The Effect of Fertility Decline on Children's Education in Costa Rica

Jing Li, UC-Berkeley
William H. Dow, UC-Berkeley
Luis Rosero-Bixby, UC-Berkeley and University of Costa Rica

Draft: September, 2013

Abstract

Costa Rica experienced a dramatic fertility decline in the 1960s and 1970s. In this paper we explore one dimension of the potential demographic dividend from this shift: the extent to which it was accompanied by quantity-quality tradeoffs leading to higher educational attainment. Specifically, we estimate the increase in secondary school attendance among children as the number of siblings decreases. We use Census data as well as survey data from the Costa Rican Longevity and Healthy Aging Study (CRELES). To address endogenous family size the analysis uses an instrumental variable strategy based on gender of the first two children, which we find significantly predicts total fertility in Costa Rican families. Results indicate that decreasing fertility strongly increases educational attainment, particularly among girls. Simulations suggest that declining fertility can statistically explain much of the rapid increase in Costa Rican secondary school attainment in the second half of the 20th century.

1. Introduction

Fertility transition and education attainment in Costa Rica

Costa Rica experienced one of the earliest and fastest fertility transitions in the developing world. The total fertility rate (TFR) fell from 7.3 to 3.7 births per woman between 1960 and 1976 (Figure 1). Only the fertility shifts in Singapore and Taiwan were faster than that observed in Costa Rica in this period (Coale 1983). After this extraordinary period, TFR continued declining although at a slower pace, reaching below-replacement levels in 2001. The most recent estimate shows a TFR of 1.83 in 2010, which is a lower rate than in the United States (1.9 births) and the lowest rate in Latin America, after Cuba which has a TFR of 1.7 births (PRB 2012).

[Insert Figure 1 here]

Period TFR is just a theoretical demographic construct that one should not expect to be immediately related to education attainment. Two mechanisms are postulated as transmission belts from fertility to children's education: (1) cohort size and (2) family size, or more precisely sibship size, which is family size from the perspective of children (Lam and Marteleto 2008). Reduced cohort size would allow public education programs to broaden coverage and to improve its quality by reducing the number of students per classroom. Reduced family size would allow parents to allocate relatively larger amount of resources on education for each child. Higher investment in human capital resulting from fertility decline is considered an important component of the so-called demographic dividend (Lee & Mason 2009).

As shown in Figure 1, cohort size and sibship size trends do not correspond exactly to TFR trends. The fertility decline ended the baby boom that was taking place in Costa Rica in the 1950s and resulted in a baby bust from 1963 to 1973. From 1975 to 1985 a second baby boom takes place as an echo of the boom of the 1950s because of the rapid growth of population in reproductive ages. In turn, sibship size is substantially larger than the TFR in the birth year of each cohort, it starts to decline for children born several years before the fall in TFR, and it declines at a more regular pace than TFR.

Elementary education (six grades) is, by law, mandatory and free in Costa Rica since 1869 (Salazar 2003). More than 90% of children complete elementary school. Studies have shown that the key factor of school attainment in Costa Rica is the dropping rate at the first year of secondary school, i.e. at ages 13 to 15 years (Estado de la Nación 2005). The proportion with some secondary education is thus an important indicator of education attainment in this country. This proportion, according to census data, has increased from 40% for those born around 1950 to 80% for those born in the 1990s (Figure 1). This progress to some extent mirrors the curve showing the decline in sibship size, except for cohorts born in the 1960s whose educational progress was interrupted by the severe economic crisis that occurred in Costa Rica in the early 1980s that made families to put their children to work and reduced government expenditures in education.

Understanding the factors that make adolescents to continue in school after seventh grade is central to policies to improve education in Costa Rica. Several studies of this topic have singled out, in addition to socioeconomic constrains, low motivation to attend school derived from lack of family support, often coupled to low education of parents, and deficiencies in the educational system (Estado de la Nación 2005). No Costa Rican study, however, mentions large family size as an obstacle for educational attainment or singles out the fertility transition as a contributing factor for the improvement of education levels in younger cohorts. We want to amend this omission by assessing the contribution made by the rapid fertility transition of Costa Rica to the improvement in education of children, focusing on the sibship effect. Documenting the contribution of fertility decline to improve education is important for understanding a mechanism of the demographic dividend and also for sounded education public policy.

Sibling number and children's education

There is a large literature examining the relationship between family size and children's educational outcomes. Many of these studies draw upon the quantity-quality (Q-Q) trade-off model (Becker and Lewis 1973) which suggests that decreased number of children in the family allows more resources to be allocated to each child which in turn increases child quality. Early empirical research supporting this theory generally finds a negative relationship between sibling size and children's educational attainment across different countries and cultural contexts (Blake 1981; Hanushek 1992; Knodel and Wongsith 1991; Rosenzweig and Wolpin 1980).

However, more recent literature on this topic has yielded mixed results. Studies have pointed out that since parents jointly determine both child quantity and quality, i.e. how many children to have and how much to invest in them, these two variables are both affected by unobserved parental preferences and other family characteristics. As a result, association between family size and children's education outcomes does not establish a causal relationship (Angrist et al. 2005). Accordingly, some studies that use more sophisticated instrumental variable (IV) approach that attempts to isolate exogenous variation in family size have found little to no relationship between family size and children's education (Angrist et al. 2005; Black et al. 2005; Caceres 2004). For instance, Black et al. (2005) use twin births as an instrument to examine the effects of family size on children's education in Norway and find the effect to be negligible. This is also the case when they control for birth order. Angrist et al. (2005) also use twin-birth as well as sibling sex-composition as instruments for family size in Israel and find no evidence of a Q-Q trade-off.

There are two important qualifications in interpreting the findings of these studies. First, these studies often attempt to purge the effect of birth order from the relationship between sibling size and education by either controlling for birth order directly (Black et al. 2005) or examining outcomes of only the first or second births (Angrist et al. 2005). However, birth order could operate as an important mechanism by which sibling size affects children's schooling if younger children in the family suffer from a disadvantage compared to their elder siblings who consumed most of the family resources. The *average* children's schooling decreases as a result and this is a consequence that policy makers are concerned about. Therefore it is debatable whether it makes sense to rule out birth order effect when empirically examining the Q-Q tradeoff.

Second, studies that find no evidence of such tradeoff almost all used data from developed countries with more comprehensive welfare system, whereas Q-Q tradeoff could be more prominent in developing countries where social resources for education are more limited (Li et al. 2008). Consistent with this argument, recent studies that use both data from developing countries and IV approach tend to confirm the negative relationship discovered between sibling size and children's schooling. Utilizing the cultural preference for sons in South Korea, Lee (2008) instruments sibling size by sex of the first child and finds adverse effect of sibling size on per-child investment in education. Similarly, Jensen (2005) also uses the sex of the first two births as instrument for number of siblings in India and shows that the number of siblings helps explain gender inequality in education as girls tend to have more siblings than boys because of son preference. In addition, Li et al. (2008) instrument family size by twin birth in examining the effect of family size on education attainment in China. Again, they find a negative effect of family size on children's education which is more evident in rural China with poor public education system.

Our study adds to the large literature in testing the Q-Q tradeoff by examining the relationship between family size and children's education in another developing country: Costa Rica.

2. Data and methods

Data

We use two data sources to examine the relationship between sibling number and children's education in Costa Rica. The first dataset comes from the Costa Rican Longevity and Health Aging Study (CRELES, or "Costa Rica Estudio de Longevidad y Envejecimiento Saludable"). It is a set of nationally representative longitudinal surveys of health and life course experiences of older Costa Ricans (Berkeley Population Center 2012). The CRELES data contain detailed demographic information of the elderly as well as their spouses and children and thus well suited for the purpose of this study. It currently consists of two birth cohorts (pre-1945 and 1945-1955) and multiple waves in each cohort. In order to ensure adequate sample size and also to cover the entire period of fertility decline, we use data from the first two waves of the pre-1945 cohort, collected in

2005 and 2007 respectively, combined with data from the first wave of the 1945-1955 cohort which was collected in 2010.

We derive our analytical sample by mainly applying the following restrictions to the CRELES data:

1) Families have at least two children, and all children in the same family have the same biological mother who is identified in the survey. This restriction is required by the IV approach we adopt. Because we use the sex composition of the mother's first two births as instruments for fertility, we need to be able to identify the sex of the two eldest children in the family.

2) Mothers were born between 1930 and 1960. This age range limit ensures that the mothers in the sample gave birth during the period of fertility decline.

3) The key variables including education, age etc. of mothers and children are non-missing.

4) Children are of age 13 or above. The restriction allows us to examine the effect of secondary school attendance for age-appropriate children, since children normally start secondary school at age 13 in Costa Rica.

After applying the above restrictions, the final analytical sample contains 9,322 children of 2,045 mothers.

One concern with the CRELES dataset is that it may not provide enough power to detect the relationship we propose. Although the sample appears to contain a sufficient number of children, it only has about 2,000 mothers (families) which is not a particularly large sample size. This may pose an issue as our key independent variable is sibling size which only varies by family.

We therefore perform our statistical analysis on an additional dataset derived from a 10% sample of the 1984 Costa Rica Census data. The Census data were collected by the National Institute of Statistics and Censuses of Costa Rica and are available from the University of Minnesota Population Center's IPUMS international web dissemination system. Apply the same restrictions as above to the Census data yields a sample of 47,336 children and 17,827 mothers. However, a major limitation of the Census data for the purpose of this study is that it only links children with their mothers if they live in the same household as their mothers at the time of the Census. We therefore restricted our sample to only families with at least two children and in which all children were living in the same household as their mother. We further require that children in the analytical sample be of age 13 to 20. The upper age limit helps minimize the possible bias from requiring that all children live in the same household as their mother, since older children were more likely to live away from their parents. The other restrictions are similar to those applied to the CRELES data. Our final analytical sample from the Census contains 9,120 children of 4,623 mothers[1].

The restriction that all children live in the same household as their mother could yield biased results if the effect of sibling size operates differently on children who stay in the

[1] The mean number of children per mother in the Census sample is lower than in the CRELES sample because of the restriction that all children live in the same household as the mother.

household before age 20 versus those who leave home before reaching 20. In the results section, we demonstrate that this restriction does not yield significant bias by comparing the coefficient estimates from the Census data to those from the CRELES data.

Measures

In this section we focus on the measures of key variables using the CRELES data. Measures using the Census data are similar. Our dependent variable is whether the child has attended at least one year of secondary school. The CRELES survey asks the elderly respondent about the highest level of schooling and the number of years at last level of schooling of each of her children. We code the dependent variable as 1 if the child has attended at least one year of secondary school, and 0 otherwise. We use secondary school attendance as the dependent variable instead of other possible measures of education because continuing education after completing elementary school represents a key decision made by Costa Rican families when children are entering teen ages. In addition, there are more than 40% of all children in both samples who did not attend any secondary school, generating sufficient variation to examine the effects of fertility on children's education.

Our key independent variable of interest is the number of siblings the child has. The CRELES survey asks all elderly respondents how many living children they have. We subtract 1 from this number to obtain the number of siblings for each child, after matching the reported number of living children with the total number of children present in the survey data for each family.

We control for several family level characteristics that potentially mediate the effect of the number of siblings on children's secondary school attendance, including mother's years of schooling, mother's age at first birth and whether the child has any sisters in the family. Butcher and Case (1994) find that women's education choices are systematically affected by whether they were raised with any sisters. Since we use the sex composition of the first two born children as instruments for mothers' fertility, controlling for whether the child has any sisters allows us to account for one potential mechanism by which the exclusion restriction could be violated. We illustrate this point in more detail in the next section. We also include dummies of child age and the canton where the parent has lived the longest[2] which absorb any year-specific and region-invariant as well as any canton-specific and time-invariant confounders. Therefore we use only within-canton and within-cohort variation to identify the effect of fertility on education.

In assessing the meaning of our results at the macro level, i.e. on the national trends of cohort's educational attainment, we estimated the time trend in family size from the perspective of children; i.e. the sibship size of children aged 14 years (Figure 1), which is the central age when families make the decision of dropping secondary school. We

---

[2] Ideally we would like to control for dummies of the child's birth canton, but it is not directly available in the survey. For robustness check using the Census data we control for dummies of the child's canton of residence five years ago, i.e. 1979.

estimated this indicator with census data on the mean and variance of surviving children of women aged 41 years for children born in 1950 to 1972 and aged 40 years for the cohorts born in 1973 to 1997. We used the following approximate relationship adapted from Lam and Marteleto (2013), which, in turn, is based on an identity proposed by Preston (1976). Our approximate formula is:

$$Sc(a) = Sw(a+m)\ [Vw(a+m) / Sw(a+m)]$$

Where: Sc is family size from child's perspective,
Sw is family size from woman's perspective,
Vw is the variance in Sw,
a is the age of children, and
m is the mean fertility age: 27 years in 1950-72 and 26 years in 1973-97.

The data sources of Sw and Vw were: 1973 Census for cohorts 1950-58, 1984 census for 1959-70 cohorts, 2000 census for 1971-85 cohorts, and 2011 census for 1986-97 cohorts

Analytic strategy

Our main analytical strategy is a two-stage least squares (2SLS) regression model. We use the sex composition of the mother's first two births as instruments for estimating the effect of having an additional sibling on the probability of the child completing at least one year of secondary school. We operationalize sex composition as a vector of two binary variables with the first one indicating whether the first two births in the family are both boys and the second one indicating whether the first two births are both girls. Sex composition has been used as instruments in previous studies (Jensen 2005; Lee 2008; Qian 2004) to examine the effect of sibling size on children's education since it is hypothesized to be unrelated to mother's fertility preferences or of unobserved variables affecting children's education (Schultz 2007). The explanatory power of these instruments depends on the extent to which sex composition of the elder children alters fertility decisions because of parental preference between having boys and girls. As shown by several surveys, starting with the 1976 World Fertility Survey, Costa Ricans have no sex preferences for their children with the exception of a preference for having balanced families with children of both sexes (DGEC & WFS 1978)

The exclusion restriction required of the instruments could be violated if sex composition of the first two births affects children's education via mechanisms other than its impact on mother's fertility decision. As mentioned before, one such mechanism may be the reference group effect proposed by Butcher and Case (1994). They find that women raised with only brothers have received on average significantly more education than women raised with any sisters, controlling for household size. On the other hand, being raised with any sisters does not have any significant impact on men's education. Their findings are consistent with the reference group model which suggests that the presence of a second daughter in the household changes the reference group for the first, as parents with only one daughter may measure her achievement on the same scale as their sons' and may provide her with an equal share of the household's educational resources (Butcher and Case 1994). Thus sex composition could affect girls' education through parental expectation for daughters and the allocation of family resources. We hence control for whether the child has any sisters in the 2SLS model.

As discussed previously, we restrict the final analytical sample to children with at least one sibling in order to apply the instruments described above. We estimate the models separately for boys and girls to capture any systematic differences in the estimated relationship by gender.

3. Results

In this section, we first present graphical evidence and descriptive statistics using our primary analytical sample from the CRELES survey data. We then show the regression results using both CRELES and Census data.

Figure 2 depicts the relationship between the number of siblings and the probability of attending secondary school by gender, with a superimposed histogram of frequency of families by sibship size. There is a negative and linear relationship between the number of siblings and the mean probability of attending secondary school, which is similar for both boys and girls. At each sibship size, girls appear to be slightly more likely than boys to have attended secondary school. The relationship becomes nosier as sibling size grows over 10 as a result of sparseness of observations.

[Insert Figure 2]

Table 1 shows descriptive statistics for children from both the CRELES and Census sample. Children in the two samples appear similar along most characteristics. A little over half of all children in both samples have attended at least one year secondary school, while girls have a slight advantage than boys in secondary school attendance. The majority of children have at least one sister, although the proportion is somewhat higher in the CRELES sample than the Census sample. On average, mothers have a little over five years of schooling which is below completion of elementary school. Mean mothers' age at first birth is about 23 in the CRELES sample and 21 in the Census sample. On the other hand, children in the census sample have on average about 1.5 fewer siblings than the Census sample, which could be explained by the additional restriction that all children have to be found in the household in the Census data. They also appear to be older with a mean birth year around 1968 compared to a mean birth year of 1974 in the Census sample. This is primarily due to the difference in time of survey between these two data sources: although we require children in both samples to be older than 13, the age of children in the Census was measured in the year 1984 while that in the CRELES sample was measured in either 2007 or 2010.

[Insert Table 1]

Table 2 presents the OLS as well as the first stage of the 2SLS regression results from both the CRELES and the Census data. Column 1 and 2 are OLS estimates for girls and boys respectively. Having an additional sibling is associated with a decrease in 4.18 percentage points in the probability of attending secondary school for girls. The effect for boys is of slightly smaller magnitude at 3.63. Both estimates are highly significant at 0.001 level. Having any sisters appears to have little effect on girls whereas it increases the probability of boys attending secondary school by 5 percentage points, with only marginal significance. Mother's years of schooling is a strong predictor of secondary school attendance and has similar effect for both boys and girls. Mother's age at first birth is significantly positively associated with secondary school attendance for boys but not girls. These results suggest that sibling size has a strong negative impact on children's education especially for girls.

Columns 3 and 4 contain OLS estimates by gender using the Census data. The coefficients on number of sibling, which is the key independent variable, appear to be somewhat larger than but qualitatively similar to the estimates using CRELES data. An additional sibling is associated with a 5.59 percentage points decrease in the probability of attending secondary school for girls and 4.48 percentage points in boys. Mother's year of schooling, another strong predictor of children's schooling also has similar coefficients between the two data sources. These estimates suggest that any sampling bias resulting from restricting the Census sample to children living in the same household as their parents are likely to be moderate.

Nevertheless, OLS estimates in general are subject to omitted variable bias because the fertility decision of mothers is likely to be endogenous to unobserved confounders that also affect children's education. We now turn to the 2SLS models with sex composition of the first two births as instruments. We focus first on the first stage models that use the covariates as well as the instruments to predict the number of siblings the child has. Columns 5 and 6 report the coefficient estimates using CRELES data. Among girls, having two eldest brothers is associated with 0.86 additional sibling on average, whereas the first two births in the family being female is associated with 0.56 fewer sibling. For boys, the first two births in the family being the same sex is associated with 0.29 to 0.72 more siblings on average. All coefficients are statistically significant at conventional level. The partial F statistic on the instruments is 34.65 in the model for female, which passed the weak instrument test (F-stats>10). On the other hand, the partial F-stat is only 6.51 in the first stage model for male, which raises the concern for weak instrument problem. This could be due to the lack of power in estimating the relationship between sibling size and education. Therefore we estimate the same models on the Census data with over 4,000 families, more than double the sample size of the CRELES data. We report the estimates from the Census data in Columns 7 and 8.

The coefficients on the instruments in the model for female are very similar to those from CRELES data, and more significant perhaps because of the increased sample size. In addition, the coefficients on the instruments for male increased somewhat from those using CRELES data especially for first two born boys. The first two births being male is now associated with 0.68 additional sibling for boys, whereas the first two births being female is associated with 1.12 more sibling. Both estimates are highly significant. Moreover, most coefficient estimates on other explanatory variables in the first stage models are also comparable across the two data sources. The fact that the two samples yield first stage coefficients with the same sign and similar magnitude (especially for female) indicate again that the Census sample is representative of the general population in the same cohorts. Since the first stage F-stats are over 50 for both male and female using the Census data which indicates strong explanatory power of the instruments, we focus on the Census sample in deriving our 2SLS estimates for the effect of sibling size on secondary school attendance.

[Insert Table 2]

Table 3 reports coefficient estimates from the second stage of the 2SLS models using the Census data. The coefficients on sibling size are highly significant for both girls and boys. Having an additional sibling decreases the probability of having attended at least one year of secondary school by almost 10 percentage points for girls and 5.94 percentage points for boys. All the other covariates appear to have highly significant effect on secondary school attendance and are of the expected sign. However, we do not claim those effects to be entirely causal or accurately estimated as our methodology focuses only on addressing the bias from estimating the effect of sibling size on education.

[Insert Table 3]

4. Discussion and conclusion

Large fertility declines such as that experienced by Costa Rica have the potential for resulting in substantial demographic dividends—but these are not automatic. We do not explore in this paper the educational supply policies that could help realize educational dividends, nor the underlying drivers of fertility decline, but we do explore the extent to which households are likely to have chosen higher educational attainment for their children as they trade-off quantity for quality. In this Costa Rican case, the quantity-quality tradeoff appears quite robust, suggesting strong demographic dividends. Sibling size has a large negative effect on the probability of attending secondary school for both boys and girls which is confirmed by the graphical evidence as well as the results from the 2SLS models estimated on both data sources. The effect is comparable in magnitude to and in some specifications larger than an additional year of mother's schooling, and is stronger among girls than boys. OLS tends to underestimate the effect of fertility on education due to confounders. The fact that the OLS and first stage estimates are qualitatively similar between two different data sources is reassuring and strengthens the validity of the 2SLS estimates using the Census data.

It would be of interest to assess the implication of the coefficient estimates on the instruments for parental preferences on children's sex in Costa Rica. Of the four coefficients on the instruments reported in the first stage models by gender from either data source, three of them are positive and statistically significant which is consistent with the hypothesis that parents prefer at least one boy and one girl, which is why they might continue to have a third child or more after having children of the same sex in the first two births. However, the coefficient on first two girls is negative in the female model, suggesting a countervailing tendency of having fewer children conditioning on the first two births being female. We further conducted a simple OLS regression analysis at the family level which predicts the probability of having a third birth by the sex composition of the first two births. We omit reporting the regression table here. It appears that having the first two births as the same sex significantly increases the probability of continue on to a third birth compared to when the first two births are one boy and one girl. The coefficient estimate on first two boys is 0. 677 with a standard error of 0.013 adjusted for clustering at the canton of residence, white the coefficient on two girls is 0.044 with a clustered standard error of 0.13. A linear F-test rejects the hypothesis that the two coefficients are equal at 0.05 level of significance. Taken together, these results suggest that Costa Rican families have a general preference of having a balanced sex among children, with some preference for having girls – a new finding for Costa Rica.

To assess the meaning of these results for the macro trends in education in Costa Rica, figure 3 shows the simulated trajectories, with a 95% confidence interval, in the proportion with secondary education of cohorts born in the 1952-1994 period. The simulation assumes that the only determinant of education that changed over time was sibship size, which followed the curve shown in Figure 1. We simulated 1,000 trajectories and the figure shows those between the 2.5 and 97.5 percentiles as a 95% confidence interval. We use in the simulation the marginal effects of family size estimated with census data with instrumental variables (IV). However, since the comparison of census-OLS estimates with CRELES estimates (Table 2) suggests that census estimates are upward biased, we introduced a downward correction of the IV-census estimates in the proportion suggested by the OLS comparison. The corrected-effects of one-less sibling on the proportion with some secondary education were 0.0731 for girls and 0.0446 for boys. We also used the standard errors from the IV-census estimate.

[Insert Figure 3]

The simulation shows that our regression estimates imply huge effects of family size on education, which might explain almost all of the 40-year improvement that occurred among women and most of the improvement in men. The predicted mean value for women born in 1994 is 75% (72% to 79% confidence interval) compared to the observed values of 41% for those born in 1951 and 79% for the 1994 cohort. The predicted mean value for men born in 1994 is 63% (59% to 66% confidence interval) compared to the observed values of 42% for those born in 1951 and 63% for the 1994 cohort. It is also worth noting that the simulation predicts as well the crossover of the curves by sex in the cohorts of the late 1950s and the faster growth of education of girls compared to boys that has taken place in Costa Rica.

These results indicate potentially large demographic dividends from Costa Rica's dramatic fertility decline, and call for further analysis of the longer-term effects on other measures of social outcomes and well-being.

References

Angrist, J. D., V. Lavy, and A. Schlosser. 2005. “New evidence on the causal link between the quantity and quality of children.” NBER working paper.

Becker, G. S. and H. G. Lewis. 1973. “On the Interaction between quantity and quality of children.” Journal of Political Economy 81(2): S279-S88.

Berkeley Population Center. 2012. "CRELES: Costa Rican Longevity and Health Aging Study" [accessed on April 7, 2012].
Available at: http://www.creles.berkeley.edu/index.html.

Black, S. E., P. J. Devereux, and K. G. Salvanes. 2005. “The more the merrier? The effect of family size and birth order on children's education.” The Quarterly Journal of Economics 120(2): 669-700.

Blake, J. 1981. “Family size and the quality of children.” Demography 18(4): 421-22.

Booth, A. L. and H. J. Kee. 2009. “Birth order matters: the effect of family size and birth order on educational attainment.” Journal of Population Economics 22(2): 367-97.

Butcher, K. F. and A. Case. 1994. “The effect of sibling sex composition on women's education and earnings.” The Quarterly Journal of Economics 109(3): 531-63.

Caceres, J. 2004. “Impact of family size on investment in child quality: Multiple births as natural experiment.” University of Maryland.

CCP & INEC (2013). Costa Rica: Estimaciones y Proyecciones de Población por sexo y edad 1950-2050. San José, Costa Rica: publicaciones del Instituto Nacional de Estadística y Censos (INEC), marzo 2013.

Coale, A. J. (1983). Recent trends in fertility in less developed countries. *Science, 221*, 828-832.

DGEC, & WFS. (1978). Encuesta Nacional de Fecundidad 1976. San José, Costa Rica: Dirección General de Estadística y Censos (DGEC) & World fertility Survey (WFS).

Hanushek, E. A. 1992. “The trade-off between child quantity and quality.” Journal of Political Economy: 84-117.

Jensen, R. 2005. “Equal treatment, unequal outcomes? Generating sex inequality through fertility behavior.” Harvard University.

Knodel, J. and M. Wongsith. 1991. “Family size and children's education in Thailand: Evidence from a national sample.” Demography 28(1): 119-31.

Lam, D. and L. Marteleto. 2008. “Stages of the demographic transition from a child's perspective: Family size, cohort size, and children's resources.” Population and Development Review 34(2): 225-52.

Lam, D., & Marteleto, L. (2013). Family Size of Children and Women during the Demographic Transition. Paper presented at the XXVII IUSSP International Population Conference, Busan, Korea.

Lee, J. 2008. “Sibling size and investment in children's education: an Asian instrument.” Journal of Population Economics 21(4): 855-875.

Lee, R., & Mason, A. (2010). Fertility, Human Capital, and Economic Growth over the Demographic Transition. European Journal on Population, 26(2) 159-182.

Li, H., J. Zhang, and Y. Zhu. 2008. “The quantity-quality trade-off of children in a developing country: Identification using Chinese twins.” Demography 45(1): 223-43.

Marteleto, L. and M. Dondero. 2013. “Maternal age at first birth and adolescent education in Brazil.” Demographic Research 28: 793-820.

Qian, N. 2004. “Quantity-quality and the one child policy: The positive effect of family size on school enrollment in China.” Department of Economics. Massachusetts Institute of Technology.

Preston, S. H. (1976). Family sizes of children and family sizes of women. Demography, 13(1), 105-114.

PRB. (2012). World Population Data Sheet. Population Reference Bureau (PRB) publications, Washington DC.

Programa Estado de la Nación. (2005). Primer Informe Estado de la Educación. San Jose, Costa Rica: Programa Estado de la Nación.

Resenzweig, M. R. and K. I. Wolpin. 1980. “Testing the quantity-quality fertility model: the use of twins as a natural experiment.” Econometrica: journal of the Econometric Society 48(1): 227-40.

Rosero-Bixby, L (1999). Interaction Diffusion and Fertility Transition in Costa Rica: Quantitative and Qualitative Evidence. In Leete, R (Editor), *Dynamics of Values in Fertility Change*. Oxford, UK: Oxford University Press (210-237).

Salazar, J. M. (2003). Historia de la Educación Costarricense. San Jose, Costa Rica: EUNED.

Shultz, P. T. 2007. “Population policies, fertility, women's human capital, and child quality.” In Handbook of development economics 4, pp. 3249-303.

United Nations. (1985). Socio-economic Development and Fertility Decline in Costa Rica. New York: United Nations (UN) Department of Economic and Social Affairs.

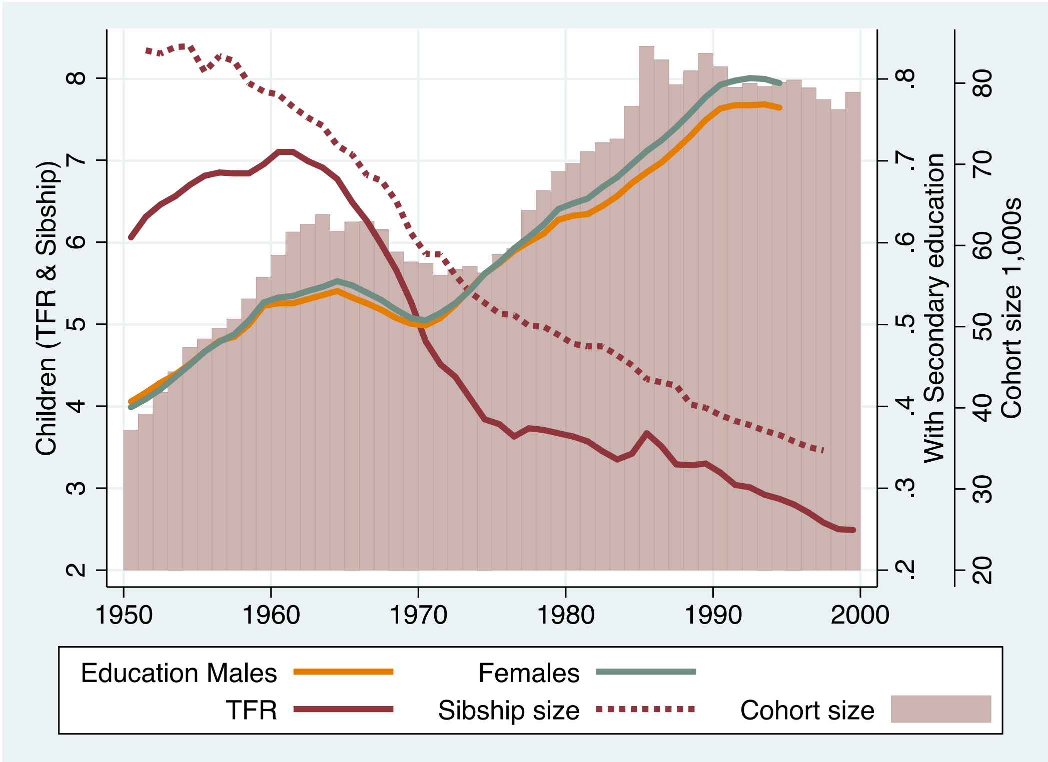

Figure 1: Total fertility rate, cohort and sibship size, and proportion with some secondary education, Costa Rican cohorts born 1950-1996

[1]

[1]Cohort size (births) and TFR source: Web page of the Central American Population Center, http://ccp.ucr.ac.cr/observa (accessed on September 3, 2013).

Proportion with secondary education source: 2011 census, population aged 14 and older by birth year, micro-data available at http://censos.ccp.ucr.ac.cr (accessed on September 3, 2013).

Sibship size: estimated with census data as explained in text.

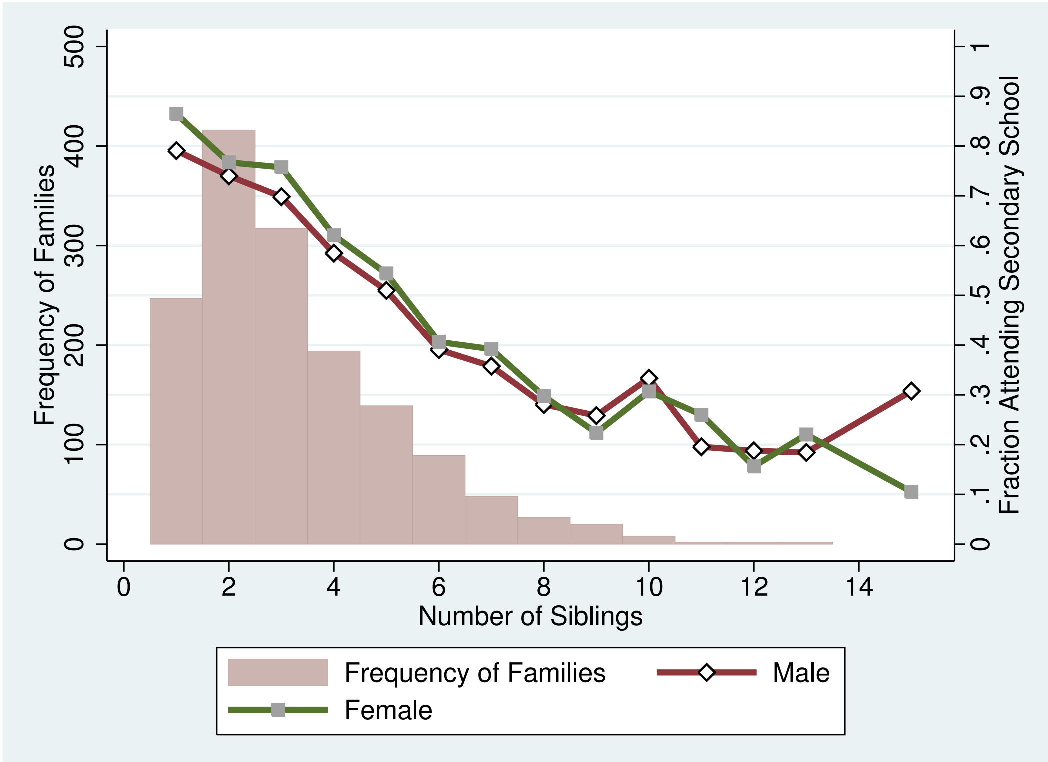

Figure 2: Mean secondary school attendance and frequency of families by number of siblings, CRELES sample

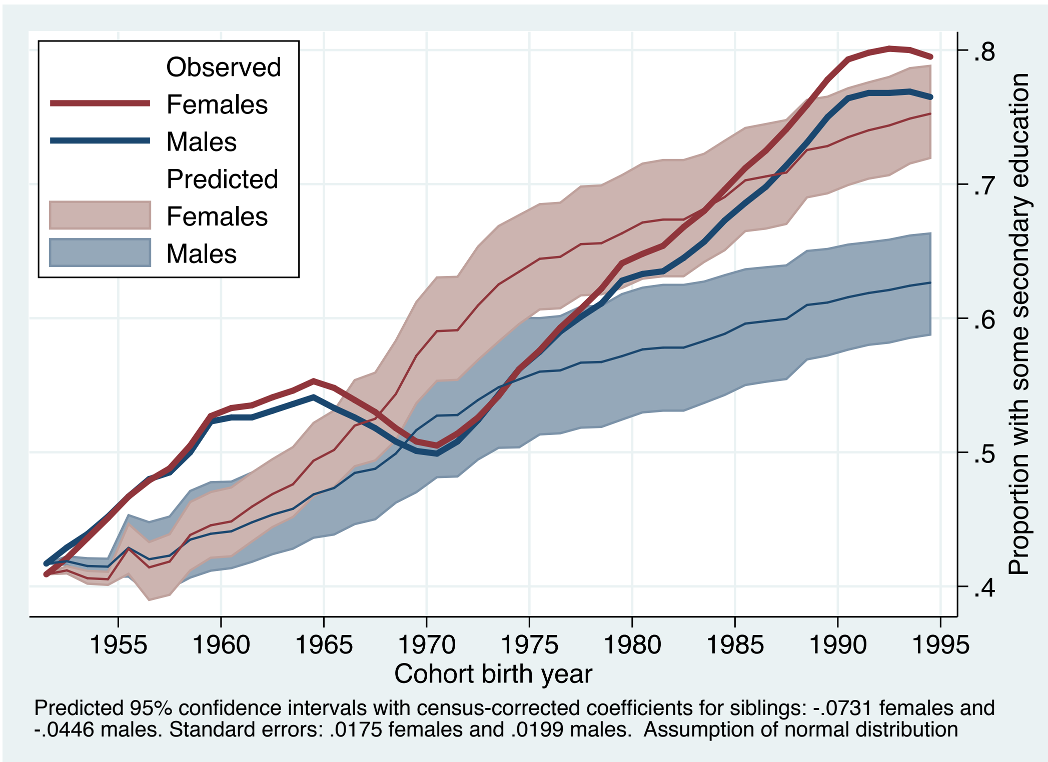

Figure 3: Proportion with some secondary education, observed and predicted from changes in sibship size, Costa Rican birth cohorts born 1951-1994

Table 1: Descriptive Statistics by Gender for Children from Different Samples

	CRELES			Census		
	Male	Female	Difference	Male	Female	Difference
Any secondary school	0.547	0.577	-0.030**	0.524	0.578	-0.054***
Number of siblings	4.690	4.739	-0.049	3.254	3.202	0.052
(standard deviation)	(2.629)	(2.714)		(1.786)	(1.738)	
Any sisters	0.906	0.888	0.018**	0.809	0.807	0.002
Child birth year	1974.0	1974.1	-0.1	1968.4	1968.6	-0.2***
(standard deviation)	(9.485)	(9.349)		(2.043)	(2.038)	
Mother's years of education	5.264	5.306	-0.042	5.808	5.884	-0.076
(standard deviation)	(3.930)	(3.918)		(3.689)	(3.639)	
Mother's age at first birth	23.192	23.116	0.076	21.283	21.365	-0.082
(standard deviation)	(5.971)	(5.895)		(4.068)	(4.096)	
# Observations of Children	4,647	4,675		4,786	4,334	
# Observations of Mothers	1,888	1,887		3,248	3,082	

Note: Mean or proportion shown. Standard deviations in parentheses.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 2: OLS and First Stage Estimates of The Effect of Sibling Number on Prob. of Attending Secondary School: CRELES and Census Data

	OLS				First Stage			
	CRELES		Census		CRELES		Census	
	Female	Male	Female	Male	Female	Male	Female	Male
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of siblings	-0.0418****	-0.0363****	-0.0559****	-0.0483****				
	(0.0041)	(0.0050)	(0.0050)	(0.0051)				
Any sisters	-0.0055	0.0553**	0.0503***	-0.0174	2.5411****	2.2066****	1.6677****	1.5736****
	(0.0194)	(0.0270)	(0.0184)	(0.0183)	(0.1415)	(0.2124)	(0.0784)	(0.0862)
Mother's year of schooling	0.0357****	0.0391****	0.0422****	0.0401****	-0.1818****	-0.1748****	-0.0849****	-0.0921****
	(0.0031)	(0.0031)	(0.0025)	(0.0036)	(0.0185)	(0.0174)	(0.0143)	(0.0151)
Mother's age at first birth	0.0023	0.0047***	0.0070****	0.0039	0.0185	0.0170	-0.0584****	-0.0649****
	(0.0015)	(0.0016)	(0.0018)	(0.0024)	(0.0203)	(0.0155)	(0.0086)	(0.0091)
First two born boys					0.8580****	0.2915**	0.9922****	0.6822****
					(0.1373)	(0.1187)	(0.1063)	(0.0771)
First two born girls					-0.5601**	0.7190***	-0.6068****	1.1180****
					(0.2196)	(0.2555)	(0.0694)	(0.1425)
Canton dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Child birth year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First stage partial F-stats					34.65	6.51	88.93	57.48
Mean of Dep Var	0.58	0.55	0.58	0.52	4.74	4.69	3.20	3.25
Number of Observations	4,675	4,647	4,334	4,786	4,675	4,647	0.4,334	4,786
R squared	0.30	0.29	0.35	0.32	0.39	0.37	0.33	0.29

Note: Standard errors adjusted for clustering at canton level. Each observation is at the child level.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 3: Second Stage Estimates of The Effect of Sibling Number on Prob. of Attending Secondary School: Census Data

	(1) Female	(2) Male
Number of siblings	-0.0977****	-0.0594***
	(0.0175)	(0.0199)
Any sisters	0.1023****	-0.0041
	(0.0271)	(0.0287)
Mother's year of schooling	0.0382****	0.0391****
	(0.0030)	(0.0044)
Mother's age at first birth	0.0043*	0.0031
	(0.0022)	(0.0028)
Canton fixed effects	Yes	Yes
Child birth year fixed effects	Yes	Yes
Mean of Dep Var	0.58	0.52
Number of Observations	4,334	4,786
R squared	0.33	0.32

Note: Standard errors adjusted for clustering at canton level. Each observation is at the child level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$