Education Accelerating the Agricultural Transition: Panel Data Analysis of Rural Mexico *Draft*

Diane Elise Charlton
Department of Agricultural & Resource Economics
University of California, Davis

April 5, 2015

Abstract

Economic theory shows that education is critical to economic development and to labor sector choice, yet there is little research to indicate whether access to education plays a role in the agricultural transformation, the stage of development when an economy shifts from primarily agriculture to non-agriculture. This paper identifies the impact of secondary school access on the probability of working in agriculture using 31 years of village-level panel data nationally representative of rural Mexico and exogenous shifts in children's access to secondary school to identify the marginal impact of a year of education on the probability of working in agriculture and the probability of migrating. The workforce from rural Mexico is currently shifting out of the agricultural sector, shortly after the federal government began investments in rural access to education. The findings show that investing in education does significantly reduce the probability of working in agriculture. However, I do not find significant impacts of education on labor mobility. These findings identify an important route by which education promotes development, indicate that advances in education can be expected to precipitate changes in the rural labor supply, and demonstrate how education can promote improved labor opportunities.

Education Accelerating the Agricultural Transition: Panel Data Analysis of Rural Mexico

Expanding job opportunities outside of the agricultural sector is critical for improving earnings and other welfare outcomes, and in many rural developing economies restricted access to education is a crucial limiting factor to obtaining non-farm work. Understanding the impacts of education on labor sector decisions can direct policy to help alleviate poverty in rural areas. Economic theory predicts that education is an essential element to economic growth (Nelson and Phelps, 1966; Mincer, 1984; Barro, 1992; Becker, Murphy and Tamura, 1994; Benhabib and Spiegel, 1994), yet there is little research to identify its role in the agricultural transformation, the stage of development when an economy shifts from primarily agriculture to non-agriculture. Understanding the role of education in the agricultural transformation can help direct policy to improve incomes of rural poor through improving school access and to prepare an economy for a smooth transition from primarily agriculture to non-agriculture as education advances. In this paper I identify the impact of an additional year of education on the probability of working in agriculture and the probability of migrating for individuals from rural Mexico between 1980 and 2010.

Rural Mexico provides a timely setting for analysis because the rural labor force is currently transitioning out of the farm sector while school provision is expanding. Taylor, Charlton and Yúnez-Naude (2012) show that the farm labor supply from rural Mexico declined between years 2002 and 2010. This may, in part, reflect recent advances in rural education. Public spending on education increased by 36 percentage points between 1995 and 2001 (Santibañes, Vernez and Razquin, 2005). I use unique household survey data nationally representative of rural Mexico that record where every household member and every child of the household head works between 1980 and 2010 along with their years of education. I create an instrument that proxies for village-level access to secondary education when individuals are 12 years old, the age when children begin secondary school, to identify the impacts of education on the probability of working in agriculture or migrating to work away from home as an adult. I look at work outcomes at age

20 and I repeat the analysis for ages 25 and 30 to see whether the effects of education grow or diminish with age. Studies show that access to non-farm work is associated with higher incomes and less income variability (Huffman, 1980; Janvry and Sadoulet, 2001; Zhang, Huang and Rozelle, 2002), and this paper shows that having a secondary education promotes access to non-farm work.

Several studies find a positive correlation between education and employment in off-farm work (Zhang, Huang and Rozelle, 2002; Huffman, 1980; Janvry and Sadoulet, 2001), but these studies do not account for the potential endogeneity of education in the labor choice model. Duflo (2000) and Foster and Rosenzweig (1996) use school construction as an instrument to identify the impacts of education on income and find significant, positive effects, but they do not distinguish between farm and non-farm labor. Joliffe (2004) measures the marginal returns to education and makes a clear distinction between the farm and non-farm sectors. However, Joliffe examines self-selected education only and does not investigate how changes in the supply of education affect labor allocation.

I contribute to two families of literature, regarding the outcomes of education on labor sector selection and migration and regarding the transition of rural developing economies out of agriculture. I use a unique variable to instrument for years of education that exploits village-cohort level changes in secondary school enrollment rates as a proxy for changes in local access to secondary school. Village level changes in school access provide good instruments for education because rural communities have little influence over when and where schools are built, so improvements in access to education within a village predict changes in the expected years of education exogenous to an individual's unobserved ability or demand for school.

I find that an additional year of education around secondary school age reduces the probability of working in agriculture at age 20 by 4.7 percentage points. This impact increases with age to over 10 percentage points by age 30. Regressing migration directly on own education shows a significant positive correlation, but the coefficient on education is not significant when I use two-stage least squares. These findings provide evidence that improving the supply of education in rural areas can improve individual and household welfare by expanding labor opportunities across sectors and that education is a critical factor in economic development through the agricultural transformation.

The paper is organized as follows. In Section I, I describe rural Mexico, including changes in the workforce and access to education. In Section II, I describe a theoretical model of the impacts of education on labor allocation between the farm and non-farm sectors. Section III describes the data and Section IV the empirical model. Section V presents the results. Section VI conducts several robustness checks. Section VII discusses the implications of the findings and Section VIII concludes.

I. The Workforce and Access to Education in Rural Mexico

Rural Mexico has entered a stage of development when the workforce is transitioning out of agriculture and non-farm production is growing. The farm work force from rural Mexico fell by 2 million, or 25 percent, between 1995 and 2010 (Charlton and Taylor, 2013). A decade or more prior to this, the federal government began efforts to improve rural access to schools. Mexico's constitution requires that basic education (currently grades 1-9) must be publicly available, free of charge, and non-religious. However, access and quality of education vary across communities and across time, and many students do not have access to basic education.

Mexico made considerable investments in rural education, particularly in the 1980s and 1990s. In 1992, the federal government increased mandatory education from the completion of primary school (grade 6) to the completion of lower-secondary school (grade 9)[1] (Rolwing, 2006). Although federally required education changed in 1992, the mandate was not effectively enforced, particularly in rural areas, so the mandate does not generate an exogenous change in expected eduction. Rather gains in education were more gradual and differed across locations. Public spending on education rose from 2.9 percent of the GNP in 1980 to 5.1 percent in 2010.[2]

Despite gains in public spending for education, improved education is not homogenous across Mexico and national school spending still lags behind most OECD countries.

[1] I refer to lower-secondary schools as "secondary" schools in the remainder of the paper
[2] http://databank.worldbank.org/data/home.aspx

Arguably, in many parts of Mexico, particularly rural areas, much of the school funding does not benefit students. A 2005 report on education spending in Mexico found that about 90 percent of the federal budget went towards teacher salaries, and in some states, as much as 98 percent. Teacher unions are strong in Mexico and salaries remain high even where teacher absenteeism is common and quality of teaching is low. In the states of Guerrero and Oaxaca, two of the poorest and most rural states in Mexico, teachers were in the classroom only about 50 percent of school days. On days when teachers were present, school hours were usually reduced by 2 to 3 hours (Santibañes, Vernez and Razquin, 2005). This suggests that limited access to education extends beyond constraints in school infrastructure and public mandates requiring students to attend school. Additionally, rural areas are likely to benefit from gains in public education more slowly than urban areas since they have less political influence.

School funding is highly centralized, so communities have little power to initiate a school-building project. The central government agency Secretaría de Educación Pública (SEP) is the largest source of school funding. In 1992, the education system was decentralized to the 32 states, but many reports contend that the decentralization was mostly administrative. For example, all primary schools must use national curriculum and nationally produced books and secondary school curriculum must receive approval from SEP. Furthermore, principals and parents do not have the authority to hire, fire, or place teachers, so there is little teacher accountability to students and parents. Since the decentralization, states gained greater authority in school placement, but state governments still rely heavily on SEP for funding, thus further limiting power at the local level to influence when and where schools are built and teachers are provided. Currently, the national government provides about 85 percent of educational funding (Santibañes, Vernez and Razquin, 2005). In 1997, SEP mandated that federal financial resources be distributed to states based on the number of schools and teachers that were decentralized in 1992. However, in 1992, many state schools operated side by side with federal schools. Consequently, states that gathered local funding for education may receive less federal support per pupil even though the demand for schools is high. In many locations, this policy effectively punished communities for gathering their own resources to meet educational demand.

The federal government prioritizes building schools in communities located farthest from existing schools and those with the highest poverty rates, yet physical infrastructure is not the only constraint to accessing education. Conversations in the field reveal that some children are denied access to the local school because of their ethnicity or religion. Physical obstructions, such as a washed out bridge, may prevent children from attending school in a nearby town. One of the major constraints for remote villages is finding teachers who are willing to live and work in the location. Limited supply of teachers and school infrastructure has been resolved in part by multi-shifting schools (providing morning, afternoon, and evening sessions) so that more students can attend school without building additional infrastructure. A system of telesecundarias, or distance learning, was implemented in the 1990s. In telesecundarias, one teacher is hired to teach all of the subjects and students watch their lessons on satellite television. Telesecundarias are most prevalent in poorer, highly rural states and student test scores tend to be lower in these schools, though other factors may be responsible for this performance gap.

The opportunity cost of a student's time studying may be another significant constraint to education for poorer households, but the Mexican government does implement programs to subsidize school attendance. For example, Oportunidades, the well-known anti-poverty program, gives 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. Since Oportunidades is a welfare program to fight poverty, qualification is targeted to the poor. However, Bobonis and Finan (2009) and Lalive and Cattaneo (2009) find that Oportunidades recipients in Mexico positively affect the school attendance of children in their communities who are ineligible for conditional cash transfers through peer effects. 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 (Skoufias, Davis and De La Vega, 2001; Schultz, 2004). Since Progresa was rolled out randomly across villages, the program's potential impacts on school attendance should not confound the results in this paper.

This paper utilizes changes in school access within villages, which I interpret from sustained changes in school enrollment rates, as a source of exogenous variation in years of education. The year that a community gains access to a school is arguably exogenous to other community trends that may impact the decision to work in agriculture. Since communities cannot control or predict the year that a school is built or access improved, access to secondary education provides a good instrument for education. Several studies indicate that improved access to education has positive impacts on years of school attendance (Duflo, 2000; Foster and Rosenzweig, 1996; Kane and Rouse, 1995; Card, 1993). Lavy (1996) observes that access to secondary education may affect primary schooling decisions as well, and Handa (2002) shows that effects of improved education persist across generations since more educated parents are more likely to send their children to school for more years.

The existing literature suggests that the effects of education can be extensive, reaching across peer groups and from one generation to the next. Accessibility of school is an important factor in determining years of education, and education is shown to have large impacts on raising incomes. I focus on the impacts of education on labor sector decisions and migration, which are important determinants of income and other welfare outcomes.

II. The Model

To illustrate how education affects the allocation of household labor between the farm and non-farm sectors, consider a utility maximizing household model. I use a household model since labor and migration decisions are often made at the household level to maximize household utility, but the model can easily be adapted to an individual decision-maker. I use a static model without risk or uncertainty. First, I assume that education is exogenous, and then I relax this assumption and use local access to education as an exogenous shift parameter to see how school access can affect labor outcomes.

II - 1 Take education as exogenously given

Assume that households maximize utility by selecting where to allocate labor, and suppose that households maximize utility by maximizing net income. Households solve

$$\max\{y_f(E, L_f, X_f, p_f) + y_n(E, L_n, X_n, p_n)\}$$

$$s.t. \quad \bar{L} \geq L_f + L_n + E$$

$$L_f \geq 0$$

$$L_n \geq 0$$

$$E \geq 0$$

where y_f is the household's farm income net input costs and y_n is the household's non-farm income net the costs of inputs. E is the years of education the household attains, L_f is the amount of labor allocated to farm work and L_n the amount of labor allocated to non-farm work, X_s is the productive inputs for sector s (farm or non-farm), and p_s is the vector of prices for sector s. Let sector inputs X_s and prices p_s be given. Each household is endowed with time $\bar{L}$ which can be allocated between farm work, non-farm work, and going to school. For the time being, I will consider education as exogenously given to see how changes in education affect the labor supply decisions.

The production functions are quasi-concave, and education and labor are complements in production. That is

$$\frac{\partial y_s}{\partial L_s} > 0$$
$$\frac{\partial^2 y_s}{\partial L_s^2} < 0$$
$$\frac{\partial^2 y_s}{\partial L_s \partial E} > 0$$

Since income is monotonically increasing in labor supplied, the time constraint is binding. Assume that non-negativity constraints are not binding so that I can find the impacts of exogenous changes in education on labor allocation across sectors. The objective function can be rewritten as

$$\max_{L_f}\{y_f(E, L_f, X_f, p_f) + y_n(E, \bar{L} - L_f - E, X_n, p_n)\}$$

Households maximize the objective function with respect to the amount of labor allocated to farm work, L_f. Let Q be the objective function. The First Order Conditions are

$$\frac{\partial Q}{\partial L_f} = \frac{\partial y_f}{\partial L_f} - \frac{\partial y_n}{\partial L_n} = 0$$

The First Order Conditions can be rewritten

$$\frac{\partial y_f}{\partial L_f} = \frac{\partial y_n}{\partial L_n}$$

That is, in equilibrium, the marginal benefit of farm labor is equal to its marginal cost in terms of lost labor in the non-farm sector.

Using the Implicit Function Theorem, I find how exogenous changes in education affect the household's allocation of labor to the farm sector.

$$\frac{\partial L_f}{\partial E} = -\frac{\frac{\partial^2 y_f}{\partial L_f \partial E} - \frac{\partial^2 y_n}{\partial L_n \partial E}}{\frac{\partial^2 y_f}{\partial L_f^2} + \frac{\partial^2 y_n}{\partial L_n^2}}$$

The sign of the expression is not certain a priori. The denominator is less than zero by diminishing marginal returns to labor in production. The sign of the numerator depends on the relative complementarity of education to labor in the farm and non-farm sectors. Households will allocate less labor to farm work as education increases if

$$\frac{\partial^2 y_n}{\partial L_n \partial E} > \frac{\partial^2 y_f}{\partial L_f \partial E}$$

That is, as education rises, households will reduce the allocation of labor to the farm sector and increase labor to the non-farm sector if the marginal gains of education to labor productivity in the non-farm sector is greater than the marginal gains of education to labor productivity in the farm sector. I expect that labor and education have greater complementarity in the non-farm sector than in the farm sector (Joliffe, 2004). Then gains in education are expected to reduce allocation to the non-farm sector.

II - 2 Endogenous Education

The analysis above assumes that education is given. However, this assumption is not realistic since education is endogenously determined from unobservable characteristics and factors. I now consider how construction of schools, located closer to the household might cause exogenous changes in education. I introduce a new term to the household's objective function. Let czE be the cost of traveling to and from school, where c is the per kilometer cost of travel, and z is the distance traveled. Assume that everyone attends some school so that the non-negativity constraint on years of education is non-binding. Households solve

$$\max_{L_f,L_n,E}\{y_f(E,L_f,X_f,p_f)+y_n(E,\bar{L}-L_f-E,X_n,p_n)-czE\}$$

The First Order Conditions are

$$\frac{\partial Q}{\partial L_f}=\frac{\partial y_f}{\partial L_f}-\frac{\partial y_n}{\partial L_n}=0$$
$$\frac{\partial Q}{\partial E}=\frac{\partial y_f}{\partial E}+\frac{\partial y_n}{\partial E}-cz=0$$

which can be rewritten as

$$\frac{\partial y_f}{\partial L_f}=\frac{\partial y_n}{\partial L_n}$$
$$\frac{\partial y_f}{\partial E}+\frac{\partial y_n}{\partial E}=cz$$

Written in this way, it is clear that at the equilibrium, the marginal productivity of farm work is equal to the marginal productivity of non-farm work, and the marginal benefit to education in terms of increased labor productivity, is equal to the marginal cost of attending school.

Now I consider what happens to education, E, and the amount of labor supplied to the farm sector, L_f, when a new school is constructed such that z decreases. I create the Jacobian matrix from the First Order Conditions and apply the Implicit Function Theorem. The Jacobian is

$$\begin{pmatrix} \frac{\partial^2 y_f}{\partial L_f^2}+\frac{\partial^2 y_n}{\partial L_n^2} & \frac{\partial^2 y_f}{\partial L_f\partial E}-\frac{\partial^2 y_n}{\partial L_n\partial E} \\ \frac{\partial^2 y_f}{\partial L_f\partial E}-\frac{\partial^2 y_n}{\partial L_n\partial E} & \frac{\partial^2 y_f}{\partial E^2}+\frac{\partial^2 y_n}{\partial E^2} \end{pmatrix}$$

The determinant of the Jacobian does not reduce into small terms, so I refer to it as $|J|$. $|J| > 0$ by the quasi-concavity of the production function.

It follows that

$$\frac{dE}{dz} = \frac{c(\frac{\partial^2 y_f}{\partial L_f^2} + \frac{\partial^2 y_n}{\partial L_n^2})}{|J|}$$

and $\frac{dE}{dz} < 0$. This means that as distance to school decreases, the household invests more in education.

Next, I see how labor allocation is expected to change with respect to distance to school.

$$\frac{dL_f}{dz} = -\frac{c(\frac{\partial^2 y_f}{\partial L_f \partial E} - \frac{\partial^2 y_n}{\partial L_n \partial E})}{|J|}$$

If $\frac{\partial^2 y_n}{\partial L_n \partial E} > \frac{\partial^2 y_f}{\partial L_f \partial E}$ then $\frac{dL_f}{dz} > 0$,which is the same condition I found for farm labor to decrease in response to exogenous increases in education. If the complementarity of education to non-farm labor is greater than the complementarity of education and farm labor, then household labor allocated to farm work is expected to decline when new schools are constructed close to the home and z decreases.

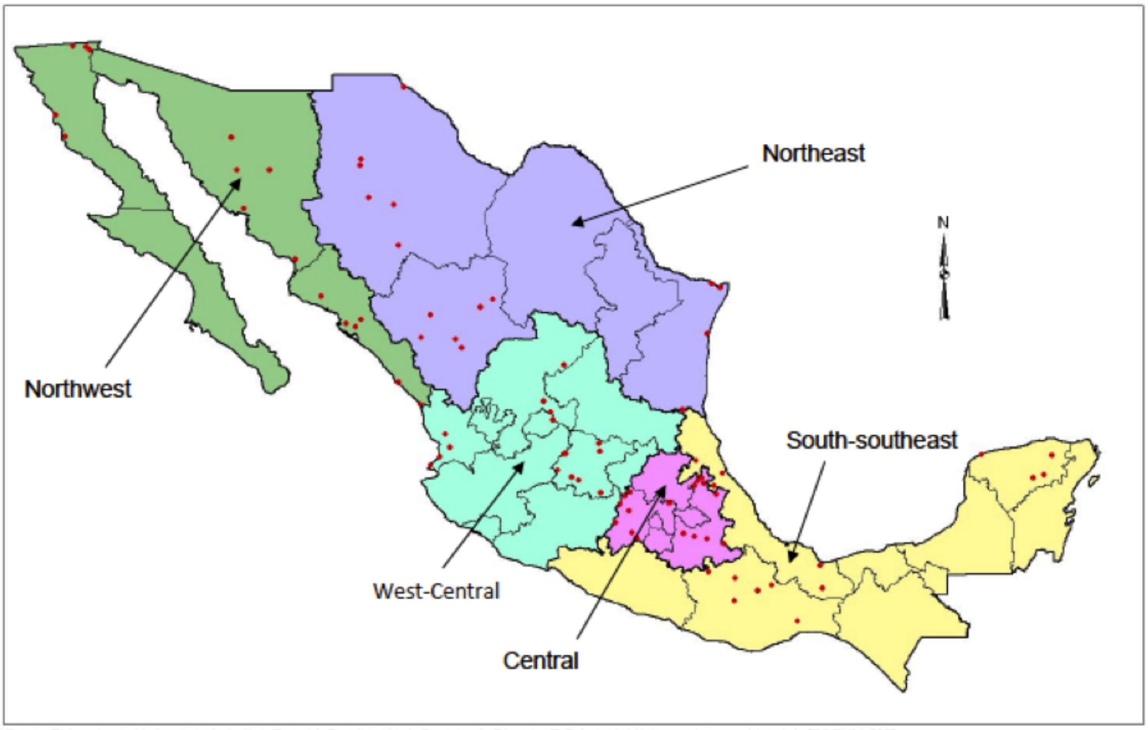
This model illustrates that building schools close to children's homes can reduce the farm labor supply in rural areas where traveling to school has traditionally been more costly. The critical assumption is that the returns to education are higher in the non-agricultural sector than in the agricultural sector.

III. Data

I use data from a nationally representative sample of rural Mexican households. The Mexico National Rural Household Survey (Spanish acronym *ENHRUM*[3]) is unique in providing retrospective panel data on individual migration from rural Mexico to both the United States and destinations within Mexico in 1980-2010.

[3] *Encuesta Nacional a Hogares Rurales de México*; Spanish acronym ENHRUM

The map in Figure 1 shows Mexico divided into five representative regions and the locations of the original ENHRUM surveys. [4]



Fuente: Elaborado en el Laboratorio de Análisis Espacial, Coordinación de Servicios de Cómputo, El Colegio de México con base en datos de la ENHRUM, 2003.

Figure 1: Map of ENHRUM Villages

The panel data come from three survey rounds: 2002, 2007, and 2010. Each round collects detailed information on migration destinations, whether migrants worked in the agricultural or non-agricultural sector, and employment status (wage-earner or self-employed) for 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. Work histories were gathered as far back as 1980 for a randomly selected group of household members and back to 1990 for all household members. Since those who do not have a work history from 1980-1990 are a random sample, the exclusion of these individuals in the earliest decade of the analysis should have no bearing on the results. Some households

[4]The surveys in the Northeast region were dropped from the 2010 survey, so I do not have data for households in this region for years 2008-2010. Some of the original localities shown in the map were dropped in the final survey round due to budget constraints or violence. The remaining sample was randomly selected to retain the integrity of national representation.

were dropped from the survey in 2010 due to budget constraints and increased violence in their communities. The method of dropping communities from the survey in 2010 maintains a nationally representative sample of rural Mexico apart from the communities dropped due to violence. The number of individuals age 20, households, and communities by survey work history period are recorded in Table 1.

Table 1. Number of Observations in each ENHRUM Survey Round

Years	Individuals (age 20)	Households	Communities
1980-2002	3,677	1,539	80
2003-2007	1,061	661	80
2008-2010	376	302	45

The first dependent variable of interest is a dummy variable equal to 1 if the individual works in agriculture at age 20. Each year, the survey records the primary sector that every household member works in for each of three locations: in the home community, migrated to another location within Mexico, and migrated to the United States. If the individual works primarily in the agricultural sector in any one of these three locations when he or she is 20 years old, then the dependent variable will be one.

Additionally, I look at the impact of education on the probability of doing self-employed agricultural work, agricultural wage work, and the probability of migrating. An individual migrates seasonally if he records working outside of his home village, either in Mexico or in the United States, and he also works in his home village when 20 years old. I define full-year migration equal to 1 if an individual reports only working outside of his home village when 20 years old. An individual works in local agriculture if he works in his home village and his primary occupation there is in the agricultural sector. Mexican agriculture refers to individuals who work in a different location in Mexico and work primarily in agriculture in that location. The same definitions apply for the non-agricultural sector in each location and they apply for each sector in the United

States.

Since I observe village-level panel data, each variable varies both within and between villages. I use within village variation to identify the model, so in Table 2, I collapse the data to the village level and take the overall, within, and between standard deviations. The within variance measures

$s_w^2 = \frac{1}{NT-1}\sum_v \sum_t (x_{iv} - \bar{x_v})^2 = \frac{1}{NT-1}\sum_v \sum_t (x_{iv} - \bar{x_v} + \bar{x})^2$.

The between variance measures $s_b^2 = \frac{1}{N-1}\sum_v (\bar{x_v} - \bar{x})^2$.

The overall variance measures $s_o^2 = \frac{1}{NT-1}\sum_v \sum_t (x_{iv} - \bar{x})^2$.

The minimum and maximum columns in Table 2 measure the minimums and maximums of x_{iv} for overall variation, $\bar{x_v}$ for between, and $(x_{iv} - \bar{x_v} + \bar{x})$ for within.

The summary statistics in Table 2 show that the mean share of 20 year-olds in a rural Mexican village who work in agriculture is 29.1 percent. Most of them are agricultural wage or salary workers (as opposed to being self-employed). The mean share that work in the non-agricultural sector is 35.9 percent. The remainder do not report working. The overall standard deviation in the share who work in agriculture is 0.361. The standard deviation between villages is 0.176, and the standard deviation within villages is 0.317. A small share of the population migrates outside of their home state for only part of the year (2.5 percent on average). A much larger share works outside of their home state for a full year (18.7 percent on average). Among those who work in their home village, most work in agriculture, and among those who migrate away from home, the majority work in the non-farm sector.

In addition to work histories, I also observe several individual and household characteristics, including years of education, gender, the number of children (age 14 and under) and the number of working-age adults (ages 15 to 65) living in the individual's household when 12 years old, whether the head of the household speaks an indigenous language, and how much land the household inherited as of 2002. These data are summarized in

Table 2. Sector and Location of Work for 20 Year-old Individuals by Village, 1980-2010

VARIABLE		mean	sd	min	max	observations
agriculture	overall	.291	.361	0	1	2,023
	between	.	.176	0	.818	80
	within	.	.317	-.527	1.24	25.3
non-agriculture	overall	.359	.373	0	1	2,023
	between	.	.157	.052	.815	80
	within	.	.338	-.382	1.31	25.3
self-employed agriculture	overall	.114	.248	0	1	2,023
	between	.	.116	0	.516	80
	within	.	.22	-.402	1.07	25.3
agriculture salary workers	overall	.179	.3	0	1	2,023
	between	.	.133	0	.479	80
	within	.	.27	-.3	1.13	25.3
seasonal migration	overall	.025	.114	0	1	2,023
	between	.	.032	0	.16	80
	within	.	.109	-.134	.988	25.3
year-round migration	overall	.187	.3	0	1	2,023
	between	.	.113	.017	.475	80
	within	.	.278	-.288	1.11	25.3
local agriculture	overall	.262	.349	0	1	2,023
	between	.	.168	0	.759	80
	within	.	.307	-.497	1.21	25.3
local non-agriculture	overall	.19	.31	0	1	2,023
	between	.	.163	.004	.808	80
	within	.	.263	-.504	1.13	25.3
agriculture elsewhere in MX	overall	.016	.102	0	1	2,023
	between	.	.03	0	.198	80
	within	.	.098	-.182	.977	25.3
non-agriculture elsewhere in MX	overall	.113	.244	0	1	2,023
	between	.	.099	0	.425	80
	within	.	.223	-.312	1.07	25.3
U.S. agriculture	overall	.019	.103	0	1	2,023
	between	.	.032	0	.191	80
	within	.	.098	-.172	.983	25.3
U.S. non-agriculture	overall	.066	.191	0	1	2,023
	between	.	.093	0	.502	80
	within	.	.17	-.436	1.02	25.3

Table 3.

Table 3. Summary of Individual and Household Characteristics

VARIABLE	mean	sd	min	max	obs
years of education	7.78	3.62	0	17	5,114
female	.453	.498	0	1	5,114
children in hh (when age 12)	2.42	2.36	0	11	5,114
adults in hh (when age 12)	2.29	2.74	0	15	5,114
indigenous language (hh head)	.139	.346	0	1	4,677
inherited land (tens of ha)	.165	1.77	0	50.7	5,114

The mean years of education in the full sample is 7.78. However, years of education differs substantially across generations, the younger generations being more highly educated than the older generations on average. Table 4 shows the years of education by age in 2010. Individuals in their twenties have expected education of 9 years while those in their fifties have expected education of only 5 years. This is an impressive rise in education in a short period of time, and it reflects the rise in access to secondary education between 1970 and 2000.

Table 4. Years of Education by Age in 2010

Age in 2010	**Mean**	**sd**	**Min**	**Max**	**Obs**
20-29	8.94	3.42	0	17	1,320
30-39	7.74	3.67	0	21	1,314
40-49	6.58	3.96	0	18	996
50-59	5.04	3.65	0	19	614

One of the factors that likely prevents many children from attending more years of school is poor access to schools. Many children in rural Mexico have to travel to other locations to attend school, which entails high costs for many families. Table 5 shows where students in ENHRUM villages, sorted by level of education, attended school in 2010.[5] 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. As the distance to school increases, attending school becomes more costly, both in the expense of travel and in the opportunity cost of time.

Table 6 records the annual expenses families reported paying for transportation to schools elsewhere in Mexico in 2010.[6] The interpretation of these data are limited since they do not include the amounts that extra-marginal families would have paid to send their children to school, the budget constraints of households, or the opportunity costs of time, which might be especially high for older children. These summary statistics suggest that distance to school may be an important factor determining years of education, as the theoretical model in Section III predicts.

Table 5. 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
Lower-secondary	frequency	6,386	3,534	124	10,044
	percentage	63.58	35.19	1.23	100
Upper-secondary	frequency	1,674	3,565	155	5,394
	percentage	31.03	66.09	2.87	100

This study identifies the impact of an additional year of school on the probability of working in agriculture. Regressing the dependent variable directly on own education, omitted variables are expected to bias the coefficient on education. Including village fixed

[5] Upper-secondary school refers to grades 10-11, 12, or 13 depending on the program.

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

Table 6. Mean Cost of School Transportation in 2010 if school is Located Outside the Home Village (2010 pesos)

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

effects will control for any unobservable location-specific factors. However, education is self-selected, so this is not expected to control for all unobservables. For example, someone who plans to work in the non-farm sector may choose to attend more years of school because the skills are more relevant for non-farm work, biasing the coefficient on education downwards (assuming that the coefficient is negative). I cannot observe innate ability, and high-ability children may attend more years of school and obtain non-farm work on account of their innate ability, biasing a negative coefficient downward as well. On the other hand, families may require their brightest, most capable children to begin working to support the household before they finish school. The opportunity cost of attending school for these children would be particularly high so they do not attend school, but if having more years of school does help individuals obtain non-farm jobs, then the potential benefit of attending school might also be greater for the high-ability children because they would be able to best apply the skills learned in school to reap financial rewards. This would bias a negative coefficient on education towards zero. Likewise, if wealthier families own more land and send their children to school for more years while their poorer neighbors, who have little land and take their children out of school to work, then wealthy children may be more likely to attend more years of school and manage their family's farm, biasing the coefficient towards zero. This shows that the sign of the bias is unknown ex-ante and an exogenous change in education is needed to isolate the variation of interest. I use changes in access to schools within villages as an instrument.

I create a proxy for school access using annual village-level enrollment rates of 12 year-old children recorded in the ENHRUM surveys each year. This is the age when children typically begin secondary school. I use sustained increases in the school enrollment rates in a village as an indicator that the village acquired access to a secondary school, likely through school construction or provision of a qualified teacher. Since education is traditionally low in these rural villages, qualified teachers are unlikely to come from within the village reducing the probability of endogenous selection based on village demand for a school.

School enrollment rates are calculated by the percentage of 12 year-old children who enroll in secondary school each year. When, for 4 consecutive years, at least 50 percent of children aged 12 attend school, then I assume that the village gained access to secondary school in the first of the 4 years. In some village-years there are no 12 year-old children in the sample (or the education of the 12 year-old children are missing). Therefore, I allow for up to 2 missing values within the stretch of consecutive years with sustained enrollment rates. If I do not observe a change in school enrollment rates for a village, then I assume that the village did not receive access to a secondary school before 2010. Table 7 summarizes the number of 12 year-old children with education data by village-year for years 1970 through 2010. There are 2.5 children per village-year on average with a range from 0 to 11.

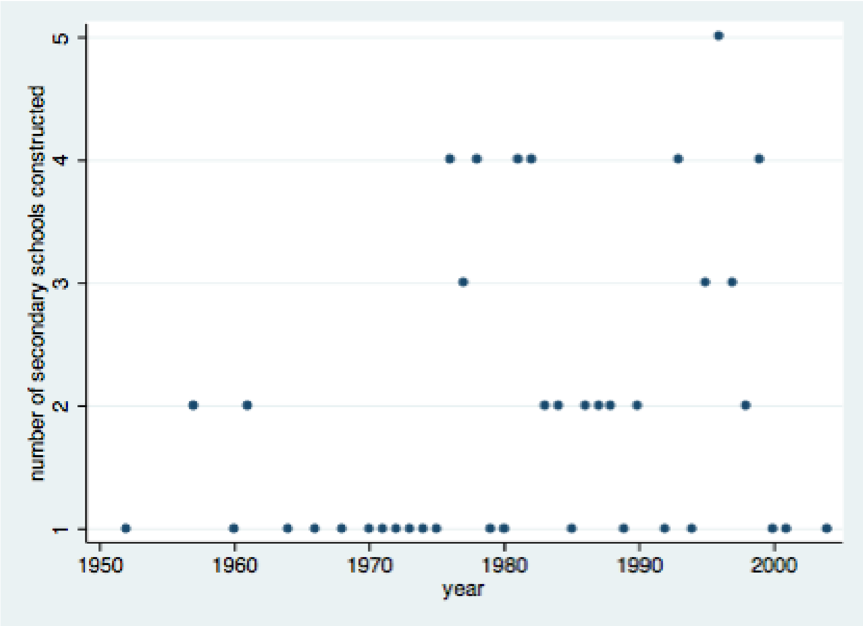
Table 7. Mean Number of 12 year-olds per Village-Year (Years 1970 through 2010)

	mean	**sd**	**minimum**	**maximum**	**observations**
Number of 12 year-olds	2.53	1.76	0	11	3,175

Figure 3 plots the number of villages where I observe changes in access to secondary schools each year using this proxy. I can observe the individual work choices at age 20 of individuals with access to secondary school if their village gained school access no later than 2002.

If observed changes in enrollment rates are a good proxy for school construction, then

Figure 3. Number of Villages that Gained Secondary School Access each Year



there should be sustained improvements in school enrollment rates in all years after the proxy turns one. I do find that the school enrollment rates are significantly higher in subsequent years to the switch. Figure 4 shows the kernel densities of secondary school enrollment rates before and after the proxy turns 1. There is a marked improvement in school enrollment rates in years subsequent to the proxy change, providing support that the proxy captures changes in school supply.

Figure 4. Kernel Density of Secondary School Enrollment Rates by Proxy for Access to Secondary School

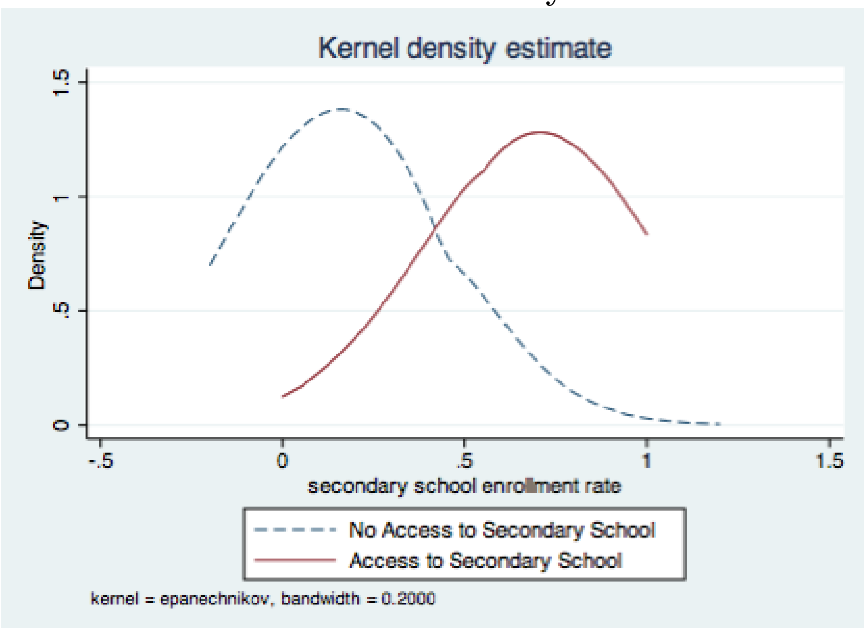


Figure 5 demonstrates the correlation between school access and mean years of education. Expected years of education are rising in years before and after schools are constructed. The x-axis in Figure 5 indicates how many years after the village gains access to a secondary school that the individual turns school age. Negative numbers indicate that the individual is too old to benefit from the school. The mean years of education jump upwards for the cohort that becomes school age in the year that the village gains school access to around 9 years of school, or the completion of lower-secondary school.

Figure 5. Mean Years of Education for Individuals who Turned 12 Before and After their Village Gained Access to Secondary School

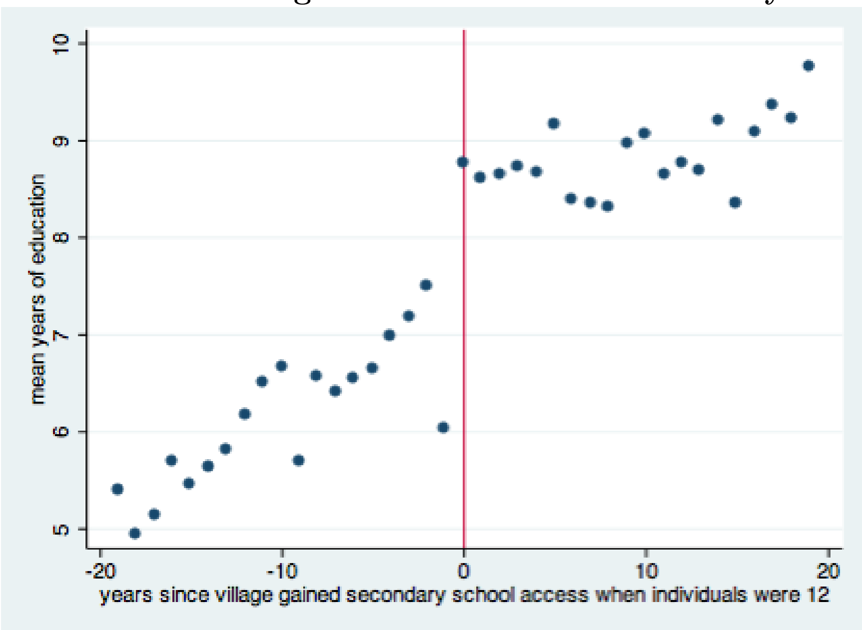
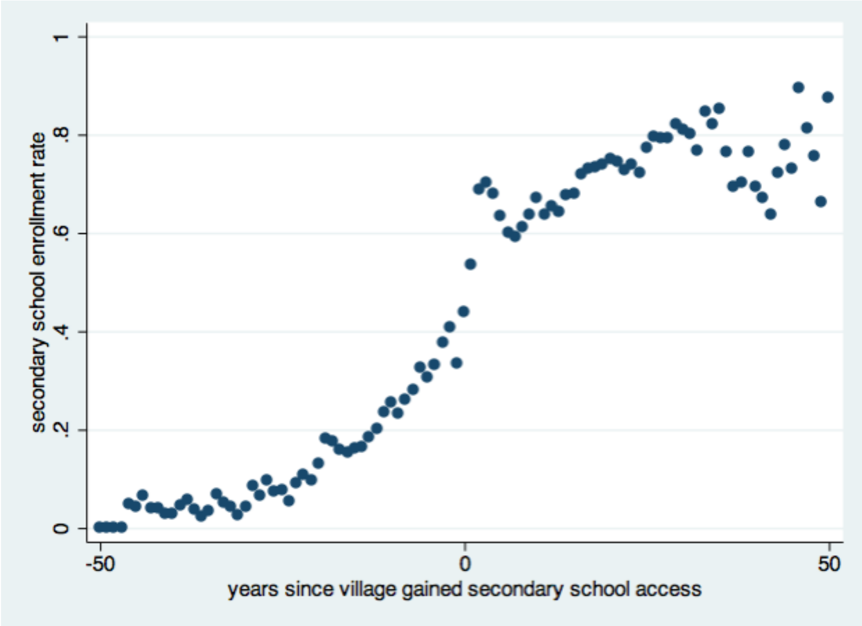


Figure 6 shows the mean secondary school enrollment rates before and after the proxy switches to one, indicating that the village gained access to secondary school.

ENHRUM includes surveys of community infrastructure in 2002 and 2007. To validate the instrument, I compare the proxy to the recorded school level located in the village in 2002 and 2007. Table 8 records the number of villages where the highest level of school located in the village is primary, lower-secondary, and upper-secondary school in 2002 and 2007 according to the ENHRUM community survey. Table 9 records the number

Figure 6. Mean Secondary School Enrollment Rate Before and After Villages Gained Access to Secondary School



of observations where the proxy and ENHRUM match on their indication that a village has or does not have a secondary school. It also records the number of observations where the proxy indicates that a village does have access to a secondary school while the ENHRUM community data indicates that a secondary school is not located in the village. Since in some villages, children can easily attend school in a neighboring village it is not surprising that there are many observations where the proxy indicates that children have access to secondary school while the ENHRUM community survey indicates that there is no secondary school in the village. It is harder to explain why the proxy would not detect access to a secondary school when a school does in fact exist in the village. Possibly the quality of teaching is low, and families choose not to send their children to school in these villages, keeping enrollment rates low. There are only 2 observations in 2007 and in 2010 where the proxy indicates that children do not have access to a secondary school and ENHRUM records that a secondary school exists inside the village.

Table 8. Highest Level of School in Village According to ENHRUM Community Survey

	2002	2007
primary	24	23
lower-secondary	47	45
upper-secondary	9	12

Table 9. Matches between Proxy and ENHRUM Community Survey Regarding Access to Secondary School

	2002	2007
proxy and ENHRUM: yes secondary school access	54	55
proxy and ENHRUM: no secondary school access	3	2
(proxy: yes) and (ENHRUM: no)	21	21
(proxy: no) and (ENHRUM: yes)	2	2
observations	80	80

IV. Empirical Approach

Many studies have attempted to measure the impacts of individual education on labor outcomes. Identification in these models is complicated by several empirical challenges.

The simplest, reduced form model regresses the outcome, whether an individual works in agriculture at age 20 on a constant and years of education. Let $Y_{i,v,t}$ be the outcome of interest. To begin, let $Y_{i,v,t}$ equal 1 if individual i from village v works in agriculture in year t, that is when he is 20 years old, and zero otherwise. Let edu_i be the years of education that i has attained.

$$Y_{i,v,t} = \gamma + \beta edu_i + \epsilon_{i,v,t}, \tag{1}$$

The coefficient β is expected to give biased and inconsistent estimates for the impact of education on the probability of working in the farm sector. Both education and work sector are self-selected variables and there are many unobservable variables likely to impact both. I can control for observed characteristics, X_i, including gender, how many children and adults lived in i's household when he was school-age, and how much agricultural land i's household has inherited. However, I expect that the location of one's home village also influences how many years i attends school and where i works as an adult. Children from villages that are better connected to urban centers likely have better access to schools and better access to non-farm jobs. I can control for unobserved, time-invariant village characteristics by including village fixed effects, λ_v, as in Equation (2).

$$Y_{i,v,t} = \alpha + \beta edu_i + \gamma X_i + \lambda_v + \epsilon_{i,v,t} \tag{2}$$

The coefficient in Equation (2) will still be biased if there are omitted variables that change over time, correlated with both education and sector selection. For example, government policies to increase mandatory years of education followed by an increase in non-farm GDP might lead to increases in both years of education and the probability of working in the non-farm sector, but the correlation is spurious. To control for simultaneous shocks, I control for year fixed effects, ϕ_t. This is similar to the differences-in-differences estimator (DD). The DD estimator measures the variation of outcomes within villages where years of education rose (or fell) relative to villages where education remained constant.

$$Y_{i,v,t} = \alpha + \beta edu_i + \lambda_v + \phi_t + \epsilon_{i,v,t} \tag{3}$$

The key assumption for DD is that education and sector choice trends would be the same across all villages absent of treatment (that is absent any changes in education within villages). However, this does not seem like a realistic assumption since I expect that trends differ across villages. Labor may transition out of agriculture more quickly in villages that are located closer to industrial centers or that have stronger migrant networks, for example. I therefore control for differences in trends across villages by controlling for village-specific trends, $\lambda_v t$.

$$Y_{i,v,t} = \alpha + \beta edu_i + \lambda_v + \phi_t + \lambda_v t + \epsilon_{i,v,t} \quad (4)$$

If, after controlling for observable characteristics, village fixed effects, year fixed effects, and village trends, i's education is as good as random, then Equation (4) will measure consistent, unbiased estimates of the expected impact of an additional year of education on labor outcomes. However, even within villages, I expect that innate ability and other unobservable individual characteristics influence years of education and labor sector outcomes. I address this concern by using exogenous changes in supply of education within villages. Let D_i be a dummy variable equal to 1 if i's village had secondary school access when i was 12 years old and 0 otherwise. I can regress the outcome directly on D_i to find the impact of having access to secondary school as a child on the probability of working in agriculture as an adult. Since individuals and their families cannot determine the year that schools are constructed in their villages, D_i generates exogenous changes in expected years of education.

$$Y_{i,v,t} = \alpha + \beta D_i + \lambda_v + \phi_t + \lambda_v t + \epsilon_{i,v,t} \quad (5)$$

Equation (5) shows the impact of having local access to secondary education on the outcome of interest (working in agriculture or migrating). To find the marginal impact of a year of education, I use school access as an instrument for education. The first stage estimates Equation 5 where years of education is the dependent variable. The second stage estimates Equation 6 where $\tilde{edu}_i$ refers to the years of education estimated in the first stage, where both stages are estimated simultaneously to adjust for more efficient standard errors.

$$Y_{i,v,t} = \alpha + \beta \tilde{edu}_i + \lambda_v + \phi_t + \lambda_v t + \epsilon_{i,v,t} \quad (6)$$

The coefficient estimated in Equation (6) is expected to yield unbiased, consistent estimates for the marginal impact of a year of education on the probability of working in agriculture.

V. Results

V - 1 Probability of Working in Agriculture Regressed on Own Education

Table 8 reports the results from regressing the dummy for working in agriculture at age 20 directly on own education. The first column includes a constant, years of education, and no additional controls. Column (2) controls for observable individual and household characteristics, including gender, the number of children under age 15 in the household when the individual was 12 years old, the number of adults in the household ages 15 to 65 when the individual was 12 years old, and how many hectares of land the household inherited as of 2002 (scaled to hundreds of hectares). Inclusion of observed characteristics shows no impact on the coefficient for education. However, the additional controls show that women are about 25 percentage points less likely to work in agriculture and those who grow up in households with more younger children are slightly more likely to work in agriculture. Once I control for village and year fixed effects, additional hectares of inherited land are associated with higher probability of working in agriculture. Column (3) includes village fixed effects, column (4) includes year fixed effects, and column (5) includes village-specific trends. After I control for village fixed effects in column (3) the magnitude of the coefficient on years of education shrinks from -2 percentage points to -1.5 percentage points, demonstrating that unobserved time-invariant characteristics of the village are correlated with educational attainment and sector choice. The coefficient changes slightly with the inclusion of year fixed effects and village time trends, but the changes are small and the coefficient remains highly significant. After including village trends in column (5), the model indicates that an additional year of education reduces the probability of working in agriculture by 1.4 percentage points.

Table 8. Probability of Working in Agriculture Regressed on Education
Linear Probability Model (Age 20)

VARIABLES	(1) agriculture	(2) agriculture	(3) agriculture Village FE	(4) agriculture Village FE Year*State FE	(5) agriculture Village FE Year*State FE Village Trends
years of education	-0.023	-0.020	-0.015	-0.013	-0.014
	(0.003)***	(0.003)***	(0.002)***	(0.002)***	(0.002)***
female		-0.252	-0.253	-0.241	-0.240
		(0.020)***	(0.019)***	(0.019)***	(0.019)***
children in hh		0.009	0.009	0.012	0.011
		(0.004)**	(0.004)**	(0.004)***	(0.004)***
adults in hh		-0.001	-0.002	0.000	0.001
		(0.003)	(0.003)	(0.003)	(0.003)
inherited land (hundreds of ha)		0.061	0.079	0.070	0.070
		(0.063)	(0.050)	(0.052)	(0.053)
Observations	5,138	5,138	5,138	5,138	5,138
R-squared	0.032	0.136	0.244	0.312	0.329

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

V - 2 Reduced Form: Regress Probability of Working in Agriculture on School Access

The coefficients reported in Table 8 are only consistent if years of education are as good as randomly assigned once I control for village fixed effects, year fixed effects, and village trends. If own education is correlated with individual unobservable characteristics that are not absorbed by the fixed effects or the control variables, then the coefficient on education will be inconsistent and biased. In Table 9, I use an exogenous proxy for education, changes in local access to secondary schools. Column (1) regresses the dummy variable for working in agriculture directly on secondary school access with no additional controls. Column (2) includes observable characteristics, and the remaining columns respectively add village fixed effects, year fixed effects, and village-specific trends.

The impact of secondary school construction is not significant when I include state-year fixed effects. However, it is large and significant after I additionally control for village trends, suggesting that confounding trends mask the effects of secondary school access when I do not control for them. The results show that exogenous gains in access to secondary school when school age reduce the probability of working in agriculture as an adult by 6.2 percentage points.

V - 3 Instrumental Variables: Two-Stage Least Squares

The reduced form regressions in Table 9 suggest that improved access to secondary schools decreases the probability of working in agriculture. The predicted mechanism of the impact is that local provision of secondary schools increases the expected years of education and additional years of education reduce the probability of working in agriculture. To identify the marginal effects of an additional year of education, I use access to secondary schools as an instrument for education.

Table 10 reports the first stage results of the impact of school access on expected years of education. The coefficient on access to secondary schools is highly significant

Table 9. Effects of School Construction on the Probability of Working in Agriculture
Reduced Form Linear Probability Model (Age 20)

VARIABLES	(1) agriculture	(2) agriculture	(3) agriculture Village FE	(4) agriculture Village FE Year*State FE	(5) agriculture Village FE Year*State FE Village Trends
secondary school access	-0.153	-0.140	-0.075	-0.022	-0.062
	(0.030)***	(0.029)***	(0.019)***	(0.025)	(0.028)**
female		-0.253	-0.255	-0.242	-0.242
		(0.020)***	(0.020)***	(0.020)***	(0.020)***
children in hh		0.009	0.009	0.012	0.012
		(0.004)**	(0.004)**	(0.004)***	(0.004)***
adults in hh		0.001	-0.001	0.000	0.001
		(0.003)	(0.003)	(0.003)	(0.003)
inherited land (hundreds of ha)		0.044	0.070	0.063	0.062
		(0.055)	(0.048)	(0.051)	(0.052)
Observations	5,138	5,138	5,138	5,138	5,138
R-squared	0.027	0.131	0.235	0.304	0.321

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

in all specifications. The coefficient becomes smaller with the inclusion of year fixed effects and village-specific trends. This is not surprising since the demand for education is rising throughout Mexico along with expanded supply of schools, leading to spurious correlation if trends are not accounted for. In the final specification I find that local access to secondary school increases the expected years of education by 1.327 years. Since secondary school is 3 years, these findings show that take-up is not complete. Some students choose not to attend school even though a local school is supplied and some students choose not to complete secondary school even though they begin. Still other students attend school in years prior to school construction, further reducing the impact of constructing schools on years of school attended.

Table 10. Effects of School Construction on Expected Years of Education

VARIABLES	(1) education	(2) education	(3) education Village FE	(4) education Village FE Year*State FE	(5) education Village FE Year*State FE Village Trends
secondary school access	2.648	2.567	2.009	1.302	1.327
	(0.180)***	(0.186)***	(0.180)***	(0.201)***	(0.257)***
female		0.193	0.166	0.100	0.103
		(0.099)*	(0.099)*	(0.100)	(0.102)
children in hh		-0.016	0.009	-0.018	-0.022
		(0.029)	(0.027)	(0.033)	(0.033)
adults in hh		0.015	-0.000	-0.016	-0.012
		(0.024)	(0.023)	(0.028)	(0.029)
inherited land (hundreds of ha)		0.920	0.619	0.543	0.573
		(0.349)**	(0.260)**	(0.260)**	(0.269)**
Observations	5,138	5,138	5,138	5,138	5,138
R-squared	0.131	0.144	0.216	0.292	0.309

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

Table 11 reports the second stage results, instrumenting for education using local access to secondary schools. The results are significant at the 5 percentage level in the specification that controls for village fixed effects, state-year fixed effects, and village-specific trends. The magnitude of the coefficient decreases when village fixed effects are included, and it rises somewhat in the fifth column compared to column (3). The results indicate that an additional year of school reduces the probability of working in agriculture by 4.7 percentage points.

The coefficient on education is larger in the IV estimate than in the naive OLS regression, and the coefficients on the control variables for individual and household characteristics are similar to those found in the reduced form model.

The test for endogeneity of education using the Durbin-Wu-Hausman test, which allows for heteroskedastic errors, shows that the IV results are significantly different from the naive OLS results. I find an F-statistic of $F(1, 79) = 2.84$ and a corresponding p-value of 0.096. A possible explanation for the large coefficients on education in the IV regressions is that access to secondary school is a weak instrument, causing bias. However, I do not find evidence of a weak instrument when I test the strength of the instrument. In the last column, which includes village and year fixed effects and village trends (and in all columns) the instrument is strong. The first-stage F-statistic in Column (5) is $F(1,79) = 27.887$, which is far above the threshold F-statistic of 10. Since the standard errors in the regression are clustered at the village level, I use Stock and Yogo's test for weak instruments, which adjusts for heteroskedastic standard errors, and I again reject the null hypothesis that the instrument is weak with a high level of confidence. These tests suggest that the large results found in the IV specification are not the product of weak instrumental variable bias.

V - 4 Migration Outcomes

Table 12 reports the results from regressing several migration-sector dummy variables on Equation (4). The results show that additional years of education are associated with greater year-round migration (i.e. the individual does not work locally at all during the

Table 11
Effects of Education on the Probability of Working in Agriculture at Age 20
Instrument for Education using Access to Secondary Schools (2SLS)

VARIABLES	(1) agriculture	(2) agriculture	(3) agriculture Village FE	(4) agriculture Village FE Year*State FE	(5) agriculture Village FE Year*State FE Village Trends
years of education	-0.058	-0.054	-0.037	-0.017	-0.047
	(0.010)***	(0.011)***	(0.009)***	(0.017)	(0.018)**
female		-0.243	-0.248	-0.241	-0.237
		(0.020)***	(0.019)***	(0.018)***	(0.017)***
children in hh		0.008	0.009	0.011	0.011
		(0.004)*	(0.004)**	(0.004)***	(0.004)***
adults in hh		0.002	-0.001	0.000	0.001
		(0.003)	(0.003)	(0.002)	(0.002)
inherited land (hundreds of ha)		0.094	0.093	0.072	0.089
		(0.064)	(0.053)*	(0.050)	(0.054)
Observations	5,138	5,138	5,138	5,138	5,138
R-squared		0.064	0.219	0.311	0.281
Test for endogeneity of education. H_0: education is exogenous					
F(1,79)	14.15***	12.65***	5.65**	0.05	2.96*
First-stage F-stats. H_0 secondary=0 using robust standard errors					
F(1,79)	216.9***	192.7***	127.04***	42.73***	27.658***
Stock and Yogo's Test for weak instruments. H_0: instrument is weak (The critical value for 2SLS of nominal 5% Wald test rejection of the null at 10% is 16.38)					
min eigenvalue stat	772.34	627.59	183.89	45.18	34.237

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

year), lower probability of working in agriculture either locally or elsewhere in Mexico, and greater probability of working in the non-farm sector elsewhere in Mexico.

Table 13 reports the IV results. The results show that additional years of education reduce the probability of working in the agricultural sector in Mexico (both locally and elsewhere in Mexico), and education increases the probability of doing local non-farm work. I find no evidence that education increases labor mobility once I instrument for education using school access.

The coefficients on the control variables indicate that gender is an important factor determining where one works. Women are less likely to work in any sector and location, and in particular they are less likely to work in the local agricultural sector and they are less likely to migrate year-round than are men. Individuals from homes with more young children when the individual is growing up are significantly more likely to migrate for the full year and they are more likely to work in the local farm sector. Finally, I find that additional hectares of inherited land increase the probability of doing local agricultural work.

Table 12. Regress Migration on Own Education
Linear Probability Model (Age 20)

VARIABLES	(1) migrate seasonally	(2) migrate all year	(3) local ag	(4) local nonag	(5) mx ag	(6) mx non-ag	(7) us ag	(8) us non-ag
years of education	-0.001	0.004	-0.011	0.001	-0.002	0.007	-0.001	-0.001
	(0.001)	(0.002)**	(0.002)***	(0.002)	(0.001)***	(0.001)***	(0.000)	(0.001)
female	-0.031	-0.097	-0.205	-0.004	-0.017	-0.015	-0.029	-0.068
	(0.006)***	(0.015)***	(0.019)***	(0.017)	(0.005)***	(0.013)	(0.007)***	(0.012)***
children in hh	0.000	0.005	0.009	0.009	-0.000	0.004	0.002	-0.000
	(0.001)	(0.004)	(0.004)**	(0.003)***	(0.001)	(0.003)	(0.001)	(0.002)
adults in hh	0.000	0.010	0.001	0.004	-0.000	0.007	-0.000	0.004
	(0.001)	(0.004)***	(0.003)	(0.002)	(0.001)	(0.003)**	(0.001)	(0.002)**
inherited land (hundreds of ha)	-0.011	-0.010	0.072	-0.029	-0.004	-0.026	0.001	0.008
	(0.005)**	(0.018)	(0.053)	(0.061)	(0.003)	(0.012)**	(0.004)	(0.010)
Observations	5,138	5,138	5,138	5,138	5,138	5,138	5,138	5,138
R-squared	0.134	0.203	0.301	0.254	0.146	0.214	0.171	0.212

Village FE, Year*State FE, and Village Trends included in each regression
Robust standard errors in parentheses, clustered at the village level
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 13. Impact of School Access on Migration at Age 20
Instrument for Education using Local School Access: Linear Probability Model)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	migrate seasonally	migrate all year	local ag	local non-ag	mx ag	mx non-ag	us ag	us non-ag
years of education	-0.002	0.027	-0.040	0.020	-0.011	0.023	0.001	0.010
	(0.007)	(0.021)	(0.016)**	(0.017)	(0.004)**	(0.016)	(0.008)	(0.013)
female	-0.031	-0.099	-0.202	-0.006	-0.016	-0.017	-0.029	-0.069
	(0.006)***	(0.015)***	(0.018)***	(0.016)	(0.004)***	(0.012)	(0.007)***	(0.011)***
children in hh	0.000	0.005	0.009	0.010	-0.000	0.004	0.002	-0.000
	(0.001)	(0.004)	(0.004)**	(0.003)***	(0.001)	(0.003)	(0.001)*	(0.002)
adults in hh	0.000	0.011	0.001	0.004	-0.000	0.007	-0.000	0.005
	(0.001)	(0.003)***	(0.003)	(0.002)*	(0.001)	(0.003)***	(0.001)	(0.002)**
inherited land (hundreds of ha)	-0.010	-0.023	0.088	-0.039	0.001	-0.035	0.000	0.002
	(0.006)*	(0.021)	(0.052)*	(0.059)	(0.005)	(0.017)**	(0.006)	(0.009)
Observations	5,138	5,138	5,138	5,138	5,138	5,138	5,138	5,138
R-squared	0.134	0.171	0.262	0.233	0.097	0.191	0.170	0.195

Village FE, State*Year FE, and Village Trends included in each regression
*** p<0.01, ** p<0.05, * p<0.1

VI. Robustness Checks

VI - 1 Labor Decisions at Ages 25 and 30

The analysis thus far investigates the labor decisions of adults when they are 20 years old. I now investigate whether these results are robust to labor decisions at older ages. Table 14 shows the results for adults at age 25. All specifications include village fixed effects, state-year fixed effects, and village-specific trends. Column (1) is the naive OLS regression that regresses the dummy for working in agriculture directly on own education. It shows that an additional year of school is associated with a 1.6 percentage point decrease in the probability of working in agriculture, similar to the results in Table 8 for 20 year-olds. Column (2) shows the results for the reduced form impact of gaining secondary school access on the probability of working in agriculture. The results are significant at the 5 percentage level and they are larger than the results found in Table 9. For 25 year-olds, having local access to secondary school reduces the probability of working in agriculture by 7.5 percentage points, 1.3 percentage points more than found for 20 year-olds. Column (3) shows the first-stage results of regressing education on secondary school access. The results differ from those found in Table 10 because the sample is different. I do not observe individuals who turn 20 near the end of the time frame. The impact of school access appears smaller in this sample, but the results are similar in magnitude and significant at the 1 percentage point level. Finally, column (4) reports the second stage results for the impact of education on the probability of working in agriculture. An additional year of school reduces the probability of working in agriculture at age 25 by 6.2 percentage points. This impact is quite a bit larger than the impact found for 20 year-olds.

The impact of education on labor outcomes at age 30 is also large. I find that an additional year of school reduces the probability of working in agriculture at age 30 by 10.4 percentage points and the results are significant at the 1 percentage level. The results are recorded in Table 15.

Table 14
Outcomes at Age 25

VARIABLES	(1) agriculture naive OLS	(2) agriculture reduced form	(3) education first stage	(4) agriculture 2SLS
years of education	-0.016			-0.062
	(0.002)***			(0.023)***
secondary school access		-0.075	1.214	
		(0.029)**	(0.286)***	
female	-0.232	-0.232	-0.008	-0.233
	(0.019)***	(0.019)***	(0.115)	(0.017)***
children in hh	0.004	0.004	-0.004	0.004
	(0.004)	(0.004)	(0.039)	(0.004)
adults in hh	0.004	0.004	-0.007	0.004
	(0.003)	(0.003)	(0.036)	(0.003)
inherited land (hundreds of ha)	0.040	0.035	0.289	0.053
	(0.019)**	(0.019)*	(0.196)	(0.022)**
Observations	4,762	4,762	4,763	4,762
R-squared	0.318	0.306	0.315	0.209

Test for endogeneity of education. H_0: education is exogenous	
F(1,79)	3.800*
First-stage F-stats. H_0 secondary=0 using robust standard errors	
F(1,79)	18.253***
Stock and Yogo's Test for weak instruments. H_0: instrument is weak	
(The critical value for 2SLS of nominal 5% Wald test rejection of the null at 10% is 16.38)	
min eigenvalue stat	21.776

All specifications include village FE, state*year FE, and village-specific trends.
Robust standard errors in parentheses, clustered at the village level
*** p<0.01, ** p<0.05, * p<0.1

VI - 2 Falsification Tests

I regress several pre-determined variables on Equation (5) to test whether access to secondary schools is indicative of other changes in the population. Table 16 reports the

Table 15
Outcomes at Age 30

VARIABLES	(1) agriculture naive OLS	(2) agriculture reduced form	(3) education first stage	(4) agriculture 2SLS
years of education	-0.015			-0.104
	(0.002)***			(0.033)***
secondary school access		-0.124	1.197	
		(0.028)***	(0.366)***	
female	-0.227	-0.227	-0.040	-0.231
	(0.018)***	(0.019)***	(0.130)	(0.017)***
children in hh	0.005	0.005	-0.003	0.005
	(0.005)	(0.005)	(0.046)	(0.007)
adults in hh	-0.001	-0.001	0.018	0.001
	(0.004)	(0.004)	(0.041)	(0.005)
married	0.031	0.032	-0.023	0.030
	(0.020)	(0.020)	(0.203)	(0.025)
inherited land (hundreds of ha)	0.105	0.098	0.426	0.142
	(0.040)**	(0.040)**	(0.246)*	(0.048)***
Observations	4,238	4,238	4,238	4,238
R-squared	0.325	0.315	0.338	

Test for endogeneity of education. H_0: education is exogenous

F(1,79) 15.47***

First-stage F-stats. H_0 secondary=0 using robust standard errors

F(1,79) 10.708***

Stock and Yogo's Test for weak instruments. H_0: instrument is weak

(The critical value for 2SLS of nominal 5% Wald test rejection of the null at 10% is 16.38)

min eigenvalue stat 13.546

All specifications include village FE, state*year FE, and village-specific trends.

Robust standard errors in parentheses, clustered at the village level

*** p<0.01, ** p<0.05, * p<0.1

results. I find no evidence that access to secondary school is correlated with any systematic changes in the village population except for the impact that it has on the number of adults in the household. However, the impact on the number of adults is intrinsic to the survey design since the survey follows the same households over time and all of the children at the beginning of the survey must grow up while at the same time, schools are being built later in the sample period. For this reason I control for adults in the household in the regressions of interest. The absence of significant impacts in Table 16 supports the model that access to secondary school is correlated with labor outcomes only through its influence on years of education.

Table 16. Balance Tests

VARIABLES	(1) children in hh	(2) adults in hh	(3) indigenous lang	(4) inherited land
secondary school access	0.015 (0.127)	0.339 (0.178)*	-0.011 (0.018)	0.041 (0.050)
Observations	5,114	5,114	4,677	5,114
R-squared	0.374	0.316	0.791	0.0699

Village FE, Year FE, and Village Trends included in each regression
Robust standard errors in parentheses, clustered at the village level
*** p<0.01, ** p<0.05, * p<0.1

Secondary school access at age 12 may not be a valid instrument if improved access to school is correlated with improved access to urban areas, factories, or non-farm employment. Presumably, those who were older than age 12 when the village gained access to the secondary school did not benefits from school access. However, if school access is correlated with access to non-farm employment (through road construction or a new bus route connecting the village to an urban area), then older individuals may shift out of the farm sector for reasons other than education. Table 17 tests whether improved access to secondary education in a village is correlated with a reduction in the farm labor supply the year that education access improves. Each column includes an indicator variable equal to 1 if the village has access to a secondary school when the individual is 20 years old. The coefficient on the indicator variable is not significantly different from zero in the first 2 columns. School access at age 20 is correlated with more years of education (Column (3)), that is, mean years of education within a village appear to rise before secondary school enrollment rates do.

Column (4) records the results of the 2-stage least squares specification, and the results are similar, though somewhat smaller, to those found before controlling for village school access at age 20. Village school access when age 20 is associated with a significantly greater probability of working in agriculture, which is the opposite effect from what is expected if school access is correlated with local access to non-farm employment. The coefficient on education shrinks to 3.8 percentage points when the indicator variable for village school access at age 20 is included. Controlling for other unobserved changes in the village that occur the year that the village gains secondary school access attenuates the estimated impact of education. However, the coefficient is still significant at the 5 percentage point level and quite a bit larger in magnitude than the OLS results, suggesting that educational attainment is a critical factor in the transition of labor out of agriculture. Communities likely demand better educational services because it improves their children's probability of finding non-farm jobs. Controlling for whether the village had a secondary school when age 20 helps isolate the impact of education apart from changing demand for education or other unobserved village-level changes that occur the year that the village gains secondary school access. This implies that these results may be more reliable than the results in Table 11, but the results in both tables are similar.

VII. Discussion

The findings indicate that increasing rural access to secondary education does accelerate the agricultural transformation, one of the critical stages of economic development. I find

Table 17
Control for Secondary School in Village at Age 20

VARIABLES	(1) agriculture naive OLS	(2) agriculture reduced form	(3) education first stage	(4) agriculture 2SLS
years of education	-0.013			-0.038
	(0.002)***			(0.018)**
secondary school access (age 12)		-0.053	1.408	
		(0.028)*	(0.246)***	
secondary school in village (age 20)	0.046	0.031	0.558	0.052
	(0.034)	(0.036)	(0.277)**	(0.030)*
female	-0.240	-0.242	0.105	-0.238
	(0.019)***	(0.019)***	(0.101)	(0.017)***
children in hh	0.011	0.011	-0.018	0.011
	(0.004)***	(0.004)***	(0.033)	(0.004)***
adults in hh	0.001	0.001	-0.017	0.001
	(0.002)	(0.003)	(0.029)	(0.002)
inherited land (hundreds of ha)	0.070	0.062	0.560	0.084
	(0.050)	(0.050)	(0.255)**	(0.050)*
Observations	5,138	5,138	5,138	5,138
R-squared	0.324	0.317	0.301	0.298
Test for endogeneity of education. H_0: education is exogenous				
			F(1,79)	1.777
First-stage F-stats. H_0 secondary=0 using robust standard errors				
			F(1,79)	34.277***
Stock and Yogo's Test for weak instruments. H_0: instrument is weak				
(The critical value for 2SLS of nominal 5% Wald test rejection of the null at 10% is 16.38)				
min eigenvalue stat				40.161

All specifications include village FE, state*year FE, and village-specific trends.
Robust standard errors in parentheses, clustered at the village level
*** p<0.01, ** p<0.05, * p<0.1

larger impacts of education on labor sector selection in the instrumental variables model than in the naive OLS regressions. This result is not intuitive since more schooling is usually correlated with greater unobserved ability, which is also expected to be correlated with higher probability of obtaining a non-farm job. A possible explanation for this unexpected finding is that school access is correlated with improved infrastructure, which also improves access to urban areas with more jobs in the non-farm sector. Omitting access to urban areas correlated with the year that villages obtain school access would bias the coefficient in the IV regression so that the results are larger in magnitude. I test this hypothesis in the falsification tests section, and I do not find evidence that individuals who were too old to benefit from improved school access were significantly less likely to work in agriculture than cohorts who turned 20 before the village gained access to the secondary school.

Another possible explanation for the large magnitude of the instrumental variables results is that the most skilled (or possibly most disciplined) children have the highest opportunity cost of attending school when the school is far from home. The most skilled children may be kept from school at a young age because their families need them to work and support the household. For these children the opportunity cost of traveling a long distance to school is higher than for other children. Being particularly skilled or hard-working, they also have the greatest marginal benefit from gaining access to secondary school since they are likely to learn more in school and apply it in the job sector. In this scenario, the children denied access to school are those with the greatest marginal benefit of attending school. This would bias the IV estimator downward (greater in magnitude). Then the true effect of education on the probability of working in agriculture would lie somewhere between the naive OLS estimate and the IV estimate.

This finding is consistent with Card (2001)'s observation that IV results are nearly always as large or larger than OLS results in studies that measure the returns to education using exogenous changes in school supply as an instrument for education. Card (2001) suggests that the estimated marginal returns to education using supply to school as an instrument are usually greater than the OLS estimates because the populations that are most affected by the change in school supply are those with higher relative returns to education and high marginal cost of going to school.

Identification in this paper depends on observable changes in school enrollment rates. Additional field work to learn when schools are built across all villages and how children travel to and from school over time can potentially strengthen the analysis. In future work I plan to investigate how education interacts with variables such as wealth, land-holdings, and gender to impact farm labor decisions and migration.

VIII. Conclusion

The findings in this paper show that policies directed towards improving access to education can accelerate the agricultural transformation. This is currently occurring in rural Mexico. The findings show that an additional year of school when a child is beginning secondary school reduces the probability of working in agriculture at age 20 by 3.8 to 4.7 percentage points. The impact appears more dramatic as individuals age. At age 30, I find a negative 10 percentage point impact of an additional year of school on the probability of working in agriculture.

Since the children initially denied access to secondary school might also be the children in the population with the greatest marginal benefits from education, the IV estimates on the impact of education might be larger than the average impact of education across the full population in rural Mexico. Nevertheless, results from the naive OLS specifications and the IV specifications confirm that education is a large, significant factor reducing the probability that individuals work in agriculture.

I find significant, positive correlation between education and the probability of working away from home in the naive regression but I do not find that additional years of education increase the probability of migration after I instrument for education. This shows that policies to improve access to secondary schools in rural areas decrease the farm labor supply but potentially have no impact on labor mobility.

Understanding education's role in the agricultural transformation has important implications for income, risk, and welfare. Much of the literature on the economic returns to education shows that both increasing education and obtaining work outside of the agricultural sector are associated with higher incomes and less income variability. If the returns to education are greater in the non-farm sector than in the farm sector, then policies that reduce the costs of going to school can help individuals to complete more years of school and find higher-paying jobs, which is likely to help alleviate poverty in rural areas.

Worldwide, as per capita income rises, the share of the working-age population employed in agriculture declines. Many potential mechanisms may drive this phenomenon. To my knowledge, this is the first paper to investigate the role of education in the agricultural transformation using a instrumental variables and data that are nationally representative of a rural population currently in transition out of agriculture. My findings show that providing secondary education in rural populations does significantly reduce the probability of working in agriculture, thus accelerating the pace of the agricultural transformation and helping agricultural households, typically confined in poverty, to gain a wider range of economic opportunities.

References

Barro, Robert J. 2001. "Human capital and economic growth." *The American Economic Review*, 91(2): 12–17.

Becker, Gary S, Kevin M Murphy, and Robert Tamura. 1994. "Human capital, fertility, and economic growth." In *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education (3rd Edition)*. 323–350. The University of Chicago Press.

Benhabib, Jess, and Mark M Spiegel. 1994. "The role of human capital in economic development evidence from aggregate cross-country data." *Journal of Monetary economics*, 34(2): 143–173.

Bobonis, Gustavo J, and Frederico Finan. 2009. "Neighborhood Peer Effects in Secondary School Enrollment Decisions." *Review of Economics and Statistics*, 91(4): 695–716.

Card, David. 1993. "Using geographic variation in college proximity to estimate the return to schooling." National Bureau of Economic Research.

Card, David. 2001. "Estimating the return to schooling: Progress on some persistent econometric problems." *Econometrica*, 69(5): 1127–1160.

Charlton, Diane, and J. Edward Taylor. 2013. "Mexicans are Leaving Farm Work: What Does it Mean for U.S. Agriculture and Immigration Policy?" *ARE Update*, 16(4).

Duflo, Esther. 2000. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." National Bureau of Economic Research, Inc NBER Working Papers 7860.

Foster, Andrew D, and Mark R Rosenzweig. 1996. "Technical change and human-capital returns and investments: evidence from the green revolution." *The American economic review*, 931–953.

Handa, Sudhanshu. 2002. "Raising primary school enrolment in developing countries: The relative importance of supply and demand." *Journal of Development Economics*, 69(1): 103–128.

Huffman, Wallace E. 1980. "Farm and off-farm work decisions: the role of human capital." *The Review of Economics and Statistics*, 14–23.

Janvry, Alain de, and Elisabeth Sadoulet. 2001. "Income strategies among rural households in Mexico: The role of off-farm activities." *World development*, 29(3): 467–480.

Joliffe, D. 2004. "The Impact of Education in Rural Ghana: Examining Household Labor Allocation and Returns On and Off the Farm." *Journal of Development Economics*, 73: 287–314.

Kane, Thomas J, and Cecilia Elena Rouse. 1995. "Labor-market returns to two-and four-year college." *The American Economic Review*, 85(3): 600–614.

Lalive, Rafael, and M. Alejandra Cattaneo. 2009. "Social Interactions and Schooling Decisions." *Review of Economics and Statistics*, 91(3): 457–477.

Lavy, Victor. 1996. "School supply constraints and children's educational outcomes in rural Ghana." *Journal of Development Economics*, 51(2): 291–314.

Mincer, Jacob. 1984. "Human capital and economic growth." *Economics of Education Review*, 3(3): 195–205.

Nelson, Richard R, and Edmund S Phelps. 1966. "Investment in humans, technological diffusion, and economic growth." *The American Economic Review*, 69–75.

Rolwing, K. 2006. "Education in Mexico." *World Education News and Reviews*, 19(3).

Santibañes, Lucrecia, Georges Vernez, and Paula Razquin. 2005. "Education in Mexico: Challenges and Opportunities." RAND Corporation.

Schultz, P T. 2004. "School subsidies for the poor: evaluating the Mexican Progresa poverty program." *Journal of development Economics*, 74(1): 199–250.

Skoufias, Emmanuel, Benjamin Davis, and Sergio De La Vega. 2001. "Targeting the poor in Mexico: an evaluation of the selection of households into PROGRESA." *World Development*, 29(10): 1769–1784.

Taylor, J. Edward, Diane Charlton, and Antonio Yúnez-Naude. 2012. "The End of Farm Labor Abundance." *Applied Economic Perspectives and Policy*, 34(4): 587–598.

Zhang, Linxiu, Jikun Huang, and Scott Rozelle. 2002. "Employment, emerging labor markets, and the role of education in rural China." *China Economic Review*, 13(2): 313–328.