

# Son Preference and Gender-Biased Breastfeeding in Pakistan

NAIMA HAFEEZ

Lahore University of Management Sciences

CLIMENT QUINTANA-DOMEQUE

University of Oxford

## I. Introduction

Breastfeeding is one of the most valuable investments made in child survival and nutrition. Prevalence estimates from the United Nations suggest that sub-optimal breastfeeding is associated with approximately 804,000 deaths annually, accounting for 11.6% of total deaths of children younger than 5 years (Black et al. 2013). Despite well-documented evidence on the benefits of breastfeeding, some children continue to be weaned early. Jayachandran and Kuziemko (2011; hereafter JK), using data from the National Family and Health Survey (NFHS) for India, found that mothers breastfeed daughters significantly less (especially if there are no elder sons) to increase the chances of becoming pregnant again with the hope to conceive a son. The contraceptive function of breastfeeding has been previously documented in the medical literature, where it has been shown that continued breastfeeding reduces the risk of future pregnancy to less than 2% in the first 6 months after delivery (Kennedy et al. 1989). The findings by JK have recently been replicated in a very different context—Egypt—by Chakravarty (2015).

In this article, we continue this line of research focusing on Pakistan, which offers an interesting case to examine son-biased preferences and complement previous studies, if only because parents' stated preference for sons is twice as large in Pakistan as it is in India: sons are preferred to daughters 10:1 in Pakistan and approximately 5:1 in India (Arnold 1997). Similar to India, there are strong cultural and social norms that place a higher value on the birth of sons. However, unlike India, Pakistan is almost entirely religiously homogeneous, with a 96.4% Muslim population (Pew Research Center 2011). Previous

We thank Elisabetta De Cao, Sonia Orefice, Simon Quinn, and two anonymous reviewers for their helpful comments and suggestions. Hafeez acknowledges financial support from the Vicky Noon Educational Foundation Oxford Scholarship. The usual disclaimers apply. Contact the corresponding author, Climent Quintana-Domeque, at [climent.quintana-domeque@economics.ox.ac.uk](mailto:climent.quintana-domeque@economics.ox.ac.uk).

Electronically published December 7, 2017  
© 2017 by The University of Chicago. All rights reserved. 0013-0079/2018/6602-0001\$10.00

studies have found that Muslim-majority countries tend to be reluctant to using artificial contraceptives (Morgan et al. 2002; Carranza 2012). In this regard, Pakistan, like Egypt, may provide more leverage to parents to realize their preference for future sons through natural methods of contraception.

Using data from the 2006–07 Pakistan Demographic and Health Survey (PDHS; NIPS 2008), we replicate the analysis of JK in India, confirming their findings: the duration of breastfeeding increases with the birth order of the child, and children born to mothers after having attained their reported ideal fertility size are breastfed increasingly longer. That is, as the family size expands, the need to use breastfeeding as a contraceptive method correspondingly increases. Moreover, we find that boys are breastfed longer than girls at every birth order and that as the share of elder male siblings increases from 0 to 1, the probability of weaning a child reduces. That is, a child is breastfed increasingly longer as the mother's preference for sons satiates. We also find that the gender gap in breastfeeding rises discontinuously as a mother attains her reported ideal fertility size. This is because closer to attaining the ideal family size, the decision to have another child is most sensitive to the gender preferences of the mother, resulting in a pronounced increase in the male-female difference in breastfeeding, which is significantly larger in our sample than that found in JK. Further, there is a noteworthy increase in the duration of breastfeeding when a mother attains her ideal number of sons, which is separate from the effect of attaining the overall ideal fertility.

Adding to JK's framework, we also examine whether there is a change in the duration of nursing on the attainment of the ideal number of daughters, but we fail to find any. Therefore, selectively increasing breastfeeding duration in response to bearing the ideal number of sons but not daughters further strengthens the findings in JK: breastfeeding is used as a contraceptive to exercise gender-differentiated birth spacing. While most of the existing literature in Pakistan analyzes the gender gap in human capital investment among existing children, we are the first to show how a future son leads a mother to selectively discriminate against her current child by gender in Pakistan.

After testing all the predictions in JK, we reestimate all the main specifications using an alternate outcome variable, namely, the mother's subjective assessment of the unwantedness of a pregnancy. We find patterns similar to those distinctly observed in the duration of breastfeeding. Not only does this provide a robustness check for the link between breastfeeding and subsequent fertility, but more importantly, it validates the role of stated preference-based measures (e.g., unwantedness of a pregnancy) as complements to revealed preference-based measures (e.g., duration of breastfeeding) or even substitutes in case the former are available but the latter are not.

The remainder of this article is organized as follows. Section II provides background context and the usefulness of breastfeeding as an informative metric to study son preference in Pakistan. Section III explains the empirical methodology. Section IV describes the data and sample selection. Section V presents the empirical results. Section VI links breastfeeding with the mother's subjective assessment of the unwantedness of her last pregnancy. Section VII concludes the article.

## II. Son Preferences around the World

### A. India, Egypt, and Pakistan

Empirical evidence from a comparative study of 44 countries using Demographic and Health Surveys (DHS; Arnold 1997) indicates that preference for sons remains particularly strong in South Asia: the average number of sons in the reported optimal family size exceeds the average number of daughters by 49% in India, 43% in Pakistan, and 26% in Bangladesh. People from India and Pakistan had lived together in the Indian subcontinent for centuries until partitioned into the sovereign states of Pakistan (later split into the Islamic Republic of Pakistan and Bangladesh) and India in 1947. A complex interplay of common socioeconomic and cultural forces has reinforced patriarchal norms that place a higher value on the birth of sons than daughters (Sathar and Kazi 2000; Gupta et al. 2003; Pande and Astone 2007; Almond, Edlund, and Milligan 2013): first, elderly parents in India and Pakistan heavily depend on children in old age and predominantly depend on males for financial and emotional support since girls marry into another family. Second, the system of patrilineal descent ensures that family assets are mostly passed on to sons and only they carry forward the family name. On the other hand, girls have to leave their parents after marriage and join the husband's family. Third, girls are often viewed as an economic burden, as parents struggle to meet rising dowry demands put forward by their in-laws for the marriage to take place.<sup>1</sup> Fourth, the desire for sons is further perpetuated by the economy of both countries, which is largely agrarian, hence favoring a family composition with multiple sons who are considered more able to reap the fruits of physical labor.<sup>2</sup>

<sup>1</sup> Recent trends show that the practice of dowry has become even more prevalent in India despite an increase in modernization (Srinivasan and Lee 2004). Anderson (2000) finds that dowry continues to prevail in Pakistan for reasons similar to those in India.

<sup>2</sup> However, the DHS reports a significant percentage of women engaged in agriculture in India (59%) and Pakistan (42%). While this may appear to be conflicting evidence, it is worth noting that in Pakistan, 24% of these women do not receive any remuneration compared to only 5% of women working in the nonagricultural sector. In India, 35% of working women are not paid in agriculture, which is more than twice as high as the percentage of men not paid in agriculture (14%). This is not surprising since 58% of the women engaged in agriculture in India are employed by family members

Preference for sons is also linked to the overall low level of female autonomy. Women in India and Pakistan generally have a low status in terms of education and literacy (55.1% of women are literate in India, while only 36% are in Pakistan; see table 1) and especially with regard to employment opportunities (36.3% are currently employed in India vs. 25.9% in Pakistan; see table 1). Low labor force participation means women bring less income to their families.<sup>3</sup> In an environment that deters women's economic empowerment otherwise, mothers seek to regain their standing in the society by the number of sons they bear.<sup>4</sup>

India and Pakistan share similar characteristics, but they are also different along many dimensions. While Pakistan is almost entirely a Muslim community, India is less monolithic in terms of the various religious communities it caters to. In this regard, it is important to highlight the replication of JK by Chakravarty (2015) using data from Egypt, a country culturally very different from India that shares similarities with Pakistan in that both have a Muslim-majority population. Chakravarty (2015) finds results similar to those by JK.

High fertility rates are a widespread phenomenon in Islamic states. The third row in table 1 shows that the total fertility rates were 4.1 in Pakistan (2006–07) and 2.7 in India (2005–06). In India, while Islam has conventionally been interpreted as not permitting birth control (Iyer 2002), the overall stance on the matter varies among different schools of thought (Obermeyer 1992). On the other hand, Hinduism regards the use of contraceptives as a personal decision outside the injunction of religion (Iyer 2002). Perhaps part of this difference is reflected in the fourth row of table 1: the percentage of married women currently using any method of family planning is 29.6% in Pakistan and 56.3% in India. Many demographic studies have found that preference for sons increases national fertility by roughly 10% (Clark 2000), since couples with daughters continue to have more children until they achieve their desired number of sons. Therefore, the observed difference in fertility rates between the two countries could possibly allude to a stronger preference for sons in Pakistan.<sup>5</sup>

---

compared to only 25% of those involved in nonagricultural work. Thus, women seem to only be assisting men in agriculture without bringing much income of their own to the household, in which case sons would continue to be preferred.

<sup>3</sup> Aslam (2009) finds that in Pakistan, total labor market returns are much higher for men despite returns to education being higher for women.

<sup>4</sup> For example, in northern India, which is culturally akin to Pakistan, a woman is called "X's mother." Gupta et al. (2003, 172) mention that "if the first [born] is not a son she will be renamed as soon as she bears a son, because being called a boy's mother is much more prestigious than a girl's mother."

<sup>5</sup> Comparing sex ratios by religion, the Census of India (2011) reveals that Hindus in India have a more skewed ratio of 931 females per 1,000 males compared to 936 among Muslims. At the outset,

**TABLE 1**  
**SUMMARY STATISTICS ON INDIA, EGYPT, AND PAKISTAN**

	India 2005–06	Egypt 2005	Pakistan 2006–07
Percentage of women who are literate	55.1	59.1	36
Percentage of ever-married women (age 15–49)			
currently employed	36.3	21.5	25.9
Total fertility rate (children per women)	2.7	3.1	4.1
Percentage of married women currently using			
any method of family planning	56.3	59.2	29.6
Median duration of breastfeeding among			
female (male) children born in the last 3 years	23.6 (25.4)	18 (19.2)	18.2 (19.5)
Ratio of female to male labor force participation			
rate (%) in 2006	42.4	27.9	24.7
Female (male) life expectancy at birth in years in 2006	65.8 (64.0)	72.0 (67.3)	65.0 (63.2)
Female (male) expected years of schooling in 2009	10.5 (11.0)	12.5 (13.1)	6.6 (7.9)

**Sources.** Bhat et al. (2007), El-Zanaty and Way (2006), NIPS (2008), and the World Bank's DataBank (<http://databank.worldbank.org>).

**Note.** Statistics are drawn from Demographic and Health Survey (DHS) reports based on the 2005–06 National Family Health Survey for India, the 2005 DHS for Egypt, and the 2006–07 DHS for Pakistan. Results are reported for the most recent year for which the surveys in the three countries were conducted roughly around the same time.

Furthermore, the Islamic inheritance law favors sons over daughters in that “only a male descendant (son or son's son) can exclude male ascendants (paternal grandfather, uncles, brothers, or nephews) from the distribution of wealth” (Carranza 2012, 6). Daughters, on the other hand, do not have the legal authority to exclude these other heirs. This rule provides a strong economic incentive to value sons more than daughters in Muslim families. While traditionally, Hindu inheritance norms also immensely favored sons over daughters (perhaps even more so), an amendment to the Hindu Succession Act was imposed in 2005 that grants daughters equal rights in ancestral property (irrespective of their date of birth). Deininger, Goyal, and Nagarajan (2010) found a positive impact of this legislation on women's physical and human capital investment: postamendment daughters are more likely to inherit land and also attain a significantly higher level of education.

Table 1 also shows that, while the overall median breastfeeding duration is higher in India, the male-female breastfeeding gap is similar in all three coun-

---

this may be taken as evidence favoring weaker son preference among Muslims. Clark (2000) shows that smaller families in India have a significantly higher fraction of sons than large families. Thus, a more balanced sex ratio in large Muslim families could actually be driven by couples (with mainly daughters) in their persistent efforts to have sons. Table 1 also shows that the ratio of female to male labor force participation rate is very different between India (42.4%) and Pakistan (24.7%). Regarding life expectancy at birth, we find similar female-male gaps in India and Pakistan (1.8 years), while the gap is much higher in Egypt (4.7 years). Finally, in terms of years of schooling, the expected gap between men and women is similar in India and Egypt (around half a year), while this difference is larger in Pakistan (where men are expected to accumulate 1.3 more years of education than women).

tries (1.8 in India, 1.3 in Pakistan, and 1.2 in Egypt). Further comparing statistics from the 2006–07 PDHS and the 2005–06 NFHS, several observations can be made: (1) while there is a drastic difference in the percentage of married women who have ever used any modern method of contraception in India and Pakistan (57.6% vs. 38.8%, respectively), the proportion who have used traditional methods is in fact higher in Pakistan (25.5% vs. 19.4% in India); (2) an analysis of contraceptive use by religion in India reveals that Muslims report the lowest percentage use of modern contraceptive methods (36.4% vs. 50.2% among Hindus) but the highest use of traditional methods among all religious groups in India; and (3) the single most common reason for not intending to use contraception in Pakistan is “up to God” (28.4%), a response completely absent in the Indian survey, whereas religious prohibition as a separate reason is cited by only 5% of women in both the samples. The cross-country comparison of the impact of religious perceptions on contraceptive use should be assessed by adding these percentages (33.4% in Pakistan vs. 5% in India) rather than treating religious prohibition as a response completely independent from “up to God.” Overall, these statistics illustrate that using breastfeeding as a contraceptive method (to possibly exercise gender-differentiated birth spacing) is particularly relevant in the context of Pakistan, where the use of natural methods of contraception and the involvement of religion in family planning are more common relative to India.

Apart from differences in religion, India and Pakistan greatly differ in terms of regional heterogeneity. Jejeebhoy and Sathar (2001) argue that region, instead of religion, plays the major conditioning role by showing that Hindu and Muslim women exercise the same autonomy once region is controlled for. There is a vast literature that highlights the differences in kinship patterns between northern and southern India and their varying effects on female autonomy (Dyson and Moore 1983; Jejeebhoy and Sathar 2001; Pande and Astone 2007). In South India, endogamous marriages (i.e., marriage between cousins or within a particular ethnic group) are much more common, which give females a stronger voice in family life.<sup>6</sup> Furthermore, it is not uncommon for daughters to provide financial support to their parents, and the dowry system is relatively less oppressive (Pande and Astone 2007). This regional variation in social practices helps somewhat mitigate the overall extent of son preference in

<sup>6</sup> *Watta sata* (the practice of bride exchange in Pakistan) is linked with increased welfare of women in marriage (Jacoby and Mansuri 2010). Cousin marriages are also extremely common in Pakistan (approx. two-thirds of marriages are between cousins) primarily due to their sociocultural benefits, but they are exploitative in nature (Hussain 1999). The evidence on the impact of consanguineous marriage on women's welfare is mixed.

India. JK find that the gender bias in breastfeeding is more pronounced in the North Indian states of Haryana and Punjab, where son preference is stronger. Culturally, Pakistan's eastern region (comprising the bulk of its population) is very similar to its geographical neighbor encompassing North and West India (Murtaza 2014). Thus, estimates of son preference from Pakistan are likely to be similar in magnitude to those from North India and therefore more pronounced than holistic estimates from India. In fact, this is what we find in our empirical results.

### ***B. Absolute vs. Conditional Son Preference in Pakistan***

Most of the work in the existing literature focuses on absolute son preference in Pakistan, whereby male outcomes are generally compared against female outcomes without controlling for the number of previous births or sex composition of elder children. More realistically, however, the extent of son preference is conditional on factors such as the fraction of elder children that are male and the attainment of the mother's ideal fertility size. JK control for both of these factors in their analysis and therefore report estimates on conditional son preference rather than absolute. To the best of our knowledge, there is only one study, Hussain, Fikree, and Berendes (2000), that studies conditional son preference in Pakistan by evaluating the effect of children's sex composition on the mother's subsequent fertility preference. They find that as the number of surviving sons increases, future pregnancies become more and more undesirable. Further, results suggest a nonexclusive form of son preference: while the overall preference for sons remains stronger, parents express the desire to have at least one daughter.

Although the work by Hussain et al. (2000) attempts to fill a gap in the literature by quantifying son preference before conception and evaluating the conditional rather than the absolute preference for sons, it is susceptible to several limitations. First, it is based on data from just one city of Pakistan and is therefore very limited in scope. Second, the survey rounds for their study were conducted long ago in 1990–91 and 1995, so there is a need to reevaluate the recent trends in data. Third, it presumes that son preference is conditional on only a single variable—children's sex composition. Following the work by JK, we also condition on the attainment of mother's ideal fertility, since the behavior of a mother toward her children, before and after the realization of her ideal family size, is not expected to be the same. Finally, for the sake of correct policy advice, it is crucial to understand the mechanism behind the male breastfeeding advantage in Pakistan. JK point out that increasing the accessibility of other contraceptives can both increase and decrease breastfeeding. However, given the factors governing low use of modern contraceptive methods in Pakistan,



the elasticity of breastfeeding duration with respect to the accessibility of alternate birth control methods is not expected to be high.

### III. Empirical Strategy

Following JK, we first estimate the following regression:

$$\text{Duration}_i = \alpha + \beta \text{Birth Order}_i + \gamma \text{Male}_i + \delta X_i + a_i + \varepsilon_i, \quad (1)$$

where the dependent variable  $\text{Duration}_i$  is the number of months child  $i$  is breastfed,  $\text{Birth Order}_i$  is the order of birth of child  $i$ ,  $\text{Male}_i$  is an indicator of the child  $i$  being a boy,  $X_i$  is a vector of controls, and  $a_i$  is child age (in months) fixed effects (from 0 to 36 months to control for censoring). The vector  $X_i$  includes the following covariates: linear and quadratic controls for the mother's and child's age,<sup>7</sup> type of place of residence (rural/urban), region of residence (dummies for provinces of Pakistan), sex of child, and dummies for the mother's level of education. Regression (1) allows us to test the following predictions:

1. *Breastfeeding duration increases with the birth order of the child.* As the number of previous births increases, the need for contraception in the form of breastfeeding increases. This effect is captured by the coefficient  $\beta$ .
2. *At any birth order, boys are breastfed longer than girls.* The contraceptive function of breastfeeding allows the mother to exercise gender-differentiated birth spacing to realize her preference for a future son. This effect is captured by the coefficient  $\gamma$ .

As emphasized by JK, finding a positive male advantage in breastfeeding in itself does not elucidate on why a mother breastfeeds her sons longer. In particular, failure to reject this prediction is not sufficient to establish that the gender breastfeeding gap is generated by exploiting the contraceptive nature of breastfeeding to realize preference for a future son. One alternate explanation that is vastly quoted in literature centers on the mother allocating more resources to her sons out of a stronger sense of care for their well-being. Following JK, we also estimate the following regression:

<sup>7</sup> JK use linear and quadratic controls for the child's birth year. The square of birth year contains very large integer values, so it is neater to instead use the variable on the age of the child (in years). This also makes the interpretation of the intercept term in our regressions more appealing. Note that since the age of the child goes from 0 to 48 months (from 0 to 4 years), we can control for age-in-month fixed effects from 0 to 36 also including the age of the child in years plus the square of the age of the child, without incurring in perfect multicollinearity.



$$\text{Duration}_i = \alpha + \beta \text{Birth Order}_i + \gamma \text{Male}_i + \rho \text{Male Share}_i + \delta X_i + a_i + \varepsilon_i, \quad (2)$$

where  $\text{Male Share}_i$  is the proportion of elder male siblings. Regression (2) allows us to test the following prediction:

3. *At any birth order, a child with a higher proportion of elder male siblings is breastfed longer.* Son preference is attenuated if a mother already has a high fraction of male children. Therefore, breastfeeding duration correspondingly increases as future pregnancy becomes unwanted in response to achieving a higher share of male children. This effect is captured by the coefficient  $\rho$ .

Recently, Chakravarty (2015) also estimated the following regression:<sup>8</sup>

$$\text{Duration}_i = \alpha + \beta \text{Birth Order}_i + \gamma \text{Male}_i + \rho \text{Male Share}_i + \sigma \text{Male}_i \times \text{Male Share}_i + \delta X_i + a_i + \varepsilon_i. \quad (3)$$

If the proportion of elder siblings that are male increases from 0 to 1, then the associated increase in breastfeeding duration is  $\rho$  months, comparing children across the same birth order and gender. Here,  $\sigma$  captures the male-female difference in breastfeeding duration as the fraction of elder male siblings increases from 0 to 1. If  $\sigma$  is negative and significant, we infer that a girl with all elder siblings that are male is breastfed longer than a boy at her place. Regression (3) allows us to test the following prediction:

4. *As the proportion of elder male siblings increases from 0 to 1, girls are breastfed longer than boys.* Failure to reject this hypothesis indicates that not all girls are equally disadvantaged in terms of breast milk intake. Further, it highlights the fact that mothers express the desire to achieve some balance in their children's sex composition.

In addition, we estimate the following regression:

$$\text{Duration}_i = \alpha + \gamma \text{Male}_i + \tau 1(\Delta \text{Ideal}_i \geq 0) + \varphi \Delta \text{Ideal}_i + \omega 1(\Delta \text{Ideal}_i \geq 0) \times \Delta \text{Ideal}_i + \delta X_i + a_i + \varepsilon_i, \quad (4)$$

<sup>8</sup> Chakravarty (2015) also uses exclusive breastfeeding as an alternate outcome. While generally DHS reports information on food and liquid intake by the baby, this is missing for our sample (NIPS 2008). Hence, we cannot replicate table IV in Chakravarty (2015).

where the variable  $\Delta\text{Ideal}_i$  captures the distance from the ideal family size and equals the birth order of child  $i$  minus the ideal fertility size reported by the mother.<sup>9</sup>  $1(\Delta\text{Ideal}_i \geq 0)$  is an indicator function equal to 1 if the mother has attained her ideal fertility and 0 otherwise. Regression (4) allows us to test the following prediction:

5. *There is a discontinuous increase in breastfeeding when a mother reaches her ideal fertility size.* Children whose mothers have reached their ideal fertility are breastfed for an extra  $(\tau + \omega \times \Delta\text{Ideal}_i)$  months, keeping all else fixed. The coefficient  $\omega$  allows children born after having reached their mother's ideal fertility to be breastfed for an increasingly longer period of time (given  $\omega$  is positive). At the cut-off point  $\Delta\text{Ideal}_i = 0$ , a discontinuous increase in breastfeeding duration is observed if  $\tau$  is positive and significant.

We also estimate the following regression:

$$\begin{aligned} \text{Duration}_i = & \alpha + \gamma\text{Male}_i + \tau\text{Male} \times 1(\Delta\text{Ideal}_i \geq 0) + \mu\text{Male}_i \\ & \times \Delta\text{Ideal}_i + \theta\text{Male}_i \times \Delta\text{Ideal}_i \times 1(\Delta\text{Ideal}_i \geq 0) \\ & + \delta X_i + a_i + \varepsilon_i, \end{aligned} \quad (5)$$

where the coefficient  $\tau$  captures the male-female difference in breastfeeding duration at  $\Delta\text{Ideal} = 0$  (i.e., just when a mother achieves her ideal fertility). Regression (5) allows us to test the following prediction:

6. *There is a discontinuous increase in male-female breastfeeding gap when a mother reaches her ideal fertility size.*

Finally, we estimate the following regressions:

$$\begin{aligned} \text{Duration}_i = & \alpha + \gamma\text{Male}_i + \tau_1 1(\Delta\text{Ideal } G_i \geq 0) + \varphi_1 \Delta\text{Ideal } G_i \\ & + \omega_1 1(\Delta\text{Ideal } G_i \geq 0) \times \Delta\text{Ideal } G_i + \tau_2 1(\Delta\text{Ideal}_i \geq 0) \\ & + \varphi_2 \Delta\text{Ideal}_i + \omega_2 1(\Delta\text{Ideal}_i \geq 0) \times \Delta\text{Ideal}_i + \delta X_i + a_i + \varepsilon_i, \end{aligned} \quad (6)$$

<sup>9</sup> The exact wording to determine the ideal family size in the PDHS is the following: "What is the ideal number of children that you would like to have in your whole life, irrespective of the number that you already have?" The question is framed in a way to ensure that the response does not entirely correlate with the number of existing children a mother has. The correlation between the ideal reported fertility and the actual family size is 0.22 in our sample.

where  $\Delta\text{Ideal } G_i$  is either  $\Delta\text{Ideal Sons}_i$ , the difference between the ideal number of sons a mother would like to have and the existing number of sons up to and including the child, or  $\Delta\text{Ideal Daughters}_i$ , the difference between the ideal number of daughters a mother would like to have and the existing number of daughters up to and including the child. Estimating (6) using  $\Delta\text{Ideal Sons}_i$  allows us to test the following prediction:

7. *Above and beyond the increase in breastfeeding duration observed on the realization of ideal total fertility, a mother breastfeeds her child significantly longer given that she has also attained her ideal number of sons.* Just when a mother attains her ideal male fertility (i.e.,  $\Delta\text{Ideal Sons}_i = 0$ ), she breastfeeds the child  $\tau_1$  months longer than the child of a mother who has not yet attained her ideal number of sons, keeping all else fixed. If  $\tau_1$  is positive and significant, we can be sufficiently confident that this effect is over and above that of attaining overall ideal fertility, since controls on overall ideal fertility are also included in the specification.

While estimating (6) using  $\Delta\text{Ideal Daughters}_i$  allows us to test the following prediction:

8. *Above and beyond the increase in breastfeeding duration observed on the realization of ideal total fertility, a mother does not breastfeed her child significantly longer on attaining her ideal number of daughters.* This holds true if  $\tau_1$  is not statistically different from 0.

This last prediction is not tested in either JK or Chakravarty.

Last, JK aimed to find the birth order at which the male-female difference in breastfeeding is expected to be most pronounced. To that end, we estimate the following regression:

$$\begin{aligned} \text{Duration}_i = & \alpha + \beta_1 \text{Birth Order}_i + \beta_2 \text{Birth Order}_i^2 \\ & + \gamma \text{Male}_i + \theta_1 \text{Birth Order}_i \times \text{Male}_i \\ & + \theta_2 \text{Birth Order}_i^2 \times \text{Male}_i + \delta X_i + a_i + \varepsilon_i. \end{aligned} \quad (7)$$

Estimating (7) allows us to test the last prediction:

9. *Male-female difference in breastfeeding is maximum at middle birth order.* At very low birth order, a mother wants to conceive again, so she weans

her children early irrespective of their gender. At very high birth order, a mother wants to stop fertility, so she breastfeeds both sons and daughters for an equally longer period of time. This implies that breastfeeding duration is most sensitive to the gender of the child at middle birth order.

#### **IV. Data: Sample Selection and Descriptive Statistics**

##### **A. Sample Selection**

To carry out the empirical analysis, we use the PDHS, the second survey conducted as part of the DHS international series.<sup>10</sup> This is the largest-ever household survey conducted in Pakistan (NIPS 2008), containing highly disaggregated information on every member of the household. We have detailed information for a total of 9,720 households. Restricting our attention to children born to ever-married 15–49-year-old women, the number of total households decreases from 9,720 to 8,798.

JK use the NFHS of India as their data source, which follows exactly the same format of the standard Demographic and Health Surveys. Therefore, using the PDHS, we have information on almost all the variables reported in the NFHS. One limitation to bear in mind, though, is that our sample size is significantly smaller than that in JK. After implementing data restrictions to ensure our analysis is comparable to that of JK (see app. A1; apps. A1, A2 are available online), our final sample comprises 6,955 children.

##### **B. Descriptive Statistics**

Tables 2A and 2B provide summary statistics for breastfeeding duration among children across two distinct dimensions: birth order and gender of the child. Following JK, we define birth order below 2 as low. Table 2A reveals that the mean breastfeeding duration is larger for higher birth order. In fact, compared to JK, the difference between low and high birth order is more pronounced in our sample.

The ideal number of children reported by the mother is also higher in our sample than in JK. The full sample mean of the variable  $\Delta\text{Ideal}$  (=birth order of present child – ideal number of children) is approximately  $-1.1$ , indicating that, on average, a mother wants one more child in order to reach her ideal fertility size. The desire for expanding family size seems to be stronger in Pakistan

<sup>10</sup> The first and third rounds of the survey were conducted in 1990 and 2012, respectively. However, none of these surveys is used, because the 1990 sample contains many missing values in the variables of interest, whereas the 2012 sample reports breastfeeding duration only for children currently breastfeeding and not for the entire sample.

**TABLE 2A**  
SUMMARY STATISTICS BY BIRTH ORDER

Variable	Birth Order $\leq 2$					Birth Order $> 2$				
	Mean	SD	Min	Max	Observations	Mean	SD	Min	Max	Observations
Months of breastfeeding	13.31	8.39	0	36	3,044	15.09	8.81	0	36	3,911
Birth order	1.49	.50	1	2	3,044	4.41	1.28	3	7	3,911
Ideal number of children	3.75	1.48	0	15	2,862	4.30	1.59	0	16	3,530
Birth order minus ideal number of children ( $\Delta$ ideal)	-2.26	1.57	-13	2	2,862	.06	1.87	-13	7	3,530
Male	.51	.50	0	1	3,044	.54	.50	0	1	3,911
Mother has at least one son	.74	.44	0	1	3,044	.96	.20	0	1	3,911
Total number of vaccinations	4.32	2.09	0	6	3,044	4.22	2.06	0	6	3,669
Age of child	1.90	1.42	0	4	3,044	1.96	1.42	0	4	3,911
Age of mother	24.94	4.31	15	45	3,044	30.84	5.22	18	49	3,911
Rural	.66	.47	0	1	3,044	.72	.45	0	1	3,911
Mother has no education	.53	.50	0	1	3,044	.69	.46	0	1	3,911
Mother has primary education	.18	.38	0	1	3,044	.15	.35	0	1	3,911
Mother has secondary education	.20	.40	0	1	3,044	.12	.32	0	1	3,911
Mother has higher education	.09	.28	0	1	3,044	.04	.21	0	1	3,911

**Note.** Data are drawn from NIPS (2008). Duration of breastfeeding (in months) is reported for children born in the last 5 years preceding the interview. Only single births are considered, who are alive at the time of the interview and have birth order at most 7. "Mother has at least one son" is defined to include the child himself. "Total number of vaccinations" is an aggregate of six binary variables for whether the child has received each of the polio 1, polio 2, polio 3, diphtheria/pertussis/tetanus (DPT) 2, DPT3, and measles vaccines, separately. Sampling weights are used.

than in India since  $\Delta$ ideal is more strongly negative at low birth order and less positive at high birth order.<sup>11</sup>

The patterns in the distribution of the remaining variables are similar to JK. The variable "mother has at least one son" has been defined to include the child himself. Total number of vaccinations is an aggregate of six binary variables for whether the child has received each of the polio 1, polio 2, polio 3, diphtheria/pertussis/tetanus (DPT) 2, DPT 3, and measles vaccines, separately. Information on vaccination coverage is not collected for every eligible child in the sample; hence, there are some missing values in this variable. As in JK, the majority of children in our sample live in rural areas—roughly 70%. This is in line with the total percentage of Pakistan's population residing in rural areas (68%; 1998 Pakistan Census). Comparing the mother's education across the range of birth

<sup>11</sup> The total number of responses on this question is less than the overall sample size. This is because a small proportion of mothers have given nonnumeric responses when asked about their ideal number of children, such as "up to God" or "don't know," so these observations (less than 1%) were eliminated for the purpose of our analysis.

**TABLE 2B**  
SUMMARY STATISTICS BY GENDER OF THE CHILD

Variable	Male					Female				
	Mean	SD	Min	Max	Observations	Mean	SD	Min	Max	Observations
Months of breastfeeding	14.73	8.88	.00	36	3,664	13.85	8.41	0	36	3,291
Birth order	3.16	1.77	1.00	7	3,664	3.11	1.77	1	7	3,291
Ideal number of children	4.03	1.62	.00	16	3,352	4.08	1.50	0	15	3,040
Birth order minus ideal number of children	-.92	2.08	-13	7	3,352	-1.02	2.09	-13	7	3,040
Male	1	0	1	1	3,664	0		0	0	3,291
Mother has at least one son	1	0	1	1	3,664	.71	.46	0	1	3,291
Total number of vaccinations	4.30	2.07	.00	6	3,460	4.22	2.08	0	6	3,117
Age of child	1.93	1.42	.00	4	3,664	1.94	1.42	0	4	3,291
Age of mother	28.33	5.75	15.00	48	3,664	28.21	5.56	15	49	3,291
Rural	.68	.47	.00	1	3,664	.70	.46	0	1	3,291
Mother has no education	.62	.48	.00	1	3,664	.62	.49	0	1	3,291
Mother has primary education	.15	.36	.00	1	3,664	.16	.37	0	1	3,291
Mother has secondary education	.15	.36	0	1	3,664	.16	.37	0	1	3,291
Mother has higher education	.07	.25	0	1	3,664	.06	.23	0	1	3,291

**Note.** See the note for table 2A.

order reveals that low birth order is associated with a higher probability of the mother having attained secondary or higher education.<sup>12</sup>

Table 2B shows that average demographic and mother characteristics are very similar across male and female populations of children. It is worth noting that the sex ratio in our final sample is more skewed (898 female children for every 1,000 male children) than that of JK (937 per 1,000). Vaccination coverage is slightly higher among sons, and sons are on average breastfed longer than daughters. The male advantage in breastfeeding is more pronounced in our sample than in that of JK.

## V. Results on Breastfeeding Duration

The outcome variable of interest is months of breastfeeding for a child at most 4 years of age. In the PDHS, this is reported for both children who were ever breastfed and those who are currently being breastfed. However, a complete spell of breastfeeding is not observed at the time the survey was conducted for the children who are still being nursed. To address this censoring problem,

<sup>12</sup> JK report the mother's education in terms of years of schooling. However, we use dummy variables for the level of the mother's education (primary, secondary, or higher). This is akin to what Chakravarty (2015) uses in his replication.

and following JK, we implement various correction procedures. First, our ordinary least squares (OLS) regressions will include child age (in months) fixed effects up until 36 months (the maximum duration of breastfeeding). Second, since OLS regressions by construction will still not give us an estimate of completed duration of breastfeeding, we will also report results from a hazard model. In the context of this analysis, hazard is defined as the event that a mother weans her child. The hazard model estimates the probability that a mother stops breastfeeding her child at time  $t$  conditional on her still nursing her child at time  $t - 1$ . In particular, we will estimate a Cox proportional hazard (CPH) model. In this model, the covariates are assumed to be multiplicatively related to the hazard rate.

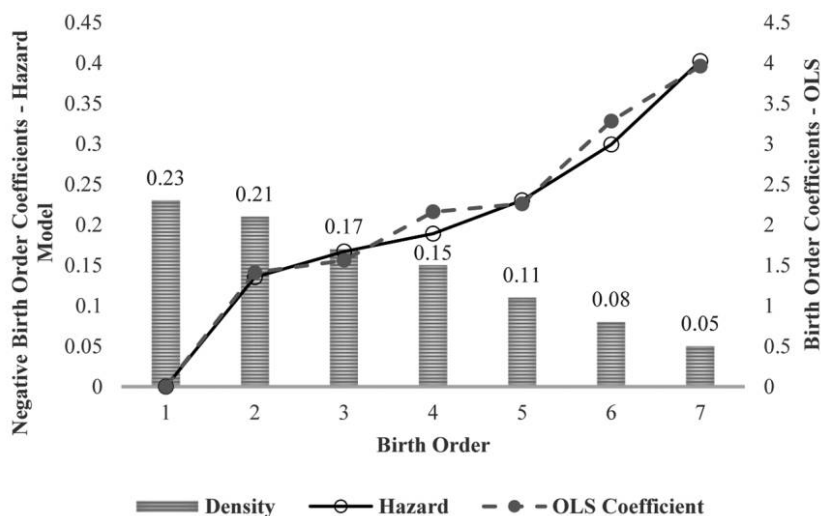
#### **A. Birth Order and Breastfeeding Duration**

There is a vast medical literature that explains the role of breastfeeding in limiting subsequent fertility (e.g., Kennedy et al. 1989), referred to as the lactational amenorrhea method. Breastfeeding as a contraceptive method is based on the fact that lactation (production of breast milk) causes amenorrhea (absence of menstrual period); breast milk production adversely affects hormones that trigger ovulation, hence preventing future pregnancy. For the purposes of our study, it is important to highlight the study for Pakistan by Kazi et al. (1995), who showed that the pregnancy rate for mothers during lactational amenorrhea is as low as 1.1% in the year following childbirth.

The relationship between the duration of breastfeeding and birth order of the child is plotted in figure 1, which reports the coefficients on seven dummy birth-order variables (from 1 to 7), where 1 is the reference category, estimated with a CPH model or OLS (in this last case controlling for child age [in months] fixed effects). To make the coefficients comparable, the signs of the CPH coefficients are reversed in figure 1. In addition, the figure also displays the histogram for the distribution of birth order. We can see that the duration of breastfeeding increases for every subsequent birth order (OLS interpretation) or, alternatively, the probability of weaning a child falls with increasing birth order (CPH interpretation). Compared with JK, the OLS point estimates are larger in magnitude.

Table 3 investigates the relationship displayed in figure 1 by means of a regression analysis. All OLS regressions include child age (in months) fixed effects to correct for right censoring in the duration of breastfeeding of children currently being nursed. Column 1 does not include any covariates. Adding covariates in column 2, the coefficient on birth order reduces from 0.6 to 0.4, roughly, but remains positive and highly significant. Therefore, a unit increase in birth order is associated with 0.4 months of additional breastfeeding, keep-





**Figure 1.** Birth order and breastfeeding duration. This figure plots the ordinary least squares (OLS) and hazard model coefficients on seven dummy birth-order variables (from 1 to 7), where 1 is the reference category, estimated with a Cox proportional hazard (CPH) model or OLS (in this last case controlling for child age [in months] fixed effects). To make the coefficients comparable, the signs of the CPH coefficients are reversed. A histogram displaying the distribution of birth order in the sample is also shown.

ing all else constant. Moreover, the coefficient is twice as large as that in JK ( $0.40 > 0.21$ ). The results from the hazard model in column 3 give us the failure probability (i.e., the probability of weaning a child in period  $t$  conditional on the mother still nursing him or her in period  $t - 1$ ). The coefficient on birth order indicates that with each subsequent birth, the probability of weaning a child falls by 4%, *ceteris paribus*. Note that the total number of observations in the hazard model is smaller than in the OLS regressions because, by construction, the hazard model eliminates observations for which the duration of breastfeeding is equal to 0. Finally, in column 4 we use observations only for children who were breastfed in the past and not currently. Therefore, the total number of observations in this column is relatively much smaller. The coefficient on birth order is highly statistically significant. It shows that the completed duration of breastfeeding increases by 0.4 months on average with a unit increase in birth order. Also, note that the size of this coefficient is comparable to that in column 2. This shows that the age-in-months fixed effects serve the purpose of correcting for the censoring bias.

### B. Gender and Breastfeeding Duration

Given the contraceptive function of breastfeeding and a general preference for sons, mothers wishing to conceive a son in the future will wean their daughters earlier. Figure 2 plots the survival curves, showing the probability of still being

**TABLE 3**  
BIRTH ORDER AND BREASTFEEDING DURATION

	OLS		Hazard	OLS
	(1)	(2)		
Birth order	.581 (.057)***	.371 (.075)***	-.037 (.008)***	.388 (.108)***
Male		.860 (.190)***	-.093 (.020)***	.981 (.276)***
Mother's education:				
Primary education		.037 (.297)	.019 (.031)	.421 (.427)
Secondary education		-.234 (.332)	.017 (.031)	.015 (.467)
Higher education		-1.226 (.457)***	.134 (.048)***	-.968 (.602)
Rural		.829 (.271)***	-.052 (.025)**	.843 (.384)**
Mother's age		-.083 (.144)	-.010 (.015)	.128 (.221)
Mother's age <sup>2</sup>		.003 (.002)	.000 (.000)	-.001 (.004)
Child's age		10.722 (.405)***	-2.005 (.039)***	9.336 (.432)***
Child's age <sup>2</sup>		-1.587 (.105)***	.337 (.008)***	-1.397 (.087)***
Punjab		-1.275 (.400)***	.025 (.044)	-1.735 (.548)***
Sindh		-.969 (.413)**	-.016 (.045)	-2.138 (.567)***
North-West Frontier Province		-.387 (.427)	-.081 (.045)*	-.971 (.588)*
Age-in-months fixed effects	Yes	Yes	No	No
Censored sample	No	No	No	Yes
R <sup>2</sup>	.37	.44		.18
Number of mothers	962	962	962	944
N	6,955	6,955	6,816	4,335

**Note.** The dependent variable is breastfeeding duration (in months) of the child born within the last 5 years preceding the survey, curtailed at 36 months. No education of the mother and the province Balochistan of Pakistan are the omitted base categories. Col. 3 reports regression results from a censored sample, where the unit of analysis is a child who was breastfed earlier but is not currently breastfeeding. All other columns report results on the full sample of children. Robust standard errors clustered by the mother are reported in parentheses. Sampling weights are used. OLS = ordinary least squares.

\*  $p < .1$ .

\*\*  $p < .05$ .

\*\*\*  $p < .01$ .

breastfed with respect to the duration of breastfeeding (alternately, age of child) for male and female children, separately.

We observe that at every value of duration of breastfeeding (alternately, age of child), and particularly between 6 and 24 months, the probability of a boy surviving breastfeeding is higher than that for a girl. The gender differential is evident at every age except at multiples of 6 months (e.g., at 12, 18, and



**Figure 2.** Gender and breastfeeding duration. This figure displays the probability of still being breastfed (survival probability) against duration of breastfeeding, separately for male and female children. Duration is top-coded at 36 months. The survival graphs have been estimated using results from the Cox proportional hazard (CPH) model in col. 3 of table 3. In evaluating the graph, all covariates (except the male dummy) are by default set at their mean values.

24 months). As pointed out in JK, this may be explained by the mother's "rounding error in reporting and an actual propensity to breastfeed up to a focal point." Compared to JK, male breastfeeding advantage in our sample opens up earlier and is relatively more apparent between 6 and 12 months. However, it also closes down earlier, as there is no apparent gender gap in breastfeeding beyond 24 months.

Correcting for the censored spells, we found that the mean of completed breastfeeding duration among boys and girls is 15.2 and 14.3 months, respectively. The mean completed durations are, however, not much different from the observed means reported in table 2B (14.7 and 13.9, respectively). On the other hand, the difference between completed and observed breastfeeding duration is much more pronounced in JK for both males and females. This is because children in our sample are weaned earlier compared to those in JK. Also, the distribution of spell durations is more evenly distributed in JK. Roughly 50% of the children in India are weaned by the age of 2 and 75% by the age of 2.5 years. On the other hand, almost the entire sample of children in Pakistan has been weaned by the age of 2 years. This shows that there is only a negligible number of observations in our sample for which breastfeeding was top-coded at 36 months, hence partly explaining why observed means are not far from corrected means.

**TABLE 4**  
GENDER AND BREASTFEEDING DURATION

	OLS		Hazard (3)
	(1)	(2)	
Male	.855 (.201)***	.860 (.190)***	-.093 (.020)***
Covariates	No	Yes	Yes
R <sup>2</sup>	.36	.44	
Number of mothers	962	962	962
N	6,955	6,955	6,816

**Note.** The dependent variable is breastfeeding duration (in months) of the child born within the last 5 years preceding the survey, top-coded at 36 months. This includes both children currently breastfeeding and children already weaned. All OLS regressions include child age (in months) fixed effects up to and including 36 months to correct for right censoring. The standard set of covariates includes birth order, dummy variables for the mother's educational attainment, linear and quadratic controls for the child's and the mother's age, type of place of residence (rural/urban), and dummy variables for region of residence (provinces of Pakistan). Robust standard errors clustered by the mother are reported in parentheses. Sampling weights are used. OLS = ordinary least squares.

\*\*\*  $p < .01$ .

Table 4 shows that sons are on average breastfed 0.86 months longer than daughters. The gender differential does not decrease in either magnitude or statistical significance by the addition of covariates, column 2. In terms of magnitude, the male advantage in breastfeeding is more than twice as large in our sample as in JK (0.86 vs. 0.39), consistent with the findings from Arnold (1997) regarding son preference: sons are preferred 10:1 in Pakistan and approximately 5:1 in India. Results from the proportional hazard model indicate that the probability of weaning a daughter in any given month is 9.3% higher than that for a son, keeping all else fixed.

### C. Sibling Sex Composition and Breastfeeding Duration

Regression results from table 5 test the third prediction. All specifications include the standard set of covariates in addition to a linear control for birth order. Results remain robust when replacing the linear control for birth order with birth-order fixed effects. Controlling for birth order ensures that the effect of the male share of elder siblings on breastfeeding duration is separate from the effect of increasing total fertility.<sup>13</sup> Column 1 in table 5 shows that

<sup>13</sup> We measure sibling sex composition as the male share of elder siblings for each child in the sample, defined to be equal to the number of elder siblings that are male divided by the total number of elder siblings. By construction, this variable will be undefined for children with birth order equal to 1, since they have no elder siblings. The omission of first-borns would make our sample nonrepresentative and, further, reduce the total sample size by a significant 23%. To deal with this concern, we assign the variable "male share of elder siblings" a value of 0 for all first-borns. First, this correction serves the

**TABLE 5**  
**SIBLING SEX COMPOSITION AND BREASTFEEDING DURATION**

	OLS		Hazard		OLS (5)
	(1)	(2)	(3)	(4)	
Male	.863 (.190)***	.859 (.190)***	-.093 (.020)***	-.092 (.020)***	1.198 (.273)***
Male share of elder siblings	.627 (.266)**		-.084 (.029)***		1.066 (.355)***
At least one elder male sibling		.512 (.249)**		-.063 (.028)**	
Male × male share of elder siblings					-.850 (.484)*
Covariates	Yes	Yes	Yes	Yes	Yes
R <sup>2</sup>	.44	.44			.44
Number of mothers	962	962	962	962	962
N	6,955	6,955	6,816	6,816	6,955

**Note.** The dependent variable is breastfeeding duration (in months) of the child born within the last 5 years preceding the survey, top-coded at 36 months. This includes both children currently breastfeeding and children already weaned. "Male share of elder siblings" is equal to the number of elder siblings that are male divided by the total number of elder siblings and is 0 for first-borns. "At least one elder male sibling" is a dummy variable equal to 1 if the child has at least one elder brother and 0 otherwise. All OLS regressions include child age (in months) fixed effects up to and including 36 months to correct for right censoring. The standard set of covariates includes birth order, dummy variables for the mother's educational attainment, linear and quadratic controls for the child's and the mother's age, type of place of residence (rural/urban), and dummy variables for region of residence (provinces of Pakistan). Robust standard errors clustered by the mother are reported in parentheses. Sampling weights are used. OLS = ordinary least squares.

\*  $p < .1$ .

\*\*  $p < .05$ .

\*\*\*  $p < .01$ .

if the male share of a child's elder siblings increases from 0 to 1, the child is breastfed on average an additional 0.63 months. Column 3 replicates these results using the CHP model. The coefficient on male share is highly statistically significant and indicates a maximum reduction of 8.4% in the probability of weaning a child if all elder siblings are male.

benefit of keeping our sample intact. Second, assigning the variable a value of 0 for first-borns also makes logical sense. Male share of elder siblings takes a value of 0 for first-borns (by imposition) as well as for children with all elder siblings of female gender (by construction). We can argue that the duration of breastfeeding should be roughly the same in both instances, because in the first case, a mother breastfeeds her first child for a shorter duration (since she wants more children irrespective of the gender of the first-born, especially in the context of high fertility rate in Pakistan), and in the second case, a mother breastfeeds her last daughter for a shorter duration (because she desires to conceive a son). Since the value of the dependent variable is expected to be the same in both instances, assigning the variable a common value of 0 is not an unreasonable formulation. Finally, we could define male share of elder siblings as the number of elder siblings that are male divided by the total number of elder siblings up to and including the child. This definition automatically assigns a value of 0 to first-borns. Our results in table 5 remain robust even if we use this definition. However, strictly speaking, this variable is not a fraction, as it can never equal 1.

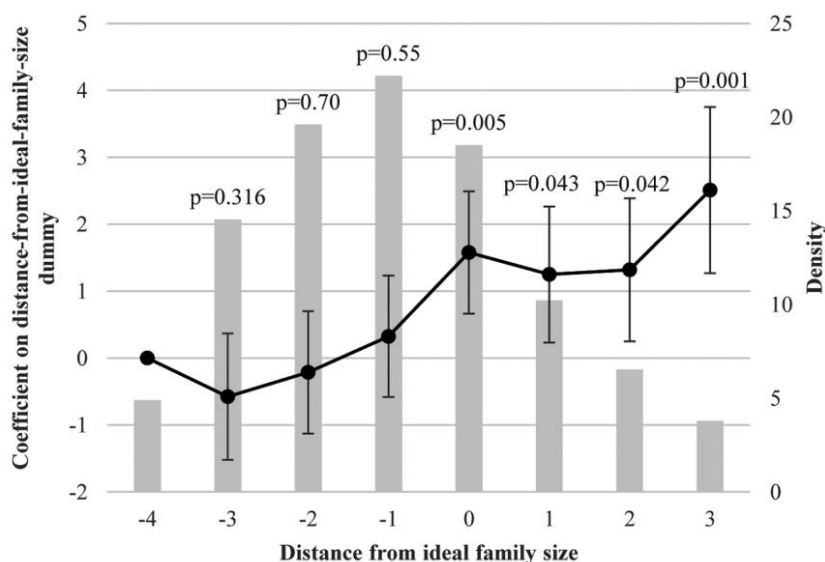
Similar to JK, we also assess the impact of the mother having at least one son on the duration of breastfeeding of the current child. The binary variable, at least one son, is defined to equal 1 if the male share of elder siblings is greater than 0 and is 0 otherwise. Column 2 shows that separate from a simple male breastfeeding advantage, a child who has at least one elder male sibling is nursed for an additional 0.5 months on average. This result remains robust under the alternate hazard model specification, column 4, where the probability of weaning a child decreases by a significant 6.3% conditional on the child having at least one elder brother. Again, the coefficients in our regression are larger than in JK. Both the overall male breastfeeding advantage and the effect of increasing male share on breastfeeding duration are more pronounced in our sample.

Further, according to the fourth prediction (Chakravarty 2015), we may expect the effect of children's sex composition on breastfeeding duration to be different for boys and girls. The results in the last column of table 5 are indeed consistent with Chakravarty (2015); girls are breastfed for an additional 0.85 months compared to boys as the ratio of elder male siblings increases from 0 to 1, keeping all else fixed. This finding fits the presumption that parents want to achieve some balance in their children's sex composition. Indeed, several studies from the Indian subcontinent corroborate this finding. For instance, Pande (2003) provides evidence that parents in India want to have at least one daughter, even though preference for sons remains persistent overall.

#### ***D. Ideal Fertility Size and Breastfeeding Duration***

Prediction 5 states that there is a sharp discontinuous increase in breastfeeding duration just when mothers reach their ideal fertility size. Figure 3 plots the coefficients on eight dummy indicators (deviations from ideal fertility size, from  $-4$  to  $3$ ), where  $-4$  is the reference category, estimated with OLS (controlling for child age [in months] fixed effects), together with their respective 90% confidence intervals and  $p$ -values.<sup>14</sup> All the coefficients below the zero deviation from the ideal family size are either 0 (for  $-4$ , by construction) or not different from 0 (all  $p > .3$ ). In other words, before reaching the ideal family size, there is no relationship between family size and breastfeeding. However, there is a sharp increase from  $\Delta\text{Ideal}_i = -1$  to  $\Delta\text{Ideal}_i = 0$ , so mothers suddenly increase breastfeeding duration as they approach their ideal fertility. In addition, all the coefficients above the zero deviation from the ideal family size are positive and statistically different from 0 (all  $p \leq .043$ ).

<sup>14</sup> As in JK, outliers are excluded to ensure that extreme values do not bias results. JK limit the variable in the range  $[-4, 4]$ , but we only consider the range  $[-4, 3]$  because  $\Delta\text{Ideal}$  is less symmetrically distributed in our sample. The mean of  $\Delta\text{Ideal}$  is  $-1$  in our sample compared to roughly  $-0.2$  in JK. Further, the histogram of  $\Delta\text{Ideal}$  in fig. 3 shows a significantly negatively skewed distribution.



**Figure 3.** Distance from ideal fertility and breastfeeding duration. This figure plots the coefficients on eight dummy indicators (deviations from ideal fertility size, from  $-4$  to  $3$ ), where  $-4$  is the reference category, estimated with ordinary least squares (controlling for child age [in months] fixed effects), together with their respective 90% confidence intervals and  $p$ -values. The omitted category in regression is the binary indicator  $1(\Delta Ideal = -4)$ , for which the coefficient is normalized to 0. Distance from ideal family size ( $\Delta Ideal$ ) is equal to the birth order of the child minus the ideal fertility reported by the mother. A histogram displaying the distribution of  $\Delta Ideal$  in the sample is also shown.

Table 6 provides further support for prediction 5 using regression analysis. Column 1 includes an indicator function for whether the mother has attained her overall ideal fertility (along with the standard set of covariates). The coefficient indicates that a mother breastfeeds her child an additional 1.2 months on reaching her ideal fertility, keeping all else fixed. Columns 2 and 3 estimate equation (4). This specification not only permits a discontinuous jump in breastfeeding duration just when mothers reach their ideal fertility (at  $\Delta Ideal_i = 0$ ) but also allows for a linear trend on either side of this discontinuity. In other words, we can determine whether breastfeeding duration is an increasing function of the distance from the ideal family size, beyond  $\Delta Ideal_i = 0$ .

The indicator variable in column 2 is highly statistically significant, indicating that a mother breastfeeds her child an extra 1.5 months once ideal fertility has been achieved. Including the standard set of covariates in the specification, this coefficient reduces to 1.1 months but remains highly statistically significant nonetheless. While the coefficient on the interaction term is not significant, it does match in sign with JK. This could be attributed to the large difference in our sample sizes, or it could simply mean an absence of increasing



**TABLE 6**  
DISTANCE FROM IDEAL FERTILITY AND BREASTFEEDING DURATION

	OLS		
	(1)	(2)	(3)
1( $\Delta\text{Ideal} \geq 0$ )	1.163 (.238)***	1.545 (.326)***	1.077 (.306)***
$\Delta\text{Ideal}$		-.046 (.097)	.076 (.092)
$\Delta\text{Ideal} \times 1(\Delta\text{Ideal} \geq 0)$		.139 (.163)	-.155 (.152)
Covariates	Yes	No	Yes
$R^2$	.44	.36	.44
Number of mothers	957	957	957
N	6,392	6,392	6,392

**Note.** The dependent variable is breastfeeding duration (in months) of the child born within the last 5 years preceding the survey, top-coded at 36 months.  $1(\Delta\text{Ideal} \geq 0)$  is an indicator function equal to 1 if the distance from the ideal family size ( $\Delta\text{Ideal} = \text{Birth Order} - \text{Ideal}$ ) is greater than or equal to 0. All regressions include child age (in months) fixed effects up to and including 36 months to correct for right censoring. Additional covariates include a gender of the child dummy, dummy variables for the mother's educational attainment, linear and quadratic controls for the child's and the mother's age, type of place of residence (rural/urban), and dummy variables for region of residence (provinces of Pakistan). Robust standard errors clustered by the mother are reported in parentheses. Sampling weights are used. OLS = ordinary least squares.

\*\*\*  $p < .01$ .

trend beyond  $\Delta\text{Ideal}_i \geq 0$  in our sample. Remarkably enough, the size of the coefficient on the indicator variable in our sample is more than twice as large as that in JK (just as the male breastfeeding advantage was in table 4).

According to prediction 6, there will be a male-female difference in breastfeeding at  $\Delta\text{Ideal}_i = 0$ . This prediction is investigated in table 7. Column 1 reports estimates without including covariates in the specification. Column 2 adds covariates, including a linear control for  $\Delta\text{Ideal}_i$ , analogous to JK incorporating  $\Delta\text{Ideal}_i$  fixed effects. Column 2 implies that the derivative of breastfeeding duration with respect to having reached ideal fertility is given by  $1.28 \text{ Male} + 0.16 \text{ Male} \times \Delta\text{Ideal}_i$ . At  $\Delta\text{Ideal}_i = 0$ , therefore, a son is breastfed 1.3 months more than a daughter. The coefficient on the third interaction term is statistically insignificant, so after having reached ideal fertility, the male-female difference in breastfeeding does not depend on how much the actual birth order exceeds the reported ideal fertility. Remarkably enough, the gender differential of 1.3 months (as mothers reach their ideal fertility) is significant and larger in magnitude than any of the effects we have studied so far (the general male advantage in breastfeeding was 0.9 months, and the male-female difference in breastfeeding duration conditional on being born after all-male elder siblings was 0.9 months). The analogous estimate from JK is 0.59 months. As a robustness check, the third column reports results from the proportional

**TABLE 7**  
**GENDER GAP IN BREASTFEEDING AT IDEAL FERTILITY SIZE**

	OLS		Hazard (3)
	(1)	(2)	
Male	-.452 (.401)	.001 (.377)	-.034 (.037)
Male $\times$ 1( $\Delta$ Ideal $\geq$ 0)	1.559 (.464)***	1.277 (.433)***	-.083 (.042)*
Male $\times$ $\Delta$ Ideal	-.541 (.134)***	-.339 (.127)***	.026 (.014)*
Male $\times$ $\Delta$ Ideal $\times$ 1( $\Delta$ Ideal $\geq$ 0)	.443 (.225)**	.159 (.210)	-.004 (.021)
$\Delta$ Ideal	.334 (.070)***	.251 (.075)***	-.036 (.008)***
Covariates	No	Yes	Yes
$R^2$	.37	.44	
Number of mothers	957	957	957
N	6,392	6,392	6,263

**Note.** The dependent variable is breastfeeding duration (in months) of the child born within the last 5 years preceding the survey, top-coded at 36 months. 1( $\Delta$ Ideal  $\geq$  0) is an indicator function equal to 1 if the distance from the ideal family size ( $\Delta$ Ideal = Birth Order – Ideal) is greater than or equal to 0. All OLS regressions include child age (in months) fixed effects up to and including 36 months to correct for right censoring. Additional covariates include dummy variables for the mother's educational attainment, linear and quadratic controls for the child's and the mother's age, type of place of residence (rural/urban), and dummy variables for region of residence (provinces of Pakistan). Robust standard errors clustered by the mother are reported in parentheses. Sampling weights are used. OLS = ordinary least squares.

\*  $p < .1$ .

\*\*  $p < .05$ .

\*\*\*  $p < .01$ .

hazard model. At  $\Delta$ Ideal<sub>*i*</sub> = 0, sons are 8.3% less likely to be weaned compared to daughters, all else fixed.

### *E. Ideal Sons vs. Ideal Daughters*

In addition to reporting the ideal number of total children a mother would like to have, the PDHS also reports the ideal number of sons and daughters, irrespective of the current sex composition of children. In the previous subsection, we found a sharp increase in breastfeeding duration just as a mother reaches her ideal fertility limit. Putting this in the context of son-biased preferences, we would like to explore whether mothers nurse their children even longer in response to having attained their ideal number of sons (prediction 7). Further, we would like to evaluate how this effect differs from having attained the ideal number of daughters (prediction 8). Testing this last prediction provides an extension to the previous work by JK and Chakravarty.

The means (standard deviation) of the ideal number of sons and daughters in our sample are 2.4 (1.4) and 1.6 (0.9), respectively. Figure 4 displays the

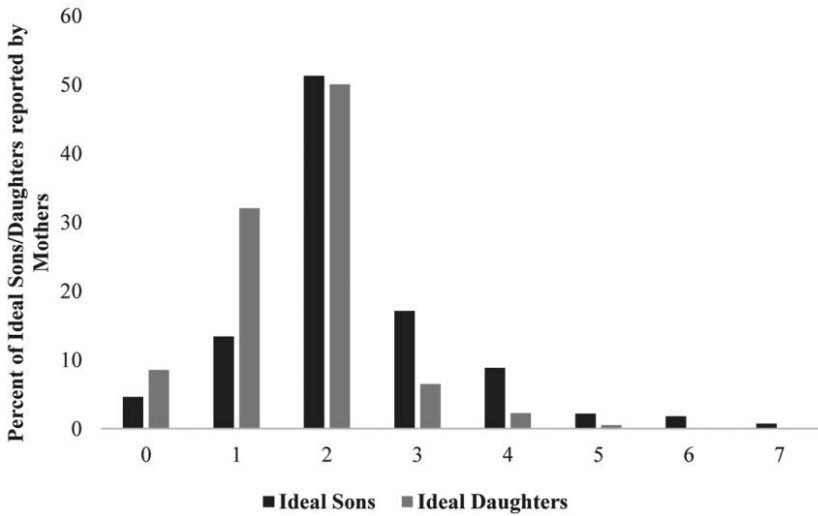


Figure 4. Frequency distribution of the reported ideal number of sons and daughters.

distribution of responses received from mothers on their reported ideal number of sons and daughters. Note that there is an overwhelming preference for sons: 91% of mothers want to have no more than two daughters, while only 69% want no more than two sons, with the remaining significant 31% expressing desiring for more.

Column 1 of table 8 shows the regression estimates from equation (6) without the controls on overall ideal fertility, with the standard set of covariates included in the equation. A mother whose present number of sons equals her ideal reported number of sons breastfeeds her child an additional 1.1 months compared to a mother who has not yet attained her ideal male fertility. The analogous estimate in the case of ideal number of daughters is statistically insignificant and relatively smaller in magnitude (col. 2). A sharp increase in breastfeeding duration just at the point  $\Delta \text{Ideal Sons}_i = 0$  and not at  $\Delta \text{Ideal Daughters}_i = 0$  illustrates that mothers want to stop future pregnancy once they have borne their ideal number of sons but not necessarily in the case of attaining their ideal number of daughters.

Column 3 allows us to quantify the effect of having attained the ideal number of sons separately from that of having reached overall ideal fertility. Above and beyond the effect of reaching the total ideal number of children, a mother breastfeeds her child an additional 0.9 months if she has also attained her ideal number of sons. Adding both of these effects together, we find that breastfeeding duration increases by 1.5 months ( $0.8 + 0.66$ ) compared to a child whose mother has achieved neither her total nor her male ideal fertility. The analogous estimate

**TABLE 8**  
**DISTANCE FROM IDEAL MALE/FEMALE FERTILITY AND BREASTFEEDING DURATION**

	OLS				Hazard	
	(1)	(2)	(3)	(4)	(5)	(6)
Male	.710 (.207)***	1.153 (.208)***	.902 (.212)***	.962 (.217)***	-.079 (.022)***	-.087 (.023)***
1( $\Delta$ Ideal Sons $\geq 0$ )	1.069 (.307)***		.863 (.335)**		-.073 (.034)**	
$\Delta$ Ideal Sons	-.098 (.106)		-.378 (.151)**		.007 (.017)	
1( $\Delta$ Ideal Sons $\geq 0$ ) $\times$ $\Delta$ Ideal Sons	.085 (.202)		.307 (.263)		-.015 (.025)	
1( $\Delta$ Ideal Daughters $\geq 0$ )		.569 (.412)		.334 (.412)		-.061 (.044)
$\Delta$ Ideal Daughters		.171 (.250)		.068 (.267)		.029 (.029)
1( $\Delta$ Ideal Daughters $\geq 0$ ) $\times$ $\Delta$ Ideal Daughters		-.144 (.272)		-.190 (.299)		.002 (.032)
1( $\Delta$ Ideal $\geq 0$ )			.656 (.340)*	1.041 (.311)***	-.055 (.036)	-.077 (.033)**
$\Delta$ Ideal			.265 (.134)**	.028 (.105)	-.022 (.015)	-.027 (.011)**
1( $\Delta$ Ideal $\geq 0$ ) $\times$ $\Delta$ Ideal			-.322 (.202)	-.057 (.174)	.031 (.020)	.016 (.019)
$R^2$	.44	.44	.44	.44		
Number of mothers	957	957	957	957	957	957
N	6,388	6,388	6,388	6,388	6,259	6,259

**Note.** The dependent variable is breastfeeding duration (in months) of the child born within the last 5 years preceding the survey, top-coded at 36 months.  $\Delta$ IdealSons is the total number of sons up to and including the child minus the ideal number of sons reported by the mother. 1( $\Delta$ IdealSons  $\geq 0$ ) is an indicator function equal to 1 if the mother has attained (or exceeded) her reported ideal male fertility. The remaining variables are defined analogously. All OLS regressions include child age (in months) fixed effects up to and including 36 months to correct for right censoring. Additional covariates include dummy variables for the mother's educational attainment, linear and quadratic controls for the child's and the mother's age, type of place of residence (rural/urban), and dummy variables for region of residence (provinces of Pakistan). Robust standard errors clustered by the mother are reported in parentheses. Sampling weights are used. OLS = ordinary least squares.

\*  $p < .1$ .

\*\*  $p < .05$ .

\*\*\*  $p < .01$ .

in JK is 0.74 months, just about half that of ours. Column 4 repeats the same exercise in the case of daughters. The only significant effect now is that of having realized ideal total fertility. Having an ideal number of daughters has no separate influence on the mother's timing of weaning her child.

The last two columns report results from the CPH model and correct for the problem of right censoring. The probability of weaning a child reduces by 7.3% once a mother has borne her ideal number of sons, keeping all else fixed. On the other hand, above and beyond the effect of reaching total ideal fertility, the effect of having attained the ideal number of daughters on breastfeeding duration is statistically insignificant in the last column.

### F. Interaction of Gender and Birth Order

Finally, we are left with prediction 9, that the male-female difference in breastfeeding peaks at middle birth order. Results are presented in table 9. OLS estimates are reported, and all specifications include the standard set of covariates with age-in-month fixed effects. The first column is a replication of JK's specification and includes birth-order fixed effects. The coefficients on the interaction terms are statistically insignificant and opposite in sign to JK, thereby rejecting the presence of an inverted u-shaped relationship between gender differential and birth order in our sample. Column 2 is a replication of Chakraborty's (2015) specification. Note that now the interaction of gender and birth order is strongly statistically significant and that of gender and squared birth order is marginally significant ( $p = .109$ ). Jointly, both variables are strongly statistically significant ( $p = .000$ ). In addition, the signs of the coefficients on interaction terms are in line with prediction 9. As birth order increases, gender differential increases but at a diminishing rate. This predicts that the gender difference in breastfeeding is likely to be most pronounced at the middle birth order.

**TABLE 9**  
GENDER GAP IN BREASTFEEDING AT MIDDLE BIRTH ORDER

	OLS		
	(1)	(2)	(3)
Male	.828 (.686)		.813 (.687)
Male $\times$ birth order	-.007 (.422)	.472 (.169)***	-.013 (.421)
Male $\times$ birth order <sup>2</sup>	.005 (.057)	-.053 (.033)	.007 (.057)
Birth order		.381 (.256)	.639 (.311)**
Birth order <sup>2</sup>		-.008 (.035)	-.039 (.041)
Birth order fixed effects	Yes	No	No
R <sup>2</sup>	.44	.44	.44
Number of mothers	962	962	962
N	6,955	6,955	6,955

**Note.** The dependent variable is breastfeeding duration (in months) of the child born within the last 5 years preceding the survey, top-coded at 36 months. All regressions include child age (in months) fixed effects up to and including 36 months to correct for right censoring. Additional covariates include dummy variables for the mother's educational attainment, linear and quadratic controls for the child's and the mother's age, type of place of residence (rural/urban), and dummy variables for region of residence (provinces of Pakistan). Robust standard errors clustered by the mother are reported in parentheses. Sampling weights are used. OLS = ordinary least squares.

\*\*  $p < .05$ .

\*\*\*  $p < .01$ .

However, Chakravarty's specification omits the gender main effect while including the interactions of gender with birth order. Adding the gender dummy in column 3, we essentially find results consistent with column 1, that is, there is no evidence for a peak in male advantage in breastfeeding at middle birth order. Thus, this is the only prediction of JK that we fail to support in our analysis.

### G. Summary

We find evidence supporting all predictions in JK (1, 2, 3, 5, 6, and 7) but 9, the additional prediction in Chakravarty (4), and our new prediction (8). In contrast to JK, who document a linear increasing trend on the right side of the discontinuity at  $\Delta \text{Ideal}_i = 0$ , we find evidence for a constant trend that indicates that the desire to curtail fertility is constant for every birth order exceeding the ideal fertility limit. However, the magnitude of the jump in breastfeeding is higher in our sample, so mothers reveal a stronger desire to limit fertility once they reach their ideal family size in Pakistan.<sup>15</sup>

The gender gap in breastfeeding is twice as large in our sample as that in JK. This difference in magnitude could be generated by two factors: (1) sampling variation due to very different sample sizes (around 7,000 observations in our study compared to more than 100,000 in JK) or (2) stronger son preference in Pakistan. Unfortunately, other available data sets in Pakistan (e.g., the 2012 PDHS; the 2011 Punjab Multiple Indicator Cluster Survey) do not report information on breastfeeding duration for children no longer nursed. Thus, it is not possible to expand our sample size to completely exclude the first factor as a potential explanation. However, we do have evidence in favor of the second factor. JK found that the male advantage in breastfeeding was largest in the northern states of India (where son preference is stronger): sons are breastfed 0.6 months longer in Punjab and Haryana compared to 0.2 months in southern states such as Kerala. Since North India forms the geographical and cultural neighborhood of East Pakistan (comprising the bulk of its population), it makes sense to expect estimates from Pakistan to be more pronounced than the overall estimates from India as a whole.

## VI. Results Using Unwantedness of Pregnancy

One of the pervasive questions in economics research is whether stated preference-based measures provide useful information on revealed preference-based measures. In this section, we estimate all the main specifications considered in this article using an alternate outcome variable: whether the child born in the last 5 years was wanted at that time ( $=0$ ) or not wanted or wanted later

<sup>15</sup> Table 1A in appendix A2 provides additional support against alternate explanations.

**TABLE 10**  
**BREASTFEEDING DURATION AND UNWANTEDNESS OF LAST PREGNANCY**

	OLS	
	(1)	(2)
Unwanted pregnancy	-1.479 (.237)***	-3.434 (.492)***
Birth order	.477 (.077)***	.306 (.083)***
Unwanted pregnancy × birth order		.564 (.124)***
Male	.843 (.190)***	.863 (.189)***
R <sup>2</sup>	.45	.45
Number of mothers	962	962
N	6,941	6,941

**Note.** The dependent variable is breastfeeding duration (in months). Unwanted pregnancy equals 1 if the last child born in the last 5 years was not wanted then and equals 0 otherwise. All regressions include child age (in months) fixed effects up to and including 36 months to correct for right censoring. Additional covariates include dummy variables for the mother's educational attainment, linear and quadratic controls for the child's and the mother's age, type of place of residence (rural/urban), and dummy variables for region of residence (provinces of Pakistan). Robust standard errors clustered by the mother are reported in parentheses. Sampling weights are used. OLS = ordinary least squares.

\*\*\*  $p < .01$ .

(=1).<sup>16</sup> Incorporating this subjective information into our analysis serves two important purposes. First, it provides a robustness check for the link between breastfeeding and subsequent fertility (table 10), but more importantly, it sheds light on the validity of stated preference-based measures (e.g., unwantedness of a pregnancy) as complements to revealed preference-based measures (e.g., duration of breastfeeding) or even substitutes in case the former are available but the latter are not (tables 11, 12).

Breastfeeding duration and unwantedness of a pregnancy can be both negatively and positively correlated. Ex post declaration of a pregnancy as unwanted could possibly be due to two reasons: (1) the mother did not want any more children or (2) the mother instead wanted a child of the other sex. In the first case, breastfeeding duration and unwantedness of a pregnancy would be positively correlated, as a mother nurses her child longer to curtail future fertility, whereas in the second case they would be negatively correlated, as the mother

<sup>16</sup> This is the same variable used in the analysis of Hussain et al. (2000). However, they used it as an ordered categorical variable, whereas we are using it as a binary variable. In our sample, 69.4% of the mothers wanted the child then, 14.8% wanted it later, and 15.8% did not want it at all. Combining the "wanted later" and "not wanted" responses together makes the number of observations more balanced in each category. Moreover, it is likely that mothers are reluctant to admit that the child was not wanted at all, blurring the distinction between wanted later and not wanted, and therefore we chose to merge these observations under a single category.



**TABLE 11**  
UNWANTED PREGNANCIES BY BIRTH ORDER, GENDER, SIBLING SEX COMPOSITION,  
AND IDEAL FERTILITY (FULL SAMPLE)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Birth order	.060 (.006)***	.057 (.006)***	.058 (.006)***				
Male	.008 (.009)	-.005 (.009)	.025 (.012)**	.006 (.009)	-.035 (.017)**	-.036 (.009)***	.047 (.010)***
At least one elder male sibling		.073 (.025)***					
Male share of elder siblings			.043 (.022)*				
Male × male share of elder siblings			-.046 (.020)**				
1( $\Delta$ Ideal $\geq 0$ )				.061 (.012)***			
$\Delta$ Ideal				.036 (.004)***	.049 (.005)***		
$\Delta$ Ideal × 1( $\Delta$ Ideal $\geq 0$ )				.023 (.011)**			
Male × 1( $\Delta$ Ideal $\geq 0$ )					.058 (.017)***		
Male × $\Delta$ Ideal					-.011 (.005)**		
Male × $\Delta$ Ideal × 1( $\Delta$ Ideal $\geq 0$ )					.027 (.013)**		
1( $\Delta$ Ideal Sons $\geq 0$ )						.026 (.017)	
$\Delta$ IdealSons						.039 (.007)***	
$\Delta$ Ideal Sons × 1( $\Delta$ Ideal Sons $\geq 0$ )						.030 (.016)*	
1( $\Delta$ Ideal Daughters $\geq 0$ )							.017 (.020)
$\Delta$ Ideal Daughters							.045 (.014)***
$\Delta$ Ideal Daughters × 1( $\Delta$ Ideal Daughters $\geq 0$ )							-.016 (.018)
Hypothesis tested	1 and 2	3	3 and 4	5	6	7	8
R <sup>2</sup>	.05	.05	.05	.06	.06	.05	.03
Number of mothers	963	963	963	958	958	958	958
N	15,370	15,370	15,370	13,880	13,880	13,876	13,876

**Note.** The dependent variable is unwanted pregnancy, which equals 1 if the last child born in the last 5 years was not wanted then and equals 0 otherwise. All regressions include dummy variables for the mother's educational attainment, linear and quadratic controls for the child's and the mother's age, type of place of residence (rural/urban), and dummy variables for region of residence (provinces of Pakistan). Robust standard errors clustered by the mother are reported in parentheses. Sampling weights are used.

\* $p < .1$ .

\*\* $p < .05$ .

\*\*\* $p < .01$ .

**TABLE 12**  
UNWANTED PREGNANCIES BY BIRTH ORDER, GENDER, SIBLING SEX COMPOSITION,  
AND IDEAL FERTILITY (BREASTFEEDING DURATION SAMPLE)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Birth order	.066 (.006)***	.061 (.006)***	.064 (.006)***				
Male	-.014 (.013)	-.038 (.015)***	-.014 (.017)	-.016 (.014)	-.069 (.022)***	-.055 (.014)***	.024 (.015)
At least one elder male sibling		.084 (.027)***					
Male share of elder siblings			.039 (.026)				
Male × male share of elder siblings			.002 (.030)				
1( $\Delta$ Ideal $\geq 0$ )				.059 (.020)***			
$\Delta$ Ideal				.031 (.005)***	.051 (.006)***		
$\Delta$ Ideal × 1( $\Delta$ Ideal $\geq 0$ )				.042 (.013)***			
Male × 1( $\Delta$ Ideal $\geq 0$ )					.045 (.025)*		
Male × $\Delta$ Ideal					-.017 (.008)**		
Male × $\Delta$ Ideal × 1( $\Delta$ Ideal $\geq 0$ )					.049 (.016)***		
1( $\Delta$ Ideal Sons $\geq 0$ )						.016 (.022)	
$\Delta$ Ideal Sons						.030 (.008)***	
$\Delta$ Ideal Sons × 1( $\Delta$ Ideal Sons $\geq 0$ )						.057 (.016)***	
1( $\Delta$ Ideal Daughters $\geq 0$ )							-.000 (.026)
$\Delta$ Ideal Daughters							.058 (.015)***
$\Delta$ Ideal Daughters × 1 ( $\Delta$ Ideal Daughters $\geq 0$ )							-.027 (.019)
Hypothesis tested	1 and 2	3	3 and 4	5	6	7	8
R <sup>2</sup>	.06	.06	.06	.07	.07	.05	.03
Number of mothers	962	962	962	957	957	957	957
N	6,941	6,941	6,941	6,382	6,382	6,378	6,378

**Note.** The dependent variable is unwanted pregnancy, which equals 1 if the last child born in the last 5 years was not wanted then and equals 0 otherwise. All regressions include dummy variables for the mother's educational attainment, linear and quadratic controls for the child's and the mother's age, type of place of residence (rural/urban), and dummy variables for region of residence (provinces of Pakistan). Robust standard errors clustered by the mother are reported in parentheses. Sampling weights are used.

\*  $p < .1$ .

\*\*  $p < .05$ .

\*\*\*  $p < .01$ .

weans the unwanted child early to try once again to attain a child of her desired sex. The negative and significant coefficient in column 1 of table 10 shows that the second effect is overpowering. Column 2 further includes an interaction term of birth order and unwantedness of a pregnancy and provides support for both cases. An unwanted child is breastfed ( $-3.4 + 0.6 \times \text{Birth Order}$ ) months longer than a wanted child, keeping all else fixed. At a high enough birth order (exceeding 6), this difference is positive (i.e., unwanted children are breastfed longer only in families comprising more than six children). Since the average total fertility in Pakistan is 4.1 (less than 6), it is not surprising that the second case dominates overall, whereby mothers shorten breastfeeding duration to be able to achieve the desired sex composition of children.

Results from table 11 inform us as to whether revealed preference for sons (breastfeeding duration) matches with stated preference for sons (unwantedness of pregnancy). The dependent variable is unwanted pregnancy as defined earlier. The same set of standard controls are included to ensure a consistent comparison with earlier results, except for child age (in months) fixed effects since the dependent variable is now no longer censored. The total number of observations is now larger than before because the outcome variable on unwanted pregnancy is reported for a larger sample. Later, we reestimate the specifications using the same sample as that used in evaluating breastfeeding duration to ensure that the results are not driven by sample selection.

Column 1 shows that the probability of delaying or not wanting a child increases with birth order, just as breastfeeding duration increased monotonically with the number of children. However, the coefficient on the male dummy varies in statistical significance and sign across specifications. Column 2 reveals that a mother who has at least one son is significantly more likely to report a higher value of the outcome variable (i.e., not wanting to have a child now) compared to a mother who does not yet have a son. Similarly, the last pregnancy is more likely to be reported as unwanted as the share of elder male children increases from 0 to 1, but a girl born after all male siblings is felt less unwanted compared to a boy (col. 3). In column 4, we observe that pregnancies become increasingly more unwanted as the mother approaches her ideal fertility size. At the cut-off point,  $\Delta \text{Ideal}_i = 0$ , the probability of reporting a pregnancy as unwanted increases if the last child is a male (col. 5). The attainment of ideal male fertility is marginally significant in column 6 ( $p = .121$ ); however, the interaction term is positive and statistically significant at the 10% level, indicating that as the mother exceeds her ideal male fertility, the probability of reporting the last pregnancy as unwanted becomes increasingly higher. On the other hand, the probability of a pregnancy reported as unwanted remains unaffected by the attainment of the ideal number of daughters (col. 7).

All the results are entirely consistent with those obtained earlier using breastfeeding duration as our dependent variable, except for the coefficient on the male dummy. Note that the coefficient changes sign and significance across specifications. If mothers interpret the question on the unwantedness of pregnancy as directly measuring their gender preferences, they might misreport to avoid stating an explicit preference for sons over daughters. This could possibly explain why the sign of the coefficient on male dummy is unstable across specifications. On the other hand, the behavioral response on breastfeeding duration is less likely to be susceptible to this reporting bias. However, the stated preference for sons emerges when analyzing more indirect measures of gender preference, such as after conditioning on the share of elder male siblings and the attainment of ideal fertility. Those born after a high proportion of elder male siblings or after having attained the mother's ideal male fertility are increasingly more likely to be reported as unwanted pregnancies. Consistent with the trends in breastfeeding, we deduce that the stated preference for a son matches closely with the revealed preference, expressed behaviorally in the number of months a child is breastfed.

To ensure our results are not confounded by sample selection bias, we replicated them while restricting the sample to the earlier observations on breastfeeding. Table 12 shows that the results remain robust except for a few minor alterations. Now, the coefficient on male share is marginally significant ( $p = .134$ ), perhaps owing to a smaller sample size, and the interaction term in column 3 is not statistically significant, indicating that mothers no longer state the desire to achieve balance in their children's sex composition. The coefficient on the interaction term in column 6 is still statistically significant, revealing that as a mother exceeds her ideal male fertility, the probability of reporting the last pregnancy as unwanted becomes increasingly large. As before, this effect is missing in the case of having borne the ideal number of daughters.

Finally, one could think that the question of the ideal number of children measures stated fertility preferences from an *ex ante* point of view, while the wantedness question measures stated fertility preferences *ex post* (i.e., given your kids so far, did you want this pregnancy?). Without a behavioral model, it is difficult to assess what we should expect to find. However, our evidence indicates that the predictions using breastfeeding (revealed preference) are borne out when using wantedness (a measure of *ex post* stated preferences).

## VII. Conclusions

Recent evidence from a birth cohort study on nearly 6,000 babies followed over 3 decades suggests that prolonged breastfeeding is associated with increased IQ and higher educational attainment and earnings in adulthood (Victora et al.

2015). Therefore, breastfeeding not only has nutritional value in early childhood but also has long-term benefits for human capital formation later in life. Weaning daughters early therefore also has the potential to adversely affect their life achievement, further augmenting gender gaps in various spheres of life.

We show that breastfeeding duration in Pakistan increases monotonically with the birth order of the child. At every birth order, boys are breastfed longer than girls. Conditioning on the elder children's sex composition, a child is breastfed longer as the share of elder male siblings increases from 0 to 1. There is a marked discontinuous increase in overall breastfeeding duration, as well as in the male-female difference in breastfeeding, as a mother approaches her target family size. Furthermore, separate from the effect of attaining overall ideal fertility, there is a sharp increase in the duration of nursing as a mother also attains her ideal number of sons. All these findings are consistent with JK, using either the OLS or the hazard model as estimation techniques. However, estimates from Pakistan are almost twice as large as those from India, perhaps alluding to a stronger preference for sons, which is consistent with the findings of Arnold (1997) regarding parents' stated son preference: sons are preferred 10:1 in Pakistan and approximately 5:1 in India. The wider gender gap in Pakistan is comparable to what JK found in the northern states of India, where son preference is much stronger than in southern states. Since North India forms the geographical and cultural neighbor of East Pakistan (that comprises the bulk of its population), it is not surprising that estimates from Pakistan are more pronounced than overall estimates from India as a whole.

Following Chakravarty (2015), we further explore whether the effect of elder children's sex composition on breastfeeding duration depends on the gender of the current child. Finding that girls are in fact breastfed longer than boys as the fraction of elder male siblings increases from 0 to 1 suggests that parents want to achieve some balance in their children's sex composition. Furthermore, we also compare the effect of having the ideal number of sons against the ideal number of daughters but fail to find any significant change in breastfeeding duration in response to the attainment of ideal female fertility. This comparison provides a richer context to examine son preference.

Finally, we find similar birth-order patterns for the mother's subjective assessment of unwanted pregnancies as we do for breastfeeding duration, highlighting the relevance of stated preference-based measures as complements (or even substitutes) of revealed preference-based measures. Of course, the advantage of having information on breastfeeding cannot be dismissed. Mitigating gender bias in breastfeeding is a way to reduce gender inequalities in health. One benefit of breastfeeding, particularly in developing countries, is protection from unsanitary conditions. JK compared breastfeeding trends with child sur-

vival patterns in India and found that mortality responds in a similar fashion to birth order, gender, children's sex composition, and the attainment of ideal fertility. Unfortunately, due to lack of sufficient data on sanitary conditions, we have not been able to investigate the mortality impact of the breastfeeding gap.<sup>17</sup> A performance index released in 2015 reveals that Pakistan has left India behind in terms of improving water and sanitation access (Cronk et al. 2015). Whether the improvement is significant enough to overcome the detrimental effects from a more than twice as large breastfeeding gender gap remains unknown but is a potential avenue for future research.

## References

- Almond, Douglas, Lena Edlund, and Kevin Milligan. 2013. "Son Preference and the Persistence of Culture: Evidence from South and East Asian Immigrants to Canada." *Population and Development Review* 39, no.1:75–95.
- Anderson, Siwan. 2000. "The Economics of Dowry Payments in Pakistan." Discussion paper, Department of Economics, Tilburg University.
- Arnold, Fred. 1997. "Gender Preferences for Children." Demographic and Health Surveys Comparative Studies no. 23 (August), Macro International, Calverton, MD.
- Aslam, Monazza. 2009. "Education Gender Gaps in Pakistan: Is the Labor Market to Blame?" *Economic Development and Cultural Change* 57, no. 4:747–84.
- Bhat, P. N. Mari, Fred Arnold, Kamla Gupta, Sunita Kishor, Sulabha Parasuraman, P. Arokiasamy, S. K. Sindg, and H. Lhungdim. 2007. *National Family Health Survey (NFHS-3), 2005–06: India*. Vol. 1. Mumbai: International Institute for Population Sciences.
- Black, Robert E., Cesar G. Victora, Susan P. Walker, Zulfiqar A. Bhutta, Parul Christian, Mercedes De Onis, Majid Ezzati, et al. 2013. "Maternal and Child Undernutrition and Overweight in Low-Income and Middle-Income Countries." *Lancet* 382, no. 9890:427–51.
- Carranza, Eliana. 2012. "Islamic Inheritance Law, Son Preference and Fertility Behavior of Muslim Couples in Indonesia." Policy Research Working Paper no. 5972, World Bank, Washington, DC.
- Census of India. 2011. "Population by Religious Community." Government of India. <http://www.censusindia.gov.in/2011census/c-01.html>.
- Chakravarty, Abhishek. 2015. "Gender-Biased Breastfeeding in Egypt: Examining the Fertility Preference Hypotheses of Jayachandran and Kuziemko (2011)." *Journal of Applied Econometrics* 30, no. 5:848–55.
- Clark, Shelley. 2000. "Son Preference and Sex Composition of Children: Evidence from India." *Demography* 37, no. 1:95–108.
- Cronk, Ryan, Jeanne Luh, Benjamin Mason Meier, and Jamie Bartram. 2015. "The WASH Performance Index Report." The Water Institute, UNC Gillings School

<sup>17</sup> The source of drinking water is reported for only 14% of the households in our sample of the PDHS. JK use the response "piped into dwelling" on the source of drinking water as a proxy for access to clean water.

- of Global Public Health, Chapel Hill, NC. <http://waterinstitute.unc.edu/wash-performance-index-report>.
- Deininger, Klaus, Aparajita Goyal, and Hari Nagarajan. 2010. "Inheritance Law Reform and Women's Access to Capital: Evidence from India's Hindu Succession Act." Policy Research Working Paper no. 5338, World Bank, Washington, DC.
- Dyson, Tim, and Mick Moore. 1983. "On Kinship Structure, Female Autonomy, and Demographic Behavior in India." *Population and Development Review* 9, no. 1:35–60.
- El-Zanaty, Fatma, and Ann Way. 2006. *Egypt Demographic and Health Survey 2005*. Cairo: Ministry of Health and Population.
- Gupta, Monica Das, Jiang Zhenghua, Li Bohua, Xie Zhenming, Woojin Chung, and Bae Hwa-Ok. 2003. "Why Is Son Preference So Persistent in East and South Asia? A Cross-country Study of China, India and the Republic of Korea." *Journal of Development Studies* 40, no. 2:153–87.
- Hussain, R. 1999. "Community Perceptions of Reasons for Preference for Consanguineous Marriages in Pakistan." *Journal of Biosocial Science* 31, no. 4:449–61.
- Hussain, R., F. F. Fikree, and H. W. Berendes. 2000. "The Role of Son Preference in Reproductive Behaviour in Pakistan." *Bulletin of the World Health Organization* 78, no. 3:379–88.
- Iyer, Sriya. 2002. "Religion and the Decision to Use Contraception in India." *Journal for the Scientific Study of Religion* 41, no. 4:711–22.
- Jacoby, Hanan G., and Ghazala Mansuri. 2010. "Watta Satta: Bride Exchange and Women's Welfare in Rural Pakistan." *American Economic Review* 100, no. 4:1804–25.
- Jayachandran, Seema, and Ilyana Kuziemko. 2011. "Why Do Mothers Breastfeed Girls Less Than Boys? Evidence and Implications for Child Health in India." *Quarterly Journal of Economics* 126, no. 3:1485–538.
- Jejeebhoy, Shireen J., and Zeba A. Sathar. 2001. "Women's Autonomy in India and Pakistan: The Influence of Religion and Region." *Population and Development Review* 27, no. 4:687–712.
- Kazi, Afroze, Kathy I. Kennedy, Cynthia M. Visness, and Talat Khan. 1995. "Effectiveness of the Lactational Amenorrhea Method in Pakistan." *Fertility and Sterility* 64, no. 4:717–23.
- Kennedy, Kathy I., Roberto Rivera, and Alan S. McNeilly. 1989. "Consensus Statement on the Use of Breastfeeding as a Family Planning Method." *Contraception* 39:477–96.
- Morgan, S. Phillip, Sharon Stash, Herbert Smith, and Karen Oppenheim Mason. 2002. "Muslim and Non-Muslim Differences in Female Autonomy and Fertility: Evidence from Four Asian Countries." *Population and Development Review* 28:515–38.
- Murtaza, Niaz. 2014. "The Indian Within." *DAWN*, March 10. <http://www.dawn.com/news/1092218>.
- NIPS (National Institute of Population Studies [Pakistan]). 2008. *Pakistan Demographic and Health Survey 2006–07*. Islamabad: National Institute of Population Studies (Pakistan).
- Obermeyer, Carla Makhoul. 1992. "Islam, Women, and Politics: The Demography of Arab Countries." *Population and Development Review* 18, no. 1:33–60.



- Pande, Rohini P. 2003. "Selective Gender Differences in Childhood Nutrition and Immunization in Rural India: The Role of Siblings." *Demography* 40, no. 3:395–418.
- Pande, Rohini P., and Nan Marie Astone. 2007. "Explaining Son Preference in Rural India: The Independent Role of Structural versus Individual Factors." *Population Research and Policy Review* 26, no. 1:1–29.
- Pew Research Center. 2011. "The Future of the Global Muslim Population." Pew, January 27. <http://www.pewforum.org/2011/01/27/the-future-of-the-global-muslim-population>.
- Sathar, Zeba Ayesha, and Shahnaz Kazi. 2000. "Women's Autonomy in the Context of Rural Pakistan." *Pakistan Development Review* 39, no. 2:89–110.
- Srinivasan, Padma, and Gary R. Lee. 2004. "The Dowry System in Northern India: Women's Attitudes and Social Change." *Journal of Marriage and Family* 66, no. 5: 1108–17.
- Victora, Cesar G., Bernardo Lessa Horta, Christian Loret De Mola, Luciana Quevedo, Ricardo Tavares Pinheiro, Denise P. Gigante, Helen Goncalves, and Fernando C. Barros. 2015. "Association between Breastfeeding and Intelligence, Educational Attainment, and Income at 30 Years of Age: A Prospective Birth Cohort Study from Brazil." *Lancet Global Health* 3, no. 4:199–205.