiangxi Province, China. *China Economic Review*, 18(4):438–455.

Appendices

Appendix 1: Variable notations, definitions & sources

FARM & HOUSEHOLD	
Source: NSS 70^{th} round	
Farm level information	
Number of on-farm diversification	Total crop & animal husbandry varieties
Total land	Total land possessed (in hectares)
Proportion of land under irrigation	Share of irrigated land
Proportion of land under purchased seeds	Proportion of land under purchased seeds
Dummy human labour	Dummy variable=1 if human labour used
Dummy fertilizer	Dummy variable=1 if fertiliser used
Dummy electricity	Dummy variable=1 if electricity used
Dummy machine repair	Dummy variable=1 if machine maintenance
Dummy machine hire	Dummy variable=1 if machine hired
Dummy cattle seed	Dummy variable=1 if expenditure made on cattle seed
Dummy goat seed	Dummy variable=1 if expenditure made on goat seed
Dummy poultry seed	Dummy variable=1 if expenditure made on poultry seed
Dummy other seed	Dummy variable=1 if expenditure made on other seed
Dummy green fodder	Dummy variable=1 if expenditure made on green fodder
Dummy dry fodder	Dummy variable=1 if expenditure made on dry fodder
Dummy veterinary	Dummy variable=1 if expenditure made on veterinary
Dummy insurance	Dummy=1 if any crop insured
Dummy outstanding credit	Dummy=1 if any outstanding institutional credit
Livestock market sale (No.)	Number of livestock sold to co-operative & govt. agency
Crops sale under MSP (No.)	Number of crops sold at MSP
Dummy tech advice (AU)	Dummy=1 if tech advice from agricultural university
Dummy tech advice (PF)	Dummy=1 if tech advice from progressive farmer
Dummy tech advice (RN)	Dummy=1 if tech advice from radio, newspaper, TV etc.
Household characteristics	
Dummy head male	Dummy=1 if head is male, 0 otherwise
Dummy head illiterate	Dummy=1 if head is illiterate, 0 otherwise
Household size	Household size
Dummy SC/ST household	Dummy=1 if household ST/SC
Avg household age	Average age of the household members
Proportion of dependents	Proportion of dependents in the household
Male female ratio	Ratio of male to female members in the household
Avg education (adults in agri)	Average education of adults engaged in agriculture
Proportion of MGNREGA workers	Proportion of household members engaged in MGNREGA works
Proportion of principal non-farm workers	Proportion of members principally in non-farm
Monthly per capita consumption	Monthly per capita expenditure of the household

Appendix 1: Variable notations, definitions & sources (*contd.*)

VILLAGE	
Source: NSS 70^{th} round	
Village proportion of non-farm workers	Proportion of households engaged in non-farm activities in the village (except itself)
Relative price of crops	Relative price of crops produced to all the similar crops produced in the village (except itself)
Relative price of animal husbandry	Relative price of livestock produced to all livestock produced in the village (except itself)
DISTRICT	
Source: Census, RBI, IMD, GoI & ICRISAT	
Proportion of SC	Proportion of SC population in the district
Proportion of ST	Proportion of ST population in the district
Proportion with middle and high schools	Proportion of villages with high & sec. schools
Proportion with colleges within 5 kms	Proportion of villages with college $<$ 5 kms
Proportion with public transport	Proportion of villages with public transport
Proportion with towns within 5kms	Proportion of villages with town $<$ 5 kms
Proportion with SHG	Proportion of villages with SHGs
Proportion of NSA to total district area	Proportion of net sown area to total area
Proportion of fallow land to total district area	Ratio of culturable waste, fallow & current fallow
Avg power available for agri use (hrs./day)	Weighted daily avg. agricultural power use
Proportion with vet hospital	Proportion of villages with a veterinary hospital
Ratio of ground and well irrigation to total NSA	Ratio of well and ground irrigation to NSA
Proportion of agri credit to total credit	Proportion of agricultural to total credit
Avg number of village manufactured goods	Avg no. of manufacture & handicrafts produced
District rainfall deviation	Rainfall deviation from normal in Visit i=1,2
District soil type	Proportion of district area under each soil types
Moisture availability index	Moisture available in the soil type of the district
Class 1 cities	Urban centres with population 100000 or above
Class 2 cities	Urban centres with population 50000-99999
Class 3, 4, 5, 6 cities	Urban centres with population 20000-49999, 10000-19999, 5000-9999, below 5000
Ratio of class 1 and 2 cities to total cities	Ratio of class 1 & 2 urban centres in the district
STATE	
Source: RBI & GoI	
Ratio of capital to total agri expenditure	Ratio of capital to total expenditure in agriculture
Total agri expenditure (Rs. million/hect)	Total capital & revenue expenditure in agriculture
Per capita state GDP	Per capita State GDP

Appendix 2: Summary table Visit 1

Variable	Obs	Mean	Std. Dev.	Min	Max
FARM					
Number of on-farm diversification activities	25,021	2.813	1.464	1	9
Proportion of land under irrigation	25,021	0.493	0.479	0	1
Dummy insurance	25,021	0.061	0.239	0	1
Proportion of land under purchased seeds	25,021	0.670	0.440	0	1
Total land	25,021	1.575	1.965	0	65.92
Dummy human labour	25,021	0.783	0.412	0	1
Dummy fertilizer	25,021	0.844	0.363	0	1
Dummy electricity	25,021	0.130	0.336	0	1
Dummy machine repair	25,021	0.994	0.080	0	1
Dummy machine hire	25,021	0.993	0.082	0	1
Dummy cattle seed	25,021	0.044	0.205	0	1
Dummy goat seed	25,021	0.019	0.136	0	1
Dummy poultry seed	25,021	0.026	0.158	0	1
Dummy other seed	25,021	0.004	0.060	0	1
Dummy green fodder	25,021	0.523	0.499	0	1
Dummy dry fodder	25,021	0.644	0.479	0	1
Dummy veterinary	25,021	0.178	0.382	0	1
Dummy outstanding credit	25,021	0.579	0.494	0	1
Crops sale under MSP (No.)	25,021	1.542	1.072	0	4
Livestock market sale (No.)	25,021	0.050	0.217	0	1
Dummy tech advice (AU)	25,021	0.017	0.129	0	1
Dummy tech advice (PF)	25,021	0.197	0.398	0	1
Dummy tech advice (RN)	25,021	0.236	0.424	0	1
HOUSEHOLD					
Dummy head male	25,021	0.936	0.244	0	1
Dummy head illiterate	25,021	0.376	0.484	0	1
Household size	25,021	5.531	2.818	1	41
Dummy SC/ST household	25,021	0.281	0.449	0	1
Avg household age	25,021	31.131	11.548	9.286	95
SD household age	25,021	17.545	6.402	0	47.38
Proportion of dependents	25,021	0.480	0.244	0	1
Male female ratio	25,021	0.520	0.159	0	1
Avg education (adults in agri)	25,021	13.919	10.970	0	169
Proportion of principal non-farm workers	25,021	0.512	0.333	0	1
Proportion of MGNREGA workers	25,021	0.033	0.179	0	1
Monthly per capita consumption	25,021	1340.100	1484.3	50	1E+05

Appendix 2: Summary table Visit 1 (*contd.*)

Variable	Obs	Mean	Std. Dev.	Min	Max
VILLAGE					
Village proportion of non-farm workers	25,021	0.456	0.346	0	0.875
Relative price of crops	25,021	4.754	39.805	0	2742
Relative price of animal husbandry	25,021	0.243	1.856	0	149.3
DISTRICT					
(i) Infrastructure					
Proportion of SC	25,021	0.188	0.089	0.01	0.553
Proportion of ST	25,021	0.110	0.169	0	0.934
Proportion with middle and high schools	25,021	0.210	0.133	0	0.882
Proportion with colleges within 5 kms	25,021	0.089	0.068	0	1
Proportion with towns within 5kms	25,021	0.212	0.217	0	1
Proportion with public transport	25,021	0.976	0.113	4E-04	1
Avg number of village manufactured goods	25,021	0.174	0.311	0	3.036
Proportion with SHG	25,021	0.672	0.272	0	1
Proportion with vet hospital	25,021	0.069	0.103	0	1
Proportion of NSA to total district area	25,021	0.563	0.205	0.042	0.981
Proportion of fallow land to total district area	25,021	0.326	0.910	0	24.79
Avg power available for agri use (hrs./day)	25,021	6.805	4.908	0	24
Ratio of ground and well irrig to total NSA	25,021	0.332	0.284	0	1
Proportion of agri credit to total credit	25,021	0.377	0.189	0.02	0.84
(ii) Urban agglomeration					
Class 1 cities	25,021	1.057	0.988	0	6
Class 2 cities	25,021	1.041	1.243	0	7
Class 3, 4, 5, 6 cities	25,021	10.876	12.886	0	96
Ratio of class 1 and 2 cities to total cities	25,021	0.200	0.172	0	1
STATE					
Ratio of capital to total agri expenditure	25,021	0.182	0.169	0.004	0.687
Total agri expenditure (Rs. million/hect)	25,021	3174.821	2031.095	1035	11170
Per capita state GDP	25,021	81061.300	40847.610	32825	2E+05

Appendix 2: Summary table Visit 2

Variable	Obs	Mean	Std. Dev.	Min	Max
FARM					
Number of on-farm diversification activities	25,021	2.652	1.890	0	9
Proportion of land under irrigation	25,021	0.506	0.488	0	1
Dummy insurance	25,021	0.030	0.171	0	1
Proportion of land under purchased seeds	25,021	0.505	0.470	0	1
Total land	25,021	1.575	1.965	0	65.923
Dummy human labour	25,021	0.611	0.488	0	1
Dummy fertilizer	25,021	0.630	0.483	0	1
Dummy electricity	25,021	0.150	0.357	0	1
Dummy machine repair	25,021	0.792	0.406	0	1
Dummy machine hire	25,021	0.792	0.406	0	1
Dummy cattle seed	25,021	0.000	0.006	0	1
Dummy goat seed	25,021	0.027	0.163	0	1
Dummy poultry seed	25,021	0.014	0.118	0	1
Dummy other seed	25,021	0.020	0.141	0	1
Dummy green fodder	25,021	0.004	0.064	0	1
Dummy dry fodder	25,021	0.611	0.488	0	1
Dummy veterinary	25,021	0.154	0.361	0	1
Dummy outstanding credit	25,021	0.579	0.494	0	1
Crops sale under MSP (No.)	25,021	1.315	1.170	0	4
Livestock market sale (No.)	25,021	0.052	0.222	0	1
Dummy tech advice (AU)	25,021	0.016	0.124	0	1
Dummy tech advice (PF)	25,021	0.160	0.367	0	1
Dummy tech advice (RN)	25,021	0.214	0.410	0	1
HOUSEHOLD					
Dummy head male	25,021	0.924	0.265	0	1
Dummy head illiterate	25,021	0.378	0.485	0	1
Household size	25,021	5.531	2.818	1	41
Dummy SC/ST household	25,021	0.281	0.449	0	1
Avg household age	25,021	31.131	11.548	9.285714	95
SD household age	25,021	17.545	6.402	0	47.37616
Proportion of dependents	25,021	0.485	0.245	0	1
Male female ratio	25,021	0.520	0.159	0	1
Avg education (adults in agri)	25,021	13.919	10.970	0	169
Proportion of principal non-farm workers	25,021	0.471	0.340	0	1
Proportion of MGNREGA workers	25,021	0.019	0.099	0	1
Monthly per capita consumption	25,021	1638.973	2210.126	0	236500
VILLAGE					
Village proportion of non-farm workers	25,021	0.502	0.338	0	0.9375
Relative price of crops	25,021	2.095	10.351	0	588.4438
Relative price of animal husbandry	25,021	0.273	1.594	0	104.8951

Table 5: SUR & 3SRI ordered probit results

	SUR		3SRI	
	Visit 1	Visit 2	Visit 1	Visit 2
FARM				
Proportion of land under purchased seeds	0.119***	0.274***		
	(3.30)	(6.37)		
Control function			-18.26***	-20.89***
			(-24.44)	(-12.83)
Proportion of land under irrigation	-0.100**	-0.0822	0.00921	-2.277***
	(-2.85)	(-1.95)	(0.24)	(-12.27)
Dummy insurance	0.0455	0.132	-0.571***	1.209***
	(0.61)	(1.44)	(-6.74)	(9.58)
Total land	0.00552	0.0100	-0.0803***	0.0286*
	(0.51)	(0.81)	(-7.32)	(2.34)
Square total land	-0.000341	-0.000388	0.00240***	0.00112***
	(-1.22)	(-1.40)	(8.15)	(3.66)
Dummy human labour	0.0530	0.126**	-0.110**	-0.205***
	(1.32)	(3.08)	(-2.60)	(-4.04)
Dummy fertilizer	-0.137**	-0.184***	-4.758***	-3.615***
	(-3.00)	(-3.62)	(-23.95)	(-13.56)
Dummy electricity	0.128**	-0.0681	-0.443***	-0.0617
	(2.67)	(-1.35)	(-8.69)	(-1.13)
Dummy machine repair	-0.189	1.598***	-1.063***	0.988**
	(-0.77)	(4.82)	(-4.06)	(3.21)
Dummy machine hire	-0.00860	1.521***	1.966***	-1.960***
	(-0.04)	(4.03)	(7.89)	(-4.26)
Dummy cattle seed	0.187*	0.398**	-0.735***	-5.045***
	(2.23)	(3.00)	(-8.45)	(-10.96)
Dummy goat seed	0.0285	-0.204*	0.304***	-0.983***
	(0.35)	(-2.01)	(3.60)	(-8.18)
Dummy poultry seed	0.0276	0.218*	-0.948***	0.827***
	(0.37)	(2.10)	(-10.19)	(7.30)
Dummy other seed	0.0371	0.297**	-0.805***	0.598***
	(0.21)	(3.23)	(-4.75)	(6.35)
Dummy green fodder	-0.0346	0.0598	0.102*	0.116
	(-0.89)	(0.29)	(2.54)	(0.61)
Dummy dry fodder	0.0925*	0.220***	-1.117***	-0.920***
	(1.99)	(4.46)	(-15.95)	(-9.12)
Dummy veterinary	-0.0371	0.156***	0.275***	-0.455***
	(-0.93)	(3.36)	(6.62)	(-6.70)
Dummy outstanding credit	1.305***	0.728***	-0.676***	-0.846***
	(18.43)	(7.79)	(-6.14)	(-5.96)

Appendix 3: SUR & 3SRI ordered probit results (*contd.)*

	SUR		3SRI	
	Visit 1	Visit 2	Visit 1	Visit 2
Crops sale under MSP	0.0314	0.143*	-0.145**	-0.790***
	(0.70)	(2.05)	(-3.28)	(-7.58)
Square of crops sale under MSP	0.0559***	-0.00920	0.172***	0.273***
	(4.99)	(-0.57)	(13.99)	(9.81)
Livestock market sale	-0.0360	0.144*	-1.060***	0.492***
	(-0.55)	(2.07)	(-13.82)	(6.34)
Dummy tech advice (AU)	0.239*	-0.146	1.482***	-2.120***
	(2.13)	(-1.38)	(11.54)	(-10.86)
Dummy tech advice (PF)	-0.112**	0.0169	-0.647***	-0.846***
	(-2.83)	(0.39)	(-15.02)	(-10.41)
Dummy tech advice (RN)	0.128***	0.0717	1.067***	0.807***
	(3.31)	(1.83)	(19.94)	(11.47)
HOUSEHOLD				
Dummy head male	-0.0819	-0.00187	-0.145*	-0.245***
	(-1.24)	(-0.03)	(-2.18)	(-3.43)
Dummy head illiterate	0.0546	0.0706	0.241***	0.0526
	(1.57)	(1.83)	(6.74)	(1.31)
Household size	0.00933	-0.00957	-0.0145	-0.0455***
	(1.18)	(-1.14)	(-1.86)	(-5.06)
Dummy SC/ST household	0.0178	-0.0226	0.330***	0.0326
	(0.49)	(-0.49)	(8.08)	(0.69)
Avg household age	0.00200	0.00273	0.00933***	-0.00996***
	(1.10)	(1.38)	(5.46)	(-4.14)
SD household age	-0.00695*	-0.00120	-0.0416***	0.0559***
	(-2.49)	(-0.39)	(-14.05)	(9.11)
Proportion of dependents	0.0448	0.0468	-0.504***	0.120
	(0.46)	(0.62)	(-5.01)	(1.55)
Male female ratio	0.00954	-0.0804	0.223*	0.0392
	(0.10)	(-0.77)	(2.26)	(0.37)
Avg education (adults in agri)	0.00125	0.00259	0.0193***	-0.00865*
	(0.36)	(0.76)	(5.34)	(-2.53)
Proportion of principal non-farm workers	-0.00155	-0.00688	0.0372	0.262***
	(-0.03)	(-0.13)	(0.73)	(4.46)
Proportion of MGNREGA workers	0.0647	0.0989	-0.292***	-1.616***
	(0.84)	(0.53)	(-3.55)	(-6.94)
Monthly per capita consumption	-1.1E-06	0.0000134	-0.0000808***	0.0000552***
	(-0.07)	(0.92)	(-5.38)	(3.38)
Square of monthly per capita consumption	1.70e-10	-2.80e-11	1.44e-10	-3.95e-10***
	(1.27)	(-0.39)	(1.17)	(-4.81)

Appendix 3: SUR & 3SRI ordered probit results (*contd.)*

	SUR		3SRI	
	Visit 1	Visit 2	Visit 1	Visit 2
VILLAGE				
Village proportion of non-farm workers	0.0225	0.0321	-0.208***	0.153**
	(0.45)	(0.57)	(-3.93)	(2.59)
Relative price of crops	-0.000356	-0.000186	-0.00330***	-0.0303***
	(-1.38)	(-0.12)	(-11.86)	(-10.85)
Relative price of animal husbandry	0.00894	0.0241**	-0.0547***	-0.0257**
	(1.48)	(2.82)	(-7.05)	(-2.84)
DISTRICT				
(i) Infrastructure				
Proportion of SC	-0.351	-0.304	-3.244***	-6.255***
	(-1.31)	(-0.94)	(-10.80)	(-11.80)
Proportion of ST	-0.330*	-0.525***	-2.644***	-1.372***
	(-2.39)	(-3.34)	(-15.34)	(-8.07)
Proportion of NSA to total district area	0.264*	0.694***	0.243	1.398***
	(2.03)	(4.85)	(1.78)	(9.16)
Proportion of fallow land to total district area	-0.00643	0.0328	0.113***	0.377***
	(-0.26)	(1.61)	(4.30)	(11.40)
Proportion with middle and high schools	-0.932***	-0.410	1.072***	1.844***
	(-5.16)	(-1.85)	(5.29)	(6.85)
Proportion with colleges within 5 kms	-0.978	16.85***	0.491	22.67***
	(-0.36)	(4.06)	(0.18)	(5.22)
Proportion with towns within 5kms	-1.376*	-3.614***	-13.66***	-5.777***
	(-2.31)	(-4.24)	(-17.66)	(-6.05)
Proportion with public transport	-0.532	0.569	-6.337***	0.866*
	(-1.49)	(1.43)	(-15.58)	(2.01)
Public transport*towns within 5 kms	1.281*	3.671***	13.53***	8.668***
	(2.00)	(4.14)	(16.78)	(8.28)
Proportion with vet hospital	0.660*	-0.161	0.702*	0.385
	(2.40)	(-0.47)	(2.56)	(1.12)
Proportion with SHG	1.162**	-0.239	-1.924***	3.019***
	(2.76)	(-0.54)	(-4.25)	(5.88)
Avg number of village manufactured goods	0.0973	0.146	1.279***	1.091***
	(0.48)	(0.85)	(6.35)	(5.70)
Square avg no. of vill manufactured goods	-0.451***	-0.196*	-0.513***	-0.292***
	(-5.39)	(-2.50)	(-6.25)	(-3.79)
Proportion SHG*avg manufactured goods	1.010***	0.334	0.450	-0.673**
	(4.53)	(1.46)	(1.96)	(-2.72)

Appendix 3: SUR & 3SRI ordered probit results (*contd.)*

	SUR		3SRI	
	Visit 1	Visit 2	Visit 1	Visit 2
Proportion of agri to total credit	-0.849***	0.177	2.480***	2.322***
	(-3.33)	(0.75)	(8.96)	(7.64)
Proportion SHG*prop agri to total credit	0.518	-0.585	-2.840***	-3.634***
	(1.36)	(-1.63)	(-7.14)	(-7.89)
Avg power for agri use (hrs./day)	0.0240*	-0.104***	-0.0177	0.247***
	(2.04)	(-8.30)	(-1.47)	(8.14)
Square avg power for agri use (hrs./day)	-0.00200**	0.00478***	0.00395***	-0.00616***
	(-3.25)	(7.69)	(5.77)	(-5.91)
Ratio of ground and well irrg to NSA	-0.217*	-0.182	0.818***	1.318***
	(-2.20)	(-1.79)	(7.61)	(8.27)
Rainfall deviation	-0.00126**	0.0000852***	0.00207***	0.000483***
	(-2.63)	(4.16)	(4.00)	(12.67)
(ii) Urban agglomeration				
Class 1 cities	-0.144**	-0.0210	-0.0329	-0.533***
	(-2.93)	(-0.38)	(-0.67)	(-7.36)
Square of class 1 cities	0.0195*	0.00360	-0.145***	0.143***
	(2.35)	(0.37)	(-13.97)	(9.41)
Class 2 cities	-0.0730	0.0486	-0.622***	0.204***
	(-1.81)	(1.16)	(-14.21)	(4.41)
Square of class 2 cities	0.00507	-0.00655	0.102***	-0.0441***
	(0.78)	(-0.92)	(14.38)	(-5.32)
Class 3, 4, 5, 6 cities	0.00251	0.0187**	0.0874***	-0.107***
	(0.42)	(3.14)	(13.10)	(-9.10)
Square of class 3, 4, 5, 6 cities	0.000	-0.0002***	-0.0003***	0.0009***
	(0.10)	(-3.46)	(-6.66)	(8.70)
Ratio of class 1 and 2 cities to total cities	0.488*	0.366	1.457***	-0.0442
	(2.52)	(1.87)	(6.93)	(-0.22)
STATE				
Ratio of capital to total agri expenditure	-0.476	1.507***	-2.325***	4.410***
	(-1.48)	(4.73)	(-6.74)	(10.77)
Total agri expenditure (Rs. million/hect)	-0.000219**	0.000142	-0.000571***	0.00229***
	(-2.75)	(1.65)	(-6.94)	(11.89)
Square total agri expenditure	2.00e-08**	-8.22e-09	7.51e-08***	-2.00e-07***
	(2.65)	(-1.03)	(9.44)	(-11.91)
Per capita state GDP	-0.00002**	0.00	0.00	0.0002***
	(-2.68)	(0.14)	(0.61)	(11.18)

Appendix 3: SUR & 3SRI ordered probit results (*contd.)*

	SUR		3SRI	
	Visit 1	Visit 2	Visit 1	Visit 2
Square per capita state GDP	7.16e-11*	-3.45e-11	-5.04e-11	-1.02e-09***
	(2.54)	(-0.99)	(-1.70)	(-11.47)
District rainfall deviation	*Included*	*Included*	*Included*	*Included*
District soil moisture availability index	*Included*	*Included*	*Included*	*Included*
Proportion of 19 soil types in the district	*Included*	*Included*	*Included*	*Included*
N	25021	25021	25021	25021
Log pseudolikelihood	-2079295.6		-2025940.3	
ρ	0.05207		0.05785	
AIC	4158983		4052271	
BIC	4160576		4053855	
LR test of independent equations	-4.0e+06		-3.9e+06	
	(Prob>chi2=1)		(Prob>chi2=1)	

Note: The first two columns are SUR estimation results, without endogeneity correction. The last two columns are endogeneity corrected 3SRI estimation results. The *control functions* for proportion of land under purchased seeds for each seasons are the share of households in the village growing exactly the same crops. Technical advice from AU refers to agricultural university, PF refers to progressive farmers, and RN refers to radio, newspaper, TV etc. Among the district level variables, average power availability for agricultural use and district rainfall deviation are available for *kharif* and *rabi* seasons separately. 19 different soil types were considered in this study (as in ICRISAT data). The proportion of land area in each district for each of these 19 different soil types were calculated.
Robust Z statistics in parentheses.
$*p < 0.1$, $**p < 0.05$, $***p < 0.001$