WORKING PAPER NO: 630

Seasonal Migration and Feminization of Farm Management: Evidence from India

S Chandrasekhar
Professor
Indira Gandhi Institute of Development Research
Gen AK Vaidya Marg, Goregaon East,
Mumbai 400065, India
chandra@igidr.ac.in

Soham Sahoo
Assistant Professor
Centre for Public Policy
Indian Institute of Management Bangalore
Bannerghatta Road, Bangalore – 560076
soham.sahoo@iimb.ac.in

Hema Swaminathan
Associate Professor
Centre for Public Policy
Indian Institute of Management Bangalore
Bannerghatta Road, Bangalore – 560076
hema.swaminathan@iimb.ac.in

Year of Publication- December 2020

Acknowledgements: S Chandrasekhar and Hema Swaminathan are grateful for support from the research initiative "System of Promoting Appropriate National Dynamism for Agriculture and Nutrition" (SPANDAN) which is supported by Global Development Grant Number OPP1029478 from Bill and Melinda Gates Foundation. Soham Sahoo acknowledges funding from the Growth and Economic Opportunities for Women (GrOW) initiative, a multi-funder partnership between the UK's Department for International Development, the Hewlett Foundation and the International Development Research Centre

Seasonal Migration and Feminization of Farm Management: Evidence from India

Abstract

Using gender-disaggregated data on land operations from India, this paper contributes to the literature on agricultural feminization by showing a relationship between seasonal or short-term migration for work and feminization of farm management. Using a nationally representative data-set covering 35,604 rural Indian households in 2013, we identify if women are taking on the role of farm managers in households with short-term migrants. Our results show that women are less likely than men to be decision makers on farms, but this dynamic changes when there is short-term migration in the household; the probability of women being decision makers on farms increases. These results are robust to concerns over omitted variables, endogeneity and sample selection issues. Our study highlights the importance of unpacking the feminization process to better understand the role of women as farm managers and the need for supporting this transition to ensure women farmers realise their full potential.

Keywords: Feminization of agriculture; Female farm managers; Seasonal migration; Agricultural households; Operational holdings; India

JEL Codes: Q1, R2, J1

1. Introduction

Using gender-disaggregated data on land operations from India, this paper contributes to the literature on agricultural feminization by showing a relationship between seasonal or short-term migration for work and feminization of farm management. The focus in this study is on managerial feminization, in contrast to feminization of the agricultural labor. In the last two decades, studies from diverse geographies have examined how women's labour market outcomes have responded to the largely male dominated processes of the expansion of non-farm opportunities (Weiliang Su, Tor Eriksson, Linxiu Zhang, Yunli Bai 2016), and out-migration (Christine lel and Ragui Assaad 2011, Mariapia Mendola and Calogero Carletto 2012) and if this has resulted in feminization of agriculture. The evidence is not conclusive (see review in Vanya Slavchevska, Susan Kaaria, and Sanna-Liisa Taivalmaa 2016), but often there is no clear distinction between women's agricultural labor and agricultural decision making.

Despite evidence from around the world that the contribution of women to agriculture is substantive, information on agricultural decision making within a household, and in particular by women, is generally not collected in agricultural surveys. As per India's Agricultural Census, the share of women operational holders has steadily increased from 12.79 per cent in 2010-11 to 13.96 per cent in 2015-16 and their share in operated area from 10.36 per cent to 11.72 per cent.[1]

The contribution of our study to the literature is twofold. First, we establish how the presence of a short-term migrant increases the probability of women being associated with the operational holding as decision makers. The importance of short-term migration affecting agricultural processes is particularly relevant for India, since the number of seasonal migrants is

[1] For cross country comparisons, FAO's Gender and Land Rights Database provides information on "land-related statistics disaggregated by gender, including the share of men and women who are agricultural holders". http://www.fao.org/gender-landrights-database/en/

significantly larger than permanent migrants in any year and it is an integral part of a household's livelihood strategy in rural India (Kunal Keshri and Ram B. Bhagat 2013). Since seasonal migrants in India are more likely to be men[2] rather than women (Tushar Agrawal and S Chandrasekhar 2016) it is but natural that this changes the roles of women who stay behind. It is in this context that data from National Sample Survey Organisation's (NSSO) Survey on Land Livestock Holdings (2013) is informative. This survey is unique in that, for the first time, information on both short-term migration and association of household members with operational agricultural holdings was collected. The survey identifies individuals as main operators (major decision maker on farm), associated operators or not associated with the holding at all. Thus, we are able to identify if women are indeed taking on the role of farm managers in households with short-term migrants. This is a superior measure of managerial feminization to the current ones in the literature that use female-headed households as a proxy for managerial feminization (Alan de Brauw, Jikun Huang, Linxiu Zhang & Scott Rozelle 2013). Second, we show that the effect of short-term migration on managerial feminization is heterogeneous and depends on household, individual and farm characteristics. Our finding highlights the importance of recognizing heterogeneity among women; an issue that was underscored by Mendola and Carletto (2012) who found that the effect of male outmigration in Albania to be different across sub-groups of women.

Our results from an ordered probit model show that women are less likely than men to be decision makers on farms, i.e., characterized as main or associated operator. But this dynamic changes when there is short-term migration in the household; the probability of women being

[2] Beginning 2004-05, the spurt in short-term internal migration has been driven by a boom in the construction industry and a majority of workers in this sector are men (Agrawal and Chandrasekhar 2016). Estimates from NSSO's Situation Assessment of Agricultural Households conducted in 2013 shows that nearly 83 per cent of short-term migrants are men. Among these, 46 per cent are household heads and 49 per cent are sons of the household head (Government of India 2014b)

decision makers on farms increases. These findings are robust to corrections for sample selection and endogeneity issues. We specifically show that any effect of omitted variables is unlikely to overturn our results. We also address the issue of reverse causality by using an instrumental variable model.

The changing role of women in agriculture is significant for agricultural policy makers. Agricultural policies and programs instead of being gender blind need to recognize that women farmers are also relevant stakeholders. The Government of India has acknowledged that "more and more females are participating in the management and operation of agricultural lands" (Government of India 2019). As women transition from being laborers to decision makers on their farms, the question that naturally arises is how best to support this transition to ensure that these women farmers can realise their full potential (Bina Agarwal 2018).

2. Related Literature

The literature on feminization of agriculture is marked by a lack of a standard definition of what constitutes feminization. Using data from four Asian countries, Caitlin Kieran, Kathryn Sproule, Cheryl Doss, Agnes Quisumbing, Sung Mi Kim (2015) discuss the challenges in constructing sex-disaggregated land indicators and decision making. The contribution by de Brauw et al (2013) and Kanika Mahajan (2019) in context of China and India, respectively, are of relevance to this paper since they discuss feminization of farm management. In absence of detailed data on decision making, the China study uses female headship as a proxy for managerial feminization in China. The India study improves on this aspect by considering a woman to be the farm manager if she is the primary decision maker regarding farm matters according to the household; however, this study does not analyze the role of migration.

The larger literature on expansion of non-farm opportunities for men and consequences of seasonal migration by men has focussed on changes in women's labor supply, work roles and changing responsibilities. These changes are mediated by several factors such as household socio-economic status, structure, location of the household as well as the social context that can play an important role in determining women's work and their participation in decision making.

The literature does not suggest a consistent pattern with respect to the impact of migration on women's work allocations. Recent studies from China explore the association between migration and feminization of agriculture and present mixed evidence, partly due to differences in the period under consideration. In examining the role of women in agriculture in the 1990s, Alan de Brauw, Qiang Li, Chengfang Liu, Scott Rozelle, and Linxiu Zhang (2008) find little evidence of either managerial or labor feminization in agriculture due to migration. While the study does not find women to be replacing men in agriculture, they suggest that in the livestock sector, there is some trend towards this. On the other hand, Chang, Hongqin; MacPhail, Fiona; and Dong, Xiao-yuan (2011) examine trends in rural China over a period of 15 years, 1991-2006 and find that migration is an important driver of feminization of agriculture for both farm and off-farm work. This trend is partially confirmed by Ren Mu and Dominique van de Walle (2011) for the 2000s with women's participation and work allocation increasing in farm work while reducing their off-farm activities, either paid or unpaid. Based on descriptive evidence, de Brauw et al (2013) corroborate that the proportion of farm work undertaken by women in China increased from 53 per cent in 1997 to 59 per cent in 2009. In the context of Egypt, Binzel and Assad (2011) find that in rural areas, women in migrant households are likely to decrease wage employment in order to accommodate a greater degree of subsistence and unpaid family work, mainly to substitute for men's labor in agriculture. This effect though, is not uniform across all age categories of women. The shift from paid to unpaid activities is also found in Albania for women when there is a current migrant in the household (Mendola and Carletto 2012). For households with a prior migrant, women were able to move from unpaid work to self-employment suggesting a role for migrant income in providing capital for entrepreneurship.

There have been fewer studies in the Indian context; further they are small sample and focus on high migration areas. A study of outmigration among rice farmers in eastern Uttar Pradesh found that women from migrant households, particularly those from nuclear households felt burdened with their increased workload (Thelma R Paris, Abha Singh, Joyce Luis and Mahabub Hossain 2005). Contrarily, a recent study in one district in West Bengal does not support feminization of agriculture (L Schenk-Sandbergen, 2018). While there is distress induced male migration, it is seasonal in nature and it has not changed gendered roles and responsibilities.

3. Data and Descriptive Statistics

We use data from NSSO's Survey on Land and Livestock Holdings 2013. The survey of 35,604 rural households covering 178,246 individuals, of whom 117,382 were in the working age of 15-65 years, was conducted in January-July 2013 (See Government of India (2014a) for details). Of the respondents, 66 per cent and 20 per cent are head of the household or spouse of the household head, respectively; nearly 10 per cent are the children of household head and 2 per cent are the spouse of the child of the household head.

The outcome variable of interest is whether a household member is associated with the household operational holding as the main operator or as another member associated with the household operational holding, or not associated with household operational holding. The main operator is the individual who takes the major decisions regarding the household operational holding over a reference period of the last 365 days. If no single individual takes major decisions, the senior most operator is identified as the main operator. Akin to other studies, a limitation is the absence of information on what these decisions are.

In 2013, we estimate that 11.8 million women and 85 million men in the age group 15-65 years are main operators. We also estimate that 104 million women and 58.7 million men in the above age group are associated operators.

A household's operational holding refers to all land (owned, leased, or possessed in some other form) where the household has undertaken some agricultural activity during the reference period. Of the estimated 156 million rural households, 68 per cent report operating land for agriculture of which 9.06 million or 9.5 per cent are female headed households. Households that are actively engaged in agriculture have more male members; on average the head is marginally older and also more likely to be better educated.

Of particular interest to this paper is the presence of short-term migrant in the household and how this affects an individual's operator status. Information is sought on whether any member of the household stayed away from the village continuously for 15 days or more for employment during last 6 months. This question refers to household members and thus, does not include permanent migrants. We estimate that 10.08 million rural households (6.5 per cent) have a short-term migrant. We highlight few key differences across households with and without a short-term migrant (Appendix Table 1). Heads of households with short-term migrants are on average younger, and with relatively lower educational qualifications (except that they are more literate) than heads of households without a short-term migrant. It is possible though that the migrant could be better educated. Migrants belong to households that possess significantly less land than other households.

4. Empirical Methods

We analyse an ordered outcome variable capturing the degree of association of each individual, in the working age population *i.e.,* those aged 15–65 years, with the household operational holding. This outcome of interest (*j*) takes one of three values: 0 if the individual is not associated with the operational holding, 1 if otherwise associated, and 2 if the individual is the main operator of the holding. The final analytic sample comprises of 92,376 individuals from households operating land.

Since the outcome variable represents a ranking of the individual's participation in agriculture, an ordered probit model is estimated with the operator status as a function of individual, household, operational land, and region level characteristics.

$$\begin{aligned} Pr(Associated_{ihds} = j) \\ = F(\beta_0 + \beta_1 STM_{hds} + \beta_2 Female_{ihds} + \beta_3 STM_{hds} \times Female_{ihds} \\ + \beta_4 X_{ihds} + \beta_5 Z_{hd} + \beta_6 V_{ds} + \mu_s) \\ j = \{0, 1, 2\} \end{aligned} \quad (1)$$

where *i*, *h*, *d*, and *s* denote individual, household, district and state, respectively. The dependent variable ($Associated_{ihds}$) takes the value 0, 1, or 2 based on operator status (*j*). The variable STM_{hds} represents whether the household has any short-term migrant ($STM_{hds} = 1$) or not ($STM_{hds} = 0$). The gender of the individual is captured by the dummy variable $Female_{ihds}$, which takes the value 1 or 0 indicating whether the individual is female or male. We are mainly interested in the interaction between STM_{hds} and $Female_{ihds}$. The coefficient of the interaction term (β_3) reflects the gender-specific difference in the effect of STM_{hd} on an individual's association with the household operational holding. If we hypothesize that the presence of short-term migrant has a positive effect on women's association with the operational holding of the household, then we would expect β_3 to be positive and significant.

The list of individual (X_{ihds}), household (Z_{hds}), and village-district (V_{ds}) level controls is given in Table 1. In the regression model, we include state fixed effects (μ_s) to allow for unobserved inter-state differences.

We briefly discuss select control variables. The construction of the individual and household variables is consistent with the Indian context. In addition to gender, the individual specific characteristics we include are age, age square, education, and relationship to head. Household level controls can be grouped into following categories: those based on demographic particulars, main source of income, and details of land possessed. Agricultural households in India are categorized based on their land-possession into marginal farmers (less than 0.4 hectare, 0.4–1 hectare), small farmers (1–2 hectares), semi-medium farmers (2–4 hectares), medium farmers (4 – 10 hectares) and large farmers (10 hectares and above).

We control for several push factors of migration at the village and district levels. We include a sub-round dummy to account for the possibility that short-term migration that occurred during the recall period of the survey, may have been affected by the seasonality of crops. Also, crop diversification helps farmers to spread the uncertainty of agricultural production over a wider portfolio of crops reducing their income volatility. Therefore, to capture any effect of cropping patterns on short-term migration, we include the Gibbs-Martin index of crop diversification at the village level. Consistent with the literature, for each household, the village-level index is calculated by excluding information from that particular household (Pratap S Birthal. Devesh Roy and Digvijay S Negi 2015). This approach obviates any endogeneity issues pertaining to household level decisions.[3] In order to control for factors affecting a district's agricultural production we control for deviation in rainfall from the long-term average rainfall, the percentage of land that is not irrigated, and the percentage of households having *Kisan Credit Card* (farmer credit card) with a credit limit of Rs. 50,000 or above. The last two variables are sourced from the Socio-Economic and Caste Census 2011.

Additional push factors we include are district rural unemployment rate and share of agricultural workers among all marginal workers in rural areas of the district. These variables are sourced from Census 2011 data, and they are lagged with respect to the decision of short-term migration in our main data. Economic theory suggests that a higher rate of rural unemployment would increase the incidence of short-term out-migration from villages. A higher share of agricultural workers among all marginal workers is also likely to cause distress-driven out-migration. In addition to these, we also control for the overall economic development. Recent studies have used satellite data on night-time lights as a proxy for regional economic growth and development. We use night-lights averaged over a district and estimate the growth rate over 2007–2011.[4]

[3]As a robustness check, we also estimate different regressions that control for the total number of crops, share of area of land used to cultivate different types of crops, and dominant crop dummies at the village level. The results are unchanged even after controlling for these different measures of cropping patterns in the village. While measuring these variables for each household, we exclude that household and calculate the average over all other households in the village. In fact, the main findings are unperturbed even if we include the measure at the household level ignoring possible endogeneity of these variables. These results are not shown but available on request.

[4] This data is recorded for every one square kilometre area worldwide by the Operational Linescan System (OLS) flown on the Defense Meteorological Satellite Program (DMSP) satellites. We downloaded the data from National Oceanic and Atmospheric Administration (NOAA, https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html) and matched it at the district level.

Identification

Since migration and labor supply decisions are likely to be determined simultaneously within the household, short-term migration is endogenous to household members' participation in agricultural activities. To address this problem, we follow two econometric approaches.

First, we use a method proposed by Emily Oster (2017) to examine the extent of omitted variable bias. This method helps us evaluate whether the endogeneity problem is so severe that it would nullify the estimated effect, within a framework which assumes that selection based on unobservables is proportional to selection based on observable covariates (J G Altonji, T E Elder, & C R Taber 2005). This method is applied on a linear probability model where we also include household fixed effects. The specification includes all individuals and the interaction between STM_{hds} and $Female_{ihds}$. Although household fixed effects subsume the effect of STM (which is a household-level variable), we are still able to identify the differential effect by gender since the interaction term has variation within a household. Inclusion of household fixed effects allows us to control for any unobservable factors at the household as well as regional level, e.g. effect of any social program in the area which may affect both migration and women's agricultural participation.

Our second strategy uses an instrumental variable framework. We use the following two variables and their interaction as instruments for STM: (a) lagged manufacturing employment in the rest of the state excluding the district where the household resides ($Manufacturing_{-ds}$), and (b) lagged rate of short-term migration in the district ($Migrantrate_{ds}$). The first variable is calculated from the NSSO survey of employment and unemployment (2004-05) while the latter from NSSO's survey of employment, unemployment and migration (2007-08).

The logic behind the first instrument is similar to the study by Francisca M. Antman (2011) where employment levels in the relevant industry at the probable destination of the migrant were used as instruments for migration. In India, a significant proportion of short-term migrants find work in other districts within the same state. We posit that a large size of the manufacturing sector employment in other districts would act as a pull factor for migration from a given district. Since this measure is constructed for each district by excluding the particular district where the household under consideration resides, it is exogenous to the household. Besides, economic conditions in other districts are unlikely to have a direct effect on women's participation in agriculture in the home district, apart from the route of migration or other factors included in the model. Thus, this variable reflects the availability of work in the destination districts and serves as a valid instrument.

The idea behind the second instrument is that a stronger migrant network exists in districts that experienced higher share of migration in the past, which could influence the current short-term migratory flows in the district (Mariapia Mendola 2012). In the existing literature, similar region-level rates of migration (preferably lagged) have been widely used as an instrument for individual and household level decision to migrate (Michael Lokshin and Elena Glinskaya 2009; Binzel and Assaad 2011; Mendola and Carletto 2012). Since we use a lagged measure of short-term migration rate in the district, it cannot be affected by current migration or labor supply decisions. Moreover, any contemporaneous relationship between this variable and an individual member's participation in agriculture can be ruled out. In a robustness exercise, we further include a control variable measuring the past share of women working in agriculture in the district and find that the results are unchanged.

As a third instrument, we include the interaction between the above two instruments. Our hypothesis is that employment opportunities in nearby districts would have a greater impact when there is a pre-existing migrant network in the district, thus the effects of the two instruments are reinforced by each other.

In addition to STM, our model includes an interaction between STM and Female which is also an endogenous variable. Since Female is exogenous, therefore, it is plausible to interact the instruments with the female dummy and use these interaction terms as additional instruments for STM*Female (Jeffrey M. Wooldridge, 2002).

Therefore, the first stage equations of the Two Stage Least Square (2SLS) model are specified as:

$$\begin{aligned} STM_{ihds} = \eta_0 &+ \eta_2 Female_{ihds} + \eta_2 X_{ihds} + \eta_3 Z_{hds} + \eta_4 V_{ds} \\ &+ \eta_5 Manufacturing_{-ds} + \eta_6 Migrantrate_{ds} \\ &+ \eta_7 Manufacturing_{-ds} \times Migrantrate_{ds} \\ &+ \eta_8 Manufacturing_{-ds} \times Female_{ihds} \\ &+ \eta_9 Migrantrate_{ds} \times Female_{ihds} \\ &+ \eta_{10} Manufacturing_{-ds} \times Migrantrate_{ds} \times Female_{ihds} + \psi_s \\ &+ u_{ihds} \end{aligned} \tag{2}$$

$$\begin{aligned} STM_{hds} * Female_{ihds} \\ &= \theta_0 + \theta_2 Female_{ihds} + \theta_2 X_{ihds} + \theta_3 Z_{hds} + \theta_4 V_{ds} \\ &+ \theta_5 Manufacturing_{-d} + \theta_6 Migrantrate_{ds} \\ &+ \theta_7 Manufacturing_{-ds} \times Migrantrate_{ds} \\ &+ \theta_8 Manufacturing_{-ds} \times Female_{ihds} \\ &+ \theta_9 Migrantrate_{ds} \times Female_{ihds} \\ &+ \theta_{10} Manufacturing_{-ds} \times Migrantrate_{ds} \times Female_{ihds} + \varphi_s \\ &+ v_{ihds} \end{aligned} \tag{3}$$

These two equations are estimated using Ordinary Least Squares (OLS), and the predicted values of the two endogenous variables are then used in the second stage of the estimation, given by Equation (4). The second stage is a linear probability model where the dependent variable is a binary indicator of whether the individual is main operator/associated with operational holding ($Associated_{ihds} = 1$) or not ($Associated_{ihds} = 0$).

$$Associated_{ihds} = \alpha_0 + \alpha_1 \widehat{STM_{hd}} + \alpha_2 Female_{ihds} + \alpha_3 \widehat{STM_{hds} \times Female_{ihds}} + \alpha_4 X_{ihds} + \alpha_5 Z_{hds} + \alpha_6 V_{ds} + \xi_s + \epsilon_{ihds} \quad (4)$$

The linear probability model works better than an ordered probit model in an instrumental variables framework for the following reasons First, there are econometric issues involved in using instrumental variable in an ordered probit framework especially when the endogenous explanatory variable is binary (in our case, STM). Second, literature suggests that it is preferable to estimate a linear model when the main interest is to obtain the marginal effects of the main explanatory variable (Joshua D. Angrist 2001; Joshua D. Angrist and Jörn Steffen Pischke, 2009).

We first estimate the model separately for the female and male sample where the interaction between STM and Female is dropped. We are interested in the sign and magnitude of the coefficient α_1 which reflects how having a STM in the household affects the probability of a male or a female individual being an operator of the land. We then estimate a pooled model and include an interaction term between STM and Female. In this pooled sample, we are interested in the estimate of the coefficient α_3 which captures the differential effect of STM by gender. The estimated standard errors are robust to heteroskedasticity and clustered at the household level. We present results from both these approaches in the following section.

5. Results

a. Ordered Probit Model

The summary statistics of the analytic sample of 92,376 working-age individuals (aged 15–65 years) from households operating land is presented in Table 1. We estimate an ordered probit model (Equation 1) separately for men and women as well as a pooled model; the marginal effects are reported in Table 2. Consistent with our expectations, we see that the presence of a short-term migrant in the household affects women's association with the operational holding. However, for men, a short-term migrant in the household is not a statistically significant determinant of the nature of involvement with operational land holding. In case of women, the probability of their being involved with the operational holding either as a main or associated operator increases by 4 percentage points.

The results from the pooled model reinforce the earlier findings. We find that women, in comparison with men, are less likely to be either associated or be the main operator of the operational holding when there is no short-term migrant in the household. However, the interaction term (STM*Female) is significant across the three outcomes. We find that in a household with a short-term migrant, the probability of a woman not being involved with the operational holding either as a main or associated operator goes down by 4.1 percentage points. This estimate is similar to that obtained in China where the probability of working on a farm is 6 per cent higher for women left behind in migrant households (Mu and van de Walle 2011). Overall, the results suggest that short-term migration is indeed associated with a greater degree of feminization of farm management.

(Table 1)

(Table 2)

Briefly turning to other variables (Appendix Table 2), we see that older men and women are more likely to be associated with the operational holding. Women are less likely to be involved in decision making pertaining to operational land when they are more educated. In contrast, only those men who have completed at least secondary education are less likely to be involved with the operational holding. More educated members are presumably looking beyond the farming sector for employment opportunities. Our finding is consistent with studies that find that more educated farmers report a greater dislike for farming than less educated ones (Pratap S Birthal, Devesh Roy, Md. Tajuddin Khan and Digvijay Singh Negi (2015); Bina Agarwal and Ankush Agrawal 2017). Agarwal and Agrawal (2017) report that among farmers who have completed at least secondary schooling barely 15 per cent report that they like farming. In contrast, among those who are illiterate nearly 48 per cent report that they like farming. Our results suggest that position in the household matters. As heads, men and women are more involved as operators on their land holdings, but not in any other role in the household (such as spouse, child of head and so on). In female headed households, both men and women are more likely to be involved with the operational holding either as associated or main operator, reflecting perhaps, a labor constraint due to the absence of an adult male member. We find a differential impact of the size of land possessed on men and women's operator status. Consistent with the literature, our results suggest that as land size increases, men are more likely to be operators, but not women. Evidence suggests that often, the land owned or operated by women may be smaller in size or of inferior quality when compared with those owned or operated by men (B B Keller, P.E. Chola and M.C. Milimo 1990 as cited in FAO 2011).

b. Assessment of Bias due to Unobservables

To examine the sensitivity of our estimates to the potential effect of unobservables, we estimate a linear probability model where the outcome indicates whether an individual has any association with the operational holding or not. The results from the linear probability models for the coefficients of interest are presented in three separate panels, each for female, male, and overall sample (Table 3). Detailed discussion of the method used is presented in Appendix A1, while the full results are in Appendix Table 3. Column (1) presents the regression with no control variables, while column (2) shows the regression with all control variables including state fixed effects for male and female samples, and household fixed effects for the overall sample. Results from the regressions with control variables corroborate our findings from the ordered probit analysis. Women have 3.9 percentage points higher probability of being associated with the operational holding when the household has a short-term migrant. The effect on men is smaller at 1.4 percentage points, and also relatively less precisely estimated. We further enrich the model for the overall sample by including household fixed effects in the set of control variables. The estimates suggest that women are 18.6 percentage points less likely than men to be an operator of the land when there is no short-term migrant from the household. However, this gender gap reduces by 4.1 percentage points when there is a short-term migrant from the household. This magnitude is quite substantial as it signifies a reduction by 22 percent over the baseline gender gap.

This analysis takes into account two parameters: δ representing the relative impact of unobservables with respect to the observable covariates on the effect of STM, and R_{max} representing the R-squared in the hypothetical regression that would control for all potential unobservables.[5] Column (3) presents the value of δ that would make the main coefficient in each of these models zero, for a given a plausible value of R_{max} . For the sample of women, the estimated δ to make $\beta = 0$ is -103, indicating that the effect of unobservables would have to be 103 times more than the effect of observables in order to make the coefficient of STM zero. Moreover, the

[5] The method is explained in the appendix.

unobservables must affect STM in the opposite direction as compared to the effect of observables. This finding practically implies that the coefficient of STM in the female sample is extremely stable and unlikely to be driven by any unobservable characteristics. For the male sample, however, the coefficient of STM is not so stable. Unobservable factors that are only 0.312 times as strong as the observable control variables would make the coefficient zero. Finally, in the overall sample where the focus is on the interaction term, the estimated value of δ for $\beta = 0$ is 2.66. Although this value is much lower than the female sample, it still suggests that the effect of unobservable factors would have to be 2.66 times stronger than the joint effect of all control variables to make the coefficient of the interaction term zero. Note that the regression on the overall sample, where the focus is on the interaction term between gender and STM, controls for all observable and unobservable characteristics at the household level as it includes household fixed effects. Therefore, the existence of omitted variables that have 2.66 times higher effect than all the included controls seems to be a very strong requirement for making the coefficient zero. In fact, according to Altonji et al. (2005), it would be reasonable to assume that the observable control variables are at least as important as potential unobservables, hence the possibility that $\delta > 1$ is less likely. This logic is reinforced by Oster (2017) that analyses the sensitivity of treatment effects found in various published studies.

(Table 3)

To summarize, we find that the effect of STM on women’s association with the operational holding is extremely unlikely to be confounded by unobservable factors especially after we have included a comprehensive set of control variables in our model. Further, the finding that women’s likelihood of operating the land increases more than men when the household has a STM is also robust.

c. Instrumental Variable Analysis

The results from the IV regressions corrected for endogeneity of a short-term migrant, are reported in Table 4. Columns (1) and (2) present the coefficients from the models estimated separately for women and men, respectively. We find that even after accounting for endogeneity, women are more likely to be associated with the operational holding in the presence of a short-term migrant. Estimating the model with a pooled sample and including the interaction term does not alter our findings (Table 4, columns 3–4). Besides, the interacted model shows that women are more likely than men to be an operator of the land when there is a STM from the household. This finding is unaltered even in the household fixed effects model.

The size of the main coefficients in the IV model is larger than what we found in the OLS model (Table 3). This is explained by the fact that the OLS model estimates the average effect of STM, while the IV model estimates the local average treatment effect (LATE). The LATE estimate obtained from the IV model measures the effect only for the sub-population of compliers – those who change their STM status due to the effect of the instruments (Angrist and Pischke, 2009). The increase in magnitude of the coefficient in the IV model, as compared to the OLS model, indicates that the effect of STM is heterogenous; especially it is larger for the population where short-term migration is induced by pre-existing migrant networks and availability of manufacturing sector employment in other districts of the state.

(Table 4)

The first stage results of the 2SLS model show that the instruments are highly correlated with the endogenous variables – the F-statistics on the joint significance of the instruments varies from 20 to 35 across the models (Table 5). The instruments also show a plausible relationship with the endogenous variables. As expected, the average marginal effects of both the instruments are significant and positive on STM. Moreover, a significant interaction term shows a positive inter-dependence in the effects: the pull-factor of manufacturing employment in other districts is amplified in the presence of a pre-existing migrant network (Appendix Figures 1 and 2). Several other diagnostic tests support validity of the 2SLS estimation. The Kleibergen-Paap rk Wald F statistics varies from 9 to 14 across the models, indicating that the models do not suffer from the problem of weak identification. The Hansen J statistics for overidentification is statistically not significant, suggesting that the instrumental variables are valid and unlikely to have any direct relationship with the dependent variable in the model.[6]

(Table 5)

d. Additional Robustness and Interpretation

We consider the possibility of sample selection due to the fact that women's involvement in operational holding is observable only for those households who have operated any land for agriculture during the last one year from the date of survey. We estimate a sample selection model while dealing with endogeneity in the 2SLS framework (Wooldridge 2002). Further details are presented in Appendix A2. We find that the results are unchanged even after considering this issue (Appendix Table 4).[7]

[6] To show that the instruments are unlikely to have any direct effect on women's involvement in operational holding, we include an additional control variable measuring the district level share of women engaged in agriculture as a skilled worker. This measure is from 2009-10, thus it is lagged with respect to the dependent variable, but measured after the time period when the instruments are measured. If the instruments directly affect women's involvement in agriculture, then this variable is likely to capture that effect. We find that our results are unperturbed, both in terms of magnitude and significance of the main coefficients, even with this inclusion in the model.

[7] In another sensitivity analysis, we also consider the possibility that our model includes some potentially endogenous variables or "bad controls". Some of our household level variables, e.g. land ownership patterns, are likely to be affected by the presence of a short-term migrant. Therefore, we re-estimate all our regressions excluding these household level covariates. Our main findings remain unchanged.

Since the instrumental variable model considers a binary outcome indicator that does not distinguish main operators from associated operators, a pertinent question is whether there is any effect of STM on women's likelihood of being the main operator. We address this issue by explicitly considering women's involvement as the main operator in models that also correct for endogeneity using the instruments. This is done either in an ordered probit framework or in linear 2SLS where the outcome variable is modified to indicate the main operator. We find that indeed STM has a positive effect on women's involvement as the main operator of the holding (Appendix Table 5).

e. Heterogeneous Effects of STM on Women

We analyse whether the effect of STM on women varies based on individual and household specific characteristics. We estimate an OLS model, which has been found to be robust to concerns of omitted variables in our previous analysis. Figures 1 and 2 present changes in the marginal effect of STM when it is interacted with some key characteristics.[8]

The effect of STM is significant and higher in magnitude for younger women below the age of 40 years. The effect is also significant and larger for a woman who is the spouse of the household head or is an unmarried daughter in the household. Majority of the short-term migrants are either male household heads or their sons. It is plausible that when the household head migrates, his wife becomes the operator of the land. However, it is interesting that an unmarried daughter, but not the

[8] When heterogeneity is analysed with respect to a continuous variable (i.e. age, number of adult male members, and area of land possessed), we also include up to third order polynomial of the continuous variable and interact them with STM as well. This makes the specification flexible to account for nonlinearity in effects.

daughter-in-law, has higher probability of operating the land when there is a STM. Besides, there is no differential effect of STM based on the education level of women (Figure 1).

It is also illustrative that despite the presence of a short-term migrant in the household, the marginal effect of STM on a woman's operational status declines with an increasing proportion of adult men in the household (Figure 2). This suggests that even when women do take on farm managerial responsibilities in migrant households, there is a mediating effect of household structure in terms of the number of adult members present to take over farm responsibilities. We also find that the effect of STM on women is significant only when the area of land possessed is less than 4 hectares. There could be several reasons why women are more likely to retreat from operator roles as household economic status improves. This can be explained by the possibility of an income effect or a status effect (it is considered beneath the household status for women to engage in farm work in any capacity) or simply that women may not be considered capable of managing large farms.

f. Situating our Findings

Our results establish that in households with a short-term migrant, women are more likely to be engaged in decision making on farms. Due to data limitations we are unable to address what decisions are taken by farm operators, the impact of managerial feminization on women's empowerment or the impact of migration on women's labor allocation. However, additional, insights on women's labor allocation are available from NSSO's Situation Assessment Survey of Agricultural Households also conducted in 2013.

In households without a short-term migrant, 41 per cent of women (not including those in education) report domestic duties as their usual primary work status. In contrast, in households with a short-term migrant, this proportion is lower at 36.7 per cent. We also observe a stark difference in the distribution of subsidiary status of women. In households without a short-term migrant, among women reporting domestic duties as their usual principal status, 14 per cent report their subsidiary status as own account worker. In households with a short-term migrant, this number is higher at 19 per cent. The industry of work for over 94 per cent of women with a subsidiary status is agriculture.

These summary statistics provide indirect evidence that during migration, albeit for short periods of time, women assume larger set of roles. This pattern is consistent with results obtained from other countries. Among other studies Mendola and Carletto (2012), who model the employment status (wage employed, paid self-employed or unpaid worker) of individuals staying behind in Albania, find instances where women increase their farm work similar to what we find in India.

6. Discussion

Beginning 2004-05, the spurt in short-term internal migration has been driven by a boom in jobs in the construction industry. Over a decade later, Government of India finally gave a policy impetus to the issue of migration, in particular circular or short-term migration, by devoting an entire chapter of its flagship pre-budget publication Economic Survey 2016-17 to the subject. They recognised that “labour migration in India tends to be circular in nature” (p. 267 Government of India 2017). The report further notes that migration is accelerating and predicts an increase in the migration rate. Since short-term migration is not expected to decline, a question of interest pertains to how it affects women from agricultural households who stay back.

A broad swath of literature on feminization of agriculture has focused almost exclusively on the trends and patterns in the proportion of women working in the agricultural sector as self-employed, unpaid help or wage labor. The skimming over of the difference between ‘feminization of agricultural labor’ and ‘feminization of farm management’ masks crucial differences in women’s roles and responsibilities in the household farm operation. In case of the former, it refers largely to the proportion of work undertaken by women on farm, while in case of latter, it includes women's participation in a range of decisions including input use, cropping decisions, sale of crops etc.

The contribution of this paper is that it provides estimates of 'feminization of farm management' in rural India and highlights how short-term migration from the household affects the probability of a woman being associated with decisions pertaining to the operational holding of the household. Despite the fact that women from the 10 million rural households with a short-term migrant are more likely to be involved in decisions pertaining to operation of land, India's national agricultural policies are still not fully aligned to creating a conducive ecosystem for women engaged in farming (Government of India 2011).

Women also tend to be less educated on average than men and face restrictions on their mobility in certain contexts, both of which could affect their ability to engage in market transactions (Olivier de Schutter 2013; Chandni Singh, Peter Dorward and Henny Osbahr 2016). A limiting factor is that in patriarchal societies, women are often not recognised as owners of land, which impairs their ability to access non-land agricultural inputs (FAO 2010). Mahajan (2019) finds that the gender of the farm manger is of relevance. This study suggests that productivity and profit in farms managed by women, as compared to men, is lower by about 11 per cent, with almost half of this productivity gap being explained by different crop choice and input usage by men and women. The productivity loss due to feminization of management is not good news for a country like India where agricultural productivity on farms is lower than the world average. Similarly, another study in the southern Indian state of Karnataka finds a gender gap in extension services with female heads less likely to benefit from these services than male heads (World Bank and IFPRI 2010). An immediate policy implication is to engender agricultural policies and programs to better identify and respond to the needs of women farmers.

An increase in farm related responsibilities of women has two plausible effects. Evidence suggests that expanding women's employment options is an important contributor to their economic empowerment; indeed it is seen that even working as unpaid family workers on the farm gives women greater bargaining power than being homemakers as it concretises women's contributions to the household. The downside is that additional responsibilities could mean that they have to work extra hours and fulfil domestic duties thereby reducing leisure time and overall wellbeing.

Due to data constraints, our study results do not directly speak about the association between migration and women's agency and their work load. Thus, the analysis also highlights the need for better sex disaggregated data in several domains -- women and men's labor supply as well as time use data to understand and monitor trends in work (paid, unpaid, across sectors). An integration of such data in migration surveys would give us an enhanced understanding of the welfare of women staying behind in the home community of the migrant. We also need disaggregated data on agricultural decision making at least as part of national level surveys administered by India's statistical agency. This data must be collected, and at the individual-level, to be able to truly understand who makes agricultural decisions and how best can policies support them.

References

Agarwal, Bina (2018) Can group farms outperform individual family farms? Empirical insights from India, World Development Volume 108, August 2018, Pages 57-73

Agarwal, Bina and Ankush Agrawal (2017), "Do farmers really like farming? Indian farmers in transition", *Oxford Development Studies*, DOI: 10.1080/13600818.2017.1283010

Agrawal, Tushar and Chandrasekhar, S. (2016), "Labour Market Outcomes of Itinerant Workers in Rural India", *Journal of International Development*, 28: 1251–1271. doi: 10.1002/jid.3126.

Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of Political Economy*, 113(1), 151-184.

Angrist, Joshua D. (2001), "Estimation of limited dependent variable models with dummy endogenous regressors: simple strategies for empirical practice", *Journal of business & economic statistics*, *19*(1), 2-28.

Angrist, Joshua D. and Jörn Steffen Pischke (2009), Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press.

Antman, Francisca M. (2011), "The intergenerational effects of paternal migration on schooling and work: What can we learn from children's time allocations?", *Journal of Development Economics,* 96: 200-208.

Binzel, Christine and Ragui Assaad (2011), "Egyptian men working abroad: Labour supply responses by the women left behind", *Labour Economics,* 18: S98–S114.

Birthal, Pratap S., Devesh Roy & Digvijay S Negi (2015) "Assessing the Impact of Crop Diversification on Farm Poverty in India," *World Development*, Vol. 72(C), pages 70-92.

Birthal, Pratap S., Devesh Roy, Md. Tajuddin Khan and Digvijay Singh Negi (2015), "Farmers' Preference for Farming: Evidence from a Nationally Representative Farm Survey in India", *The Developing Economies*, 53: 122–134. doi:10.1111/deve.12072

Chang, H., MacPhail, F. and Dong, X. 2011. The feminization of labor and gender work-time gap in rural China. *Feminist Economics*, 17(4): 93–124.

de Brauw Alan, Jikun Huang, Linxiu Zhang and Scott Rozelle (2013), "The Feminisation of Agriculture with Chinese Characteristics", *Journal of Development Studies*,49(5): 689-704.

de Brauw, Alan, Qiang Li, Chengfang Liu, Scott Rozelle, and Linxiu Zhang (2008), "Feminisation of Agriculture in China? Myths Surrounding Women's Participation in Farming", *China Quarterly*, 194: 327–348.

de Schutter, Olivier (2013), "The Agrarian Transition and the 'Feminization' of Agriculture", Food Sovereignty: A Critical Dialogue. International conference (Conference Paper #37), Yale University, September 14-15, 2013.

Food and Agriculture Organisation (FAO). (2011), "Gender Differences in Assets", ESA Working Paper No. 11-12. Agricultural Development Economics Division.

Food and Agriculture Organisation (FAO). (2010), The State of Food and Agriculture 2010-11.Women in Agriculture: closing the gender gap for development. Rome: FAO.

Government of India (2019) Agriculture Census 2015-16, All India Report on Number and Area of Operational Holdings. http://agcensus.nic.in/document/agcen1516/T1_ac_2015_16.pdf

Government of India (2017) Economic Survey 2016-17, Ministry of Finance, Department of Economic Affairs, Economic Division.
https://www.indiabudget.gov.in/budget2017-2018/es2016-17/echapter.pdf

Government of India (2014a), Key Indicators of Land and Livestock Holdings in India, Ministry of Statistics and Programme Implementation, Government of India.

Government of India (2014b), Key Indicators of Situation of Agricultural Households in India. Report No NSS KI(70/33), National Sample Survey Organisation, Ministry of Statistics and Programme Implementation, Government of India.

Government of India (2011), XII Five-Year Plan Report of the Working Group on Women's Agency and Empowerment, Ministry of Women and Child Development, New Delhi.

Keller, B.B., P.E. Chola and M.C. Milimo (1990), "Women and agricultural development", In A.P. Wood, S.A. Kean, J.T. Milimo and D.M. Warrren (Eds.). The dynamics of agricultural policy and reform in Zambia. Ames, IA: Iowa State University Press.

Keshri, Kunal and Ram B. Bhagat (2013), "Socioeconomic Determinants of Temporary Labour Migration in India", *Asian Population Studies*, 9(2): 175–195.

Kieran Caitlin, Kathryn Sproule, Cheryl Doss, Agnes Quisumbing, Sung Mi Kim (2015) Examining gender inequalities in land rights indicators in Asia, Agricultural Economics 46 (2015) supplement 119–138

Lokshin, Michael and Elena Glinskaya (2009), "The Effect of Male Migration on Employment Patterns of Women in Nepal", *The World Bank Economic Review*, 23(3): 481 –507.

Mahajan. Kanika (2019) Back to the plough: Women managers and farm productivity in India, World Development. Volume 124, December 2019

Mendola, Mariapia (2012), "Rural out-migration and economic development at origin: A review of the evidence", *Journal of International Development*, *24*(1), 102-122.

Mendola, Mariapia and Calogero Carletto (2012), "Migration and gender differences in the home labour market: Evidence from Albania", *Labour Economics*, 19(6): 870–880.

Mu, Ren and Dominique van de Walle. (2011), "Left Behind to Farm? Women's Labour Re-allocation in Rural China." *Labour Economics,* 18: S83–S97.

Oster, E. (2017). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 1-18.

Paris, Thelma R., Abha Singh, Joyce Luis and Mahabub Hossain (2005), "Labour Outmigration, Livelihood of Rice Farming Households and Women Left Behind: A Case Study in Eastern Uttar Pradesh", *Economic and Political Weekly*, 40(25): 2522–2529.

Schenk-Sandbergen, L. (2018). De-feminisation of Agricultural Wage Labour in Jalpaiguri, West Bengal. *Economic and Political Weekly*, *LIII*(25), 46–53.

Singh, Chandni, Peter Dorward and Henny Osbahr (2016), "Developing a holistic approach to the analysis of farmer decision-making: Implications for adaptation policy and practice in developing countries", *Land Use Policy,* 59 (2016) 329–343.

Slavchevska, Vanya, Susan Kaaria, and Sanna-Liisa Taivalmaa (2016) Feminization of Agriculture in the Context of Rural Transformations: What is the Evidence?, Working Paper, World Bank

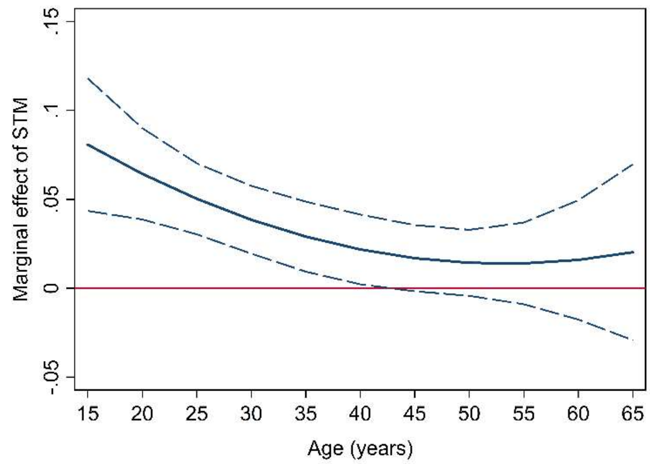
Su, Weiliang, Tor Eriksson, Linxiu Zhang, Yunli Bai (2016) Off-farm employment and time allocation in on-farm work in rural China from gender perspective, China Economic Review 41 (2016) 34–45

Wooldridge, Jeffrey M. (2002), Econometric Analysis of Cross Section and Panel Data, MIT Press, Cambridge, Massachusetts.

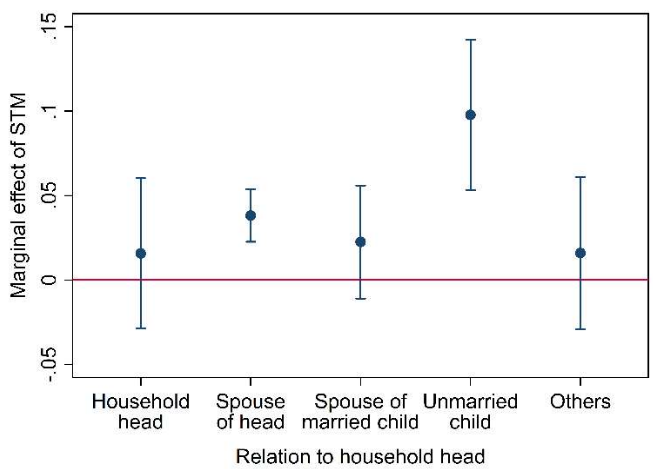
World Bank and IFPRI (International Food Policy Research Institute) (2010), *Gender and governance in rural services: Insights from India, Ghana, and Ethiopia.* Washington, D.C.

Figure 1: Heterogeneity analysis based on individual characteristics

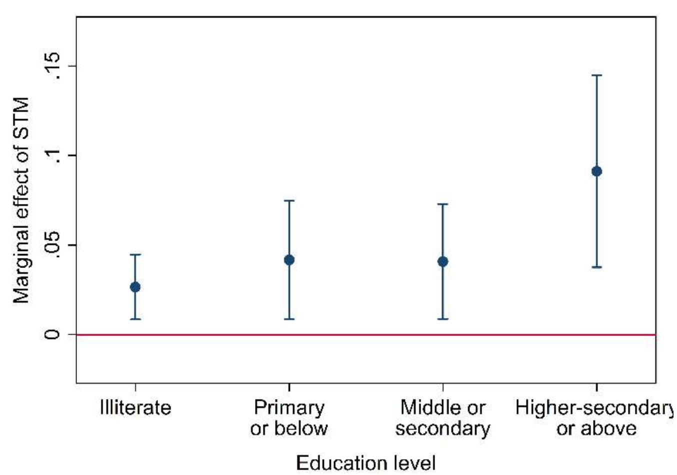
A. Age (years)



B. Relationship to household head

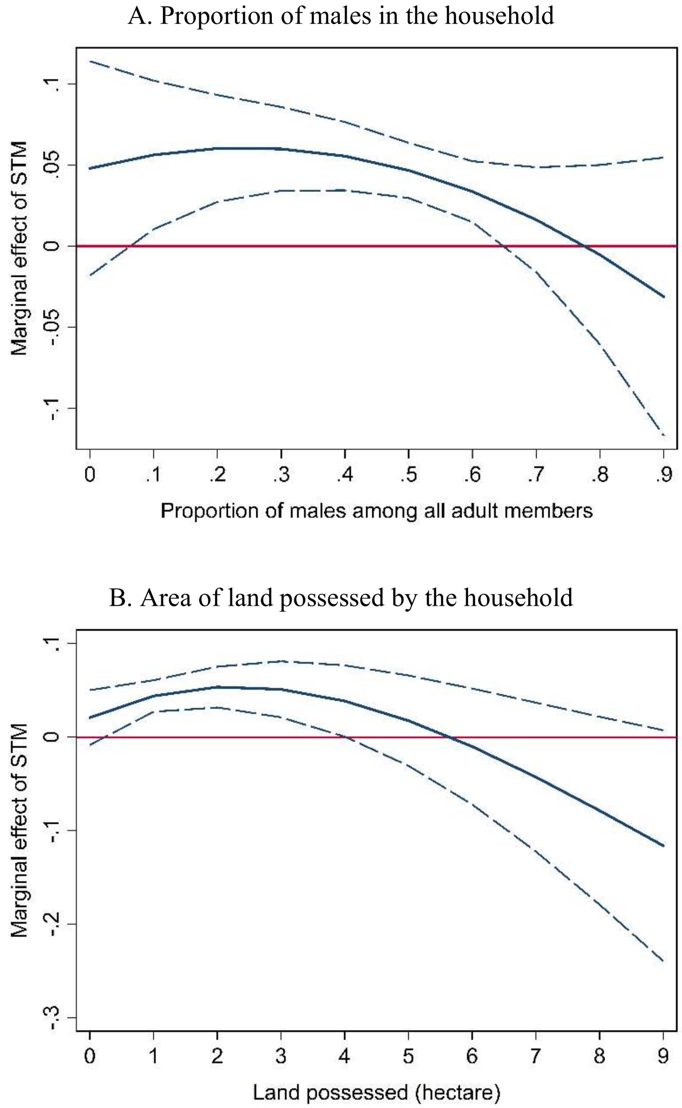


C. Education level



Note: Each graph plots how the marginal effect of STM on the probability that women is an operator of the land varies based on the woman's characteristics. 90 percent conficence intervals are also provided.

Figure 2: Heterogeneity analysis based on household characteristics



Note: Each graph plots how the marginal effect of STM on the probability that women is an operator of the land varies based on household characteristics. 90 percent conficence intervals are also provided.

Table 1: Sample characteristics

Variables	**Obs.**	**Mean**	**SD**
Individual characteristics			
Main operator of the holding	92,376	0.253	0.435
Associated operator of the holding	92,376	0.499	0.500
Female	92,376	0.495	0.500
Age	92,376	35.34	14.44
Square of age	92,376	1,458	1,123
General education: Primary or lower	92,376	0.208	0.406
General education: Middle	92,376	0.192	0.394
General education: Secondary	92,376	0.147	0.354
General education: Higher secondary or above	92,376	0.174	0.379
Spouse of head	92,376	0.228	0.419
Married child	92,376	0.107	0.309
Spouse of married child	92,376	0.113	0.317
Unmarried child	92,376	0.216	0.411
Grandchild	92,376	0.0278	0.165
Father/Mother/Father-in-law/Mother-in-law	92,376	0.0188	0.136
Brother/Sister/Brother-in-law/Sister-in-law/Other relatives	92,376	0.0465	0.211
Servants/Employees/Other non-relatives	92,376	0.00168	0.0409
Household characteristics			
Household has a short-term migrant	92,376	0.0698	0.255
Whether household head is female	92,376	0.0701	0.255
Age of household head	92,376	52.10	12.66
Head's education: Primary or lower	92,376	0.270	0.444
Head's education: Middle	92,376	0.158	0.364
Head's education: Secondary	92,376	0.117	0.321
Head's education: Higher secondary or above	92,376	0.114	0.318
Household size	92,376	6.607	3.634
Dependency ratio: 0-5 children in total household size	92,376	0.0873	0.124
Dependency ratio: 6-14 children in total household size	92,376	0.142	0.165
Proportion of males aged 15 years or above	92,376	0.397	0.165
Average age of the household members	92,376	30.43	9.119
Main income source: non-agriculture	92,376	0.0525	0.223
Main income source: wage/salary	92,376	0.144	0.352
Main income source: other	92,376	0.0172	0.130
Caste: SC	92,376	0.122	0.327
Caste: ST	92,376	0.190	0.392
Caste: OBC	92,376	0.408	0.492
Religion: Muslim	92,376	0.0939	0.292
Religion: Christian	92,376	0.0626	0.242
Religion: Other	92,376	0.0414	0.199
Land possessed [0.4, 1)	92,376	0.175	0.380
Land possessed [1, 2)	92,376	0.336	0.472
Land possessed [2, 4)	92,376	0.233	0.423

Land possessed [4, .)	92,376	0.0789	0.270
Total land leased out	92,376	0.0349	0.342
Share of leased in land (of total possessed)	92,376	0.0815	0.237
Share of area possessed >= 1 season but < 1 year	92,376	0.0236	0.136
Share of area possessed >= 1 year but < 2 years	92,376	0.0205	0.126
Share of area possessed >= 2 years	92,376	0.943	0.210
Share of plot area outside village but within district	92,376	0.106	0.259
Share of plot area outside district but within state	92,376	0.00277	0.0448
Share of plot area outside state	92,376	0.00151	0.0366
Livestock units equivalent[a]	92,376	1.475	1.924
Sub-round dummy	92,376	0.502	0.499
Region characteristics			
Crop diversification index (village level)	92,376	0.22	0.225
Percentage of land unirrigated in district	92,376	44.67	19.61
Percentage of households with KCC	92,376	3.664	3.986
Growth rate of night lights	92,376	0.069	0.074
Rural unemployment rate	92,376	0.107	0.055
Rural share of agricultural workers among marginal workers	92,376	0.754	0.130
Average rainfall deviation	92,376	-21.41	41.12
Past manufacturing employment in rest of the state (log)	91,207	14.10	1.737
Rate of short term migration	92,376	0.0579	0.166
Proportion of households with land in the district	92,376	0.461	0.175

Source: National Sample Survey 2013 data on Land and Livestock Holding for all variables except: percentage of un-irrigated land in the district, percentage of households having *Kisan* credit card, and proportion of households with land in the district (Socio-Economic and Caste Census 2011), rainfall deviation (Indian Meteorological Department), past manufacturing employment in rest of the state (National Sample Survey 2004-05) and rate of short term migration (National Sample Survey 2007-08).

[a] The animal unit equivalent is constructed following the method given in the Manual on Cost of Cultivation Surveys by Central Statistical Organization of India: http://mospi.nic.in/Mospi_New/upload/manual_cost_cultivation_surveys_23july08.pdf.

Table 2: Marginal effects from ordered probit model for different types of association with the operational holding – all individuals (15-65 years)

	(1) Not associated	(2) Associated with other members	(3) Main operator
	Female		
STM	-0.040***	0.031***	0.009***
Observations	45,729	45,729	45,729
	Male		
STM	-0.005	0.0004	0.004
Observations	46,647	46,647	46,647
	All		
STM	0.005	-0.0005	-0.005
	(0.005)	(0.0007)	(0.005)
Female	0.158***	-0.056***	-0.102***
	(0.004)	(0.001)	(0.003)
Female*STM	-0.046***	0.020***	0.026***
	(0.009)	(0.003)	(0.006)
Observations	92,376	92,376	92,376
Other control variables	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes

Standard errors (clustered at the household level) in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Assessing the effect of unobservables on the main coefficient of interest

	(1)	(2)	(3)
	No-control	All-control	δ for β=0 given R_{max}
	Female (R_{max} = 0.77)		
STM	0.038***	0.039***	-103.034
	(0.009)	(0.010)	
Observations	45,729	45,729	
R-squared	0.0004	0.245	
	Male (R_{max} = 0.64)		
STM	-0.016**	0.014*	-0.312
	(0.007)	(0.008)	
Observations	46,647	46,647	
R-squared	0.0001	0.260	
	All (R_{max} = 0.73)		
Female	-0.165***	-0.186***	
	(0.003)	(0.005)	
STM * Female	0.038***	0.041***	2.66
	(0.011)	(0.012)	
Observations	92,376	92,376	
R-squared	0.036	0.531	

Note: Standard errors (clustered at the household level) in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Column (1) presents regression without any control variable. Column (2) presents regression with all control variables, including state fixed effects for male and female sample, and household fixed effects in the sample of all individuals. For male and female samples, R_{max} is calculated by estimating a regression that includes household fixed effects. For the overall sample, R_{max} is calculated from a regression that controls for household-by-gender fixed effects.

Table 4: 2SLS estimates of the effect of STM on individual's association with operational holding (binary dependent variable of whether associated or main operator of the operational holding)

	(1)	(2)	(3)	(4)
Variables	Female	Male	All	All (with household fixed effects)
STM	0.932***	0.489***	0.447***	
	(0.232)	(0.162)	(0.115)	
STM * Female			0.502***	0.717***
			(0.181)	(0.191)
Female			-0.119***	-0.128***
			(0.022)	(0.020)
Observations	45,141	46,066	91,207	90,075
Other control variables	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes
Household fixed effects	No	No	No	Yes
Weak identification test				
Kleibergen-Paap rk Wald F	12.19	8.809	14.23	12.49
Overidentification test Hansen J	2.933	0.812	6.312	2.715
Overidentification test p value	0.231	0.666	0.177	0.257

Robust standard errors (clustered at the household level) in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5: First stage of 2SLS estimates

	(1)	(2)	(3)	(4)	(5)
	Female	Male	All		All (with household fixed effects)
Variables	STM	STM	STM	STM * Female	STM * Female
Past manufacturing employment in rest of the state (log)	0.022**	0.016*	0.019**	-0.000	
	(0.010)	(0.010)	(0.009)	(0.002)	
Past rate of short-term migration in district	-0.254***	-0.191***	-0.208***	0.008	
	(0.063)	(0.058)	(0.058)	(0.005)	
Past manufacturing employment in rest of the state (log) * Past rate of short-term migration in district	0.026***	0.020***	0.021***	-0.001**	
	(0.005)	(0.004)	(0.004)	(0.000)	
Female * Past manufacturing employment in rest of the state (log)			-0.000	0.022**	0.023***
			(0.013)	(0.010)	(0.009)
Female * Past rate of short-term migration in district			-0.031	-0.277***	-0.256***
			(0.085)	(0.063)	(0.058)
Female * Past manufacturing employment in rest of the state (log) * Past rate of short-term migration in district			0.004	0.028***	0.026***
			(0.007)	(0.005)	(0.004)
Observations	45,141	46,066	91,207	91,207	90,075
R-squared	0.045	0.044	0.044	0.068	0.547
Other control variables	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes
Household fixed effects	No	No	No	No	Yes
F stat for joint significance of instruments	32.87	23.52	28.26	20.01	34.86
Average marginal effects					
Past manufacturing employment in rest of the state (log)	0.023**	0.017*	0.02***		
	(0.01)	(0.009)	(0.007)		
Past rate of short-term migration in district	0.117***	0.095***	0.106***		
	(0.02)	(0.012)	(0.008)		

Robust standard errors (clustered at the household level) in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix A

Appendix A1. Assessment of Potential Bias due to Unobservables

Since the decision of short-term migration is likely to be endogenous, our regression includes a comprehensive set of control variables to account for the factors driving this migration decision. Yet, the possibility of omitted (unobserved) variables cannot be ruled out. If an omitted variable is correlated with both STM and the dependent variable, then the estimated coefficient of STM can be biased. We follow a strategy developed by Altonji et al. (2005) and later extended by Oster (2017) to assess the extent of such omitted variable bias in our regression. The method seeks to infer the potential effect of unobservables by examining the effect of observable control variables on the coefficient of interest. To explain further, let us consider the following simplified regression framework:

$$Y = \beta X + \gamma Z + W \tag{2}$$

Suppose that X is the main variable and β is the main coefficient of interest, Z represents all observed control variables and W includes all the unobserved factors. The method proposed by Altonji et al. (2005) relies on the "proportional selection assumption" which posits that the relationship between X and unobservables is proportional to the relationship between X and observables. The degree of this proportionality is given by δ. Formally, this assumption is reflected in the following relationship:

$$\frac{Cov(X,W)}{Var(W)} = \delta \frac{Cov(X,\gamma Z)}{Var(\gamma Z)} \tag{3}$$

Under this assumption, it is possible to estimate the omitted variable bias by looking at how the main coefficient (β) changes when control variables are added in the regression. Oster (2017) extends this method to consider the fact that not all control variables have equal relevance. Hence, her analysis incorporates the movement in the R-squared value of the regression, along with the movement in β, to deduce the bias due to unobservables. Following this method, the estimated omitted variable bias is a function of two parameters: δ and R_{max} where R_{max} is defined as the R-squared from the hypothetical regression that controls for all factors – both observables and unobservables.

In the context of our analysis, we seek to understand how large the influence of unobservables would have to be to completely wash away the effect of STM on women's association with the operation holding. Therefore, we estimate the value of δ that would make $\beta = 0$, given a reasonable value of R_{max}.

Since this method is applicable only in linear regression, therefore we estimate a linear probability model where the outcome variable is a binary indicator that takes the value 1 if an individual operates the land as either a main operator or an associated operator, and it takes the value 0 if the individual is not associated with the operational holding. As in the previous model, we estimate the regression separately for samples of females, males, and all individuals.

When the regression is estimated on male and female sample separately, the main variable of interest is STM which indicates whether the household has a short-term migrant. Since this is a household level variable, therefore any possible endogeneity in this variable is likely due to unobserved factor that also varies at the household level. Inclusion of household fixed effects would completely subsume the effect of any such unobservable variable. Hence, following Oster's (2017) suggestion, we argue that a plausible value of R_{max} in this case is given by the R-squared of the regression that includes household fixed effects.[i] This value is 0.77 and 0.64 respectively, for the female and male samples. On the other hand, the main variable of interest in the regression on all individuals is the interaction term between STM and gender because it shows whether women are more likely than men to be operators when there is a STM in the household. This interaction term varies between male and female and across households. Therefore, for this model a plausible value of R_{max} is obtained by estimating the regression including household-by-gender fixed effects. R-squared from this augmented model comes out to be 0.73. We use these values of R_{max} and estimate how large δ would have to be to completely nullify our main findings.

Appendix A2. Robustness Check Considering Sample Selection

In our sample, nearly 24 per cent of households did not operate any land for agriculture during the last 365 days from the date of survey. For these households, the dependent variable indicating who operates the agricultural land is not relevant, and hence, not defined. Given the systematic differences observed, we have a non-random sample of households for whom an individual member's association with the operational holding is reported. If there are unobservable characteristics, e.g. household's affluence, preferences etc., that determine the probability of operating any land and also which member of the household is associated with the operational holding, then we have a sample selection problem; not accounting for selectivity will result in biased and inconsistent estimates.

We follow Wooldridge (2002) to deal with the problem of endogeneity and sample selection in our empirical model. Both of these problems are considered in a single framework. We first estimate a selection equation using a probit model:

$$\begin{aligned} Pr(Opland_{ihds} = 1) & \\ = \Phi(&\gamma_0 + \gamma_2 Female_{ihds} + \gamma_2 X_{ihds} + \gamma_3 Z_{hds} + \gamma_4 V_{ds} \\ &+ \gamma_5 Manufacturing_{-ds} + \gamma_6 Migrantrate_{ds} \\ &+ \gamma_7 Manufacturing_{-ds} \times Migrantrate_{ds} \\ &+ \gamma_8 Manufacturing_{-ds} * Female_{ihds} \\ &+ \gamma_9 Migrantrate_{ds} * Female_{ihds} \\ &+ \gamma_{10} Manufacturing_{-ds} \times Migrantrate_{ds} \times Female_{ihds} \\ &+ \gamma_{11} Land_{ds} + \delta_s) \end{aligned} \quad (2)$$

The dependent variable is a binary indicator of whether the individual belongs to a household having any operational land. $\Phi(.)$ is the cumulative distribution function of the standard normal distribution. All explanatory variables from the main model (Equation 1) are included except the two endogenous variables, namely short-term migration status and its interaction with the female dummy. The selection equation should also contain another identifying variable which affects the probability of operating agricultural land, but does not affect individual's association with operational holding. We use district level percentage of households with land ($Land_{ds}$) as the identifying variable for selection. Similar to Heckman (1979), we estimate the Inverse Mills Ratio or IMR (ratio of the estimated standard normal density and cumulative distribution function) from this equation. In the next step, the main model is estimated using 2SLS method, with the full set of instruments and IMR as an additional control variable.[ii]

The results from the 2SLS model, corrected for selectivity and for endogeneity are presented in Appendix Table 4. The findings corroborate those obtained in our earlier analyses. The effect of STM is positive for both men and women, but it is significantly higher for women, thereby reducing the gender gap. These findings support our hypothesis that women get more involved in agricultural decision making when there is a short-term migrant in the household. The first stage of the 2SLS model correcting for selectivity is presented is Appendix Table 6, the selection equation is presented in Appendix Table 7. The identifying variable, *i.e.,* proportion of landed households in the district, is a significant determinant of whether household operates land or not.

Additional References

Lenze, J., & Klasen, S. (2016). Does women's labor force participation reduce domestic violence? Evidence from Jordan. *Feminist Economics*.

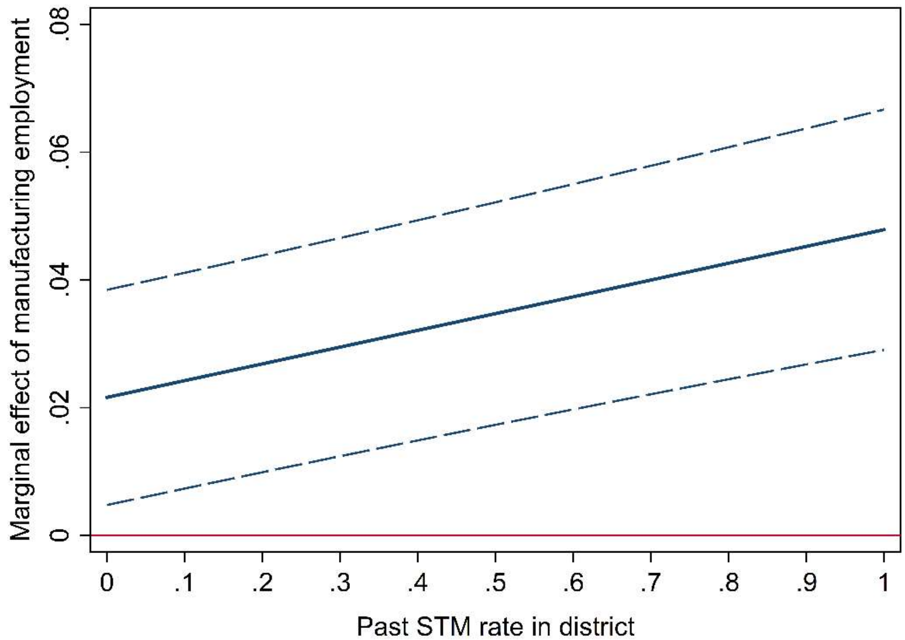
Roodman, D. (2011). Fitting fully observed recursive mixed-process models with cmp. *The Stata Journal*, *11*(2), 159-206.

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

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

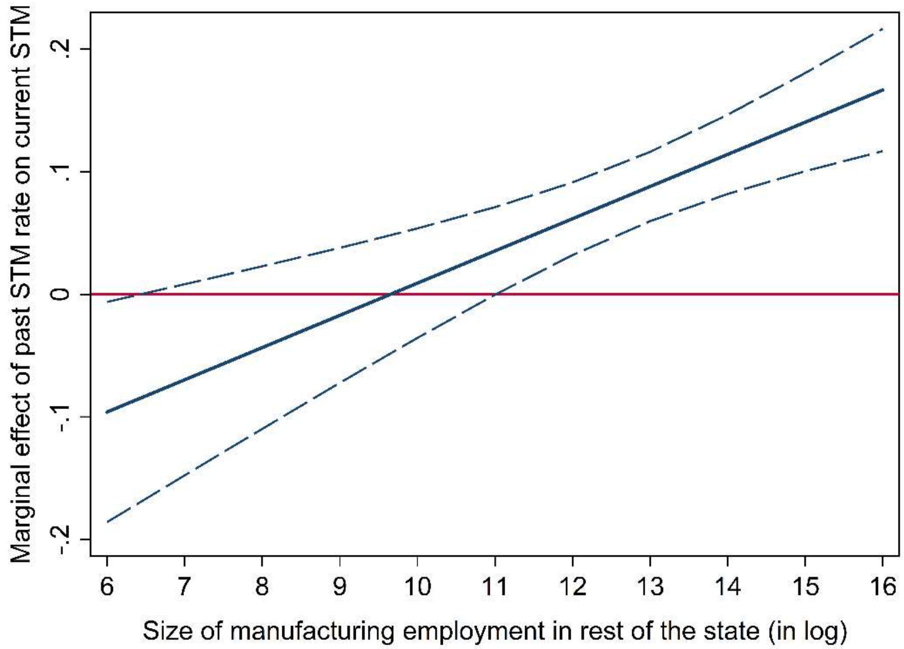
Appendix A

Appendix Figure 1: First stage illustration – the effect of past manufacturing employment (in logarithm) in the rest of the state on current STM varies with past STM rate in district



Note: The figure is obtained from the first stage regression of the 2SLS model. The dashed line denotes 90 percent confidence interval.

Appendix Figure 2: First stage illustration – the effect of past STM rate in district on current STM varies with the size of past manufacturing employment (in logarithm) in the remaining districts of the state



Note: The figure is obtained from the first stage regression of the 2SLS model. The dashed line denotes 90 percent confidence interval.

Appendix Table 1: Comparison of characteristics depending on whether household has a short-term migrant

Variables	**Household with STM**	**Household without STM**	**Difference**	**p-value (two tail)**
Proportion of male members in the household	0.368	0.373	-0.006	0.199
Average age of the household members	27.287	31.04	-3.753	0.000
Whether household head is female	0.113	0.114	-0.001	0.891
Age of household head	46.904	48.569	-1.664	0.000
Head's education: Illiterate	0.453	0.359	0.094	0.000
Head's education: Below primary	0.141	0.141	0	0.977
Head's education: Primary	0.112	0.127	-0.015	0.041
Head's education: Middle	0.14	0.155	-0.015	0.065
Head's education: Secondary	0.084	0.108	-0.025	0.000
Head's education: Higher secondary	0.034	0.055	-0.021	0.000
Head's education: Graduate or above	0.037	0.056	-0.019	0.000
Household size	5.767	5.03	0.737	0.000
Dependency ratio: 0-5 children in total household size	0.291	0.252	0.039	0.000
Dependency ratio: 6-14 children in total household size	0.045	0.067	-0.023	0.000
Income source: agriculture	0.519	0.609	-0.09	0.000
Income source: non-agriculture	0.093	0.091	0.002	0.727
Income source: wage/salary	0.332	0.245	0.087	0.000
Income source: other	0.055	0.055	0.001	0.893
Caste: Other	0.204	0.264	-0.06	0.000
Caste: SC	0.178	0.16	0.018	0.027
Caste: ST	0.234	0.178	0.056	0.000
Caste: OBC	0.384	0.398	-0.014	0.210
Religion: Hindu	0.784	0.807	-0.023	0.010
Religion: Muslim	0.136	0.095	0.041	0.000
Religion: Christian	0.047	0.057	-0.011	0.041
Religion: Other	0.033	0.041	-0.008	0.088
Land possessed	0.991	1.213	-0.222	0.000
Total land leased out	0.047	0.036	0.01	0.174
Share of owned land out of total possessed	0.894	0.886	0.007	0.269
Share of leased in land out of total possessed	0.082	0.092	-0.01	0.091
Share of homestead land	0.303	0.277	0.026	0.005
Share of area possessed <1 season	0.011	0.016	-0.005	0.076
Share of area possessed >= 1 season but < 1 year	0.022	0.023	-0.001	0.834
Share of area possessed >= 1 year but < 2 years	0.019	0.021	-0.002	0.589
Share of area possessed >= 2 years	0.947	0.94	0.007	0.162
Share of plot area within village	0.912	0.918	-0.006	0.233
Share of plot area outside village but within district	0.083	0.078	0.005	0.333
Share of plot area outside district but within state	0.003	0.003	0.001	0.534
Share of plot area outside state	0.002	0.002	0.001	0.436
Number of observations	2,115		31,357	

Note: The analysis is based on 33,472 households which have at least one member in the working age of 15-65 years and hence are included in our analytical sample.

Appendix Table 2: Marginal effects from ordered probit model for different types of association with the operational holding – full model estimates

Variables	Female			Male		
	Not associated	Associate operator	Main operator	Not associated	Associate operator	Main operator
	(1)	(2)	(3)	(4)	(5)	(6)
STM	-0.040***	0.031***	0.009***	-0.005	0.0004	0.004
	(0.009)	(0.007)	(0.002)	(0.004)	(0.0004)	(0.004)
Individual characteristics						
Age	-0.030***	0.024***	0.007***	-0.024***	0.002***	0.022***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
Square of age	0.0004***	-0.0003***	-0.0001***	0.0003***	-0.00003***	-0.0003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
General education: Primary or lower	0.017***	-0.013***	-0.004***	-0.015***	0.001***	0.014***
	(0.006)	(0.004)	(0.001)	(0.005)	(0.000)	(0.004)
General education: Middle	0.036***	-0.028***	-0.008***	-0.001	0.0001	0.001
	(0.007)	(0.005)	(0.001)	(0.005)	(0.0004)	(0.004)
General education: Secondary	0.074***	-0.058***	-0.016***	0.014***	-0.001***	-0.013***
	(0.007)	(0.006)	(0.002)	(0.005)	(0.000)	(0.004)
General education: Higher secondary or above	0.128***	-0.100***	-0.028***	0.039***	-0.004***	-0.036***
	(0.008)	(0.006)	(0.002)	(0.005)	(0.001)	(0.004)
Spouse of head	0.563***	-0.440***	-0.124***	0.332***	-0.030***	-0.302***
	(0.017)	(0.015)	(0.003)	(0.056)	(0.007)	(0.051)
Married child	0.666***	-0.520***	-0.146***	0.399***	-0.036***	-0.363***
	(0.026)	(0.022)	(0.005)	(0.007)	(0.005)	(0.005)
Spouse of married child	0.707***	-0.552***	-0.155***	0.440***	-0.039***	-0.401***
	(0.020)	(0.017)	(0.004)	(0.014)	(0.006)	(0.013)
Unmarried child	0.748***	-0.584***	-0.164***	0.438***	-0.039***	-0.399***
	(0.020)	(0.017)	(0.004)	(0.007)	(0.006)	(0.006)
Grandchild	0.826***	-0.645***	-0.181***	0.512***	-0.046***	-0.466***
	(0.026)	(0.022)	(0.006)	(0.011)	(0.007)	(0.010)
Father/Mother/Father-in-law/Mother-in-law	0.723***	-0.564***	-0.159***	0.442***	-0.039***	-0.402***
	(0.022)	(0.019)	(0.004)	(0.013)	(0.006)	(0.012)
Brother/Sister/Brother-in-law/Sister-in-law/Other relatives	0.695***	-0.542***	-0.153***	0.422***	-0.038***	-0.385***
	(0.020)	(0.017)	(0.004)	(0.007)	(0.006)	(0.006)
Servants/Employees/Other non-relatives	0.760***	-0.593***	-0.167***	0.441***	-0.039***	-0.402***
	(0.071)	(0.056)	(0.015)	(0.024)	(0.006)	(0.021)
Household characteristics						
Whether household head is female	-0.045***	0.035***	0.010***	-0.079***	0.007***	0.072***
	(0.011)	(0.009)	(0.002)	(0.005)	(0.001)	(0.005)
Age of household head	0.00009	-0.00007	-0.00002	-0.001***	0.0001***	0.001***
	(0.0003)	(0.0003)	(0.000)	(0.000)	(0.000)	(0.000)
Head's education: Primary or lower	0.013**	-0.010**	-0.003**	0.009***	-0.001***	-0.008***

	(0.006)	(0.005)	(0.001)	(0.003)	(0.000)	(0.003)
Head's education: Middle	0.031***	-0.024***	-0.007***	0.015***	-0.001***	-0.014***
	(0.007)	(0.006)	(0.002)	(0.004)	(0.000)	(0.003)
Head's education: Secondary	0.032***	-0.025***	-0.007***	0.015***	-0.001***	-0.013***
	(0.008)	(0.006)	(0.002)	(0.004)	(0.000)	(0.004)
Head's education: Higher secondary or above	0.046***	-0.036***	-0.010***	0.017***	-0.002***	-0.016***
	(0.009)	(0.007)	(0.002)	(0.005)	(0.000)	(0.004)
Household size	0.007***	-0.006***	-0.002***	0.008***	-0.001***	-0.007***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
Dependency ratio: 0-5 children in total household size	-0.109***	0.085***	0.024***	-0.057***	0.005***	0.052***
	(0.026)	(0.021)	(0.006)	(0.016)	(0.002)	(0.014)
Dependency ratio: 6-14 children in total household size	-0.071***	0.056***	0.016***	-0.034***	0.003***	0.031***
	(0.021)	(0.016)	(0.005)	(0.012)	(0.001)	(0.011)
Proportion of males aged 15 years or above	0.136***	-0.107***	-0.030***	0.052***	-0.005***	-0.047***
	(0.021)	(0.017)	(0.005)	(0.012)	(0.001)	(0.011)
Average age of the household members	-0.002***	0.002***	0.0005***	-0.001**	0.0001**	0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Main income source: non-agriculture	0.014	-0.011	-0.003	0.071***	-0.006***	-0.064***
	(0.011)	(0.009)	(0.002)	(0.006)	(0.001)	(0.006)
Main income source: wage/salary	-0.049***	0.038***	0.011***	0.073***	-0.006***	-0.066***
	(0.007)	(0.005)	(0.002)	(0.004)	(0.001)	(0.004)
Main income source: other	-0.026	0.021	0.006	0.034***	-0.003***	-0.031***
	(0.017)	(0.013)	(0.004)	(0.011)	(0.001)	(0.010)
Caste: SC	-0.053***	0.041***	0.012***	0.001	-0.0001	-0.001
	(0.008)	(0.006)	(0.002)	(0.004)	(0.0003)	(0.004)
Caste: ST	-0.090***	0.070***	0.020***	-0.017***	0.001***	0.015***
	(0.008)	(0.006)	(0.002)	(0.004)	(0.000)	(0.004)
Caste: OBC	-0.039***	0.030***	0.008***	-0.0001	0.00001	0.0001
	(0.006)	(0.005)	(0.001)	(0.003)	(0.0002)	(0.003)
Religion: Muslim	0.035***	-0.027***	-0.008***	0.005	-0.0004	-0.004
	(0.009)	(0.007)	(0.002)	(0.004)	(0.0002)	(0.004)
Religion: Christian	0.006	-0.005	-0.001	-0.005	0.0004	0.005
	(0.014)	(0.011)	(0.003)	(0.007)	(0.001)	(0.006)
Religion: Other	0.045***	-0.035***	-0.010***	0.017**	-0.002**	-0.015**
	(0.015)	(0.012)	(0.003)	(0.008)	(0.001)	(0.007)
Land possessed [0.4, 1)	0.029***	-0.023***	-0.006***	-0.036***	0.003***	0.033***
	(0.007)	(0.006)	(0.002)	(0.004)	(0.001)	(0.004)
Land possessed [1, 2)	0.031***	-0.024***	-0.007***	-0.041***	0.004***	0.037***
	(0.007)	(0.006)	(0.002)	(0.004)	(0.001)	(0.004)
Land possessed [2, 4)	0.034***	-0.027***	-0.008***	-0.042***	0.004***	0.038***
	(0.008)	(0.006)	(0.002)	(0.005)	(0.001)	(0.004)
Land possessed [4, .)	0.051***	-0.040***	-0.011***	-0.043***	0.004***	0.040***
	(0.011)	(0.009)	(0.002)	(0.006)	(0.001)	(0.005)

Total land leased out	-0.001	0.001	0.0003	0.005	-0.0004	-0.004
	(0.007)	(0.005)	(0.002)	(0.003)	(0.0003)	(0.003)
Share of leased in land out of total possessed	0.002	-0.001	-0.0003	0.003	-0.0002	-0.002
	(0.010)	(0.008)	(0.002)	(0.006)	(0.001)	(0.005)
Share of area possessed >= 1 season but < 1 year	-0.055**	0.043**	0.012**	-0.009	0.001	0.008
	(0.027)	(0.021)	(0.006)	(0.013)	(0.001)	(0.012)
Share of area possessed >= 1 year but < 2 years	-0.055**	0.043**	0.012**	-0.011	0.001	0.010
	(0.027)	(0.021)	(0.006)	(0.014)	(0.001)	(0.013)
Share of area possessed >= 2 years	-0.035	0.027	0.008	0.011	-0.001	-0.010
	(0.021)	(0.017)	(0.005)	(0.010)	(0.001)	(0.009)
Share of plot area outside village but within district	0.018**	-0.014**	-0.004**	-0.014***	0.001***	0.013***
	(0.009)	(0.007)	(0.002)	(0.004)	(0.000)	(0.004)
Share of plot area outside district but within state	0.026	-0.020	-0.006	-0.029	0.003	0.026
	(0.050)	(0.039)	(0.011)	(0.024)	(0.002)	(0.022)
Share of plot area outside state	0.106*	-0.082*	-0.023*	-0.024	0.002	0.022
	(0.064)	(0.050)	(0.014)	(0.032)	(0.003)	(0.029)
Livestock units equivalent	-0.009***	0.007***	0.002***	-0.004***	0.0003***	0.003***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
Sub-round dummy	-0.008**	0.007**	0.002**	-0.006***	0.0005**	0.006***
	(0.004)	(0.003)	(0.001)	(0.002)	(0.0002)	(0.002)
Region characteristics						
Crop diversification index (village)	-0.069***	0.054***	0.015***	-0.001	0.00006	0.001
	(0.010)	(0.008)	(0.002)	(0.005)	(0.0004)	(0.005)
Percentage of land unirrigated in district	-0.001***	0.001***	0.0002***	-0.0001	0.00001	0.0001
	(0.000)	(0.000)	(0.000)	(0.0001)	(0.000007)	(0.0001)
Percentage of households with KCC	-0.001	0.001	0.0002	0.0004	-0.00004	-0.0004
	(0.001)	(0.001)	(0.0002)	(0.0004)	(0.00004)	(0.0004)
Growth rate of night lights	-0.139***	0.108***	0.030***	-0.030*	0.003	0.027*
	(0.032)	(0.025)	(0.007)	(0.018)	(0.002)	(0.016)
Rural unemployment rate	-0.294***	0.230***	0.065***	-0.060	0.005	0.055
	(0.089)	(0.069)	(0.020)	(0.046)	(0.004)	(0.042)
Rural share of agricultural workers among marginal workers	-0.114***	0.089***	0.025***	-0.081***	0.007***	0.074***
	(0.027)	(0.021)	(0.006)	(0.014)	(0.002)	(0.013)
Average rainfall deviation	0.0001	-0.00009	-0.00002	-0.0001***	0.000008**	0.0001***
	(0.0001)	(0.00006)	(0.00002)	(0.00004)	(0.000003)	(0.00003)
Observations	45,729	45,729	45,729	46,647	46,647	46,647
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors (clustered at the household level) in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table 3: Full model estimates of OLS regression of whether an individual is an operator (main/associated) or not

Variables	(1) Female	(2) Male	(3) All	(4) All
STM	0.039***	0.014*	0.006	
	(0.010)	(0.008)	(0.008)	
Female			-0.188***	-0.186***
			(0.005)	(0.006)
STM * Female			0.042***	0.041***
			(0.012)	(0.014)
Age	0.033***	0.035***	0.033***	0.041***
	(0.001)	(0.001)	(0.001)	(0.001)
Square of age	-0.0004***	-0.0004***	-0.0004***	-0.0005***
	(0.000)	(0.000)	(0.000)	(0.000)
General education: Primary or lower	-0.025***	0.038***	-0.006	0.009
	(0.006)	(0.007)	(0.005)	(0.006)
General education: Middle	-0.054***	0.019**	-0.023***	0.007
	(0.008)	(0.008)	(0.005)	(0.007)
General education: Secondary	-0.096***	-0.013	-0.057***	-0.016**
	(0.009)	(0.008)	(0.006)	(0.007)
General education: Higher secondary or above	-0.159***	-0.046***	-0.091***	-0.040***
	(0.010)	(0.009)	(0.007)	(0.008)
Spouse of head	-0.014	-0.064	0.011**	0.029***
	(0.014)	(0.075)	(0.005)	(0.006)
Married child	-0.142***	-0.047***	-0.049***	-0.000
	(0.024)	(0.012)	(0.009)	(0.011)
Spouse of married child	-0.159***	-0.114***	-0.132***	-0.070***
	(0.017)	(0.027)	(0.011)	(0.013)
Unmarried child	-0.206***	-0.130***	-0.157***	-0.094***
	(0.018)	(0.014)	(0.010)	(0.013)
Grandchild	-0.306***	-0.302***	-0.297***	-0.192***
	(0.027)	(0.024)	(0.019)	(0.022)
Father/Mother/Father-in-law/Mother-in-law	-0.239***	-0.241***	-0.247***	-0.270***
	(0.021)	(0.029)	(0.014)	(0.017)
Brother/Sister/Brother-in-law/Sister-in-law/Other relatives	-0.150***	-0.085***	-0.107***	-0.079***
	(0.018)	(0.011)	(0.009)	(0.011)
Servants/Employees/Other non-relatives	-0.217**	-0.094	-0.116**	0.043
	(0.087)	(0.058)	(0.052)	(0.055)
Whether household head is female	0.054***	0.042***	0.059***	
	(0.013)	(0.010)	(0.007)	
Age of household head	-0.0004	0.0004	-0.0002	
	(0.0004)	(0.0004)	(0.0003)	
Head's education: Primary or lower	-0.016**	-0.029***	-0.014***	
	(0.007)	(0.007)	(0.005)	
Head's education: Middle	-0.043***	-0.039***	-0.035***	
	(0.008)	(0.008)	(0.006)	
Head's education: Secondary	-0.046***	-0.028***	-0.033***	
	(0.009)	(0.009)	(0.007)	
Head's education: Higher secondary or above	-0.063***	-0.029***	-0.046***	
	(0.010)	(0.010)	(0.008)	
Household size	-0.008***	-0.009***	-0.008***	
	(0.001)	(0.001)	(0.001)	
Dependency ratio: 0-5 children in total household size	0.175***	0.163***	0.166***	

	(0.029)	(0.020)	(0.019)
Dependency ratio: 6-14 children in total household size	0.102***	0.007	0.057***
	(0.023)	(0.016)	(0.015)
Proportion of males aged 15 years or above	-0.077***	-0.042**	-0.066***
	(0.023)	(0.017)	(0.015)
Average age of the household members	0.002***	0.001*	0.001***
	(0.000)	(0.000)	(0.000)
Main income source: non-agriculture	-0.053***	-0.096***	-0.077***
	(0.012)	(0.010)	(0.009)
Main income source: wage/salary	0.001	-0.075***	-0.040***
	(0.007)	(0.006)	(0.005)
Main income source: other	-0.019	-0.039**	-0.029**
	(0.018)	(0.019)	(0.014)
Caste: SC	0.058***	0.006	0.035***
	(0.009)	(0.007)	(0.006)
Caste: ST	0.113***	0.030***	0.072***
	(0.009)	(0.006)	(0.006)
Caste: OBC	0.047***	0.004	0.027***
	(0.007)	(0.005)	(0.005)
Religion: Muslim	-0.050***	-0.004	-0.025***
	(0.011)	(0.008)	(0.007)
Religion: Christian	-0.011	0.010	-0.003
	(0.016)	(0.011)	(0.010)
Religion: Other	-0.060***	-0.027**	-0.046***
	(0.017)	(0.013)	(0.011)
Land possessed [0.4, 1)	0.001	0.035***	0.017***
	(0.008)	(0.007)	(0.006)
Land possessed [1, 2)	0.001	0.049***	0.024***
	(0.008)	(0.006)	(0.006)
Land possessed [2, 4)	-0.006	0.054***	0.022***
	(0.009)	(0.007)	(0.006)
Land possessed [4, .)	-0.026**	0.057***	0.013
	(0.013)	(0.009)	(0.009)
Total land leased out	-0.002	-0.007	-0.005
	(0.008)	(0.006)	(0.006)
Share of leased in land out of total possessed	-0.004	-0.004	-0.004
	(0.011)	(0.008)	(0.008)
Share of area possessed >= 1 season but < 1 year	0.068**	0.028	0.048**
	(0.031)	(0.020)	(0.021)
Share of area possessed >= 1 year but < 2 years	0.068**	0.034	0.051**
	(0.031)	(0.021)	(0.020)
Share of area possessed >= 2 years	0.033	0.002	0.016
	(0.025)	(0.016)	(0.016)
Share of plot area outside village but within district	-0.017	0.021***	0.002
	(0.010)	(0.007)	(0.007)
Share of plot area outside district but within state	-0.030	0.078**	0.024
	(0.058)	(0.032)	(0.035)
Share of plot area outside state	-0.132*	0.070	-0.030
	(0.078)	(0.047)	(0.052)
Livestock units equivalent	0.012***	0.007***	0.010***
	(0.002)	(0.001)	(0.001)
Sub-round dummy	0.014***	0.008**	0.011***
	(0.005)	(0.004)	(0.003)

Crop diversification index (village)	0.076***	0.010	0.044***	
	(0.012)	(0.008)	(0.008)	
Percentage of land unirrigated in district	0.001***	0.0003**	0.001***	
	(0.000)	(0.0001)	(0.000)	
Percentage of households with KCC	0.001	-0.001	0.0004	
	(0.001)	(0.001)	(0.001)	
Growth rate of night lights	0.164***	0.049	0.095***	
	(0.040)	(0.030)	(0.027)	
Rural unemployment rate	0.336***	0.087	0.208***	
	(0.101)	(0.072)	(0.068)	
Rural share of agricultural workers among marginal workers	0.161***	0.135***	0.149***	
	(0.031)	(0.023)	(0.021)	
Average rainfall deviation	-0.0001	0.0001**	0.00001	
	(0.0001)	(0.00005)	(0.0001)	
Constant	0.192***	0.213***	0.308***	0.110***
	(0.054)	(0.040)	(0.036)	(0.027)
Observations	45,729	46,647	92,376	92,376
R-squared	0.245	0.260	0.245	0.531
State fixed effects	Yes	Yes	Yes	Yes
Household fixed effects	No	No	No	Yes

Robust standard errors clustered at the household level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix Table 4: Second stage estimates from 2SLS with selection correction (binary dependent variable of whether individual is associated with/main operator of operational holding)

	(2)	(3)	(1)
Variables	Female	Male	All
STM	0.758***	0.383***	0.467***
	(0.156)	(0.103)	(0.118)
Female			-0.199***
			(0.010)
Female * STM			0.230**
			(0.117)
Inverse Mills Ratio	0.021	0.006	0.016
	(0.017)	(0.015)	(0.013)
Observations	45,729	46,647	92,376
Other control variables	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes
Weak identification test (Kleibergen-Paap rk Wald F)	9.279	18.43	14.25
Overidentification test (Hansen J)	2.930	0.402	3.689
Overidentification test p value	0.231	0.818	0.297

Bootstrapped standard errors (clustered at the household level) in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table 5: Probability of women's involvement with the operational holding as the main operator

Variables	(1) 2SRI	(2) Recursive	(3) 2SLS	(4) 2SLS
STM	0.040***	0.026***	0.032	0.375**
	(0.010)	(0.007)	(0.035)	(0.185)
Observations	45,100	45,141	45,141	17,101
Other control variables	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes

Robust standard errors (clustered at the household level) in parentheses. Standard errors in the model in column (1) are bootstrapped. *** p<0.01, ** p<0.05, *p<0.1. The regressions consider only the female sample. Column (1) shows estimation from ordered probit model with two stage residual inclusion method to tackle endogeneity, following Terza et al. (2008), Wooldridge (2015), and Lenze and Klasen (2017). Column (2) shows estimate from ordered probit model that uses a recursive simultaneous equation system to address endogeneity, following Roodman (2011). Columns (3) and (4) uses a standard 2SLS method with linear probability models considering a binary dependent variable. Columns (3) considers the dependent variable taking the value 1 if an individual is the main operator, 0 otherwise. In column (4), the dependent variable takes the value 1 if the individual is main operator, and 0 if she is not an operator (i.e. it only considers movement between not-operator and main-operator). In all the models in columns (1) to (4), the same set of instrumental variables, as discussed in the paper, are used.

Appendix Table 6: First stage results of 2SLS with selection correction

	(1)	(2)	(3)	(4)
	Female	Male	All	
Variables	STM	STM	STM	Female * STM
Share of construction workers	0.096***	0.090***	0.096***	-0.021***
	(0.023)	(0.024)	(0.022)	(0.006)
Rate of short-term migration	0.088***	0.070***	0.070***	-0.006
	(0.018)	(0.017)	(0.016)	(0.004)
Female * Share of construction workers			-0.006	0.139***
			(0.010)	(0.019)
Female * Rate of short-term migration			0.019**	0.102***
			(0.008)	(0.015)
Inverse Mills Ratio	0.025**	0.011	0.019*	0.012**
	(0.012)	(0.014)	(0.012)	(0.006)
Other control variables	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes
Observations	45,729	46,647	92,376	92,376
R-squared	0.046	0.046	0.046	0.062

Standard errors (clustered at the household level) in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table 7: Coefficient estimates from the selection equation (probit) for whether household operated any land for agriculture

	(1)	(2)	(3)
	Female	Male	All
Variables	Opland	Opland	Opland
Proportion of households with land	0.348***	0.404***	0.371***
	(0.132)	(0.140)	(0.129)
Other control variables	Yes	Yes	Yes
Observations	56,060	56,377	112,437
State fixed effects	Yes	Yes	Yes

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

[i] Since the variation of STM is also at the household level, therefore it is not possible to separately identify the effect of STM in a regression that includes household fixed effects. However, the purpose of this regression is only to get a plausible estimate for R_{max}.

[ii] Each of the last three equations involved in the 2SLS estimation includes Inverse Mills Ratio as an explanatory variable. Therefore, to avoid the problem of generated regressor, the standard errors, which are clustered at the household level, are bootstrapped (Wooldridge, 2002).