Gille, Véronique

Conference Paper

Education spillovers in farm productivity: empirical evidence in rural India

Proceedings of the German Development Economics Conference, Berlin 2011, No. 31

Provided in cooperation with:
Verein für Socialpolitik

Education spillovers in farm productivity: empirical evidence in rural India

Véronique Gille*

Centre d'Économie de la Sorbonne, University Paris 1 Panthéon Sorbonne

May 31, 2011

Abstract

Empirical evidence of education spillovers in agriculture in developing countries are scarce and focus on specific channels. This paper provides evidence of such spillovers in rural India, by evaluating the overall impact of education of neighbors on farm productivity. I use cross-sectional data from the India Human Development Survey of 2005. I test the presence of education spillovers by using a Cobb-Douglas production function. Neighbors are defined at the village, caste and occupation level using spatial econometric tools. The complementarity between neighbors and household's education is also tested by adding an interaction term in the specification. The results show that education spillovers do exist: one additional year in the mean level of education of neighbors increases households' farm production by 1.5% *ceteris paribus* . Moreover, the impact of neighbors' education increases with the household level of education. This paper shows the importance for policy makers of taking into account education spillovers and policies' complementarity when facing political trade-offs. It is one of the few to underline that education externalities do not only exist in urban contexts and that education spillovers do not only occur between workers of the manufacturing and service sectors. There are also spillovers in sectors considered as more traditional such as agriculture.

Keywords: Education externalities, Rural India, Farm productivity

JEL Classification Numbers: D13, I25, O12, Q12

*Address: 106-112 bvd de l'Hôpital, Maison des Sciences Économiques, 75647 Paris cedex 13 FRANCE, telephone: +33631210602, e-mail:*veroniquegille@gmail.com*

1 Introduction

In the last twenty years, India has massively invested in education. Its efforts and accomplishments in terms of education have hugely increased with the launching of the Sarva Shiksha Abhiyan program in 2002, which aimed at providing primary education to all 6 to 14 years old children by 2010. Even if this goal has not yet been reached, the number of children out of school has been reduced from 25 million in 2003 to 8.1 million in 2009 (source: World Bank).

This process is considered of major importance for India because as mentioned by the World Bank "Education is one of the most powerful instruments for reducing poverty and inequality. Education is equally key to enhance India's competitiveness in the global economy. Therefore, ensuring access to quality education for all, in particular for the poor and rural population, is central to the economic and social development of India.".

The impact of education on productivity, growth and more generally on development in India has been widely asserted by researchers [Psacharopoulos and Patrinos, 2002]. Nevertheless, education does not induce growth only because it improves individual productivity. Education is a key issue for development also because it has positive externalities, notably in terms of learning spillovers. Consequently education has a higher social return than its private return.

Having a clear idea of education social returns in a rural India is important in terms of education policy. The increase in schooling supply in India came along an increasing part of this supply provided by private schools [Desai et al., 2009]. If social returns of education are high, private financing of education is not optimal. The goal of this paper is consequently to assert the existence of education externalities in rural India, by estimating the overall impact of neighbors' education level on household farm productivity in a first time, and by looking at the complementarity between household education and neigbhors education in a second time. We use spatial econometrics tools to evaluate the spillover effect while taking into account social interactions in Indian villages.

Education externalities are not a new idea in the literature [Marshall, 1890, Lucas, 1988]. Theoretically, there are several potential channels leading to human capital having a higher social than private return. They can be classified into two types. The first channel is a broad one: education has a higher social return at a community level such as a city or a State because education reduces the probability of getting involved in activities which produce negative externalities. It also increases the probability of engaging in activities with positive externalities [Moretti, 2003].

The second channel is what is called a spillover effect. It occurs at the individual level. As Kremer [1993] assumes in his O-ring theory, the human capital of a worker may have a marginal return which grows with the human capital of other workers. In the context that we are concerned with in this paper, that is to say rural households in developing countries, how can the education level of neigbhors increase own productivity? The reasons are three-fold. First, there can be learning spillovers in technology adoption. This phenomena has been widely asserted in the literature [Foster and Rosenzweig, 1995, Besley and Case, 1994, Munshi, 2004, e.g]: as more educated people are more prone to adopt new technologies, and as neighbors are highly influenced by their neighbors behavior in technology adoption, households which have neighbors with higher level of education may be more prone to adopt new technologies. Second, neighbors can influence the efficiency with which one is using its inputs, if we make the hypothesis that all farmers are not on the production frontier [Weir and Knight, 2007]. Third, neighbors can also have an "allocative effect", which is the efficiency with which farmers choose their inputs or output given their relative prices [Kumbhakar and Lovell, 2003].

Existing literature mainly focuses on education externalities in cities or in firms in developed countries. To my knowledge there are only three papers which study education externalities in rural areas in developing countries. The first one is a paper written by Appleton and Balihuta [1996]. They look at the external effect of education on agricultural productivity in rural Uganda by introducing the average level of education of neighbors in the production function of farmers. They find that the average primary schooling of other farm workers in the area significantly raises own productivity. It is a first attempt to take into account neighbors' education impact. However, as they underline themselves, their results are limited by their data. As neighbors education is calculated at the community level, they cannot control for omitted community effects, which may upward bias the impact of neighbors education level. Furthermore, due to data constraints, their definition of neighborhood is very broad. Weir and Knight [2007] estimate average and stochastic production frontier with neighbors' education as control variable in rural Ethiopia. They control for village fixed effects. They also find a positive effect of neighbors' education on average production, but they don't find any impact of neighbors' education on farmers' efficiency. Asadullah and Rahman [2009] use the same methodology to study external returns of education in agriculture in Bangladesh. They control for village fixed effect while defining neighbors at a lower level called "Bari". They find no evidence of external returns of education on farm productivity. However, their results may be driven by their neighborhood definition: in the sample design, their is only two households selected by "Bari". So when Asadullah and Rahman [2009] look at the external impact of education, they just look at the impact of the other household selected in the "Bari". This definition of neighbors allows for village dummies but may be too restrictive to capture any external effect.

The contribution of my paper to this literature is threefold. First, it is the first paper to study the presence of education externalities in rural India. To my knowledge, only specific channels have been studied, such as the impact of neighbors behavior on technology adoption. Second, it takes into account social interactions in the definition of neighborhoods in villages which have not been accounted for in this literature. Third, it allows for complementarities between household's education level and neighbors education level.

According to the results, it seems that there are education spillovers in rural India and that these spillovers increase with the own level of education. One additional year in the mean level of education of neighbors increases in mean farm production by 1.5%.

The paper is organized as follows. Section 2 describes the theoretical framework on which the empirical specification is based, the empirical strategy with the neighborhood definition and the econometric issues and the data used. Section 3 presents the results with three different specifications. Section 4 tests the robustness of the results by testing some hypotheses made and section 5 concludes.

2 Empirical methodology and issues

2.1 Theoretical Framework

The empirical strategy is based on the theoretical model of Lucas [1988] where the productivity of a worker depends on the aggregate human capital of his coworkers. Suppose a Cobb-Douglas production function where the output y of a household farm i is :

$$y(t) = Ak(i)^{\alpha}h(i)^{\beta}n(i)^{\gamma}l(i)^{\delta} \tag{1}$$

Where A is aggregate productivity, k is the physical capital, h is the amount of human capital of the household, n is the number of workers, l is the land.

The fact the the productivity of a household farm depends on the surrounding human capital can be captured by allowing A to vary depending on this surrounding human capital. In other words, according to Lucas [1988], A can be written as

$$A = BH^{\Delta} \tag{2}$$

where H is the average human capital in the neighborhood of a household farm. Δ measures consequently the external effect human capital. If we reformulate the equation in logs, we have:

$$lny_i = lnB + \Delta lnH + \alpha lnk_i + \beta lnh_i + \gamma lnn_i + \delta lnl_i \qquad (3)$$

2.2 Empirical specification

The influence of neighbors' human capital on farm productivity is estimated following equation (3). Consequently the equation estimated is

$$lny_{ivr} = B + \Delta W_i E + \alpha k_i + \beta lnh_i + \gamma lnn_i + \delta lnl_i + \theta X_i + \lambda Z_v + c_r + u_i \qquad (4)$$

Where y_i is aggregated farm production of household i, k_i is a dummy variable which is equal to one if the household owns any agricultural equipment, h_i is the mean level of education of the most educated man and woman of the household, n_i is the number of days worked on the land, l_i is the amount of land cultivated, X_i is a set of other control household variables which influence productivity. Z_v is a set of village-level variables, c_r is a regional fixed effect and u_i is a household error term which control for other determinants of household's productivity.

The level of education of neighbors is taken into account by using spatial econometric tools: W_i is the ith row of a matrix W which allocates to each household its neighbors. More precisely, each element w_{ij} of W is defined as follows: $w_{ij} = \theta$ with $\theta = 0$ if i and j are neighbors and $w_{ij} = 0$ otherwise. $w_{i,i} = 0$ for all i to exclude the household's level of education from the calculation of neighbors' educational level. E is a column vector whose elements represent the level of education of each household in the database. As each row W_i is normalized such that $W_i = 1$ for all i, $W_i E$ is an average of neighbors' education level.

2.3 Neighborhood definition

One issue in estimating neighborhood impact is neighborhood definition. Goux and Maurin [2007] underline that "distant neighbors have less influence than close ones" and that using a too broad definition of neighbors can lead to an underestimate of the influence of close neighbors. To control for this bias we test several definitions of neighborhood. These definitions are based on geographic contiguity, caste group membership and occupation.

The first matrix we use to define the neighborhood is a simple contiguity matrix: according to the literature on neighborhood effect in agriculture in India [Foster and Rosenzweig, 1995, Munshi, 2004], the social unit where interactions occur is the village. So we define neighbors as people from the same village. Furthermore we assume that each neighbor has the same influence on farm productivity. In other words, in a first time, $w_{i,j} = 1$ if i and j are from the same village, $w_{i,j} = 0$ otherwise. The consequence of this construction is that each neighbor's observation has equal weights: $w_{i,j} = w_{i,k}$.

For the second matrix we use a definition of neighborhood based on social groups. Due to a lack of details in our data, we cannot define very precisely social groups. Nevertheless, the specific social organization of India in castes helps us understanding social interactions occurring in villages.

Srinivas [1962] defines a caste as an "hereditary, endogamous group which is usually localised. It has a traditional association with an occupation, and a particular position in the local hierarchy of castes. Relations between castes are governed, among other things by the concepts of pollution and purity, and generally maximum commensality i.e. interdining occurs within the caste". Castes are traditionally specialized in a specific occupation which make the different groups interdependent from each other. So castes groupes interact between each others. Nevertheless, castes at the bottom of the traditional hierarchy are considered as impure and are ostracized by people from other castes in the village. Consequently they are less likely to interact with people from other castes[1].

Consequently, in this matrix $w_{i,j}$ is equal to 1 if i and j are from the same village and from the same social group, 0 otherwise. Again $w_{i,i} = 0$ for all i and the rows of the matrix are normalized.

The third matrix we use is a matrix based on neighbors' occupation. i and j are considered as neighbors if they do the same occupation. $w_{1,i,j} = 1$ if i and j are from the same village and the main occupation of j is agriculture, 0 otherwise.

Finally, we compute a fourth matrix which is a compilation of the two previous ones: i and j are considered as neighbors if they are from the same caste and they both cultivate land.

2.4 Identification strategy

A second issue when estimating neighbors impact is the effect mentioned by Manski [1993] as the "correlated effect": unobserved variables which impact farm productivity can be correlated to neighbors' education and the impact may be overestimated.

To solve this issue two solutions are used: in a first specification we add village-level variables and regional dummies. Regional dummies control for omitted variables at an aggregate level. The region is a statistical entity which has been constructed by the National Sample Survey Organization (NSSO) to conduct its surveys. Regions are composed of several districts of the same State and they are homogeneous in their "agro-climatic conditions and socio-economic features" [Murthi et al., 2001]. Consequently regional dummies control in particular for geological omitted variables, such as quality of land (dryness, fertility, etc.) or climate variables. Village level variables control for correlated effects at a local level. I add three variables which may influence village education level through the supply side because people have better access to schools, or through the demand side. The first variable is the distance of villages to the nearest city to proxy for the isolation of each village. The second

[1] For more information on castes, please refer to DeliÃ¨ge [2004].

variable which is the proportion of household having electricity in the villages also captures the isolation of the village. And the third variable proxies for credit access in the village: it is a dummy variable which is equal to 1 if there is a bank branch office or a credit cooperative in the village.

Nevertheless, adding village-level variables and regional fixed effects may not be sufficient to get rid of the correlation with the error term. So in a second time we add village dummies. But this specification may not efficiently estimates the coefficient of of neighbors' education variable because this variable may have a small within-village variance. In particular, when we define neighbors as households from the same village, the variable "education of neighbors" is the mean level of education of other households in the village. So this variable varies for each household only depending on its own level of education[2]. It is the same when neighbors are defined as households who also cultivate land. But definitions of neighbors as households from the same caste or from the same caste and occupation allow for more within village variance.

Even if adding village dummies is the best specification we can do to suppress the endogeneity, we must keep in mind that estimation with village dummies does not solve the problem of time-dimension unobserved characteristics. Consequently the coefficient associated with neighbors' education level may be a little upward biased.

Under the hypothesis that migration has not as a consequence to sort people according to factors influencing their farm productivity, this model can be consistently estimated with OLS estimates [Gallo, 2000]. This hypothesis is credible because in India, there is a really low rate of migration. According to the Indian Census of 2001, 72.4% of the rural population was born in their place of residence. The migrants are mainly women for reason of marriage.

[2] Imagine there are 3 households in the village which have respectively 0, 3 and 6 years of education. Neighbors education for the first household is $(3+6)/2$, for the second household it is $(0+6)/2$ and for the third household it is $(0+3)/2$. So for each household the variable "education of neighbors" only varies of the household level of education divided by the number of neighbors. The more the number of neighbors, the less within village-variance of this variable.

If we only consider men, 89.1% were born in their current place of residence. For this reason, many authors choose to ignore the migration problem in rural India [Banerjee et al., 2007, Foster and Rosenzweig, 1995, Anderson, 2005]. We follow this work in our study. Nevertheless, this hypothesis is discussed in part 4.

2.5 Data: India Human Development Survey

The database used to evaluate human capital spillovers is the India Human Development Survey. This survey was jointly conducted in 2005 by researchers from the University of Maryland and the National Council of Applied Economic Research, New Delhi. It took place in all States and Union Territories (UT) of India, with the exception of the Islands of Andaman and Nicobar and Lakshadweep. These places have not been surveyed because of their small population. Across these 33 States and UT, 41,554 households in 1503 villages and 971 urban neighborhoods were interviewed.

The goal is to evaluate the impact of neighbors' human capital in agriculture, so I only keep rural households. Since the survey's drafting, 19 villages of the sample have been classified as urban zones by the 2001 National Census of India. We take the census definition to define rural households. This leads to 26,734 households.

The education of neighbors variable is calculated with the whole set of rural household. That is to say I don't only consider the impact of neighbors who also cultivate land but the impact of all neighbors in the village. Out of these 26,734 rural households, 43 have missing values for the educational level. So the education of neighbors is calculated with 26,691 households.

For the equation estimation, I only consider rural households who cultivate land that is to say 14,298 households. Out of these households, 4034 have missing values for one of the variable[3]. Consequently the estimation is made on 10,264 households. The number of obser-

[3] In particular, due to logistical constraints the interviewers were only able to complete 1454 village questionnaires, resulting in 49 villages being omitted. According to the survey managers, there were no consistent pattern to these omissions.

vations slightly decrease when the definition of neighbors becomes more restrictive because some households in the database don't have neighbors from the same caste or neighbors who also cultivate land.

Variables used are described in table 1 and in table 5 in the appendix.

Table 1: DESCRIPTIVE STATISTICS

	mean	sd	min	max
Production (in log)	7.16	1.37	1.50	13.65
HH Education	5.08	4.05	0	15
Neighbors' Education	4.43	2.07	0	11.53
Days worked (in log)	5.68	.83	0	8.33
Land cultivated (in log)	1.03	1.10	-6.21	5.97
Inputs (in log)	7.76	1.33	3	13.64
Agr. equipment	.095	.293	0	1
Irrig. land	.61	.49	0	1
Nb of seasons	1.73	.60	1	3
Distance to town	14	10.91	1	85
% of HH with elect	63.1	35.01	0	100
Credit access	.46	.50	0	1
Observations	10264			

3 Results

3.1 Specification with village-level variables and regional dummies

The results of the estimation with village level variables and regional dummies are given in table 2. Column (1) shows the results for the estimation where neighbors are defined

Table 2: RESULTS WITH VILLAGE-LEVEL VARIABLES

	(1) Village	(2) Caste	(3) Occupation	(4) Caste and Occupation
HH Education	0.0109***	0.0103***	0.0107***	0.0103***
	(0.00213)	(0.00215)	(0.00207)	(0.00211)
Neighbors' Education	0.0176**			
	(0.00782)			
Same Caste group		0.0180***		
Neigh's educ level		(0.00544)		
Same Occupation			0.0155**	
Neigh's Educ level			(0.00610)	
Same Occupation and Caste				0.0147***
Neigh's Educ level				(0.00485)
Days worked (in log)	0.0520***	0.0543***	0.0524***	0.0519***
	(0.0170)	(0.0171)	(0.0172)	(0.0175)
Land cultivated (in log)	0.547***	0.548***	0.548***	0.551***
	(0.0183)	(0.0184)	(0.0184)	(0.0185)
Inputs (in log)	0.391***	0.387***	0.391***	0.387***
	(0.0184)	(0.0185)	(0.0185)	(0.0188)
Agr. equipment	0.107***	0.106***	0.101***	0.102***
	(0.0285)	(0.0285)	(0.0285)	(0.0286)
Irrig. land	0.276***	0.274***	0.276***	0.279***
	(0.0285)	(0.0284)	(0.0285)	(0.0288)
Nb of seasons	0.0876***	0.0879***	0.0870***	0.0867***
	(0.0257)	(0.0257)	(0.0258)	(0.0262)
Distance to town	-0.00260*	-0.00257*	-0.00266*	-0.00255*
	(0.00150)	(0.00150)	(0.00150)	(0.00154)
% of HH with elect	0.000641	0.000716	0.000671	0.000743
	(0.000573)	(0.000562)	(0.000574)	(0.000575)
Credit access	0.0316	0.0314	0.0305	0.0302
	(0.0284)	(0.0284)	(0.0283)	(0.0286)
Constant	2.049***	2.101***	2.063***	2.126***
	(0.466)	(0.472)	(0.463)	(0.482)
Observations	10264	10173	10211	9927
r2	0.751	0.752	0.751	0.752

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

as households from the same village. Column (2) defines neighbors as households from the same village and from the same caste. Column (3) gives the results for the estimation where neighbors are defined as household also cultivating land in the village. Finally column (4) shows the results for the most restrictive definition of neighbors: neigbhors are households from the same caste and who also cultivate land.

Whatever the definition of neighbors is, the impact of neighbors education on farm production is significantly positive. Moreover, the coefficients are very stable across the different specifications: one year increase in the mean level of education of neighbors seems to increase farm production between 1.5 and 1.8% *ceteris paribus* .

In the four specifications, almost all the control variables are statistically significant at a 0.1% level and have the expected signs. The three factors of production, land, labor and inputs are significant and increase production. The addition of their coefficient give an idea of the returns to scale: here it seems that the returns to scale are constant, as the sum of the three coefficient is around 1 in the four specifications. This result is in conformity with the literature on India [Munshi, 2004]. Not surprisingly, production also increases with agricultural equipment, irrigation and the number of seasons where the land is cultivated. In the village level variables, only the variable "Distance to a city" has an impact: as expected it decreases productivity. But having electricity in the village and the variable "credit access" have no impact. So these village-level variables don't seem to control efficiently for the correlation between the error term and our variable of interest the education of neighbors.

3.2 Specification with village dummies

Table 3 shows the results of the estimation with village dummies. As anticipated, the two definitions of neighbors which have by construction a low within village variance are not anymore significant. However, when neighbors are defined as households from the same caste in the same village or households from the same caste and who also cultivate land in

Table 3: RESULTS WITH VILLAGE DUMMIES

	(1) Village	(2) Caste	(3) Occupation	(4) Caste and Occupation
HH Education	0.0114***	0.0105***	0.00976***	0.0109***
	(0.00358)	(0.00211)	(0.00287)	(0.00213)
Neighbors' Education	0.0253			
	(0.0495)			
Same Caste group		0.0175**		
Neighbors' Educ level		(0.00711)		
Same Occupation			-0.00225	
Neighbors' Educ level			(0.0198)	
Same Occupation and Caste				0.0130**
Neighbors' Educ level				(0.00658)
Days worked (in log)	0.0728***	0.0740***	0.0728***	0.0710***
	(0.0155)	(0.0156)	(0.0155)	(0.0159)
Land cultivated (in log)	0.610***	0.611***	0.610***	0.620***
	(0.0198)	(0.0199)	(0.0197)	(0.0191)
Inputs (in log)	0.307***	0.304***	0.307***	0.300***
	(0.0183)	(0.0183)	(0.0182)	(0.0182)
Agr. equipment	0.130***	0.132***	0.130***	0.135***
	(0.0283)	(0.0283)	(0.0282)	(0.0282)
Irrig. land	0.282***	0.278***	0.282***	0.281***
	(0.0332)	(0.0330)	(0.0331)	(0.0335)
Nb of seasons	0.0891***	0.0903***	0.0892***	0.0865***
	(0.0242)	(0.0241)	(0.0241)	(0.0244)
Constant	3.232***	3.289***	3.364***	3.347***
	(0.276)	(0.142)	(0.172)	(0.143)
Observations	10475	10383	10416	10124
r2	0.836	0.837	0.835	0.838

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

the village, the spillover effects persists with the village dummies and its level is not really different than in the previous estimation: when neighbors are defined as households from the same caste and the same occupation, one additional year of education in the neighborhood increases farm production of 1.3%. When neighbors are defined as people from the same caste only the impact is of 1.8%.

3.3 Interaction effects

Till now, the assumption I make is that the education of neighbors is a substitute to the household level of education. But the impact of the education of neighbors may actually depends on the level of education of the household. To check this issue, I add an interaction term between the level of education of the household and the mean level of education of the neighborhood. The estimation is made only with the two definitions of neighbors which stayed significant when adding village dummies. The results are shown in table 4.

The table only shows the coefficient of the variables of interest, household's education, neighbors' education and the interaction term between these two variables. We can see that in the four specifications, the coefficient of the household education and of neighbors education is not anymore statistically different from 0. However what really matters is the joint significance of each of these two coefficients with neighbors education level. As we can see in the bottom of the table, the coefficients of household's education level and of the interaction terms are jointly significant as well as the coefficients of neighbors's education level and of the interaction term. The interpretation is then that neighbors' education level has an impact on farm production and that this effect increases with the level of education of the household. The impact of neighbors education for some level of household education are shown in the bottom of table 4. When a household is not educated (0 years of education), the education level of its neighbors has no impact. But from 3 years of education onwards, production increases with neighbors education (except if we consider only neighbors from

Table 4: INTERACTION TERMS

	Same caste	Same caste Village Dummies	Same caste and Agr	Same caste and Agr Village Dummies
HH educ	-0.00189	-0.00182	-0.00351	-0.00280
	(0.00483)	(0.00465)	(0.00513)	(0.00495)
educ same caste	0.00528	0.00557		
	(0.00716)	(0.00826)		
HHeduc * educ same caste	0.00252***	0.00256***		
	(0.000966)	(0.000949)		
educ same caste and agr			0.00148	0.00128
			(0.00666)	(0.00765)
HHeduc * educ same caste and agr			0.00262***	0.00266***
			(0.000929)	(0.000939)
Observations	10173	10371	9927	10114
r2	0.752	0.837	0.753	0.838
Joint significance of γ and Θ_1	0.000	0.000	0.000	0.000
Joint significance of Θ_1 and Θ_2	0.000	0.002	0.000	0.003
Neighbors impact				
education HH : 0 ans	0.005	0.006	0.001	0.001
education HH : 3 ans	0.013**	0.013*	0.009*	0.009
education HH : 5 ans	0.018***	0.018**	0.014***	0.015**
education HH : 10 ans	0.030***	0.031***	0.027***	0.028***
education HH : 15 ans	0.043***	0.044***	0.039***	0.041***

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

γ, Θ_1 and Θ_2 are respectively the coefficients of HH education, interaction term and neighbors education the same caste and who also cultivate land, who have an impact only when a household have 5 years of education at least). When the household has the maximum number of years of education (15 years), one additional years of education in the mean level of education of the neighborhood increases farm production by 4%.

4 Migrations robustness check

The results obtained are based on the hypothesis that migration is too small to be taken into account. Actually, even if the rate of migration is very low, if people choose to migrate close to people they look like according to factors influencing farms' productivity, it leads to an overestimation of the neighborhood influence. All the robustness checks are only shown with neighbors considered as households from the same caste because according to the results it is the best definition of neighborhood. All the estimations are with village dummies.

As our data let us know how long each household has lived in the village, we can differentiate between "old households" and "new households". So we estimate our regression without new households. The caste-weighted education of neighbors variable is re-calculated excluding "new households". Column (1) of table 6 in the appendix is the result of the benchmark regression. Column (2) excludes people who arrived in the last two years before the survey. Column (3) takes a larger definition of "new people": it excludes people who arrived in the last ten years in the village. The results show that whatever the definition of new households is, when we eliminate these people from the sample, the results stay the same. The spillover effect is between 1.6% and 1.7%.

But this specification takes only into account the situation where the whole household has moved. Actually, the principal source of migration from village to village is women migration for marriage purpose [Rosenzweig and Stark, 1989]. To check if this kind of migration can create biases in our results, I evaluate the knowledge spillover without taking into account women's level of education with the reduced sample without "new households"[4].

[4]New households here are households who are in the village since less than three years.

The results are shown in column (4) of table 6. Neighbors' education coefficient is not anymore significant. However, as there are village dummies in this estimation, this lack of significance may be due to a low within village variation. To check if it is the case, I just keep villages where there are more than ten households which cultivate land. In this new specification showed in column (5), the coefficient of neighbors education level is significant at a 10% level. Neighbors education level seems to have an impact, even when I do not consider women. Consequently it seems that migration is not a real issue.

5 Conclusion

A huge emphasis has been put on education policies in developing countries in the last ten years, notably thanks to the Millennium Development Goals. There is no more need to prove the beneficial impact of education policies on a social point of view and on revenues and growth. But education has a broader impact than its private one: education also produces externalities. In particular, at the individual level, education may have a spillover effect: the education level of your coworker/neighbor may influence your productivity. This has been empirically tested in firms in developed countries but there is a lack of evidence for developing countries, in particular for rural workers. Though, if there are education spillovers, it has policy implications especially when there is a trade-off with other policies.

In this paper, my goal is to test and evaluate the existence of education spillovers in a rural context in India. If there are spillovers, we expect that the education level of the neighborhood has a positive impact on households' farm productivity.

I use cross-sectional data from the India Human Development Survey of 2005. To be sure that the definition of neighborhood does not drive the results, we test four different definitions. For the four specifications we find that education spillovers do occur. This results are persistent when adding village dummies for two definitions of neighbor : one additional year in the mean educational level of neighbors seems to increase in mean farm productivity by 1.7%. The robustness checks confirm these results. Moreover, this impact increases with household's level of education.

These findings underline three important concerns. First, education externalities do not only exist in urban contexts and education spillovers do not only occur between workers of the manufacturing and service sectors. There are also spillovers in sectors considered as more traditional such as agriculture.

Second, these findings confirm (again) the choice of improving education in developing countries: giving a child education will certainly provide him greater revenues but it may also provide his neighbors greater revenues, because they will be more productive! Consequently, education has a multiplicative effect, which politics should not forget while deciding resources allocation.

Finally, it opens the way to further research. In particular, this paper does not explore the channels through which this spillover effect happens. The impact of the education of neighbors on farm productivity may be alternatively or conjointly due to the direct impact of neighbors' education on productivity, through learning spillovers for example, or to the indirect impact of neighbors' education level through neighbors' productivity. Further research is needed to clarify this issue. Moreover, as underlined by Manski [1993], "social effects might be transmitted by distributional features other than the mean". This could also be a room to explore.

References

Siwan Anderson. Caste as an impediment to trade. *University of British Columbia Mimeo*, 2005.

Simon Appleton and Arsene Balihuta. Education and agricultural productivity: Evidence from uganda. *Journal of International Development*, 8(3):415–444, 1996.

M. N. Asadullah and S. Rahman. Farm productivity and efficiency in rural bangladesh: the role of education revisited. *Applied Economics*, 41(1):17–33, 2009.

Abhijit Banerjee, Lakshmi Iyer, and Rohini Somanathan. Public action for public goods. Technical report, National Bureau of Economic Research, Inc, February 2007.

Timothy Besley and Anne Case. Diffusion as a learning process: Evidence from HYV cotton. Technical report, Princeton University, Woodrow Wilson School of Public and International Affairs, Research Program in Development Studies., May 1994.

Robert DeliÃ¨ge. *Les castes en Inde aujourd'hui*. Presses Universitaires de France - PUF, June 2004.

Sonalde Desai, Amaresh Dubey, Reeve Vanneman, and Rukmini Banerji. Private schooling in india: A new educational landscape. *India Policy Forum*, 5:1–38, 2009.

Andrew D Foster and Mark R Rosenzweig. Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of Political Economy*, 103(6): 1176–1209, 1995.

Julie Le Gallo. Econométrie spatiale 1 -Autocorrélation spatiale. Technical report, LATEC, Laboratoire d'Analyse et des Techniques EConomiques, CNRS UMR 5118, Université de Bourgogne, June 2000.

Dominique Goux and Eric Maurin. Close neighbours matter: Neighbourhood effects on early performance at school. *Economic Journal*, 117(523):1193–1215, 2007.

Michael Kremer. The O-Ring theory of economic development. *The Quarterly Journal of Economics*, 108(3):551–75, 1993.

Subal C. Kumbhakar and C. A. Knox Lovell. *Stochastic Frontier Analysis*. Cambridge University Press, March 2003. ISBN 9780521666633.

Robert Jr. Lucas. On the mechanics of economic development. *Journal of Monetary Economics*, 22(1):3–42, 1988.

Charles F Manski. Identification of endogenous social effects: The reflection problem. *Review of Economic Studies*, 60(3):531–42, 1993.

Alfred Marshall. *Principles of economics*. Macmillan and co., limited, 1890.

Enrico Moretti. Human capital externalities in cities. Technical report, National Bureau of Economic Research, Inc, April 2003.

Kaivan Munshi. Social learning in a heterogeneous population: technology diffusion in the indian green revolution. *Journal of Development Economics*, 73(1):185–213, 2004.

M. Murthi, P. V. Srinivasan, and S. V. Subramanian. Linking indian census with national sample survey. *Economic and Political Weekly*, pages 783–792, 2001.

George Psacharopoulos and Harry Anthony Patrinos. Returns to investment in education : a further update. Technical report, The World Bank, September 2002.

Mark R Rosenzweig and Oded Stark. Consumption smoothing, migration, and marriage: Evidence from rural india. *Journal of Political Economy*, 97(4):905–26, 1989.

Mysore Narasimhachar Srinivas. Caste in modern india and other essays, 1962.

Sharada Weir and John Knight. Production externalities of education: Evidence from rural ethiopia. *Journal of African Economies*, 16(1):134–165, 2007.

Appendix

Table 5: DESCRIPTION OF VARIABLES

Variable	**Description**
Dependant variable	
Production	Household's farm agricultural production: total production in rupees
Educational Level	
HH Education	Household's level of education: mean of the years completed at school (0-15) of the most educated female adult (> 21 years) and of the most educated male adult
Neighbors' Education	Neighbors education level: Mean of the level of education of other households in the same village
Caste-weighted Neigh's Educ level	Mean of the level of education of other households in the same village (depending on the caste)
Occup-weighted Neigh's Educ level	Mean of the level of education of other households in the same village (depending on the occupation)
Control variables	
Days worked	Number of hours worked per year per household on the field in logs
Agr. Equipment	Scored 1 if the household owns tractors, threshers or bio-gas plants
Irrig. Land	Scored 1 if the land is irrigated
Fertilizer p.a.	Quantity of fertilizer used per acre
Nb of seasons	Number of times the land is harversted in a year
Distance to town	Distance of the household's village to a town
% of HH with elec	Percentage of households having electricity in the village
Credit access	Scored 1 if the village has bank offices or credit cooperatives

Table 6: MIGRATION ROBUSTNESS CHECK

	(1) Baseline	(2) > 2 years	(3) > 9 years	(4) Men > 2 2 years	(5) Men > 2 years
HH Education	0.0105***	0.0102***	0.0103***		
	(0.00211)	(0.00211)	(0.00211)		
HH17 ed5 Highest				0.00717***	0.00804***
male 21+ ed				(0.00170)	(0.00191)
Same Caste group	0.0175**	0.0173**	0.0164**	0.00750	0.0142*
Neigh's educ level	(0.00711)	(0.00709)	(0.00707)	(0.00585)	(0.00724)
Village dummies	YES	YES	YES	YES	YES
Observations	10383	10364	10303	10051	7465
r2	0.837	0.837	0.838	0.837	0.826
ll	-8582.9	-8556.1	-8472.6	-8249.1	-6317.6

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$