Exogenous variation in family size: Sex composition of children and the third birth

Suhishan Bhandari*

Abstract

The paper investigates fertility preferences using Nepal Demographic and Household Surveys (NDHS) data from 1996 to 2022. Owing to the assumption that sex distribution at birth is random and exogenous to fertility, the causal influence of sex composition on family size is estimated. The findings reveal a persistent preference for sons, with families having two daughters more likely to conceive an additional child than those with at least one son. Wealthier and urban women exhibit stronger son preferences, especially since the early 2000s.

Keywords : Sex preference, Fertility, Demographic Trends

JEL Classification: I10, J11, J13, J16

*Kathmandu University School of Arts, Hattiban, Lalitpur

1 Introduction

Son preference in fertility is a widely documented phenomena in developing countries like Vietnam (Haughton and Haughton, 1999), India (Pande and Astone, 2007; Das, 1987), China (Zheng et al., 2018; Graham et al., 1998), Nigeria (Milazzo, 2014), and in different parts of the developing world of Central Asia and Eastern Europe (Filmer et al., 2008). The paper aims to scrutinize son preference in the context of Nepal, with the hypothesis in question: Do families having two daughters exhibit higher proclivity to have additional children than those who have had at least one son?

A family may exhibit gender-based fertility preferences for a number of reasons. First, theoretically, economic factors with strong disincentives to raise daughters such as dowry requirement in marriages, bridal migration and thereby, low ROI (return on investment) on human capital for daughters, old-age support provided by sons, and discrepancies in labor market opportunities may induce families to prefer at least one son (Monica Das Gupta and Hwa-Ok, 2003). Second, cultural preferences such as strong patrilineal family systems, society's view on gender roles, shifts in women's power over her life-cycle (Gupta, 1996), may determine skewed gender preferences.[1]

Nepal echoes these afore-mentioned findings of prevalent son preferences in fertility as exemplified by unanimous findings in Brunson (2010); Leone et al. (2003); Stash (1996); Karki (1988) and Sapkota et al. (2019). The study by Sapkota et al. (2019) uses a multinomial logit model to assess the impact of son preference in fertility using Nepal Demographic and Health Survey (NDHS) data from 2016. The study finds considerable heterogeneity of son and daughter preferences, whereby less-educated, poorer and rural women are more likely to prefer a son than educated, richer and urban women. This is one of only a handful of studies that employs an empirical approach other than descriptive survey results, which has been found to be the norm in Nepal. Worldwide, the general approach to investigate sex-preference fall into one of three categories: quantitative survey results, examination of sex ratios, and/or examination of parity progression using risk/hazard models (like Cox Proportional Hazard Models). Given the insufficiency of mere descriptive statistics in this endeavour, parity progressions (like that propounded by (Ben-Porath and Welch, 1976) and hazard models have been a popular choice, but these methods aren't quite capable of establishing causality.

In response to these considerations, the empirical approach adopted by this paper is a continuation from Angrist and Evans (1998)'s seminal study on fertility preferences and female labour market outcomes. Although female labour market outcomes is beyond the scope of this research, we still have the opportunity to analyze the causal effect of sex composition on

[1] Extensive analysis on the determinants of son preference in the context of rural India is done by Pande and Astone (2007)

continued fertility.[2]

There have been various studies that have adopted Angrist and Evans (1998)'s approach. While the original Angrist and Evans (1998) paper showed that mothers in USA exhibited a preference for balanced sex composition, similar research in Korea (Lee, 2008) showed that conceiving two girls as first two children increases the likelihood of further childbirth. However, similar studies in Australia (Moschion, 2013), UK (Iacovou, 2001) and Latin America (Cruces and Galiani, 2007) echo a preference for balanced sex composition (a preference for both sons and daughters). According to Iacovou (2001) and findings from Williamson (1983) in an analysis of over 50 countries, a marked preference for sons is found among developing nations while the developed nations are shown to prefer balanced sex-composition of children. This is further impetus to look into the nature of sex preference in Nepal, given its developing status, as well as it unique trajectory of demographic transition in the last 30 years. Thus, utilizing the theoretical assumption that sex composition of children is randomly assigned and thereby exogenous to fertility decisions, we look at the causal impact of sex composition on continued fertility in Nepal.

The logical implication of continued fertility, besides the obvious matter of exogenous increase in family size, is that if underlying son preference does indeed induce further childbirth in households, then a two-girl sibling household may find itself in a completely different economic predicament than a household with at least one boy because the former will be more likely to add another sibling. Simply, if families follow a stopping rule based on the gender composition of their children, then, girls will have more siblings than boys, on average (Barcellos et al., 2014). This may mean that boys and girls live in households with different observed and unobserved characteristics. While the implication of this conjecture is beyond the scope of this study, it does provide a basis/motivation to look into son-biased stopping rules.

2 Data and Descriptive Statistics

Data Source

For the purpose of this study, I use Demographic and Health Survey for Nepal (NDHS), provided by United States Agency for International Development (USAID). The data offers comprehensive information on population demographics, health indicators, and key socioeconomic factors through nationally representative surveys. The data is amenable to the current research

[2]As part of their 2SLS approach, while the second least squares estimation is of minimal relevance here, the results of the first OLS (namely: regression of proclivity to have additional children on sibling sex composition) is pertinent.

Table 1: Summary Statistics (Mean) for all Women Aged 15-49 from all NDHS iterations since 1996

VARIABLES	**1996**	**2001**	**2006**	**2011**	**2016**	**2022**
Mean Children Ever Born	3.45	3.30	2.44	2.10	2.02	1.86
Proportion with more than 2 children	0.592	0.581	0.429	0.367	0.348	0.307
Education in Single Years	1.14	1.56	3.18	4.55	5.15	6.09
Brahmin/Chhetri	0.338	0.311	0.337	0.408	0.355	0.335
Hindu	0.874	0.875	0.866	0.854	0.870	0.850
Age at first birth	19.10	19.18	19.27	19.53	19.55	19.83
Observations	8,429	8,554	10,789	12,641	12,819	14,824

due to a) its provision of key fertility related information such as: clearly specified birth order as separate columns with sex and DOB of each child, maternal information such as total children ever born and dates for every birth; and b) socioeconomic data: religion, ethnicity, education, family size, wealth index, and so on. Similarly, examining the successive 5-year intervals of NDHS from 1996 to 2022 allows us to analyze our findings over time and assess their temporal robustness.

Descriptive Statistics

Before moving on to the subsequent empirical work in section III, table 1 presents preliminary demographic information on women aged 15 to 49, for each of the past six NDHS. Table 2 presents demographic information for women aged 15-49 who had two children or more.

Table 1 reveals a consistent decline in the average number of children per woman reflecting decreased fertility rates as per the UN's population estimates for Nepal (UN, 2022). This trend is mirrored in family size, with fewer women having more than two children. Corresponding to UNESCO's observations of an increasing adult literacy rate among women (UNESCO, 2023), the dataset shows a positive trajectory in women's education. However, notably, the age at first birth variable indicates that, on average, women in the NDHS sample experience their first childbirth as early as 19.

Table 2 provides a focused analysis on women with two or more children in the context of the current paper. The mean children ever born statistic exhibits a tautological increase compared to Table 1. Additionally, women in this group tend to give birth to their first child marginally earlier in life. The average birth interval between the first and second child remains consistent

Table 2: Summary Statistics (Mean) for Women Aged 15-49 who had two or more than two children from all NDHS iterations since 1996

VARIABLES	**1996**	**2001**	**2006**	**2011**	**2016**	**2022**
Mean Children Ever Born	4.42	4.20	3.88	3.50	3.33	3.06
Education in Single Years	0.86	1.18	1.73	2.63	2.99	3.94
Proportion with Two boys	0.264	0.264	0.254	0.266	0.252	0.257
Proportion with Two Girls	0.238	0.247	0.244	0.235	0.230	0.246
Brahmin/Chhetri	0.33	0.31	0.33	0.41	0.35	0.33
Hindu	0.87	0.87	0.87	0.86	0.88	0.85
Age at First Birth	18.97	19.06	19.09	19.25	19.19	19.39
Age at Second Birth	21.65	21.76	21.81	22.15	22.25	22.69
Observations	6,335	6,451	6,436	7,106	7,169	8,163

at approximately 2-3 years. Further, if a woman has two children, there is a higher likelihood of having two boys rather than two girls, aligning with the expected sex ratio estimates propounded by WHO. The educational profile of women with two or more children reveals a consistent trend across NDHS iterations, indicating that they have, on average, fewer completed education years compared to the general population of women (Table 1)

3 Empirical Strategy

We want to see whether the propensity to have more than two children differs by sex-composition of the first two children. The following regression is used to to link *More Than Two* (y_i) variable with *Two Girls* (x_i), for every NDHS iteration from 1996 to 2022 A.D. (t).

$$y_{it} = \alpha_0 + \alpha_1 x_{it} + \alpha_2 \mathbf{Z}_{it} + \varepsilon_{it}$$

y_i is a categorical variable for *More Than Two* (1 if the mother had more than two children and 0 if not).

x_i is a categorical variable for *Two Girls* where, if s_i is a female first born categorical variable, and s_2 is a second born female,

$$TwoGirls = s_1 * s_2$$

Z_{it} is a vector of parent/household-level covariates (mother's age, years of education, mother's age at first birth, mother's age at second birth) as well as dummies for demographic indicators such as Brahmin/Chhetri, religion (Hindu, Muslim, Buddhist, Christian and others), urban/rural status, and wealth index quantiles as provided by the NDHS.[3]

Note on Empirical Specification

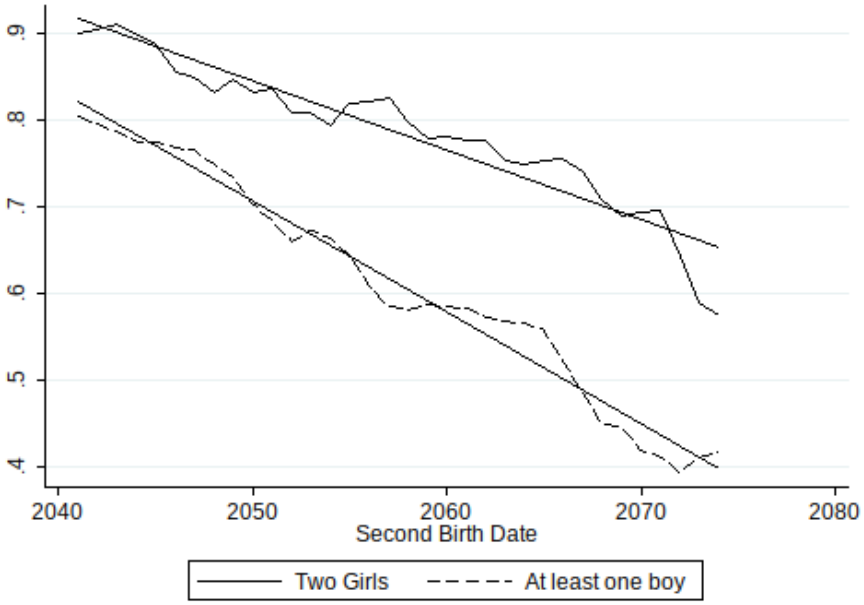
The theoritical justification for the empirical strategy is jointly motivated by parity progression considerations on Ben-Porath and Welch (1976) and the usage of a similar regression as intrument in Angrist and Evans (1998). Suppose a family already has $n_y \geq 1$ children and they are trying to decide how many additional children to have. If the family prefers at least one boy i.e. if a family is boy-biased, then an all girl sibling composition reduces the utility from n_y. This in turn increases the chances that parents will try to have additional children because the marginal utility for that additional child has increased.[4]

The dependent variable $MoreThanTwo$ (y_i) is essentially a transitional probability metric (the probability of $(n+1)$ children given n), for parity progression 2 to 3 i.e. $n = 2$. The specific choice of $n = 2$ is in part due to Nepal's decreasing fertility rates and inclination towards smaller families, which means that fewer and fewer families progress to parities higher than 3. This does mean, however, that estimates for earlier DHS iterations (mainly from 1996 to 2006 A.D.) may not be theoretically robust, given their existent higher fertility rates. Parity 2, then, seems to stand out as the best compromise, and thereby maybe the reason of its wide adoption in the literature. A way to visualize this would be to group women into distinct cohorts based on the year in which they birthed their second child, and plot it against the transitional probability from 2 to 3 children.

[3] According to the DHS Program : The wealth index is a composite measure of a household's cumulative living standard. The wealth index is calculated using easy-to-collect data on a household's ownership of selected assets, such as televisions and bicycles; materials used for housing construction; and types of water access and sanitation facilities. The wealth index places individual households on a continuous scale of relative wealth. DHS separates all interviewed households into five wealth quintiles to compare the influence of wealth on various population, health and nutrition indicators

[4] This theoretical analysis is largely influenced by similar analysis on Angrist and Evans (1998)

Figure 1: Proportion of women who had more than two children by respective birth cohorts i.e. the year they birthed their second child



Note : All six DHS iterations were pooled together to formulate this graph. We only take the years 2040 B.S. (1983 A.D.) to 2075 B.S. (2018 A.D.) because this encapsulates more than 90 percent of our sample. The line plot is a three-year moving average plot, and the straight lines are lines of best fit.

The figure shows that while the overall propensity to have more than 2 children is decreasing over time for both groups, the rate of decrease is much slower for families whose first two children are girls. While this may, in fact, be a result of our preferred choice of the second parity, it acts as a basic intuition/foundation for understanding the result from the regression in the later sections.

Note on Causality

Since x_i (sex composition of children at birth) is assumed to be exogenously determined, the regression of *More Than Two* on *Two Girls* can be causally interpretated. To check this, we can compare some demographic characterisitics of people who had first two children as girls and those who had at least one boy i.e. $TwoGirls = 1$ against $TwoGirls = 0$. Table 2 reports *Two Girls* contrasts for mother's age, age at first birth, ethnic and religion indicators as *Brahmin/Chhetri* and *Hindu,* and wealth indices for all NDHS iterations from 1996 onwards.

Table 3: Differences in mean (Standard error) for demographic variables by *TwoGirls*.

Variable	**1996**	**2001**	**2006**	**2011**	**2016**	**2022**
Age	0.216	0.334	-0.048	-0.411*	0.265	-0.042
	(0.237)	(0.229)	(0.231)	(0.212)	(0.214)	(0.196)
Age at second birth	0.0226	0.0014	0.054	0.065	0.207**	0.201**
	(0.103)	(0.094)	(0.095)	(0.097)	(0.096)	(0.096)
Brahmin/Chhetri	-0.005	0.017	-0.0027	0.0034	0.0058	0.0018
	(0.014)	(0.013)	(0.136)	(0.013)	(0.013)	(0.012)
Hindu	-0.014	0.0014	-0.0053	0.0135	0.012	-0.0095
	(0.0096)	(0.009)	(0.0095)	(0.009)	(0.0091)	(0.0089)
Years of education	-0.073	0.052	-0.0114	0.025	0.036	0.201*
	(0.078)	(0.077)	(0.093)	(0.103)	(0.108)	(0.11)

Note : Standard Errors are reported in parentheses. The samples are the same as reported in Table 2

Out of 30 estimates, only four samples show contrasts significantly different from 0, out of which the difference of 0.2 of woman's age at second birth is merely a difference in 2.4 months and a difference of 0.4 in mother's age is a difference of almost 5 months. Similarly, a difference of 0.2 in years of education is also a difference of an additional 2.4 months in school. Since the variables showcased in Table 3 are used as covariates in the main regression, such differences would not impact the robustness of our estimates.

4 Main Results

The discussion of main results is disaggregated into their respective NDHS iteration. A comprehensive summary is provided subsequently.

NDHS 1996

The results linking sex mix and fertility for 1996 NDHS iteration is reported in Table 4. Women with two daughters are estimated to be 3 percentage points more likely to have a third child while those with two sons exhibite a negative correlation with further childbearing. This mirrors the findings from Leone et al. (2003), albeit using a different empirical approach. An important covariate, years of education, shows that each additional year of schooling reduces the likelihood of a third child by 2 percentage points. Column (3) reports statistically significant heterogeneities, with urban women as large as 8 percentage points less likely to have a third child. However, among women with two daughters, urban women are shown to be 7 percentage points more likely to further childbearing than their rural counterparts.

NDHS 2001

Table 5 reports the results linking sex mix and fertiltiy for 2001 DHS iteration. Mirroring the findings from 1996, the estimates show that women with two daughters are 5 percentage points more likely to have a third child (from Column 2), while those with two boys are 4.7 percentage points less likely to further childbearing. Education continues to exert a dampening effect on women's propensity for additional childbirth. Similarly, urban women are 8 percentage points less likely to have a third child than rural women.

A notable addition is that of Muslim women. When compared to Hindu women who had two daughters, Muslim women are estimated to be 10 percentage points less likely to bear a third child, after adjusting for covariates. However, it's important to note the limited sample size of Muslim women in the 2001 NDHS dataset (355), which raises concerns about the reliability of this estimate. I hypothesize that gender-based fertility preferences might be obscured for Muslim women due to their already high fertility rates, whereas such preferences are more evident among Hindu women, who typically have fewer children on average (with mean children ever born estimates of 3.26 for Hindu women and 3.9 for Muslim women)

NDHS 2006

Table 6 reports the results linking sex-mix and fertiltiy for 2006 DHS iteration. OLS estimates show that women with two daughters are as high as 10 percentage points more likely to have a third child, while the effect of two sons on further child-bearing, is negative. Regarding heterogenieties, estimates show that Muslim women are 7 percentage points more likely to have a third child (in general) that Hindu women. Wealth, measured by wealth index quantiles, inversely affects the propensity for a third child across middle to richest quantiles. Similarly, increased schooling decreases the likelihood by 2-3 percentage points. However, compared to the poorest quantile, the richest quantile of women, if they have two daughters, exhibit 11 percentage points increased likelihood to have a third child.

NDHS 2011

The 2011 NDHS findings mirror past trends: having two daughters increases the likelihood of a third child by 11 percentage points, while having two sons decreases it by nearly 7 percentage points. Each additional year of schooling reduces the likelihood by 3 percentage points. Muslim women are 14 percentage points less likely than Hindu women to have a third child if they have two daughters. Wealth plays a significant role: compared to the poorest quantile, poorer to richest quantiles are 9-20 percentage points more likely to have a third child if their first two children were girls. As evident by the statistical insignificance of two girls estimate in column (3), the effect of two daughters on further childbearing is entirely explained by wealth index quantiles.

NDHS 2016

Table 8 from the 2016 NDHS reveals that women with two daughters are 16 percentage points more likely to have a third child, a higher estimate compared to previous iterations. Urban women with two girls are 7 percentage points more likely to have a third child than rural women, while Muslim women are 13 percentage points less likely than Hindu women. Consistent with previous years, rich and richest wealth index quantiles show an 8 and 11 percentage point increase in likelihood of further childbearing after two girls.

This is different from the findings reported by Sapkota et al. (2019), where urban status and wealth is shown to have deterred son-preference, while Muslim women were shown to have higher son preferences. The differences in results may be attributed to two important considerations : a) Sapkota et al. (2019) look at *desired* gender based fertiltiy preferences whereas

this paper looks at fertility *ex post*, b) this paper uses interaction terms in the regression to extract difference-in-difference estimates, which differs in interpretation from standard dummy-variable usage.

NDHS 2022

In the latest iteration of NDHS in 2022, Table 9 highlights that women with two daughters are nearly 19 percentage points more likely to have a third child, marking the largest effect observed across six DHS iterations. Education remains a deterrent to fertility. Furthermore, the three richest wealth index quantiles exhibit a higher propensity for further childbearing compared to the poorest quantile.

Table 4: OLS Estimates for More Than Two Children equation for 1996 NDHS iteration

	(1)	(2)	(3)
Two Girls	0.043***	0.033***	0.021
	(0.012)	(0.013)	(0.014)
Two Boys		-0.028**	-0.028**
		(0.012)	(0.012)
Boy Second		-0.006	-0.006
		(0.012)	(0.012)
Years of education		-0.021***	-0.018***
		(0.002)	(0.002)
Urban			-0.079***
			(0.016)
TwoGirls X Urban			0.074**
			(0.033)
Observations	6323	6323	6323
Adjusted R-squared	0.002	0.278	0.281
Covariates	No	Yes	Yes
Dummy and Interaction	No	No	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: The dependent variable is *MoreThanTwo*. The covariates are mother's age, years of education, mother's age at first birth, mother's age at second birth. Column (3) includes dummies such as Brahmin/Chhetri, religion (Hindu, Muslim, Buddhist, Christian and others), and urban/rural status and the interaction of *TwoGirls* and religion, urban/rural status.

Table 5: OLS Estimates for More Than Two Children equation for 2001 NDHS iteration

	(1)	(2)	(3)
Two Girls	0.065***	0.053***	0.052***
	(0.012)	(0.012)	(0.014)
Two Boys		-0.047***	-0.047***
		(0.012)	(0.012)
Boy Second		-0.001	-0.002
		(0.012)	(0.012)
Years of education		-0.028***	-0.025***
		(0.002)	(0.002)
Urban			-0.082***
			(0.015)
TwoGirls X Muslim			-0.101**
			(0.050)
Observations	6451	6451	6451
Adjusted R-squared	0.004	0.307	0.311
Covariates	No	Yes	Yes
Dummy and Interaction	No	No	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: The dependent variable is *MoreThanTwo*. The covariates are those specified in Section 3. Column (3) includes dummies such as Brahmin/Chhetri, religion (Hindu, Muslim, Buddhist, Christian and others), and urban/rural status and the interaction of *TwoGirls* and religion, urban/rural status.

Table 6: OLS Estimates for More Than Two Children equation for 2006 NDHS iteration

	(1)	(2)	(3)
Two Girls	0.132***	0.108***	0.073***
	(0.013)	(0.013)	(0.024)
Two Boys		-0.088***	-0.087***
		(0.013)	(0.013)
Boy Second		0.016	0.017
		(0.013)	(0.013)
Years of education		-0.030***	-0.022***
		(0.001)	(0.002)
Muslim			0.071**
			(0.030)
Wealth Index (middle income)			-0.041**
			(0.016)
Wealth Index (Rich)			-0.070***
			(0.017)
Wealth Index (Richest)			-0.166***
			(0.019)
TwoGirls X Richest (Wealth Index)			0.117***
			(0.036)
Observations	6436	6436	6436
Adjusted R-squared	0.016	0.341	0.350
Covariates	No	Yes	Yes
Dummy and Interaction	No	No	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The covariates and dummies are those specified in Section 3. Additional inclusion is that of wealth-index quantiles and its interaction term.

Table 7: OLS Estimates for More Than Two Children equation for 2011 NDHS iteration

	(1)	(2)	(3)
Two Girls	0.180***	0.114***	0.013
	(0.013)	(0.013)	(0.024)
Two Boys		-0.068***	-0.071***
		(0.012)	(0.012)
Boy Second		-0.040***	-0.036***
		(0.013)	(0.013)
Years of education		-0.032***	-0.019***
		(0.001)	(0.002)
Two Girls X Muslim			-0.149**
			(0.065)
Two Girls X Wealth Index (poorer)			0.093***
			(0.032)
Two Girls X Wealth Index (middle income)			0.121***
			(0.033)
Two Girls X Wealth Index (rich)			0.133***
			(0.034)
TwoGirls X Wealth Index (richest)			0.202***
			(0.036)
Observations	7106	7106	7106
Adjusted R-squared	0.025	0.361	0.390
Covariates	No	Yes	Yes
Dummy and Interaction	No	No	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The covariates and dummies are those specified in Section 3. from 2006 onwards, additional inclusion is that of wealth-index quantiles and its interaction term.

Table 8: OLS Estimates for More Than Two Children equation for 2016 NDHS iteration

	(1)	(2)	(3)
Two Girls	0.200***	0.165***	0.086***
	(0.013)	(0.013)	(0.025)
Two Boys		-0.062***	-0.063***
		(0.013)	(0.012)
Boy Second		-0.011	-0.009
		(0.013)	(0.012)
Years of education		-0.034***	-0.025***
		(0.001)	(0.001)
Two Girls X Muslim			-0.138***
			(0.050)
Two Girls X Urban			0.070***
			(0.023)
Two Girls X Wealth Index (rich)			0.079**
			(0.034)
TwoGirls X Wealth Index (richest)			0.111***
			(0.037)
Observations	7169	7169	7169
Adjusted R-squared	0.030	0.368	0.392
Covariates	No	Yes	Yes
Dummy and Interaction	No	No	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: The dependent variable is *MoreThanTwo*. The covariates are those specified in Section 3. Column (3) includes dummies such as Brahmin/Chhetri, religion (Hindu, Muslim, Buddhist, Christian and others), urban/rural status and wealth index and the interaction of *TwoGirls* and religion, urban/rural status, and wealth index.

Table 9: OLS Estimates for More Than Two Children equation for 2022 NDHS iteration

	(1)	(2)	(3)
Two Girls	0.245***	0.188***	0.115***
	(0.012)	(0.013)	(0.021)
Two Boys		-0.062***	-0.063***
		(0.012)	(0.012)
Boy Second		-0.028**	-0.025**
		(0.012)	(0.012)
Years of education		-0.030***	-0.021***
		(0.001)	(0.001)
Two Girls X Muslim			-0.243***
			(0.055)
Two Girls X Wealth Index (middle income)			0.130***
			(0.029)
Two Girls X Wealth Index (rich)			0.142***
			(0.031)
TwoGirls X Wealth Index (richest)			0.231***
			(0.037)
Observations	8162	8162	8162
Adjusted R-squared	0.045	0.362	0.388
Covariates	No	Yes	Yes
Dummy and Interaction	No	No	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: The dependent variable is *MoreThanTwo*. The covariates are those specified in Section 3. Column (3) includes dummies such as Brahmin/Chhetri, religion (Hindu, Muslim, Buddhist, Christian and others), urban/rural status and wealth index and the interaction of *TwoGirls* and religion, urban/rural status, and wealth index.

Summary and Discussion

In accordance to our research question, the estimates on women's proclivity to have more than two children by sex composition of the first two children, for all six NDHS iterations, show a persistent boy-bias. Women with two daughters are estimated to be more likely to have a third child than those with two sons. This finding is similar to the boy-bias reported by Lee (2008) in Korea, but is in contrast to those exhibited by Moschion (2013); Iacovou (2001); Cruces and Galiani (2007) in Australia, UK and Latin America respectively.

The estimates for our regression of *TwoGirls* on *MoreThanTwo* exhibit an increasing trend from 1996 to 2022, with the highest estimate observed in the latest iteration. However, these differences might be overestimated due to two reasons: a) we only look at parity progression from 2 to more than 2, neglecting other transitions and b) the higher fertility rate in 1996 might have masked gender-based fertility preferences i.e. families expecting to have, for example, eight children, would not be worried that their first two are girls, even if they have a son preference.

Considering the estimation that wealthier households exhibited stronger son preferences, my hypothesis is that in a quantity-quality tradeoff paradigm (Becker and Lewis, 1973), the realization of underlying tastes (preference for boys) may have been made possible by decrease in the real price of child-bearing (wealth). For instance, lets take two couples with identical tastes: both prefer sons, but possess differential wealth: couple A is wealthier than couple B. If both couple A and B had two daughters, and if the price of raising children is exogenous and market-determined, couple A would have the financial means necessary to attempt a third child than couple B. However, we must bear in mind that a household's preference of sons is not always exclusively a matter of taste but also of expected future returns and old age support, a custom prevalent in South Asia (Monica Das Gupta and Hwa-Ok, 2003).

Limitations and Considerations

The paper aims to establish a causal relationship between sex composition and subsequent fertility. However, there are three notable limitations that warrant discussion:

1. Sex ratios at birth are assumed to be representative of the natural human sex-ratio at birth (105 boys born 100 girls (Orzack et al., 2015)). A study on Conditional Sex Ratio (CSR) for second-born children where the first-born was female reports that after legalization of abortion in 2002 A.D. in Nepal, sex-selective abortion became more common most among educated, richer and urban women (Frost et al., 2013). This mirrors our findings

of higher boy-biasedness among urban and wealthier groups. Similarly, another nationally representative study, estimates that CSR was significantly higher than biological Sex Ratio at Birth (SRB) and purports sex-selective abortion as the dominant cause (Pradhan et al., 2019). Therefore, consideration of the impact of sex-selective abortion on the *TwoGirls* estimate presents an avenue for future investigation.

2. While the paper operates under the assumption of random sex distribution at birth, there exists evidence suggesting variability in the probability of male births within the population (Ben-Porath and Welch, 1976). Moreover, multiparous women (women who have given birth more than once) are more likely to have prenatal knowledge of the sex of their offspring (Pradhan et al., 2019).

3. Although the NDHS includes ethnicity indicators, their consistency across all surveys varies, which is why the paper refrains from using ethnicity as a dummy variable. Incorporating considerations of sex preferences within a multi-ethnic context could offer a more comprehensive understanding of gender-based fertility preferences, particularly given the diverse fertility associated with different ethnicities.

Conclusion

The paper sheds light on the prevalence of son preference in fertility dynamics in Nepal. We found that families with two daughters are more inclined to have additional children, indicating a persistent preference for sons. We also found that wealthier and urban women exhibit stronger preference for sons compared to poorer and rural women, espcially after the turn of the millenium. The incorporation and statistical significance of wealth, religion and urban status heterogneities underscores a need to consider socio-economic and cultural factors in understanding these preferences. Moving forward, it's crucial to look deeper into the impact of sex-selective abortion and explore the diverse influences of ethnicity on fertility preferences.

References

Angrist, J. D. and Evans, W. N. (1998). Children and their parents' labor supply: Evidence from exogenous variation in family size. *The American Economic Review*, 88(3):450–477.

Barcellos, S. H., Carvalho, L. S., and Lleras-Muney, A. (2014). Child gender and parental investments in india: Are boys and girls treated differently? *American Economic Journal: Applied Economics*, 6(1):157–189.

Becker, G. S. and Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of political Economy*, 81(2, Part 2):S279–S288.

Ben-Porath, Y. and Welch, F. (1976). Do sex preferences really matter? *The Quarterly Journal of Economics*, 90(2):285–307.

Brunson, J. (2010). Son preference in the context of fertility decline: limits to new constructions of gender and kinship in nepal. *Studies in family planning*, 41(2):89–98.

Cruces, G. and Galiani, S. (2007). Fertility and female labor supply in latin america: New causal evidence. *Labour Economics*, 14(3):565–573.

Das, N. (1987). Sex preference and fertility behavior: A study of recent indian data. *Demography*, 24(4):517–530.

Filmer, D., Friedman, J., and Schady, N. (2008). Development, modernization, and son preference in fertility decisions. *World Bank Policy Research Working Paper*, (4716).

Frost, M. D., Puri, M., and Hinde, P. R. A. (2013). Falling sex ratios and emerging evidence of sex-selective abortion in nepal: evidence from nationally representative survey data. *BMJ open*, 3(5):e002612.

Graham, M. J., Larsen, U., and Xu, X. (1998). Son preference in anhui province, china. *International Family Planning Perspectives*, 24(2):72–77.

Gupta, M. D. (1996). Life course perspectives on women's autonomy and health outcomes. *Health Transition Review*, pages 213–231.

Haughton, J. and Haughton, D. (1999). Son preference. *Health and Wealth in Vietnam: An Analysis of Household Living Standards*, pages 97–199.

Iacovou, M. (2001). Fertility and female labour supply. Technical report, ISER Working Paper Series.

Karki, Y. B. (1988). Sex preference and the value of sons and daughters in nepal. *Studies in Family Planning*, 19(3):169–178.

Lee, J. (2008). Sibling size and investment in childrenâs education: An asian instrument. *Journal of Population Economics*, 21:855–875.

Leone, T., Matthews, Z., and Zuanna, G. D. (2003). Impact and determinants of sex preference in nepal. *International family planning perspectives*, pages 69–75.

Milazzo, A. (2014). Son preference, fertility and family structure: Evidence from reproductive behavior among nigerian women. *World Bank Policy Research Working Paper*, (6869).

Monica Das Gupta, Jiang Zhenghua, L. B. X. Z. W. C. and Hwa-Ok, B. (2003). Why is son preference so persistent in east and south asia? a cross-country study of china, india and the republic of korea. *The Journal of Development Studies*, 40(2):153–187.

Moschion, J. (2013). The impact of fertility on mothers' labour supply in australia: Evidence from exogenous variation in family size. *Economic Record*, 89(286):319–338.

Orzack, S. H., Stubblefield, J. W., Akmaev, V. R., Colls, P., Munné, S., Scholl, T., Steinsaltz, D., and Zuckerman, J. E. (2015). The human sex ratio from conception to birth. *Proceedings of the National Academy of Sciences*, 112(16):E2102–E2111.

Pande, R. P. and Astone, N. M. (2007). Explaining son preference in rural india: the independent role of structural versus individual factors. *Population Research and Policy Review*, 26:1–29.

Pradhan, E., Pearson, E., Puri, M., Maharjan, M., Maharjan, D. C., and Shah, I. (2019). Determinants of imbalanced sex ratio at birth in nepal: evidence from secondary analysis of a large hospital-based study and nationally-representative survey data. *BMJ open*, 9(1):e023021.

Sapkota, J. B., Paudel, D. B., Neupane, P., and Thapa, R. B. (2019). Preference for sex of children among women in nepal. *Global Social Welfare*, 6:69–78.

Stash, S. (1996). Ideal-family-size and sex-composition preferences among wives and husbands in nepal. *Studies in Family Planning*, pages 107–118.

UN (2022). World population prospects 2022.

UNESCO (2023).

Williamson, N. E. (1983). Parental sex preferences and sex selection.

Zheng, Z., Gu, B., and Gietel-Basten, S. (2018). Fertility preferences in china. *Family Demography in Asia: A Comparative Analysis of Fertility Preferences*, pages 109–120.