The relationship between fertility and children's education in Brazil: a natural experiment using twin data

Ana Paula de Andrade Verona

Abstract

The purpose of this work is to assess the role of fertility or number of children in determining the investments on the children's human capital in Brazil. The decline in fertility and the increase in children and adolescents' education observed during the last decades in Brazil suggest a potential trade-off between these two variables The main contribution of this paper is to employ a natural experiment using twins' data as a proxy for exogenous fertility. From it, it is possible to notice estimates of the exogenous fertility impact on the education of children. These estimative are uncontaminated by the simultaneity bias. The database used is a sub-sample of the Brazilian National Household Sample Survey (Pesquisa Nacional por Amostra de Domicílios - PNAD), conducted by the Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatística - IBGE). Data from the 1984, 1985, and 1996 to 2001 PNAD were used, with the exception of 2000. The results indicate primarily the existence of a trade-off between (exogenous) fertility and the education of children in Brazil. Secondly, it was found that this negative relation has lost strength or magnitude during the last fifteen years. Moreover, it is also seen that such trade-off is greater in situations deemed to be "unfavorable".

Key words: natural experiment, twins, exogenous fertility and children education.

1- Introduction

During the last decades Brazil has undergone rapid and profound demographic changes. In this process, a rapid decline in fertility has been observed, which consequently changed the age composition of the population. The significant decline in school-age children and youths stands out among these changes.

A reduction in the proportion of children and youths can contribute to better access to education. In the decade of the 1990s, Brazil provided almost universal access to education for children between 8 and 14 years of age, as measured by school enrollment. Furthermore, secondary school education expanded in an accelerated rhythm, one of the most positive results presented by the Brazilian education system at the end of the decade of the 1990s[1]. The fertility decline and the increase in education for children and youths suggest a potential trade-off between these two variables.

The motivation for the current study is to evaluate the role of fertility within the family in the determination of investments in children's human capital. The principal contribution of this article is to employ a natural experiment using twin data as a proxy for exogenous fertility. From these data it was possible to perceive the estimated impact of exogenous fertility on children's educational achievement, without contamination of simultaneity bias.

2 – Theoretical model of the relationship between fertility and education.

2.1- The demand for children

The demand for children is influenced by an ample set of variables that makes its analysis considerably complex. The traditional view argues that children compete for family resources, but, however, do not contribute to the acquisition of these resources [2]. Many of these resources are transferred in order to cover expenditures for food, health, hygiene, education, shelter, and leisure, among others. This article recognizes the importance of the range of possible investments that parents make in raising their children; however, it is limited to analyzing parents' investments in formal education.

The economic perspective, on which this article draws, considers the maximization of the utility function of the family. Based on this consideration, one can affirm that parents seek to invest in their children in an efficient manner, investing equally or not among them. Parents make their decisions about the distribution of education among their children based on expectations of the return of human capital invested in each child [2]. This perspective uses an

interaction between quantity and quality of children. The first is the number of children that a couple decides to have and the second is the formation of human capital for each child, that the parents expect to provide, including principally, investments in schooling (formal and informal), in health, and in dedication of time and attention.

3 – A review of causality and the simultaneity bias

Casual analysis is inserted in a larger problem based on the fundamental counterfactual notion of a unit of analysis, such as individual, state or country [3]. This notion becomes clear when considering the following situation: an individual "i" has two possible, or potential outcomes, Y_{1i} e Y_{0i}, where Y_{1i} is the result if the individual experiences a particular event or makes a determined decision, and Y_{0i} is the result if he does not experience the event, maintaining all else constant. The difference between the two (Y_{1i} - Y_{0i}) is the casual effect of the event or action. Only one of the outcomes can be observed, since the individual cannot do both things at the same time. The unobserved result is called counterfactual [3].

In practice, a researcher can learn more about counterfactual results if she considers the experiment to be random. In much research, principally in the medical field, it is believed that the best evidence about counterfactual results can be observed through random experimentation. The lack of random experimentation is one of the principal reasons that econometric studies are always less convincing than research in other fields based more in experiments [4].

3.1 – The problem of endogeneity and the identification strategy.

The response variable, Y_i, is affected by observed and unobserved characteristics (immeasurable). The endogeneity problem arises because the event (or action) is also affected by immeasurable characteristics. This occurs because it is not possible to totally explain the determinants or motives that lead an individual to make a decision or to have some experience and not others (for example, some people smoke, study, get married, while other individuals do not). If these unobservable characteristics are correlated with the non-observable characteristics that affect the response variable, a spurious correlation will be established between the event (or action) and the response variables Y_i [3].

This being so, a basic point in the literature about causal effect is that it cannot be estimated without some type of hypothesis or restriction, given the inherent non-observation of the counterfactual result. The majority of the important causal problems in economic and population research arise when a variable that affects Y_i is endogenous, that is, when the values

of the variable can vary in unobservable ways. This problem produces an incorrect estimation of the causal effect in Y_i, corresponding thus to a non-adjusted regression coefficient. This means that hypotheses or identification strategies must be made [3].

One of the identification strategies considered is the natural experiment. This consists of a heterogeneous set of means that employ "natural" occurrences considered random, such as those observed in laboratory experiments[4]. In other words, natural experiments "are cases where different sets of individuals are treated in such a manner that, if not random by design, they will be in practice" (p. 112) [5]. Basically, the natural experiment is an identification strategy that uses random variables related to climate or biology, for example, as the explanatory variable.

One objection of natural experimentation is in relation to its validity. The type of exogenous variable employed in this experiment can illustrate a trade-off between internal and external validity of random tests. The internal validity is assured when the experiment is truly random and when there is no contamination between the experimental and control groups. External validity is perceived when the estimate can be generalized for a larger population than that under the effect of the experiment. The natural experiment (at least that which uses twin data) represents an extreme attempt to maximize the internal validity (of exogeneity), which many times results in a significant loss of external validity (possibility of generalization).

The problem arising from the lack of the possibility of generalization, in experiments that use twin data, occurs because it is necessary to acquire sub-samples of families with multiple births (a rare event). This type of sub-sample can have (very) different characteristics than those observed in the total sample. Therefore in a natural experiment even if the internal validity is assured (exogenous impact) the results cannot be generalized to the larger population [3].

3.2 -Causality and the simultaneity bias in the case of fertility and education

When one pretends to do causal analysis between two variables that are endogenous, a simultaneity bias can be observed. Simultaneity is a classic argument for correlation between the error term and the explanatory variables [6].

A specialized literature that analyzes the impacts of family size on educational production of children has found deficiencies that can compromise the interpretation of estimates [7]. Within these, there exists one that is observed when it is assumed that the decision of family size is not related to parents' preferences regarding amount of resources (time and money, for example) that they expect to invest in each child's education. This extreme hypothesis is not accepted by many

theories of family behavior, and results in an estimate that exaggerates a possible negative impact of the number of siblings on children's education. In fact, using an economic model, one finds that education (quality) as well as number of children (quantity) are generally related to the same determining variables (income, household goods, mother's education, etc.). Thus, it is extremely difficult to separate the impacts of the determinants of fertility on education, and, by the same token, separate the determinants of education on demand for children.

When considering that the choices of quality and quantity of children, in a household, are thus determined simultaneously by the parents, that is, controlled concurrently within the model, it is assumed that the total investment in human capital for each child as well as the level of fertility of the parents, are endogenous variables. In this case, one can observe the presence of a bias due to the occurrence of endogenous decisions, that are made conjointly and that are influenced by common determinants. This statistical bias is called simultaneity bias. The estimates of the impact of family size on human capital investment in children can be compromised given the existence of this bias.

In models where fertility and education are determined in the same time, the direction of the simultaneity bias depends on the how size of family in relates to the costs of education, and of schooling in relation to family size. In general, the bias will be less in low-income countries, in rural areas, where levels of education are low[7]. Probably, in these situations, the decision about family size is not substantially affected by decisions regarding educational investment in children, which makes the simultaneity bias weak or even non-existent. In other words, in this case, the fertility can be considered a "less" endogenous variable. On the other hand, the simultaneity bias should be stronger among more educated women, in urban regions and, of course, in times of low fertility.

Due to the possibility of simultaneity bias in the estimate of the relationship between fertility and children's education, an attempt was made to eliminate or minimize this bias by using twin birth data. In the twin birth at least one of the children was not planned or expected. A family that deals with an unplanned birth must feel compelled to readjust their spending, having to include among them resources for the health of the child, as well as other costs arising from the growth and raising of the child. Using fertility as an exogenous variable, through a proxy such as twin births, the estimate between the quality and quantity of children can be seen more clearly, or at least, present less of a simultaneity bias[8]. Thus, the objective or motivation of

this study is to measure the exogenous impact of multiple births (twins) on the educational achievement of the children, who at the time of the study were between 7 and 14 years of age.

In Brazil the use of natural experiments is still not widely used. Thus, the application of twin data, by itself, is a relevant contribution of this study. There exists but one study that evaluates the Brazilian case using twin birth data in first pregnancy, to measure the impact of exogenous fertility in the insertion of women in the labor market [9].

4 – Data

The database used is a sub-sample of the Brazilian National Household Sample Survey (Pesquisa Nacional por Amostra de Domicílios - PNAD), conducted by the Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatística - IBGE). Data from the 1984, 1985, and 1996 to 2001 PNAD were used, with the exception of 2000. These data were grouped into two distinct databases. The first contains only the 1984 and 1985 PNAD and the second 1996 to 2001. The main reason for separating the data and constructing these two datasets was to make comparisons between levels of fertility, in two different contexts. Households were maintained in the sample that had at least one non-twin child, between the ages of 7 and 14. The sample was filtered according to the age of the non-twin child living in the household. A total of 1,191 and 1,859 twin births were found (in the first and second database, respectively), that corresponds to 0.55% and 0.42% of the total of live births, surviving or not.

Households with triplets were also identified in all of the surveys. In total, 19 triplet births were found in households that had at least one non-twin child between the ages of 7 and 14. However, these cases were excluded in the attempt to minimize the case of fertility treatments. When a couple decides to have this treatment, the exogenous variation, observed with the birth of twins, could no longer exist. This is because the incident of twins and triplets becomes expected or controlled within the household, which would make fertility an endogenous variable. The objective to not consider households with triplets was to minimize the chance of finding births as a result of these treatments. Regardless of this care, it is supposed that in developing countries, such as Brazil, fertility treatment is still not popular, not being accessible to all social classes. In other words, fertility treatments are not considered so representative to the point of invalidating the natural experiment using twin data as an exogenous variation.

Finally, the resulting 1984, 1985, and 1996 to 2001 PNAD samples had a total of 57,054 and 149,177 households with at least on non-twin child between 7 and 14 years of age, respectively.

4.1- Descriptive analysis

This section provides the descriptive analysis of the analytical samples used. Table I shows some of the socio-economic characteristics of the households that have, at least, one non-twin child between the ages of 7 and 14 and child of head of household, in households with and without twin births (total sample). Table 1 shows the samples of the 1984, 1985 and 1996 to 2001 PNADs. It is important to point out that in the descriptive analysis and the multivariate results sample weights adequate for this type of sample were used.

Table I – Socio-economic characteristics of households with at least one non-twin child between 7 and 14 years of age, in households with and without twin births (total sample). Samples for 1984-1985 and 1996-2001.

Household Characteristics	1984-1985	1996-2001
Sex of head of household (%)		
Man	88.88	83.60
Woman	11.12	16.40
Condition (%)		
Urban	82.26	81.60
Rural	17.74	18.40
Have Spouse (%)		
Yes	89.04	86.35
No	10.96	13.65
Regions (%)		
Southeast	35.24	32.41
South	18.10	17.71
Center	9.93	10.96
Northeast	29.15	31.10
North	7.59	7.83
Groups of mother's education		
Illiterate	22.01	14.06
1 to 3 years schooling	24.61	17.01
4 years schooling	26.22	16.14
5 to 8 years schooling	13.43	27.23
9 or more years schooling	13.73	25.56
Mean mother's years schooling	4.18	5.84
	(3.87)	(4.25)
Mean Births	4.73	3.57
	(3.01)	(2.51)
Total of observations	57,054	149,179

Source: 1984, 1985 and 1996 to 2001 PNAD's. Obs: Standard error in parentheses.

Table 1 shows that in the two distinct periods (1984 and 1985, and from 1996 to 2001), the man is the head of the household in the large majority of the cases. Next, one can perceive that the household condition, whether urban or rural, practically did not vary between the two periods analyzed. Observe that in the two samples more than 80% of the households are located in urban areas. Another important characteristic is the presence of a spouse in the household. It is possible to perceive, from Table 1, a small increase households characterized by the absence of a spouse, between the two samples. Also, considering the five major regions of Brazil, the distribution of households does not present great variation.

The three last variables in Table 1 are in regard to the mother in residence in the household. In the first sample, more than 70% of the mothers had between 0 and 4 years of schooling. In the sample of the 1996 to 2001 PNAD's, more than half of the mothers have 5 or more years of schooling. This change is, by far, the greatest between the two PNAD samples. The second measure, that also represents mother's education, is the mothers' mean years of

schooling. The last variable represented in Table 1 is the mean number of births of women living in the household. Note that there is a reduction of 1.2 children between the samples.

Table 2– Socio-economic characteristics of households with twin births and with at least one non-twin child, between 7 and 14 years of age. 1984-1985 and 1996-2001 Samples.

Household characteristics	1984-1985	1996-2001
Sex of head of household (%)		
Man	90.60	85.85
Women	9.40	14.15
Condition (%)		
Urban	76.66	75.90
Rural	23.34	24.10
Have spouse (%)		
Yes	90.68	88.27
No	9.32	11.73
Regions (%)		
Southeast	33.42	30.23
South	15.20	15.49
Center	9.74	10.49
Northeast	32.91	34.97
North	8.73	8.82
Groups de mother's education		
Illiterate	28.04	18.83
1 to 3 years schooling	27.62	20.39
4 years schooling	25.86	17.16
5 to 8 years schooling	8.98	25.82
9 or more years schooling	9.49	17.81
Mean mother's education	3.40	4.93
	(3.45)	(4.04)
Mean births	7.06	5.60
	(3.23)	(2.97)
Total of observations	1,191	1,859

Source: 1984, 1985 and 1996 to 2001 PNAD's. Obs: Standard error in parentheses.

Table 2 shows the households with twin births and with at least one non-twin child between the ages of 7 and 14. The principle socio-economic characteristics of these households are shown in the two samples considered.

According to Table 2, the distribution of the socio-economic variables in households with twin births is not very different from that observed for households in the total sample. Only the maternal characteristics show marked differences. One can observe a relative disadvantage of mothers of twins, in the 1985 and 1985 PNAD's, more than 80% of these mothers have between 0 and 4 years of schoolings. In the sample of 1996 to 2001 PNAD's, the total of mothers of twins with 5 or more years of schooling is almost 44%. These sums, if compared to those in Table 1, show that the mothers of twin are, on average, less educated. Another difference between Table 1 and Table 2 is the mean number of births. The second table shows that mothers

of twins present, on average, a greater number of births, in the two samples of only twin mothers than in the samples of total households (see Table 1).

5 – Methodology

As was previously mentioned, the objective of this article is to observe the impact of an exogenous variation in fertility on the educational attainment of children, between 7 and 14 years of age. The fertility and child's education variables used in the natural experiment are the rate of twins (TG) and the standardized index of mean educational attainment of children 7 to 14 (IDE). These variables have already been previously developed and employed[10].

The independent variable, TG (rate of twins), that is the ratio between the total twin births and the total pregnancies, both within the household, was constructed from information about fertility of the supposed mother of the household. The denominator, that is, the total number of pregnancies of the mother in each household is equal to the number of live births, surviving or not, in the case of non-multiple births, or equal to the number of children minus one, in the cases where the mother had live twin births. This subtraction is necessary because obviously with the birth of twins, in spite of the birth of two children, implies only one pregnancy.

The dependent variable is the index IDE. This is a sum of the educational attainment of the children, between 7 and 14 years of age, in each household, in relation to the mean attainment of the sample, at the same ages. The constructed index is presented below:

$$IDE = \frac{1}{N}\sum_{i=1}^{n}\frac{e_{ix}}{\overline{e_x}}$$

Where N is the total number of children at ages 7 to 14 in the household; e_{ix} is the educational attainment of child i at age x, measured by years of schooling; $\bar{e}_x$ is the average educational attainment of all the children, at age x, in the total sample.

This index is household data and uses as an education variable years of completed schooling. The standardized index is a continuous variable. It allows for the use of ages in the interval of 7 to 14 years, without worrying about the age effect or with the effect of the grade in which the child is enrolled at the time of the study.

An ordinary least squares (OLS) regression equation will be used. In considering the fertility based on twin birth data, it is believed that the variation in the number of children is exogenous. Supposing, then, that the occurrence of twins is random, and not correlated with any variable, one has the following equation:

$$X_i = \alpha + \lambda G_i + u_i \quad (1)$$

where

G corresponds to the rate of twins.

In the equation (1), the coefficient λ represents the impact of a measure of exogenous quantity of fertility (G) in school attainment of children between 7 and 14 years of age (X). This estimate should not be biased, since the hypothesis of no correlation between the regressor and the unobserved factors was assumed $E(u_i / G_i) = 0$.

In addition to the rate of twins, other control variables were also included, such as: a dummy variable of sex of head of household (woman =1 and man=0), a dummy variable for the condition of the household (rural=1 and urban=0), four dummy variables for the major regions of Brazil (southeast is the reference category), a dummy variable for mother's education (0 to 4 years of schooling =1 and 5 or more = 0). Also, as a control, dummy variables were included for each of the different periods. In the case of the 1984 and 1985 sample, only one dummy variable was used (1985 PNAD=1 and 1984 PNAD=0), and in the case of the 1996 to 2001 sample, four dummy variables were used (1996 PNAD is the reference category). The relative estimates for each control variable used as well as that for the twin rate are presented in the appendix.

6 – Results

Table 3 shows the result of the estimates of the impact of the twin rate (variation of exogenous quantity of children) on the educational attainment of youths in the household that had at least one non-twin child between 7 and 14 years of age. This table shows the coefficients of the twin rate in the two distinct samples. Their sign is negative.

Table 3 – Coefficients of TG, for both total samples.

Variable	1984 and 1985 PNAD's	1996 to 2001 PNAD's
Rate of Twins	-0.4019 **	-0.1122 *
	(0.1423)	(0.0633)
Adjusted R^2	0.0767	0.0562
Total observations	57,054	149,177

Source: 1984, 1985 and 1996 to 2001 PNAD's. Own Construction
Obs: Standard Error in parentheses.
Controls used: mother's education, condition and head of household, major regions of Brazil, and dummies for time periods.
Level of significance: *** = 1%, ** = 5%, * = 10%.

In Table 3, the twin rate presents a negative impact on the measure of childrens' educational attainment, in greater magnitude for the sample for the years 1984 and 1985. From this result, one can say that the trade-off observed between quality and quantity of children is more expressive in the "older" sample, characterized by a larger mean number of births than that observed in the 1996 to 2001 PNAD's. The coefficient of the twin rate (controlling for the other socio-economic variables) is almost four times greater in the 1984 and 1985 sample than in that of 1996 to 2001. One can note, then, the existence of a temporal effect on the trade-off between quality and quantity of children, since in the two different moments in time, the twin rate, an measure of exogenous quantity of children, controlling for the same variables shows very different coefficients. However, it is important to point out that during this time interval the access to education increased, which could also have contributed to this temporal effect on the trade-off between quality and quantity of children.

The results of the coefficients of each variable included in the regression equation can be found in the appendix.

6.1 – Interactions

An attempt was made to observe conjointly the impact of the twin rate and each independent variable. For example, the effect of the occurrence of twins could be stronger if, in addition, the mother is illiterate or the household is located in a less developed region. This behavior is known as an interaction, and is usually captured by the inclusion of products, of the variables chosen to interact, in the regression model.

Table 4 shows the results of the three interactions performed. The first interaction was between the twin rate and a dummy variable of mother's education (0 to 4 years of schooling is the reference category); the second interaction (called extreme case 1) is between the twin rate, the dummy variable of mother's education and a dummy variable for the northeast region (northeast=1 and the other regions=0); and the third interaction (called extreme case 2) is between the twin rate, the dummy variable of mother's education, the dummy variable for the northeast region and a dummy variable for rural household condition (rural=1 and urban=0). These are the three "unfavorable" socio-economic contexts considered in this article.

Interactions were also performed separately between the twin rate and other socio-economic variables. Of all of these, only the coefficients presented in Table 4 were significant,

with the exception of those from the 1984 and 1985 PNAD's sample. All of the estimated coefficients, including those of the interactions, were placed in the appendix.

Table 4 also shows the estimates of the coefficients for the pure TG variable (without the interaction of any variable), that was also included in the regression with interactions. The effect of this variable will be called the principal effect and that of the interaction the differential effect of TG. The sum of these is the total effect.

The interactions in the 1984 and 1985 PNAD's are not significant, that is, the impact of the twin rate is not different between the distinct groups or "unfavorable" socio-economic contexts considered. But, however, when examining the principal effect of the twin rate, this period presents a larger coefficient (in magnitude) and more negative (see Table 4).

The interactions in the 1996 to 2001 PNAD's sample are significant, that is, the impact of the twin rate varies between the distinct groups and socio-economic contexts considered "unfavorable". But when examining the principal effect of the twin rate, this period presents a smaller coefficient (in magnitude) and it is not significant. In this period, the TG has a weak impact on children's education, if considering the total sample and not the groups separately.

Table 4- Coefficients of TG with interactions, for both total samples.

Interactions and TG	1984 and 1985 PNAD's	1996 to 2001 PNAD's
TG*mother's education[1]	0.2622 (0.3121)	-0.3952*** (0.1163)
TG	-0.5836*** (0.2634)	0.0384 (0.0927)
Adjusted R^2	0.0767	0.0562
TG*mother's education * Northeast[2]	0.1819 (0.2920)	-0.5155*** (0.1341)
TG	-0.4252*** (0.1591)	-0.0531 (0.0701)
Adjusted R^2	0.0767	0.0562
TG*mother's education* Northeast*rural[3]	0.0660 (0.4218)	-0.5633*** (0.1676)
TG	-0.4055*** (0.1486)	-0.0820 (0.0663)
Adjusted R^2	0.0767	0.0562
Total observations	57,054	149,177

Source: 1984, 1985 and 1996 to 2001 PNAD's. Own Construction
Obs: Standard Error in parentheses. Controls used: mother's education, condition and head of household, major regions of Brazil, and dummies for time periods.
Level of significance: *** = 1%, ** = 5%, * = 10%.

1- Mother's education is a dummy variable whose reference is group of mothers with 0 to 4 years of schooling, in relation to those with 5 years or more.

2- Northeast is a dummy variable whose reference is this region.

3- Rural is a dummy variable whose reference is this condition of household, in relation to urban.

It is easy to imagine that in the period of less fertility, such as from 1996 to 2001, in relation to 1984 and 1985, the trade-off between quality and quantity of children would be observed only in some specific groups, in which the fertility is still relatively high. And still, possibly in the years of 1984 and 1985, undesired fertility (measured by TG) still was high enough not to affect in any differentiated form the different social groups considered. Finally, it is possible to suppose that, in the older period the effect of "unfavorable" socio-economic conditions, observed separately, was not important, but that they came ot be in the more recent period, between 1996 and 2001.

The interpretation of the coefficients in Table 4 is simple. For example, the additional effect of the measure of exogenous quantity of children (TG) and the dummy variable of mother's education (mothers with 0 to 4 years of schooling =1) is given by the coefficient – 0.3952. That is, an increase of one unit of the interaction between TG and the dummy variable of mother's education, results in a decrease of 0.3952 of the same unit, in the index of mean school attainment, keeping the other variables constant. When to interaction is added the effect of the household being in the Northeast, the negative effect on the index increases to 0.5155. And,

when, to the interaction is also added the effect of the household being located in a rural area, the negative effect increases to 0.5633.

In addition, the efficacy of the twin rate is greater in situations in which, or for whom, controlling fertility is expensive or more difficult. This is because, when access to contraceptive methods is very limited, the level of undesired fertility probably will be high, which will minimize or weaken the relationship between the twin rate and the level of desired fertility (N*). In the "unfavorable" socio-economic situations constructed above, such as low maternal education, in rural households and in the Brazilian Northeast, the TG coefficient appears more negative also due to a greater observed efficacy in the use of this measure of exogenous quantity of fertility.

In performing these interactions, the impact of the measure of exogenous quantity of family size in "unfavorable" socio-economic situations was sought, in which the mean number of births was high. It was found that the estimated coefficients increased in their negative value in situations or periods of higher fertility or mean number of births.

7- Conclusion

Understanding the relationship between family size and the investment in human capital of children is not an easy or unimportant task. Many areas, such as the social sciences and economics, recognize the relevance of this study. In some countries, in which this association has been researched, the results are varied, finding a positive, negative, mixed or even no relationship between the two variables cited above, without the possibility of determining a single and precise answer.

This study makes a contribution to the important discussion of the relationship between fertility (as an exogenous component) and children's education, and for the promotion of public policy. Levels of fertility are one of the principal determinants of variations in growth rates (or rates of decline) of the future work force. And human capital, invested today in our children, be it through family decisions or by way of public agencies is the principal agent in shaping the adquired and used skills by this future labor force. These two factors are of utmost importance for economic development. In addition, the levels of education that children attain (in qualitative as well as quantitative terms) in the present, particularly for girls (who are one portion of future mothers) will greatly influence the health and survival of the next generations. Furthermore, detailed studies of the relationship between fertility and education could facilitate decisions

related to the levels and quality of education, the structure of the educational system, and even to the division of resources between education and other areas, that is indispensable when looking to the well-being of the population and economic growth.

Still, through information on the relationship between education and fertility (exogenous), decisions about public policy can become more rational. Some of these policies could be in the areas of education, worker's employment and also population health.

Finally, knowing whether, within the household, the number of siblings has an impact on the children's education can be an important tool in the making of public policy that aids in the promotion of the universalization of school enrollment and improvements in the rates of educational attainment. This study also can aid in the promotion of specific programs, such as, for example, "bolsa-escola" (school-stipends).

8 –Bibliográphic References

1 - Castro MHG. As desigualdades regionais do sistema educacional brasileiro. In: Henriques, R. organizador. Desigualdade e pobreza no Brasil. Rio de Janeiro: Editora IPEA, 2000; p. 425-458.

2 – Becker GS. The demand for children. In: Becker GS. The treatise on the family. Cambridge; London: Editora Harvard University; 1981. p. 93-112.

3 - Moffit R. Remarks on the analysis of casual relationships in population research**.** Johns Hopkins University. http://www.econ.jhu.edu/People/Moffitt/causal.pdf

4 - Angrist J, Krueger A. Empirical strategies in labor economics. In.: Ashenfelter O, Card D. (Eds.) The handbook of labor economics. Amsterdam: Editora North-Holland, 1998. v.3A, Chapter 23.

5 - Deaton A. The analysis of household surveys: a microeconomic approach to development policy. Econometric issues for survey data. Maryland: The Johns Hopkins University, 1997. p. 63-131.

6 - Gujarati DN. Econometria básica. 3.ed. São Paulo: Editora Makron Books; 2000.

7 - Kelley AC. The consequences or rapid population growth on human resource development: the case of education. In: Ahlburg DA., Kelley AC., Mason KO (organizadores). The impact of population growth on well-being in developing countries. Berlin: Editora Springer; 1996. p. 67-137.

8 - Montgomery MR., Lloyd C. The consequences of unintended fertility for investments in children: conceptual and methodological issues. New York: Editora The Population Council, 1996. 28 p. (Working Papers, n.89)

9 - Pazello ET. A maternidade e a mulher no mercado de trabalho. [Tese de doutorado] São Paulo: Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo; 2004.

10 - Rosenzweig MR, Wolpin KI. Testing the quantity-quality fertility model: the use of twins as a natural experiment. Econometrica,1980; 8, 1, 227-240.

APPENDIX

Table 1 - Estimates of all the variables included in the model, and with of the interactions. 1984 and 1985 PNAD's sample.

Variables	Without interations		Interation of TG And mother's education		Interation of TG And regions[1]		Interation of TG and status		Extreme case 1		Extreme case 2		With all interactions	
	Coef.	S. E.	Coef.	S. E.	Coef.	S. E.	Coef.	S. E.	Coef.	S. E.	Coef.	S. E.	Coef.	S. E.
Twin Rate	-0.40**	0,14	-0.58***	0.26	-0.46***	0.17	-0.44***	0.14	-0.43***	0.16	-0.41***	0.15	-0.67**	0.27
Maternal Education (0 to 4 years = 1)	-0.61***	0,02	-0.61***	0.02	-0.60***	0.02	-0.61***	0.02	-0.61***	0.02	-0.61***	0.02	-0.61***	0.02
Head of household (Woman = 1)	-0.13***	0,02	-0.13***	0.02	-0.13***	0.02	-0.13***	0.02	-0.13***	0.02	-0.13***	0.02	-0.13***	0.02
Status (Rural = 1)	-0.36***	0,02	-0.36***	0.02	-0.33***	0.02	-0.36***	0.02	-0.36***	0.02	-0.36***	0.02	-0.33***	0.02
Regions														
North	-0.33***	0,02	-0.33***	0.02	-0.38***	0.02	-0.33***	0.02	-0.33***	0.02	-0.33***	0.02	-0.38***	0.02
Northeast	-0.31***	0,02	-0.31***	0.02	-0.38***	0.01	-0.31***	0.02	-0.31***	0.02	-0.31***	0.02	-0.38***	0.01
South	0.28***	0,02	0.28***	0.02			0.28***	0.02	0.28***	0.02	0.28***	0.02		
Center	-0.04	0,02	-0.04	0.02			-0.04	0.02	-0.04	0.02	-0.04	0.02		
(Southeast is omitted)														
PNAD85 (PNAD 1984 =1)	0.06***	0,02	0.06***	0.02	0.04**	0.02	0.06***	0.02	0.06***	0.02	0.06***	0.02	0.04**	0.02
Constant	1.66***	0,02	1.66***	0.02	1.72***	0.02	1.66***	0.02	1.66***	0.02	1.66***	0.02	1.72***	0.02
TG*mother's education			0.26	0.31									0.29	0.30
TG*northeast					0.38	0.33							0.39	0.34
TG*north					-0.58	0.26							-0.56**	0.25
TG*status							0.18	0.45					-0.01	0.46
Case 1									0.18	0.29				
Case 2											0.07	0.42		
Adjusted R^2	0.08		0.08		0.08		0.08		0.08		0.08		0.07	
Total observations	57,054		57,054		57,054		57,054		57,054		57,054		57,054	

Source: 1984, 1985 and 1996 to 2001 PNAD's. Own Construction
Level of significance: *** = 1%, ** = 5%, * = 10%.
1- In this interaction, the regions ommitted were: Southeast, South and Center.

Table 2- Estimates of all of the variables included in the model, and with all of the interactions. 1996 to 2001 PNAD's sample.

Variables	Without Interactions		Interation of TG And mother's education		Interation of TG and regions		Interation of TG and status		Extreme case 1		Extreme case 2		With all of the interactions	
	Coef.	S. E.	Coef.	S. E.	Coef.	S. E.	Coef.	S. E.	Coef.	S. E.	Coef.	S. E.	Coef.	S. E.
Twin Rate	-0.11**	0.06	0.04	0.09	-0.15**	0.07	-0.15**	0.07	-0.,05	0.07	-0.08	0.07	-0.03	0.09
Maternal Education	-0.27***	0.01	-0.27***	0.01	-0.27***	0.01	-0.27***	0.01	-0.27***	0.01	-0.27***	0.01	-0.27***	0.01
(0 to 4 years = 1)														
Status	-0.15***	0.01	-0.15***	0.01	-0.14***	0.01	-0.15***	0.01	-0.15***	0.01	-0.15***	0.01	-0.14***	0.01
(Rural = 1)														
Head of household	-0.06***	0.01	-0.06***	0.01	-0.07***	0.01	-0.06***	0.01	-0.06***	0.01	-0.06***	0.01	-0.07***	0.01
(Woman = 1)														
Regions														
Northeast	-0.23***	0.01	-0.23***	0.01	-0.28***	0.01	-0.23***	0.01	-0.22***	0.01	-0.23***	0.01	-0.28***	0.01
North	-0.16***	0.01	-0.16***	0.01	-0.21***	0.01	-0.16***	0.01	-0.16***	0.01	-0.16***	0.01	-0.21***	0.01
Center	0.05***	0.01	0.05***	0.01			0.05***	0.01	0.05***	0.01	0.05***	0.01		
South	0.17***	0.01	0.17***	0.01			0.17***	0.01	0.17***	0.01	0.17***	0.01		
(Southeast is ommitted)														
PNAD97	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01
PNAD98	0.02***	0.01	0.02***	0.01	0.02***	0.01	0.02***	0.01	0.02***	0.01	0.02***	0.01	0.03***	0.01
PNAD99	0.08***	0.01	0.08***	0.01	0.08***	0.01	0.08***	0.01	0.08***	0.01	0.08***	0.01	0.08***	0.01
PNAD2001	0.13***	0.01	0.13***	0.01	0.13***	0.01	0.13***	0.01	0.13***	0.01	0.13***	0.01	0.13***	0.01
(PNAD 1996 =1)														
Constant	1.21*	0.01	1.21***	0.01	1.26***	0.01	1.21***	0.01	1.21***	0.01	1.21***	0.01	1.26***	0.01
TG*mother's education			-0.40***	0.12									-0.48***	0.13
TG*northeast					0.16	0.17							0.18	0.19
TG*north					-0.06	0.18							0.00	0.18
TG*status							0.23	0.20					0.39*	0.23
Case 1									-0.52***	0.13				
Case 2											-0.56***	0.17		
Adjusted R^2	0.06		0.06		0.06		0.06		0.06		0.06		0.06	
Total observations	149,177		149,177		149,177		149,177		149,177		149,177		149,177	

Source: 1984, 1985 and 1996 to 2001 PNAD's. Own Construction
Level of significance: *** = 1%, ** = 5%, * = 10%.
1- In this interaction, the regions ommitted were: Southeast, South and Center.