DETERMINANTS OF HOUSEHOLDS EXPENDITURE
IN BASIC EDUCATION IN COLOMBIA

A Thesis
submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Master of Public Policy
in Public Policy

By

Ana Maria Rojas Villamil, B.A

Washington, DC
April 13, 2012
DETERMINANTS OF HOUSEHOLDS EXPENDITURE IN BASIC EDUCATION IN COLOMBIA

Ana Maria Rojas, B.A.
Thesis Advisor: Omar Robles, PHD

ABSTRACT

The importance of early childhood education is indisputable. In the last decades developing countries have prioritized their budget towards education but despite those efforts, institutional investment in education still does not translate in higher enrollment rates in early education. Public investment can provide educational facilities, quality, and other institutional measures but only household’s investment will enable its utilization. This study pursued to identify poor household’s constraints to investment in basic education in Colombia, to understand how educational policy actions should be focalized. The OLS and fixed effects analysis for 2008 and 2003 estimates evidence the gap between urban and rural households, suggesting that the educational attainment of the parents is the most robust determinant of expenditure in early education and that data for single-mothers as the most vulnerable group is not conclusive enough. Policy targeting this population should be revised and resources prioritized towards specific programs to improve educational attainment of parents and decentralization measures. This kind of research provides empirical evidence for Colombia and would be useful for policy makers who seek a general understanding of educational regulatory policy and its implications to national development.
The research and writing of this thesis is dedicated to my beautiful mom and my loving father.

Con mucho amor,
Ana Maria Rojas
# Table of Contents

1. INTRODUCTION ................................................................................................................... 1
2. LITERATURE REVIEW ........................................................................................................ 4
3. CONCEPTUAL FRAMEWORK ............................................................................................ 6
   3.1 Determinants .................................................................................................................. 6
      3.1.1 Household’s characteristics ...................................................................................... 6
      3.1.2 Household’s head characteristics ............................................................................. 7
      3.1.3 Region Fixed Effects ................................................................................................. 7
4. DATA AND METHODS ........................................................................................................ 8
   4.1 Data Source ..................................................................................................................... 8
   4.2 Methodology .................................................................................................................... 9
      4.2.1 Household’s characteristics .................................................................................... 10
      4.2.2 Household’s head characteristics .......................................................................... 10
      4.2.3 Region Fixed Effects ............................................................................................... 10
5. RESULTS .............................................................................................................................. 11
   5.1 Descriptive Results .......................................................................................................... 11
   5.2 Regression Results ........................................................................................................... 14
      5.2.1 Household Size ........................................................................................................ 14
      5.2.2 Urban Households .................................................................................................. 15
      5.2.3 Female households .................................................................................................. 15
      5.2.4 Marital Status .......................................................................................................... 16
      5.2.5 Household's Head Level of Education .................................................................... 17
      5.2.6 Household's Head Age ............................................................................................ 17
6. DISCUSSION ........................................................................................................................ 18
   6.1 Validity Threats .............................................................................................................. 18
   6.2 Statistical Limitations .................................................................................................... 19
   6.3 Policy implications ......................................................................................................... 20
7. CONCLUSION ....................................................................................................................... 22
8. APPENDICES ....................................................................................................................... 24
1. **INTRODUCTION**

In the last decade early education issues have particularly captures the attention of policymakers after several studies\(^1\) have been showing a consistent and close relationship between school attainment and economic development. Early education has become one of the global policy priorities and the way of establishing equal educational opportunities for children around the world\(^2\). International initiatives such as Education for All (EFA) and the education Millennium Development Goals, have promote developing countries to prioritize their budget towards education investment to improve educational infrastructure, quality of education, and access. Despite all these efforts, these countries report that there are no significant changes in economic development and still 77 million children all over the world are not enrolled in preschool and primary education\(^3\).

Colombia is an interesting case, where the prioritization of education expenditures doesn’t match with higher levels of primary education enrollment. Following international trends policymakers in the country have focused on the allocation of resources in institutional efforts of provision of education (infrastructure, quality and access). In 2006 education expenditure was arround11% of the total of government expenditure, 4.8% of the country’s GDP, a very important share considering Colombia is a middle-income country (Graph 1).

\(^2\) UNESCO, 2008
\(^3\) Department of Education and Child Development, Fact Sheet - Education and Training Reform Act 2006.
The public expenditure per capita as a percentage of the GDP in 2006 was estimated in 19.2%. But participation rates still remained very low for this same year. In the case of preschool education net enrollment ratio was estimated in 34.6% for a preschool age population of 2,685 children and for primary education 88.5% for a primary school age population of 4,568 children⁴.

Why this inconsistency between government’s action and policy outcomes? Why so many children remain outside the education system? Are resources and policy actions addressing the demand side of the early education issue in Colombia? Public investment can provide

---

educational facilities, quality, and other institutional measures but only household’s investment will enable its utilization.

Ignoring household’s expenditure in education restrictions or enhancers can be costly for educational policy in the long run. Some affirmative actions have been part of social policy in Colombia targeting “vulnerable groups” for resource allocation and policy actions, but how is that vulnerable population identified when addressing the early education issue?

Several household’s characteristics, which could be called social, cultural, educational, occupational and other factors, might influence the nature and quantum of investments that households make in the education of their children, and define where policy actions should be prioritized to. Unfortunately, there is not much research on the extent of household expenditure on education in Colombia or on the determinants of household’s expenditure decisions.

The purpose of the present study is to examine various dimensions under which Colombian household’s take decisions about investing in early education. Identifying poor household’s constraints to investment in basic education in Colombia could help understanding how educational policy actions should be focalized. This analysis is worth pursuing as empirical evidence of the low level of Colombian government engagement with educational policy in a developing country where educational reforms have failed throughout the years.
2. LITERATURE REVIEW

Becker (1964) developed the theory of human capital development, arguing that when investment in education is undertaken it will provide skills to the population who will be capable of working, and will eventually lead to economic growth. The World Development Report (1984) covered much of the earlier literature concerning the impact of education on lowering fertility and improving maternal health. Behrman (1990), in a review of human resources and poverty found strong evidence for the impact of maternal years of schooling on child health. Since 1990 this literature has continued to confirm in more detailed ways the positive impact of primary education on the health of households across the generations.

Also, the importance of early childhood education and care programs on the development of children has been studied in many empirical studies around the world (Hendey and Steuerle, 2010). These studies show evidence of the effect of educational policy in improving economic welfare and health, reducing inequalities and promoting more democratic political systems.

On the other hand, research on household expenditures on education is very limited. The issue has not attracted wide attention of researchers so far. However, broadly within the framework of household’s decision-making behavior models, understanding how parents allocate resources across children has been researched in economics. Becker (1967, 1981) examined an individual maximization model, where decisions regarding investment in education are mainly made on the basis of efficiency considerations; while Behrman, Pollak and Taubman (1982) introduced a ‘family’ model with several other considerations, including equity between several children of
the family, sons and daughters, younger and older children, etc., (or prejudices and biases, e.g., say discrimination against girls), concluding that decisions regarding investment in education in general and more particularly at lower levels of education are made by families and rarely by the individual concerned. Hence, family/household expenditure function is considered appropriate.

McMahon (1984) developed a future-oriented family utility function to explain why families invest in education in USA. His investment demand and supply functions included variables on expected nonmonetary returns, family disposable income, tax subsidies, student loans, family size (number of brothers and sisters), order of birth, and the demand function was estimated with the help of academic scores, and schooling level of parents. Ability of the children in studies and mother’s education were found to be very important. Williams (1983) tried to explain the trends in private expenditures on education in Australia with the help of government expenditures, real price index of the cost of education, real personal disposable income and the demographic term.

More recently and closely related to this thesis study, Tunali, I. (2000) studied the determinants of school attainment of boys and girls in Turkey including individual, household and community factors, introducing fixed effects for regions, and Tilak (2000) made a brief analysis of determinants of household expenditure on education in India, to explain elasticity between institutional and private expenditures, also using households’ demographic characteristics.
3. CONCEPTUAL FRAMEWORK

This study pursues an statistical analysis for Colombian household’s expenditure in education, introducing household’s cultural and socio-economic characteristics as determinants of families behavior when investing in education of children (preschool and primary education), with a particular emphasis on the household’s head characteristics, and including fixed effects for regions, to control for regional differences that might affect investment behavior. Decisions regarding investment in education in general and more particularly at lower levels of education are made by families and rarely by the individual concerned. Hence, family-household expenditure function is considered appropriate in the present context.

3.1 Determinants

Household’s expenditure decisions depend on various factors related to the household characteristics, but in the case of expenditure decisions on the education for the smallest children, particularly factors related with the household family structure characteristics.

3.1.1 Household’s characteristics:

One of the main characteristics inherent to a household is the number of members of the household. The expected relationship between education expenditure and the size of the household is ambiguous. On one hand, resources have to be distributed between more members, reducing the availability for education expenditures, but on the other hand, probably more people contribute to the household budget, which might positively affect allocation of extra resources in basic education.
Other exogenous variables that are inherent to the household might affect the expenditures in education: whether a household is located in an urban area for instance, since urban centers concentrate more educational institutions, increasing competition in the provision and facilitating access. Also benefits related with low transportation costs, and better quality of the education.

3.1.2 Household’s head characteristics:

In general, policy tend to address affirmative actions to single-mothers in many countries, because they are considered a very vulnerable group in the population, and many resources are focalized to help their children. Colombia is not an exception, but do the numbers support this general assumption? Socio-demographic characteristics of the household’s head like gender, marital status, level of education, and age, should be examined. Do married parents invest more in education than single parents? Do single women invest less in the education of their children than single men? Do older parents invest more in education? How does the level of education of the household head affect their decision to invest in education?

3.1.3 Region Fixed Effects:

Although high levels of enrollments have been achieved at the primary school level for both boys and girls in much of Colombia (APENDIX 1), substantial regional differences might remain⁵. The political and administrative division of the country implies a decentralization of regulations and resources, which might affect household’s investment decisions.

---

4. DATA AND METHODS

4.1 Data Source

The analysis uses the National Households Survey, conducted by the Colombian National Department of Statistics (DANE) in 2008. This data covered 13,600 homes in a geographical area for National Total, Cities, Capital and 13 Metropolitan areas. This study takes a subsample of 3,013 households, to simplify the analysis and focus on basic education expenditure variables in “poor households”. The subsample meets the following criteria:

- One family households: Because most of the analysis focused on the household’s head characteristics, household’s with more than one head were excluded to simplify the analysis.
- Households with children under 14 years: Because the study is only seeking to draw conclusions about expenditure on basic education, this is the appropriate group that generally would be attending child-care, preschool and primary school.
- Households that have expenditures in education higher than 0, and below 3 million pesos (1,500 dollars) per child: This allows using the logarithm functional form of expenditures, and establishes a limit that focused on the poor population.

Unfortunately the 2008 Survey did not follow the same households that the previous National Households Survey in 2003. Despite this fact, for aggregate comparative purposes, this analysis also includes data from the 2003 survey which has detailed observations for children under 5 years old. This data covered 24,000 homes in a geographical area for of nine regions. This study takes a subsample of 2,059 households by the same criteria that 2008 data.
4.2 Methodology

*Multivariable regression analysis with regional fixed effects:* This study estimates an expenditure function to identify the effect of the household’s specified characteristics, including fixed effects analysis to control for regional differences. The conceptual model is expressed as a functional relationship that relates expenditures to its determinants:

**Equation 1**

\[
l(\text{exed/children})_i = f(x_i)
\]

where \(l(\text{exed/children})\) refers to the logarithm of the household expenditure on basic education (preschool and primary school) divided by the number of children under 14 years old that belong to the household. The \(X\) denotes the set of independent variables specified as determinants in the model. The equation takes the following functional form:

**Equation 2**

\[
l(\text{exed/children})_i = \beta_0 + \beta_1 \text{size}_i + \beta_2 \text{urb}_i + \beta_3 \text{female}_i + \beta_4 \text{marr}_i + \beta_5 \text{hed}_i + \beta_6 \text{hage}_i + \beta_7 \text{hage}_i^2 + \mu_i
\]

When introducing fixed effects the equation takes the following form:

**Equation 3**

\[
l(\text{exed/children})_i = \beta_0 + \beta_1 \text{size}_i + \beta_2 \text{urb}_i + \beta_3 \text{female}_i + \beta_4 \text{marr}_i + \beta_5 \text{age}_i + \beta_6 \text{age}_i^2 + \beta_7 \text{ed}_i + \beta_8 \text{reg1} + \beta_9 \text{reg2} + \beta_{10} \text{reg3} + \beta_{11} \text{reg4} + \beta_{12} \text{reg5} + \beta_{13} \text{reg6} + \beta_{14} \text{reg7} + \beta_{15} \text{reg8} + \mu_i
\]
where independent variables are defined as:

4.2.1 Household’s characteristics:

$urb_i =$ Categorical variable for whether the household $i$ is in a urban area

$size_i =$ Number of children in the household $i$

4.2.2 Household’s head characteristics:

$female_i =$ 1 if the household’s head is a female, 0 if not, for household $i$

$married_i =$ 1 if the household head is married, 0 otherwise, for household $i$

$age_i =$ Age of the household’s head, for household $i$

$educ_i =$ Level of education of the household’s head for household $i$

0 = none, 1 = prekinder, 2 = kinder, 3 = preschool, 4 = primary, 5 = High School, 6 = Technician, 7 = Technologic, 8 = Undergraduate, 9 = Graduate, 10 = higher

4.2.3 Region Fixed Effects:

9 regions:

Atlantic (reference group), 1 = Oriental, 2 = Central, 3 = Pacifica, 4 = Bogotá, 5 = San Andrés, 6 = Amazonia Orinoquia, 7 = Antioquia, 8 = Valle
5. RESULTS

5.1 Descriptive Results

Column (1) on Table 1 shows the multivariable regression of all the household characteristics and family structure variables described in this study for 2008 on the log of expenditure/ number of children under 14; column (2) introduces the regional fixed effects on the same equation. On the other hand columns (3) and (5) show the separate effect for females and males on the analysis, while (4) and (6) show fixed effect for each gender respectively.

Table 1: Household's Expenditure in Education in Colombia 2008

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Household's Characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>-0.1</td>
<td>-0.1</td>
<td>-0.0943</td>
<td>-0.0962</td>
<td>-0.1051</td>
<td>-0.1047</td>
</tr>
<tr>
<td></td>
<td>[0.010]**</td>
<td>[0.010]**</td>
<td>[0.016]**</td>
<td>[0.016]**</td>
<td>[0.014]**</td>
<td>[0.014]**</td>
</tr>
<tr>
<td>Urban</td>
<td>0.1654</td>
<td>0.1408</td>
<td>0.2233</td>
<td>0.1802</td>
<td>0.1172</td>
<td>0.103</td>
</tr>
<tr>
<td></td>
<td>[0.046]**</td>
<td>[0.048]**</td>
<td>[0.070]**</td>
<td>[0.074]**</td>
<td>[0.083]**</td>
<td>[0.064]**</td>
</tr>
<tr>
<td>Household's Head Characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.0483</td>
<td>0.0447</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.042]</td>
<td>[0.042]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>0.0047</td>
<td>0.0022</td>
<td>0.0364</td>
<td>0.0332</td>
<td>-0.0236</td>
<td>-0.0268</td>
</tr>
<tr>
<td></td>
<td>[0.042]</td>
<td>[0.042]</td>
<td>[0.064]</td>
<td>[0.064]</td>
<td>[0.057]</td>
<td>[0.057]</td>
</tr>
<tr>
<td>Educational Attainment</td>
<td>0.0363</td>
<td>0.0367</td>
<td>0.0425</td>
<td>0.048</td>
<td>0.0307</td>
<td>0.0294</td>
</tr>
<tr>
<td></td>
<td>[0.012]**</td>
<td>[0.013]**</td>
<td>[0.019]**</td>
<td>[0.020]**</td>
<td>[0.017]**</td>
<td>[0.017]**</td>
</tr>
<tr>
<td>Age</td>
<td>0.0038</td>
<td>0.0036</td>
<td>0.0017</td>
<td>0.0012</td>
<td>0.0055</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>[0.008]</td>
<td>[0.008]</td>
<td>[0.012]</td>
<td>[0.012]</td>
<td>[0.010]</td>
<td>[0.010]</td>
</tr>
<tr>
<td>R²</td>
<td>0.035</td>
<td>0.035</td>
<td>0.033</td>
<td>0.033</td>
<td>0.038</td>
<td>0.038</td>
</tr>
<tr>
<td>Observations</td>
<td>3013</td>
<td>5669</td>
<td>1407</td>
<td>1407</td>
<td>1606</td>
<td>1606</td>
</tr>
</tbody>
</table>

Notes: The table shows OLS estimates in (1), (3) and (5) and fixed effects in (2), (4) and (6). Standard Errors are in brackets. A single asterisk denotes significance at the 10% level, two asterisks denotes significance at the 5%.
Table 2 shows the separate analysis for expenditures made by households in childcare and preschool vs. expenditures addressed to primary school. In this case, the OLS regression of the family structure variables for 2008 is on the log of expenditure on preschool education/children under five in column (1), and the log of expenditure on primary school/children under fourteen (3), including fixed effects columns (2) and (4) in each case.

Table 2: Household's Expenditure in Education in Colombia 2008 (Preschool vs. Primary)

<table>
<thead>
<tr>
<th>Variable</th>
<th>log (Expenditure Pre-school/ch &lt;5)</th>
<th>log (Expenditure Primary/ch&lt;14)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fixed Effects</td>
<td>Fixed Effects</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td><strong>Household's Characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>-0.1445</td>
<td>-0.1129</td>
</tr>
<tr>
<td></td>
<td>[0.030]**</td>
<td>[0.010]**</td>
</tr>
<tr>
<td>Urban</td>
<td>0.1532</td>
<td>0.145</td>
</tr>
<tr>
<td></td>
<td>[0.147]</td>
<td>[0.046]**</td>
</tr>
<tr>
<td><strong>Household's Head Characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.0492</td>
<td>0.0481</td>
</tr>
<tr>
<td></td>
<td>[0.133]</td>
<td>[0.042]</td>
</tr>
<tr>
<td>Married</td>
<td>-0.251</td>
<td>0.0253</td>
</tr>
<tr>
<td></td>
<td>[0.133]**</td>
<td>[0.042]</td>
</tr>
<tr>
<td>Educational Attainment</td>
<td>0.0134</td>
<td>0.0345</td>
</tr>
<tr>
<td></td>
<td>[0.037]</td>
<td>[0.013]**</td>
</tr>
<tr>
<td>Age</td>
<td>0.0207</td>
<td>0.0051</td>
</tr>
<tr>
<td></td>
<td>[0.023]</td>
<td>[0.008]</td>
</tr>
<tr>
<td>Age Squared</td>
<td>-0.0001</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td>[0.000]</td>
<td>[0.000]</td>
</tr>
<tr>
<td>R²</td>
<td>0.066</td>
<td>0.044</td>
</tr>
<tr>
<td>Observations</td>
<td>407</td>
<td>2873</td>
</tr>
</tbody>
</table>

Notes: The table shows OLS estimates in (1) and (3) and fixed effects in (2) and (4). Standard Errors are in brackets. A single asterisk denotes significance at the 10% level, two asterisks denotes significance at the 5%.
Table 3 shows the OLS analysis for the 2003 dataset, which only includes information of expenditures on education of children under 5 years old. The multivariable regression of the family structure variables on the log of expenditure/children under 5 in column (1); the linear correlation between these variables and the quadratic function for the age of the household’s head, column (2) introduces the regional fixed effects on the same equation. On the other hand columns (3) and (4) show the separate effect for females and males on the analysis, while (4) and (6) show fixed effect for each gender respectively.

Table 3: Household's Expenditure in Education in Colombia 2003

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1) Fixed Effects</th>
<th>(2) Fixed Effects</th>
<th>(3) Male Fixed Effects</th>
<th>(4) Female Fixed Effects</th>
<th>(5) Male Fixed Effects</th>
<th>(6) Female Fixed Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Household's Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>-0.0617</td>
<td>-0.0106</td>
<td>-0.0763</td>
<td>-0.0072</td>
<td>-0.0453</td>
<td>-0.0108</td>
</tr>
<tr>
<td></td>
<td>[0.016]**</td>
<td>[0.014]</td>
<td>[0.022]**</td>
<td>[0.020]</td>
<td>[0.024]**</td>
<td>[0.021]</td>
</tr>
<tr>
<td>Urban</td>
<td>-0.8139</td>
<td>-0.1313</td>
<td>-0.7695</td>
<td>-0.0954</td>
<td>-0.9243</td>
<td>-0.2192</td>
</tr>
<tr>
<td></td>
<td>[0.059]**</td>
<td>[0.061]**</td>
<td>[0.068]**</td>
<td>[0.072]**</td>
<td>[0.121]**</td>
<td>[0.0610]**</td>
</tr>
<tr>
<td><strong>Household's Head Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-0.1575</td>
<td>-0.129</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.123]</td>
<td>[0.109]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>0.1384</td>
<td>0.1699</td>
<td>0.2333</td>
<td>0.2082</td>
<td>0.1033</td>
<td>0.149</td>
</tr>
<tr>
<td></td>
<td>[0.130]</td>
<td>[0.115]**</td>
<td>[0.223]</td>
<td>[0.199]</td>
<td>[0.163]</td>
<td>[0.143]</td>
</tr>
<tr>
<td>Educational Attainment</td>
<td>0.1184</td>
<td>0.0765</td>
<td>0.1074</td>
<td>0.0635</td>
<td>0.1441</td>
<td>0.0984</td>
</tr>
<tr>
<td></td>
<td>[0.025]**</td>
<td>[0.022]**</td>
<td>[0.029]**</td>
<td>[0.026]**</td>
<td>[0.049]**</td>
<td>[0.043]**</td>
</tr>
<tr>
<td>Age</td>
<td>0.0128</td>
<td>-0.0014</td>
<td>0.0371</td>
<td>0.0294</td>
<td>-0.0167</td>
<td>-0.0359</td>
</tr>
<tr>
<td></td>
<td>[0.015]</td>
<td>[0.013]</td>
<td>[0.019]**</td>
<td>[0.017]**</td>
<td>[0.045]</td>
<td>[0.021]**</td>
</tr>
<tr>
<td>Age Squared</td>
<td>-0.0001</td>
<td>0.00001</td>
<td>-0.0003</td>
<td>-0.0003</td>
<td>0.0001</td>
<td>0.00001</td>
</tr>
<tr>
<td></td>
<td>[0.0001]</td>
<td>[0.0001]</td>
<td>[0.0002]**</td>
<td>[0.0001]**</td>
<td>[0.0002]**</td>
<td>[0.0002]**</td>
</tr>
<tr>
<td>R²</td>
<td>0.131</td>
<td>0.087</td>
<td>0.078</td>
<td>0.087</td>
<td>0.109</td>
<td>0.075</td>
</tr>
<tr>
<td>Observations</td>
<td>2059</td>
<td>2059</td>
<td>1427</td>
<td>2059</td>
<td>632</td>
<td>632</td>
</tr>
</tbody>
</table>

Notes: The table shows OLS estimates in (1), (3) and (5) and fixed effects in (2), (4) and (6). Standard Errors are in brackets. A single asterisk denotes significance at the 10% level, two asterisks denotes significance at the 5%.
5.2 Regression Results

The analysis of household level data provides important insights in understanding the determinants of household expenditures on education. Despite the households in the 2008 survey are not the same ones in the 2003 survey, the regression analysis for both years adds consistency to some of the results. The number of observations is considerable, and the findings show that family structure factors influence the levels of early education expenditures. Also, some strong regional effects are evidenced, revealing differences in the social structures across regions in Colombia, and some differences between subsamples of preschool vs. primary school expenditures. Lastly, some gender gaps are found that persist over time.

5.2.1 Household Size:

For the 2008 households, controlling for the number of members of the family show an on average decrease of approximately 10% on the expenditure on education. This outcome is significant at the 5% level and doesn’t vary much when introducing the regional fixed effects. For the preschool sample it shows an effect if 14%, while for primary school is 11%. These outcomes support the argument that fewer resources are available for educational investment in the case of Colombian households.

For the 2003 households, the negative effect is much smaller though. An effect of 6% decrease on average, and when including the fixed effect it drops to 1% decrease. This effect is mainly explained by female-headed household’s behavior. Aside of the evidenced differences between the 2003 and the 2008 surveys, these results might be showing a very strong regional heterogeneity present in 2003 that might have been moderated through the years through educational homogenization and standardization reforms related to enrollment and cost of basic education at a National Level⁶.

5.2.2 Urban Households

The regression analysis for 2008 shows that urban households increase expenditures on education in 16% more compared to rural households. This outcome is expected since urban centers are more likely to provide more child education options and facilitate access. When controlling by regions the effect is moderated to 14% probably because in most regions rural areas are less insulated from urban centers, facilitating access to education centers and closing the gap. The gap between male and female households though is considerable. While males in urban households are estimated to increase the expenditure in almost 22%, females only do it in 11%. This might be explained in part by an income gender gap. Also, a regional effect is present in this case: the figures drop to 18 and 10% respectively.

On the other hand, for 2003 preschool education expenditures show a statistically significant positive 83% variation for urban households. But, this effect drops dramatically to 13% when controlling by region characteristics, which suggest a lot of heterogeneity among the different regions access to basic education centers. In the separate gender analysis the strong regional effect persist: a 76% effect for males’ drops to 9.5%, while for females it changes from 92% to 21%. The difference between the numbers of 2008 can be partly explained again by the homogenization and standardization reforms in 2006, which moderates the fixed effects for the 2008 data.

5.2.3 Female households

Female-headed households in 2008 have a positive but not significant effect in the overall sample of around 5% increase in expenditures on education. When breaking the sample between preschool education expenditures and primary school expenditures, introducing fixed effects reduce the effect to 3% only for preschool expenditures, which might translate in the restrictions for mothers in some regions when deciding to invest in the very first years of education of their children.
On the other hand, the 2003 data, though with no statistical significance, show a negative effect predicting a decrease of 15% on expenditures for children under 5, a consistent result with current policy actions that often address female-headed households. Probably the situation has changed for female household heads in the latter years, through government support and educational reforms. The increasing employment rates for women in Colombia, might also explain this transition.

5.2.4 Marital Status

Married household’s heads show a small positive effect of circa 0.4% on the expenditure in the education of their children in 2008. This effect is statistically not significant for the complete sample, but when breaking the sample, preschool expenditures show a significant negative effect of 25%. It could be argued that married parents with preschoolers behave differently since one of the parents can stay with the small children at home, and provide some kind of basic cognitive skills to their children, while the other parent works, which is not likely for single parent households. On the other hand, in the 2003 sample the effect for children under 5 years old has a different direction. The fixed effect shows a positive effect of 17% on expenditures in education and a big gap between male and female headed households, the were not present in the 2008 dataset. The analysis then shows inconsistent outcomes, from one year to the other.

Literature supports a positive incidence of stable family structures in educational outcomes. On the other hand, the probability of two incomes contributing to the household expenditure decisions should increase expenditures. But, some other considerations can be made at this point: the way of reporting marital status in a public survey might have misleading outcomes. In the first place, in a strongly catholic country like Colombia, marital status has a number of social

---

7 *Familias en Acción* is a program run by the Colombian government as part of the RAS (Red de Apoyo Social or Social Support Network) and is financed by a loan from the World Bank and the IADB. It is a Conditional Cash Transfer program, consisting of cash transfers to poor families conditional on children attending school and meeting basic preventive health care requirements. The program targets mainly female-headed households.
connotations, which might affect the answer of its respondents. Second, in poor households particularly, people remain legally married (high cost of divorce), but in reality their couples don’t live in the household or contribute at all in the households budget. Third, the change in the methodology and data collection strategy of the statistics Colombian National Department of Statistics might be reflecting the introduction of measurement errors, and other statistical problems to the 2008 dataset.

5.2.5 Household's Head Level of Education:

The level of education of the household’s head is measured in terms of years of schooling, and has a significant positive effect of 3.5% on the expenditures in basic education in 2008. When controlling by regional characteristics the effect persist, while the data show that women with higher educational attainment invest in basic education one percent point less than more educated men. When dividing the sample between preschool and primary education expenditures, a positive effect of around 1.5% for preschool and a positive significant effect of 3.5% for primary expenditures. The effect is also positive and significant for the 2003 dataset, 11%, which again shows a large regional effect, dropping to 7% and a gender gap of around 4%, where more educated females spent more than their male counterparts.

Literature supports this outcome widely arguing that education increases the awareness of the benefits of education. The results here are consistent to show that the educational attainment of the parents has a positive effect on educational outcomes of their children under 14.

5.2.6 Household's Head Age:

The household’s head age has no statistically significance for any of the 2008 subsamples considered in this analysis. For the 2003 data though, a significant positive outcome is found for male headed households, of around 4% increase for children under 5 years old and a moderated regional effect. These results though don’t really give a conclusive outcome as a determinant of expenditure in education.
6. DISCUSSION

6.1 Validity Threats
Some validity considerations should be addressed before discussing the results. In the first place, using two different years that don’t follow the same households might have misleading interpretations for external validity. The most important differences identified between the 2003 and the 2008 surveys and should be reminded throughout the results analysis. For instance, the 2003 survey includes only 30% female-headed households, compared to a 53% of the 2008 sample. On the other hand the 2003 survey includes 73% married household heads, compared to a 50% in 2008, and a much less educated sample, 1.27 levels of education on average, vs. 3.5 levels of the 2008 sample. Lastly, the 2003 survey reports on average a slightly younger sample of household heads, 41 years old, compared to 48 years in the 2008 case (APPENDIX 2).

On the other hand, for internal validity concerns, using a sub-sample that only includes households that have reported expenditures of at least $1 in basic education might be introducing a bias in the estimations, because every household with children is not being included in the sample, probably the most vulnerable households, those unable to invest at all in this type of expenditures are being left out.

Also the fact that only one family households are included in the analysis, might threat internal validity, those households might have bigger problems when investing in early education, but because most of the analysis focused on the household’s head characteristics, household’s with more than one head would have limited the analysis.
6.2 Statistical Limitations

**Heteroscedasticity**

From a cross-sectional analysis, heteroscedasticity could be expected, as varied types of households living in various parts of the country are sampled together. Using the semi-log regression equation (1) of the expenditure function, a post-examination of the estimated residual squares didn’t exhibit any systematic pattern, meaning no heteroscedasticity was present in the model.

**Multicolinearity**

On the other hand, the results show very low $R_j^2$ values. Multicolinearity can be a problem. This means there isn’t a lot of variation that is independent from the other determinants considered in the equation. Some authors have suggested a formal detection-tolerance or the variance inflation factor (VIF) for multicollinearity:

**Equation 4**

\[
\text{tolerance} = 1 - R_j^2, \quad \text{VIF} = \frac{1}{\text{tolerance}}
\]

where $R_j^2$ is the coefficient of determination of a regression of j on all the other variables. A tolerance of less than 0.20 or 0.10 and/or a VIF of 5 or 10 and above indicates a multicollinearity problem\(^8\). (APPENDIX 3)

---

\(^8\) O'Brien 2007


6.3 Policy implications

This analysis pursued to identify poor household’s constraints to investment in basic education in Colombia, to understand how educational policy actions should be focalized. In general, policy should address population that is struggling with allocating resources toward child education. The OLS estimates of the regression analysis and the fixed effects analysis helped identifying major determinants of household expenditures on education.

Rural-Urban balance of education opportunities

Concentration of economic activities, schools and training institutions in urban centers is a perennial problem in Colombia. Despite differences between 2003 and 2008 outcomes, in general data shows a persisting gap remains of 15% on average, this suggest that decentralization policies could have an important effect in household expenditures on education. If quality is improved in rural schools, and more prepared teachers are incentivized to move to rural areas, parents might also be incentivized to spend on the education of their children. Also, employment opportunities in rural areas might increase household’s budget, and more resources become available to allocate for children education. Other important items of household expenditures on education have to do with transaction costs of living in rural areas like transportation fees.

Gender oriented policy

Existing social policy in Colombia targets mainly single-mother households for eligibility in income-transfers and in-kind benefits, but in many households, as suggested before, single females can’t/won’t probe their “real” marital status, or are not necessarily the household’s children mothers. Some part of the data for 2003 in this analysis supports this policy efforts showing that significant gender differences exist and they are normally against female-headed households. But on the other hand, results are not really conclusive since they lack statistical significance and consistency for the 2008 sample. Further analysis of intra-household inequality and gender gap considerations should be studied by the government to reallocate policy programs, if the situation have change for female household heads en the latter years.
Parents Educational Attainment

The most consistent and statistically significant outcome in this analysis is the effect of educational attainment of households head educational attainment in expenditures on education. The effect though appears to be moderate. Further policy action should be focalized on increasing adult’s educational attainment, not only from the early education interventions, but also through initiatives targeting youth and adult education attainment. Educational attainment might also have employability positive outcomes for parents.

Other considerations

From the statistical point of view, the importance of the incidence of family structure in many social outcomes, further efforts should be pursue to collect information about marital status, one of the key variables in for this kind of policy analysis.
7. CONCLUSION

Many developing countries have prioritized their budget towards education investment to improve educational infrastructure and quality of education. In Colombia, despite several institutional efforts of public investing in early education, they still don’t translate into household’s budgetary considerations to invest on early education. Identifying how different factors affect expenditure in education decisions in households is crucial to do the proper evaluation of the efficiency and quality of the existing programs to promote education, implementation strategies and the role of regulatory mechanisms and incentives structures.

This study estimates an expenditure function to identify the effect of some household’s cultural and socio-economic characteristics as determinants of family’s behavior when investing in education of their children. The OLS estimates allow many important conclusions. In the first place, the concentration of economic activities, schools and training institutions in urban centers show the perennial problem in Colombia, evidence of a gap of 15% percent between urban and rural households is founded. Decentralization policies might have an important effect in household expenditures on education. If quality is improved in rural schools, and more prepared teachers are incentivized to move to rural areas, parents might also be incentivized to spend on the education of their children. On the other hand, policy actions already targets single-mother households, but evidence is really not robust enough to identify this group as the most vulnerable population among poor households in Colombia. Further analysis of intra-household inequality and gender gap considerations should be studied by the government to reallocate policy programs.
Findings in this analysis suggest that the educational attainment of the parents is the most statistically significant determinant of their expenditure in the education of their children. Policy addressing this factor should be prioritized through specific programs to improve educational attainment of parents and in general adult population.

Also, from the statistical point of view, the importance of the incidence of family structure in many social outcomes, further efforts should be pursued to collect information about many more household characteristics, to achieve proper statistical analysis. For instance, the marital status of the household head, one of the key variables for this kind of policy analysis, appears to have a lot of measurement concerns. Adequate surveys and methods are needed to make policy conclusions. How can programs nowadays address single parents with no reliable data about this households really struggling with access to basic education for their children.

This kind of research provides empirical evidence for Colombia in this field for the first time. Hopefully this analysis would be useful for policy makers who seek a general understanding of educational regulatory policy and its implications to national development when promoting and structural educational reform. In some cases though, identifying determinants is not revealing all the factors behind access, opportunities and availability of resources, but instead cultural practices and propensity of poor households to under invest in education. Further studies in the field are needed to contribute generate a better empirical picture of the complexities of households restrictions and behavior when investing in education. For instance gender gap studies regarding fertility, educational attainment and labor market participation.
8. APPENDICES

8.1 APPENDIX 1  Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>2003 Households</th>
<th>2008 Households</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2003 Households</td>
<td>2008 Households</td>
</tr>
<tr>
<td></td>
<td>2003 Households</td>
<td>2008 Households</td>
</tr>
<tr>
<td></td>
<td>Children under 5</td>
<td>Preschool</td>
</tr>
<tr>
<td><strong>Mean (SD)</strong></td>
<td><strong>Mean (SD)</strong></td>
<td><strong>Mean (SD)</strong></td>
</tr>
<tr>
<td><strong>Household's Characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expenditure/Number of children</td>
<td>$170,036 (268,381)</td>
<td>$108,840 (188,091)</td>
</tr>
<tr>
<td>Size</td>
<td>5.44 (2.39)</td>
<td>5.2 (2.20)</td>
</tr>
<tr>
<td>Urban (1 for urban, 0 for rural)</td>
<td>0.79 (0.40)</td>
<td>0.71 (0.45)</td>
</tr>
<tr>
<td><strong>Household's Head Characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.31 (0.46)</td>
<td>0.52 (0.49)</td>
</tr>
<tr>
<td>Married (1 if married, 0 other)</td>
<td>0.73 (0.44)</td>
<td>0.49 (0.50)</td>
</tr>
<tr>
<td>Level of education (1-7)</td>
<td>1.27 (1.48)</td>
<td>3.56 (1.81)</td>
</tr>
<tr>
<td>Age</td>
<td>41.8 (13.70)</td>
<td>46.5 (15.40)</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>2059</td>
<td>407</td>
</tr>
</tbody>
</table>

Note: The table shows means with standard deviations below in parentheses.
8.2 APPENDIX 2 Regressions Specifications

The analysis includes the following specifications; the numbers on the equations correspond to the columns in Table 1:

(1) OLS

\[ l(exed \text{ / children}) = \beta_0 + \beta_1 \text{size} + \beta_2 \text{urb} + \beta_3 \text{fem} + \beta_4 \text{marr} + \beta_5 \text{ed} + \beta_6 \text{age} + \beta_7 \text{age}^2 + \mu_i \]

(2) Regional Fixed Effects

\[ l(exed \text{ / children}) = \beta_0 + \beta_1 \text{size} + \beta_2 \text{urb} + \beta_3 \text{fem} + \beta_4 \text{marr} + \beta_5 \text{age} + \beta_6 \text{age}^2 + \beta_7 \text{ed} \\
+ \beta_8 \text{reg}1 + \beta_9 \text{reg}2 + \beta_{10} \text{reg}3 + \beta_{11} \text{reg}4 + \beta_{12} \text{reg}5 + \beta_{13} \text{reg}6 + \beta_{14} \text{reg}7 + \beta_{15} \text{reg}8 + \mu_i \]

(3) OLS female households

\[ l(exed \text{ children} \_fem) = \beta_0 + \beta_1 \text{size} + \beta_2 \text{urb} + \beta_3 \text{marr} + \beta_4 \text{age} + \beta_5 \text{age}^2 + \beta_6 \text{ed} + \mu_{fem} \]

(4) OLS male households

\[ l(exed \text{ children} \_mal) = \beta_0 + \beta_1 \text{size} + \beta_2 \text{urb} + \beta_3 \text{marr} + \beta_4 \text{age} + \beta_5 \text{age}^2 + \beta_6 \text{ed} + \mu_{mal} \]
8.3 **APPENDIX 3 Statistical Limitations**

*Multicollinearity*: Some authors have suggested a formal detection-tolerance or the variance inflation factor (VIF) for multicollinearity:

\[
\text{tolerance} = 1 - R_j^2, \quad \text{VIF} = \frac{1}{\text{tolerance}},
\]

where \( R_j^2 \) is the coefficient of determination of a regression of explanator \( j \) on all the other explanators. A tolerance of less than 0.20 or 0.10 and/or a VIF of 5 or 10 and above indicates a multicollinearity problem.

<table>
<thead>
<tr>
<th>Calculations</th>
<th>2003 Households</th>
<th>2008 Households</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Children under 5</td>
<td>Children under 14</td>
</tr>
<tr>
<td></td>
<td>OLS</td>
<td>FE</td>
</tr>
<tr>
<td>( R_j^2 )</td>
<td>0.131</td>
<td>0.087</td>
</tr>
<tr>
<td>Tolerance = (1-( R_j^2 ))</td>
<td>0.869</td>
<td>0.913</td>
</tr>
<tr>
<td>VIF = 1/Tolerance</td>
<td>1.15075</td>
<td>1.09529</td>
</tr>
<tr>
<td>Multicollinearity</td>
<td>NO</td>
<td>NO</td>
</tr>
</tbody>
</table>
9. REFERENCES


  www.dfid.gov.uk/pubs/files/edcostandanedpaper02.pdf


