The role of the previous offspring's sex composition in the transition to the third child among Indian mothers born between 1966-1985

Eleonora Trappolini, Patrizia Farina, and Laura Terzera
University of Milan-Bicocca (Italy)

Introduction and Theoretical framework

The preference for sons is a widespread phenomenon especially in Asia, where many "missing girls" are counted every year. India is one of the countries most affected by this preference across class, caste, State religion, and women's education. According to Chao and Yadav estimates (2019), around 15 million female births were missing between 1990 and 2016. Close to 400 thousand female births are missed in India annually, amounting to about three per cent of female births (UNFPA, 2020). The number of missing female births as a per cent of the female is highest in the northern and western regions. Also, the World Population Report released in June 2020 reported that one in three girls missing globally due to sex selection is from India, adding up to 46 million (UNFPA, 2020).

Material and non-material factors are behind the perceived different values of sons and daughters. From a cultural point of view, as the Indian society is patrilineal, the family line depends upon male offspring. In a 2012 survey, 65% of women expect to live only with sons in old age, even if the difference among communities are relevant (UNFPA, 2020). Some ceremonies - as ancestor worship - are performed by sons, the only ones entitled to carry on the family name and a good afterlife of parents. Furthermore, the patrilocal nature of the family implies the co-residence of at least a son in supporting elderly parents. In contrast, the daughter is less valued since she resides with the father or mother-in-law after marriage. Daughters are also perceived as too expensive due to weddings' dowry and the financial burden, especially among poorer families. However, as regards non-material benefits that the mothers of sons would acquire within the household rather than mothers of daughters, Zimmermann L. (2018) shows, with a quantitative approach, no evidence in this sense. Mothers do not benefit from the birth of a son with larger decision-making powers in the household. The benefits also do not depend on the child's age, which is inconsistent with a channel predicting a better bargaining position for women with adult sons who start to take over the management of the household.

To meet the preference for sons, Indian families implement several strategies to contain girls: sex selection at birth and postnatal discrimination that increase the infant mortality rate. A recent study by Rai et al. (2019) provides evidence that sons are less likely to die because of better nutrition and health care than girls, and this behaviour is more substantial for sons born after sisters. The son preference has an impact on both the size and gender composition of siblings in the family (UNFPA, 2020) and on the increase of maternal mortality at reproductive ages due to the repeated and closely spaced pregnancies until the ideal number of sons has been reached (Milazzo, 2018).

At least two specific measures can reveal the degree and intensity of son preferences. The postnatal measure is represented by differences in infant mortality in favour of boys deriving from less health and nutrition care toward the daughters. The other measure, prenatal, is represented by the sex ratio at birth (SRB) when it is significantly higher than the natural value of 105-106 males for 100 females. The higher value of SRB derives mainly from sex-selection abortion if efficient methods to identify the sex of the fetus are available (Boongarts, 2013; Guillmoto, 2012).

While the relevance of the under-five infant mortality rate of boys and girls is converging towards low and very similar levels (World Bank, 2021), the SRB, on the other hand, although declining, still shows hugely high values, especially in some regions of the country (Yadava et al., 2020). From the 1990s, SRB has increased, reaching a peak of 113.6 in 2001 (Zehn et al., 2013). After then it declined to 109.9 in 2011, reaching 111.3 and 112 respectively in West and North states. According to the most recent estimates, SRBs are likely to continue their downward trend until 2025, even if it will stop to around 109 in the worse area of the country, a value far from the 105/106 "norm" (Kaur et al., 2016).

In recent years, the pattern of the distorted SRB in India has been discussed, and several common determinants have been underlined. One of them refers to the connection with the sex composition of the offspring already born. The SRB is very high at higher birth orders and among those who have no son, especially in the Northern, Western, and Central regions. Still, it is commonly elevated in India if the family already has one or more daughters (UNFPA, 2020).

Several authors (e.g. Das Gupta and Bhat, 1997; Guilmoto, 2007) pointed out a relevant fertility decline where the son-preference is strong. In such circumstances, the gender bias could intensify since the family would ensure that sons were born within the desired smaller number of children. This is why women with no previous male births were more likely to have a son than women who already had at least one (Singh et al., 2021). This effect has been relevant in previous decades in India as fertility declined faster than son preference.

The distortion of SRB could also result from the interconnection between caste and wealth, as well education and empowerment of women.

As regards economic prosperity - highly correlated to the social structure of caste - Kaur et al. (2016) pointed out two contradictory conclusions about the relationship between SRB and wealth. The optimist approach points that prosperity leads to an improvement in the SRB as people's perceptions of the value of girls and boys begin to close down. Conversely, the pessimistic approach points that sex selection will rise as socio-economic circumstances improve. The poor tend to shape the composition of a smaller family through recently acquired means and access to sex selection technology, preferring more sons to protect the depletion of family wealth. The effect of education on SBR is also controversial. Some studies emphasize the protective enforcement of education (Pande, 2006; Kaur et al., 2016); others do not. The former assesses that women's education is the most significant factor in reducing son preference.

On the contrary, Das Gupta's (1987) and Jha et al.'s (2006) studies confirm the relevance of education on the decline of child mortality. Women with high education tend to use sex selection more than less educated women. The different results could also depend on the framework of the analysis. They could change according to the birth order or enlighten the single effect instead of the intersection among education, wealth, and family size demand (Singh, 2021).

Since the topic analysed is gender discrimination, the roles assumed by the two parents as man and woman, are considered a crucial factor and, therefore, a gendered approach can clarify the hazy results. For example, Robitaille and Chatterjee (2018) show that sex-selective abortions are more used if both spouses or only fathers prefer male children, while sex-selective abandonment is used if only mothers prefer male children. Regarding an economic perspective, Craigie and Dasgupta (2017) show how a reduction in the gender wage gap helps to reduce the preference for male and female children.

To sum up, most of the literature yields contradictory results in highlighting the positive association between economic and social development with gender discrimination represented by the distortion of SRB, presumably the result of parents' strategies to have at least one son or fewer daughters in a more or less gender unequal context.

Moreover, according to Dubuc and Singh (2018), the SRB is not an appropriate indicator for assessing behavioural change because a small proportion of sex-selective procedures suffices to distort the SRB significantly.

Since the literature shows mixed results about the role of background characteristics, particularly education and wealth, we intend to contribute to the debate on this issue by focusing on the relevance of such structural variables. By applying a sequence analysis technique, which allows to consider timing of events simultaneously, the study aims to measure the pattern of the sex sequence of previous births on the transition to the third child among Indian mothers born between 1966 and 1985.

Data and methods

Data

We used the National Family Health Survey 4 (NFHS-4) carried out in 2015-2016, a nationally representative household survey that provides information on population, health, and nutrition for India covering all 29 states and 7 union territories. The primary objective of the survey is to provide essential data on health and family welfare, as well as data on emerging issues in these areas. The

NFHS-4 sample is a stratified two-stage sample. The primary sampling units were villages in rural areas and the census enumeration blocks in urban areas and were selected with a probability proportional to the size within each stratum. The interview was carried out with the technique of computer-assisted personal interview (CAPI)[1].

We used the Woman's Questionnaire that collects information from all eligible women aged 15-49 (N = 699,686 women), who were asked questions on a large variety of topics: socio-demographic, reproduction, family planning, maternal and child health, marriage and sexual activity, fertility preferences, husband's background and woman's work, HIV/AIDS, other health issues, and domestic violence.

For the purpose of our study, we restricted our sample to women belonging to the birth cohorts from 1966 to 1985 with at least 2 children. The final sample comprises 250,280 women. Henceforth, all the analyses and results shown refer only to this subsample.

Before proceeding, we point out a few shortcomings of the study. First, we do not have information about what has happened before the first child. Second, the sequences were retrospectively built, and based on self-assed information. Thus, we cannot rule out the possibility of under-reporting of female births, especially if they suffered infanticide.

Methods

We used sequence analysis (Abbott and Tsay, 2000) in order to analyse the sex composition of the offspring already born. At the time of writing this draft, we run the analysis on a stratified random sampling taking a 10% sample from each State due to computational reasons[2], thus, the subsample is composed of 25,028 women. In addition, we conducted separate analysis by birth cohort (1966-1970; 1971-1975; 1976-1980; 1981-1985).

---

[1] For further details, see https://dhsprogram.com/pubs/pdf/FR339/FR339.pdf.

[2] Results are robust to different random sample extraction. In the final version of the paper we will use the all dataset.

In our study, we defined as the sequence an ordered string of 'states' representing the childbearing history of women. This technique was successfully employed to analyse migrant working trajectories (e.g. Fuller 2014; Barbiano di Belgiojoso and Ortensi 2019), family formation among migrants and union dissolution (e.g. Barbiano di Belgioso and Terzera 2018; Mikolai and Kulu 2019).

This method approaches the analysis with a life course perspective and can identify sequence models considering two essential elements to define the sequence: the timing of events (period of observation) (Billari and Piccarreta, 2005) and the state space (the set of all possible states, namely 'sex sequences').

As for sex sequences in India, some studies analyse the sex composition of the offspring (e.g. Gellatly and Petrie, 2017 Yadav et al., 2020); however, these studies do not consider the timing of events. Indeed, the first study to apply sequence analysis technique to this topic was the one by Farina and Terzera (2015).

With respect to the timing, we considered 15 years of observation achieving sequences that are 30 semesters (time-unit) long. In this regard, we analysed women's fertility history from 15 to 30 years-old for all cohorts considered. This choice is based on the fact that mothers' mean age at 1st child and at 3rd child is similar across all birth cohorts (around 21 for the 1st child and 25 for the 3rd, respectively) (Table 1A in the Appendix), what changes is the share of mothers who have their 1st and 3rd child (in general), compared to mothers who have their children within 30 years-old (Table 2A in the Appendix).

The state space was the self-declared child-birth taking also into account the child's sex. In this context, we could identify seven states that formed the state space coded as follows: no birth (coded 0); 1 son (S); 1 daughter (D); 2 sons (SS); two daughters (DD); first son then daughter (SD); first daughter then son (DS). Children who died within the first year of life were not considered if followed by another childbirth.

To compare each sequence with all others, we employed the Optimal Matching Analysis (OMA)[3] to obtain a distance matrix[4] that reflected the similarity in the frequency, order and location of states. The distance between two sequences is the number of operations (deletion, insertion, and substitution) to transform one sequence into another. Each operation has a cost. The distance between two sequences is thus the minimum total cost for transforming one sequence into another (Billari, 2001; Gauthier et al., 2009). The lower this distance, the higher the similarity between the two sequences. By default, insertion and deletion cost 1, while substitution costs 2.

Social studies usually propose alternative definitions for 'costs' (Kleinepier et al., 2015). We decided to adopt an ad hoc substitution cost matrix, as previously done by Fuller and Martin (2012), analysing the sex sequence of previous births (Table 1).

Table 1: Substitution cost matrix for the sequence analysis.

	0	**S**	**D**	**DD**	**SD**	**DS**	**SS**
0	0	1	0.50	0.75	1.50	1.50	2
S		0			0.50		1
D			0	0.75		1	
DD				0			
SD					0		
DS						0	
SS							0

Notes: 0 = no birth; S = son; D = daughter

In order to identify the different types of life courses (Aassve et al., 2007), which are interpreted as profiles associated with the particular characteristics of each sequence, we performed a cluster analysis on the distance matrix obtained by the OMA using Ward's algorithm. We opted for a seven cluster solution. We determined the number of clusters by means of the Duda and Caliński methods,

[3] The sequence analysis was performed in Stata 16 using a package implemented by Brisnsky-Fay, Kohler, and Luniak (2006).

[4] For further details, see Abbot and Tsay (2000), which is one of the available methods.

also taking into consideration the number of cases for each cluster with the aim of maximising how informative the classification was (Halpin, 2016).

Finally, since we assume that the sex and timing sequence of previous births affects the probability of a third birth, in order to examine the relationship between the sex sequences of previous births and the probability of a third birth, as well as the determinants of third birth, we used logistic regression models.

To introduce the variable "sex sequence of previous births", which corresponds to the clusters identified with the cluster analysis, in the logistic model, as previously mentioned, we conducted separated analyses by birth cohorts for computational reasons.

Dependent variable

The dependent variable included in the logistic regression is a dichotomous variable indicating the birth of the third child.

Main explicative variable and independent variables

The main explicative variable is the "sex sequence of previous births' cluster". This variable is composed of seven categories, which correspond to the seven clusters that we labelled considering the sex composition, the timing both of the interval between births (medium or short time) and when women started childbearing. The categories are: (1) SS -early[5]- with short birth interval –*reference*, (2) DD with short birth interval, (3) DS -early- with short birth interval, (4) DS -late[6]- with medium birth interval, (5) SS -late- with medium birth interval, (6) SD with short birth interval, (7) SS or SD -much later- with short birth interval.

As independent variables, we included in the model certain structural variables as suggested by the literature: civil status (Never married, Currently married –*reference*, Widowed, Divorced/Separated); type of place of residence (Urban –*reference*, Rural); area of residence (North –*reference*, East, West,

[5] It refers to when women started childbearing. In this case, it means that women started childbearing early.
[6] It refers to when women started childbearing. In this case, it means that women started childbearing late.

South, Central, North-East); education (Low educated *–reference*, High educated); religion (Hinduism *–reference*, Muslimism, Christianism, Other religion) and wealth (Poor *–reference*, Middle, Rich).

To control for infant mortality (i.e. if a woman had experienced the death of a child in the first year of life) we also included in the model the continuous variable ‘children died’.

Results

Descriptive results

Table 1 shows the distribution of women in the sample by number of children and by five-year birth cohort. Across birth cohorts, the share of women with no child, one child and two children is higher among the youngest cohorts (1976-1980 and 1981-1985), while among the latter, the share of those who have three and four and more children is lower than the oldest ones (1966-1970 and 1971-1975). Nonetheless, it should be bear in mind that such a result is related to the fact that at the time of the survey, the oldest birth cohorts had reached the end of their reproductive history.

Table 1: Distribution of all women in the subsample by number of children and birth cohort

	1966-1970	**1971-1975**	**1976-1980**	**1981-1985**	**Total**
0	4.1	4.1	4.9	7.5	5.3
1	8.3	8.6	9.0	14.7	10.4
2	26.3	28.7	33.6	37.3	32.0
3	25.1	26.4	26.8	24.3	25.7
4+	36.2	32.2	25.7	16.2	26.6
N. of women	*6,839*	*7,682*	*9,211*	*9,545*	*33,277*

Source: Authors’ elaboration on NFHS-4 data.

The main structural and reproductive characteristics of the subsample composed of those women with at least two children are shown in Table 2.

Almost all women in the subsample are married or have been married. Despite the share of women with high education (secondary or tertiary) increased over time, a relevant quota of women still has low education (72.4% in the oldest cohort), while among the youngest cohort this percentage drops to 53.3%. The level of wealth index changed across birth cohorts: the percentage of women declaring ‘poor’ wealth index increases; and those declaring ‘rich’ wealth index decreases. 28% of women live in an urban context and these characteristics appear almost unchanged across birth cohorts as well as the State distribution, with around 27% of women living in the North.

Table 2: Characteristics of women with at least two children by birth cohort, India 2015-2016

Variables	**1966-1970**	**1971-1975**	**1976-1980**	**1981-1985**	**Total % (mean)**
Civil status					
Never married	0.1	0.0	0.0	0.1	0.1
Currently married	89.2	92.1	94.3	97.2	93.5
Widowed	9.9	6.9	4.7	2.1	5.6
Divorced/Separated	0.9	1.1	0.9	0.6	0.9
Education					
Low educated	72.4	67.1	59.8	53.3	62.3
High educated	27.6	32.9	40.2	46.7	37.7
Wealth index					
Poor	37.0	39.6	42.9	45.8	41.8
Middle	20.9	21.6	20.6	20.8	20.9
Rich	42.1	38.8	36.4	33.4	37.3
Type of residence					
Urban	28.8	28.9	28.3	26.5	28.0
Rural	71.2	71.1	71.7	73.5	71.9
State					
North	25.9	26.8	27.4	26.7	26.8
East	17.2	16.9	17.5	18.0	17.4
West	13.3	15.6	13.2	14.1	14.1
South	15.9	15.0	15.2	13.9	14.9
Central	13.2	12.4	11.7	12.2	12.3
North-East	14.5	13.3	15.0	15.1	14.5
Religion					
Hinduism	77.4	76.1	74.8	76.0	76.0
Muslimism	11.4	12.8	13.5	12.9	12.7
Christianism	6.8	6.6	7.1	6.9	6.9
Other religion	4.4	4.5	4.6	4.2	4.4
Mother's age at first birth[a]					
<20	39.6	42.6	42.6	45.1	42.6
20-24	51.3	50.3	49.6	47.2	49.5
25-30	9.1	7.1	7.8	7.7	7.9
Total number of women	*5,162*	*5,928*	*7,064*	*6,874*	*25,028*
Mean number of children	(3.6)	(3.5)	(3.2)	(2.9)	(3.3)
Total number of children	*18,686*	*20,468*	*22,711*	*19,856*	*81,721*

Notes: The table shows percentages and should be read in columns. [a] Excluding women with no child.
Source: Authors’ elaboration on NFHS-4 data.

Most of women in the subsample follow/declare the Hinduism religion, and these characteristics are similar across all birth cohorts. As regards mother’s age at first birth, the Table shows that most women started childbearing within 24 years-old and, and in any case, around 8% women have started it between 25-30 years old, with no relevant differences among birth cohorts. The (total) mean number of children is 3.3, reaching its maximum, 3.6, among women born between 1966-1970.

From Table 3 it is possible to deduce the role of the sex sequence of the first two births. This is particularly evident comparing the sequences of two sons (SS) and two daughters (DD): while for the former the share of women who have two children is consistently higher than that of women who have already had a third child, the fertility behaviour is reversed (and more intensely) when the first two children are female. Furthermore, at a descriptive level, there are no relevant differences across birth cohorts.

Table 3: Gender sequences of births for women with at least two children in the subsample by birth-cohort, India 2015-2016

	1966-1970		**1971-1975**		**1976-1980**		**1981-1985**	
Sequence	***At least two***	***Two***	***At least two***	***Two***	***At least two***	***Two***	***At least two***	***Two***
DS	24.3	24.1	24.1	26.8	23.9	28.4	23.9	27.9
SD	24.9	25.5	24.8	26.6	23.9	28.3	23.9	26.8
DD	25.2	9.1	27.0	8.9	30.4	9.9	31.9	12.2
SS	25.6	41.3	24.1	37.7	21.8	33.4	20.3	33.1
N. of women	*3,790*	*1,372*	*4,194*	*1,734*	*4,599*	*2,465*	*3,812*	*3,062*

Notes: The table shows percentages and should be read in columns. (0) means no child; (S) son; (D) daughter.
Source: Authors' elaboration on NFHS-4 data.

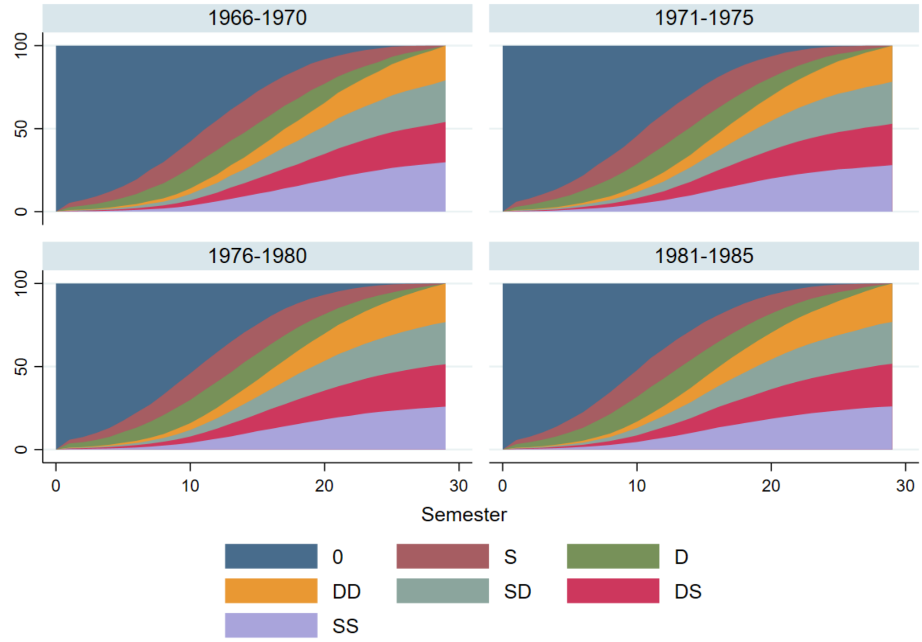
Sequence analysis and cluster analysis

The chronogram in Figure 1 shows the distribution of states over time by birth cohort. Different colours represent different states. For example, the sex composition represented by two daughters (DD) is marked by yellow, while two sons (SS) is marked by light violet.

The figure shows that birth sequences for women who have at least two children are similar across birth cohorts, in terms of timing and sex composition. For example, the share of DD has slightly increased from 20.9% for the cohort 1966-1970 to 23.1% for the cohort 1981-1985. The proportion of women with one son and one daughter (SD or DS, regardless the order) floats around 25.0%. Finally, the share of SS has slightly decreased from 29.8% for the cohort 1966-1970 to 26.0% for the cohort 1981-1985.

Next, we conducted cluster analysis to understand whether there are distinct groups of sex compositions. This analysis identifies seven different groups, and we labelled each cluster taking into account the distribution of the colours along the x-axis, and the timing of the interval between births (long, medium or short time) (Figure 2)[7].

Figure 1: Chronograms of the sex composition of the offspring over time, by birth cohort



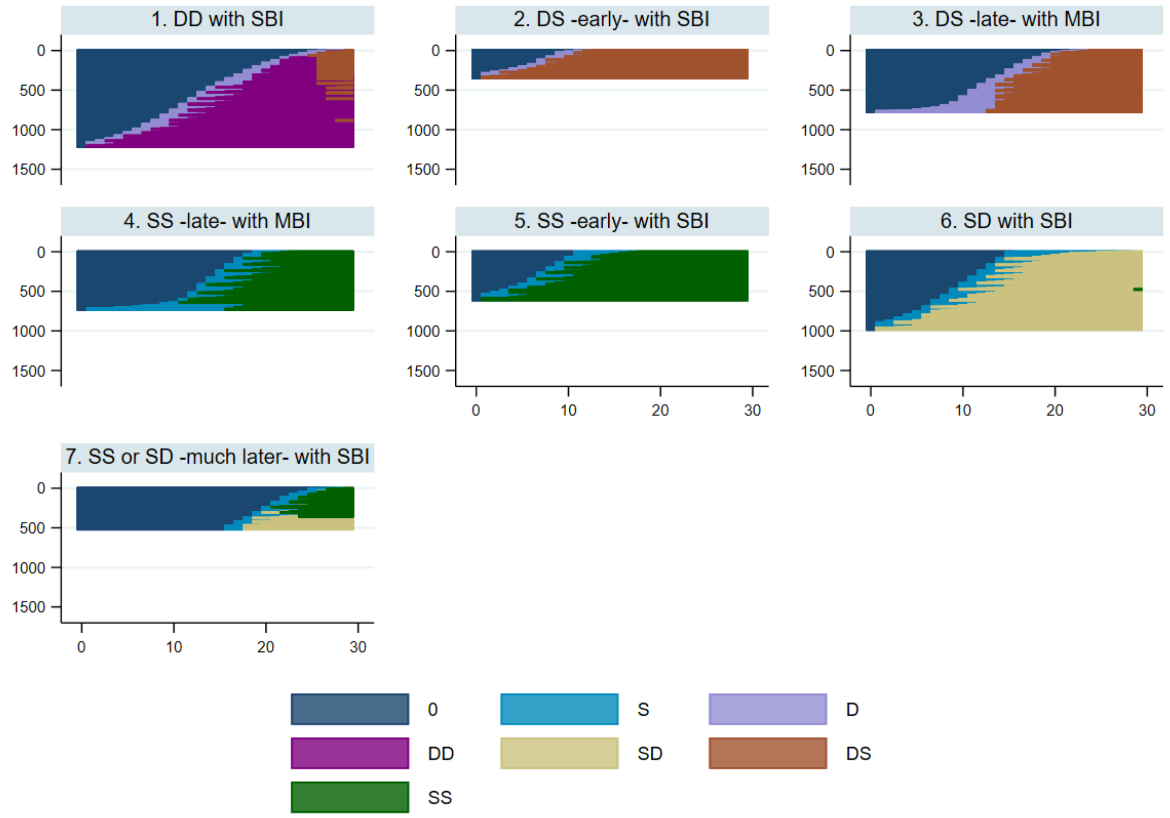
Notes: **S** = son, **D** = daughter.
Source: Authors' elaboration on NFSH-4 data.

Analysing the clusters of the first birth cohort (1966-1970), the clusters 1 and 6 represent the group of women who have a daughter as a second child after either a daughter (1) or a son (6) characterised by a short time interval between births without a particular timing trend. The clusters 2 and 3 are represented by the DS sex sequences with differences in terms of the timing of events. While in the cluster 2 the time interval between birth is short and most of women started childbearing early and have the second child (a son) around the 12th semester (about 21 years old), in the cluster 3 women started childbearing late and the time interval between birth is medium. As regards clusters 4 and 5 are characterised by the SS sex composition. In cluster 4 women started childbearing late and with a medium time interval between births, while in cluster 5 women started childbearing early and have

[7] We show the cluster analysis only for the first birth cohort because for the others the pattern is quite similar.

the second son around the 16th semester (about 23 years old). Finally, cluster 7 is dominated by women who started childbearing much later than the other clusters and who have as first child a son and as second child another son or a daughter with a short time interval between births.

Figure 2: cluster analysis of the sex composition of the offspring over time. 1966-1970 birth cohort



Notes: **SBI** = short time interval between births; **MBI** = medium time interval between births; **S** = Son; **D** = daughter.
Source: Authors' elaboration on NFSH-4 data.

Logistic model for the transition to the third child

Table 4 confirms the relation between the transition to the third child and the sex sequence of the offspring already born. Results display that for all birth cohorts, the sex sequence of previous births, in particular the presence of a daughter in the family, is strongly associated to the probability of having a third birth. Actually, the sex sequence "DD with short birth interval" reports the higher ORs than the other sequences, especially among the youngest women (OR = 6.32, p-value <.001). Another important factor is represented by when women start childbearing. Those who start earlier have a higher propensity to have a third child with respect to those who start later. Indeed, the timing of events is one of the two crucial factors in determining the propensity to the final fertility, thus, this, moreover, is a sign of the group which is more likely to have more children in general.

Education and wealth are negatively associated with transition to higher orders; higher level of education, as well as reporting higher level of wealth decreases the probability of having a third child.

Table 4: Adjusted ORs for the transition to the third child, India 2015-2016

Variables	1966-1970	1971-1975	1976-1980	1981-1985
Sex sequence of previous birth's cluster (ref. SS -early- with SBI)				
DD with SBI	1.13	2.19***	2.24***	6.32***
DS -early- with SBI	2.95***	1.20*	1.38**	1.23*
DS -late- with MBI	0.87	0.22***	0.37***	0.51***
SS -late- with MBI	0.46***	0.37***	0.36***	0.47***
SD with SBI	1.08	1.23	1.17	1.61***
SS or SD -much later- with SBI	0.29***	0.48***	0.46***	0.65***
Education (ref. Low educated)				
High educated	0.52***	0.42***	0.45***	0.43***
Wealth (ref. Poor)				
Middle	0.95	0.88	0.71***	0.65***
Rich	0.66***	0.55***	0.48***	0.35***
Religion (ref. Hinduism)				
Muslimism	2.10***	2.25***	2.28***	2.62***
Christianism	1.18	2.61***	2.67***	2.40***
Other religion	0.97	0.74*	0.73*	0.91
Place of residence (ref. Urban)				
Rural	0.96	0.99	1.04	0.87*
Civil status (ref. Currently married)				
Never married	1.87	-	-	1.39
Widowed	0.80*	0.92	0.68**	0.66*
Divorced/Separated	0.46*	0.60	0.41**	0.54
Area of residence (ref. North)				
East	0.82*	0.76**	0.90	0.68***
West	0.82*	0.80*	0.72***	0.65***
South	0.30***	0.30***	0.30***	0.28***
Central	0.92	1.13	0.90	0.61***
North-East	0.83	0.60***	0.68***	0.66***
Children died	0.97	1.19*	1.13	1.11

Notes: SBI = short time interval between births; MBI = medium time interval between births; S = son; D = daughter.
Source: Authors' elaboration on NFSH-4 data.

Religion as well is related to the probability of having a third child. For all birth cohorts, following/declaring the Muslimism and Christianism, increases such propensity compared to those following/declaring the Hinduism religion. Unsurprisingly, being currently married is positively associated with the probability of having a third child. Finally, for all birth cohorts, women residing in the North of India have a higher probability for the transition to the third birth than women residing in other part of the State.

Discussion

By using the sequence analysis technique, the study aims to analyse the effect of the sex composition of the offspring already born and to explore what are the determinants for the transition to the third child among Indian mothers born between 1966-1985.

Women's reproductive histories seems to be the result of personal strategies that imply a choice about the sex and number of children to be welcomed in the family.

The findings of this study highlight the importance of two elements in shaping the probability of having a third child: the sex composition of previous births and the timing of events. Indeed, we find that the presence of at least one son in the family is fundamental for Indian parents for all birth cohorts. Mothers without males are much more likely to have another child compared to mothers with at least one son, and this pattern is particularly evident in the youngest cohort (1981-1985). In addition, we observe that women who start childbearing earlier have a higher propensity for the transition to higher orders, conversely, those who start later are less likely to have a third child.

The results also detect the well-known effect of some structural variables for the transition to the third birth. For all birth cohorts we observe that high educated women are less likely to have a third birth, as well as those who report higher level of wealth index. Other relevant factors are represented by the religion and the area of residence. As regards the former, those who follow/declare the Hinduism religion are less likely to have a third child compared to the others; while for the latter, women residing in the North of India are more likely to have a third child, especially for the youngest cohort (1981-1985), since such result is confirmed for all areas of the State. These results are in line with the study by Farina and Terzera (2015), where using the same approach and the NFHS-3, authors analysing the effect of sequences on the probability of having a fourth child among Indian mothers born between 1956-1960.

References

1. Aassve, A., Billari, F. C., & Piccarreta, R. (2007). Strings of adulthood: A sequence analysis of young British women's work-family trajectories. *European Journal of Population/Revue européenne de Démographie*, 23(3), 369-388.
2. Abbott, A., & Tsay, A. (2000). Sequence analysis and optimal matching methods in sociology: Review and prospect. *Sociological Methods & Research*, 29(1), 3-33.
3. Billari, F. C. (2001). Sequence analysis in demographic research. *Canadian Studies in Population [ARCHIVES]*, 439-458.
4. Billari, F. C., & Piccarreta, R. (2005). Analyzing demographic life courses through sequence analysis. *Mathematical Population Studies*, 12(2), 81-106.
5. Bongaarts, J. (2013). The implementation of preferences for male offspring. *Population and Development Review*, 39(2), 185-208.
6. Chao F., Yadav A. (2019) Levels and Trends in the Sex Ratio at Birth and Missing Female Births for 29 States and Union in *Foundations of Data Science* 1 (2), 177-196.
7. Craigie, T. A., & Dasgupta, S. (2017). The gender pay gap and son preference: evidence from India. *Oxford Development Studies*, 45(4), 479-498.
8. Das Gupta, M. D., Chung, W., & Shuzhuo, L. (2009). Evidence for an incipient decline in numbers of missing girls in China and India. *Population and Development Review*, 35(2), 401-416.
9. Das Gupta, M., & Mari Bhat, P. N. (1997). Fertility decline and increased manifestation of sex bias in India. *Population Studies*, 51(3), 307-315.
10. Dubuc, S., & Sivia, D. S. (2018). Is sex ratio at birth an appropriate measure of prenatal sex selection? Findings of a theoretical model and its application to India. *BMJ Global Health*, 3(4), e000675.
11. Farina P., & Terzera L. (2015). Sequences of Birth History among Indian Mothers born between 1956-1960. *Indian Journal of Applied Research*, 5 (7), 315-317.
12. Fuller, S., & Martin, T. F. (2012). Predicting Immigrant Employment Sequences in the First Years of Settlement 1. *International Migration Review*, 46(1), 138-190.
13. Gauthier, J. A., Widmer, E. D., Bucher, P., & Notredame, C. (2009). How much does it cost? Optimization of costs in sequence analysis of social science data. *Sociological Methods & Research*, 38(1), 197-231.
14. Gellatly, C., & Petrie, M. (2017). Prenatal sex selection and female infant mortality are more common in India after firstborn and second-born daughters. *J Epidemiol Community Health*, 71(3), 269-274.
15. Guilmoto C. Z. (2007). Characteristics Of Sex Ratio Imbalance in India, and Future Scenarios. Paper presented at the 4th Asia Pacific Conference on Reproductive and Sexual Health and Rights, UNFPA.
16. Guilmoto, C. Z. (2012). Sex Imbalances at Birth, UNFPA Asia, and Pacific Regional Office http://www.unfpa.org/publications/sex-imbalances-birth.
17. Jha, P., Kesler, M. A., Kumar, R., Ram, F., Ram, U., Aleksandrowicz, L., & Banthia, J. K. (2011). Trends in selective abortions of girls in India: analysis of nationally representative birth histories from 1990 to 2005 and census data from 1991 to 2011. *The Lancet*, 377(9781), 1921-1928.
18. Kaur R, Bhalla S. S., & Agarwal M.K. Ramakrishnan P. (2016). Sex Ratio at Birth The role of Gender, Class, and Education UNFPA.
19. Kaur R., & Kapoor T. (2021) The Gendered Biopolitics of Sex Selection in India in *Asian Bioethics Review*, 13, 111–127.
20. Kleinepier, T., de Valk, H. A., & van Gaalen, R. (2015). Life paths of migrants: A sequence analysis of Polish migrants' family life trajectories. *European Journal of Population*, 31(2), 155-179.
21. Milazzo A. (2018) Why are adult women missing? Son preference and maternal survival in India In *Journal of Development Economics*, 134, 467-484.

22. Pande, R., & Malhotra A. (2006). Son Preference And Daughter Neglect In India - What Happens To Living Girls?. New Delhi: International Center for Research on Women.

23. Raj, N. Johnsa, McDougal L, Trivedi A., Bharadwaj P. Silverman J., Kumar K., Ladusingh L., & Singh A. (2019) Associations Between Sex Composition of Older Siblings and Infant Mortality in India from 1992 to 2016 in *EclinicalMedicine*, 14,14-22

24. Robitaille, M. C., & Chatterjee, I. (2018). Sex-selective abortions and infant mortality in India: The role of parents' stated son preference. *The Journal of Development Studies*, 54(1), 47-56.

25. Singh, A., Kumar, K., Yadav, A. K., James, K. S., McDougal, L., Atmavilas, Y., & Raj, A. (2021). Factors Influencing the Sex Ratio at Birth in India: A New Analysis based on Births Occurring between 2005 and 2016. *Studies in Family Planning*, 52(1), 41-58.

26. UNFPA (2020) Sex Ratio at Birth in India Recent Trends and Patterns, Report prepared by Kulkarni P. for *United Nations Population Fund.*

27. Yadava A., Anandb A., Singhc R., & Jungari S. (2020). Sex composition and son preference in India and major states: Evidence from the National Family Health Survey-4 (2015-16) in *Clinical Epidemiology and Global Health*, 8, (4), 1140-1146

28. Zhen, G., Das Gupta, M. & Shuzhuo L. (2013) "Missing Girls" in China and India: trends and policy impacts, Maryland Population Research Center, Working Paper MPRC-2013-01 http://www.popcenter.umd.edu

29. Zimmermann, L. (2018). It'sa boy! Women and decision-making benefits from a son in India. *World Development*, 104, 326-335.

APPENDIX

Table 1A: Mothers' mean age at 1st child and at 3rd child

	1st child	**3rd child**
1966-1970	21.08	24.84
1971-1975	20.77	24.62
1976-1980	20.85	24.53
1981-1985	20.98	24.39

Table 2A: Share of mothers who have the 1st/3rd child (all) and those who have the 1st/3rd child within 30 years old by birth cohort

1st child	**1966-1970**		**1971-1975**		**1976-1980**		**1981-1985**	
	All mothers	**Mothers aged ≤ 30**	**All mothers**	**Mothers aged ≤ 30**	**All mothers**	**Mothers aged ≤ 30**	**All mothers**	**Mothers aged ≤ 30**
0	3.71	7.25	3.72	6.94	4.67	7.57	7.19	8.66
1	96.29	92.75	96.28	93.06	95.33	92.43	92.81	92.33

3rd child	**1966-1970**		**1971-1975**		**1976-1980**		**1981-1985**	
	All mothers	**Mothers aged ≤ 30**	**All mothers**	**Mothers aged ≤ 30**	**All mothers**	**Mothers aged ≤ 30**	**All mothers**	**Mothers aged ≤ 30**
0	34.33	44.83	37.41	45.90	43.65	50.23	55.53	57.88
1	65.67	55.17	62.59	54.10	56.35	49.77	44.47	42.12