Ciencias Económicas 31-No.1: 2013 / 9-29 / ISSN: 0252-9521

RETURNS TO EDUCATION, COHORTS AND BUSINESS CYCLE IN COSTA RICA: 1987-2008

Diego Rojas[1]

Recibido: 23/04/2013 Aceptado: 06/05/2013

ABSTRACT

A pseudo-panel approach is used to estimate the returns to schooling (RTS) in Costa Rica. This approach ameliorates the "ability" bias due to the correlation between the level of education and non-observable characteristics of the individual. We found that RTS are higher for older samples.

Once we study the behavior of the RTS -using different settings and estimators- we analyze their correlation with Deaton's year, cohort and age effects. We found that the income of younger cohorts is greater than the income of older cohorts, once experience and short run fluctuations of the economy are accounted for. This difference in income between generations is explained by differences in levels of education. Other factors that differ between generations seem to be less important to explain their income differences. Finally, we present preliminary evidence suggesting that short run fluctuations of the GDP affect in a greater extent those with less education.

KEY WORDS: PSEUDO PANEL, RETURNS TO EDUCATION, COHORTS

RESUMEN

La técnica de pseudo-panel es utilizada para estimar los rendimientos a la educación en Costa Rica. Este enfoque disminuye el sesgo de "habilidad" debido a la correlación entre el nivel de educación y características no observables del individuo. Encontramos que los rendimientos a la educación incrementan a medida que la muestra utilizada incluya personas de mayor edad.

Una vez que estudiamos los rendimientos a la educación -para diferentes muestras y utilizando diferentes estimadores- analizamos su relación con los efectos año, cohorte y edad de Deaton. Encontramos que el ingreso de cohortes jóvenes es mayor que el ingreso de cohortes viejas, una vez que hemos controlado por la experiencia y las fluctuaciones económicas. Esta diferencia en el ingreso entre generaciones es explicada por los diferentes niveles de educación. Otros factores que difieren entre generaciones parecen ser de menor importancia para explicar dicha diferencia en ingresos. Finalmente, presentamos evidencia preliminar que sugiere que las fluctuaciones de corto plazo en el PIB afectan en mayor medida a aquellos con menor educación.

PALABRAS CLAVES: PSEUDO PANEL, RENDIMIENTOS A LA EDUCACIÓN, COHORTES

1 Banco Mundial, USA; rojas.luisdiego@gmail.com, Código Postal WASHINGTON-DC, 20015.
I would like to thank Daniel Lederman for his guidance, comments and contributions to this paper. I would also like to thank an anonymous referee for helpful comments. The findings, interpretations, and conclusions expressed in this paper are entirely those of the author. They do not necessarily represent the views of the World Bank or any other organization with which the author may be affiliated. Any errors or omissions are the author's responsibility. The author is a Consultant with the Office of the Chief Economist for Latin America and the Caribbean of the World Bank.

1. INTRODUCTION

The relationship between education and income has been widely studied for different countries. For instance, Psacharopoulos(1993), Psacharopoulos & Ng(1992) and Psacharopoulos & Patrinos(2004) present estimates of the return to education for a wide range of countries. Most of these studies use cross-sectional data to estimate the return to schooling (RTS), although recently the use of panel and pseudo-panel data has increased.

It is well known that when characteristics of the individual (such as ability) are not taken into account and they are correlated with the level of education, the ordinary least square (OLS) estimation of the RTS is biased. To correct this "ability" bias, two common methods had been employed; the use of panel data and instrumental variables. The main problem when using instrumental variables is to find proper instruments. The main problem using panel data is the lack of such data (particularly, for developing countries).

A third option is to use series of cross-sections in the form of a pseudo-panel. When analyzing the return to schooling, this data structure permits to reduce the ability bias that is present in a single cross-section. Moreover, this approach can be applied in many developing countries in which it is common to find series of household income surveys that change the sample every year.

Pseudo-panel data has other advantages over panel data. As Deaton (1997) points out, pseudo-panel does not suffer from attrition because it is constructed from new samples every year; also cohort data is likely to be less susceptible to measurement error than panel data, for "the quantity that is tracked is normally an average and the averaging will nearly always reduce the effects of measurement error" (Deaton (1997), p-120).

Given these advantages, pseudo-panels have recently been used to estimate the returns to schooling. For instance, Warunsiri and McNown (2009) use pseudo-panel techniques to calculate the return to schooling in Thailand. Dickerson et al (2001) use a similar approach to calculate the returns to education in Brazil, while Kaymak(2008) uses cohort data to calculate the return to schooling in the US.

In spite of its advantages, the use of pseudo-panel data carries a set of difficulties when it is applied to the estimation of the return to schooling. In particular, when this approach is followed, the results (a) can be sensitive to the pseudo-panel setting and (b) attenuation bias might be present due to sampling errors. Moreover, the estimation might suffer from various types of selectivity bias due to gender, retirement and self-employment decisions[2]. In our application we examine (a) and (b) with less discussion upon the selectivity bias.

Our estimates are based on Deaton and Paxon's (1994) income decomposition. Their methodology decomposes income in three effects: age, cohort and year. Generally, once education is accounted for, the age effect is associated with experience, the year effect is related to macroeconomic fluctuations and the cohort effect is related to particular characteristics of a group (or cohort). In our application, the pseudo-panel is constructed using Costa Rican Household Surveys from 1987 to 2008. In this construction, a set of cohorts is defined (by year(s) of birth) and they are followed through each survey.

Different settings of the pseudo panel and different estimators are calculated. We use three settings for the construction of the pseudo-panel. In our first example, we define each cohort by a single year of birth and we keep, in every year, those individuals between a minimum and a maximum age. This means that every year a new cohort "appears" and the oldest cohort "disappears" from the data set.

In a second example, a fixed set of cohorts are chosen and they are followed through all the years. In this case no cohorts are added or subtracted in any year, but the range of ages observed changes year by year. In particular, each year we use an older sample.

2 These (selectivity) biases are also present when using cross-section to estimate the returns to schooling.

In a third example, the cohorts are defined by groups of people who are within a range of five years of age. For example, a cohort is defined for those who are between 15 and 19 years old in a particular year. These grouped cohorts are used to increase the number of observations within a cohort, which is recommended to reduce the measurement error bias. In spite of this advantage, in this setting, it is difficult to identify the year and age effects.

In each example, we estimated the returns to schooling by OLS and weighted least squares (WLS). We also calculated the estimator proposed by Devereux (2007a). This last estimator takes into account the measurement error present in the pseudo-panel.

Once we study the behavior of the RTS under different settings and estimators, we analyze the relationship between the RTS and the year, cohort and age effects mentioned above. We found that the RTS are higher for older samples of people. This effect is greater in the cross-sectional estimates than in the pseudo-panel estimates. Therefore, it is important to define a criterion to choose the age range to be used.

Moreover, we found that the income of younger cohorts is greater than the income of older cohorts, once experience and short run fluctuations of the economy are accounted for. This difference in income between generations is explained by differences in their levels of education. Other factors that differ between generations seem to be less important. Finally, we present preliminary evidence suggesting that short run fluctuations on income affect in a greater extent those with less education.

The rest of the paper is divided in three parts. The next section presents the methodology and different estimates of the RTS for Costa Rica. The following section describes the relationship between years of schooling and the year, age and cohort effect. The conclusions are presented in the last section.

2. METHODOLOGY

We used Deaton's (1997) and Deaton & Paxon's(1994) methodology to analyze three effects on earnings: age, year and cohort. We also study their relationship with years of schooling. This methodology is suitable for analyzing Costa Rican household income, given the lack of a panel data set at individual level. Instead of a panel data set, the Costa Rican National Institute of Statistics performs a household survey each year that changes the sample year after year. Deaton's methodology lets us link this set of cross-sectional data through time.

In this methodology -instead of a single individual- a group of people is followed. This group of people -or cohort- is defined by year of birth. For instance, those aged 15 in 1987 make up a cohort; those aged 16 in 1987 make up another cohort and so on. To follow a variable for these cohorts through time, the mean of the variable is calculated for the members of the cohort and this average is linked to the average -of the same variable- of those one year older in the next survey. For instance, the income of the cohort which was 15 in 1987 is represented by the average income of those aged 15 in 1987. One year later (in 1988), the income of this cohort is defined by the average income of those aged 16 in 1988. This lets us construct a data set with observations for cohorts in different years. This type of data array is known as a pseudo-panel or synthetic panel.

Pseudo-panel data has been used to analyze different variables (in particular household income) when panel data is not available[3]. For instances, Deaton and Paxon (1994) constructed a pseudo-panel using 15 consecutive household income and expenditure surveys, of Taiwan. Warunsiri and McNown (2009) used a pseudo-panel to estimate the return to education for Thailand, while Dickerson et al (2001) took a similar approach using data from Brazil.

3 For some applications using pseudo-panels see for example: Browning et al (1985), Gourinchas and Parker(2002), Alegre and Pou(2008) and Propper et al (2001).

As Deaton (1997) points out, pseudo-panels have some advantages over most panels. Cohort data does not suffer from attrition (as do most panels) because they are constructed from new samples every year. Also, they are likely to be less susceptible to measurement error than panel data, because the quantity being tracked is an average and the averaging will, nearly always, reduce the effect of measurement error (Deaton 1997).

To apply this pseudo-panel approach, we follow the Mincerian tradition and assume that earnings are explained by the educational level and experience. Consider the following empirical formulation:

$$(1)\quad \ln(y_{it}^{*}) = \omega + s_{it}^{*}\rho + X_{it}^{*}\beta + v_i + \varepsilon_{it}$$

Where S_{it}^{*} refers to the educational level of the individual i in year t, v_i refers to characteristics of the individual i (as ability), and X_{it}^{*} refers to others income determinants. If v_i was correlated with s_i^{*} the least squares estimation of (1) (without including the individual effect) would capture this correlation, and would be biased. If the covariance between ability and education was positive, the OLS estimation of (1) would overestimate the return to schooling. If the covariance was negative, the OLS estimation of (1) would underestimate the RTS . The covariance is usually assumed to be positive. Individuals with higher ability tend to acquire higher levels of education. Although, Warunsiri and McNown (2009) suggested that this covariance can be negative, since people with higher ability face a higher opportunity cost of schooling. This higher opportunity cost might lead to a negative correlation between ability and schooling.

Taking means over each cohort in (1) we obtain:

$$(2)\ \ln(y_{ct}) = \omega + s_{ct}\rho + X_{ct}^{*}\beta + v_c + \varepsilon_{ct}$$

In (2) the sub-index c represents cohorts. As Kaymak (2008), Warunsiri and McNown (2009) noted, this averaging, eliminates the ability bias, provided that is orthogonal to s_{ct}.

To relate (2) with Deaton and Paxon's(1994) decomposition, we assume that v_c is orthogonal to s_{ct} and that X_{ct}^{*} is composed by age, cohort and year effect. Then, we can rewrite (1) as:

$$(3)\ \ln(y_{it}^{*}) = \omega + s_{it}^{*}\rho + \alpha_{at} + \theta_c + \varphi_t + \varepsilon_{it}$$

and averaging by cohort, we obtain the cohort version of (3):

$$(4)\ \ln(y_{ct}) = \omega + s_{ct}\rho + \alpha_{at} + \theta_c + \varphi_t + \varepsilon_{ct}$$

Where $\ln(y_{ct})$ represents the mean of the logarithm of the income y for those members of the cohort c in the year t; θ_c refers to the cohort effect, α_a refers to the age (a) effect and φ_t refers to the year (t) effect. Moreover, c is defined as the age of the cohort in some particular year t_0. In other words, $c = a - (t\text{-}t_0)$.

In order to identify the three effects -θ_c, α_a and φ_t- Deaton (1997) and Deaton & Paxon(1994) propose to use a set of dummy variables of cohorts, age and years. Given the (linear) relationship between these three variables, the regression is not possible (even when one dummy variable is dropped from each set). To solve this problem, Deaton and Paxon (1994) propose attributing time trends to age and cohort effects, and let the year effect to capture cyclical fluctuations that average

zero in the long run. Thus, the variable that captures the year effect must be orthogonal to trend and sum zero. To accomplish such decomposition, Deaton (1997) suggested regress $\ln(y_{ct})$ on a set of dummies for cohorts (excluding the first), for age (excluding the first) and a set of *T-2* year dummies defined as:

$$(5)\ d_t^* = d_t - [(t-1)d_2 - (t-2)d_1]$$

Where *t=3,..,T* and d_t equals 1 if the year is t and 0 otherwise. Following this approach, we identify the cohort, year and age effects by means of the regression:

$$(6)\ \ln(y_{ct}) = \omega + s_{ct}\rho + A\alpha + C\theta + Y\varphi + \varepsilon_{ct}$$

Where A is a matrix of age dummies, C is a matrix of cohort dummies and Y is a matrix with *T-2* columns each one defined by (5).

Moreover, note that (3) refers to the relationship between income and others covariates in the population. As Deaton (1997, 1985) noted, if (3) refers to the sample, when the average is taken, the cohort effect will vary over time, because in each new sample (survey), there is a different set of individuals. If (3) refers to the population, and we assume that the cohort population is fixed in time, the average will be taken for the same population every year and the cohort fixed effect will not vary over time. Instead of the population version of (4), only the sample data is observed. Then, as Deaton (1985) proposes, we can use the sample data to estimate (4), but the covariates must be viewed as measured with errors.

We assume that year, cohort and age variables do not suffer from measurement error, while years-of-schooling does. Then, when the years of education is included as explanatory variable of the income, a measurement error bias is induced, and a correction must be employed.

Therefore, the use of pseudo-panel data permits us to reduce the ability bias, but it induces a measurement error bias. Deaton (1985) noted that the population means are not observed, but the sample means are an error-ridden estimator, with variances that can be estimated from the data. Then using error-in-variables techniques Deaton(1985) suggested an estimator for the parameters in (4), which is consistent when the number of cohorts tend to infinity. Verbeek and Nijman (1993) showed that Deaton's estimator is biased when T is small (even if the number of cohorts is large). They proposed an estimator which does not suffer from bias due to a small number of sample periods. Devereux (2007a) showed that Deaton's(1985) estimator is biased on finite samples and proposed a different error-in-variables estimator (which is unbiased in small samples). To represent these estimators, let us denote the population variables by y^* , X^* and their correspondent sample values by y , X. Include all the independent variables into the X matrix of (4) and let n_{ct} be the size of cohort c in year t, while K stands for the number of columns of X. As Devereux (2007a,2007b) showed, assuming that

$$\begin{pmatrix} y_{ct} - y_{ct}^* \\ X_{ct} - X \end{pmatrix} \sim IID\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \frac{1}{n_{ct}}\begin{pmatrix} \sigma_{00} & \sigma' \\ \sigma & \Sigma \end{pmatrix}\right)$$

all the above mentioned estimators are represented by:

$$\beta_\gamma = \left(\sum_c \sum_t n_{ct} X_{ct} X'_{ct} - \gamma\hat{\Sigma}\right)^{-1}\left(\sum_c \sum_t n_{ct} X_{ct} y_{ct} - \gamma\hat{\sigma}\right)$$

When γ=0, β_γ, is the WLS estimator; when γ=G=CxT, β_γ, is Deaton's estimator; when $\gamma = \left(\frac{T-1}{T}\right) G$, β_γ , is Verbeek and Nijman's estimator and when γ=G–K–1, β_γ is Devereux's estimator. The last three estimators eliminate the variance due to the measurement error when using the sample estimates of σ and $\sum$. As Verbeek (2008) noted, when the number of observations per cohort tends toward infinity, both σ and $\sum$ tend toward zero, as well as their estimator, then β_γ is asymptotically equivalent to the WLS estimator. In other words, when the size of the cohorts is large enough, the measurement error bias is small and can be ignored. It is for this reason that most empirical studies ignore the measurement error problem and use standard estimators.

There is not a way to know if the size of the cohort is large enough. For example, Verbeek and Nijman (1992) suggest cohort sizes of 100, 200 while Devereux (2007b) suggests that n_{ct} must be as large as 2000. To increase the number of observations in each cohort, Deaton(1997) suggests to redefine cohorts by groups of birth-year, for example, defining a single cohort for those aged between 15 and 19 in t_0 instead of having five different cohorts. In the next section, we use these approaches to estimate (4), using Costa Rican household surveys.

3. THE CASE OF COSTA RICA

Several studies have calculated the returns to schooling in Costa Rica[4]. For instances, Trejos and Gindling (2005) estimate the cross-sectional RTS from 1980 to 1999. This period covers an economic crises and the succeeding recovery of the economy. The recovery is associated with a structural adjustment program. An important part of this program was a comprehensive trade and financial liberalization. As explained in Robbins and Gindling(1999), these reforms induce an increase in the relative demand for skilled workers, increasing the returns to schooling. Our study covers the period from 1987 to 2008. In this period similar results are expected since the process of liberalization is reinforced and a set of economic measures were put in place in order to increase the foreign direct investment. As a result of these measures an important number of foreign firms were established, demanding relative more skilled workers.

Trejos and Gindling (2005) find a reduction of the returns to schooling from 1980 to 1983, associated with the economic crisis. Between 1983 and 1999 the RTS are stable, ranging from 8.5% to 9.3%. Their results for this period are similar to our estimates[5]. Moreover, we find an increase in the returns to schooling at the beginning of the 1990s and 2000s. These results are consistent with the findings in Robbins and Gindling(1999). They argue that trade liberalization in Costa Rica led to an increase in the relative demand for skilled workers, increasing the returns to schooling. Their results are consistent with the "skill-enhance-trade" hypothesis, whereby liberalization increases the physical capital (through imports), which rises the demand for skilled workers. At the beginning of the 1990s the process of trade liberalization is accelerated in Costa Rica, while in the 2000s Costa Rica has experienced an important increase in foreign direct investment, increasing the demand for skilled workers.

To analyze our pseudo-panel approach we use the cross-section estimates of the returns to schooling as a benchmark. These estimators suffer (at least) from two types of bias: an ability bias and an attenuation measurement error bias.

To calculate the return to schooling we use the household surveys collected by the Costa Rican National Institute of Statistics and Census from 1987 to 2008. The sample includes males whose income, hours worked and educational level are observed in each survey. The years of

4 See for example, Funkhouser(1996, 1998, 1999), Psacharopoulos(1992), Trejos and Gindling(2004), Gindling and Trejos (2005).

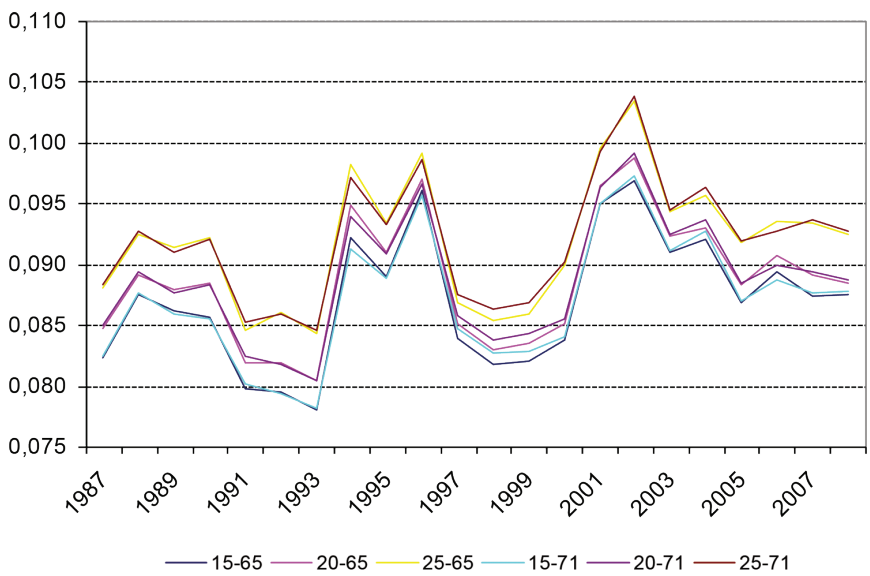
5 See Graph 1.

schooling range from zero (no education) to 19 years. Moreover, we estimate the returns to schooling for different age-groups sets. These sets include those individuals aged between 15-65, 20-65, 25-65, 15-71, 20-71 and 25-71. For each year, we assume that income is a function of age, age square and the level of education, i.e

$$\ln(y_i) = \omega + \alpha_1 age_i + \alpha_2 age_i^2 + \rho s_i + \varepsilon_i$$

where y_i and s_i represent the hourly income and years of schooling of individual i, respectively. The monthly income figure was divided by 4.33 to obtain the weekly income figure and then this figure was divided by the reported weekly hours worked to arrive at an hourly income figure. The hourly income figured was adjusted by the consumer price index of the respective month in which the survey was collected. Graph 1 shows the (cross-sectional) returns to schooling.

GRAPH 1
MINCERIAN RETURNS TO SCHOOLING: DIFFERENT AGE GROUPS, COSTA RICA



Source: Author's calculations. Graph 1 shows the cross-sectional returns to schooling. Each line represents a different age group.

For groups represented in Graph 1, the cross-sectional returns to schooling, on average, range from 8.7% to 9.2%. We observe that the samples which include the younger groups have lower returns. On the other hand, when the oldest group (65-71) is included (keeping the youngest age constant) there is no change in the estimated.

To analyze these results, we estimate the returns to schooling using a pseudo-panel approach. To construct our (first) pseudo-panel, we restrict the sample to those born between 1937 and 1972, i.e to those aged between 15 and 50 in 1987. First, we defined each cohort by a single year-of- birth. This means that, in each new survey, the sample increases in age by one unit. Individuals aged 15 are observed only in 1987 (because they are 16 years old in 1988, 17 years old in 1989 and so on). Individuals aged 16 are observed in 1987 and 1988 only. At the same time, in 1987

we do not observe individuals aged 51 or older while in 1988 we observe individuals aged up to 51 years old, in 1989 up to 52 years old and so on until 2008 when we observe individuals aged 71. In other words, when we fix the cohorts, we use an older sample in the next survey.

Once the cohorts are defined, we take the logarithms of the hourly income and average this figure by cohort. This average identifies the cohort income in the estimations. Similarly, the average years of schooling identifies the level of education of the cohort.

Using these variables, we estimate (4). Note that in the cross-sectional estimates (above), we imposed a polynomial form to the age effect while in this cohort approach we use a set of dummy variables. Moreover, in the pseudo-panel approach the ability bias is reduced, so the cross-sectional estimate of the return to schooling is expected to be higher than the cohort (or pseudo-panel) estimate.

We estimate (4) without including the cohort dummy variables. The income is a function of the years of schooling, age, and year effects only. When we fix the cohort to those aged between 15 and 50 in 1987, the OLS estimate of the return to education is 7.9%. Similar results are found when the cohorts are fixed to those aged between 20 and 50 in 1987 and also for those aged between 25 and 50 in 1987. The returns to education in these cases are 7.8% and 8%, respectively[6]. Note that these RTS are lower than the cross-sectional estimates, and they do not include the cohort dummy variables.

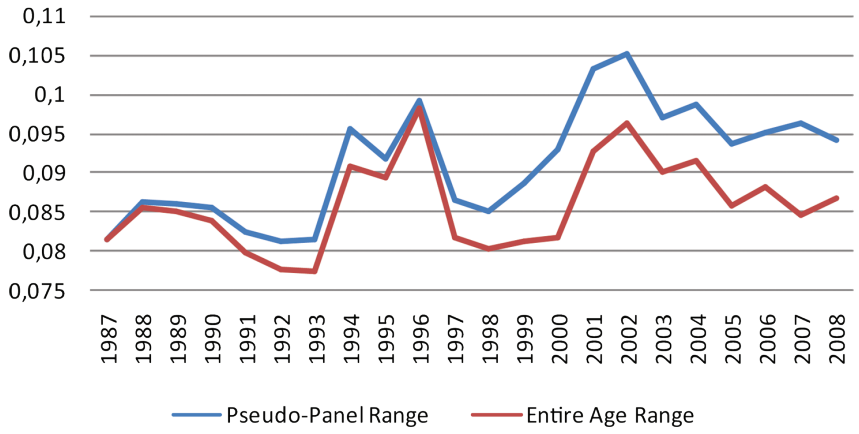
People have different incomes in part due to different levels of education, experience, and other "endogenous decisions"; but their income also differs in part due to the cohort they belong to. Each generation grows up at a different moment in time, experiencing different public policies, different levels of public and private investment, thusly having different opportunities, etc. For example, in Costa Rica, we have observed that younger cohorts have higher levels of education than older cohorts. This phenomenon might be due to a greater public investment in education, an increase in wealth in the younger generations, and an intergenerational change in preferences on education, among others. When the cohort effect is omitted, this covariance between ``the generation'' and the level of schooling is captured in the return to education and therefore it is biased. The fact that younger generations have higher levels of education causes a bias in the estimated return to education, when the cohort effect is not included in the regression.

For this reason, we include the cohort dummy variables in the regression. Our (OLS) pseudo-panel estimate is 9.5% when it is calculated using cohorts aged between 15 and 50 in 1987. The average cross-section estimate when we use those aged between 15 and 50 for every year is 8.6%. The respective estimates when we use cohorts between 20 and 50 are 9.3% and 8.8%; and when we use cohorts between 25 and 50 the estimates are 9.45% and 9.16%. Therefore when we include the cohort effect, the RTS increases, and it is higher than the cross-sectional returns. This can be explained by higher levels of education in succeeding generations.

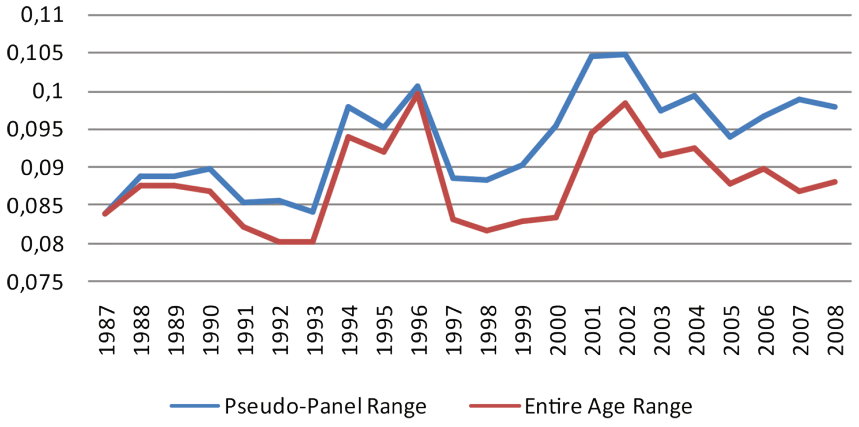
As we have pointed out, the cross-section RTS increases if we use an older sample. In the cohort case, we follow a constant set of cohorts; which means that each new sample includes individuals who are older. As older individuals seem to have higher returns to schooling, the pseudo-panel estimate might be capturing this sample selection. To exemplify this point, each of the Graphs 2, 3 and 4 show the cross-section estimates in two cases (each one represented by one line in the graph). In the first case, the return to schooling is calculated using the completed age range (indicated) in each year. In the other case, for each year we change the sample to include those who appear in the pseudo-panel (i.e the minimum and the maximum age are increased by one unit each year). Two observations are relevant, (1) when moving between graphs we notice that the older samples have higher returns to schooling and (2) when we use people aged between 15 and 50 or those aged between 20 and 50 in 1987, the point estimate difference is getting larger every year.

6 The WLS are similar.

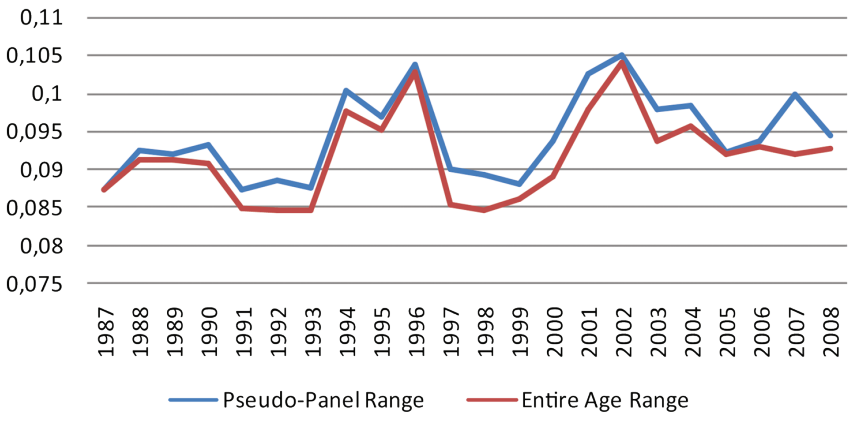
GRAPH 2
MINCERIAN RETURNS TO SCHOOLING: AGE RANGE 15-50



GRAPH 3
MINCERIAN RETURNS TO SCHOOLING: AGE RANGE 20-50



GRAPH 4
MINCERIAN RETURNS TO SCHOOLING: AGE RANGE 25-50



Source: Author's calculations. Graph 1 shows the cross-sectional returns to schooling. Each line represents a different age group.

The (OLS) pseudo-panel estimate is 9.5% when calculated using cohorts aged between 15 and 50 in 1987. The average cross-section estimate, when we use those aged between 15 and 50 every year is 8.6%, but when we use those who appear in the pseudo-panel, the average cross-section estimate is 9.1%. The respective estimates when we use cohorts between 20 and 50 are 9.3%, 8.8% and 9.3% and when we use cohorts between 25 and 50 are 9.45%, 9.16% and 9.43%. The average cross-sectional estimate is closer to the pseudo-panel estimate when the same individuals are used for both estimations. The above observations suggest that, when older samples are used, the estimate returns are higher.

Given this relationship between the sample and the cross-sectional estimates, we change our pseudo-panel data set to maintain the same age range every year. Instead of keeping cohorts fixed, we fix the age range. This means that, the pseudo-panel is formed of different cohorts each year because we have to drop the oldest cohort and we have to include a new cohort each year.

We use three age ranges; 15-65, 20-65, 25-65, and we run two sets of estimations with them. One set without including the cohort dummy variables and one set including these dummies. When excluding the dummy variables, the respective OLS estimates for the three age groups are 8.32%, 8.30% and 8.27% (and the WLS estimates are 8.45%, 8.41% and 8.3%). When we include the cohort dummy variables, the OLS estimates for those three age ranges are: 9.2%, 9.5% and 9.5% (The WLS estimates are 9.2%, 9.7% and 9.7%).

These returns are higher than the average cross-sectional return (see Graph 1) and are close to the pseudo-panel estimates when the cohorts are fixed. Then, in spite of the higher cross-section return when using older samples the two pseudo-panel estimates are close.

The next table summarizes the returns to schooling for all the cases we have analyzed.

TABLE 1
RETURNS TO SCHOOLING OLS ESTIMATES

	Pseudo-Panel Approach				Cross-Sectional Approach	
	Excluding Cohort		*Including Cohort*		*(Average 1987-2008)*	
Age/Cohort Group	Fixed Cohort	Fixed Age	Fixed Cohort	Fixed Age	Psdo-Pan Age Range	Entire Age Range
15-50	7.9		9.5		9.1	8.6
20-50	7.9		9.4		9.3	8.8
25-50	8.0		9.5		9.4	9.2
15-65		8.3		9.2		8.7
20-65		8.3		9.5		8.9
25-65		8.3		9.5		9.2
15-72						8.7
20-72						8.9
25-72						9.2

Source: Author's calculations.

We observe that the smallest as well as the highest returns to schooling (RTS) are found when the pseudo-panel approach is used. The smallest estimates of the RTS are found when the cohort dummy variables are not included in the regression, and the highest estimates are found when they are included. Then, the difference between the smallest and highest estimates is due to the correlation between the years of schooling and the cohort effect. In other words, in Costa Rica, the fact that younger generations tend to have higher average years of education, might bias the return to schooling when this inter-generational relationship is not account for.

We have found that:

a) Samples of older people have shown slightly higher returns to schooling. This increase in the RTS is greater when the cross-sectional approach is used than when the cohort approach is used. Therefore, it is important to define a criterion to choose the age range to be included, particularly when using the cross-sectional approach.
b) When a fixed group of cohorts are followed through different surveys, the sample used each year gets older and therefore younger individuals are not taken into account when the return is calculated. This can induce a selection bias, although, in Costa Rica, the estimate returns are very similar to those found when the age range is fixed.
c) The returns to education are higher when the cohort dummy variables are included in the regression. This is evidence of an ``important'' correlation between the inter-generational process and the years of education. In Costa Rica, younger cohorts have higher average years of education than older cohorts. This correlation might bias the estimated returns to education when this inter-generational process is not taken into account.

We are going to fix the age range. Therefore we have to carefully select the age range and deal with the measurement error bias. In the next subsection, we will focus on this measurement error bias and we will examine the selection of the age range later.

4. DEALING WITH SAMPLING ERRORS

To deal with the measurement error bias, errors-in-variable techniques had been applied in Deaton (1985), Verbeek & Nijman (1993) and Devereux (2007a) to find asymptotically consistent estimators in the case of pseudo-panels. Deaton (1985) noted that the population cohort means are not observed, but the sample means are error-ridden estimators, with variances that can be estimated from the survey data. Then, he proposed an estimator for the parameters in (4), which is consistent when the number of cohorts (C) tends to infinity. Verbeek and Nijman (1993) showed that Deaton's estimator is biased when T is small. Therefore they proposed an estimator which does not suffer from bias due to a small number of sample periods. Devereux (2007a) also showed that Deaton's (1985) estimator is biased in finite samples and proposed another error-in-variables estimator (which is unbiased in small samples).

When we include people aged between 15 and 65 the OLS and WLS estimators of the return to schooling are both 9.2%, while the Devereux estimator is 10.2%. When people aged between 20 and 65 are selected, the OLS estimator is 9.5%, the WLS is 9.7% and the Devereux estimator is 12.0%. In the case when the age sample range from 25 to 65, the OLS estimator and WLS are similar to those in the previous case (9.5% and 9.7%, respectively) while the Devereux estimator is 11.3%. As expected, the Devereux estimator is higher.

On the other hand, for the reasons mentioned above, some empirical studies ignore this measurement-error-bias correction when pseudo-panel data are used[7]. In general, when the cohort size is large, the measurement error bias can be small enough to be ignored. Verbeek and Nijman (1992) suggested that a cohort size of 100-200 might be large enough. Our three pseudo-panels do not fulfill this (empirical) requirement. The cohort size, in the first pseudo-panel (aged 15-65), ranges from 19 to 378. Moreover, 28% of the cohort-age observations have a cohort size smaller than 100. In the second pseudo-panel (aged 20-65) 29% of the observations have a cohort size smaller than 100 and in the third data set (aged 25-65) 33% have a "small'' cohort size.

Given that the cohort sizes are small (particularly for the older cohorts) we (also) defined cohorts by 5 year-of-birth groups, fixing the age range at 15-65. In this case, the minimum cohort size is 185. The OLS return to schooling is 11.35%, the WLS is 11.07% and the Devereux is 14.53%.

Even in this setting, the cohort size might be small. Devereux (2007b) showed that having 100 or 200 observations per group might not be enough to ignore the bias[8]. Devereux (2007b) also showed that when the cohort size is too small, the bias-corrected estimators might be extremely variable and "identification may be sufficiently poor that no estimator provides reliable estimates" (Devereux, 2007b).

Although the problem of small cohort size might be solved by grouping cohorts, this construction makes it difficult to isolate the year effect from the age effect. Note that in the sample, an age-year pattern emerges when we define cohorts by 5 year-birth groups. We divide the age range 15-65 into groups, each group having 5 ages. Since the age of the group is defined as the average age, in a single year, we observed just 9 or 10 ages, and those ages are spaced by 5 years. For example, if age 17 is observed in a given year, the ages 22, 27, 32, 37, 42, 47, 52, 57 and 62 will be observed as well. Each age is observed only 4 or 5 times as we observed a given age only every five years.

Then, in a given year, we observed ages spaced by 5 years and we observed a particular age only every five years. This structure makes it difficult to split the age and year effects. We will illustrate this in the next section.

Age, Year and Cohort Effects: Choosing the Cohort Setting and the Age Range

In this section we will analyze Deaton's income decomposition and its relationship with the years of schooling. Our first step is to select the pseudo-panel setting and the age range. Single birth-year cohorts or 5 birth-year cohorts can be used. The former have a smaller cohort size than the latter. A larger cohort size lessens the measurement error bias so a 5 birth-year cohort might be preferred. On the other hand, as we noted above, this cohort setting imposes a particular structure on the ages observed each year[9]. In spite of that, we apply Deaton and Paxon's (1994) decomposition to the income using 5 birth-year cohorts.

Graph 5 shows the age effect using these 5 birth-year cohorts, when the age range is fixed between 15 and 65 and the WLS estimator is used[10].

---

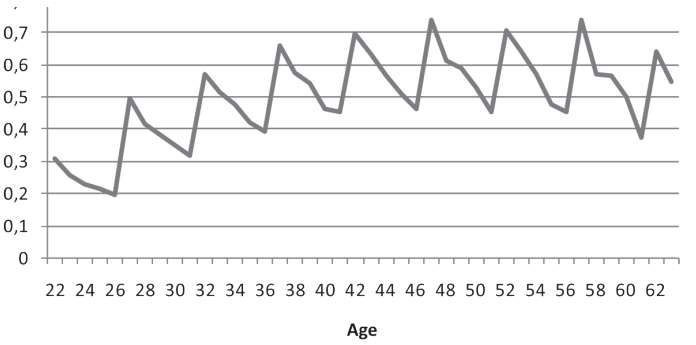
7 See for example Warunsi and McNown(2009), Dickerson et al (2001), Browning et al (1985) and Alegre and Pou(2008).

8 Although he recognizes that the required number of observations depends on the specific application.

9 In a given year only some ages are observed, and a given age is observed only every five years.

10 In this case the years of schooling are not part of the decomposition, and then the WLS estimator equals the Devereux estimator, as long as we assume that age, cohort and year do not present measurement error.

GRAPH 5
INCOME AGE EFFECT: 5 BIRTH-YEAR COHORTS



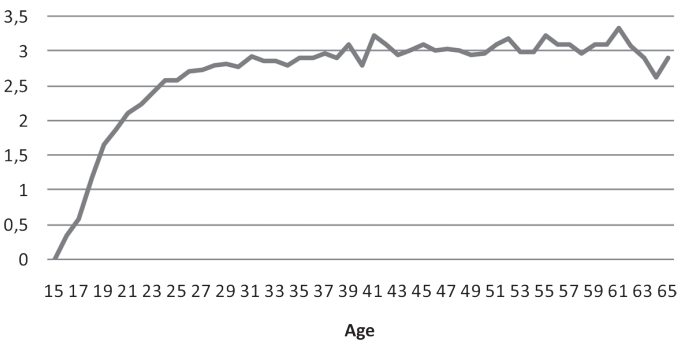
Source: Author's calculations.

We observe in Graph 5 that the age effect has a cyclical behavior. A trough appears when the age is 26 and every five years afterwards. As we noted above, all those ages share the same set of years in which they are observed. In this case, the troughs are observed in 1987, 1992, 1997, 2002 and 2007. Costa Rica's GDP growth shows a similar pattern. Costa Rica's GDP growth has troughs, particularly in 1991, 1996 and 2001; it seems that these troughs occur one year before a trough occurs in the age effect. Thus, the cyclical behavior of the age effect seems to be related to these cyclical fluctuations in the GDP growth[11]. Because we want to isolate these cyclical fluctuations in the year effect this panel setting is not used. Instead we define cohorts by a single birth-year. As mention above, this set has a small cohort size, and there is the possibility of measurement error bias in our estimation of the return to schooling.

Once the pseudo-panel setting is selected, we must choose the age range. As mentioned above, the selection of this range is important given the higher returns are observed in older samples. To select the minimum age, the same age, cohort and year decomposition is applied to the years of schooling. The age effect is presented in Graph 6. This effect flattens out at age 24. Since the retirement age in Costa Rica is 65, the minimum age chosen is 24 and the maximum age is 65.

11 This seems to be due to the structure imposed on the pseudo-panel when we group cohorts. In the case of Costa Rica, there is a match between this imposed structure and the behavior of GDP growth.

GRAPH 6
YEARS OF SCHOOLING: YEAR EFFECT



Source: Author's calculations.

We choose as the initial age, the point where the age effect for the years of schooling flattens because we do not want to include (in average) people who are working and, at the same time, acquiring more years of schooling. Thus, we fix the age range between 24 and 65 in each survey and we define the cohorts by a single birth-year. Using these parameters, in the next section, we analyze the cohort, age, year effect and their relationship with the years of schooling. But, as the cohort size is small, the results might suffer from an attenuation bias due to a measurement error in the years of schooling.

Cohort, age and Year Effect and their relationship with the years of schooling

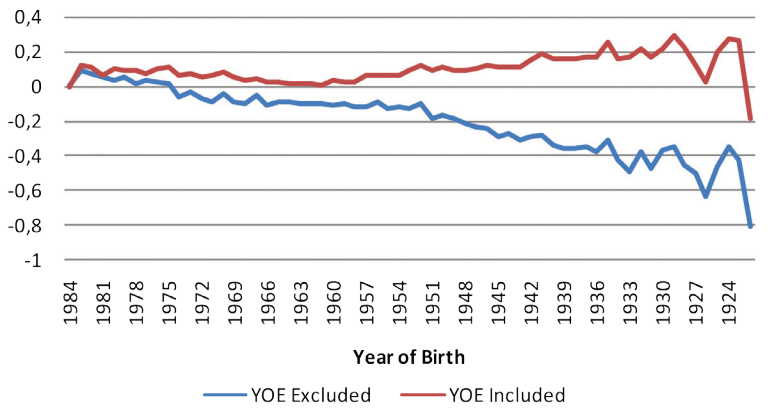
In this section Deaton's age, year and cohort effects are calculated. These effects are calculated including and excluding the years of education (YOE) from the regression. When the years of education are included, the WLS return to schooling is 9.66% and the Devereux estimator is 11.5%. To calculate these effects the age range is fixed between 24 and 65 and the cohorts are defined by a single birth-year. Then, the pseudo-panel has 924 observations. The dummy variables for the youngest cohort (those born in 1984) and for age 24 are dropped.

Graph 7 shows the cohort effect. The cohort effect reflects those factors that are particular to the cohort (i.e fixed in time); i.e. those that are exogenous to the cohort, but endogenous to the inter-generational change. These factors include changes in technology, improvements in education quality, different public policies in each generation, initial wealth, etc.

Graph 7 shows that (at the same age) younger cohorts earn more than older cohorts when only age and year effects are accounted for. For example, at the same age, the cohort born in 1984 earns 40.5% more than those born in 1940, 30.77% more than those born in 1944 and 9% more than those born in 1964. This effect flattens out for those born between 1957 and 1965. These cohorts include those who were between 15 and 23 years old in 1980. This particular group was negative affected by the economic disruptions that occurred in Costa Rica in 1980. In that year, Costa Rica suffered a balance of payment crisis and public expenditures were reduced including those for schooling. For example, Montiel, Ulate, Peralta and Trejos (1997:43) stated that during the 1980 crisis high school was the educational level most affected by the reduction in public expenditures. Even in 1996 (when their article was written) real public expenditures for education

had not reached their level in 1980. This reduction in public expenditures did not affect the cohorts older than 23 in 1980 because between 1950 and 1979 the expenditures of the Ministry of Public Education were increasing as a percentage of GDP (Montiel, Ulate, Peralta and Trejos, 1997:21-22).

GRAPH 7
COHORT EFFECT



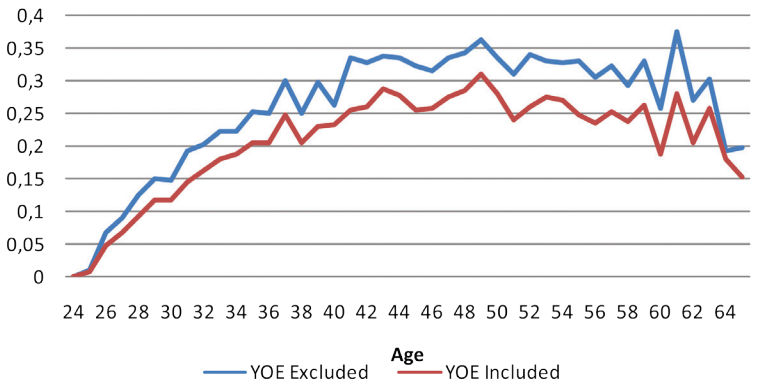
Source: Author's calculation.

Graph 7 also shows the cohort effect when the average educational level is included in the regression. When the YOE are included, a different pattern emerges: the cohort effect flattens out. This means that, once education is accounted for, there is no evidence of a difference in income between cohorts. This change in the cohort effect pattern reflects the correlation between the YOE and the cohort effect. As mention above, in Costa Rica, younger cohorts have higher average YOE than older ones.

We find that younger cohorts earn more than older cohorts when only age and year effects are accounted for. When education is taken into account, there seems to be no differences in income between younger and older cohorts. We conclude that an important part of the difference in income between generations of Costa Ricans can be explained by the higher average YOE that new generations possess.

Graph 8 shows the age effect. This graph shows that there is an increasing income before individual turned 41. After age 41, the age effect on income flattens out. When the years of schooling are included in the regression, the increase in the age effect curve is less steep. When, education is excluded, an individual that is 41 years old earns 40% more than when he was 24. When education is account for, this percentage falls to 29%.

GRAPH 8
AGE EFFECT

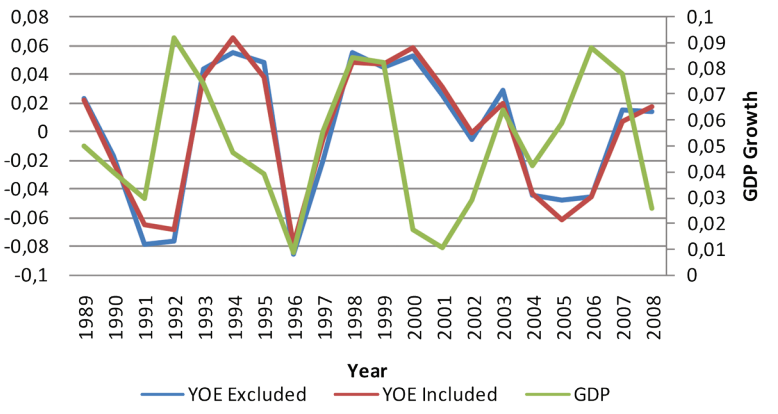


Source: Author's calculations.

Graph 9 shows the year effect. This graph includes three lines. One plots annual GDP growth rates, based on data from the Central Bank of Costa Rica and the World Bank's World Development Indicators (WDI). The Central Bank data is used after 1992 and the WDI data before 1992. The Central Bank of Costa Rica does not have GDP growth data before 1992 because they did a revision of the national accounts that year. Therefore, the data before and after are not comparable.

The other two lines in Graph 9 represent the estimated temporary shocks to income by year. One line is a plot of the year effect without accounting for the average years of schooling. The other line corresponds to the year effect after accounting for education. These two lines showing temporarily shocks to the economy demonstrates how the year effect mimics fluctuations in the economy. This is evidence of the flexibility of income in the face of temporary shocks.

GRAPH 9
YEAR EFFECT



Source: Author's calculations.

Moreover, the difference between the year effect including and excluding the YOE is small. If this difference is interpreted as an omitted variable bias, it could be explained by the correlation between the YOE and the year effect. When the YOE is decomposed into the three effects, for those aged between 24 and 65, the year effect is not significant.

This result might be expected. Assume that the sample does not change with the economic fluctuations and is composed of those who have finished their education. Then, when the economy experiences a slowdown this group cannot change their YOE. This gives us a small correlation between the YOE and the year dummy variables, and a small difference between the year effect including and excluding the YOE. On the other hand, if a slowdown in economic activity tends to affect, for example, those with less YOE, the year effect could be correlated with the YOE. If a slowdown in economic activity causes an increase in the unemployment rate of those with less YOE, then, those who are working would have a higher average YOE, and the correlation between YOE and the year dummy variables would be higher.

As preliminary supporting evidence of this hypothesis, Table 2 shows the employment rate by educational levels, from 1990 to 2007. The employment rate tends to behave like the GDP growth rate for those individuals that do not have a formal education and, to a lesser degree, for those who have only completed the primary level. The employment rate for the other groups as measure by educational level has a small correlation with GDP growth.

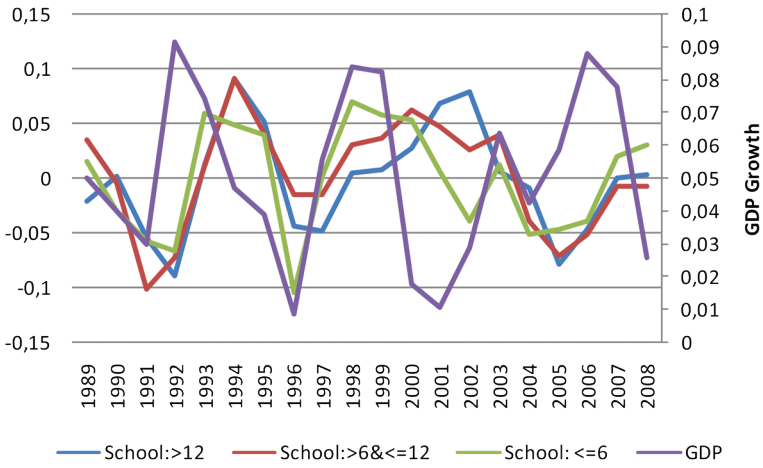
TABLE 2
MALES OCCUPATIONAL RATE BY EDUCATIONAL LEVELS: COSTA RICA 1990-2007

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Males	73.8	71.3	71.5	72.3	72.7	72.5	69.7	70.9	72.3	71.4	69.6	69.8	69.1	69.0	69.0	70.2	70.2	70.8
No Education	60.4	56.8	56.7	55.6	53.7	58.5	54.4	53.1	57.1	52.6	50.0	55.6	55.7	54.4	55.1	51.3	49.5	53.7
Primary Incomplete	67.5	63.0	63.2	63.4	66.5	61.7	60.7	60.8	60.1	60.4	58.0	57.2	55.9	55.8	56.1	58.3	56.8	54.7
Primary Complete	82.0	79.5	79.9	80.7	81.3	81.1	76.9	79.2	80.0	78.4	75.9	77.4	75.6	75.3	75.9	76.7	77.5	77.5
Secondary Incomplete	63.6	62.5	62.1	62.5	62.8	62.9	62.8	63.0	64.3	62.5	62.4	59.2	60.0	59.3	58.1	60.5	61.1	62.3
Secondary Complete	85.1	85.4	84.7	85.1	83.4	83.2	80.9	82.2	84.6	87.2	82.2	83.5	81.4	82.5	81.1	83.6	82.4	81.3
Tertiary Incomplete	71.2	71.8	68.3	73.0	75.6	77.4	67.4	72.1	77.4	73.9	72.9	74.1	73.1	71.3	72.7	72.4	72.5	77.1
Tertiary Completed	84.4	85.9	81.9	84.6	82.0	84.0	83.3	84.4	85.8	87.0	86.8	86.1	88.4	86.9	88.6	87.9	90.1	89.8
Ignored	81.2	74.5	73.8	83.9	77.0	82.4	74.7	85.5	81.3	82.7	85.5	88.0	79.2	85.0	82.4	80.3	82.6	91.3

Source: Costa Rican Household Surveys.

Although the employment rate does not seem to move with GDP growth, the level of income does, as shown in Graph 9. To analyze this co-movement for the different educational levels, we calculated Deaton's decomposition for three groups. The first group includes those with six or less years of education, the second includes those with more than six YOE but less than 12, and the third group includes those with more than 12 YOE. The year effects for these groups are shown in Graph 10.

GRAPH 10
YEAR EFFECTS: DIFFERENT YOE GROUPS



Source: Author's calculations.

The year effect for those with six or less YOE behaves closer to the GDP growth than the year effect of those in the other two groups[12]. This suggests that economic fluctuations are associated with variations in the income of those with less YEO.

In addition to the measurement error and ability bias refer to above, our first set of estimations suffer from selection bias. When young individuals were included in the sample, we did not take into consideration that an important part of this group had decided to go to school full time and not to work. This decision, between working and studying, causes a selection bias. Something similar occurs when the groups of cohort are fixed. When the cohorts are fixed, the pseudo-panel includes those aged between 15 and 50 in 1987. Those aged 50 in 1987 were 71 years old in 2008. The members of this cohort can retire and do not participate in the labor market. Then, the inclusion of older cohorts can induce a ``retirement'' bias.

12 Moreover, the year effect, for those with more than 12 YOE, moves in a countercyclical way after 1999. Although, for this group, the year dummy variables are not (jointly) significant.

5. CONCLUSIONS

In this study the return to education in Costa Rica was calculated using a pseudo-panel approach. To apply this method, different cohort settings were used. Specifically, three settings were examined. In the first case a fixed group of cohorts where used, while in the second a case, a fixed age range were used. In the first case, the sample used is older every year, while in the second case; the age range is kept constant. Although the cross-sectional estimates show higher returns for older samples, when the pseudo-panel approach is applied, the returns are similar. This means that the pseudo-panel estimates are more robust to changes in the age range of the sample than the cross-sectional estimates.

In the third case, cohorts were defined by 5 birth-year groups. This setting increases the cohort size, reducing the measurement error bias, but it makes difficult to split the year and age effects. In this setting, the same ages are observed in a specific year, and each age is only observed every five years. In Costa Rica, GDP growth has a trough every five years (between 1991 and 2001), therefore the age effect shows a cyclical behavior that should be attributed to the year effect. Then, this setting, while ameliorate the measurement error bias, makes it difficult to isolate the year effect from the age effect.

On the other hand, the pseudo-panel return seems to be smaller than the average cross-section return when the cohort dummy variables are not included in the regression, i.e when the cohort effect is omitted. When the cohort effect is included, the return to schooling calculated using the pseudo-panel approach is higher than the average (estimated) cross-section return. This can be explained by the covariance between the average years of education and the cohort effect. In Costa Rica, succeeding cohorts, have higher levels of education.

This relationship between cohorts and years of schooling can also be viewed when the decomposition of Deaton and Paxon's (1994) is performed. When this decomposition only includes year, age and cohort dummy variables it seems that younger cohorts earn more than older cohorts. When the YOE is included in the regression, the cohort effects flattens out. Thus, the differences between incomes of different cohorts seem to be explained by the higher average levels of education of the younger cohorts.

On the other hand, the year effect suggests that economic fluctuations (in the period analyzed) tend to be more related to variations of income than to variations in the employment rate. Preliminary evidence suggests that the employment rate and income of those with less education moves with GDP growth. Thus, people with more years of education experience less fluctuation of their income along the business cycle.

6. REFERENCES

Alegre J. and L. Pou. 2008. ``El Consumo y la Tasa de Ahorro Privados de los Hogares Españoles: Una Descomposición de los Efectos Edad y Cohorte››. Investigaciones Económicas, XXXII(1):87-121.

Card D. 1994. ``Earnings, Schooling and Ability Revisited››. NBER Working Paper Series No. 4832.

Browning M.; A. Deaton. and Margaret I. 1985. ``A Profitable Approach to Labor Supply and Commodity Demands Over The Life-Cycle››. Econometrica, 53(3):503-543.

Deaton A. 1985. ``Panel Data from Times Series of Cross-Sections››. Journal of Econometrics, 30: 109-126.

Deaton A. 1997. The Analysis of Household Surveys: A Microeconometric Approach. World Bank and John Hopkins University Press.

Deaton A and Christina P. 1994. ``Saving, Growth, and Aging in Taiwan›› in David A. Wise, ed., Studies in the Economics of Aging, Chicago University Press for the National Bureau of Economic Research.

Devereux P. 2007a. ``Improved Error-in-Variables Estimators for Grouped Data››. Journal of Business & Economic Statistics, 25(3):278-287.

Devereux P. 2007b. ``Small-Sample Bias in Synthetic Cohort Models of Labor Supply››. Journal of Applied Econometrics, 22: 839-848.

Dickerson A.; F. Green and J. Saba. 2001. ``Trade Liberalization and Returns to Education: A Pseudo-Panel Approach››. Studies in Economics No 0114, Department of Economics, University of Kent.

Funkhouser E. 1999. ``Cyclical Economic Conditions and School Attendance in Costa Rica'', Economics of Education Review, 14:31-50.

Funkhouser E. 1998. ``Changes in the Returns to Education in Costa Rica", Journal of Development Economics, 57:289-317.

Funkhouser E. 1996. ``The Urban Informal Sector in Central America: Household Survey Evidence", World Development, 24(11):1737-1751.

Gindling T.H and J.D Trejos. 2005. ``Accounting for Changing Earnings Inequality in Costa Rica, 1980-99'', The Journal of Development Studies, 41(5): 898-926.

Gourinchas P. and J. Parker. 2002. ``Consumption over the Life Cycle››. Econometrica, 70(1): 47-89.

Kaymak B. 2008.``Enssays in Education, Selection and Wage Dynamics››. Doctoral Thesis, University of Rochester, New York

Montiel N; A. Ulate; L. Peralta y J.D Trejos. 1997. ``La educación en Costa Rica: Un solo sistema?››. IICE, Serie de Divulgación Económica No 28. Febrero.

Propper C.; H. Rees and K. Green. 2001.``The Demand for Private Medical Insurance in the UK: A Cohort Analysis››. The Economic Journal, 111(471), Conference Papers, C180-C200.

Psacharopoulos G. 1993. ``Returns to Investment in Education: A global Update››. Policy Research Working Paper WPS1067, Latin America and the Caribbean Region, The World Bank.

Psacharopoulos G and Y. Ng. 1992. ``Earnings and Education in Latin America: Assessing Priorities for Schooling Investments››. Policy Research Working Paper WPS1056, Latin America and the Caribbean Region, The World Bank.

Psacharopoulos G. and H. Patrinos. 2004.``Returns to Investment in Education: a Further Update››. Education Economics, 12(2): 111-136.

Robbins D. and T.H. Gindling. 1999. ``Trade Liberalization and the Relative Wages for More-Skilled Workers in Costa Rica", Review of Development Economics, 3(2):140-154.

Trejos J.D and T.H Gindling. 2004. ``Inequality in Central America in the 1990s", CEPAL Review, 84: 175-196.

Verbeek M.2008.``Pseudo-Panels and Repeated Cross-Sections››, in L. Matyas and P. Sevestre, eds, ``The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice››, Third Edition, Springer, Berlin, 370-383.

Verbeek M and T. Nijman. 1993. ``Minimum MSE Estimation of a Regression Model with Fixed Effects from a Series of Cross-Sections››. Journal of Econometrics, 59: 125-136.

Verbeek M and T. Nijman. 1992. ``Can Cohort Data be Treated as Genuine Panel Data››. Empirical Economics,17:9-23.

Warunsiri S. and R. McNown.2009. ``The Return to Education in Thailand: A Pseudo-Panel Approach››. IBS Working Paper University of Colorado at Boulder.