GENDER PREFERENCES AND THE INTRAHOUSEHOLD RESOURCE ALLOCATION: EVIDENCE FROM PERU

Víctor Saldarriaga[†]

Abstract

Labor market outcomes have often been the core of discussion when referring to the economics of gender discrimination. However, little attention has been paid to the mechanics of gender preferences and parental resource allocation at intrafamilial level. Human capital investments made by parents during the childhood and youth transition of children can have direct effects on labor market outcomes of individuals, henceforth, explaining most of the wage differentials observed when comparing men and women in adulthood.

Accounting for those differences in parental preferences for gender could become difficult because for the presence of many other observed and not observed omitted characteristics of children which can affect in a distinct way both boys and girls. This paper uses two different non-parametric approaches in order to measure the unequal concern shown by parents with regard to the gender differences among children in Peru. Contrarily to the general belief of favoritism for boys, results show that girls have slightly better educational outcomes than boys. Nonetheless, differences found appear not to be statistically different from zero.

I. Introduction

There is a widespread wisdom that employers show preferences for males over females when hiring employees, even after taking into consideration similarities in observable characteristics such as schooling, age and working experience[1]. This situation is often interpreted as gender discrimination in the labor market. However, in many countries, females also receive less investment in human capital formation than males during childhood.

Researchers consider three different mechanisms in which gender differences can make parents allocate their resources unequally among their children. First, if parents take into account the observed gender wage differentials in the labor market and want to reinforce children's endowments, then it would be optimal to invest more in boys rather than girls when expected returns are higher for males. The same insight is applicable when parents are concerned in compensating children's endowments: in this particular case, parents are likely to invest more in girls when employers exhibit preferences in hiring and paying males rather than females. Second, parents may respond to differences in raising prices across gender. Third, even when having the same outcomes and raising prices for both boys and girls, parents may show preferences for one sex when the utility reported from boys'(girls') outcomes is higher than that reported from girls'(boys'). Consequently, considering these three factors mentioned above, parents can alter their behavior and favor boys or girls in their human capital investment decisions.

Latin American countries are not far away from this reality. By early and mid 90's, men in many Latin American countries had better educational outcomes then women. For example, the illiteracy rate was above five percentage points when comparing the population of men and women aged 15 or more in countries like Bolivia, El Salvador, Guatemala, Mexico, Nicaragua and Peru. In previous decades, the illiteracy rate in these countries and other Latin American countries reached a difference of more than ten percentage points. A similar picture could be seen when comparing indicators such as years of schooling and school enrollment.

It was not until the second half of the 90's when the educational outcomes gap between men and women in Latin America began to close. Because the educational policy became more inclusive and expanded the coverage, women began to show better results by the beginning of the new century. On the other hand, the arrival of new models of intervention such as the conditional cash transfers (CCT's) had its own positive effects on the demand for education. A peculiarity of this type of policy interventions is that in Mexico, the CCT's benefits differed according to sex and age of the beneficiaries during secondary education.

In this context, this paper returns to the discussion about how much closer the gender educational gap has became. Specifically, I would like to answer if preferences for giving boys a better education comparing to girls still persist at the family level for the case of Peru.

In order to answer this question, the identification strategy is based in two non-parametrical approaches that allows me to obtain the counterfactuals for making

[†] Assistant Researcher at the Group for the Analysis of Development – GRADE.

[1] See Oaxaca [1973] for the United States, Birdsall and Sabot [1985] for Brazil, Nicaragua, Kenya and Tanzania and Ñopo, Atal and Winder [2010] for eighteen countries in Latin America.

the comparisons of three educational outcomes: school attendance, an indicator of being left back in education and the grade or schooling year for age score (normative years of schooling) in a sample of children aged 6 to 17. In the first technique, I use the matching methodology for identifying identical boys and girls in a set of imposed covariates. The second technique is related with the use of male-female twins as counterfactuals within the families. In contrast to the matching methodology, the twin comparisons allow me to remove family fixed effects and make identification more accurate. Contrarily to the popular belief, girls are found to have slightly better educational outcomes than boys, especially in school attendance and the "held back" indicator. These results show a different behavior in rural areas. However, all differences in educational outcomes are found not to be statistically different from zero.

This paper is organized as follows: section II describes the previous empirical findings in the literature related to gender preferences within the family, section III introduces the theoretical considerations, section IV describes the identification strategy, section V comments the data and descriptive statistics, section VI presents the results and a detailed discussion about the identification strategy, its advantages and limitations and finally section VII summarizes and concludes.

II. Previous Empirical Findings

The economic literature related to gender preferences within the family was initially documented by Ben-Porath and Welch [1976]. In this study, authors conclude that the tendency to have more children depends on the sex composition of earlier children: parents are likely to have more children when firstborn children are girls rather than boys. Later, Behrman, Pollak and Taubman [1986] developed a theoretical framework for explaining the mechanics of parental preferences which may favor girls or favor boys after generalizing their original parental resource allocation model introduced in 1982. In the generalized model, the authors dropped out the assumption that parents only value returns to education in the labor market, allowing them to operate also through the marriage market. Although the empirical findings support the relevancy of marriage market outcomes on the allocation of human capital investments among children, the authors found no evidence that parental preferences favor boys.

Recently, some researchers have turned their attention to study the effects of child gender on the time allocation of parents and family stability. For example, Dahl and Moretti [2004] and Lundberg [2005] found that parents with girls are more likely to be divorced and that boys are more likely than girls to be living with their fathers. According to the latter, it is possible to think that parental preferences for sons may be driven by parents instead of mothers. However, it has also been found that more highly educated fathers are likely to reduce their hours of work while their wives maintain the same working hours when they have a very young son rather than a daughter. Because this is an indicator of decreasing the household specialization, it can be argued that sons may receive less care than girls.

In the case of developing countries, a recent investigation made by Grant and Behrman [2010] shows that, despite the gap in school entry favoring boys, girls tend to progress faster than boys and also have equal or greater schooling attainment than boys. These findings are consistent with the trajectory of gender gaps in developed regions, where girls are found to be ahead boys in educational outcomes (i.e.: higher secondary school and college completion rates). To conclude, Gertler and Glewwe [1992] found that, in the case of rural Peru, parents place more value in sending boys to secondary school than on sending girls. This result seems to arise due to the fact that returns to secondary education for girls in rural Peru are not very strong. For this reason, dividing the sample by urban and rural areas may be more relevant in the empirical analysis.

III. Theoretical Considerations

To study the existence of gender preferences within the family, I base my analysis on the Behrman, Pollak and Taubman [1982] model (BPT) of parental preferences and provision for progeny.

In this model, parents are assumed to have separable welfare functions defined over the expected human capital dependent (HCD) incomes of each of the *n* children:

$$w = W\left(Y_1, Y_2, \ldots, Y_n\right),$$

where *W* is the parental welfare function, Y_i is the expected HCD income of the *i-th* child and *n* is the number of children. Additionally, parents face restrictions of raising resources, mainly in the human capital formation of their children. Education and other human capital investment prices are assumed to be fixed and are denoted by P_E and P_X respectively, so the family constraint could be written as follows:

$$\sum_{i=1}^{n}\left(P_E E_i + P_X X_i\right) \leq R,$$

where E_i and X_i are the quantities of education and other human capital investments assigned to the *i-th* child. The remaining constraints are the HCD income production functions of each of the *n* children:

$$Y_i = Y_i\left(E_i, X_i, S_i, G_i\right),$$

where S_i denotes the sex of the *i-th* child and G_i represents its "natural" and genetic endowments [2].

Parents are supposed to maximize their welfare function considering the family resource constraint and the HCD income production functions of their children. I return later with a detailed discussion about the HCD income production functions. Assuming an interior solution for this constrained maximization problem, the first order conditions for the parental human capital investment in education of the *i-th* child is:

$$\frac{\partial W}{\partial Y_i}\frac{\partial Y_i}{\partial E_i} - \lambda P_E = 0,$$

where λ is the family specific Lagrangian multiplier for the nominal budget constraint of resources devoted to investments in the human capital formation of family's children. Expressing the first order conditions in ratio form yields:

$$\frac{\partial W/\partial Y_i}{\partial W/\partial Y_j} = \frac{\partial Y_j/\partial E_j}{\partial Y_i/\partial E_i},$$

which is the usual tangency condition required for maximizing the parental welfare function with regard to the HCD income possibility frontier for the *i-th* versus the *j-th* child.

Under the assumption that the HCD income possibility frontier arising from comparing child *i* and child *j* is symmetric with respect to the origin, the tangency point which maximizes the parental resource allocation problem depends exclusively in the shape adopted by the parental welfare function. The former assumption implies that the expected HCD income of each child equals that of his/her sibling. I now discuss in detail the HCD income production function and the influence of gender preferences in the parental welfare function.

HCD Income Production Function

So far I have mentioned the arguments involved in the expected HCD income production function of children. Considering the questions posed in this paper, I focus the discussion in the sex, S_i, and "natural" and genetic endowments, G_i, of the children. In the context where parents allocate the same quantities of education and other human capital investments for each child (e.g.: children attend to the same school and enjoy the same educational benefits given by their parents), the sex and "natural" and genetic endowments can have a direct impact on the equity of the expected HCD income among children so parents may response to those effects and reallocate the family resources with the aim of achieving equity in the distribution of the expected HCD incomes of their children. This means that parents can be able to trade off the HCD income of a given child when the expected HCD income of he/she is beyond an arbitrary threshold that boundaries the accepted inequality of the expected HCD incomes of all children. This threshold, of course, can be equal or different among children. The latter implies that parents show unequal concern with respect to their progeny.

However, it is difficult to distinguish between unequal concern and gender preferences when the existing difference is exclusively driven by sex differences across children. For example, let us consider the case in which parents know about the gender wage differentials in the labor market. In a two-child family where one child is male and the other is a female, it is possible that parents want to compensate the expected HCD incomes of both children so they allocate additional resources to the girl with regard to the boy. In this particular scenario, assuming that the gender wage gap is large enough, one can observe that parents trade more education or other human capital investments (say, a better quality school, particular teachers, languages and computer lessons, etc.) in benefit of the girl child, but because of the desired equity in the distribution of children expected HCD incomes they want to achieve, hence, there is not a, strictly speaking, gender preference that would favor girls.

Additionally, parents may consider other factors in the expected HCD income production function of children.

[2] In a later paper, the authors consider the expected HCD income of spouses and the weights that parents place to the expected earnings of the child's spouse as arguments for the HCD income production functions of children. For simplicity, I do not consider those arguments but will be discussed opportunely.

For instance, if parents take into account the marriage market outcomes (say, assortative mating with regard to education, socio-economic status and even beauty) and expected spouses earnings, it is possible that all these arguments unbalance the distribution of the expected HCD income of children. If parents respond to these differentials, then they will reallocate inputs on those "disadvantaged" children in order to achieve equity. Nonetheless, it is possible that these responses would benefit one of both sexes more than the other so that this reflects unequal concern of parents with respect to gender.

On the other hand, let us consider the differences in "natural" and genetic[3] endowments among children. If a child is genetically better endowed (innate abilities), then it is possible to observe that the amount of time or investments spent in him/her gets reduced because parents can either increase their working hours or allocate more resources to consumption as they know that the child has more chances to *succeed* in adulthood. Moreover, parents can reallocate their resources in favor of those less endowed children with the objective of getting tantamount expected HCD incomes. There exist, of course, many other stories that can be tell with respect to differences in genetic endowments of children. Let us consider now the case of "natural" endowments. In a given family, if one child is better "naturally" endowed than the rest of his/her siblings (say because the mother had a better nutritious diet during pregnancy that results in a better birthweight of the child[4]), then it is possible that this results in greater expected HCD income for that child with respect to his/her siblings. Hence, if parents observe those differences across expected HCD incomes of their children, then it is possible that they would want to compensate those differences.

Recalling all these considerations, some conditions and assumptions are required for a correct identification of the unequal concern show by parents when considering gender differences among siblings. Before setting the context through identification is plausible, I discuss briefly some insights regarding the parental welfare function.

Gender Preferences and Parental Welfare Function

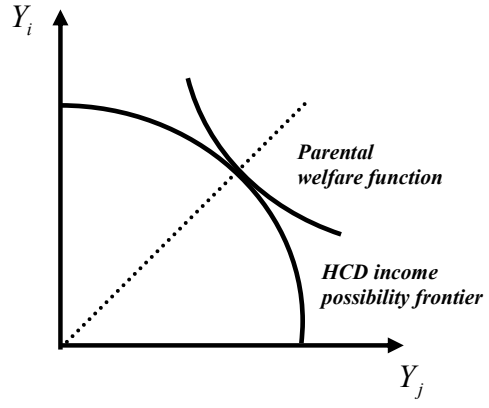
When referring to the parental welfare function, there exist mainly three different sources of variation. The first one is related to the thresholds parents define in order to achieve equity in the expected lifetime earnings of children. As mentioned above, these thresholds can be fixed or may vary across children.

The second source of variation corresponds to weights placed by parents favoring few children of their progeny. In this way, the relative success of one child with respect to another can produce that parents feel more or less satisfaction, even if the expected HCD income of the children are exactly the same. This idea is directly related with parental preferences for spoiled children and is also applicable when considering gender differences.

The last source of variation corresponds to risk aversion shown by parents. Then, if parents are very risk averse, household resources will be entirely allocated to the "disadvantaged" child in order to achieve equity in the expected lifetime earnings of offspring. Another possibility is that parents would like to reinforce child "advantages" with respect to his/her siblings, so that resources will be allocated favoring the more skilled child. Finally, in the middle of both extremes, there exists the context where parents are risk neutral, so that they allocate the same amount of inputs to every children in the family without considering differences in expected HCD incomes of their progeny.

All these ideas can be explained graphically. To be better understood, let us consider the outline established in figure 1. In this particular case of a two-child family, child *i* and child *j* have exactly the same expected lifetime earnings, so any differences in outcomes is exclusively driven by parental preferences.

FIGURE 1
OPTIMUM PARENTAL INTRAFAMILIAL RESOURCE ALLOCATION OF HUMAN CAPITAL INVESTMENTS



As pointed out by Behrman, Pollak and Taubman: "[…] parental preferences exhibit `equal concern´ for all children if the welfare function is symmetric in its

[3] See Becker [1981] for further discussions.
[4] See for example Behrman and Rosenzweig [2004].

treatment of all children so that the corresponding indifference map is symmetric around the 45º line"[5]. In a situation when the thresholds are identical for all children so that parents are willing to trade off HCD income among their children above the common threshold, parents exhibit "unequal concern" favoring boys when weights are higher for boys rather than girls, so they value HCD income above that threshold more for boys than for girls. This situation makes the parental welfare function move towards the side of the favored gender away from the 45º line.

IV. Identification Strategy

Consider the following equations describing the human capital formation of two opposite-sex children within a family:

$$y_{mi} = \alpha + \beta s_{mi} + \lambda_{mi} + \delta_i + \varepsilon_{mi} \qquad (1a)$$

$$y_{fi} = \alpha + \beta s_{fi} + \lambda_{fi} + \delta_i + \varepsilon_{fi}, \qquad (1b)$$

where *m* and *f* indexes for the male and female child, respectively, of the *i*-th family. In both equations, the dependent variable y is a measure of child investment or educational outcomes, s_{mi} represents the inputs allocated for the male child, s_{fi} represents the inputs allocated for the female child, λ_{mi} and λ_{fi} accounts for the child specific endowments respectively, δ_i is a family fixed effect term and finally ε_{mi} and ε_{fi} are independent and identically distributed error terms with zero mean accounting for all other omitted factors.

Since I am interested in proving whether there exists or not differences in the inputs allocated for each child with regard to its gender, then, assuming that the correlation between s_{mi}, s_{fi} and their corresponding error terms is zero and, further, that the distribution of endowments across siblings follows a random generating process, a simple strategy of differentiating equation (1b) from equation (1a) will led to unbiased estimates of the parameter β[6]. For instance, the differentiated equation can be written as follows:

$$\Delta y_i = \beta \Delta s_i + \eta_i \qquad (2)$$

Equation (2) suggest that the unit of analysis is now the family *i* rather than the boy or girl belonging to that family. The error term η_i contains not only the differences of the error terms of equation (1a) and (1b) but also the differences of each child's endowments which are supposed to be randomly generated. If the inputs allocated for boys are equal than those allocated for girls within the same family, then it would not be reason to observe any differences across human capital outputs when comparing by gender. However, this strategy requires that (i) there exist families with both boys and girls and (ii) there are not systemic differences across siblings unless the gender variation, so the problem becomes in finding the counterfactual of a given children with respect to its opposite sex.

One possible way to overcome this problem is by matching each child with another one which shares exactly the same set of observed characteristics. Ñopo [2008] introduces this technique in order to explain gender differences in wages. As the author points out, this methodology does not require the usage of any parametric estimation, but split the observed wage gap in four different additive elements: differences in male and female supports, explained differences in the returns to each gender characteristics and lastly, unexplained differences in the wage gap.

Although I use the same methodology of comparison described above, I focus my analysis only in the common support differences in human capital investment between boys and girls. In this way, I group children according to a set of variables including geographical area of residence, age, household characteristics and parental background, dropping out those children which group includes only one of both sexes. Since there is no reason to think why boys can differ from girls after the matching process has been made, mean tests will be applied in the common support sample in order to prove the existence of unequal concern shown by parents with regard to the gender differences across siblings. Additionally, I use bootstrapping so that standard errors are corrected for the potential presence of possible discontinuities in the distribution of the common support sample.

Even though the methodology described seems to solve the "counterfactual problem", I remain skeptic about how accurate the resulting parameter can be under the matching methodology. I have two reasons to consider why there might be bias in β. First, the matching methodology relies in measurable characteristics provided in the survey. Yet this methodology ignores omitted factors that can be covariating with the error term, hence, biasing the estimated parameter. Second,

[5] Behrman, Pollak and Taubman [1982], pp. 55-56.

[6] Equations (1a) and (1b) do not consider the possibility of self investments in human capital formation. Since I am focusing in parental investments during childhood, self investments appear not to be relevant in this particular analysis.

as described in equations (1a) and (1b), heterogeneity among the families and children endowments are particularly important in the estimation of β. This condition implies that each child has to share exactly the same endowed skills (e.g.: abilities and other genetic factors) which can only be ensure when comparing relatives in a particular family.

As discussed in Rosenzweig and Wolpin [1980] and Rosenzweig and Wolpin [2000], monozygotic twins can be used to eliminate the contamination introduced by genetic endowments. In contrast, this is not the case of dizygotic twins, who share about fifty percent of their genes. Differences in dizygotic twins can be directly observed when considering the gender of the children, for example[7]. In spite of the comparisons discussed above, I use dizygotic twins as counterfactuals firstly because of the gender variation needed in the identification strategy and secondly because siblings can approximate in a better way both the familial endowments and the "natural" skills inherited from parents. Mean tests differentiating by gender will be applied in the sample obtained from twins in order to estimate the parameter of interest β.

At this point, it is necessary to delimit the interpretation of the resulting parameter when using twins as counterfactuals. First of all, for the reason that twins are not a usual event during pregnancy, the estimated parameter is, strictly, what the *jargon of econometrics* calls a "local average treatment effect" (LATE). Second, due to the particular scenario in which the comparisons are going to be done, special considerations about the heterogeneity involved in this neighbour effect have to be taken into account when interpreting the parameter. Third, inferences must be restricted only to the special case of twins since the evaluation is delimited to this particular neighbour effect.

To conclude this section, three additional assumptions are needed for a correct identification of parental preferences towards children sex. First, parents face identical raising prices for all children. Second, I assume that parents are interested only in equal distribution of the expected HCD income that children produce themselves, so additional wealth generated through the marriage market is irrelevant in the parental resource allocation decisions. Third, it is required that parents do not consider the gender wage differentials in the labor market as an argument when predicting the HCD income of their children.

Having discussed the limitations, considerations and power of inference of both identification strategies, I turn my attention in describing the data and commenting the results.

V. Data

For the purpose of identification, data is taken from the Peruvian 1992, 1996, 2000, 2004-2008 and 2009 Demographic and Health Surveys (DHS). The sample of interest is composed of children who were at home by the time when the interview was taken and whose mother were alive and responded to the women individual questionnaire. The latter restriction allows me for identifying whether the woman had or not twins at any parity. Since the available scholar information is more extensive for children in elementary and secondary school transition, I restrict the sample for those children reporting ages between 6 and 17 years, which results in 181145 children.

Table 1 summarizes the composition of the common support sample used for the matching methodology. Here after I will refer to this sample as the "matched sample".

Exactly the fifty percent of the matched sample are boys. In this sample, only 6.80% of the children report not currently attending to school, so it is feasible to expect no difference when comparing school attendance status of boys and girls. For this reason, two more outcomes will be considered. First, I construct an indicator of normative years of schooling taking into account the age of the child and the maximum year of school attained. Because children usually begin school at the age of 6 or 7 and school transition lasts 11 years, then they are supposed to finish the elementary at the age of 11 or 12 and the secondary at the age of 16 or 17. In this sense, I am able to identify how many years the child has been "delayed" in school, being zero the maximum value this variable adopts, which corresponds to "no delay". Notice, however, that this variable does not allow me to identify whether the child is delayed in the school or has left school.

The second variable is a dummy variable that indicates whether the child has been "held back" depending on the results of the score of the normative years of schooling. This variable takes the value of 0 when the normative years of schooling score is zero and takes the value of 1 in any other case. With these two additional outcomes, I am able to discuss both the

[7] There exists the possibility of observing gender differences in monozygotic twins. However, this phenomenon seems to be very unlikely. Having a negligible likelihood of observing gender differences in monozygotic twins and no further information in the survey, I consider only dizygotic twins when referring to the use of twins as counterfactuals hereafter.

extensive and the intensive margins of schooling "drop outs" and delays when comparing boys and girls.

TABLE 1
DESCRIPTIVE STATISTICS: VARIABLES USED FOR THE MATCHING METHODOLOGY

Variables	Mean	Standard Deviation
Child characteristics		
Age	9.82	2.68
Birth order	2.55	1.68
Parental background		
Mother's age	34.88	5.68
Mother's schooling	8.10	4.34
Mother works	64.86	-
Household characteristics [a]		
Rural area	38.99	-
Number of household members	5.84	1.75
Number of siblings (including child)	3.70	1.87
Sex of household head: male	25.09	-
Wealth index: quintile 1	19.19	-
Wealth index: quintile 2	18.82	-
Wealth index: quintile 3	24.11	-
Wealth index: quintile 4	15.87	-
Wealth index: quintile 5	22.02	-
Survey		
1992	18.45	-
1996	15.99	-
2000	15.74	-
2004-2008	32.96	-
2009	16.86	-
Total children		813

a/ Includes geographical region of residence.

For the reason that the matching methodology requires that each child has exactly the same characteristics described by the included variables, the resulting matched sample is about 0.04% of the original pooled sample, which corresponds to 813 observations. Table 2 summarizes the matched sample distribution by gender.

TABLE 2
SAMPLE DISTRIBUTION BY GENDER (MATCHING)

Survey	Boys (%)	Girls (%)	Total
1992	16.72	17.00	16.86
1996	33.17	32.75	32.96
2000	15.72	15.76	15.74
2004-2008	15.72	16.26	15.99
2009	18.67	18.23	18.45
Total	100.00	100.00	100.00

On the other hand, the pooled database describes that almost 2.32% of Peruvian women between 15-49 years old who have ever had a child, had multiple births (monozygotic and dizygotic twins) and 0.76% of the same population had male-female dizygotic twins. When restricting the sample to male-female twins between 6 to 17 years old, this results in 826 observations with exactly the same number of boys and girls. This does not necessary mean that the sample is composed only by two-child twins, but also by triplets so that the sex balance is achieved. Table 3 describes the distribution of the "twin sample" by gender.

TABLE 3
SAMPLE DISTRIBUTION BY GENDER (TWINS)

Survey	Boys (%)	Girls (%)	Total
1992	20.84	20.1	20.47
1996	19.85	20.82	20.34
2000	23.24	23.24	23.24
2004-2008	20.82	21.07	20.94
2009	15.25	14.77	15.01
Total	100.00	100.00	100.00

As one can notice, one potential drawback of this analysis is that it relies in a small sample. Because the "local effects" are robust in large samples, the resultant parameters should be treated with caution. Despite the sample size, further limitations of the estimated "local" parameters will be discussed when commenting the results.

VI. Results

A. Matching

The third column of the table 4 shows the resultant parameters when using matching for identification. The matching results show no difference between boys and girls in any of the three educational outcomes used.

Even when comparing in urban and rural subsamples, there exists no differences across boys and girls. This is not surprising when considering the school attendance, because of the high rates of enrollment and school attendance especially in the primary education. Overall, almost 94% of the population aged 16 to 17 attends to school and this rate is not very different across urban and rural population.

When considering the "held back" indicator, it is possible to see that around 43% of the children in the matched sample are delayed in school. This ratio is slightly lower for girls than for boys, however, when considering the normative years of schooling score, it is possible to see that the mean of delay is higher (in absolute terms) for girls rather than boys. In the reminder, I do not find a significant difference in either of the two outcomes, so it is not possible to argue in favor of the existence of unequal concern shown by parents with respect to the gender of their children when using the matching methodology.

TABLE 4
MATCHING RESULTS

	Boys	Girls	Difference
School attendance			
Total	0.928	0.936	-0.008
	[0.013]	[0.012]	[0.035]
Urban	0.943	0.956	-0.013
	[0.015]	[0.013]	[0.013]
Rural	0.904	0.905	-0.001
	[0.024]	[0.023]	[0.043]
"Held back" indicator			
Total	0.435	0.402	0.032
	[0.025]	[0.024]	[0.073]
Urban	0.36	0.331	0.03
	[0.031]	[0.030]	[0.067]
Rural	0.551	0.516	0.035
	[0.040]	[0.040]	[0.070]
Normative years of schooling			
Total	-1.215	-1.301	0.086
	[0.109]	[0.122]	[0.326]
Urban	-0.976	-1.065	0.089
	[0.132]	[0.149]	[0.322]
Rural	-1.589	-1.675	0.087
	[0.185]	[0.207]	[0.335]

Standard errors in brackets. Bootstrapped standard errors.
Total observations: 813
Urban observations: 495
Rural observations: 318

Before commenting the twin results, there are some other considerations I would like to discuss about the matching methodology. First, even when the resultant parameters obtained from the common support sample are not very different from zero, this does not mean that parents show no preferences for the gender of their children. One argument in favor of this insight is that I am not able to observe what is happening out of the common support sample, so it might be that girls or boys are grouped in the left side tail of the distribution when plotting any of these three educational outcomes by gender. If so, it would be hard to think about a reasonable explanation why this could be happening. In contrast to the labor market where there are "traditional occupations" for men and women (i.e.: housekeepers, pre-primary teachers, builders, etc.), the educational market is not restrictive with regard to the gender, so the only reasonable explanation is that families favors one of both sexes.

The second consideration is the span of variation within the common support sample. This idea is directly related with the size of the sample. Although the matching methodology implies a gain in accuracy of the parameters of interest, it also implies a reduction of the sample size and is decreasing with the number of variables considered for the matching. In this way the variance could be affected by the reduction of observations at the same time that the degrees of freedom are reduced. In this situation, there exists the possibility of falling in type 2 error.

Finally, as mentioned above, the resultant parameters depend on the variables considered for the matching. This could be a problem because I want to identify shifts in parental preferences when changing the gender of children. The matching methodology is able to equate, at most, resource constraints across different families because wealth can be approximated by using observed variables. Nonetheless, parental preferences refer exclusively to "tastes", which are difficult to be computed. Last, the matching methodology compares different families, so there may be a combination of different "tastes" across compared families that could result in confounding factors when estimating the parameters.

B. Twins

The last column of table 5 shows the estimated parameters of the differences between inputs allocated for male-female twins. Surprisingly, notice that the means of the school attendance and the "held back" indicator show that twins are doing relatively worst than the matched sample. Yet, this is not the case when considering the normative years of schooling score.

One possible explanation for the former is that multiple births narrow down the resources allocated for each child. Because multiple births shift the family size upwards, parents must distribute household resources among a larger number of children. Therefore, it is expected that multiple births have deleterious effects on educational outcomes of the twins themselves and the rest of their siblings.

When comparing male-female twins, the difference between educational outcomes is not large enough to be considered significant for supporting the hypothesis of existence of unequal concern shown by parents.

First, it seems that the rate of attendance is higher for girls. When decomposing by urban and rural areas, it is possible to see that only in rural areas boys have higher rate of school attendance than girls. Still, school attendance outcomes, even when splitting the sample by urban and rural areas, are very close to zero. Mean tests cannot reject the null hypothesis of non-statistic significance of the differences.

TABLE 5
TWIN RESULTS

	Boys	Girls	Difference
School attendance			
Total	0.862	0.876	-0.014
	[0.017]	[0.016]	[0.012]
Urban	0.892	0.9	-0.008
	[0.020]	[0.020]	[0.029]
Rural	0.823	0.845	0.022
	[0.028]	[0.027]	[0.039]
"Held back" indicator			
Total	0.464	0.448	0.016
	[0.025]	[0.024]	[0.017]
Urban	0.394	0.381	0.013
	[0.032]	[0.032]	[0.045]
Rural	0.552	0.533	0.02
	[0.037]	[0.037]	[0.052]
Normative years of schooling			
Total	-0.684	-0.688	0.003
	[0.087]	[0.086]	[0.061]
Urban	-0.58	-0.602	0.022
	[0.098]	[0.105]	[0.143]
Rural	-0.818	-0.797	-0.021
	[0.153]	[0.142]	[0.209]

Standard errors in brackets.
Total observations: 826
Urban observations: 460
Rural observations: 363

In the case of the "held back" indicator, results show that there is less proportion of girls delayed in school than boys. This difference tends to be higher in rural areas. However it is not large enough to be statistically different from zero. Lastly, like the previous results, there is no difference between boys and girls when considering the normative years of schooling score.

In spite of the findings, it is appropriate to consider the heterogeneity of the population used for the identification. When comparing male-female twins, one of the implicit assumptions is that parents have not changed their behavior after the multiple births had taken place. Instead, parents do change their behavior after conceiving a multiple birth. Arguments in favor of this position can be supported by the detrimental effects in school attendance and the delay indicator found when comparing the twin versus the matched sample. Besides, having male-female twins can reinforce changes in the parental behavior by forcing parents to distribute resources equally among their twin siblings. One possible explanation is that parents with unequal concern would never want to show their preferences because of the bad reputation this would imply. Then, even when parents have preferences for children's sex, they would never exhibit those preferences with twins because of the notorious differences in inputs allocated for each child when referring to identical siblings.

To sum up, there exist no difference between boys and girls when using the twin sample. However, using male-female twins can complicate the identification of parental gender preferences among children if there is a selection bias in the sense that parents change their behavior after conceiving a multiple birth.

VII. Summary and Conclusions

This paper attempts to provide evidence which supports the existence of unequal concern with regard to the gender shown by parents when investing in human capital formation of children in Peru.

In this context, the "treatment" has been defined as being a girl. Because investments in progeny are made primarily in the elementary and secondary school transition of the children, the span of analysis has been limited to children aged 6 to 17. However, the major problem is to identify the counterfactuals of the treated population. Because individuals have different characteristics such as genetic endowments and parental background that can enhance the educational outcomes of the children, the identification strategy relies in two different non-parametrical approaches: one-to-one matching and the male-female twins experiment.

The matching methodology allows balancing household observable characteristics such as family size, financial boundaries, neighborhood effects and parents' schooling. Nonetheless, it fails in balance child's specific endowments such ability and implicitly assumes that comparable families have the same preferences with respect to the resource allocation and the sex of their progeny. An alternative strategy is to use male-female dizygotic twins as counterfactuals within the families. This strategy allows removing family fixed effects, so it makes the identification more accurate. However, since multiple births are not common in families (around 2.30% of women aged 15-49 have had multiple births) and male-female twins appears with less frequency, the size of the sample can affect the results.

In spite of these considerations, I have not found statistical significant effects when comparing boys and girls in educational outcomes. Specifically, the comparisons rely in three particular outcomes: school attendance, a "held back" indicator which measures school progression and the normative years of schooling score, which measures the degree of delay in

school. Even when considering urban and rural population alone, the differences seem to be very close to zero. Despite the use of male-female twins can increase the precision of the estimated parameters, it occurs that multiple births can shift parental preference towards hiding gender preferences among their progeny.

As a result of the discussion about the followed methodology and its consequences in bias in the parameters of interest, I believe that parametric approaches can give additional insights about parental preferences for boys or girls. Furthermore, it is necessary to look upon other outcomes such as health, malnutrition, antenatal care and even gender selection of the unborn.

Finally, the economics of discrimination literature should expands its field of research and consider intrafamilial resource allocation mechanics, especially in developing countries like Latin American countries.

REFERENCES

Atal, Juan Pablo, Hugo Ñopo and Natalia Winder, "New Century, Old Disparities: Gender and Ethnic Wage Gaps in Latin America", *IDB Working Paper Series No. 109*, IDB, 2010.

Becker, Gary, *A Treatise on the Family*, Cambridge, Mass.: Harvard University Press, 1981.

Behrman, Jere and Mark Rosenzweig, "Returns to Birthweight," *Review of Economics and Statistics,* 86 (2004), 586-601.

Behrman, Jere, Robert Pollak and Paul Taubman (BPT), "Parental Preferences and Provision for Progeny," *Journal of Political Economy*, 90 (February, 1982), 52-73.

Behrman, Jere, Robert Pollak and Paul Taubman, "Do Parents Favor Boys?," *International Economic Review*, 27 (February, 1986), 33-54.

Ben-Porath, Yoram and Finis Welch, "Do sex Preferences Really Matter?," *Quarterly Journal of Economics*, 90 (May 1976), 285-307.

Birdsall, Nancy and Richard Sabot, *Labor Market Discrimination in Developing Economies*, Washington: World Bank, 1985.

Dahl, Gordon, and Enrico Moretti, "The Demand for Sons: Evidence from Divorce, Fertility, and Shotgun Marriage," *National Bureau of Economic Research* Working paper No. 10281, 2004.

Gertler, Paul and Glewwe, Paul, "The Willingness to Pay for Education for Daughters in Contrast to Sons: Evidence from Rural Peru," *The World Bank Economic Review*, 6 (January 1992), 171-188.

Grant, Monica and Jere Behrman, "Gender Gaps in Educational Attainment in Less Developed Countries," *Population and Development Review*, 2010.

Lundberg, Shelly, "Sons, Daughters, and Parental Behavior," *Oxford Review of Economic Policy*, 21 (September, 2005), 340-356.

Ñopo, Hugo, "Matching as a Tool to Decompose Wage Gaps," *Review of Economics and Statistics*, 90 (2008).

Oaxaca, Ronald, "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review*, 14 (October, 1973), 693-709.

Rosenzweig, Mark and Kenneth Wolpin, "Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment," *Econometrica*, 48 (1980), 227-240.

Rosenzweig, Mark and Kenneth Wolpin, "Natural 'Natural Experiments' in Economics," *Journal of Economic Literature*, 38 (2000), 827-874.

Rosenzweig, Mark and T. Paul Schultz, "Market Opportunities, Genetic Endowments, and Intrafamily Resource Distribution: Child Survival in Rural India," *American Econmic Review*, 72 (September, 1982), 803-815.