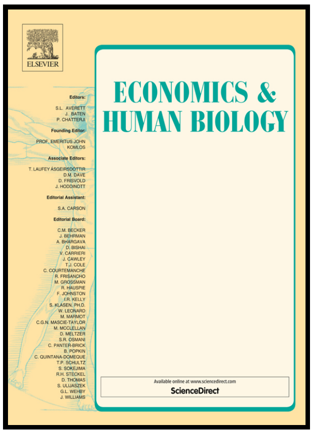
Journal Pre-proof



Parental gender preferences over three centuries: Evidence from Argentina

Fernando Antonio Ignacio González

PII: S1570-677X(23)00101-6
DOI: https://doi.org/10.1016/j.ehb.2023.101320
Reference: EHB101320

To appear in: *Economics and Human Biology*

Received date: 23 May 2023
Revised date: 4 October 2023
Accepted date: 18 November 2023

Please cite this article as: Fernando Antonio Ignacio González, Parental gender preferences over three centuries: Evidence from Argentina, *Economics and Human Biology,* (2023) doi:https://doi.org/10.1016/j.ehb.2023.101320

Parental gender preferences over three centuries: Evidence from Argentina

Fernando Antonio Ignacio González

Universidad Nacional de Misiones, Posadas, Argentina

fernando.gonzalez@fce.unam.edu.ar

Abstract

In this paper I examine the evolution of parental gender preferences in Argentina (i.e., parents who prefer a certain gender composition in their children). To do this, I use census microdata that spans the 19th, 20th, and 21st centuries. The estimation strategy exploits the plausibly random assignment in the gender of children.

The results show a persistent preference for a mixed gender composition (i.e., having at least one boy and one girl) instead of children of the same gender. This translates into an increase in the probability of having a third child, conditional on already having two children of between 9%-23% for those couples who have children of the same gender -in relation to couples with children of opposite genders-. These preferences are heterogeneous over time and have important implications in terms of fertility (i.e., the reduction of these mixed gender preferences -in favor of greater gender-neutrality- could contribute to reducing the number of children per couple). In addition, the findings of this work support the empirical literature that uses the gender composition of the first two children as an instrumental variable to study the impact of fertility on labor participation.

Keywords: parental gender preferences, fertility, population census, Argentina

1. Introduction

Since the pioneering works of Iacovou (2001) and Angrist & Evans (1998) we know that mixed gender parental preferences are frequent (i.e., parents who prefer children of opposite genders -at least one boy and one girl- instead of children of the same gender). The existence of these preferences has implications for fertility and labor supply. As shown by Iacovou and Angrist & Evans, parents with two children of the same gender are more likely to have a third child -and thus try to achieve the desired gender composition- in relation to parents with two children of opposite genders. This results in a greater number of children throughout life (Hammoudeh, 2017) and, ultimately, in less labor force participation. In this paper I provide novel evidence on the evolution of these preferences for a period that extends over three centuries (XIX, XX and XXI).

The above implies that the gender composition of the first two children can be used as an instrumental variable to estimate the causal impact of fertility on labor participation (Iacovou, 2001; Angrist & Evans, 1998). Thus, parental gender preferences are used as a source of exogenous variability in household size. In econometric terms, this translates into the incorporation of a dummy variable that takes the value 1 for those couples with two first children of the same gender. This estimation strategy was later extended by Cruces & Galiani (2007) and Tortarolo (2013) when examining Latin American countries. The authors coincide in reporting an increase of between 3 and 5 percentage points in the probability of having a third child in those parents with two children of the same gender -in relation to those with two children of opposite genders-. For the gender composition of children to be a valid instrument, it must be exogenous (i.e. parents must not be able to influence the gender of their children) and, indeed, the existence of stable mixed gender parental preferences must be verified. In this work, I provide evidence in favor of the validity of this instrument.

The existence of mixed gender parental preferences may respond to multiple reasons (Trivers & Willard, 1973; Rozenweig & Wolpin, 2000; Gass *et al*., 2006; Li, 2021; Gabay-Egozi *et al*., 2022; Goli *et al*., 2022) as: genetic diversity (some studies suggest that parents may prefer having children of both sexes to maximize genetic diversity in their offspring and therefore achieve a more robust gene pool and have better chances of survival for the family line); sibling dynamics (parents may define preferences based on their perception of how sibling dynamics work -e.g., some parents believe that children of different genders may get along better or have less rivalry); economic factors (parents may believe that children of different genders have diverse needs and, therefore, require a broader range of resources and support); educational and occupational aspirations (parents might have specific educational or occupational aspirations for their children based on gender stereotypes. They may believe that children of different genders are better suited for particular careers or roles).

Although the existence of mixed gender parental preferences and their impact on labor market results is known, there is little evidence about the evolution of these preferences. Usually, studies on the subject examine a few decades. In this regard, it is critical to know its evolution over an extended period to identify reversals or changes. An exception to the above comes from the work of Jones, Millington & Price (2023). These authors analyze the evolution of parental gender preferences for the case of the United States from microdata that extend between 1850-2019. Their findings show that preferences for a mixed gender composition in children have intensified since the second half of the 20th century. Extending the analysis to the developing world is then relevant.

In this context, in this paper I analyze the impact that parental gender preferences have on fertility in the case of Argentina. To do this, I use household microdata from multiple census waves that span the period 1895-2010. This allows me to analyze the evolution of these preferences over an extended period. Argentina -third largest economy in Latin America- is a relevant case study. This country experienced rapid population growth during the 19th century as a result of its policies to attract immigrants -especially Europeans-. The migratory flow was substantially reduced towards the second half of the 20th century. Argentina is a country that has also experienced wide economic and welfare fluctuations in recent decades (González et al. 2018; 2021; 2022), including a sustained reduction in the infant mortality rate and the fertility rate (Figure A.1). This makes it possible to examine changes in parental gender preferences in the face of different compositions of the population (natives *vs*. immigrants) and, therefore, analyze whether these preferences respond to an essentially local or imported phenomenon.

To the best of my knowledge, this work adds value to the literature on gender preferences and fertility in three respects. First, this is the first work to analyze this topic for Argentina. Typically, the literature has focused either on developed countries or on Asian countries. Second, this work differs from most previous studies by examining a period that spans three centuries (19th, 20th, and 21st). Usually, previous literature has concentrated on examining periods of a few decades. An exception to this is the work of Jones *et al*. (2023). Third, this work incorporates a new check when comparing between natives *vs.* immigrants. This allows us to inquire about whether parental gender preferences are a local or an imported phenomenon.

This work is inserted within the literature that examines parental gender preferences (Dahl & Moretti, 2008; Maurin & Moschion, 2009; Mu & Zhang, 2011; Angelov & Karimi, 2012; González, 2018; Goli *et al*., 2022; Tan *et al*., 2023; Jones *et al*., 2023; Huang *et al*., 2024). This work dialogues with the literature that examines parental gender preferences over long periods (Jones *et al*., 2023) and especially for developing countries (Cruces & Galiani, 2007; Tortarolo, 2013). It is important to highlight that although this work evaluates the existence and stability of mixed gender parental preferences, these are not the only type of preference that may exist. Abundant evidence has shown that in certain contexts parents may have strong preferences for one gender -i.e., only boys or only girls- (Dahl & Moretti, 2008; Kolk & Schnettler, 2012; Huang *et al*., 2024).

Hereinafter, section 2 describes the information sources, while section 3 presents the estimation strategy. Section 4 describes the main results of the work and, finally, section 5 details the conclusions.

2. Sources of information

In this paper I use multiple census microdata waves as sources of information. This includes microdata for years 1895, 1970, 1980, 1991, 2001, and 2010. The microdata of the waves between 1970-2010 come from the Integrated Public Use Microdata Series platform ([IPUMS], 2020). The microdata of wave 1895 come from the records digitized by the Gino Germani Research Institute ([IIGG], 2015). This source also includes microdata from wave 1869. This wave is excluded from the analysis since it does not include a variable that allows identifying people within the same household (household identifier).

From these sources of information, it is possible to identify the gender and age of each person, the relationship with the head of the household (partner, child, etc.), their marital status (single, married, cohabiting, divorced or widowed) and years of education, among other variables. The 1895 and 2001 censuses do not allow the number of years of education of each person to be identified -although they do include a literacy variable (1 if they know how to read and write, 0 otherwise). Furthermore, the 1895 census does not include a variable with kinship relationships within each household. This is estimated based on the self-reported gender, age and number of children. The 2010 census does not include marital status and therefore here the sample cannot be restricted to married or cohabiting persons.

Based on these sources of information, I construct a sample of families, following a similar criterion to that used by Jones *et al*. (2023), for which the mother has stopped having children, no child has left the parental household, and it is possible to link the children with her parents. To guarantee the foregoing, I restrict the sample by retaining the households in which the mother is the head of the household or the partner of the household's head, the mother is married or cohabiting, the eldest child is 17 years of age or younger (to minimize the chances of including households with children who have already left the parental household), the youngest child is 5 years of age or older (to minimize the chances of including households that could have more children), and the mother is 45 years of age or younger (an age in which the eldest child is probably still living with the parents). Other authors have incorporated more conservative sample cuts: Dahl & Moretti (2008) restrict the maximum age of the child to 12 years. Unlike the microdata from Argentina, those used by Dahl & Moretti do not allow us to know exactly the number of children that each mother and father has had, and this is an advantage of the present work in terms of security in identifying the number of children. As shown below, the results of this work do not change when these criteria vary.

Taken together, these restrictions seek to ensure that the number of children identified in each household matches the total number of children a mother has had. The number of children identified in each household according to the previous procedure and the self-reported number of children (for each mother and father) present discrepancies in 0.07% of the households, on average, considering those census waves in which the kinship and the number of children variables are simultaneously available (1970, 1980, 1991 and 2001). This is an advantage of the Argentine census microdata -unlike other countries- and it gives me greater confidence in the number of children identified in each household.

Table 1, below, presents basic descriptive statistics of all the waves considered. The table reflects several stylized facts of the last decades in Argentina -and in more general terms, in developing countries-: the fertility rate has dropped substantially, the population has aged, and schooling levels have increased. This coincides with a decrease in the proportion of foreigners in the country.

Table 1: Descriptive statistics for households in Argentina (1895-2010)

	1895	1970	1980	1991	2001	2010
Number of children	2.42	2.05	2.21	2.32	2.25	2
Age of mother	34.39	36.54	36.11	36.3	36.61	35.53
% of inmigrants	30.75	9.27	8.07	3.91	0.52	5.52
Years of education[a]		7.5	6.82	9.26		17.21
Two boys[b]	27.27	26.3	25.87	25.97	26.47	25.85
Two girls[b]	24.42	23.26	23.83	23.74	24.05	23.94
A boy and a girl[b]	48.31	50.44	50.3	50.28	49.48	49.64

Source: own elaboration based on IIGG and IPUMS. Note: [a] The number of years of education is not available in the 1895 and 2001 censuses. [b] Two boys/girls (households where the first two children are boys/girls) is a proportion estimated among households with at least two children.

3. Methodology

The estimation strategy is based on a fixed effect model that allows me to control for unobserved heterogeneity -equation 1-. This equation allows me to estimate the chances of having a third child, conditional on the first two children being boys or the first two being girls -the omitted category here is two children of opposite genders-.

$$Have\ Third_i = \beta_0 + \beta_1 Two\ boys_i + \beta_2 Two\ girls_i + X_i + \mu_i \qquad (1)$$

where $Have\ Third_i$ is a dummy variable that takes the value 1 if mother *i* has a third child and 0 otherwise. $Two\ boys$ and $Two\ girls$ are dummy variables that take value 1 if mother *i* has had two first male children or two first female children, respectively. X_i is a vector of covariates that includes age, years of education, and fixed effects by district of residence. These variables attempt to control for other determinants that could affect the probability of having a third child. μ_i is the error term. The coefficient β_1 reflects the increase in the chances (in percentage points) of having a third child given that the first two children are boys -in relation to a mother with two first children of the opposite genders-. An analogous interpretation (but in the case of the first two female children) corresponds to the coefficient β_2. Equation 1 is estimated considering mothers with at least two children.

Equation 1 is re-estimated by implementing multiple robustness checks. First, I show that the results are robust to the exclusion of controls and fixed effects. Second, instead of estimating for each year separately, I consider a pooled model that includes time fixed effects (Table A.3). Third, I disaggregate the estimates between native Argentines and immigrants to identify possible heterogeneous effects on preferences based on country of origin. Fourth, I allow variations in the inclusion criteria in the sample (different maximum ages for the mother and children-). Fifth, I consider the total number of children as the dependent variable -instead of a dummy variable that identifies the third child. This robustness check (the number of children as the dependent variable) is informative about the possible impact of mixed gender parental preferences on fertility. However, it has limitations. It is not possible to distinguish from other simultaneous preferences (in the minimum number of children, family size, among others). In this regard, Table A.1 shows that there are no indications of differences regarding these preferences.

The identification strategy of this work assumes that the gender of the children is randomly assigned. That is, parents cannot influence the gender of their children. In this regard, Table A.1 presents descriptive statistics for the three types of mothers (those with the first two male children, the first two female children, and the first two children of opposite genders). From this the characteristics of all groups of mothers are similar to each other with differences in the means of less than 1%. This provides evidence in favor of randomness in the gender assignment of children.

It is important to highlight that although this work evaluates the existence and stability of mixed gender parental preferences, these are not the only type of preference that may exist. Abundant evidence has shown that in certain contexts parents may have strong preferences for one gender - i.e., only boys or only girls- and these preferences could vary between subgroups (Dahl & Moretti, 2008; Kolk & Schnettler, 2012). The existence of different types of gender preferences is empirically relevant since they can give rise to different types of stopping rules (Blanchard & Lippa, 2007; Blanchard, 2022; Kabátek *et al*., 2022; Baland *et al*., 2023).

The stopping rule refers to a behaviour by which parents continue childbearing till they reach a specific number of children of a given gender. The presence of stopping rules can lead to instrumental births (i.e., having children until the desired gender composition is reached) and selective abortions (i.e., interruptions of pregnancies of the undesired gender). In the case of preferences for sons, the existence of instrumental births predicts that girls will have a greater number of younger siblings than boys. In the case of selective abortions, boys will have a greater number of older siblings than girls. The opposite is true in case of preferences for girls. In both cases, it is expected that sex ratios that emerge are substantially far from their natural level (105-106). In this regard, Figure A.2 shows that the sex ratio for children between 0-1 years old in Argentina (104) has been similar to its natural level (105-106) in the analyzed period. This is an indication that, for Argentina, there is no single gender preference (i.e., preference for a son or a daughter).

4. Results

Table 2 presents the results that arise from estimating equation 1 for all the waves of microdata available for Argentina (1895-2010). The results confirm that there are gender preferences. That is, parents are not indifferent to the gender of their children. This arises from the fact that the coefficients of $Two\ boys$ and $Two\ girls$ are significantly different from zero in most of the years analyzed. That is, the probability of having a third child (conditional on already having two) differs between couples according to the gender of the first two children.

Between 1970-2010, as shown in Table 2, parents show a clear preference for having at least one child of each gender. Indeed, the probability of having a third child increases between 2 and 5 percentage points when the first two children are of the same gender -in relation to a couple that has two children of opposite genders-. This represents an increase of between 9% and 23% in relation to the mean of the dependent variable for the omitted category.

The exception to the above is observed in 1895. As shown in Table 2, for that year the coefficients of interest are not significantly different from zero. That is, the conditional probability of having a third child is independent of the gender of the first two children.

Another interesting result that emerges from Table 2 is the difference between the estimated coefficients for $Two\ boys$ and $Two\ girls$. The equality of means test shows that there are no significant differences between 1895-1970 and between 2001-2010. Only between 1980-1991 do significant differences appear. In these decades there seems to be a slight preference for sons -the coefficient for *Two girls* is greater in absolute value than that of *Two boys*-: the probability of having a third child increases more when the first two are girls. All the above shows the evolution of parental gender preferences, highlighting the importance of considering series long enough to capture these changes.

Table 2: Effect of the gender composition of the children on the likelihood of having a third child

Dep: dummy for having third child	1895	1970	1980	1991	2001	2010
Two boys	-.0852525	.0514741***	.0422174***	.0273542***	.0252959***	.0220977***
	(.0827041)	(.0103517)	(.0054584)	(.0034365)	(.0040995)	(.003203)
Two girls	-.0718697	.050845***	.0546997***	.0360855***	.027297***	.0274724***
	(.0669143)	(.0109017)	(.0061355)	(.0037132)	(.0040888)	(.0032609)
Test for Two boys=Two girls (p-value)	0.6333	0.6773	0.0616	0.0493	0.6914	0.1526
Omitted Y mean	.2705	.2198	.2407	.2883	.2759	.2438
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	664	11,173	53,084	117,284	68,505	98,417
R^2	0.4154	0.0447	0.0151	0.0083	0.0127	0.0098

Source: Own elaboration based on IIGG and IPUMS. Standard errors, in parentheses, are clustered at the district of residence level. * significant at 10%, ** significant at 5%, *** significant at 1%. Two boys/girls takes value 1 if the first two children are boys/girls and 0 otherwise.

The above results are robust to the exclusion of control variables and fixed effects. Table A.2 presents these results and shows that in this case the increases in the conditional probability of having a third child range between 9% and 24% for couples with two children of the same gender -in relation to couples with children of opposite genders.

Table 3, below, presents the results that arise from re-estimating equation 1 by considering the total number of children as dependent (instead of a dummy that identifies the third child). The results are robust to this specification and extend previous results: couples with two first children of the same gender have, on average, a greater number of children throughout their lives - compared to couples with two first children of opposite genders-. As in Table 2, the exception to the main result is given by the microdata from 1895.

Table 3: Effect of the gender composition of the children on the number of children

Dep: number of children	1895	1970	1980	1991	2001	2010
Two boys	.1009786	.0936599***	.0690459***	.0735177***	.0626305***	.0358637***
	(.2219381)	(.0203078)	(.0095712)	(.0076235)	(.0088943)	(.0069165)
Two girls	-.0726809	.110284***	.114116***	.099577***	.0800327***	.0555928***
	(.2111141)	(.0194267)	(.0116515)	(.0097145)	(.0089723)	(.0068741)
Omitted Y mean	.2705	.2198	.2407	.2883	.2759	.2438
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	664	11,173	53,084	117,284	68,505	98,417
R^2	0.4407	0.1253	0.0660	0.0498	0.0792	0.0378

Source: Own elaboration based on IIGG and IPUMS. Standard errors, in parentheses, are clustered at the district of residence level. * significant at 10%, ** significant at 5%, *** significant at 1%. Two boys/girls takes value 1 if the first two children are boys/girls and 0 otherwise.

Table 4 shows the changes in the conditional probability of having a third child according to the origin of the mother (native Argentine *vs*. immigrant). From this the gender preferences of Argentine mothers have remained more stable over time. Indeed, between 1970-2010 it is observed that native Argentine mothers show strong gender preferences for having at least one boy and one girl. In the case of immigrant mothers, a reduction in these preferences is observed from 2001. From that year on, the estimated coefficients cease to be significantly different from zero. Additionally, I show that parental gender preferences are heterogeneous among immigrants (Table A.4). Thus, Asian immigrants present parental preferences for the same gender (instead of mixed). Furthermore, the difference in absolute value between both coefficients (Two boys and Two grils) is greater in the case of immigrants from Europe and North America. This indicates that a person's regional origin is important in terms of parental gender preferences. In other words, these preferences differ between natives and immigrants.

Table 4: Effect of the gender composition of the children on the likelihood of having a third child according to origin

	Dep: dummy for having third child	1895	1970	1980	1991	2001	2010
Native	Two boys	-.200525	.0580233***	.0420103***	.0286669***	.0252644***	.0219972***
		(.1516331)	(.0112978)	(.0058446)	(.0043196)	(.0041345)	(.003232)
	Two girls	-.1270147	.047941***	.0539044***	.0345541***	.0273699***	.0259464***
		(.1302176)	(.0115664)	(.0062467)	(.0048874)	(.0040916)	(.0034344)
Immigrant	Two boys	.062161	-.0166249	.0408624**	.0400118***	.0023327	.0235122
		(.1347384)	(.0254623)	(.0190142)	(.0147065)	(.1008398)	(.0171297)
	Two girls	-.0107566	.0801847**	.0608481**	.0426933**	.0703762	.0596131***
		(.1126351)	(.040259)	(.0268891)	(.0170935)	(.0957006)	(.0162641)
	Controls	Yes	Yes	Yes	Yes	Yes	Yes
	Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Source: own elaboration based on IIGG and IPUMS. Standard errors, in parentheses, are clustered at the district of residence level. * significant at 10%, ** significant at 5%, *** significant at 1%. Sample sizes and R^2 are omitted for simplicity and are available upon request from the author. Two separate regressions are estimated for each column. Two boys/girls takes value 1 if the first two children are boys/girls and 0 otherwise.

Table 5 shows that the results reported in tables 2 to 4 are robust by allowing for variations in the inclusion criteria in the sample. This includes the maximum age of the mother (which becomes more restrictive, from 45 to 40 years), the age of the eldest child in the household (reduces from 17 to 15 years) and the age of the youngest child (increases from 5 to 7 years). These changes, in all cases, further restrict the sample allowing for less inclusion error. From this it follows that the increase in the conditional probability of having a third child is not the result of the inclusion criteria used in this study.

Table 5: Effect of the gender composition of the children on the likelihood of having a third child by varying the inclusion criteria

		1895	1970	1980	1991	2001	2010
Age of mother<41	Two boys	-.0835904	.0512034***	.046757***	.0277647***	.0213469***	.0198833***
		(.096517)	(.0113654)	(.0061633)	(.0040544)	(.0045118)	(.0035278)
	Two girls	-.0854555	.0576411***	.0581658***	.032055***	.0202363***	.0253084***
		(.082047)	(.0123025)	(.0069098)	(.0043338)	(.0046107)	(.0038802)
Age of youngest child>6	Two boys	-.0505932	.0579563***	.0419608***	.030093***	.0253613***	.0258951***
		(.1398949)	(.0109353)	(.0064415)	(.0042411)	(.0053356)	(.0044416)
	Two girls	-.1148271	.0547685***	.0583885***	.0461483***	.0323814***	.0311452***

		(.1145517)	(.0118056)	(.0063456)	(.0041259)	(.0048611)	(.0039406)
Age of oldest child<16	Two boys	-.1208785	.0518841***	.0372471***	.0305977***	.0268978***	.0227598***
		(.1003996)	(.011638)	(.0060079)	(.0035812)	(.0042451)	(.0035581)
	Two girls	-.0736469	.049007***	.0557801***	.0397128***	.0296886***	.0276265***
		(.0884207)	(.0119812)	(.0063942)	(.0048203)	(.0046995)	(.0038344)
	Controls	Yes	Yes	Yes	Yes	Yes	Yes
	Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Source: own elaboration based on IIGG and IPUMS. Standard errors, in parentheses, are clustered at the district of residence level. * significant at 10%, ** significant at 5%, *** significant at 1%. Sample sizes and R^2 are omitted for simplicity and are available upon request from the author. Three separate regressions are estimated for each column. Two boys/girls takes value 1 if the first two children are boys/girls and 0 otherwise.

The results reported in this work are in line with previous evidence for other countries. Jones *et al*. (2023) coincide in reporting, for the United States, that the conditional probability of having a third child increases by 2 percentage points -in relation to couples with two first children of opposite genders-. This increase is substantially lower than that reported in this study (Table 2) -which was around 5 percentage points-. This is indicative of greater parental gender preferences in Argentina in relation to the United States. Consistent results are reported for Latin American countries -increase between 3 and 5 percentage points- (Cruces & Galiani, 2007; Tortarolo, 2013) and in France -increase of 4 percentage points- (Maurin & Moschion, 2009). For Sweden, Angelov & Karimi (2012) report a more modest increase of 0.14 percentage points -which denotes parents are more gender-neutral there-.

More generally, the paper's findings showed a substantial change in parental gender preferences between 1895 and 1970: at the beginning of the period there were no gender preferences, while at the end there was a clear preference for a mixed gender composition of children. It is interesting to discuss this point. First, the sample size of the 1895 microdata is small. This can be problematic in the sense that its low power may not identify any significant difference (although when it exists). Second, a possible explanation for this heterogeneity in preferences could be found in the work of Rozenweig & Wolpin (2000). Here the authors argue that in a low-income context, a greater parental gender neutrality could prevail, since having children of the same gender reduces child-rearing costs. That is, children of the same gender are more likely to share items such as clothing and footwear. These items can represent a substantial portion of low-income household spending.

Although I cannot test the above hypothesis -due to lack of income information in Argentine microdata-, this could be a reasonable explanation: during the 19th century, Argentine parents were gender-neutral given a context of lower income and, later, with the growth in income experienced during the first half of the twentieth century were able to opt for a mixed gender composition of their children.

Finally, it is important to consider the possible mechanism identified in this work: immigration. The results of this work show that the parental gender preferences of immigrants can differ substantially from those of native Argentines. Thus, the extensive migratory flows that the country experienced during the second half of the 19th century and the first half of the 20th century can help explain the heterogeneous evolution of these preferences in the Argentine case. This potential mechanism has been explored in greater detail for the case of China (Huang *et al*., 2024).

5. Conclusions

Throughout this work I have examined the existence of parental gender preferences in Argentina throughout a period that spans the 19th, 20th and 21st centuries. In particular, I estimated the impact that the gender composition of the first two children has on the probability of having a third child. The findings show that, by preferring to have at least one child of each gender (a boy and a girl), the probability of having a third child increases between 9%-23% in couples whose first two children are of the same gender (two boys or two girls) -in relation to couples with two first children of opposite genders-. These preferences have shown changes over time and according to the parents' country of origin (native Argentines *vs.* immigrants).

Taken together, the results reported here show the presence of clear parental gender preferences. That is, Argentine parents are not gender-neutral and prefer to have at least one child of each gender instead of several children of the same gender. This has important implications in terms of fertility and population growth. Indeed, since parents prefer children of both genders, this implies that they have a greater number of children until they reach the desired gender combination (this is consistent with what is reported in Table 3). In turn, if these gender preferences are reduced, there will be a drop in fertility. This is consistent with what is reported in tables 2 and 3 and with the substantial drop observed in Argentina in terms of fertility in the last decade.

On the other hand, the findings of this work support the empirical economics literature that examines the impact of fertility on female labor participation and resorts to an instrumental variable approach based on the gender composition of the children (Angrist & Evans, 1998; Cruces & Galiani, 2007; Angelov & Karimi, 2012). This literature is based on the existence of mixed gender parental preferences (preferences for at least one boy and one girl). Thus, in general terms, the results reported in this work validate the above. However, it is recommended to test the existence of these preferences in each case (something that was not verified for Argentina in 1895).

In the future, it appears relevant to be able to evaluate the changes that have occurred in terms of parental gender preferences, throughout the last decade, with the appearance of the feminist movement. Perhaps these preferences could mutate towards preferences for girls instead of a mixed composition. In addition, it is interesting to be able to evaluate differences in these preferences between age ranges and according to religion. Finally, it is valuable to be able to extend the analysis to changes in the use of time from these parental gender preferences.

Appendix

Table A.1: Descriptive statistics of mothers according to the gender of their first two children

	Two boys	Two girls	A boy and a girl
Age of mother	36.43	36.42	36.36
Years of education	11.42	11.32	11.33
Age of oldest child	12.98	12.94	12.92
Age of youngest child	8.16	8.08	8.24
Number of people at household	4.79	4.81	4.74

Source: own elaboration based on IIGG and IPUMS.

Table A.2: Effect of the gender composition of children on the likelihood of having a third child when excluding controls and fixed effects

Dep: dummy for having third child	1895	1970	1980	1991	2001	2010
Two boys	-.0551397	.0494982***	.0417004***	.0274048***	.0253107***	.0222675***
	(.0444324)	(.0099368)	(.0055131)	(.0033518)	(.0040737)	(.0031975)
Two girls	-.0335926	.0539821***	.0543965***	.0360913***	.0270282***	.0278089***
	(.0407129)	(.0104071)	(.0061607)	(.003676)	(.004095)	(.0032769)
Omitted Y mean	.2705	.2198	.2407	.2883	.2759	.2438
Controls	No	No	No	No	No	No
Fixed effects	No	No	No	No	No	No
N	664	11,173	53,084	117,284	68,505	98,417
R^2	0.0031	0.0036	0.0030	0.0012	0.0008	0.0008

Source: own elaboration based on IIGG and IPUMS. Standard errors, in parentheses, are clustered at the district of residence level. * significant at 10%, ** significant at 5%, *** significant at 1%. Two boys/girls takes value 1 if the first two children are boys/girls and 0 otherwise.

Table A.3: Effect of the gender composition of the children on the likelihood of having a third child with pooled data

Dep: dummy for having third child	1	2	3	4
Two boys	.1452449***	.1439604***	.1405492***	.1392115***
	(.0020354)	(.0020541)	(.0019611)	(.0019512)
Two girls	.151992***	.1507255***	.1469988***	.1456828***
	(.0020906)	(.0020944)	(.0021756)	(.0021673)
Omitted Y mean	.2561	.2561	.2561	.2561
Controls	No	Yes	No	Yes
Fixed effects	No	No	Yes	Yes

N	483,307	483,307	483,307	483,307
R^2	0.0322	0.0334	0.0444	0.0458

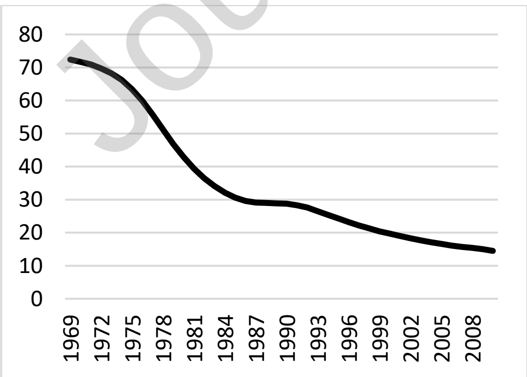
Source: own elaboration based on IIGG and IPUMS. Standard errors, in parentheses, are clustered at the district of residence level. * significant at 10%, ** significant at 5%, *** significant at 1%. Two boys/girls takes value 1 if the first two children are boys/girls and 0 otherwise.

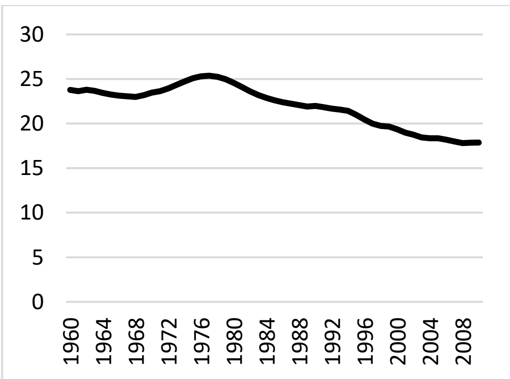
Table A.4: Effect of the gender composition of the children on the likelihood of having a third child according to region of origin with pooled data

Dep: dummy for having third child	Natives	Immigrants	Asian immigrants	Europe+NA immigrants	ALC immigrants
Two boys	.034577***	.0206355***	-.0763512***	.0214727***	.0212037***
	(.0006743)	(.0023856)	(.018576)	(.0037364)	(.0031322)
Two girls	.0390197***	.0619638***	-.0844521***	.1107863***	.0375147***
	(.0006927)	(.0024781)	(.0179732)	(.0038547)	(.0032762)
Test for Two boys=Two girls (p-value)	0.0000	0.0000	0.6965	0.0000	0.0000
Controls	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes

Source: own elaboration based on IIGG and IPUMS. Standard errors, in parentheses, are clustered at the district of residence level. * significant at 10%, ** significant at 5%, *** significant at 1%. Sample sizes and R^2 are omitted for simplicity and are available upon request from the author. Separate regressions are estimated for each column. Two boys/girls takes value 1 if the first two children are boys/girls and 0 otherwise. NA is the acronym for North America and ALC for Latin America and the Caribbean

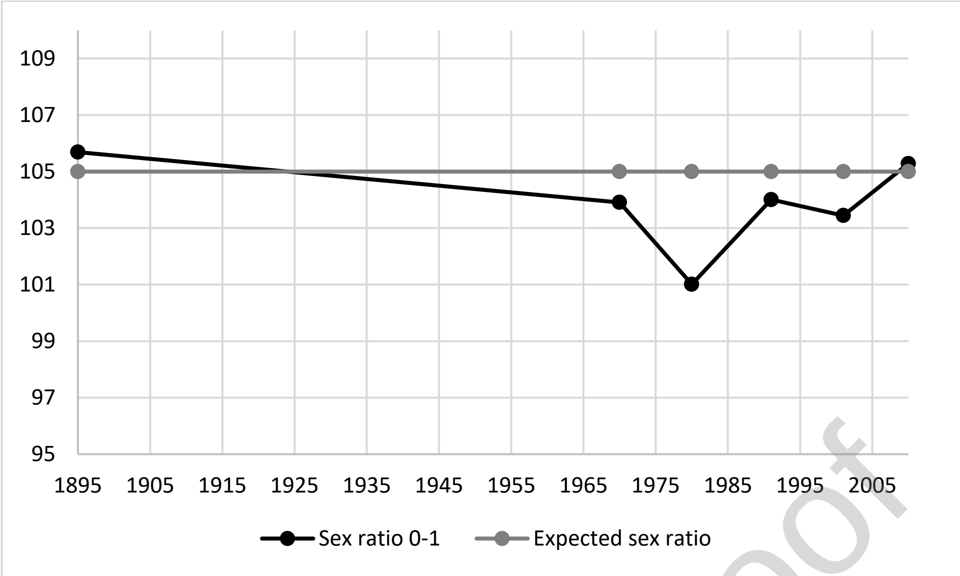
Figure A.1: Mortality (left) and fertility (right) rates in Argentina





Source: own elaboration based on World Bank Development Indicators Database. Note: The mortality rate refers to mortality in children between 0-5 years per 1000 live births. The fertility rate refers to the number of live births per 1,000 people.

Figure A.2: Sex ratios in Argentina



Source: own elaboration based on IIGG and IPUMS. Note: sex ratios are estimated for children between 0-1 years.

References

Angelov, N.; Karimi, A. (2012). Mothers' income recovery after childbearing. Working paper 20, IFAU. https://www.ifau.se/globalassets/pdf/se/2012/wp12-20-mothers-income-recovery-after-childbearing.pdf

Angrist, J.; Evans, W. (1998). Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size. American Economic Review, 88(3), 450-477.

Baland, J.; Cassan, G.; Woltrin, F. (2023). Sex-selective abortions and instrumental births as the two faces of the stopping rule. New measures and world evidence. Working paper 18014, CEPR.

Blanchard, R. (2022). A Novel Method for Studying the Effect of Older Brothers on Sexual Orientation and Its Robustness to Stopping Rule Distortions. Journal of Sex Research, 59(6). https://doi.org/10.1080/00224499.2021.1984379

Blanchard, R.; Lippa, R. (2007). Birth Order, Sibling Sex Ratio, Handedness, and Sexual Orientation of Male and Female Participants in a BBC Internet Research Project. Archives of Sexual Behavior, 36. https://doi.org/10.1007/s10508-006-9159-7

Cruces, G.; Galiani, S. (2007). Fertility and female labor supply in Latin America: New causal evidence. Labour Economics, 14(3), 565-573.

Dahl, G.; Moretti, E. (2008). The Demand for Sons. Review of Economic Studies, 75(4), 1085-1120. https://doi.org/10.1111/j.1467-937X.2008.00514.x

Gabay-Egozi, L.; Nitsche, N.; Grieger, L. (2022). In Their Footsteps or Shadow? Gender Differences in Choosing a STEM Major as a Function of Sibling Configuration and Older Sibling's Gender and Math Ability. Sex Roles, 86, 106-126. https://doi.org/10.1007/s11199-021-01255-0

Gass, K.; Jenkins, J.; Dunn, J. (2006). Are sibling relationships protective? A longitudinal study. The Journal of Child Psychology and Psychiatry, 48(2), 167-175. https://doi.org/10.1111/j.1469-7610.2006.01699.x

Goli, S.; Mavisakalyan, A.; Rammohan, A.; Vu, L. (2022). Conflicts and son preference: Micro-level evidence from 58 countries. Economics & Human Biology, 46. https://doi.org/10.1016/j.ehb.2022.101146

González, F.; Dip, J. (2022). The Impact of Educational Assortative Mating on Income Inequality: Evidence from Argentina. Journal of Population and Social Studies, 30.

González, F.; Dip, J. (2022). Whom Is It *Fair* to Subsidise? Evidence from Argentina. Studies in Microeconomics. https://doi.org/10.1177/23210222221105

González, F.; Santos, M. (2018). Pobreza multidimensional urbana en Argentina: ¿Reducción de las disparidades entre el Norte Grande Argentino y Centro-Cuyo-Sur? (2003-2016). Cuadernos de Economía, 39(81). https://doi.org/10.15446/cuad.econ.v39n81.76486

González, F.; Santos, M.; London, S. (2021). Convergencia y desarrollo en la Argentina urbana (2003-2016). Regions and Cohesion, 11(1), https://doi.org/10.3167/reco.2021.110105.

González, L. (2018). Sex selection and health at birth among Indian immigrants. Economics & Human Biology, 29. https://doi.org/10.1016/j.ehb.2018.02.003

Hammoudeh, W. (2017). The effect of son preference on Palestinian fertility. The Lancet, 390. https://doi.org/10.1016/S0140-6736(17)32038-X

Huang, Z.; Jiang, X.; Sun A. (2024). Fertility and delayed migration: How son preference protects young girls against mother–child separation. Journal of Development Economics, 166. https://doi.org/10.1016/j.jdeveco.2023.103191

Iacovou, M. (2001). Fertility and female labor supply. Working paper 19, Institute for Social and Economic Research.

Instituto de Investigaciones Gino Germani (2015). Microdatos censales 1895. http://censos1869-1895.sociales.uba.ar/

Integrated Public Use Microdata Series (2020). International microdata: Version 7.3 [dataset]. Minneapolis Population Center: IPUMS. https://doi.org/10.18128/D020.V7.3

Jones, T.; Millington, M.; Price, J. (2023). Changes in Parental Gender Preference in the United States: Evidence from 1850–2019. Working paper. https://www.toddrjones.com/papers/Gender_Preference_most_recent.pdf

Kabátek, J.; Perales, F.; Ablaza, C. (2022). Evidence of a Fraternal Birth Order Effect on Male and Female Same-Sex Marriage in the Dutch Population: A Reply to Blanchard, Semenyna, Gómez Jiménez & Vasey. The Journal of Sex Research, 59(6), 697-703. https://doi.org/10.1080/00224499.2021.2002798

Kolk, M.; Schnettler, S. (2012). Parental status and gender preferences for children: is differential fertility stopping consistent with the Trivers-Willard hypothesis? Journal of Biological Science, 45(5). https://doi.org/10.1017/S0021932012000557

Li, W. (2021). The "miseries" of sex imbalance: Evidence using subjective well-being data. Journal of Development Economics, 151. https://doi.org/10.1016/j.jdeveco.2021.102634

Maurin, E.; Moschion, J. (2009). The social multiplier and labor market participation of mothers. American Economic Journal: Applied Economics, 1(1), 251-272. https://doi.org/10.1257/app.1.1.251

Mu, R.; Zhang, X. (2011). Why does the Great Chinese Famine affect the male and female survivors differently? Mortality selection versus son preference. Economics & Human Biology, 9(1). https://doi.org/10.1016/j.ehb.2010.07.003

Rozenweig, M.; Wolpin, K. (2000). Natural "Natural Experiments" in Economics. Journal of Economic Literature, XXXVIII, 827-874.

Tan, C.; Tan, Z.; Zhang, X. (2023). The intergenerational legacy of the 1959–1961 Great Chinese Famine on children's cognitive development. Economics & Human Biology, 51. https://doi.org/10.1016/j.ehb.2023.101300

Tortarolo, D. (2013). Female labor force supply and fertility: Causal evidence for Latin America. Revista de Economía Política de Buenos Aires, 13, 27-65.

Trivers, R.; Willard, D. (1973). Natural Selection of Parental Ability to Vary the Sex Ratio of Offspring. Science, 179(4068), 90-92. https://doi.org/10.1126/science.179.4068.90

Declaration of interest statement

None

Highlights

- In this paper I examine the evolution of parental gender preferences in Argentina (i.e., parents who prefer a certain gender composition in their children).
- I use census microdata that spans the 19th, 20th, and 21st centuries. The estimation strategy exploits the plausibly random assignment in the gender of children.
- The results show a persistent preference for a mixed gender composition (i.e., having at least one boy and one girl) instead of children of the same gender.
- This translates into an increase in the probability of having a third child, conditional on already having two children of between 9%-23% for those couples who have children of the same gender -in relation to couples with children of opposite genders-.