

NBER WORKING PAPER SERIES

INTERNATIONAL MIGRATION, REMITTANCES, AND SCHOOLING:  
EVIDENCE FROM EL SALVADOR

Alejandra Cox

Manuelita Ureta

Working Paper 9766

<http://www.nber.org/papers/w9766>

NATIONAL BUREAU OF ECONOMIC RESEARCH

1050 Massachusetts Avenue

Cambridge, MA 02138

June 2003

Paper presented at the Inter-American Seminar in Economics organized by the National Bureau of Economic Research and ITESM, Monterrey Campus – Mexico, November 15-16 2002. The authors benefited from detailed comments from Bonnie Palifka and from valuable suggestions by Sebastian Edwards, William W. Gould, Audrey L. Light, Kevin M. Murphy, Wayne Strayer, Finis Welch and workshop participants at Ohio State University and Texas A&M University. Manuelita Ureta gratefully acknowledges financial support for this project from The Bush School of Government and Public Service. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

©2003 by Alejandra Cox Edwards and Manuelita Ureta. All rights reserved. Short sections of text not to exceed two paragraphs, may be quoted without explicit permission provided that full credit including © notice, is given to the source.

International Migration, Remittances, and Schooling: Evidence from El Salvador

Alejandra Cox Edwards and Manuelita Ureta

NBER Working Paper No. 9766

June 2003

JEL No. F22, I22, I28, J24, O15

**ABSTRACT**

We examine the effect of remittances from abroad on households' schooling decisions using data for El Salvador. Following the massive war-related emigration of the 1980's, remittances became a significant source of household income throughout the 1990's. We use the Cox proportional hazard model to examine the determinants of school attendance. Measuring income from a source that is uncorrelated with parental schooling— remittances—, we find that remittances have a large, significant effect on school retention. We estimate that while household income net of remittances has a small, though significant, impact on the hazard of leaving school in rural and urban areas, remittances have a much larger impact on the hazard of leaving school. In urban areas, the effect of remittances is, at its smallest, 10 times the size of the effect of other income. In rural areas, the effect of remittances is about 2.6 times that of other income. Our finding is of interest in that it suggests that subsidizing school attendance, particularly in poor areas, may have a large impact on school attendance and retention, even if parents have low levels of schooling.

Alejandra Cox Edwards Manuelita Ureta

Department of Economics Texas A&M University

California State University

1250 Bellflower Blvd

Long Beach, CA 90840

Globalization is generally studied from the point of view of international trade, finance or migration. One aspect of globalization that is less understood is that of remittances. We examine the effect of remittances on households' schooling decisions using data for El Salvador. El Salvador has significantly expanded its educational coverage after the civil war that ended in 1989. In 1997, 75 percent of rural children and 92 percent of urban children aged 12 were enrolled in school. The recent higher rates of retention in school are a remarkable accomplishment for a country that has 30 percent illiteracy rate among adults aged 40 to 50, and 25 percent illiteracy among adults aged 30 to 40. The record of expansion of the educational system in El Salvador during the 1960's and 1970's was dismal in comparison with the record of other developing countries. Schultz (1989) notes that in the period from 1960 to 1981 El Salvador was the only country to report a decline in expected years of schooling, from 5.63 to 5.06 years, according to the *1984 World Development Report* (table 25).

There are two dimensions of the case of El Salvador that make it unique and interesting. First, the institutional setting for public policy in the educational sector is such that the expansion of school facilities is primarily driven by the active participation of parents in the allocation of public and private funds. Parents have played a leading role financing the expansion of private schools, which have led the urban expansion. To expand primary school coverage in rural areas, a priority of the government in the 1990's, the Ministry of Education allocates resources to parents' associations enabling them to hire teachers and buy materials.

Second, following the massive out migration of the 1980's, remittances from abroad have become a significant source of household income throughout the 1990's.<sup>1</sup> In 1997, 14 percent of rural and 15 percent of urban households received remittances from relatives or friends living abroad. Income from remittances expands the household budget

---

<sup>1</sup> The literature has produced a number of studies that look into the flow of remittances, its persistence (see for example, Funkhouser (1992) and Montes (1990)) and its impact on the economy (see for example Lopez (1991)).

constraint allowing some families to find it optimal that their children acquire more schooling than they would acquire otherwise.

□

Our sample is a cross-section of 14,286 individuals aged 6 to 24, from the 1997 Annual Household Survey. This is a nationally representative sample collected by the National Statistics Office (Digestyc). We use the Cox proportional hazard model to estimate the impact of characteristics of the individual and the family—including remittances and income separately—on the hazard of dropping out of school. We find that remittances have a significant impact on school retention.

□

The use of the Cox proportional hazard model to study school attainment with cross-sectional data is an innovation.<sup>2</sup> The model poses some challenges, which we address, and at the same time it offers several advantages over its alternatives. First, the Cox proportional hazard model makes use of all the available information in observations for children or young adults who are enrolled in school at the time of the survey (9,547 of the 14,286 in our sample) as well as for those who are no longer enrolled. One popular alternative is to study school completion among older individuals, who are expected to have finished their formal education. An analysis of school attainment based on the sample of individuals aged 25 and over would be especially undesirable in El Salvador where there have been recent efforts to expand school enrollment throughout the country. Second, the model yields estimates of survival functions, allowing us to establish, for example, that the 1980-83 urban birth cohort enjoyed the largest improvement in school attainment relative to the cohort immediately prior, and that the 1984-87 rural cohort experienced, by far, the largest improvement relative to the cohort that preceded it. Third, the model allows very flexible specifications, and we find, for example, that the

---

<sup>2</sup>—There are two papers that use proportional hazard models that came to our attention recently. Inchauste (2000) analyzes Bolivian data and uses a hazard model to estimate the probability that children quit school in period  $t+1$  given that they attended school in period  $t$ . The hazard estimates are weighted by the probability of being constrained (which is driven by variables such as “no financial resources”, school being “too far away”, and “having to work.”) Brown and Park (2001) use proportional hazard models to estimate the impact of several variables (including credit constraints) on the likelihood of dropping out of school at the primary and secondary level. Neither of these papers discusses the advantages or shortcomings of using Cox proportional hazard models as we do here.

effect of parental schooling on the hazard of leaving school falls as the level of the child's schooling increases.

In what follows we begin with a brief discussion of the literature followed by a description of the institutional setting of the school system in El Salvador; we then describe the data and methodology and present our results.

### **Theoretical Issues**

Assuming that parents care about their children's well being and that they can borrow to finance productive investments, optimal investment in schooling is a function of net rates of return. These in turn are a function of benefits and costs. Differences in rates of return arising from exogenously determined, individual-specific traits, such as innate ability, likely provide a partial explanation for observed differences in completed schooling across members of a given family. Differences in rates of return due to location-specific factors, such as distance to school, likely explain part of the difference in average completed schooling across residents of different towns. Differences in rates of returns driven by economy-wide conditions, such as the labor market valuation of schooling, likely explain some of the difference in school completion levels across countries.

If parents are unable to borrow to finance investments in education, all of the above forces still apply, but households' decisions are going to be constrained by their own resources. The Becker-Tomes (1976) model suggests that the amount poor parents are willing to invest in the child's education will be smaller than the optimum under a non-binding borrowing constraint, but this amount will be increasing in parental income up to the point at which the marginal return to investment is equal to the rate of interest. When there is more than one child at home, parents also decide the allocation of resources across children (Behrman, Pollack, Taubman, 1991).

The literature on the schooling of children in developing countries has emphasized the role of family income constraints in explaining differences in school attainment. The empirical work has found that family income, parental schooling, sex, residence, and family size are important determinants of school attainment. Behrman and Wolfe (1984) use Nicaraguan data to show that family background (place of birth) and parental schooling, particularly the mother's schooling, are important determinants of children's school attainment. Parish and Willis (1993) used Taiwanese data and found that the effects of parental income, and the size and composition of the family were strongest among older cohorts and poorer households. They interpret this evidence as indicating that older cohorts and poorer households are more credit constrained than younger cohorts and richer households. Lillard and Willis (1994) used Malaysian data to study the effects of parents' and grandparents' schooling on children's progress through the educational system. Their data allows them to identify separately the effect of parental schooling from the effect of other time-varying variables on children's school attainment. They find that at least two-thirds of the impact of parental education on children's schooling is a direct consequence of parent schooling while the remaining one-third can be attributed to unmeasured factors that influence educational attainment of parents and children.

In the absence of longitudinal data, there are two problems when attempting to isolate the impact of genetic ability and the household budget constraint. First, parental schooling may capture simultaneously genetic ability, motivation, and the capacity to generate income. Second, current income is a poor measure of the household budget constraint, particularly in the context of important variability in current income relative to permanent income, leading to a downward bias in the estimated effect of household income in children's schooling (Haveman and Wolfe 1995). This ambiguity is problematic for it weakens the support for income transfers as a means to improve opportunities for children whose parents lack schooling. Studies that shed light on the marginal effects of the parental income versus schooling variables are especially valuable.

We examine a cross section of households in 1997, measure the school attainment of individuals aged 6 to 24, and estimate the impact of household characteristics on the hazard of dropping out of school. Included in the household characteristics we measure are local indicators, parental schooling, income and remittances (where the last two are measured at the time of the survey). The 1980's out migration wave from El Salvador was primarily motivated by political factors. The peace accord was signed in 1992, and migration flows after 1992 are more likely to be explained by economic factors, including the presence of a migrant population in the United States. Therefore, our observed sample may include households receiving remittances because they have a one-time political exile relative or friend living abroad, and households receiving remittances because they sent a family member to work and remit income from abroad. Remittances received by the first type of household are arguably equivalent to a randomly assigned income transfer, while remittances received by the second type are not.

If remittances are randomly assigned across households, our analysis provides an unusual opportunity to identify separately the effect of income from that of parental schooling. If migration has been driven primarily by economic reasons and remittances cannot be viewed as the result of random assignment, our study still makes the important contribution of focusing attention on the effect of remittances on human capital formation and quantifies the effect on school retention rates throughout primary and secondary schooling. To our knowledge, this is the first study to examine the effects of remittances on an important household decision using micro-level data.

We argue that remittances are closer to a randomly assigned transfer, and that their measured effect is a cleaner estimate of the impact of relaxing a household's budget constraint on school retention rates than is the measured effect of household income—which is typically correlated with parental schooling. We focus on three characteristics of households: family income, parental schooling, and the number of school-age children. We separate rural and urban households because of the significant differences in income and schooling across the rural-urban divide. Comparing the location of recipient and non-recipient households along the distribution of these variables, we show

that all groups are likely to receive remittances. In addition, we report the median amount of remittance and show that it is remarkably stable across the various groups.

□

### **Household Characteristics and Remittances**

□

Table 1 reports the percentage of households receiving remittances that fall in each decile of the income distribution of non-recipient households, separately for urban and rural areas. Non-recipient households represent 85% of urban and 86% of rural households. Column 2 in each panel of Table 1 reports the percentage of recipient households that fall in each decile when we measure pre-remittance income only. In both regions, close to 30% of households fall in the first decile of the distribution, and the remaining households are roughly evenly distributed throughout the other 9 deciles. The over-representation of households in the lowest decile of the income distribution does not imply that recipient households are significantly more likely to be poor households. These households have fewer members employed than non-recipient households so their income-generating capacity in the local economy must be lower than it is for non-recipient households.<sup>3</sup> Column 3 in each panel shows the median remittance amount received by households in each decile. Clearly, the remittance amount is independent of the pre-remittance income of the recipient households: the within-decile median amount is typically 875 colones (100 U.S. dollars) or half that amount. Of course, the addition of a remittance to household income changes the location of recipient households along the income distribution of non recipient households. Column 4 in each panel reports the percentage of recipient households that fall in each decile of the income distribution of non-recipient households by post-remittance income. In urban areas, the percentage of recipient households that fall in each decile is remarkably close to 10 percent, suggesting that households that have relatives abroad who send remittances come from every segment of the distribution of income. There is no evidence of a pattern of self-selection where poor households are predominantly likely to receive funds from relatives living abroad.

---

<sup>3</sup> In urban areas, the average recipient household has 1.35 workers against 1.7 in non-recipient household. In rural areas, the corresponding averages are 1.24 and 1.66.



□

In rural areas, recipient households are slightly under-represented in the three lowest deciles of the income distribution. Note that the median remittance amount is the same in rural and in urban areas because it is largely determined by the expatriates' earnings in the United States. But since incomes are considerably lower in rural than in urban areas, remittances have a more pronounced effect on rural household income that shows up in an under-representation of recipient households in the three lowest deciles, and a slight over-representation in the top two deciles. But the important point to note is that we find no pattern of self-selection where recipient families are predominantly likely to come from one segment of the distribution of household income.

□

Table 2 shows that 28% of all households in El Salvador are female-headed. While recipient households are more likely to be female-headed, 25% of non-recipient households are also female-headed. An alternative measure of this relationship (not shown) is that 23% of female-headed households receive remittances against 11% of male-headed households. Of course, political exiles and those emigrating for economic reasons are typically male, so the higher incidence of female-headed households among recipient households is unsurprising and sheds no light on the issue of which type of migration dominates the data.

□

Table 3 reports the distribution of urban and rural households by the schooling of the adult female in the household—head of household or spouse. The choice of the female adult to represent each household is deliberate, and it allows a cleaner comparison across households that may or may not have an adult male present. Migration patterns in El Salvador have resulted in about 30% female-headed households, and female headship is more common among recipient households. Therefore, a comparison of parental characteristics between recipient and non-recipient households is better done controlling for adult female characteristics. The urban and rural distributions of households across adult female schooling levels are fairly close for non-recipient and recipient households, though recipient households are somewhat over-represented in the lowest schooling category, 0 to 3 years of schooling. If political exiles came from all walks of life in El

Salvador and the exiles represent the lion's share of those working abroad, the educational distribution of the wives left behind would be similar to that of wives in non-recipient households. The fact that recipient households are somewhat over-represented in the lowest schooling categories suggests that migrants for economic reasons may have represented an ever increasing fraction of migration flows in the five years leading up to the 1997 survey (assuming that they come from the low end of the educational distribution and that assortative mating dominates).

□

Table 4 organizes urban and rural households according to the number of school age children (children aged 6 and older). We find no systematic differences in the distribution of recipient and non-recipient households by the presence of school-age children. Finally, note that the median value of the remittance amount is 875 colones (US\$100) in most cells in all three tables.<sup>4</sup>

□

### **Schooling in El Salvador**

□

The educational system in El Salvador is organized along four levels. Children aged 4 to 6 attend pre-school. Children aged 7 to 15 attend primary education which is divided in three cycles of three years each, from first to ninth grade. Children aged 16 to 18 attend secondary education which goes from tenth to twelfth grade. Table 5, based on 1997 data, shows the enrollment of children in the school system, by age categories.

□

---

<sup>4</sup> The survey collects information on remittances, their amount and their use. Respondents typically quote the remittances received in dollars, and these amounts are translated into local currency (colones) by the interviewers, using the official exchange rate. We could have also considered using the information collected on expenditures and in particular, the amounts of remittances that are said to have been spent on education. Nevertheless, we decided not to use the latter variable for two reasons. First, there is a significant discrepancy between the total amount of registered remittances and the total expenditure that they afford, with total expenditure being about 30 percent higher than the recorded amount received. We believe that expenditures are more likely to be measured with error, since there are several items involved. Second, the amount of remittances spent on education says nothing about the fact that remittances spent on other items will free resources that may be ultimately spent on education. We therefore preferred to stay within the tradition of economic analysis and keep tabs on what families do about education rather than on the amounts they say they spend towards education.

There has been a rapid expansion in school enrollment throughout the 1990's (Table 6). Note the significant increases in rural pre-school and primary coverage, and urban pre-school and secondary coverage. The levels of school attainment of recent birth cohorts in El Salvador are significantly higher than those of older cohorts. This gain is the outcome of an important national effort, considering current standards of living in El Salvador. Table 7 reports school attainment levels for older individuals. We see that no more than 40 percent of the parents of current school-aged children completed primary education.

Enrollment levels have yet to reach the equivalent of "universal primary coverage:" of the young aged 12, only 80 percent of rural residents and 90 percent of urban residents are enrolled in primary school. As Table 5 suggests, rates of retention in primary education are relatively high, but there is slow progress through grades. Once children get to the third cycle of primary schooling, 45 percent are in the grade corresponding to normal progress, 30.5 percent are behind, and 19 percent have abandoned the system (see MINED, 1997).

The survey records answers to the question "Why is (a particular person) not attending school?" Among primary school age children, the most frequent answers are "too expensive" and "age related reason." The first one suggests that the family's budget constraint has a major impact on the decision to drop out of school. The second answer hints at the role played by the child's ability and/or the quality of schools, or may be an indication that the child never enrolled or started school late.

#### *Demand Driven Expansion*

El Salvador's Community-Managed Schools Program (popularly known by the acronym, EDUCO, or *Educación con Participación de la Comunidad*) is an innovative program of public support for community managed schools. A prototype of today's EDUCO schools emerged in the 1980's when the civil war left a void in public education. Some communities took the initiative to organize their own schools, administered and financially supported by a family-based association. While these early attempts were constrained by the low rural income base, they revealed a strong latent demand for

education, as well as a desire to participate in the governance of schools. In 1991, El Salvador's Ministry of Education (MINED), with the support of international agencies, chose the implementation of the same prototype as the main method of expanding educational coverage in rural areas.

The present EDUCO schools are managed autonomously by an elected Community Education Association (Asociación Comunal para la Educación or ACE) drawn from the parents of the students. In EDUCO schools, the ACE performs a central role of administration and management: ACEs are contracted by MINED to deliver a given curriculum to an agreed upon number of students. ACEs are charged with the close monitoring of teachers' performance, hiring and firing of teachers, and equipping and maintaining the schools.<sup>5</sup>

### **Data and Methodology**

The 1997 National Household Survey (Encuesta de Hogares de Propósitos Múltiples EHPM) has national representation and covers 8,387 families. We focus on the schooling attendance and family characteristics of individuals aged 6 to 24, about 40 percent of the population. Based on the survey data we can describe a number of important characteristics of each child's family. We have information on housing conditions, the age, schooling, income and labor force status of family members, the number of household members who are currently living abroad, and the value of remittances sent by expatriates.

Our hypothesis is that parents make schooling decisions for their children on the basis of expectations of the returns on these investments. Factors that influence the expected

---

<sup>5</sup> The EDUCO program has become a very productive model for the modernization for the entire Ministry of Education (Jacir de Lovo, 1997). Jimenez and Sawada (1998) compared student achievement on standardized tests and school attendance of rural students in EDUCO schools with the achievement and attendance of students in traditional schools. Their analysis controls for student characteristics and selection bias, using an exogenously determined formula for targeting EDUCO schools as an instrumental variable. They find that the rapid expansion of rural schools through EDUCO has diminished student absences and has not had an adverse effect on student achievement.

return to schooling are the cost of school attendance—out-of-pocket and opportunity costs—and the anticipated rewards in the labor market for skills acquired through formal schooling. Parents will compare the expected returns from investing in the schooling of their children to the returns from alternative investments that will also increase a child's human capital, like nutrition, health care, clothing and shelter. Poor families may find it optimal to choose quite low levels of completed schooling for their children in order to enjoy a given level of consumption of food, health services, etc. Low levels of completed schooling imply relatively low out-of-pocket expenses for schooling and enable the child to start working at an early age (the legal working age in El Salvador is 10 years of age). That is, in a poor country like El Salvador, the family's budget constraint may play an important role in a family's decision on school attendance of their children.<sup>6</sup>

Parents' expectations of the rewards to skill in the labor market may depend on the parents' own schooling levels. More educated parents may be better informed about the employment opportunities and wages available to those with given schooling levels, or may themselves enjoy the rewards the labor market offers workers with above average education. Therefore, parental schooling may play a role in a family's choice of completed schooling for their children.

Other factors that affect the perceived costs and benefits of schooling relate to the availability of schools offering the required grade levels within a reasonable commuting distance from the household. We do not know the exact location of each household or the geographical distribution of schools offering given grade levels. We do know, however, if a household resides in a rural area—where we expect less access to schools—and if the dwelling has access to neither electricity nor water, suggesting the household is located in a particularly remote area, or is an extremely poor household.

A model of parental schooling decisions for their children would not be complete without reference to another two factors. First, it is reasonable to expect that parents' decisions on the schooling of older children are made sequentially, taking into account the child's school experience. When a child is retained in grade, the return to continued investment in the child's schooling is instantly reduced. Retention in grade raises the cost of completing the same grade level with no change in benefits. As we discuss below, we have access to a single cross-section for El Salvador. The survey has very good information on school related behavior, but we do not know at what age an individual started school or if a person—in or out of school—was ever retained in grade. While we would like to control for the age of children and young adults in a model of the probability of their continued school attendance, we cannot do it.

The second factor that likely plays a role in parental decisions regarding a child's schooling is the composition of the household. The number and ages of siblings bears heavily on the family's resources. In order to capture fully the budget constraint of a family, income is clearly not enough. We know the composition of each household at the time of the survey, but we do not know it at any other time. In particular, we do not know it at the various stages in a child's life when the composition of the household may have played an important role in determining whether the child would continue on to the next cycle of education, for example the second cycle (grades 4 through 6).

### *The Statistical Model*

Our sample is a cross-section of 14,286 individuals aged 6 to 24. We know the sex, age and school attainment of the individual, characteristics of the family, place of residence, and properties of the dwelling where they live. We assume that children of school age who are not enrolled in school have dropped out of school, and use the Cox proportional hazard model to estimate the impact of characteristics of the individual and the family on the hazard of dropping out of school.

---

<sup>6</sup> Compulsory free schooling is a way to overrule parental doubts on the potential benefits of education. Nevertheless, in El Salvador as in most developing countries, compulsory primary education is not enforced, and schooling is not free or readily available to all children.

□

The choice of the proportional hazard model is somewhat unusual in the modeling of school attainment levels, and warrants some discussion. The Cox proportional hazard model is tailored for the analysis of survival-time data, and, clearly, the number of “years of completed schooling” does not correspond one-to-one with calendar time because children can be retained in grade. The question, then, is whether the proportional hazard framework is acceptable for the analysis of school attainment. We will argue it is.

□

The hazard framework requires that we choose one of two possible “outcomes” for each individual in the sample: the individual is enrolled in school (“right-censored”) or is not enrolled (“failed”). We have a single cross-section and no retrospective information on enrollment behavior. Thus, we have no choice but to treat every individual who is not enrolled in school on the day of the survey as if they have “failed.”

□

The proportional hazard model adapted to the case in question assumes that the observed fraction of the population that dropped out after grade  $t$ , relative to those that completed grade  $t$ , is:

□

$$h(t) = h_0(t) * \exp\{x_i'\beta\}$$

□

where  $h_0(t)$  is the baseline hazard of leaving school after grade  $t$ , which is left unspecified and is estimated;  $x_i'$  is a vector of covariates, and  $\beta$  is the vector of parameters to be estimated. The crucial assumption in the Cox proportional hazard model is that the effect of the covariates is proportional over the entire base line.

□

The Cox proportional hazard model has several features that make it an attractive statistical framework for the problem at hand. The most obvious advantage is that it exploits all the available information in observations that are right-censored, that is, observations for children or young adults who are enrolled in school at the time of the survey (9,547 of the 14,286 in our sample). We can avoid the unappealing practice of limiting the sample to individuals aged 25 and over who are expected to have finished

their formal education. In the case of El Salvador, this practice would be especially undesirable since it is only recently that the government has engaged in a bold effort to expand grade availability in rural areas and increase school enrollment generally throughout the country. A focus on older birth cohorts would give us no insight into recent educational trends in El Salvador.

□

Another desirable feature of the Cox proportional hazard model is that it readily yields an estimate of the underlying baseline hazard function, enabling us to identify the grade levels where dropout rates are concentrated, net of the effect of measured determinants of school completion. This information may be useful for policy makers seeking to further expand school enrollment.

□

In the empirical analysis, we include everyone in the sample aged 6 to 24, who does not have missing information. Note that the sample *includes* individuals who have never attended school. This highlights another convenient feature of the Cox proportional hazard approach. Since time, in and of itself, is not of the essence for the analysis, we model the determinants of “never enrolling in school” much like we model the decisions regarding continuing to attend school. That is, we view “never enrolled” as the first stage of the schooling process. Everyone in the sample “enters the study” (in the jargon of biostatistics) through this stage which has an arbitrarily chosen length of time. Then, there is a probability that a child will not “fail” and continues on to enroll in the first grade. Put differently, the hazard of “failing” in the “not enrolled” interval simply refers to the hazard of never enrolling in school at all.<sup>7</sup>

□

Some of the older individuals in the sample have completed more than 12 years of schooling. There are few such cases and we are primarily interested in school attainment in primary and secondary school. So we focus our attention on grades 1 through 12 by

\_\_\_\_\_

<sup>7</sup> The distinction between failure to enroll and drop-out hazard has generally been overlooked. A notable exception is Filmer and Pritchett (2000). Interestingly enough, their results show that the enrollment profiles of the poor differ across countries but fall into distinctive regional patterns: in some regions the poor reach nearly universal enrollment in first grade, but then drop out in droves leading to low attainment



truncating an individual's completed schooling at 12 years if he or she has more than 12 years of completed schooling, and treat the observation as right censored.

□

Note that we are examining a cross section with individuals of different ages, some currently in school and some who dropped out of school in the past. We have a single cross section, and find ourselves in the unenviable position of having to limit the vector of covariates to factors that we can argue convincingly can be treated as time-invariant. Some determinants of the parents' schooling choices for their children are clearly grade-level-invariant, for example the child's sex and the schooling levels of the parents. Our empirical specification also includes an indicator variable equal to one if the household has no access to water or electricity, and a set of variables that measure income and the presence and amount of remittances received from abroad.

□

#### *Operational Definitions of the Covariates*

Before turning to details of the specification we estimate, we list the set of covariates and give their operational definitions. The covariates include an indicator variable for sex (equal to 1 if male). We control for a child's sex to allow for the possibility that parents expect different returns to investment in schooling for boys and girls, all else equal.

□

We include an indicator variable for a household's lack of access to water and electricity (equal to 1 if there is no access) to serve as a proxy for local conditions. Presumably, lack of access to these basic services is correlated with low population density and long distances to schools, factors that increase the cost of attending school, and lower the returns. While we realize that access to basic services is not strictly time invariant, we expect it to be nearly so and include it in our vector of covariates. Even though internal migration is common, it is far less common for an entire family to change residence, especially in the rural areas where families typically own the piece of land where they

---

---

(typical of South America), while in other regions the poor never enroll in school (typical of South Asia and Western/Central Africa).

reside. Also, internal migration is more likely among young adults that are starting a family.<sup>8</sup>

Parental schooling has been found to be significant in explaining the school attainment of children. Often times, the effect of father's schooling is different from that of mother's schooling. We chose not to test for differences in the effect of paternal and maternal schooling levels because in El Salvador close to 30 percent of households have absent fathers. An attempt to test for the separate effects of paternal and maternal schooling results in a loss of too many observations, and runs the serious risk of yielding a highly selected, non-representative sample of households. We created a variable, "Parental Schooling," equal to the higher of the two parents' schooling, if both are present, or the schooling of the parent who is present when one is absent.

The remaining three covariates pertain to household income and the presence and amount of remittances. We include the survey variable "household income," measured in colones per month, to capture the budget constraint of the child's household. As other studies have found, we expect that higher incomes will lower the hazard of leaving school. Two questions are of interest. First, does household income influence school attainment? Second, does income from remittances have the same effect as household income from work in the local economy? To answer the last question we exclude the remittance amount from the measure of household income and include it as a separate covariate.<sup>9</sup>

---

<sup>8</sup> The 1997 Household Survey we use does not have information on internal migration. Issues related to migrant families can be examined with the 1998 Household Survey that includes a migration module.

<sup>9</sup> The standard of living in urban areas of El Salvador is significantly higher than in rural areas, so it is unclear how to measure correctly the impact of an additional colon from remittances on a household's budget constraint. We experimented with an alternative metric for household income and remittances. In the alternative specification, the variable "Location in the income distribution (net of remittance)" is measured in integers ranging from 1 to 20 that correspond to the location (in intervals of 5 centiles) of a given income level in the income distribution, excluding remittances. Thus "Location in the income distribution (net of remittance)" will be equal to 1 if the household income is in the bottom 5 centiles of the income distribution for the area (urban or rural) where it resides. We then added to each recipient household's income, the income from remittances. The location on the income distribution will tend to worsen for those without remittances, and will improve for those with remittances. The change in location is measured in units of 5 centile points and ranges from 0 (for no change) to 19 (for a change from the lowest to the highest 5 centiles in the distribution). "Change in location due to remittance" will be equal to

Income from remittances will also expand the family's budget constraint, allowing some families to move closer to the optimum amount of schooling for their children. We note that the typical amount of remittances is 100 dollars, or 875 colones. As Table 8 shows, the median remittance represents a more significant fraction of incomes of rural than urban families: remittances represent 46 percent of household incomes for the median rural recipient and 27 percent for the median urban recipient household. On average, remittances make up 49 percent and 37 percent of income for rural and urban recipient households.

In addition to controlling for household income and income from remittances, we include an indicator for the presence of income from remittances (equal to 1 if the household receives remittances from abroad). We believe that the massive out-migration in El Salvador during the 1980's was driven primarily by political rather than economic factors. But we want to allow for the possibility that households that receive remittances differ systematically, for unobserved reasons, from other households. In fact, even if the initial stimulus for migrating was political, families receiving remittances have already experienced direct benefits from migration. One can argue that these families attach a higher value to the migration option for their own children than do other families. This effect alone may well cause differences across families in the expected returns from schooling, and in the hazard of leaving school. More generally, the indicator variable will capture any additional effect of remittances on children's schooling that acts through channels other than the budget constraint, and any systematic differences in attitudes

---

---

2 if the family's income *including the remittance* falls 2 intervals to the right of its initial location in the income distribution, that is, when excluding the amount of the remittance.

The hazard model that uses the metric described above yields a slightly higher value for the log-likelihood function for the urban sample, and a slightly lower value for the rural sample. Regarding the remaining coefficient estimates, they are quite invariant to the choice of metric for income and remittances. We implemented a test proposed by Weesie (1999) that is a generalization of Hausman's test, of the null hypothesis that the remaining coefficients are equal in the two specifications. (Note that this is not a nested hypothesis.) For the urban sample the test-statistic is equal to 12.41 and is distributed chi-squared with 6 degrees of freedom. We fail to reject the null hypothesis; the p-value for the test is 0.0534. For the rural sample the test-statistic is equal to 13.08 and is distributed chi-squared with 8 degrees of freedom. We fail to reject the null hypothesis; the p-value for the test is 0.1090.

toward the schooling of children across families that do and those that do not receive remittances.

□

It should be noted that the appropriate variable that influences investment on children is permanent household income. What we use is a crude proxy of the variable, which is household income in one particular year separated in two components, remittances and the rest. Arguing that income and the presence of remittances are time-invariant covariates is perhaps a bit heroic. While it is probably the case that income mobility in El Salvador is quite low, a question arises with respect to the stability through time of remittances vs. other income. If remittances are more stable, they are a better measure of permanent income relative to other observed income. We return to these issues as we examine the results.

□

There are two variables we intended to use as covariates, “year of birth” and an indicator variable for residence in the rural areas. The reason for controlling for year of birth is simply to pick up any systematic differences between cohorts. The differences could be due to the implementation of educational reform by the government or any effects that the civil war of the 1980’s had on the operation of the schools or children’s ability to attend school. For reasons discussed in detail in the Appendix, our preferred specification is one where we stratify on year of birth and we estimate separate models for the urban and rural samples.

□

We estimate separate models for the urban and rural samples allowing separate effects for the four segments of the baseline hazard, and we stratify each sample on 5 strata according to year of birth. The segments are “never enrolled,” 1<sup>st</sup> through the 6<sup>th</sup> grade, 7<sup>th</sup> through the 9<sup>th</sup> grade, and 10<sup>th</sup> through the 12<sup>th</sup> grade.<sup>10</sup> The four oldest strata

---

<sup>10</sup>We divided the baseline hazard into the segments listed in text based on the following considerations. Primary education is divided into three cycles of three grades each. Therefore, when distance to school becomes a problem, it is likely it will happen when a child is ready to enroll in the next cycle of primary education or secondary education. In addition, public education becomes relatively scarce at the third cycle of primary education, and close to half of secondary education is private. This suggests that out-of-pocket costs of additional years of schooling beyond the sixth grade rise relative to schooling at younger ages. In addition, in the early stages of estimation we divided the hazard further, allowing for separate

combine individuals born in four-year intervals: 1972 to 1975, 1976 to 1979, 1980 to 1983, and 1984 to 1987. The youngest stratum is for individuals born between 1988 and 1990.

□

### **Estimation Results**

□

To provide a context for the estimation results, we begin by presenting Kaplan-Meier estimates of the survival functions in Figure 1. It is immediately apparent that children living in urban areas have a significantly higher probability of attaining any given level of school completion than rural children do. For example, rural children have a probability of .60 of completing the 6<sup>th</sup> grade, compared to over .85 for urban children. The gap only widens at higher-grade levels. Urban children have a probability of .70 of completing the 12<sup>th</sup> grade, compared to only .30 for rural children. The graphed functions also highlight the more common exit points from schooling for children. The survival functions show large declines at the 3<sup>rd</sup>, 6<sup>th</sup>, and 9<sup>th</sup> grades, especially in rural areas.

□

Of the individuals in our estimation sample (individuals aged 6 to 24 who have no missing values), 55 percent of those in urban areas live in households that have 2 or more children in the sample. The corresponding figure is 65 percent in the rural sample. To allow for correlation in the “residuals” of family members, that is, to allow for a household effect, we compute robust estimates of variance by clustering observations on a family identifier.

□

The estimates of the determinants of the hazard of dropping out of school are summarized in Tables 9 and 10. The (exponentiated) coefficient for the indicator variable “Male” is significantly different from one in urban areas, where males are 27 percent more likely to drop out of school (or never enroll in school at all). One possible

---

---

effects for the 1<sup>st</sup> through the 3<sup>rd</sup> grade, and the 4<sup>th</sup> through the 6<sup>th</sup> grade. But we did not reject the null hypothesis that the segments could be combined into one covering the 1<sup>st</sup> through the 6<sup>th</sup> grade.

explanation for this finding is that boys have better labor market opportunities than do girls, and thus face higher opportunity costs of attending school.

□

In rural areas, the effect of sex is not proportional over the entire baseline hazard. Boys have a higher hazard than do girls of never enrolling in school and of leaving school after the first two cycles of primary education. The estimated hazard ratio is 1.16, or a 16 percent higher hazard. This effect is qualitatively similar to the estimated effect of sex in urban areas. What is puzzling is that for grades 1 through 6, the hazard of leaving school is 10 percent *lower* for boys than girls in rural areas. The coefficients in question are statistically significantly different from one another, and we do not have an explanation for this pattern of effects.

□

Interestingly, the dampening effect of higher parental schooling on the hazard of leaving school declines monotonically as a child reaches ever higher-grade levels. In urban areas, each year of additional parental schooling lowers the hazard that a child leaves school (or never enrolls) by 20 percent, while the child is attending one of the first two cycles of primary education. Once the child is enrolled in the third cycle, 7<sup>th</sup> through the 9<sup>th</sup> grade, an additional year of parental schooling only lowers the hazard by 14 percent. And the effect on the hazard for secondary education is lower still: 7 percent. The estimated coefficients are statistically significantly different for the three segments of the hazard.

□

In rural areas much the same pattern is observed, though the estimated effects of parental schooling on the hazard of leaving school are smaller than they are in urban areas. An additional year of parental schooling has the largest estimated effect on the hazard that a child will never be enrolled in school: the hazard drops 16 percent with each additional year of parental schooling.<sup>11</sup> For children attending grades 1 through 6, an additional

---

<sup>11</sup> This may appear to be too large of an effect. But average parental schooling in rural areas is close to three years of schooling. Thus, the child of a parent with 3 years of schooling has a hazard of never enrolling in school that is only 59 percent ( $\exp(-.177 \times 3) = .588$ ) as high as the hazard for a child whose parents have no schooling.

year of parental schooling lowers the hazard of leaving school by 13 percent. For youngsters attending grades 7 through 12, the effect drops to 8 percent.

□

Despite the previous discussion, we cannot conclude that parental schooling is a weightier factor for children's schooling in urban than in rural areas. As seen in Figure 1, the survival function is significantly higher in urban than rural areas; correspondingly, the hazard function is significantly lower in urban than rural areas. The combined effects of the coefficient estimates for parental schooling and the level of the baseline hazard functions is illustrated in Figure 2. We graph the estimated survival functions evaluated at the 10<sup>th</sup> centile, the 1<sup>st</sup> quartile, the median, the 3<sup>rd</sup> quartile and the 90<sup>th</sup> centile of the distribution of parental schooling, separately for urban and rural areas. We set income at the median for each area, and graph the estimated baseline survival functions for boys belonging to the 1976-1979 birth cohorts (all other covariates are set to zero). The graph for urban areas appears to have only 4 survival functions. The reason for this is that the survival functions for the third quartile and the 90<sup>th</sup> centile overlap completely. Clearly, the effect of parental schooling on the hazard of leaving school in urban areas is concentrated in the bottom half of the parental schooling distribution. Children whose parents have above average schooling levels have survival probabilities close to 1, when evaluated at the median household income.

□

The graph for rural areas has 4 survival functions because the 10<sup>th</sup> decile and the 1<sup>st</sup> quartile of the distribution of parental schooling is the same: 0 years of schooling. Recall that the estimated coefficients of parental schooling are smaller for rural parents, but the hazard of leaving school is considerably higher in rural areas. As a result, the overall effect of parental schooling in rural areas is quite important, and operates over the entire distribution of parental schooling.

□

The coefficient on lack of access to water or electricity is significant in both regions and much larger in urban areas. Lacking access to basic services increases the hazard of leaving school (or never enrolling) by 150 percent in urban areas. In rural areas, lacking access to basic services increases the hazard of never enrolling in school by 86 percent,

the hazard of leaving school while enrolled in the 1<sup>st</sup> through the 6<sup>th</sup> grade by 33 percent, and there is no effect on the hazard beyond the 6<sup>th</sup> grade. While 38 percent of rural households lack water and electricity, only 5 percent of urban families lack water and electricity. Therefore, this large difference in the estimated coefficients only applies to the 5 percent of urban families that live in extremely poor conditions.

□

Household income has a proportional effect over the entire baseline hazard for the urban areas. This suggests that the household's budget constraint plays a significant role in the choice of schooling level for children, and the effect does not diminish as the children attain ever-higher grade levels. The (exponentiated) estimated coefficient in urban areas is .996, indicating that an increase in household income of 100 colones lowers the hazard that a child will never enroll in school or will leave school once enrolled by 0.4 percent. A movement from the first to the third quartile of the urban income distribution (1,517 to 4,583 colones) lowers the hazard by 12 percent. The estimated effect of income is weaker in rural areas. Household income has no effect on the hazard of leaving school beyond the 6<sup>th</sup> grade. Yet, household income lowers the hazard that a child will never enroll in school or will leave school before reaching the 6<sup>th</sup> grade. The point estimate is .993, indicating that a movement from the first to the third quartile of the rural income distribution (750 to 2,225 colones) lowers the hazard by 10 percent.

□

To assess the practical importance, as opposed to the statistical significance, of household income we graph the baseline survival functions for boys in the 1976-79 birth cohorts. The functions, presented in Figure 3, are evaluated at the 10<sup>th</sup> centile, 1<sup>st</sup> quartile, median, 3<sup>rd</sup> quartile, and 90<sup>th</sup> centile of the distribution of household income, separately for urban and rural areas. Also, the functions correspond to median parental schooling: 6 years in urban areas, and 2 years in rural areas. Figure 3 makes it abundantly clear that household income has a very small effect on the survival functions, even though the effect is more pronounced for rural households. We also examined the survival functions evaluated at the 1<sup>st</sup> quartile of the distribution of parental schooling. The survival functions are of course lower than those graphed in Figure 3, but they convey the same message:



household income has a very modest effect on the hazard of leaving school, regardless of the level of parental schooling.

□

This result is consistent with previous literature. Haveman and Wolfe (1995) review a number of studies on the determinants of years of schooling based on U.S. data and report that income elasticities have been found to be low and estimated in a wide range, from .02 to .2. They believe this result is likely explained by the measurement error in the household income variable. At the same time, Hill and Duncan (1987) carefully measure household income and report that a 10 percent increase in household income, after controlling for a number of other variables, is associated with an increase in school attainment of less than 1 percent.

□

Now we turn our attention to remittances. As we explained earlier, we estimate separate effects for income from remittances and from other sources to examine the possibility that the source of the income matters for the choice of schooling levels of children. Economic theory suggests that the source of income should not matter—a dollar is a dollar.

□

We estimate a significant and large effect of income from remittances. In urban areas, the median remittance (875 colones) lowers the hazard that a child will never enroll in school, or will leave school while enrolled in the 1<sup>st</sup> through the 6<sup>th</sup> grade, by 54 percent ( $\exp(-.0877 \times 8.75) = .46$ ). The estimated effect for the remaining segments of the hazard, grades 7<sup>th</sup> through 12<sup>th</sup>, is .965, or 27 percent when evaluated at the median remittance amount. So the effect is half as high for children beyond the 6<sup>th</sup> grade as for children below the 7<sup>th</sup> grade. Note that the estimated coefficients for the two income effects are statistically significantly different from one another. More importantly, there is an enormous difference in the magnitude of the effects depending on the income source: the effect of remittances is, at its smallest, 10 times the size of the effect of other income.<sup>12</sup>

---

<sup>12</sup> We estimated a hazard model where the effect of income was allowed to differ between households that do and do not receive remittances. For the urban sample, the test statistic for the null hypothesis that the

□

In rural areas, the median remittance lowers the entire baseline hazard by 14 percent. That is, the hazard of a child leaving school (or failing to enroll) is 14 percent ( $\exp(-.0171 \times 8.75) = .86$ ) lower if the child's family receives the median remittance amount. The estimated coefficients for the two income effects are statistically significantly different from one another, and the magnitudes of the effects are not close: the effect of remittances is 2.6 times higher than the effect of other income.

□

If income from remittances is more stable than other sources of income, a likely situation for rural households, income from remittances is a better proxy for permanent income than is other income and this may explain the difference in estimated effects.<sup>13</sup> Another possible explanation for the difference is that remittances recipients may exhibit a higher propensity to spend on their children's schooling out of remitted funds than other funds, perhaps because the expatriate family member has made it a condition for the financial support. In any case, whether the source of the income makes a difference is an empirical issue, regardless of our ability to tell a story that accounts for the observed behavior.

□

The presence of remittances in the urban areas has no effect on the hazard of leaving school, *after controlling for the amount of the remittance*, so the variable "Receives remittance" does not appear among the covariates. This means that urban households that receive remittances from abroad do not differ systematically from other urban households with regards to the schooling of their children, though the funds from remittances have a vastly larger effect on the hazard of leaving school than funds from other sources.

□

---

effect of income from other sources is equal for both types of households is equal to 0.11, with one degree of freedom, and a p-value of 0.742. Consequently, we fail to reject the null hypothesis of same effect. The same result applies for rural households, in which case the test statistic is equal to 1.58, with a p-value equal to 0.209.

<sup>13</sup> Edwards (2000) reports that the fraction of households that receive remittances has been remarkably stable –between 14.4 and 15.6 percent– between 1992 and 1997. This is the short time period for which remittances income is measured in the national household survey that measures living standards.

The presence of remittances in rural households has a large effect on the hazard of leaving school. Children in those households have a 24 percent lower hazard. That is, in rural areas the source of income matters less than in urban areas, but the mere existence of remittances as an income source tends to reduce the hazard of leaving school. Rural families receiving remittances appear, then, to differ systematically from other rural families in that, all else constant, their children show a lower hazard of leaving school, at all grade levels.

Overall, the effect on the hazard of leaving school of receiving a remittance of 100 U.S. dollars (the median amount in both areas) is greater for children in urban than in rural households. In urban areas a remittance of 875 colones (100 dollars) lowers the hazard by 54 percent for children below the 7<sup>th</sup> grade, and 27 percent for youngsters beyond the 6<sup>th</sup> grade. In rural areas, the same remittance amount, combined with the effect of the presence of the remittance, lowers the hazard by 25 percent at all grade levels. To gauge whether the effect of remittances is of practical importance, in Figure 4 we graph the baseline survival functions for boys in the 1976-79 birth cohorts, evaluated at no remittance and at the median remittance of 875 colones. The functions are evaluated at the median of parental schooling and household income, separately for urban and rural areas. Note that in the case of rural children, the survival function showing the effect of the median remittance includes the effect of the presence of the remittance itself.

In urban areas, a remittance of 875 colones has a small effect on the baseline survival function. But note that the survival function evaluated at the median parental schooling and median household income is very high at all grade levels. There isn't much room for the remittance to have an effect. And yet, it has the effect of raising the survival function to near 1. In rural areas, the effect of a remittance in the median amount is substantial and operates at all grade levels.

#### *Cohort Trends*

The estimation of the Cox proportional hazard model stratified by birth cohort yields estimates of the underlying baseline hazard and survival function for each birth cohort.

We report the survival functions corresponding to our preferred specification in Tables 11 and 12, and plot the baseline hazards and the survival functions for the four earliest birth cohorts in Figures 5 through 8. Following the same procedure used for the previous figures, we graph baseline functions where parental schooling and household income are set to their median values for each area. The survival and hazard functions, then, apply to boys living in dwellings with access to water or electricity and whose household does not receive a remittance from abroad.

□

The first row of figures in Tables 11 and 12, labeled Grade Level 0, present the value of the survival function for the outcome “enrolls in school.” That is, the value .971 (oldest birth cohort in the urban areas) implies that a child has a .971 probability of enrolling in the first grade. It also implies that, for that birth cohort, we estimate that .029, or 2.9 percent, of that cohort never enrolled in school.

□

The survival functions for urban and rural children born in 1988 through 1990 are lower than the functions for earlier cohorts. This does not mean that grade attainment levels are declining for recent birth cohorts. Recall that we include everyone between the ages of 6 and 24 in the sample. A high fraction of children aged 6 and 7 are not enrolled in school in 1997. This is not a surprise. The distribution of ages for children enrolled in the first grade shows many children aged 7 and 8 and even older children. We observe the children born in 1990 when they are 7 years old, and a large fraction of this group is not enrolled in school yet, which explains the estimates of the survival functions for this cohort.

□

The survival functions reveal immense progress in school retention rates for recent birth cohorts in El Salvador. Traditional “exit” points have been the highest grade in each three-year cycle of primary education: the 3<sup>rd</sup>, 6<sup>th</sup>, and 9<sup>th</sup> grades. The estimate of the survival function for the 4<sup>th</sup> grade increases from .902 for those born in the years 1972-75 to .941 for those born in 1984-87, in urban areas. The improvement is even more pronounced in rural areas. The estimate increases from .445 to .797, implying a near

doubling of the probability that rural children will enroll in the 4<sup>th</sup> grade or beyond, bringing the survival probabilities much closer to those observed in urban areas.

□

Retention rates beyond the 6<sup>th</sup> grade show even more improvement. In urban areas, the estimate of the survival function for the 7<sup>th</sup> grade increases from .808 for the 1972-75 cohorts to .929 for the 1984-87 cohorts. The improvement is even more pronounced in rural areas. For grade level 7, the estimate of the survival function increases from .192 for the earliest cohort to .773 for children born in 1984-87, a four-fold increase, and one that makes the probability of enrolling in the 7<sup>th</sup> grade in rural areas considerably closer to those observed in urban areas.

□

The probability that children will enroll in secondary education has increased almost 20 percentage points in urban areas, and increased three-fold in rural areas, between the cohorts born in 1972-75 and those born in 1980-83. For the 1980-83 cohorts, the probability is .884 in urban and .333 in rural areas.

□

Figures 5 and 6 make evident the timing of the improvements highlighted above. Focusing on the survival function at the 6<sup>th</sup> grade, in urban areas we see a steady, gradual rising of the survival probability as we move from the earliest to the most recent cohorts. In contrast, in rural areas we see little improvement for the 1976-79 cohorts compared to the 1972-75 cohorts. But then we see a big upward shift of the survival function for those born in 1980-83, and an even bigger shift for the 1984-87 cohort. More generally, the survival functions for the urban areas suggest that the 1980-83 cohorts enjoyed the largest improvement relative to the cohort immediately prior. In rural areas, the 1984-87 cohorts experienced, by far, the largest improvement relative to the cohort that preceded it. Clearly, improvements in retention rates were delayed in rural compared with urban areas.

□

There is another interesting aspect to the timing of the improvement in school retention rates. Note that the children born in the years 1980 to 1983—the cohorts enjoying the largest relative improvement in urban areas—were aged 6 to 9 in 1989, the year that the

civil war ended. Those children were at the beginning of their school careers, and it appears that they benefited greatly from the ending of the civil war. Put differently, the civil war probably disrupted significantly the functioning of schools and impoverished a large number of families with the unfortunate consequence that many children of schooling age in the decade of the 1980's received very little schooling. The delayed improvement in school retention rates in rural areas perhaps can be explained by the severity of the effects of the civil war and the fact that it took longer for things to go back to normal in the rural areas. Another contributing factor may be the higher rate of retention in grade in rural than urban areas.

□

Finally, Figures 7 and 8 present the estimated baseline hazards. The hazards for the urban area show that improvement in retention rates has been concentrated in the traditional "exit" points. There has been scant progress up through the 5<sup>th</sup> grade, but the hazard has decreased significantly at the 6<sup>th</sup> and 9<sup>th</sup> grades, as we move from the earliest to the more recent cohorts. In sharp contrast, in rural areas the improvement has occurred at all grade levels.

□

### **Concluding Remarks**

□

We study school attainment with cross-sectional data using the Cox proportional hazard model. This technique is particularly attractive to examine data for El Salvador because there has been a significant expansion in school attainment in the recent years. The estimates reveal immense progress in school retention rates for recent birth cohorts in El Salvador.

□

The estimated survival functions for the urban areas suggest that the 1980-83 birth cohorts enjoyed the largest improvement relative to the cohort immediately prior. In rural areas, the 1984-87 cohorts experienced, by far, the largest improvement relative to the cohort that preceded it. Clearly, improvements in retention rates were delayed in rural compared with urban areas. The cohorts enjoying the largest relative improvement in urban areas were aged 6 to 9 in 1989, the year the civil war ended. These children

were luckier than their older siblings were; they benefited from government efforts to increase enrollment in primary schooling, especially in rural areas. These efforts were a central component of the peace process.

□

The key determinant of school retention is parental schooling. The evidence of strong correlation between parental schooling and school survival both in urban and rural areas is especially interesting in El Salvador, because of the significant differences in school availability between urban and rural areas. It suggests that the task of increasing school retention would generally take several generations, and would continue in spite of the expansion of school facilities.

□

Among the factors that affect the costs and benefits of schooling we are particularly interested in the effect of income and the possible impact of remittances. The evidence from El Salvador sends a powerful message regarding the impact of the budget constraint on school retention rates. We find that budgets have a small, though significant, impact on the hazard of leaving school in rural and urban areas. But we find that remittances have a much larger impact on the hazard of leaving school. In urban areas, the effect of remittances is, at its smallest, 10 times the size of the effect of other income. In rural areas, the effect of remittances is about 2.6 times that of other income. In addition, the presence of remittances—irrespective of amounts—lowers the hazard of leaving school in rural areas but has no effect in urban areas.

□

Our finding that parental schooling has a significant effect on the hazard of leaving school is consistent with previous literature. As others have argued, parental schooling may simultaneously capture genetic ability, motivation, and the capacity to generate income. This ambiguity is problematic for it weakens the policy implications for improving opportunities for children who lack schooling. We measure income from a source that is not directly correlated with parental schooling—remittances, and find that remittances have a larger effect than other types of income on school retention. This suggests that relaxing the budget constraint of poor households does have an effect on children's school attainment, even if parents have low levels of schooling.

## References

□  
Al-Sanarrai Samer; and Tessa Peasgood (1998): "Educational Attainments and Household Characteristics in Tanzania," *Economics of Education Review* Vol. 17, No. 4, pp. 395-417.

□  
Becker, Gary (1991): *A Treatise on the Family Enlarged Edition* Cambridge: Harvard University Press

□  
Becker, Gary and Nigel Tomes (1976): "Child Endowments and the Quality and Quantity of Children," *Journal of Political Economy* Vol. 87, No. 4, pt.2: S143-S162.

□  
Behrman, Jere R. and Barbara L. Wolfe (1984), "The Socioeconomic Impact of Schooling in a Developing Country," *The Review of Economics and Statistics* Vol. 66, pp. 296-303.

□  
Blake, Judith (1989): *Family Size and Achievement*. Berkeley: University of California Press

□  
Brown, Phillip and Albert Park (2001): "Education and Poverty in Rural China," *Economics of Education Review* (forthcoming)

□  
Cameron, Stephen V. and James J. Heckman (1998): "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males," *Journal of Political Economy* Vol. 106, No. 2, pp. 262-333.

□  
Chuang, Hwei-Lin (1997): "High School Youth's Dropout and Re-Enrollment Behavior," *Economics of Education Review* Vol. 16, No. 2, pp.171-86.

□  
Cochrane, Susan H. and Dean T. Jameson (1982): "Educational Attainment and Achievement in Rural Thailand," in *New Directions for Testing and Measurement: Productivity Assessment in Education #15* edited by Anita A. Summers. San Francisco: Jossey-Bass.

□  
Edwards, Alejandra Cox (2000): "Alternativas de Política Económica para agilizar la creación de empleo," Chapter 3 in *Crecimiento con Participación: Una estrategia de desarrollo para el siglo XXI*. Vol. II Fusades. San Salvador.

□  
Filmer, Deon and Lant Pritchett (2000): *The Effect of Household Wealth on Educational Attainment around the World: Demographic and Health Survey Evidence*. The World Bank.

□  
Funkhouser, Edward (1992): "Migration, remittances and economic adjustment," in G. Borjas and R. Freedman eds. *Immigration and the work force* (University of Chicago Press).



Grambsch, P. M. and T. M. Therneau (1994): "Proportional hazards tests and diagnostics based on weighted residuals," *Biometrika* Vol. 81, pp. 515-26.

Haveman, Robert and Wolfe, Barbara (1995): "The Determinants of Children's Attainments: A Review of Methods and Findings," *Journal of Economic Literature*, 33: 1829-78.

Hill, Martha, and Duncan, Greg J. (1987): "Parental Family Income and the Socioeconomic Attainment of Children," *Social Science Research*, 16 (1): pp 39-73.

Inchauste, Graciela (2000): "Educational Choices and Educational Constraints: Evidence from Bolivia," IMF Working Paper WP/00/42 March.

Jacir de Lovo, Ana Evelyn (1997) *Reforma Educativa y Decentralización: Aspectos Institucionales* (Documento Fusades).

Jimenez, Emmanuel and Yasuyuki Sawada (1998): "Do Community-Managed Schools Work? An Evaluation of El Salvador's EDUCO Program" The World Bank Research Department, February.

King, Elizabeth M. and M. Anne Hill (1993): *Women's Education in Developing Countries: Barriers, Benefits and Policies* Baltimore: The John Hopkins University Press for The World Bank.

Knight, John B. and Richard H. Sabot (1990): *Education, Productivity, and Inequality: The East African Natural Experiment*. Baltimore: The John Hopkins University Press for the World Bank.

Lillard, Lee A. and Robert J. Willis (1994), "Intergenerational Educational Mobility: Effects of Family and State in Malaysia," *The Journal of Human Resources* Vol. 29, No. 4, pp. 1126-66.

Lopez, Jose Roberto, et al. (1991): "Small Business Development in El Salvador: The Impact of Remittances," *Migration, remittances, and small business development: Mexico and Caribbean Basin countries* Series on Development and International Migration in Mexico, Central America and the Caribbean Basin, vol 4, Boulder and Oxford: Westview Press.

MINED (Ministry of Education) (1997). *Dinámica de la Matricula Estudiantil de Educación Parvularia, Básica y Media correspondiente al periodo 1986-1996*. San Salvador (May).

Montes, Segundo (1990): "Las remesas que envían los Salvadoreños de Estados Unidos. Consecuencias Sociales y Económicas. Instituto de Investigaciones, Universidad Centroamericana Jose Simeon Canas, San Salvador.

Parish, William L. and Robert J. Willis (1993): "Daughters, Education, and Family Budgets: Taiwan Experiences," *The Journal of Human Resources* Vol. 28, No. 4, pp.863-98.

□

Parson, Donald O. (1975): "Intergenerational Wealth Transfers and the Educational Decisions of Male Youth," *Quarterly Journal of Economics*, pp 605-17.

□

Schultz, T. Paul (1989): "Education Investments and Returns," in *Handbook of Development Economics*, vol. 1, edited by Hollis Chenery and T.N. Srinivasan. Amsterdam: North-Holland.

□

STATA (1999). *Stata Statistical Software: Release 6.0*. College Station, TX: Stata Corporation.

□

Weesie, Jeroen (1999): "Seemingly Unrelated Estimation and the Cluster-Adjusted Sandwich Estimator," *STATA Technical Bulletin Reprints*, Vol. 9, pp. 231-48.

□

Wolfe, Barbara and Jere Behrman (1984): "Who is schooled in Developing Countries? The Roles of Income, Parental Schooling, Sex, Residence and Family Size" *Economics of Education Review* Vol. 3, No. 2, pp. 231-45.

**Table 1. Location of Households Receiving Remittances in the Income Distribution of Households that Do Not Receive Remittances**

Urban Households				Rural Households			
Income distribution of non recipient households Decile	Households receiving remittances			Income distribution of non recipient households Decile	Households receiving remittances		
	% of households in given decile (income before remittances)	Median remittance amount (colones)	% of households in given decile (income plus remittances)		% of households in given decile (income before remittances)	Median remittance amount (colones)	% of households in given decile (income plus remittances)
1	27.4	875	6.3	1	36.5	875	3.1
2	9.9	875	9.2	2	8.4	525	5.1
3	9.6	586	10.5	3	9.9	438	6.0
4	10.8	875	9.2	4	6.2	875	12.1
5	5.6	438	11.6	5	5.0	438	10.2
6	7.9	586	10.3	6	7.4	875	9.6
7	5.7	875	9.6	7	4.3	438	11.3
8	8.1	875	10.4	8	4.1	875	10.4
9	9.8	875	14.1	9	8.3	438	14.4
10	5.2	875	8.8	10	10.1	875	17.8

**Table 2: Fraction of Female Headed Households among Non-Recipient, Recipient, and All Households, in Rural and Urban Areas**

Urban Households			
% of Female Headed Households	Non-recipient households %	Recipient households %	All households %
	27.4	45.9	30.2

Rural Households			
% of Female Headed Households	Non-recipient households %	Recipient households %	All households %
	22.5	41.7	25.2

**Table 3. Distribution of Non-Recipient and Recipient Households, By Schooling of the Adult Female**

Urban Households			
Years of schooling of the adult female	Non-recipient households %	Recipient households %	Median Remittance (colones)
0 - 3	35.3	48.7	875
4 - 6	20.2	20.5	875
7 - 9	16.1	9.7	875
9 - 12	17.4	14.9	875
13 +	11.1	6.2	875

Rural Households			
Years of schooling of the adult female	Non-recipient households %	Recipient households %	Median remittance (colones)
0 - 3	75.3	80.3	875
4 - 6	17.5	14.5	656
7 - 9	5.2	3.8	875
9 or more	2.0	1.5	438

**Table 4. Distribution of Non Recipient and Recipient Households, by the Number of School-Age Children**

<b>Urban Households</b>			
Number of school-age children	Non-recipient households %	Recipient households %	Median remittance (colones)
none	32.2	36.4	875
1	23.3	26.4	875
2	23.3	18.1	438
3 or more	21.1	19.1	875

<b>Rural Households</b>			
Number of school-age children	Non-recipient households %	Recipient households %	Median remittance (colones)
none	28.0	31.0	875
1	19.4	26.6	875
2	18.8	15.8	875
3 or more	33.8	26.7	875

**Table 5. Placement of Children in the School System, Percentages by Age Group, 1997**

Age	Target grade	No school	In grade	Behind	Dropout	Total
4 – 6	Pre-school	64.2	35.8			100
7 – 9	I cycle of primary (1-3)	13.1	83.9	1.8	1.2	100
10 – 12	II cycle of primary (4-5)	6.5	60.2	28.6	4.7	100
13 – 15	II cycle of primary (6-9)	5.5	45.2	30.5	18.9	100
16 – 18	Secondary (10-12)	7.5	30.0	22.4	40.1	100

**Table 6. Annual Growth in School Registration, by Grade Level**  
(Average for 1990-96)

Level	Target age	Total			Urban			Rural		
		Total	Public	Private	Total	Public	Private	Total	Public	Private
Pre-K	4 - 6	20.7	24.0	12.6	10.1	8.6	12.6	53.1	57.5	11.9
1 - 9	7 - 15	4.2	3.9	5.9	1.5	0.2	6.7	7.5	7.7	-0.6
10 - 11	15 - 18	11.7	8.0	16.4	11.7	8.0	16.4			

**Table 7. Distribution of School Attainment in 1997, by Age**

Age	No school	Primary I	Primary II	Primary III	Secondary	Post-Sec	Total
24 - 25	12.9	15.2	22.6	23.8	19.7	5.7	100
26 - 30	14.3	15.9	22.1	18.5	20.3	9.0	100
31 - 35	17.8	17.7	19.9	16.9	16.6	11.1	100
36 - 40	22.2	18.7	21.3	13.9	14.2	9.7	100
41 - 45	24.2	21.3	23.2	12.0	13.0	6.3	100
46 - 50	31.4	25.2	21.9	7.2	9.4	4.9	100
51 - 55	34.1	25.2	20.5	7.8	7.1	5.3	100
56 - 60	41.6	24.8	16.9	5.2	5.3	6.2	100
61 - 65	46.2	25.9	16.1	4.3	3.7	3.8	100
66 - 70	53.1	24.8	12.3	2.7	4.1	3.1	100
71 +	58.0	20.8	14.0	3.0	2.3	2.0	100
Total	32.5	19.1	19.8	12.4	11.2	5.1	100

**Table 8. Household Income and Remittances in 1997,**  
**By Region (full sample, in Colones)**

	Urban	Rural
Average household income	4,105	1,768
Per capita average household income	1,081	419
Recipient households (% of all)	15%	14%
Median remittance (among recipients)	875	875
Average remittance (among recipients)	1,183	909

**Table 9. Estimates of the Determinants of the Hazard of Leaving School, Urban Areas**  
(Cox Proportional Hazard Stratified by Birth Cohort)

	Specification that uses income (in hundreds of Colones)		
Covariate; <i>segment of the hazard over which the effect applies, if it does not apply to the entire hazard</i>	Hazard Ratio	Robust st. error	z-stat.
Male	1.274	.058	5.32
Parental schooling; <i>never enrolled, 1<sup>st</sup>-6<sup>th</sup> grade</i>	.801	.011	-15.98
Parental schooling; <i>7<sup>th</sup>-9<sup>th</sup> grade</i>	.860	.012	-10.69
Parental schooling; <i>10<sup>th</sup>-12<sup>th</sup> grade</i>	.927	.009	-8.10
No access to water or electricity	2.553	.272	8.81
Income (net of remittance)	.996	.001	-3.47
Remittance amount; <i>never enrolled, 1<sup>st</sup>-6<sup>th</sup> grade</i>	.916	.016	-5.08
Remittance amount; <i>7<sup>th</sup>-12<sup>th</sup> grade</i>	.965	.008	-4.48
(Born in 1988-1990) × (Parental schooling; <i>never enrolled, 1<sup>st</sup>-6<sup>th</sup> grade</i> )	1.111	.021	5.59
(Born in 1988-1990) × (Remittance amount; <i>never enrolled, 1<sup>st</sup>-6<sup>th</sup> grade</i> )	1.084	.024	3.58
Log-likelihood	-11,843.387		
Number of observations	7,382		

□  
□

**Table 10. Estimates of the Determinants of the Hazard of Leaving School,  
Rural Areas  
(Cox Proportional Hazard Stratified by Birth Cohort)**

□

Covariate; <i>segment of the hazard over which the effect applies, if it does not apply to the entire hazard</i>	Hazard Ratio	Robust st. error	z-stat
Male; <i>never enrolled, 7<sup>th</sup>-12<sup>th</sup> grade</i>	1.156	.054	3.07
Male; <i>1<sup>st</sup>-6<sup>th</sup> grade</i>	.902	.046	-2.02
Parental schooling; <i>never enrolled</i>	.839	.015	-10.06
Parental schooling; <i>1<sup>st</sup>-6<sup>th</sup> grade</i>	.874	.013	-9.41
Parental schooling; <i>7<sup>th</sup>-12<sup>th</sup> grade</i>	.923	.015	-4.83
No access to water or electricity; <i>never enrolled</i>	1.856	.138	8.29
No access to water or electricity; <i>1<sup>st</sup>-6<sup>th</sup> grade</i>	1.334	.088	4.36
Income (net of remittance); <i>never enrolled, 1<sup>st</sup>-6<sup>th</sup> grade</i>	.993	.002	-3.44
Income (net of remittance); <i>7<sup>th</sup>-12<sup>th</sup> grade</i>	1.000	.002	0.05
Remittance amount	.983	.011	-1.57
Receives remittance	.764	.088	-2.35
Log-likelihood	-19,443.516		
Number of observations	6,904		

□



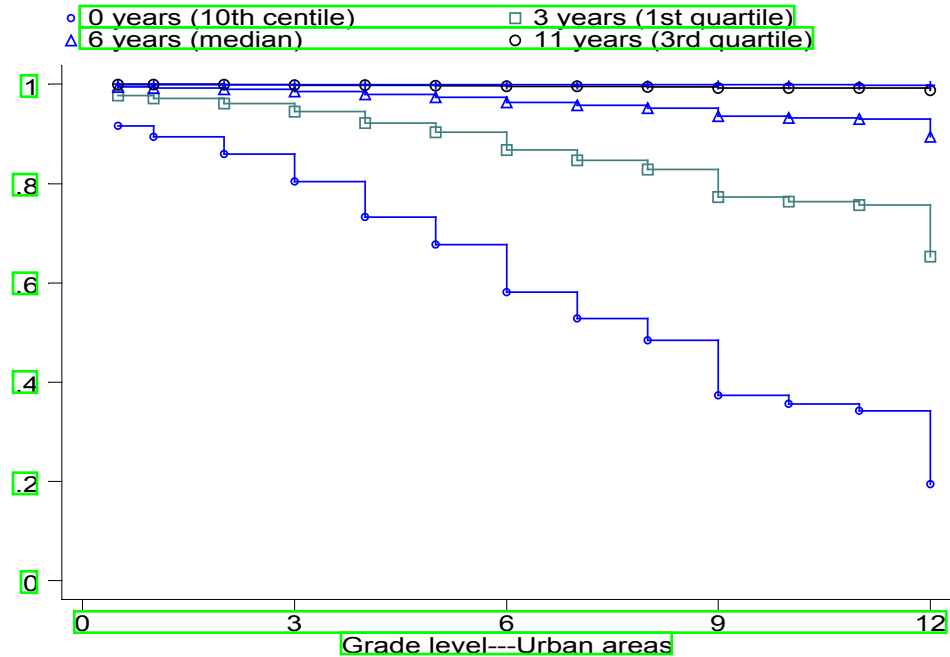
**Table 11. Estimated Survival Functions, by Birth Cohort  
Urban Areas**

	Birth Cohort				
Grade Level	1972-75	1976-79	1980-83	1984-87	1988-90
0	.971	.976	.982	.965	.919
1	.962	.970	.976	.960	.915
2	.946	.959	.964	.951	.912
3	.926	.942	.948	.946	.912
4	.902	.918	.932	.941	.912
5	.879	.898	.916	.936	
6	.829	.861	.901	.929	
7	.808	.838	.892	.929	
8	.792	.819	.884	.929	
9	.699	.761	.884		
10	.684	.752	.862		
11	.672	.743	.856		
12	.491	.635	.856		

**Table 12. Estimated Survival Functions, by Birth Cohort  
Rural Areas**

	Birth Cohort				
Grade Level	1972-75	1976-79	1980-83	1984-87	1988-90
0	.832	.854	.896	.891	.645
1	.787	.814	.856	.863	.634
2	.679	.714	.781	.837	.630
3	.552	.601	.701	.816	.613
4	.445	.493	.617	.797	.613
5	.353	.393	.539	.793	
6	.217	.276	.441	.773	
7	.192	.246	.420	.773	
8	.174	.229	.408	.773	
9	.110	.150	.333		
10	.101	.143	.333		
11	.095	.140	.333		
12	.035	.071	.333		





□

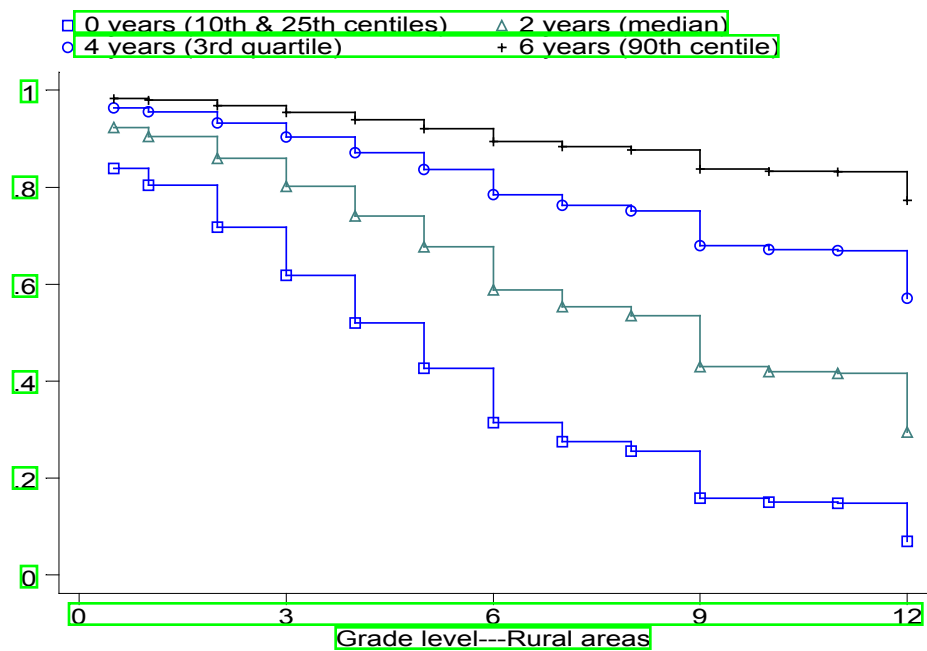
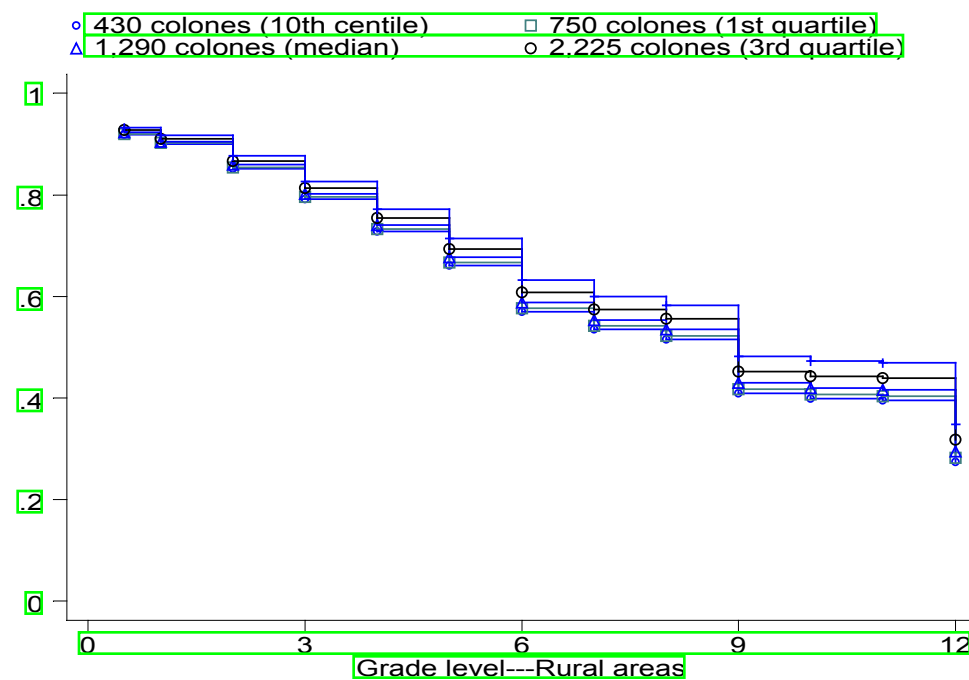
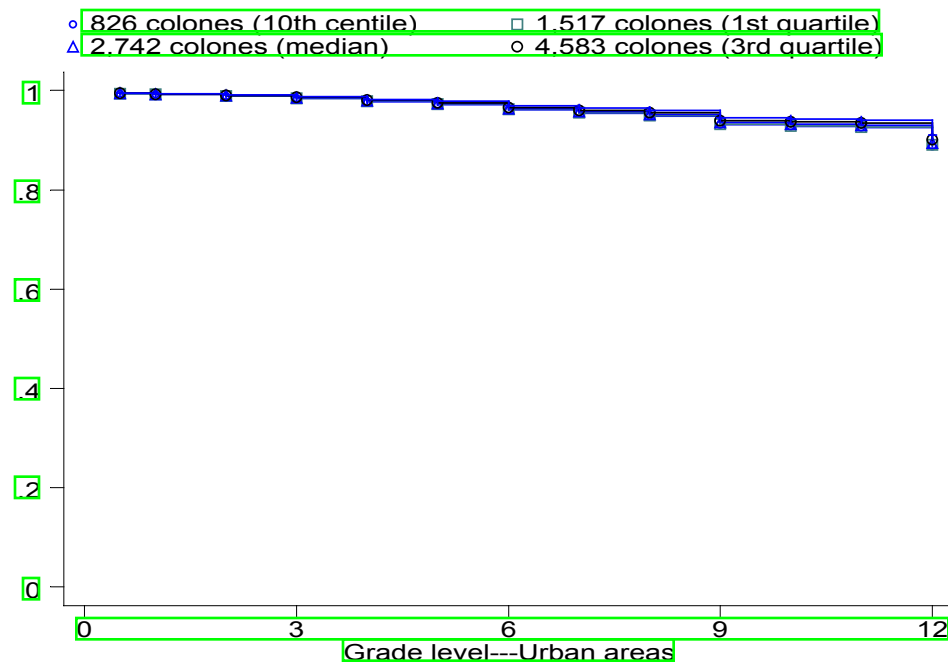
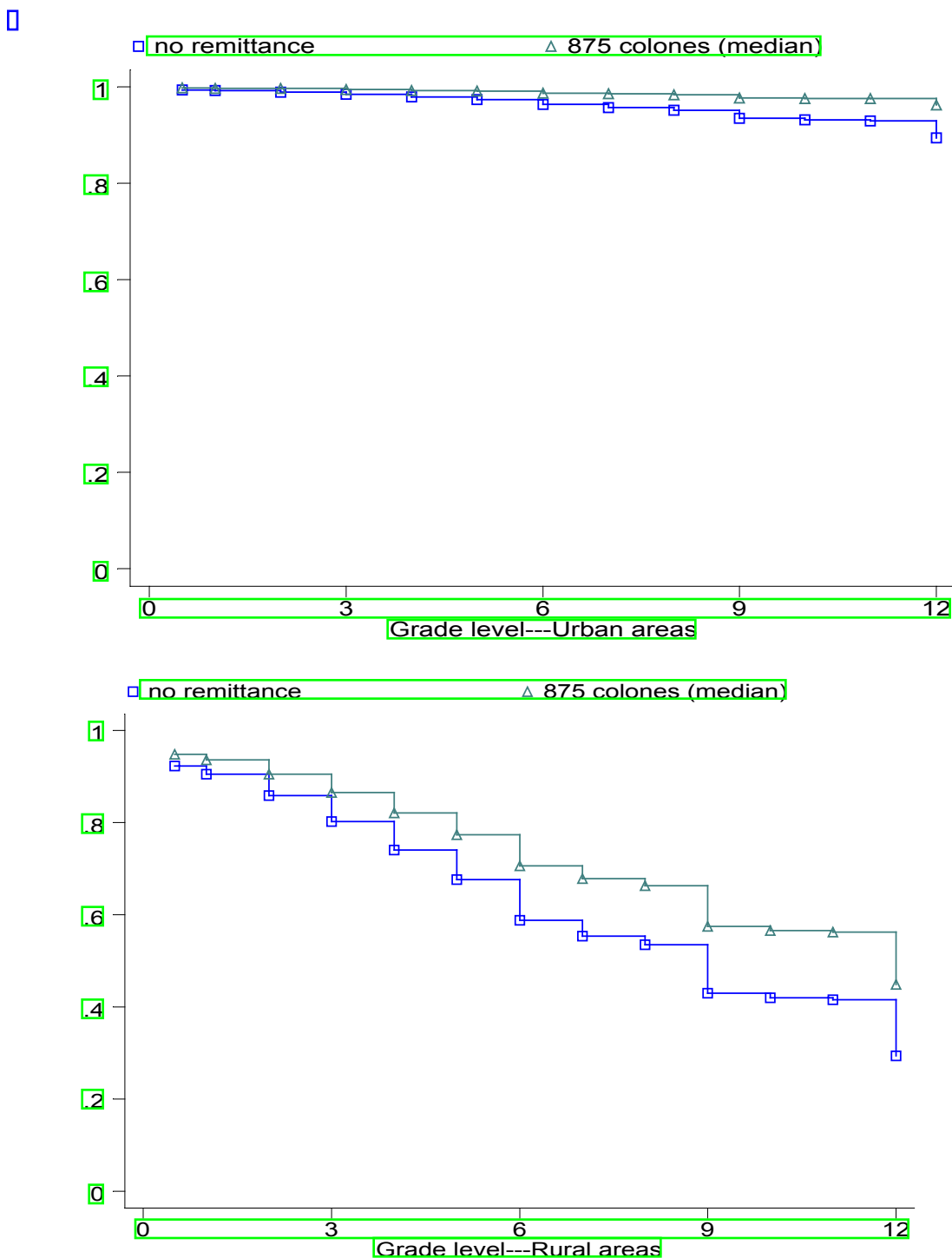


Figure 2. Estimated survival functions, by level of parental schooling, for urban and rural areas. (Household income is set at the median for each area. The functions correspond to the 1976-79 birth cohorts)



1  
 2  
 3

**Figure 3. Estimated survival functions, by level of household income, for urban and rural areas. (Parental schooling is set at the median for each area. The functions correspond to the 1976-79 birth cohorts)**



**Figure 4. Estimated survival functions, with and without remittances in urban and rural areas. (Household income and parental schooling are set at the median for each area. The functions correspond to the 1976-79 birth cohorts)**

□  
□

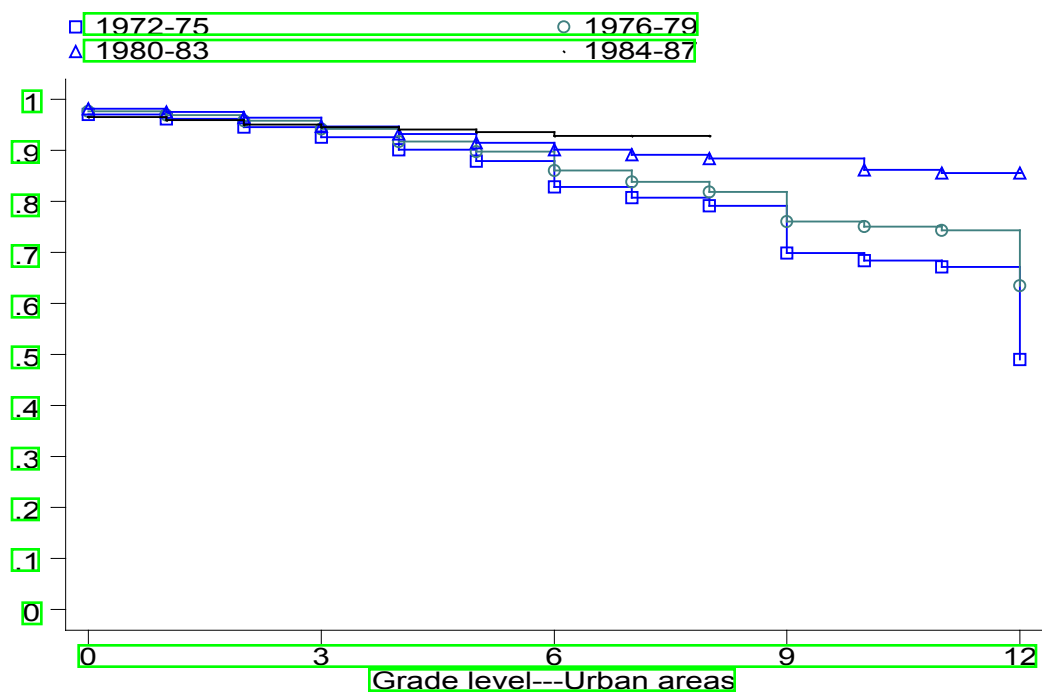


Figure 5. Estimated survivor functions, by birth cohort

□  
□

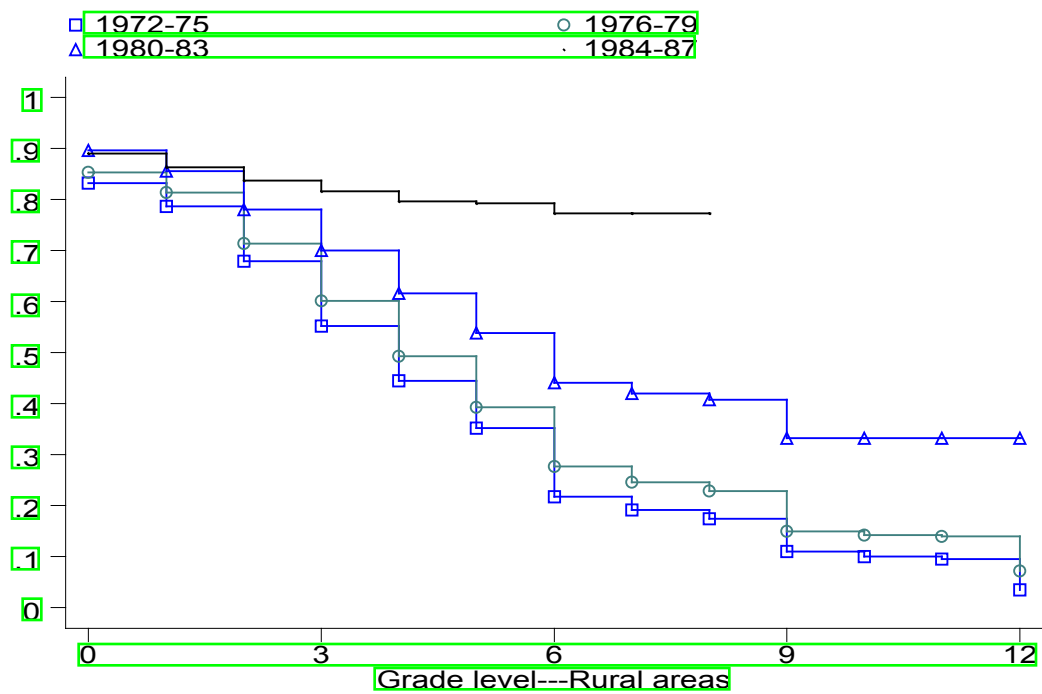


Figure 6. Estimated survivor functions, by birth cohort

□

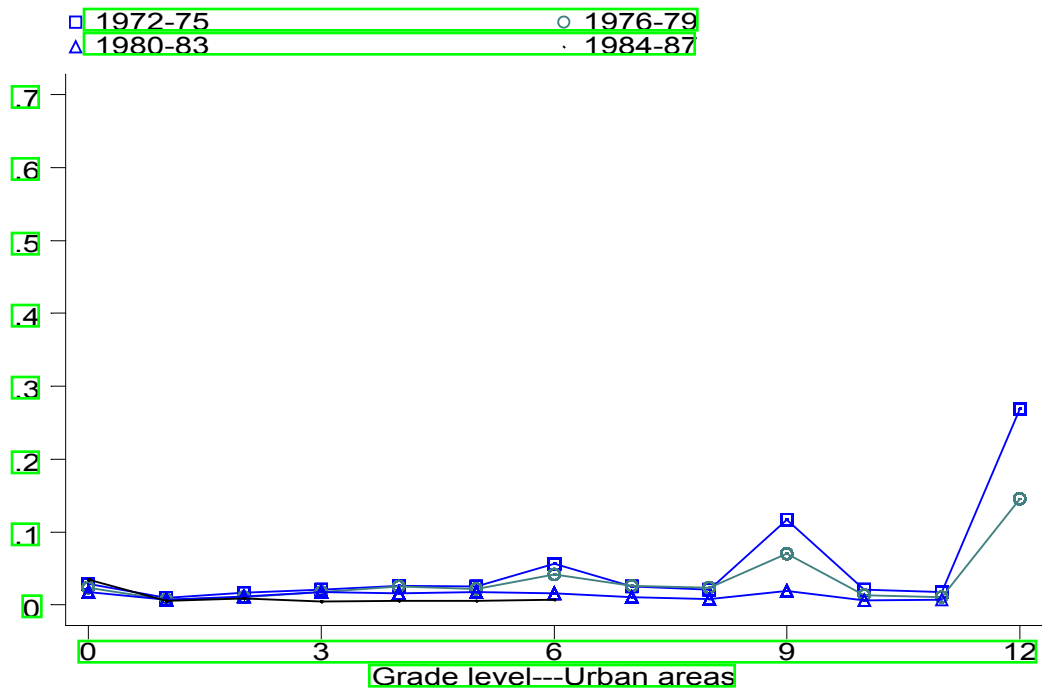


Figure 7. Estimated baseline hazards, by birth cohort

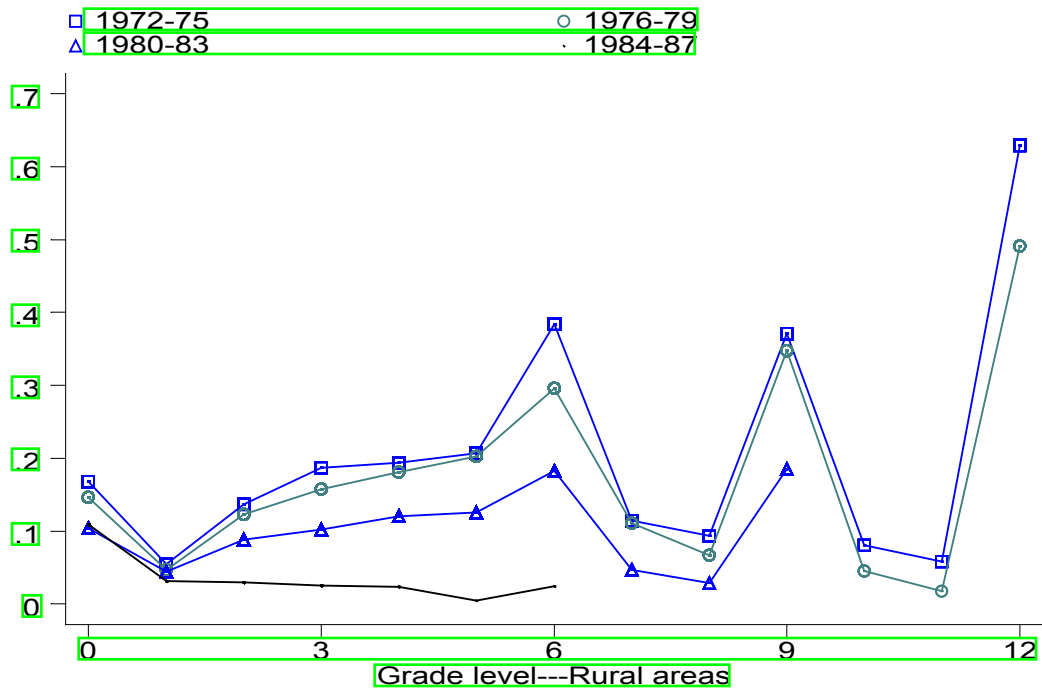


Figure 8. Estimated baseline hazards, by birth cohort

## **APPENDIX**

□

### **Specification of the Statistical Model**

□

We analyze separately the behavior of rural and urban households. The key reason being the presence of factors we cannot control for that affect the demand for schooling in rural areas. Schools are more readily available in urban areas; the living conditions in urban areas are much better than in rural areas, and the organization of the economy, particularly in the way it affects the costs of attending school is less favorable to schooling in rural than urban areas. The separate analyses are not driven by the assumption that perceived benefits from schooling differ between rural and urban areas. El Salvador is a relatively small and densely populated country; the fact that internal migration is significant does not justify that assumption. Nevertheless, we tested the null hypothesis that the urban and rural samples can be combined and the estimated coefficients are equal for the two samples. We reject the null hypothesis, with a p-value for the test statistic equal to 0.00. As a second check, we tested the null hypothesis that, while the underlying baseline hazards differ for the urban and rural samples, the effect of the covariates on the hazards are the equal for the two samples. Once again, we reject the null hypothesis, with a p-value for the test statistic equal to 0.00.

□

The crucial assumption in the Cox proportional hazard model is that the hazard ratio is proportional over the entire base line. For example, if the hazard of leaving school is 30 percent higher for boys than for girls, the proportional hazard assumption implies that this ratio is the same in the first grade, the fifth grade, or at any grade level. In our search for a parsimonious specification, we paid special attention to the validity of the proportionality assumption. We tested the assumption for each individual covariate and conducted a global test.<sup>14</sup>

□

---

<sup>14</sup> Grambsch and Therneau (1994) proposed a scaled adjustment of the Schoenfeld residuals and a global test of the proportionality assumption based on the adjusted residuals. The global test and a chi-squared test for individual covariates are implemented in STATA 6.0 (1999).



The testing revealed that the proportionality assumption was invalid for at least one covariate in the urban and the rural samples. We relaxed the assumption of proportionality by allowing the covariates to have different effects over four segments of the baseline hazard. The segments are “never enrolled,” 1<sup>st</sup> through the 6<sup>th</sup> grade, 7<sup>th</sup> through the 9<sup>th</sup> grade, and 10<sup>th</sup> through the 12<sup>th</sup> grade.<sup>15</sup> When we repeated the tests we found that the variable “year of birth” continued to fail the proportionality test. Consequently, we estimated separate models for the urban and rural samples allowing separate effects for the four segments of the baseline hazard, and we stratified each sample on individual year of birth. Inspection of the estimated baseline hazard and survival function for each stratum revealed that it is reasonable to combine the observations into 5 strata according to year of birth. The four oldest strata combine individuals born in four-year intervals: 1972 to 1975, 1976 to 1979, 1980 to 1983, and 1984 to 1987. The youngest stratum is for individuals born between 1988 and 1990.

The next step in the specification search involved testing the null hypothesis that the effect of the covariates is the same for all the strata. We estimated a fully interacted model, where the effect of each covariate was allowed to differ between strata and segments of the baseline hazard, separately for the urban and rural samples. Then, we estimated the model that imposes the constraints contained in the null hypothesis. For the urban sample, the likelihood ratio test results in rejection of the null hypothesis. The test statistic is equal to 97.69 with 69 degrees of freedom, and a p-value equal to 0.013. When we allow the effects to differ for the youngest stratum, those born between 1988 and 1990, we fail to reject the null hypothesis that the effects of the covariates are equal

---

<sup>15</sup>We divided the baseline hazard into the segments listed in text based on the following considerations. Primary education is divided into three cycles of three grades each. Therefore, when distance to school becomes a problem, it is likely it will happen when a child is ready to enroll in the next cycle of primary education or secondary education. In addition, public education becomes relatively scarce at the third cycle of primary education, and close to half of secondary education is private. This suggests that out-of-pocket costs of additional years of schooling beyond the sixth grade rise relative to schooling at younger ages. In addition, in the early stages of estimation we divided the hazard further, allowing for separate effects for the 1<sup>st</sup> through the 3<sup>rd</sup> grade, and the 4<sup>th</sup> through the 6<sup>th</sup> grade. But we did not reject the null hypothesis that the segments could be combined into one covering the 1<sup>st</sup> through the 6<sup>th</sup> grade.

for the remaining 4 strata. The test statistic is equal to 55.79 with 58 degrees of freedom, and a p-value equal to 0.558.<sup>16</sup>

□

We continued testing nested hypotheses aimed at simplifying the specification as much as possible. Finally, we reduced the original specification that had 93 separate effects to one with 10 effects. The covariates in our preferred specification for the urban sample are:

1. an indicator for sex
2. parental schooling—first two segments of the hazard
3. parental schooling—third segment of the hazard
4. parental schooling—fourth segment of the hazard
5. an indicator for lack of access to basic services
6. household income
7. remittance amount—first two segments of the hazard
8. remittance amount—last two segments of the hazard
9. an indicator variable for the youngest stratum interacted with parental schooling for the first two segments of the hazard
10. an indicator variable for the youngest stratum interacted with the remittance amount for the first two segments of the hazard.

□

Note that we failed to reject the null hypothesis that the presence of remittances (after controlling for the remittance amount) has no effect on the hazard of leaving school, and therefore, we drop it from the specification for the urban sample.

□

---

<sup>16</sup> We use six covariates: indicators for sex, lack of access to basic services, and presence of remittances, parental schooling, location on the income distribution, and change in the location due to the remittance. We allow the effects to differ for four segments of the hazard. Therefore, in principle, the set of estimated coefficients for each stratum has 24 ( $6 \times 4$ ) coefficients. Note that that two tests described in text only differ by 11 degrees of freedom. The reason is that there are no observations in the later segments of the baseline hazard for the younger birth cohorts—no one born in 1988 is in the 7<sup>th</sup> grade or beyond by 1996, the survey year. Also, there are no observations with “no access to basic services” falling in the second segment of the hazard. Consequently, we can only estimate 11 separate parameters for the youngest stratum, which explains the difference of 11 degrees of freedom between the tests.

The individual tests of the null hypothesis of proportionality for each of the ten variables listed fail to reject the null hypothesis. The largest test statistic is for household income. It is equal to 2.68, and is asymptotically distributed chi-squared with 1 degree of freedom, so the p-value for the test is equal to 0.102. When we perform the global test of the null hypothesis of proportionality of effects over all segments of the baseline hazard, the test statistic is equal to 4.38, with 10 degrees of freedom, and a p-value equal to 0.929, and we fail to reject the null hypothesis.

□

For the rural sample, we fail to reject the null hypothesis that the effects of the covariates are equal for all strata. The statistic for the likelihood ratio test is equal to 82.08, distributed with 66 degrees of freedom, and the p-value is equal to 0.088. But we reject the null hypothesis that the covariates have the same effect on all four segments of the baseline hazard. Through sequential testing we arrived at the simplest specification of the model that does not violate the proportionality assumption of the Cox proportional hazard model. There are eleven covariates in the specification:

1. an indicator for sex—first, third and fourth segments of the hazard
2. an indicator for sex—second segment of the hazard
3. parental schooling—first segment of the hazard
4. parental schooling—second segment of the hazard
5. parental schooling—third and fourth segments of the hazard
6. an indicator for lack of access to basic services—first segment of the hazard
7. an indicator for lack of access to basic services—second segment of the hazard
8. household income—first and second segments of the hazard
9. household income—third and fourth segments of the hazard
10. remittance amount
11. an indicator for the presence of remittances

□

We failed to reject the null hypothesis that lack of access to basic services, for the third and fourth segments of the hazard (that is, the 7<sup>th</sup> through the 12<sup>th</sup> grade), has no effect on the hazard of leaving school.

The individual tests of the null hypothesis of proportionality for each of the ten variables listed fail to reject the null hypothesis. The largest test statistic is for parental schooling for the second segment of the hazard. It is equal to 3.10, and is asymptotically distributed chi-squared with 1 degree of freedom, so the p-value for the test is equal to 0.078. When we perform the global test of the null hypothesis of proportionality of effects over all segments of the baseline hazard, the test statistic is equal to 12.14, with 11 degrees of freedom, and a p-value equal to 0.353, and we fail to reject the null hypothesis.

□

We conclude the discussion of our specification search noting that “Parental schooling” is the only covariate with a different effect on the hazard of leaving school at different grade levels, both in the urban and the rural samples.

□

The pattern of effects of parental schooling estimated for the urban and rural samples has intuitive appeal. As a child progresses through the school system, the child’s own school experience will weigh more heavily, and the parents’ schooling will weigh less heavily, in the evaluation of the returns to additional schooling for the child. In the sociology of education literature, the declining effect of “family background” variables (typically parental schooling and family income) on ever-higher school grade transitions has been interpreted as “dynamic selection bias.” In other words, family background correlates with unobserved factors like ability, and as the group of students who attain ever-higher grade levels becomes smaller and smaller it also becomes increasingly less representative of the population of individuals of that age because they are of above-average ability. The effect of family background variables declines because the correlation of these variables with unobserved ability fades away as the sample of “survivors” gets smaller and average ability in the sample increases.

□

Cameron and Heckman (1998) argue that the standard model used in the sociology of education literature, a logit specification, together with the data typically used to estimate it require ad hoc distributional assumptions before inferences regarding dynamic selectivity bias can be made. They estimate a logit model and an ordered discrete-choice

model with corrections for unobserved heterogeneity. The logit results are of interest because they allow a comparison of parameter estimates with and without corrections for unobserved heterogeneity.

□

Comparing the estimates for variables capturing parental schooling and family income reported in Cameron and Heckman's Tables 1 and 4, it is clear that the selectivity correction makes scant difference for primary and secondary schooling. The effect of correcting for selectivity is concentrated in graduating from high school and higher schooling levels. Moreover, in their preferred specification, the ordered discrete-choice model, the estimates of the density of "ability" at various schooling levels suggest that the density for high school graduates is only slightly to the right of the baseline density. We interpret Cameron and Heckman's results as suggesting that dynamic selectivity bias is unimportant in primary and secondary schooling. Indeed, in most countries the curriculum is set at a level of difficulty that allows near universal graduation from primary school and very high graduation rates from secondary school.

□

For several reasons, we have elected not to implement a correction for selectivity. First, the evidence discussed in the paragraph above suggests that dynamic selectivity bias is a minor problem for the analysis of grades 1 through 12, the grades we examine. Second, our results do not conform to the typical pattern found in logit models: our estimate of the effect of family income declines for higher grade levels in rural but not in urban areas. Third, assume we found a lowering of the hazard of leaving school in grades 1 through 6 and an increase in the hazard of leaving school after the 6<sup>th</sup> grade over time. It is possible that the increase in the hazard beyond the 6<sup>th</sup> grade would be due to a "worsening" of the pool of students because of higher retention rates in grades 1 through 6. We could conclude erroneously that matters have gotten worse in grades 7 through 12. But we have found no such pattern. We estimate improvements in retention rates at all grade levels as we move from the oldest to the most recent birth cohorts.