Building Schools: Shrinking Farm Labor Supply
Draft

Diane Elise Charlton and J. Edward Taylor
Department of Agricultural & Resource Economics
University of California, Davis

April 23, 2014

I. Introduction

Throughout the world, economic development is correlated with a rapid decline in the agricultural labor force. As the per capita income rises, the share of the population working in agriculture quickly diminishes. There is yet little research investigating the underlying mechanisms behind this phenomenon. In this paper, I investigate the role of human capital investments in the transition off the farm in rural Mexico. Does improved access to schools reduce the probability that individuals do farm work?

I investigate this question using nationally representative household panel data from rural Mexico and local school supply data provided by the federal government. The probability of working in agriculture from rural Mexico is downward trending between years 1980 and 2010. Simultaneously younger generations in rural Mexico are completing more years of education than previous generations did. The average years of school completed for rural Mexicans ages twenty to twenty-nine in 2010 is almost nine years while the average years of schooling for ages fifty to fifty-nine is only five years. This rapid increase in education across generations likely affects the rural-born labor force in many aspects, from the average productivity and wages of workers within a given labor sector to the share of workers who select to work in each labor sector.

One of the primary challenges in measuring the impacts of education on labor market outcomes is that individual educational attainment is usually endogenous. The expected returns to education in the labor sector where a child believes he will work in the future will influence how many years the child goes to school, giving rise to bias from reverse causality. Educational attainment is also likely correlated with unobservable individual characteristics, such as ability or discipline, which are correlated with labor sector decisions. Identification in a model measuring the effects of education on labor sector choice must then depend on an exogenous change in education seen over time. Rural Mexico saw a wide expansion of schools throughout the 1990s and 2000s. I use variation in access to schools within communities across cohorts as a means of identification. I measure the impacts of building primary, secondary (jr. high equivalent), and upper-secondary (high school equivalent) schools in one's home village on the probability that an individual works in the farm sector as an adult. If the returns to education are greater in the non-farm sector and construction of schools in one's home village makes education less costly, then theory predicts that more individuals will invest in higher levels of education and work in the non-farm sector as schools become more accessible.

I investigate the marginal effect of improving local access to schools on the farm labor supply using village-fixed effects and state-by-year fixed effects. I find that access to secondary school education within one's own village significantly reduces the probability of working in agriculture while I find no effect of building primary or upper-secondary schools in the local village.

These findings have important implications for the farm sectors in Mexico and the United States, which both depend on labor supply from rural Mexico. As education continues to rise in rural Mexico, fewer young adults will choose to work in the farm sector. The agricultural industry will have to adapt by seeking out different sources of labor (presumably from a country and region where education is not on the rise), or by learning to produce more with fewer workers. The latter seems like a more viable solution for the long run. By increasing mechanization in the farm industry, the average productivity of workers will rise, increasing wages. Technological advances in the farm industry will increase the returns to education within farm jobs, thereby reducing the demand for farm labor overall and also helping the farm sector to retain more workers with higher levels of education.

II. Literature

Several studies indicate that improved access to education has positive impacts on the average years of schooling attended. **?** find that increased availability of schools in India increased the levels of education and **?** also finds that levels of education increased for cohorts in villages in Indonesia that benefitted from a massive school construction program in the 1970s. These studies further find that school construction had a positive impact on wages for the students that benefited from the increased school supply. These results are not limited to developing countries. Empirical evidence from the United States also shows that proximity to higher levels of education has significant positive impacts on college attendance (**??**). The abundance of evidence across countries suggests that the physical location of schools relative to where individuals live is a strong indicator of total years of school attendance.

? observes that access to higher levels of education may, in fact, affect schooling decisions at a very young age. His results from an empirical study in rural Ghana demonstrate that improved access to secondary schools has a significant impact on primary school enrollment. He notes that the marginal returns to a primary school education measured from urban wage data in Ghana are near zero while the marginal returns to secondary education, in contrast, are substantial. This suggests that if a family does not have access to affordable secondary education, they may not enroll their children in primary education either. Furthermore the costs of education in rural areas may be primarily indirect, including the opportunity costs of time spent in travel to and from school. Consequently, if a child can only attend primary school, and the opportunity cost of time is high, then families may not enroll their children in school at all. In fact, Lavy finds that the probability of ever attending primary school with respect to the distance to a primary school is 0.10, and the elasticity with respect to the distance to a secondary school is 0.30. The cross-elasticity of primary school enrollment with respect to secondary school availability is considerably greater than the direct elasticity with respect to primary school availability. These results demonstrate access to higher levels of education have implications from the beginning of an individual's education. The impact of building a secondary school might be much larger than expected if students would not otherwise enroll in primary school.

Another study from rural Ghana investigates how school access affects farm and non-farm wages. **?** is one of the few studies on the marginal returns to education that makes a clear distinction between the farm and non-farm sectors with predictive consequences for sector selection. Joliffe attempts to measure the returns to education in the farm and non-farm sectors in rural Ghana using a two-step selection model. The first step of the model estimates whether individuals work in the farm or non-farm sector. The second step estimates the effect of an additional year of schooling on income conditional on sector selection. The results indicate that the returns to education are greater in the non-farm sector. Consequently, as education increases households allocate more labor to non-farm work. Joliffe examines self-selected education only however and does not investigate how changes in the supply of education affect labor allocation.

In addition to supply, the demand for higher education is also an important determinant of school attainment, and consequently wages. **?** uses empirical evidence from Mozambique to demonstrate that programs to increase adult literacy also increase children's school enrollment by shifting the demand for schooling outwards. There has been a significant shift in demand for schooling throughout Mexico in recent years. In the nineteen-nineties, Mexico began a campaign to increase school attainment in poor communities through demand-side interventions. The government program Progresa gives cash transfers to eligible families conditional on children's school attendance and take-up of health services. The program was designed to reduce poverty and it thus targets the poorest communities and families in the country. The program was implemented using a random roll-out design and studies indicate that the program was effective at both targeting the poorest families and at increasing school attainment (**??**). The outward shift in demand for schooling generated by Progresa along with outward shifts in supply likely both impact labor allocation of the affected generations. In this study, I focus on the supply-side effects of education. I control for shifts in demand for education by using village fixed effects and village-specific time trends. Since Progresa was rolled out randomly across villages, it should not confound the results focused on school supply in this study once the controls are included.

The above papers provide important insights into the effects of school supply and demand on the take-up of schooling and on wage outcomes, but they do not illuminate the effects of school supply on sector choice. This paper seeks to investigate how local school supply affects the probability of working in agriculture, and ultimately what role education plays in the transition of populations out of farm work.

III. Data

I use panel data on individuals from a nationally representative sample of rural Mexican households along with community-level data on school supply. The Mexico National Rural Household Survey (Spanish acronym *ENHRUM*[1]) is unique in providing both retrospective and panel data on migration from rural Mexico to both the United States and destinations within Mexico. Rural localities were selected using a sampling frame designed by Mexico's national census office (Instituto Nacional de Estadística, Geografía e Informática (INEGI)) to be nationally and regionally representative. Rural localities were defined as having populations between 500 and 2,499 inhabitants in 2002.

The panel data come from three survey rounds: 2002, 2007, and 2010. Each round collected detailed information on migration destinations, whether migrants worked in the agricultural or non-agricultural sector, and employment status (wage-earner or self-employed) for all family members, including the household head, his/her spouse, all others living in the household, and children of the household head and spouse living outside the household in the year prior to each survey. The survey gathered the annual work histories of every member of the household with retrospective data going back to 1980. For each

[1] *Encuesta Nacional a Hogares Rurales de Mexico*; Spanish acronym ENHRUM

household member, the primary occupation (that which earned the most income or where he/she worked the most hours) is recorded for a job in the home village, elsewhere in Mexico, and in the United States. Thus, each member can record up to three occupations in a given year if he/she worked in three different locations.

It should be noted that the sample size differs somewhat between survey rounds. Due to the costs of gathering retrospective data, work histories were not gathered for a random sample of individuals selected from each household in 2002. Work histories back to 1990 were collected for all individuals in the household in 2007. Consequently, a random selection of work histories are missing for years 1980-1990. Households were additionally dropped from the survey in 2010 due budget constraints and increased violence in certain communities. Communities were dropped from the survey in 2010 so as to maintain a nationally representative sample of rural Mexico however (with the possible exception of characteristics that may be correlated with violence in some extreme cases). The number of working age individuals, households, and communities for which ENHRUM has work history data in each survey round are recorded in Table 1.

Table 1. Number of observations in each ENHRUM survey round

	Individuals	Households	Communities
2002	4,678	1,704	80
2007	6,325	1,517	80
2010	4,136	862	45

Summary statistics for individuals are reported in Table 2. The sample is restricted to working age adults (those between ages 15 and 65 each year). Reported statistics include the mean probability that an individual works in the agricultural or in the non-agricultural sectors, the mean age, percentage male, the mean household size (averaged across individuals rather than households), the percentage of individuals that are married, and the mean years of education reported for each individual in the ENHRUM survey. I also report the between and within variation of each of these variables throughout the panel. A little more than a quarter of the sample work in agriculture throughout the panel (28.9 percent). The within standard deviation of 0.259 indicates that some individuals do switch in or out of the agricultural sector. Some of this variation is caused by individuals entering the work force for the first time after the age of fifteen. A larger fraction of the worker-year observations are in the non-agricultural sector (31.5 percent). The within standard deviation is a little higher in the non-agricultural sector as well. The mean years of education completed across all individual-years is almost six years, or the completion of primary school. The within standard deviation is greater than zero since some individuals continue going to school after the age of fifteen.

Table 2. Individual Summary Statistics

		mean	sd	min	max	obs
agriculture	overall	.289	.453	0	1	97,029
	between	.	.3	0	1	10,308
	within	.	.259	-.679	1.26	9.41
non-agriculture	overall	.315	.465	0	1	97,029
	between	.	.308	0	1	10,308
	within	.	.31	-.653	1.28	9.41
age	overall	32.5	12.8	15	65	211,651
	between	.	11.7	15	65	10,854
	within	.	7.39	17.5	47.5	19.5
male	overall	.491	.5	0	1	211,651
	between	.	.5	0	1	10,854
	within	.	0	.491	.491	19.5
household size	overall	6.14	3.44	1	20	211,651
	between	.	3.09	1	20	10,854
	within	.	1.45	-2.7	14.9	19.5
married	overall	.177	.382	0	1	211,651
	between	.	.317	0	1	10,854
	within	.	.147	-.791	1.14	19.5
years of education	overall	5.96	4	0	21	206,637
	between	.	3.76	0	17.9	10,566
	within	.	1.05	-5.17	17.1	19.6

Table 3 shows summary statistics at the community level based on responses to questions about farm land, migration, and remittances in the first and second ENHRUM survey rounds. Household responses were averaged to community level responses, and then summarized. Of interest in Table 3 is that there is wide variation across communities in the farm land quality (proxied by the percentage of farm land that is level), languages spoken by the household head, U.S. migration rates, and remittances received from households from the U.S. or elsewhere in Mexico. These differences are indicated by the between standard deviation. There is some variation within communities over time, but this standard deviation is usually smaller than the between standard deviation.

We can see that most of the farm land in the sample is not level. This varies widely across communities however. In some communities all farmers reported that none of their land was level while in other communities all farmers reported that all of their land was level. Hiring workers to do farm work is more common in some communities than others, indicated by the range of average farm wages paid by households across communities. Few household heads speak English while a substantial percentage do speak an Indigenous language, and in some communities all of the household heads speak an Indigenous language.

Income sources also vary widely across communities. The mean income from government transfers is generally greater than the mean income from U.S. remittances. It is apparent that U.S. migration is more common in some communities than in others. The mean number of migrants per household ranges from zero to at least four. In communities where migration does occur, it is apparent from the maximum U.S. remittances, that remittances can be quite large. Remittances from elsewhere in Mexico are lower on average as well as at the maximum. These statistics are suggestive that the communities in rural Mexico differ from one another in many characteristics and they likely have many more unobservable differences that may affect the educational opportunities in each community and the speed at which the community transitions out of agriculture.

Table 3. Community Characteristics

		mean	sd	min	max	obs
Percent of farm land	overall	39	24.7	0	100	158
that is level	between	.	21.8	0	87.5	79
	within	.	11.7	-11	89	2
Farm wages paid	overall	673	1,272	0	9,367	205
by household	between	.	1,034	0	4,833	80
	within	.	813	-3,991	5,581	2.56
Percent household heads	overall	3.86	5.45	0	23.5	80
that speak English	between	.	5.45	0	23.5	80
	within	.	0	3.86	3.86	1
Percent household heads	overall	17.3	33	0	100	160
that speak Indigenous lang.	between	.	33.1	0	100	80
	within	.	2.13	9.33	25.3	2
US migrants	overall	.783	.796	0	4.06	80
per household	between	.	.796	0	4.06	80
	within	.	0	.783	.783	1
Mean household	overall	1,581	3,308	0	32,045	205
remittances from Mexico	between	.	2,275	0	12,561	80
	within	.	2,326	-9,053	21,065	2.56
Mean household	overall	4,637	8,725	0	81,125	205
remittances from US	between	.	6,734	0	36,538	80
	within	.	5,346	-20,826	49,224	2.56
Mean government transfer	overall	5,128	4,048	0	19,382	205
income per household	between	.	2,911	652	14,202	80
	within	.	2,875	-2,719	16,871	2.56

*Wages and incomes are in 2002 Mexican pesos.

In addition to the ENHRUM survey data, I also use annual community-level data from Mexico's Secretaría de Educación Pública (SEP), which indicates the level of the most advanced school located inside the community. SEP provided these data upon request and these are the most comprehensive data on school supply that are centrally collected. Schools are not always immediately registered with the federal government however. Consequently, some of the schools in these data may have been constructed prior to the year indicated. These data serve as a good proxy for local school supply if not completely accurate with regard to timing of school establishment. School levels include preschool (usually two years), primary school (grades 1-6), secondary school (grades 7-9), and high school (grades 10-11, 12, or 13 depending on the program). It is possible that primary schools in some of the communities, in some years, did not offer all six grades since teachers and resources may have been scarce. I have no data indicating the maximum grade taught in each school, so I assume that schools offered all grades that are standard for a primary or secondary school. Table 4 indicates the number and percentage of locations in the ENHRUM survey where the most advanced school in the village is a preschool, a primary school, a secondary school, or an upper-secondary/high school in years 1990 and 2010 [2]. Figure 1 shows how the local school supply in rural Mexico changed over time.

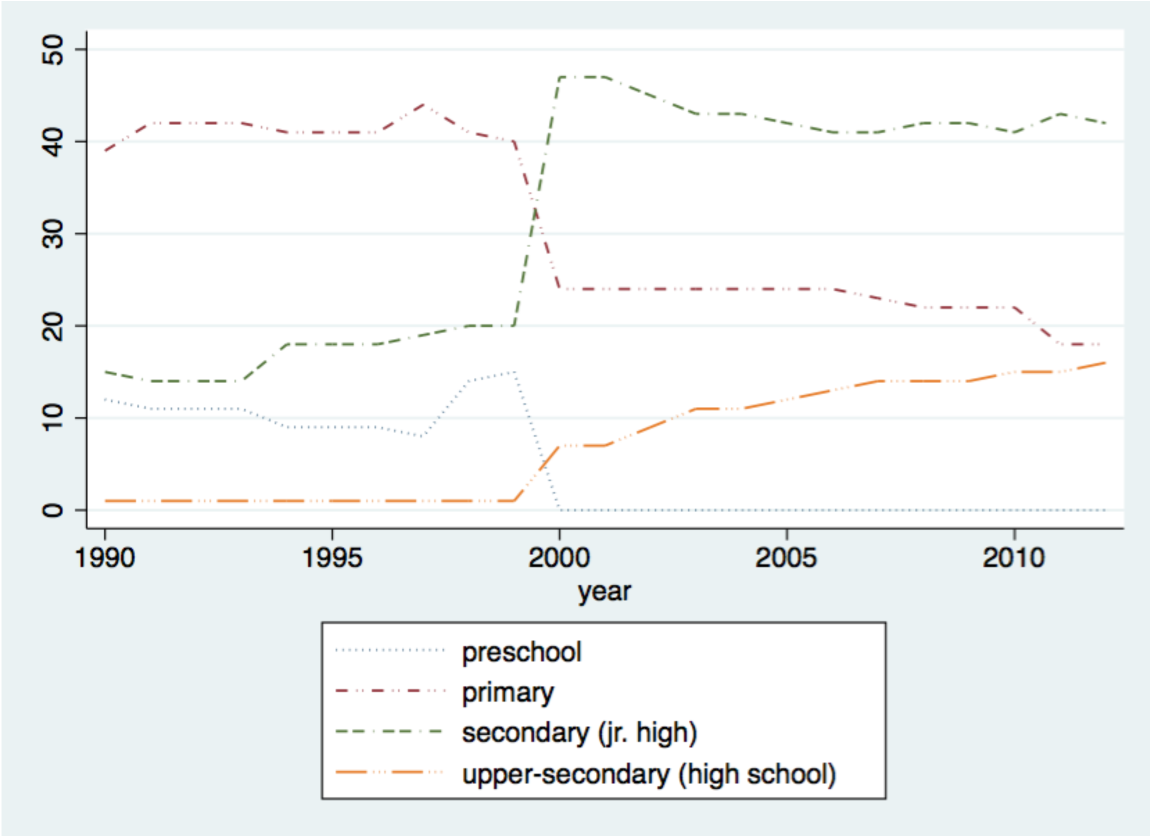
Table 4. Most advanced school in village

	1990		**2010**	
School	**Frequency**	**Percentage**	**Frequency**	**Percentage**
Preschool	12	17.91	0	0
Primary	39	58.21	22	28.21
Secondary	15	22.39	41	52.56
Upper-secondary	1	1.49	15	19.23

I combine the data from SEP with the survey data to create dummy variables indicating whether a primary school was located in an individual's community when the individual was six years old, a secondary school when the individual was twelve years old, or an upper-secondary school when the individual was fifteen. Many individuals in my sample were school-age in years prior to 1990, the first year for which SEP recorded the school supply in the communities. Therefore, if the first recorded school in the community in 1990 was a preschool or primary school, then I assume that a preschool or primary school was located in the community in all previous years. If the first recorded school in a community was a secondary school or more advanced, I report the school supply in all previous years as missing values and consequently individuals who were school age in that village before that year are dropped from the analysis. Fortunately, for this analysis there are only 15 villages with a secondary school in 1990 and only one village with an

[2] Data are missing for some villages in years prior to 1998. I report the percentage of villages for which I have data in 1990.

Figure 1. Most advanced school in each village, 1990-2010



upper-secondary school. By assuming that some villages had primary schools in years previous to the first recorded school, the results for the impact of primary schools on the decision to work in agriculture may be biased toward zero. This assumption provides more observations to see the impacts of building secondary schools (and more advanced schools) inside an individual's village.

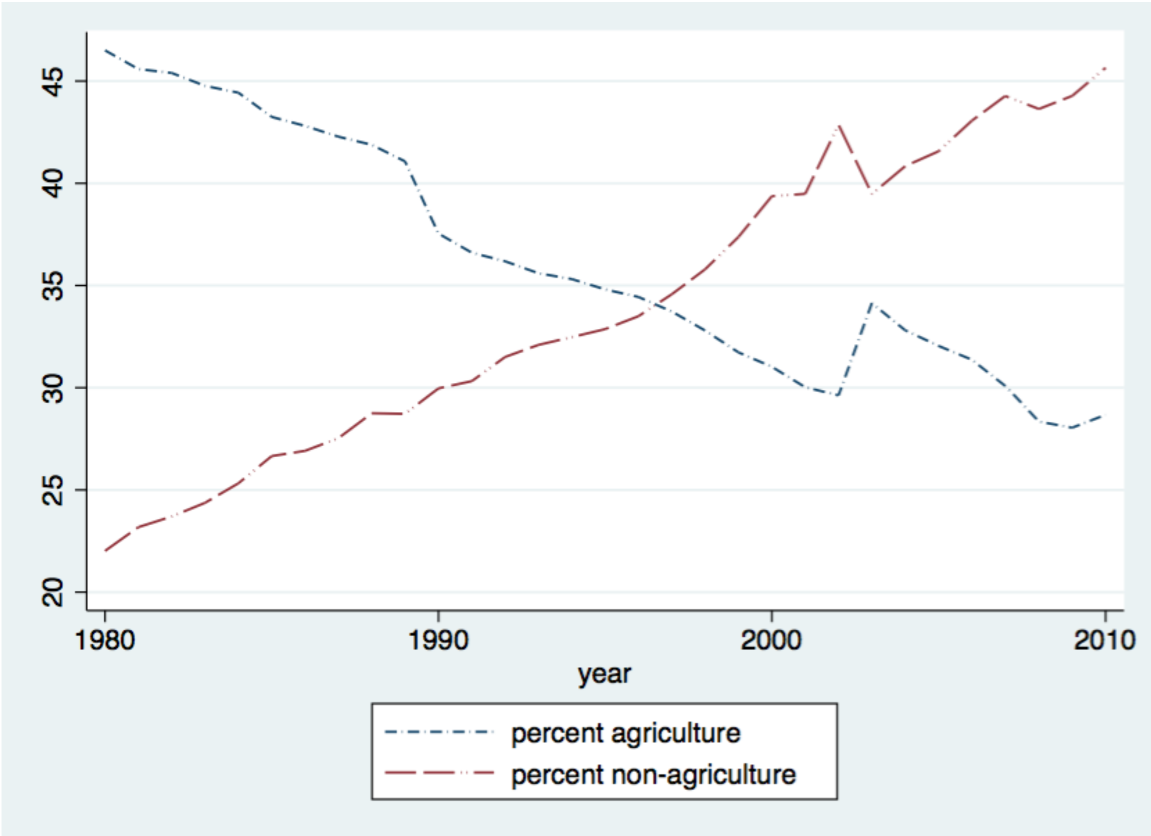
IV. Decreasing farm labor force and rising education

In previous work, I find a declining trend in the probability of working in agriculture for individuals from rural Mexico in years 1980-2010 (**?**). The probability of working in agriculture is decreasing by approximately .07 percentage points each year. This translates to a decline in the supply of agricultural labor from rural Mexico of 11,200 persons each year. [3] Figure 2 shows the percentage of the working-age population that worked in the agricultural and non-agricultural sector over time from 1980 through

[3]Calculation based on a rural working age population of 16 million persons in 2010

2010. Aside from one upward jump in 2003, the percentage of the sample working in the agricultural sector decreased steadily throughout this period while the percentage working in the non-farm sector rose at a similar rate. The share of the workforce in the farm sector surpassed that of the non-farm sector as early as 1996 and the gap between these two shares continued to expand throughout the sample period.

Figure 2. Percent of reported workforce in agriculture and non-agriculture



Simultaneous to the decline in the probability of working in agriculture, the education in rural Mexico was rapidly increasing. Table 5 shows the mean years of education completed by different age groups from rural Mexico in 2010. Adults in their twenties complete almost four years of school on average beyond that completed by adults in their fifties. Many factors likely contribute to the rise in education of younger generations in rural Mexico, and the federal government has taken an active role to increase the years of education throughout Mexico. This role includes both supply-side and demand-side efforts. On the demand side, the federal government implements the well-known antipoverty program, Oportunidades, giving cash transfers to families conditional on children's school attendance and regular health check-ups. Oportunidades (formerly called Progresa) began in 1997. It was initially offered only to households in randomly selected villages for impact evaluation. Since finding a significant impact on school enrollment, the program became nationally available for qualified households.[4]

Table 5. Years of Education by Age in 2010

Age in 2010	Mean	sd	Min	Max	Obs
20-29	8.8	3.42	0	17	2,376
30-39	7.9	3.69	0	21	2,456
40-49	6.9	4.01	0	20	1,901
50-59	5.26	3.71	0	19	1,219

In addition to providing financial incentives for families to send their children to school, national educational policies have also increased years of compulsory schooling for children. In 1992, compulsory schooling rose from the completion of primary school (grade six) to completion of lower secondary school (grade 9) (?). However, despite federal mandates, many children still do not complete primary or secondary school.

On the supply-side, the government rose its commitment to building schools in communities formerly excluded from education in the 1990s. Government spending on educational institutions in Mexico increased from 5.6% of the GDP in 1995 to 6.5% in 2005.[5] Government policy mandates that school infrastructure investments target communities of high population, poverty levels, and exclusion rates. Yet many communities still lack access to schools. In part this is because investment in school infrastructure has not yet caught up with the need. In some communities, a washed out bridge or ongoing feuds between neighboring communities may prevent children from attending school even though to federal policy makers it appears that a school is located just few kilometers away.

[4]Since Oportunidades is a welfare program to fight poverty, qualification is targeted to the poor.

[5]OECD. Centre for Effective Learning Environments (CELE): Mexico invests in a public education facilities plan. http://www.oecd.org/edu/innovation-education/centreforeffectivelearningenvironmentscele/mexicoinvestsinapubliceducationfacilitiesplan.htm. accessed November 14, 2013

Consequently, policies to expand the school supply has made education more accessible and more affordable for children in some communities but not all.

While most villages have at least a primary school located inside the community, many still do not have secondary schools. Table 6 shows where students in ENHRUM villages, sorted by level of education, attended school in 2010. It shows whether they attended school in their home village, elsewhere in Mexico, or in the United States. As expected, as students advance in their studies, a much greater share travel to other locations to attend school. Table 7 shows the expenses families reported paying for transportation to schools elsewhere in Mexico in 2010.[6] These data do not allow one to understand what these costs represented to families with school-age children however. They do not include the costs that extra-marginal families would have paid to send their children to school, how budget constrained households were, or the opportunity costs of time, both for time traveling to and from school and time spent in school. The opportunity cost of time might be especially high for older children who might otherwise be working if they were not in school.

Table 6. Where students attended school in 2010 by education level

Type of School		Local	Elsewhere in Mexico	U.S.	Total
Primary	frequency	18,135	1,550	124	19,809
	percentage	91.55	7.82	0.63	100
Secondary	frequency	6,386	3,534	124	10,044
	percentage	63.58	35.19	1.23	100
High School	frequency	1,674	3,565	155	5,394
	percentage	31.03	66.09	2.87	

How accessible a school is geographically likely plays a large role in determining how many years a student continues attending school. A child is probably more likely to

[6]To put these expenses in some sort of context, the national mean daily minimum wage in Mexico in 2010 was 56 pesos.

Table 7. Mean cost of school transportation in 2010 if school is located outside of home village 2010 pesos

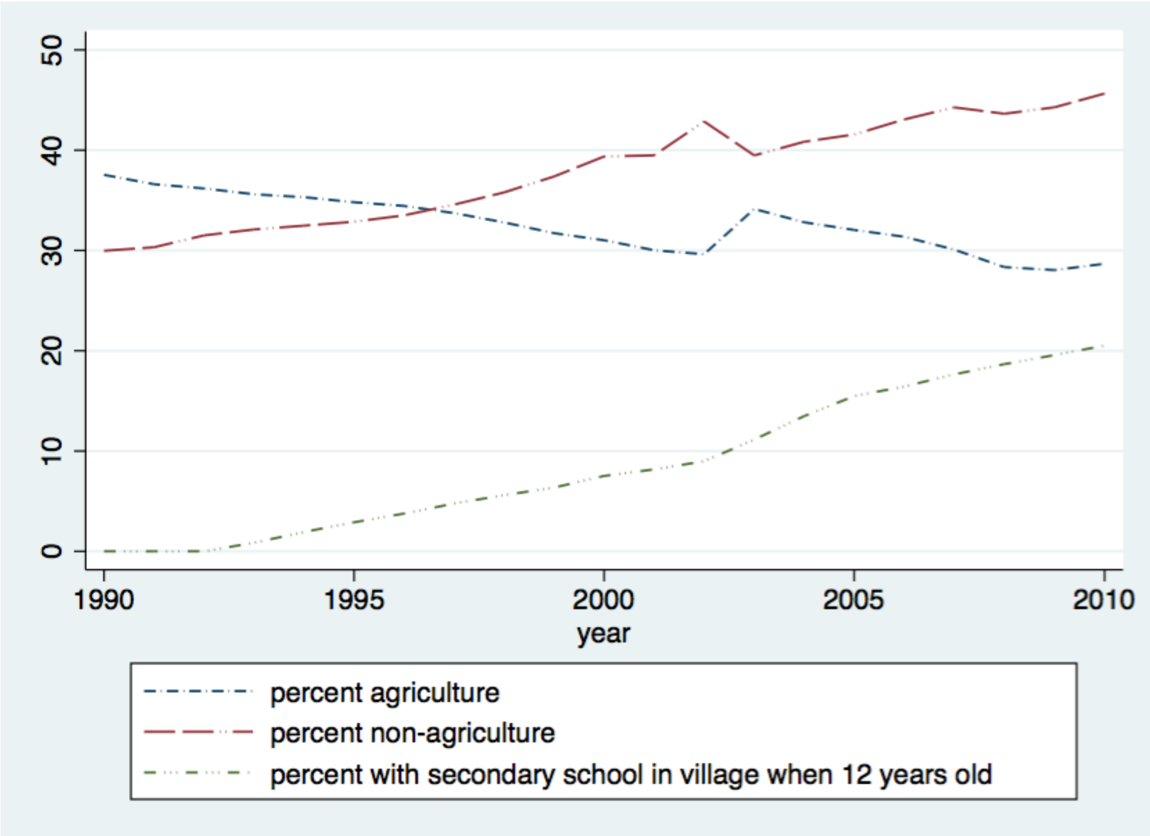
Type of School	mean cost	sd	minimum	maximum	observations
Primary	64.89	74.08	0	345.17	1,271
Secondary	92.67	106.55	0	690.34	3,410
High School	144.38	262.99	0	2,157.32	3,410

complete a secondary education if there is a secondary school located inside his village. These educational gains likely, in turn, influence labor force decisions as the children come of age to begin working. Figure 3 shows the percentage of working-age individuals who had access to a secondary school in their home village when they were twelve years old (the age when children would begin secondary school if they were on a standard track) along with the percentage who worked in the agricultural and non-agricultural sectors between 1990 and 2010. The percentage of individuals with a local secondary school rises apace with the percentage working in the non-agricultural sector and opposite the direction of those working in the agricultural sector. The negative trend in the probability of working in agriculture from rural Mexico simultaneous to rising availability of more advanced schools in rural locations leads to the question of whether there is a causal relation between access to school and labor sector decisions.

V. The Model

To examine how local school supply affects the probability of working in agriculture, I develop a simple two-period model. In the first period, individual i can work in the farm sector and invest in non-farm skills by going to school. Attending school entails both the opportunity cost of time and a cost due to inconvenience and travel. In the second period, individuals must choose whether they will work in the agricultural sector or the non-agricultural sector. There are no returns to education in the agricultural sector, but there is a wage premium increasing in the hours worked in agriculture the previous period.

Figure 3. Percent of workforce with secondary school in home village, working agriculture, and working non-agriculture



Assume that individuals maximize their expected discounted sum of utilities across the two periods by choosing how many years to invest in non-farm skills (attending school) and which sector (agriculture or non-agriculture) to work in during the second period.

Everyone is endowed with 1 unit of time each period. In the first period, i decides what fraction of his time to invest in school, denoted $s_{i,1}$, and he will spend the remainder of his time working on the farm.

Non-farm human capital depends on the fraction of time that individual i invests in skills the previous period according to the function $\theta_i(s_{i,t-1})$, which expresses non-farm human capital in terms of labor efficiency units. θ_i differs across individuals as some individuals may be more adapt to learning non-farm skills than others. Let $\theta_i'(s_{i,t-1}) > 0$ and $\theta_i''(s_{i,t-1}) < 0$. That is, non-farm skills are increasing in school investment with diminishing marginal returns.

The cost of going to school is forgone agricultural wages in period 1 along with an additional cost that measures the inconvenience of going to school, the cost of transportation to school, and even the cost of boarding at the school if it is located far away. Let $d_{i,t}$ be a function of the distance to the school and inconvenience of traveling there. The cost function is defined $C_{i,t} = c(s_{i,t}, d_{i,t})$, where

$$\frac{\partial c(s_{i,t}, d_{i,t})}{\partial s_{i,t}} > 0$$
$$\frac{\partial c(s_{i,t}, d_{i,t})}{\partial d_{i,t}} > 0$$
$$\frac{\partial^2 c(s_{i,t}, d_{i,t})}{\partial s_{i,t} \partial d_{i,t}} > 0$$

That is, the cost of attending school is increasing in the fraction of time (or number of years) and individual attends school and in the distance traveled to school. Years of attendance and distance traveled to school positively interact to increase cost.

I assume wages in both sectors are known ex-ante and there is no uncertainty about wages and the returns to education or work experience. The base wage in the agricultural sector is w^a. In period 2, individuals can earn a wage premium for sector-specific experience, $\rho^a(1 - s_{i,t-1})$, which is increasing in the time worked in the agricultural sector the previous period. One might think of this as increased speed from one harvest season to the next for workers that do piece-rate work, experience-based promotions, or learning by doing. Non-farm wages are increasing in skills and previous years of experience. In period 2, individuals can choose to work in the non-agricultural sector instead of the agricultural sector. The base wage in the non-agricultural sector is w^n and the skill premium for education is equal to $\theta(s_{i,t-1})$.

Let $\pi_{i,t}$ be the net income that individual i earns in period t, and let $u_t(\pi_{i,t})$ be i's utility in period t. Utility each period is increasing in net income. It is twice differentiable and concave. Further assume that

$$\lim_{\pi \to 0} u_t'(\pi_{i,t}) \to \infty$$

$$\lim_{\pi \to \infty} u_t'(\pi_{i,t}) \to 0$$

This implies that i will always work in the first period to have positive net income. Since there are only 2 periods in the model individuals will not invest in education in period 2.

$$\text{Let } a_{i,2} = \begin{cases} 1 & \text{if } i \text{ works in agriculture in period } t = 2 \\ 0 & \text{otherwise} \end{cases}$$

Incomes each period are expressed

$$\pi_{i,1} = (1 - s_{i,1})w^a - c(s_{i,1}, d_{i,1})$$

$$\pi_{i,2} = a_{i,2}[w^a + \rho^a(1 - s_{i,1})] + (1 - a_{i,2})[w^n + \theta(s_{i,1})]$$

Individuals choose $s_{i,1}$, and $a_{i,2}$ to maximize the sum of utilities each period. Let β be the discount rate. The optimization problem is

$$\max_{s_{i,1}, a_{i,2}} u_1[(1 - s_{i,1})w^a - c(s_{i,1}, d_{i,1})] + \beta u_2[a_{i,2}[w^a + \rho^a(1 - s_{i,1})] + (1 - a_{i,2})[w^n + \theta_i(s_{i,1})]]$$

$$s.t. \ s_{i,1} \geq 0, \ s_{i,1} \leq 1, \ a_{i,2} \geq 0, \ a_{i,2} \leq 1$$

Since wages are linear in time worked each period, individuals will choose one sector or the other in period 2, devoting all their time to that occupation. An individual is indifferent between occupations only if the wages are equal across sectors. Since the optimal choice of $a_{i,2}$ is a corner solution, I consider separately the optimization problems for $a_{i,2} = 1$ and $a_{i,2} = 0$.

Case 1: Work in agriculture in both periods ($a_{i,2} = 1$)

$$\max_{s_{i,1}} u_1[(1 - s_{i,1})w^a - c(s_{i,1}, d_{i,1})] + \beta u_2[w^a + \rho^a(1 - s_{i,1})]$$

$$s.t.\ s_{i,1} \geq 0,\ \ s_{i,1} \leq 1$$

It is easy to see that the objective function is maximized at the corner solution where $s_{i,1} = 0$. For any value of $s_{i,1} > 0$ individual i would lose income in period 1 due to the opportunity cost of time and the cost of travel and inconvenience, and i would lose income in period 2 due to a reduced experience premium.

Case 2: Work in non-agriculture in period 2 ($a_{i,2} = 0$)

$$\max_{s_{i,1}} u_1[(1 - s_{i,1})w^a - c(s_{i,1}, d_{i,1})] + \beta u_2[w^n + \theta_i(s_{i,1})]$$

$$s.t. s_{i,1} \geq 0, s_{i,1} \leq 1$$

The objective function cannot be solved at $s_{i,1} = 1$ since the marginal utility of income approaches infinity as income approaches zero in any period. I first consider the interior solution. The first order conditions are

$$u_1'()[-w^a - c'_{s_{i,1}}(s_{i,1}, d_{i,1})] + \beta u_2'()[\theta_i'(s_{i,1})] = 0$$

where $c'_{s_{i,1}}()$ denotes the partial derivative of the cost function with respect to $s_{i,1}$

The FOC can be rewritten as

$$\frac{u_1'()}{\beta u_2'()} = \frac{\theta_i'(s_{i,1})}{w^a + c'_{s_{i,1}}(s_{i,1}, d_{i,1})}$$

Although this model cannot be solved explicitly, it is possible to do some comparative statics to see how changes in parameters affect the optimal time spent in school, given that i works in the non-agricultural sector in period 2. Any parameter changes that in turn raise or lower the utility derived in the first and second cases impact how many individuals select the agricultural and non-agricultural sector in period 2. This helps inform how the parameters in the model affect the transition out of farm work.

Comparative statics

Consider what would happen if a new school were built in i's village and he no longer had to travel to another town to go to school. This would decrease $d_{i,1}$ in the model. Since $\frac{\partial^2 c(s_{i,t}, d_{i,t})}{\partial s_{i,t} \partial d_{i,t}} > 0$, then $c'_{s_{i,1}}(s_{i,1}, d_{i,1})$ would decrease as well and the marginal utility of income in period 1 would decrease since total income in period 1 has risen. Wages are set, so in order to meet the first order conditions in *Case 2*, i must invest more time in school, thus raising $s_{i,1}$ which increases $\theta_i(s_{i,1})$ and decreases $\theta'_i(s_{i,1})$. After the adjustment, the net income in both periods will rise, in the first period due to the decrease in $d_{i,1}$, and the second period as a result of the increase in $\theta_i(s_{i,1})$.

In *Case 1* where i works in agriculture in both periods, i does not invest any time in education and can obtain no benefit from education so its optimal solution would be unchanged. However, individuals who marginally preferred working in agriculture both periods will prefer *Case 2* after the school construction because the optimized utility in both periods increases in *Case 2* with the decrease in the cost of attending school.

This illustrates that building schools and reducing the inconvenience of going to school can reduce the farm labor supply in rural areas where attending school may have a high cost of travel. The critical assumption in this model is that the returns to education are higher in the non-agricultural sector than in the agricultural sector. This assumption is consistent with other empirical research (for example, see **?**). In the next section, I develop a framework for testing whether the construction of schools in rural Mexico reduced the probability of working in agriculture over the years 1980 through 2010.

VI. Identification

Many studies have attempted to measure the impacts of individual education on labor outcomes. Identification in these models is complicated by several empirical challenges. To identify the impact of schooling on the farm labor supply, I begin with a simple reduced form regression and I address the confounding factors that would likely bias the results in this model one at a time.

The simplest model of the relation between education and the farm labor supply would be to regress a variable for working on agriculture in year t on the years of education that the individual has completed. Given the nature of panel data, I control for autocorrelation and a time trend, even in the reduced form model. The actions of individual i in any given year are likely correlated with his actions in previous years. I control for this autocorrelation by including lagged dependent variables and a time trend.

$$a_{i,t} = \gamma_0 + \sum_{j=1}^{\rho} \gamma_j a_{i,t-j} + \beta edu_{i,t} + \mu t + \epsilon_{i,t} \tag{1}$$

where $a_{i,t}$ is a dummy variable equal to one if individual i works in agriculture in year t, $edu_{i,t}$ are the number of years of education that i has completed, t is an annual time trend, and $\epsilon_{i,t}$ includes all unobservable variables correlated with i's decision to do farm work in year t. The coefficient on years of education in (**??**) likely suffers from omitted variable bias however. For example, if individuals with innate ability for non-farm work also have greater innate ability for school, then they are likely to complete more years of school and their abilities will also help them acquire work in the non-farm sector, thus biasing β downwards (assuming $\beta < 0$). Likewise if an individual is determined to find a job in the non-farm sector and he is convinced that more years of education will help him get a job or improve his wages, then he is more likely to pursue additional years of education and β will be biased downwards due to reverse causality. The problem lies in the fact that both education and labor sector are self-selected variables and many unobserved factors influence both decisions regarding education and decisions regarding the labor sector. To identify the effects of education on labor sector choice, I must use a proxy for education that is exogenous to the decision for working in agriculture.

For this reason I look at the impacts of local school supply on sector selection rather than the number of years of educational attainment. Building schools in a community is not an individual decision. Individuals have little to no control over the level of schooling offered in their village nor when schools will be built. This reduces the potential for biased results due to self-selection. Equation **??** replaces the variable $edu_{i,t}$ with a vector of dummy variables indicating whether there was a primary school in individual i's home village when he was 6 years old (the typical age when children in Mexico begin primary school), whether there was a secondary school in i's village when he was 12 years old, and whether there was a upper-secondary school when he was 15 years old.

$$a_{i,t} = \gamma_0 + \sum_{j=1}^{\rho} \gamma_j a_{i,t-j} + \beta s_{i,t} + \mu t + \epsilon_{i,t} \tag{2}$$

Although individuals and their families cannot choose when or where schools are built, the coefficient β in (**??**) is still susceptible to omitted variable bias. Communities may vary intrinsically in ways that correlate school acquisition with a reduction in farm work. The village characteristics summarized in Table 2 show how diverse the communities in the sample are. Many of these variables (and other unobservable characteristics) are likely correlated with school supply and the propensity to work in agriculture. It would be impossible to control for all village characteristics that might potentially bias the β. Instead, I control for unobservable village characteristics by using village fixed effects. Consequently, identification in this model is based only on variation in school supply within each village (and not between villages).

$$a_{i,t} = \gamma_0 + \sum_{j=1}^{\rho} \gamma_j a_{i,t-j} + \beta s_i + \mu t + \lambda_v + \epsilon_{i,t} \tag{3}$$

Results from (**??**) might still be biased if state or regional policies promote non-agricultural growth and industrialization simultaneous to developing school infrastructure. In this instance, the trend in the local educational opportunities will be correlated with localized trends in the non-farm opportunities. Consequently, it might appear that advances in the availability of education enhance the probability that individuals do not work in agriculture when, in reality, access to new economic opportunities is driving the trend out of the farm sector. Since confounding unobservable village characteristics might change over time, village fixed effects will not control for these endogenous characteristics. I address this potential problem by controlling for state-by-year fixed effects.

$$a_{i,t} = \gamma_0 + \sum_{j=1}^{\rho} \gamma_j a_{i,t-j} + \beta s_i + \mu t + \lambda_v + \sigma_{s,t} + \epsilon_{i,t} \tag{4}$$

One might further be concerned that families move to villages based on the school supply or proximity to non-farm job opportunities. Such self-selection is of low concern in this dataset however. Families are unlikely to move to ENHRUM villages in search of better educational opportunities since the school levels in these villages are relatively low. Many villages in the sample have school supply capped at lower secondary (or junior high equivalent), primary, or even preschool levels. Furthermore, the work histories of individuals that migrate out of ENHRUM villages are recorded with their home village as long as at least one person from their household remains in the village of origin during the subsequent survey rounds. There is little attrition due to migration of entire households over time.

VII. Results

Table 8 reports the results from the empirical model described in Section V. Standard errors are clustered at the individual level. The first column is the reduced form regression of the probability of working in agriculture regressed on its lags, a time trend, and individual completed years of education. The coefficient on years of education is significantly less than zero and quite large, but as described in the Identification section, this coefficient is likely downward biased due to correlation with omitted variables. In column II, I use a proxy for education, the school supply in the home village when the individual was school-age. I include dummy variables for having a primary school, secondary school, and upper-secondary school inside the village when the individual was 6, 12, and 15 years old respectively. The coefficient on primary school is positive and statistically significant at the 10 percentage level. However, this specification uses variation across villages and the coefficients likely reflect many unobserved characteristics that differ across villages.

Column III includes village fixed effects, thus isolating within village variation. The coefficient on secondary school is significantly less than zero and indicates that individuals are approximately 1.7 percentage points less likely to work in agriculture if they had access to a secondary school inside their village when they were 12 years old. The results are robust to inclusion of village fixed effects and state-by-year fixed effects. The reduced probability of working in agriculture from having a secondary school is in addition to a negative trend in the probability of working in agriculture of .07 percentage points each year on average. Given a rural working-age population of 16 million people in 2010, a negative trend of .07 percentage points means that the farm labor force is decreasing by 11,200 persons each year. This means that if there were a secondary school in every rural village in Mexico, the expected farm labor supply would decrease by an additional 272,000 persons. In 2010, only 72% of the villages in the ENHRUM sample had a secondary school. When scaled by the rural working-age population of Mexico each year, the impacts of the trend and of building more advanced schools in rural villages on the farm labor supply is quite large.

Table 8. Percentage Impacts on the Probability of Working in the Farm Sector

VARIABLES	(1) Completed Years of Education No FE	(2) School in Village No FE	(3) School in Village Village FE	(4) School in Village State-year FE & Village FE
years of education	-0.212			
	(0.013)***			
School in Village:				
primary school		0.329	-2.459	-2.540
		(0.179)*	(2.701)	(2.728)
secondary school		-0.558	-1.512	-1.691
		(0.590)	(0.765)**	(0.773)**
upper-secondary		0.941	0.751	0.716
		(1.171)	(1.436)	(1.451)
t (year)	-0.038	-0.067	-0.077	
	(0.008)***	(0.010)***	(0.010)***	
L.agriculture	0.802	0.784	0.779	0.784
	(0.009)***	(0.012)***	(0.012)***	(0.012)***
L2.agriculturet	0.107	0.107	0.102	0.104
	(0.009)***	(0.012)***	(0.012)***	(0.012)***
Observations	127,237	85,324	85,324	85,324
Number of id	7,953	5,650	5,650	5,650
R-squared within	0.502	0.492	0.492	0.497
R-squared between	0.960	0.947	0.947	0.948
R-squared overall	0.833	0.827	0.827	0.829

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

VIII. Conclusion

The farm labor supply from rural Mexico is on a downward trend and improvements in education are accelerating the transition out of agriculture. As access to secondary schools in rural Mexico improves, fewer children grow up to work in the farm sector. Educational opportunities in rural Mexico are growing and will continue to improve. As attending school becomes less costly to rural residents, the probability that the youth choose to work in agriculture decreases. Worldwide countries transition out of agriculture as their economies develop. These findings show that access to education is an important factor in that transition. Expanding access to education is a fundamental piece in economic development. Improvements in education are unlikely to reverse, which further suggests that the impacts of building schools on the trend out of agriculture will be long-lasting.

The transition out of agriculture in rural Mexico is of particular importance because both the U.S. and Mexico have historically depended on an elastic supply of farm labor from rural Mexico. These findings show that the farm labor supply from rural Mexico is declining, the expansion of more advanced schools in rural Mexico is speeding the transition out of agriculture, and the schools are not likely to disappear. The U.S. and Mexican agricultural industries will have to adapt by switching to less labor-intensive crops, seeking out new sources of labor (presumably from a country and region where education is not on the rise), or by using more mechanized farm practices that require fewer workers. The most viable solution in the long-term appears to be increased mechanization, which will cause the average productivity of workers to rise, increasing wages. Technological advances in the farm industry will increase the returns to education within farm jobs, thereby reducing the demand for farm labor overall and also helping the farm sector to retain more workers with higher levels of education.