Gender-biased fertility preferences may decrease fertility: evidence from a counterfactual analysis.*

Guilhem Cassan [†] Milan Van Steenvoort [‡]

October 24, 2023

Abstract

Gender-biased fertility preferences are prevalent in many societies. They are commonly thought to be a driver of fertility. We show here that this is not necessarily true: switching from gender-biased to gender-neutral fertility preferences can increase fertility. The magnitude and the sign of the variation in fertility depends on the choice of gender-neutral fertility counterfactual. We illustrate our findings using data from the Indian Demographic and Health Surveys and show that gender-biased fertility preferences can lead to a change ranging from -15% to 23% of total excess fertility or -4% to 6% of total fertility.

JEL : J13, J16
Keywords: stopping rule, gender discrimination, fertility

*We thank Jean Marie Baland and François Woitrin for useful suggestions. We thank the audiences of UNamur. Research on this project was financially supported by the Excellence of Science (EOS) Research project of FNRS O020918F.

[†]CEPR, DEFIPP, CRED, CAGE, CEPREMAP, University of Namur, Namur, Belgium
[‡]Maastricht University, Maastricht, Netherlands

1 Introduction

Excess fertility, a household's fertility extending beyond its desired family size, is prevalent in many societies. Gender-biased fertility preferences are often seen as one of the drivers of excess fertility. They are also viewed as slowing down fertility transition (Arnold, 1985; Cleland et al., 1983; Bairagi and Langsten, 1986; Chowdhury and Bairagi, 1990; Arnold, 1992; Clark, 2000).[1] It has indeed been demonstrated that gender-biased fertility preferences do lead to larger than desired families (Sheps, 1963; Yamaguchi, 1989; Basu and De Jong, 2010; Baland et al., 2023). Therefore, it is commonly assumed that gender-biased preferences must lead to an *increase* in excess fertility (Park, 1983; Arnold, 1985; Mutharayappa et al., 1997; Larsen et al., 1998). This is of particular importance because additional pregnancies are an important cost to mothers (Hardee et al., 2004; Bahk et al., 2015; Milazzo, 2018) and are associated with lower infant and child health outcomes (Gipson et al., 2008; Shah et al., 2011; Hall et al., 2017).

In this paper, we argue that this reasoning may not always be correct: gender-biased fertility may lead to larger than desired families without increasing excess fertility. In fact, we show that gender-biased fertility preferences can decrease rather than increase excess fertility. We make a simple logical point: for gender-biased preferences to increase excess fertility, excess fertility should be higher under gender-biased preferences than under gender-neutral preferences, all other things equal. That is, fertility under gender-biased preferences is to be compared to a counterfactual under which preferences are not gender-biased. The definition of a gender-neutral counterfactual is far from obvious. We show here that a counterfactual analysis proves that gender-biased preferences do not always increase excess fertility. In fact, under a plausible gender-neutral counterfactual, excess fertility can decrease because of gender-biased preferences.

We develop a simple framework and methodology to analyze the contribution of gendered preferences to excess fertility in various counterfactual scenarios. We implement this methodology to Indian data and show that gendered preferences contribute to total excess fertility, but that both the sign and the magnitude of this contribution are ambiguous. Our estimates range from a decrease of total excess fertility of 15% to an increase of 23%.

[1] Cleland et al. (1983) for example write " gender preferences may sustain higher levels of childbearing than would be the case if the sex of children was a matter of indifference".

For total fertility, the respective proportions are -4% and +6%.

We decompose gender-biased preferences into two components: *composition* and *intensity*.[2] The *composition* component relates to cases in which parents want strictly more boys than girls.[3] The main intuition behind the composition effect is that the probability of reaching the desired gender composition of children at the desired family size depends on the gender imbalance of the desired composition. For a given desired family size, the more imbalanced the desired gender composition, the lower the probability of reaching that composition at the desired family size (Sheps, 1963). For instance, think of parents whose desired total number of children is two. If they are indifferent to the gender composition of their children, then they will reach their ideal gender composition in 100% of cases. If they instead want one son and one daughter, they have a 50% chance of reaching their desired gender composition at their desired family size. And if they want two sons, they only have a 25% chance of reaching their desired number of sons upon reaching their desired family size (under the simplifying assumption that the probability to have a son is 50%). So, at the society level, the distribution of desired gender compositions will determine the proportion of families who do not reach their desired gender composition upon reaching their desired family size and are in a position to have excess fertility. Note the ambiguity in the definition of a gender-unbiased fertility composition. Would unbiased parents want the same number of each gender, or would they be indifferent to the gender of their children (Williamson, 1976)? Nevertheless, irrespective of the definition of the gender-unbiased counterfactual, the proportion of families who do not reach their desired gender composition upon reaching their desired family size and are in position to have excess fertility is always higher when the composition of preferences is biased, but the extent to which this proportion differs from a gender-unbiased counterfactual depends on the definition of the later.

The *intensity* component relates to the difference in parents' willingness to go above their desired fertility when, having reached their desired family size, they have not reached their desired number of sons compared to when they have not reached their desired number of daughters. Think of the example above of a family desiring one son and one daughter. While

[2]Note here that these concepts are different from the decomposition proposed by Coombs et al. (1975) into a preference for a family size and a preference for a gender composition.

[3]More precisely, it relates to the case in which parents want a larger proportion of boys than the natural probability for a boy to be born.

the composition of their preferences may not be biased, their intensity may be if they are more willing to have additional children in case they do not reach their desired number of sons compared to when they do not reach their desired number of daughters. Note here that it is again not obvious what a non-biased behaviour would be. Indeed, in an unbiased world, the intensity should be equal for sons and daughters. But should the willingness to pursue fertility when a daughter is missing increase to that of when a son is missing? Or, to the contrary, should that linked to a son being missing decrease to that of daughters? Both scenarios[4] lead to very different consequences for excess fertility. In the first case, excess fertility will increase, while in the second, it will decrease.[5]

We believe that this decomposition of gender-biased preferences clarifies the counterfactual reasoning necessary to assess the impact of gender-biased preferences on excess fertility. We develop an analysis of the consequences of the removal of gender-biased preferences on excess fertility. We take a counterfactual approach: we compute excess fertility in a gender-unbiased world, and compare it to excess fertility under observed (gender-biased) preferences, using Indian data on gender preferences and fertility as an illustration. As discussed above, there is ambiguity about the definition of gender-unbiased preferences. In terms of *composition*, would parents want the same number of each gender, or would they be indifferent (Williamson, 1976)? In terms of intensity, would parents want to increase the intensity of their preference for daughters to that for sons, or the contrary? These different definitions of gender unbiased preferences lead to vastly different conclusions in terms of excess fertility. We scrutinize the different possible combinations of counterfactuals, and show that the conclusions on the link between excess fertility and gender-biased fertility preferences are extremely sensitive to the choice of counterfactual.

Our paper contributes to the very rich literature on the link between gender-biased preferences and fertility behavior. Arnold (1985); Chowdhury et al. (1993) study fertility choices of families with different compositions of already born children at a given family size. They estimate a counterfactual gender-neutral fertility by attributing the lowest observed fertility

[4]There is of course a continuity of alternative possibilities in between these two extremes

[5]Note here that this point is different from the one underlined by McClelland (1979); Widmer et al. (1981) which discuss the fact that sex biased preferences may decrease fertility if the expected disutility to have a child of the undesired gender is higher than the expected utility to have a child of the desired gender.

behavior to all families within a family size cell. To our knowledge, these studies were the first to explicitly compute counterfactual gender-neutral fertility. We improve on this approach by distinguishing the two components of gender preferences and their specific contribution to excess fertility, and by underlining that attributing the lowest observed fertility is one extreme assumption among a very large set of alternatives. Several studies measure the extent to which families continue to have children at a given family size when they seem to be missing boys with respect to their desired gender composition (Repetto, 1972; Park, 1976, 1983; Das, 1987; Altindag, 2016; Kim and Lee, 2020). Clark (2000) predicts how the distribution of gender composition of families should look like in India under generalized gender-biased preferences and shows that observed distributions of gender composition indeed align with the theoretical predictions. However, there is in general no mention of what fertility should look like in a gender-neutral counterfactual. A more theoretical literature has also studied mechanisms linked to gender-biased fertility practices. For instance, Yamaguchi (1989) and Basu and De Jong (2010) show that families tend to be larger under gender-biased preferences, as daughters will have more siblings. Baland et al. (2023), in their study of the stopping rule, show that daughters tend to have more younger siblings than sons. Sheps (1963) influential study has demonstrated how gender bias in the desired composition may lead to excess fertility. Those efforts do not distinguish between the *composition* or *intensity* mechanism, nor are they comparing excess fertility under gender-biased preferences to a gender-unbiased counterfactual. They therefore can not inform about the effect of gender-biased preferences on excess fertility.

This paper is organised as follows. Section 2 presents our simple framework of gender preferences and discusses the implications of the definition of the gender-neutral counterfactuals for excess fertility. Sections 3 and 4 take this framework to the data. Using Indian DHS data, we show the extent to which estimates of excess fertility can vary depending on the definition of the gender-neutral counterfactual. Section 5 analyses the evolution of the effect of gender-biased fertility preferences on excess fertility over time. Section 6 concludes.

2 A simple framework of gender preferences

In the following, we discuss from a theoretical perspective what gender-biased composition and gender-biased intensity mean and what they imply in terms of excess fertility. We make explicit how excess fertility would change if fertility preferences were unbiased. In particular, we illustrate that the way in which"unbiased" preferences are defined drastically matters. In several realistic counterfactual scenarios, the removal of the gender bias in fertility preferences leads to an increase in excess fertility.

We describe here the simple framework of gender composition and fertility preferences used in the paper. Families have preferences over an ideal number of boys $b^* \geq 0$, an ideal number of girls $g^* \geq 0$ and an ideal family size $n^* \geq b^* + g^*$. Family can grow to a maximum family size of $\bar{N} \geq n^*$. That is, as in one of the models of Baland et al. (2023), families have lexicographic preferences in number of boys, girls and family size. Ideal family size, ideal number of boys and ideal number of girls are ranked in terms of priority by the family. Ideally, families want to reach b^* boys and g^* girls in n^* births. If they favor the ideal number of sons to the ideal family size, they will be willing to go beyond n^* births if the number b of boys born among the first n^* births is below b^*. In this framework, we call $max(N - n^*, 0)$ the excess fertility, with N the total number of births upon fertility completion. That is, excess fertility is the number of births in excess of the ideal number of births. We now detail how gender-biased composition and gender-biased intensity can affect excess fertility.

2.1 Gender-biased composition of fertility preferences

The intuition of the impact of gender-biased composition on excess fertility is as follows: the probability of reaching the desired gender composition of children at the desired family size depends on the gender imbalance of the desired composition. The more imbalanced towards a certain gender the composition is, the higher the probability of not reaching the desired number of children of that gender at the desired family size. This is because the more balanced the gender preferences, the larger the probability to obtain the desired gender composition at n^*.

Think of the gender of children as a lotery and of n^* as the number of draws in that lotery, with p the probability to have a boy. The closer b^*/g^* is to $p/(1-p)$, the higher the probability to reach b^* and g^* in n^* draws. That is, if the desired gender composition is close to the natural gender composition, parents maximize their probability to obtain their desired gender composition in the minimum number of draws. The further away b^*/g^* is from $p/(1-p)$ - i.e., the more gender-biased the desired composition is - the less likely parents are to attain their desired gender-composition at their desired family size.

Take the case where $n^* = 2$ and $p = 0.5$. Families have lexicographic preferences over family size and gender composition, and always favour gender composition over family size.[6] We assume away gender-biased intensity - which will be discussed in the next subsection - by imposing that families who do not reach either b^* or g^* among the first n^* births are willing to have an infinite number of births in order to have a child of the desired gender. That, is, with $p = 0.5$, families who do not reach their desired gender composition at rank 2 will on average have an excess fertility of 2 for each child missing for the ideal number of children of each sex to be born.[7] Note that there are 3 possible gender compositions at rank 2: {bg, gg, bb} with probabilities {0.50, 0.25, 0.25}.

Suppose that families have *gender-balanced* preferences regarding the gender composition of their children: $\{b^*, g^*\} = \{1, 1\}$. At rank 2, 50% of families have reached their desired gender composition and stop their fertility, while 50% of families are short of one gender, and will have on average 2 additional births. Overall, in this society, the average number of children per family is 3, and each family has on average 1 excess birth.

Suppose now, instead, that families have *gender-biased* preferences towards boys, such that $\{b^*, g^*\} = \{2, 0\}$. That is, families want a family of two boys instead of a family of one boy and one girl. At $n^* = 2$, the families who have {bb} will stop having children since they reached their desired gender composition of children and, hence, 25% of the families will have two children. However, families who have {bg} or {gg} at $n^* = 2$ have not reached their desired number of boys. The 25% of {bg} families will have, on average, four children. The 25% of {gg} families will have, on average, six children. As a result, in this society, the average number of children per family is 3.5, and each family has on average 1.5 excess births.

The average number of children per family differs depending on the gender composition of the fertility preferences. On average, each family has 1 excess birth when the desired composition is gender-balanced and 1.5 excess births when the desired composition is gender-biased. Note that in both scenarios, there are excess births. Hence, the number of excess births that can be attributed to the gender-biased composition is only the difference between the two scenarios, that is, only 0.5 excess births can be attributed to the gender-biased composition.

[6]This is obviously not always the case in real life. What matters is that, on average, a sufficiently large share of families favour gender composition over family size for the pattern that we discuss here to emerge on aggregate.

[7]See Yamaguchi (1989); Baland et al. (2023).

However, note how we implicitly defined the gender-neutral composition of fertility preferences as gender-balanced preferences, that is, a desire for a gender composition corresponding to the natural sex ratio. An alternative definition for a gender-neutral composition of fertility preferences is a situation in which parents are indifferent to the gender of their children. In such a situation, all families would always reach their desired gender composition at their desired family size. Under this alternative counterfactual, 1.5 excess births can be attributed to gender-biased composition. That is, the number of excess births attributed to gender-biased composition increases by 200% when moving from a gender-balanced counterfactual to a gender-indifferent counterfactual.

2.2 Gender-biased intensity of fertility preferences

We detail here the intuition behind gender-biased intensity's impact on excess fertility.

Take the case of a family having reached its ideal family size but not its ideal gender composition. This family can either stop having children, thereby not reaching its ideal gender composition, but staying at its ideal family size; or the opposite. Gender-biased intensity of preferences means that families are more likely to favor gender composition over ideal family size if boys rather than girls are missing.[8]

Take again the case of families who want the same number of boys b^* and of girls g^*, for an ideal family size of n^*. For $n^* = 2$, such families want one boy and one girl. As above, for a probability $p = 0.5$ of having a boy, the composition of children is {bg, gg, bb} with probabilities {0.50, 0.25, 0.25}.

As discussed, if such families have gender-neutral intensity, such that they are willing to have an infinite number of children to reach the number of boys and girls that they desire, excess fertility is on average 1. Suppose now that families have gender-biased intensity of preferences. Only families missing a boy are willing to have an infinite number of children, while fam-

[8]Note here that this example assumes that, in expected terms, the benefit of having a child of the desired gender more than compensates for the cost of having a child of the undesired gender. As discussed in McClelland (1979), when this is not the case, gender-biased intensity would lead families missing the gender they prefer to pursue fertility less often than families missing the gender they favour less. In our framework, we take a revealed preferences approach, where what matters is the actual behaviour. That is, we would classify such families as having a gender-biased intensity towards girls rather than boys, since they pursue their fertility to reach their desired number of girls more often than to reach their desired number of boys.

ilies missing a girl prefer to stop at their ideal family size of 2. Now, only 25% of families will reach a size of 4 while 75% have a family size of 2, for an average family size of 2.3. On average, in the population, excess fertility is 0.3. Finally, let's now assume that families prefer staying at their ideal family size compared to reaching their ideal number of children of any gender. This is an other case of unbiased intensity. Here, however, the average family size is 2, and excess fertility is 0.

So what is the excess fertility caused by gender-biased intensity? In the first counterfactual, excess fertility effectively decreases because of gender-biased intensity, from 1 to 0.3. In the second counterfactual, excess fertility increases from 0 to 0.3. The definition of the gender-neutral counterfactual for intensity therefore not only changes the magnitude of excess fertility, but it can also change its sign.

It is not obvious which counterfactual should be favored over the other: in a gender-neutral intensity world, should the propensity to continue having children when the ideal number of girls is not attained "increase" to that of boys, or should the propensity to continue having children when the ideal number of boys is not attained "decrease" to that of girls? Both counterfactuals seem equally reasonable, but have vastly different consequences in terms of excess fertility attributable to gender-biased preferences. To the best of our knowledge, the existing literature did not consider this channel. This may be due to the fact that societies in which gender preferences become more equal often also see total fertility decrease. However, ceteris paribus, this channel may lead to more rather than less excess fertility when gender preferences equalize.

3 A decomposition of excess fertility by type of gender bias

We discuss here how we estimate the contribution of gender bias to excess fertility. We explain how we decompose excess fertility into its intensity and composition components, as well as how we define alternative gender-neutral counterfactuals. We use real world data to illustrate how the definition of gender-neutral composition and intensity can profoundly alter our understanding of the role of gender-biased preferences in explaining excess fertility and get a sense of the magnitudes involved.

3.1 Data

We use data from the Indian Demographic and Health Surveys (DHS) of 1992, 1998, 2006 and 2015, to illustrate the gender-biased intensity and composition effects on excess fertility. These surveys document, for every interviewed woman, her birth history up to the date of the interview, including the birth order and gender of each child. In addition, it elicits the desired family size as well as the desired gender composition of children. Moreover, the questionnaire contains a question asking each woman if she desires to continue her fertility at the time of the interview. We restrict our sample to women who declare desiring no more children[9]. Table 1 reports the descriptive statistics. Our sample consists of 254,922 women. On average, each woman gave birth to 3.15 children and desired an average of 2.47 children - among which, on average, 1.26 boys, 0.99 girls and 0.23 children of either gender. 87 % of all women in our sample reached or surpassed their desired family size. These data provide us with information on desired fertility, desired gender composition and actual fertility and gender composition. This allows us to measure the extent to which fertility choices are affected by gender-biased composition preferences and gender-biased intensity, and compute counterfactuals for what excess fertility would look like under various definitions of gender-neutrality of preferences. We describe our methodology below. Note that we are well aware that declared fertility preferences may suffer from bias. In particular, mothers may tend to declare that their actual fertility corresponds to their desired fertility. We do not think this is of particular concern. Indeed, this bias would work against finding any gender-bias in declared preferences and against finding any effect of gender-biased preferences on actual fertility. That, despite this bias, we do find important magnitude speaks to the importance of the mechanisms that we discuss.

[Table 1 about here.]

3.2 Decomposition of excess fertility by cause

The approach that we follow is akin to a Blinder-Oaxaca decomposition (Blinder, 1973; Oaxaca, 1973). We decompose excess fertility into two com-

[9]Note that our analysis can also be performed using the preferences of all women in the DHS. However, our restriction allows us to provide a more meaningful order of magnitude of our effects: we estimate the excess fertility due to gender-biased preferences for women who declared having finished their fertility and compare it to the total excess fertility of these women.

ponents: gender-biased composition of children (which affects the probability to reach the ideal gender composition at n^*) and gender-biased intensity (which affects the propensity to continue having children above ideal family size conditional on not having reached the ideal gender composition at n^*).

To isolate these components, we model excess fertility as driven by families who have not reached their desired gender composition of children at n^*. The prevalence of these families in the population will be driven by the prevalence of gender-biased composition. We differentiate between families who have not reached their desired number of boys and families who have not reached their desired number of girls, therefore allowing gender-intensity to affect excess fertility.[10]

Excess fertility can therefore be written as:

$$Excess\_Fertility_{i,c} = \quad \alpha_i * B_c + \beta_i * G_c \tag{1}$$

With $Excess\_Fertility_{i,c}$ the total number of births in excess of the desired family size in the society for an intensity i and a composition c. B_c (G_c) is the number of families who have not reached their desired number of boys (girls) at n^* for a composition c. α_i (β_i) is the average number of births in excess of n^* that families have when they did not reach their ideal number of boys (girls). i and c can be either unbiased (u) or biased (b). There is a similarity between this approach and the Blinder-Oaxaca methodology: excess fertility is decomposed into difference in average characteristics (B_c and G_c) and differences in "returns" to those characteristics (α_i and β_i).

Note that if families have gender-neutral intensity, then $\alpha_i = \beta_i = \bar{k}$. Therefore, when families have a gender-biased intensity, $\alpha_i = \beta_i + \gamma$, where γ measures the extent to which families will go above their desired fertility more when boys are missing compared to girls. As discussed in the previous section, it is unclear whether, under gender-neutral intensity, parents would want to decrease the intensity of their preference for boys to that for girls, or the contrary? That is, in the first intensity counterfactual, $\bar{k} = \beta_i$, whereas in the second intensity counterfactual, $\bar{k} = \alpha_i = \beta_i + \gamma$. Hence, excess fertility under gender-neutral intensity and gender-neutral composition can be written:

[10] We leave out considerations related to other causes of excess fertility, such as child mortality, access to contraceptive methods or different preferences across parents. While these are obviously important mechanisms, they are of no importance in the mechanisms we discuss.

$$Excess\_Fertility_{u,u} = \quad \bar{k} * B_u + \bar{k} * G_u. \tag{2}$$

Excess fertility under gender-biased intensity and gender-neutral composition can be written:

$$Excess\_Fertility_{b,u} = \quad (\beta_i + \gamma) * B_u + \beta_i * G_u \tag{3}$$

In this framework, the excess fertility caused by gender bias in *intensity* is $Excess\_Fertility_{b,u} - Excess\_Fertility_{u,u}$. With $\bar{k} = \beta_i$, this equals $\gamma * B_u$. But with $\bar{k} = \alpha_i = \beta_i + \gamma$, this equals to $-\gamma * G_u$. That is, depending on the counterfactual, the total effect may change sign.

Similarly, excess fertility under gender-unbiased intensity and gender-biased composition can be written as:

$$Excess\_Fertility_{u,b} = \quad \bar{k} * B_b + \bar{k} * G_b \tag{4}$$

Here, $Excess\_Fertility_{u,b} - Excess\_Fertility_{u,u} = \bar{k} * (B_b - B_u + G_b - G_u)$ is the excess fertility caused by gender bias in *composition*.

Finally, excess fertility under both gender-biased composition and gender-biased intensity is given by:

$$Excess\_Fertility_{b,b} = \quad (\beta_i + \gamma) * B_b + \beta_i * G_b \tag{5}$$

In the first intensity counterfactual ($\bar{k} = \beta_i$), excess fertility caused by both types of biases can therefore be written as:

$$\begin{aligned} Excess\_Fertility_{b,b} - Excess\_Fertility_{u,u} = \quad & \underbrace{\beta_i * (B_b - B_u + G_b - G_u)}_{Composition\,Effect} \\ + \quad & \underbrace{\gamma * B_u}_{Intensity\,Effect} \\ + \quad & \underbrace{\gamma * (B_b - B_u)}_{Interaction\,Effect} \end{aligned} \tag{6}$$

In the second intensity counterfactual ($\bar{k} = \alpha_i = \beta_i + \gamma$), excess fertility caused by both types of biases can be written as:

$$
\begin{aligned}
Excess\_Fertility_{b,b} - Excess\_Fertility_{u,u} = \ & \underbrace{(\beta_i + \gamma) * (B_b - B_u + G_b - G_u)}_{Composition\ Effect} \\
& - \underbrace{\gamma * G_u}_{Intensity\ Effect} \\
& - \underbrace{\gamma * (G_b - G_u)}_{Interaction\ Effect}
\end{aligned} \tag{7}
$$

Note that these two expressions are not the sum of the pure intensity effect of Decomposition 3 and the pure composition effect of Decomposition 4. As in any Oaxaca-Blinder decomposition, a third term is present, $\gamma * (B_b - B_u)$ or $-\gamma * (G_b - G_u)$: it measures the excess births due to the gender-biased intensity for the additional number of families not reaching their desired number of boys (girls) because of the gender-biased composition. We refer to this mechanism as the *interaction effect*.

To compute counterfactuals based on this simple methodology, we therefore require β_i, γ, B_b, B_u, G_b and G_u. The following section discusses how we compute them.

3.2.1 Measuring preferences over composition

We turn to real world data to measure gender-biased preferences.

To measure B_i and G_i, we rely on the ideal gender composition and family size declared by mothers in the DHS. That is, our data gives b^*, g^* and n^* for each family. For each triplet $\{b^*,\ g^*,\ n^*\}$ we can compute the probability to reach both b^* and g^* at the ideal family size n^*.[11] For example, for $\{b^*, g^*, n^*\} = \{1, 1, 2\}$, 49.9% of families would reach their ideal gender composition at their ideal family size. Among the 50.1% remaining, 27 percentage points are missing a girl and 23 percentage points are missing a boy.

Figure 1 presents how preferences are distributed in our sample for families who desire up to 6 children ($\geq$ 99% of our sample). At every ideal

[11] To compute these probabilities, we need an additional assumption on p, the probability that a birth is a male. We take the sex ratio at birth observed in the data: 52%.

family size, the ideal proportion of boys is always larger than than that of girls. This is particularly striking for families desiring only one child, of which only 11.04% desire a girl, but this remains true at all desired family size: the share of desired girls oscillates between 35 and 45 percent for ideal family sizes of two onwards, while the desired share of boys oscillates between 45 and 60 percent.

Based on this data, we compute the number of families B_b and G_b who do not reach b^* or g^* at family size n^* by simply doing a weighted sum of probabilities. In total, 27.31% of families do not reach their desired number of boys (b^*) and 20.60% of families do not reach their desired number of girls (g^*) at their desired family size (n^*). These will be our gender-biased composition baseline.

[Figure 1 about here.]

3.2.2 Defining gender-unbiased composition

There are two possible definitions of gender-unbiased composition. First, families can be gender-indifferent: the only objective is n^*. Therefore, gender indifferent families always reach their ideal gender composition at n^*, $B_u = G_u = 0$, and there is no excess fertility. Second, families may want an equal number of boys and girls, holding n^* constant. To compute excess fertility in this second counterfactual, we hold the distribution of n^* constant, but impose gender balance in the number of children for which parents express a preference over the gender.[12] We can then compute B_u and G_u. Both will take positive values: because parents have a specific desired gender composition, some of them will not reach that composition upon reaching n^*.

Table 2 presents the share of families not reaching their desired gender composition at n^*, by definition of preferences. With gender-biased composition as observed in the data, 27.31% of all families do not reach their desired number of boys and 20.60% do not reach their desired number of girls. The percentage of families not reaching their desired number of boys decreases to 22.03% if preferences were gender-balanced. However, the percentage of families not reaching their desired number of girls increases to 25.21% with

[12] If $b^* + g^*$ is even, $b^* = g^* = (b^* + g^*)/2$. If $b^* + g^*$ is uneven, 50% of families desire $b^* = (b^* + g^*)/2 + 0.5$ and $g^* = (b^* + g^*)/2 - 0.5$ and 50% of families desire the opposite. In Appendix A.1, we define a plausible alternative definition of balanced fertility preferences and reproduce our analysis. The implications remain qualitatively identical.

gender-balanced composition. Interestingly, the total proportion of families not reaching their desired gender composition of children (either boys or girls) does not change substantially. This is due to two counteracting factors. On the one hand, families with gender-balanced composition desire less boys on average, which increases the probability of reaching the number of desired boys. On the other hand, families with gender-balanced composition want a higher number of girls, making them less likely to attain this objective. Finally, with gender-indifferent composition, all families reach their desired composition.

[Table 2 about here.]

3.2.3 Measuring preferences over intensity and defining gender unbiased intensity

We now turn to the estimation of gender-biased intensity of fertility preferences (γ in our model). In contrast to preferences over composition, we do not directly observe this information in the data. However, we do observe the actual fertility behaviour, on top of the ideal fertility. We can therefore measure the extent to which families go beyond n^* when they do not reach b^* compared to when they do not reach g^*.[13]

We run the following regression:

$$\begin{aligned} Excess\_Fertility_i = \beta_0 \quad & +\beta BirthMissing_i \\ & +\gamma BirthMissing_i * BoyMissing_i \\ & +\delta_i + \omega_i + \epsilon_i \end{aligned} \tag{8}$$

$Excess\_Fertility_i$ is the number of births after the last desired birth n_i^* for mother i, $BirthMissing_i$ is a dummy that equals 1 if the mother has not reached her desired sex composition of children at the last desired birth n_i^*, $BoyMissing_i$ is a dummy that equals 1 if the mother has not reached her desired number of boys. δ_i comprises fixed effects for the year of birth and the birth order of the child. ω_i comprises fixed effects for the year of birth of the mother, the age of mother at birth n_i^*, the educational attainment, the religion of the mother, the state in which the mother lives and whether the mother belongs to a scheduled tribe or a scheduled caste. Standard errors are clustered at the primary sampling unit and the regression is weighted using the sample weight of each mother. We restrict our sample to all births

[13] Note that at n^*, by definition, parents reach their desired number of children for at least one gender.

that occurred at least 80 months[14] before the date of the interview to allow sufficient time to observe a subsequent birth.

The coefficients of interest are β and γ. β measures the number of additional births that mothers on average have when they did not reach their ideal number of girls. γ is the additional number births that mothers on average have when they did not reach their ideal number of boys rather than girls.

[Table 3 about here.]

Table 3 reports the results from equation 8. Families not reaching their desired number of girls have, on average, $\hat{\beta} = 0.045$ (p-value < 0.001) additional births. We find that families not reaching their desired number of boys have, on average, $\hat{\gamma} = 0.649$ (p-value < 0.001) more subsequent births, compared to mothers not reaching their desired number of girls.

We can directly plug these results in our model to measure gender-biased intensity as $\alpha_i = \hat{\beta} + \hat{\gamma}$ and $\beta_i = \hat{\beta}$. In addition, we can also use them to compute gender-neutral counterfactuals. As already discussed, a gender-neutral intensity is such that the number of additional births when b^* is not reached is the same as when g^* is not reached. These results therefore bound the potential counterfactual such that $\hat{\beta} \leq \alpha_i = \beta_i = \bar{k} \leq \hat{\beta} + \hat{\gamma}$. In other words, there is a continuum of gender-neutral intensity counterfactuals, but they are bounded by our results. The lower bound is such that the intensity for both genders is "lowered" to that of girls: $\alpha_i = \beta_i = \hat{\beta}$. The upper bound consists in "increasing" intensity for both genders to that of boys $\alpha_i = \beta_i = \hat{\beta} + \hat{\gamma}$. We use these two bounds as counterfactuals for gender-neutral intensity.

4 Counterfactual analysis

We can now use the model and preferences presented in the previous section to estimate the impact of gender-biased fertility preferences on excess fertility. We have two counterfactuals for each type of bias: gender-balanced and gender-indifferent for composition and "boys to girls" and "girls to boys" for intensity. We therefore have four counterfactuals for excess fertility under gender-neutral preferences that we can compare to the excess fertility obtained under the observed, gender-biased, preferences. That is, we calcu-

[14] 95^{th} percentile of the interval between the last and second to last birth.

late:
$$\frac{ExcessFertility_{b,b} - ExcessFertility_{u,u}}{Fertility}, \tag{9}$$

Where $Fertility$ is either the total fertility or the total excess fertility in our sample. Figure 2 presents how much excess fertility can be attributed to gender-biased preferences based on the different counterfactual used, expressed as a share of the total (excess) fertility of our sample. We can see that the effect is highly dependant on the choice of counterfactual. When we define gender-neutral preferences as having a balanced composition and increase the intensity for girls to that of boys, the excess fertility that can be attributed to gender-biased preferences is negative: total excess fertility is in fact 15.51% lower with gender-biased preferences than what it would be under gender balanced composition and "girls to boys" intensity. On the contrary, when we define gender-neutral preferences as having a balanced composition and decrease the intensity for boys to that of girls, the excess fertility that can be attributed to gender-biased preferences is positive - total excess fertility is increased by 21.33% when moving from a gender balanced - boys to girls counterfactual to the observed biased preferences. The highest effect is found when gender-neutral families are indifferent with respect to the composition, with a 23.92% increase in total excess fertility. Obviously, this effect is independent of the choice of intensity counterfactual. Note also how the magnitude does not change much between the indifferent counterfactuals and the balanced - boys to girls counterfactual. This underlines the importance of the intensity component in determining both the size and the sign of the effect of gender-biased preferences. Finally, when referring to total fertility rather than total excess fertility, the respective shares are -4.09%, 5.62% and 6.31%.

[Figure 2 about here.]

We now decompose these results into intensity, composition and interaction effects. Table 4 below presents how we compute each of these based on the different definitions of our counterfactuals. In the remainder of this paper, we express all our results as a share of the total excess fertility in our sample. Naturally, the interpretations are identical when the results are expressed as a share of total fertility (only the magnitudes are affected since total fertility is by definition larger than total excess fertility).

[Table 4 about here.]

Figure 3 displays this decomposition of excess fertility.

[Figure 3 about here.]

Composition counterfactual defined as balanced. Under this counterfactual for composition, the conclusion vastly differs depending on the definition of the intensity counterfactual. Indeed, if the intensity counterfactual is defined as "girls to boys", then excess fertility is *higher* in the gender-neutral counterfactual. In fact, the gap between the excess fertility without and with biased preferences is equivalent to 15.51% of the total excess fertility. In other words, excess fertility decreases thanks to gender-biased preferences. When intensity is defined as "boys to girls", then the conclusion is opposite, and gender-biased preferences increase excess fertility by an amount equivalent to 21.33% of the total excess fertility.

The component driving the ambiguity of results is the intensity mechanism. Its contribution changes from +17.18% to -19.66% of the total excess fertility. It is the component with the largest contribution to the total effect, and the only component whose sign changes depending on the counterfactual. Both the interaction and the composition effect have unambiguous effects on excess fertility. The interaction effect always imply that gender bias leads to excess fertility. Even in the "girls to boys" counterfactual, the interaction effect shows that there are less excess births in a gender-neutral counterfactual. Indeed, when gender composition is balanced, a larger share of families do not reach their ideal number of girls than when gender composition is biased ($G_u > G_b$). Therefore, if intensity is set to "girls to boys" the interaction component is positive (3.60%). This is also true if the intensity is set to "boys to girls" (4.12%), as there are less families not reaching their ideal number of boys in the gender-neutral counterfactual ($B_u < B_b$).
The composition component is also always positive and is of very small magnitude (0.55% and 0.04% if the intensity is set to "girls to boys" and "boys to girls", respectively): the total proportion of families not reaching their gender composition ($B_i + G_i$) is not strongly affected, even though the gender distribution of missing births changes ($G_u > G_b$ and $B_u < B_b$).

Composition counterfactual defined as indifferent. Under this alternative composition counterfactual, we would draw very different lessons. Irrespective of the intensity counterfactual, we estimate that gender-biased preferences lead to an increase of 23.92% of total excess fertility. In addition, the decomposition of this effect across components differs substantially. First, by definition, there is no intensity effect: in this counterfactual, $G_u = B_u = 0$, so the pure intensity effect is always zero. Second, the composition component is orders of magnitude larger than under the balanced composition counterfactual. This is, again, because $G_u = B_u = 0$. But its magnitude is also drastically affected by the definition of the intensity counterfactual (an increase in excess fertility equivalent to 2.62% and 39.98% of the total excess fertility if the intensity is set to “boys to girls” and “girls to boys”, respectively). Under the “girls to boys” counterfactual, the families not reaching their desired composition have many additional births, while the opposite is true under the “boys to girls” counterfactual. Finally, the interaction effect is orders of magnitude larger under the indifferent composition than under the balanced composition. In addition, and in contrast to the balanced composition counterfactual, its sign changes as a function of the intensity counterfactual (a change in excess fertility equivalent to +21.30% and -16.06% of the total excess fertility if the intensity is set to “boys to girls” and “girls to boys”, respectively), since there is always a much larger share of families not reaching their desired gender composition, irrespective of the gender, under the biased composition.

5 Evolution over time

In the previous section, the parameters of our model were not allowed to change over time. We now relax this assumption. This allows us to understand how excess fertility has evolved. We define four cohorts: mothers born before 1960, mothers born between 1960 and 1969, mothers born between 1970 and 1979 and mothers born after 1979. We reproduce the previous methodology for each cohort.

First, we estimate the evolution of the gender differential in the intensity of fertility preferences. We run the OLS regression from Equation 8 for each cohort. Figure 4 displays, for each cohort, the $\hat{\beta}$ and $\hat{\gamma}$ coefficients. We observe a continuous increase in the gender differential ($\hat{\gamma}$) over time, whereas the intensity for girls ($\hat{\beta}$) is decreasing over time. That is, the bias in intensity is growing over time (see Appendix A.2 for respective estimation results).

[Figure 4 about here.]

Second, we estimate the evolution of the distribution of families not reaching both b^* and g^* at n^*. We simulate the distribution based on the declared preferences in the same manner as previously[15].

[15]However, we estimate the sex ratio at birth for each cohort and determine the proba-

Figure 5 presents the evolution of the share of families not reaching their desired composition. We observe that the proportion not reaching the desired number of boys decreased (see Appendix A.3 for a detailed analysis of this evolution), whereas the opposite is true for girls. We observe a slight decrease in the number of families not reaching their desired number of boys or girls when we impose gender-balanced preferences.

[Figure 5 about here.]

We combine those findings to measure the extent to which bias in preferences contributes to excess fertility over time, and how each component contributes to it. Figure 6 and 7 present the results, defining the composition counterfactual as balanced and indifferent, respectively. Overall, the bias in preferences have increased - in absolute value - excess fertility, expressed as a share of the total excess fertility of our sample, irrespective of the counterfactual. Again, however, the sign of this contribution changes depending on the counterfactual used. We now discuss each of these counterfactuals and their decomposition in more detail.

Composition counterfactual defined as balanced. Under this definition of gender-neutral composition, the sign of the effect depends on the definition of the intensity counterfactual. When defined as “boys to girls”, the bias in preferences leads to excess fertility. This excess fertility is growing over time. This, in turn, is due to the increase in the gender differential in intensity. The interaction effect remains relatively stable, since the increase in the gender differential in intensity over time is partially offset by the decrease in the additional number of families not reaching their desired number of boys due to a gender-biased composition of fertility preferences at their desired family size (see Figure 5). The composition effect remains marginal.

When intensity is set as “girls to boys”, as previously, bias in preferences lead to a decrease in excess fertility. This decrease becomes more important over time. This is mainly driven by the intensity effect, while both the interaction and composition effect remain small and stable.

Composition counterfactual defined as indifferent. Under this composition counterfactual, gender-biased preferences always lead to an increase in excess fertility. This increase is relatively stable over time, but this masks

bility to reach the desired gender composition of children based on the cohort-specific sex ratio. This allows us to account for the increasing usage of sex-selective abortion.

important evolution in the contribution of the composition and interaction components, which cancel each other out. When intensity is set as “boys to girls”, the interaction effect increases over time, due to the increase in the gender differential. The composition effect decreases over time and tends towards zero for the latest cohort. This is caused by the simultaneous decrease in the intensity for girls over time (see Appendix 5) and by the decrease in the share of families not reaching their desired number of boys.

When intensity is set as “girls to boys”, the composition effect increases over time, which is explained by the increase in the intensity for boys over time. This evolution is, however, counteracted by the interaction effect, which is negative and decreases over time, similarly driven by the increase in the gender differential in intensity over time.

[Figure 6 about here.]

[Figure 7 about here.]

6 Conclusion

In this paper, we study the effect of gender-biased fertility preferences on excess fertility, highlighting the importance of the definition of gender-neutral fertility preferences. We identify two main mechanisms through which gender-biased fertility preferences affect excess fertility, an intensity and a composition mechanism. For each mechanism, we define two different counterfactuals: defining the gender-neutral intensity as the intensity for boys or for girls and a gender-balanced or a gender-indifferent composition as gender-neutral composition counterfactual. We illustrate both mechanisms, using the different gender-neutral counterfactuals, in the context of India using data from four waves of the Demographic and Health Surveys.

The results of our empirical illustration suggest that the effect of gender-biased fertility preferences on excess fertility are sensitive to the definition of gender-neutral fertility preferences. Changing the definition of gender-neutral intensity leads to opposite effects of gender-biased fertility preferences on excess fertility when we define a gender-balanced composition as the gender-neutral composition. When we define a gender-indifferent composition as the the gender-neutral composition, we find identical effects of gender-biased fertility preferences on excess fertility for both definitions of gender-neutral intensity, but the effects are driven by different mechanisms.

Our empirical illustration remains based on declared fertility preferences, hence its validity in capturing true fertility preferences is arguably subject to some shortcomings (see the discussion by Pritchett and Summers (1994)). However, the aim of our empirical is to illustrate the importance of the definition of gender-neutral counterfactuals rather than to provide precise estimates. This paper therefore advances the discussion on the effect of gender-biased fertility preferences on excess fertility by addressing and illustrating the importance of explicitly defining and discussing the gender-neutral counterfactual of fertility preferences.

References

Altindag, O. (2016). Son preference, fertility decline, and the nonmissing girls of Turkey. *Demography*, 53.

Arnold, F. (1985). Measuring the effect of sex preference on fertility: The case of Korea. *Demography*, pages 280–288.

Arnold, F. (1992). Sex preference and its demographic and health implications. *International Family Planning Perspectives*, 18(3).

Bahk, J., Yun, S.-C., Kim, Y.-m., and Khang, Y.-H. (2015). Impact of unintended pregnancy on maternal mental health: a causal analysis using follow up data of the panel study on Korean children (pskc). *BMC pregnancy and childbirth*, 15(1):1–12.

Bairagi, R. and Langsten, R. L. (1986). Sex preference for children and its implications for fertility in rural Bangladesh. *Studies in Family Planning*, 17(6).

Baland, J.-M., Cassan, G., and Woitrin, F. (2023). Sex-selective abortion + instrumental birth = stopping rule. new measures and world evidence. *CEPR working paper.*

Basu, D. and De Jong, R. (2010). Son targeting fertility behavior: Some consequences and determinants. *Demography*, 47(2):521–536.

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human resources*, pages 436–455.

Chowdhury, A., Bairagi, R., and Koenig, M. A. (1993). Effect of family sex composition on fertility preference and behaviour in rural Bangladesh. *Journal of Biosocial Science*, 5(4).

Chowdhury, M. K. and Bairagi, R. (1990). Son preference and fertility in Bangladesh. *Population and Development Review*, 16(4).

Clark, S. (2000). Son preference and sex composition of children: Evidence from India. *Demography*, 37(1).

Cleland, J., Verall, J., and Vaessen, M. (1983). Preferences for the sex of children and their influence on reproductive behaviour. *Comparative Studies*, (27).

Coombs, C. H., Coombs, L. C., and McClelland, G. H. (1975). Preference scale for number and sex of children. *Population Studies*, 29.

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

Gipson, J. D., Koenig, M. A., and Hindin, M. J. (2008). The effects of unintended pregnancy on infant, child, and parental health: a review of the literature. *Studies in family planning*, 39(1):18–38.

Hall, J. A., Benton, L., Copas, A., and Stephenson, J. (2017). Pregnancy intention and pregnancy outcome: systematic review and meta-analysis. *Maternal and child health journal*, 21(3):670–704.

Hardee, K., Eggleston, E., Wong, E. L., Hull, T. H., et al. (2004). Unintended pregnancy and women's psychological well-being in Indonesia. *Journal of biosocial Science*, 36(5):617–626.

Kim, S. and Lee, S.-H. (2020). Son preference and fertility decisions: Evidence from spatiotemporal variation in Korea. *Demography*, 57.

Larsen, U., Chung, W., and Das Gupta, M. (1998). Fertility and son preference in Korea. *Population Studies*, 52.

McClelland, G. (1979). Determining the impact of sex preferences on fertility: a consideration of parity progression ratio, dominance, and stopping rule measures. *Demography*, 16(3).

Milazzo, A. (2018). Why are adult women missing? Son preference and maternal survival in India. *Journal of Development Economics*, 134.

Mutharayappa, R., Choe, M. K., Arnold, F., and Roy, T. (1997). Son preference and its effect on fertility in India. *National Family Health Survey Subject Reports*.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International economic review*, pages 693–709.

Park, C. B. (1976). Lifetime probability of additional births by age and parity for American women, 1935–1968: A new measurement of period fertility. *Demography*, 13(1):1–17.

Park, C. B. (1983). Preference for sons, family size, and sex ratio: an empirical study in Korea. *Demography*, 20(3).

Pritchett, L. and Summers, L. H. (1994). *Desired fertility and the impact of population policies*, volume 1273. World Bank Publications.

Repetto, R. (1972). Son preference and fertility behavior in developing countries. *Studies in Family Planning*, 3(4).

Shah, P. S., Balkhair, T., Ohlsson, A., Beyene, J., Scott, F., and Frick, C. (2011). Intention to become pregnant and low birth weight and preterm birth: a systematic review. *Maternal and child health journal*, 15(2):205–216.

Sheps, M. C. (1963). Effects on family size and sex ratio of preferences regarding the sex of children. *Population Studies*, 17(1):66–72.

Widmer, K. R., McClelland, G. H., and Nickerson, C. A. (1981). Determining the impact of sex preferences on fertility: a demonstration study. *Demography*, 18(1).

Williamson, N. (1976). Sons or daughters. *A crosscultural survey of parental preferences*, pages 149–165.

Yamaguchi, K. (1989). A formal theory for male-preferring stopping rules of childbearing: Sex differences in birth order and in the number of siblings. *Demography*, 26(3):451–465.

A Appendix

A.1 Alternative definition for balanced composition of fertility preferences

We provide an alternative definition of gender-balanced preferences and show that the qualitative implications of our analyses remain identical. The magnitude of the effects, however, is affected by the choice of definition of gender-balanced preferences. We consider the following alternative definition, for a family who has gender-preferences over a number n^* of children:

$$b^* = g^* = \begin{cases} n^*/2, & \text{if } n^* \text{ is even} \\ (n^*-1)/2, & \text{if } n^* \text{ is odd} \end{cases} \tag{10}$$

That is, if a family has gender preferences over an even number of children, then we define a gender-balanced desired gender composition as wanting half the children to be boys, and the other half to be girls. Note that this is identical to our previous definition of a gender-balanced composition. If however, a family has gender preferences over an odd number of children, we subtract one child from the odd number of children and define a gender-balanced desired gender composition as wanting half the remaining children to be boys, and the other half of the remaining children to be girls. Since families who have gender preferences over an odd number of children essentially have one additional birth to reach their desired gender composition, the share of families not reaching their desired number of boys and girls is lower under this definition of gender-balanced preferences (17.09% and 21.14%, respectively) compared to our previous definition (21.67% and 26.18%, respectively). Figure 8 displays the total effect, decomposed by mechanisms, of biased fertility preferences on excess fertility, similar to Figure 3.

[Figure 8 about here.]

Given that the share of families not reaching their desired number of boys and girls is lower under the alternative definition compared to the previous definition, the intensity effect must also be lower. Nevertheless, the main qualitative implication of our analysis, that moving from "boys to girls" to "girls to boys" intensity counterfactual reverses the sign of the total effect, remains unchanged. In the first case, excess fertility is higher when families have gender-biased rather than gender-neutral preferences whereas in the second case, excess fertility is lower when families have gender-biased rather than gender-neutral fertility preferences.

A.2 Gender differential in intensity over time

[Table 5 about here.]

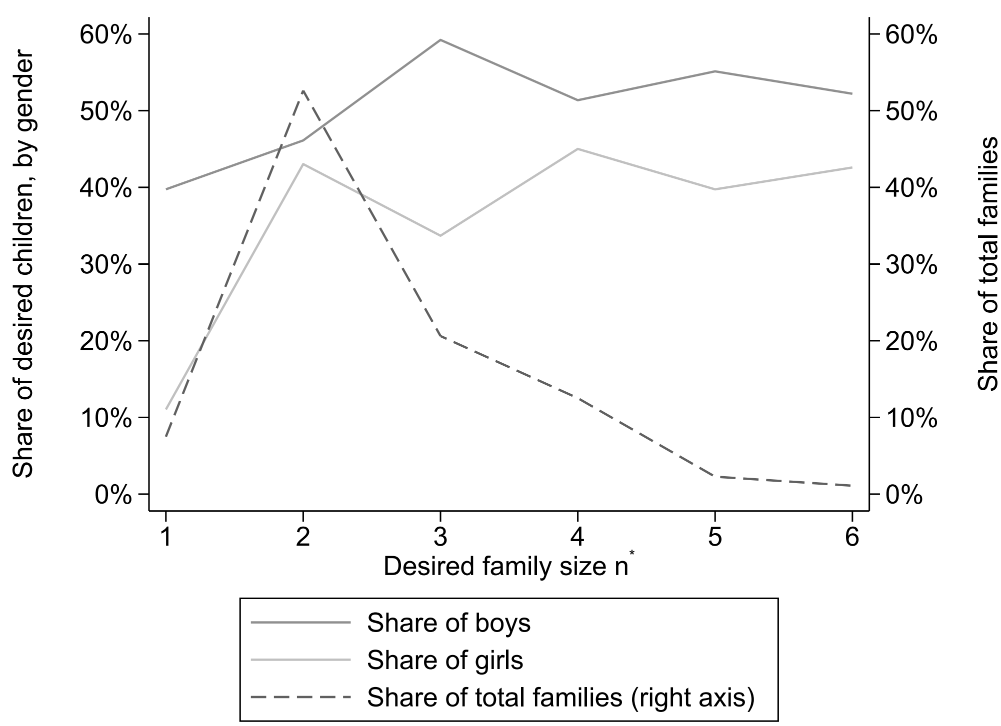
A.3 Evolution of the number of families not reaching their desired number of boys at their desired family size over time

In this section, we detail the evolution of the number of families not reaching their desired number of boys b^* at their desired family size n^* over time. Two factors explain this evolution. Firstly, the increase in the sex ratio at birth (i.e., the probability that the born child is a boy) over time, from 50.85% to 52.68% between the first and last cohort (intermediate sex ratios at birth are 51.11% and 51.31% for the second and third cohort, respectively). Secondly, the decrease in the share of desired boys in the total desired family size (b^*/n^*), from 52.87% to 50.30% between the first and last cohort (values are 51.56% and 50.78% for cohorts 2 and 3, respectively). Figure 9 displays the evolution of the average desired share of children by gender in the total family size, and the average desired family size, over time. The evolution of the number of families not reaching their desired number of boys b^* at their desired family size n^* over time can therefore not be fully accounted for by the spread of sex-selective abortion strategies.

[Figure 9 about here.]

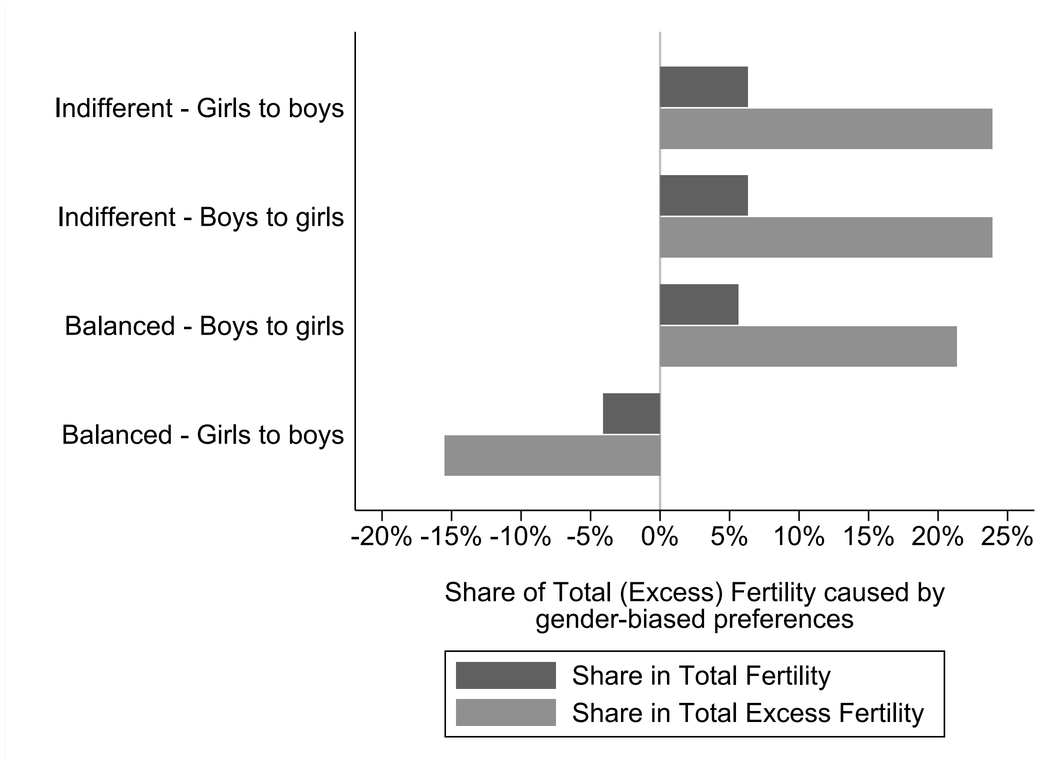
Figures

Fig. 1 Share of desired boys and girls by desired family size.



Reading: There are 20.61% of families desiring 3 children. These families desire 59.22% of boys and 33.69% of girls.

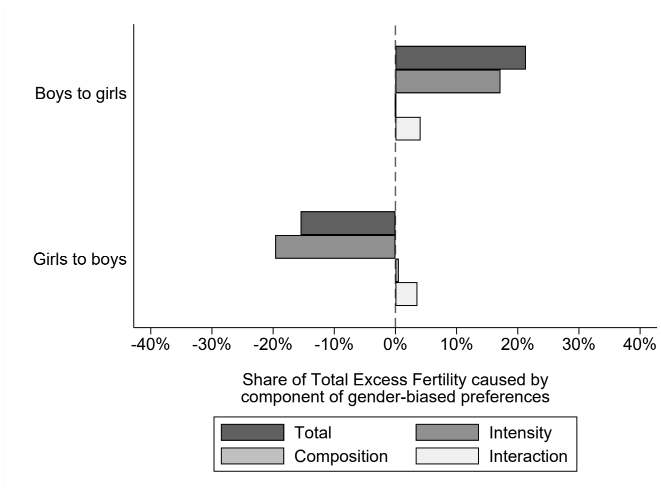
Fig. 2 Excess fertility caused by gender-biased preferences



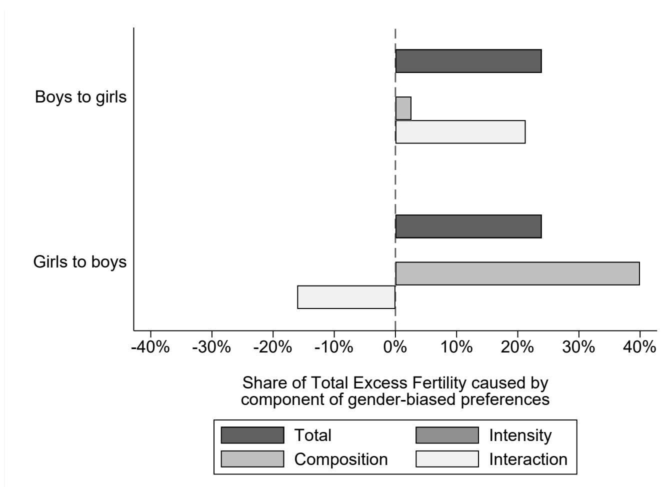
Reading: Under the balanced - girls to boys counterfactual, total excess fertility would increase by 15.51% compared to the gender-biased case.

Fig. 3 Decomposition of excess fertility caused by gender-biased preferences

(A) Composition counterfactual: *Balanced*

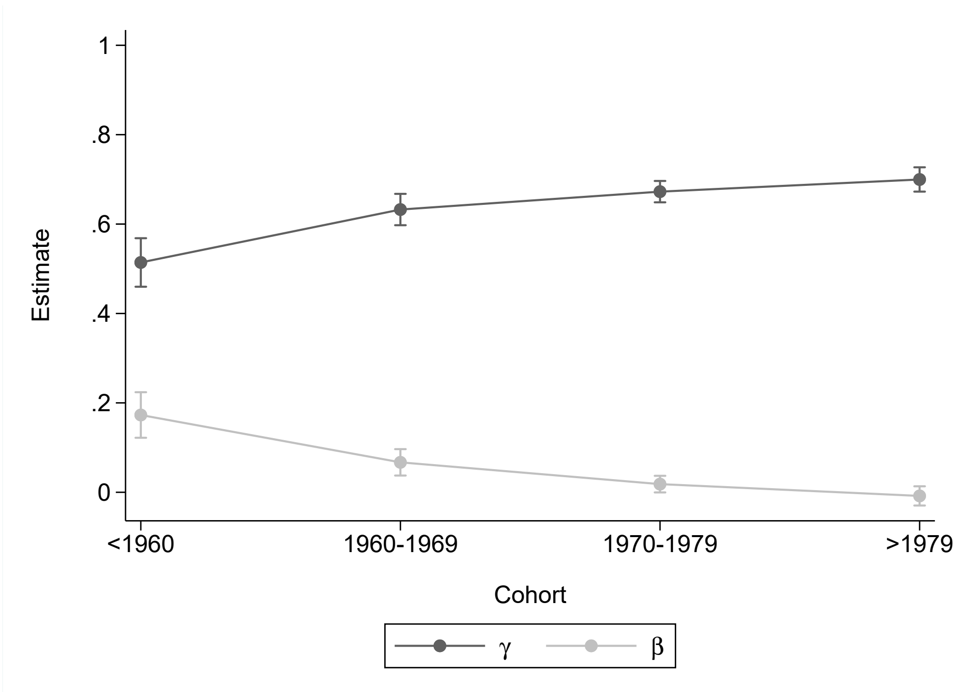


(B) Composition counterfactual: *Indifferent*



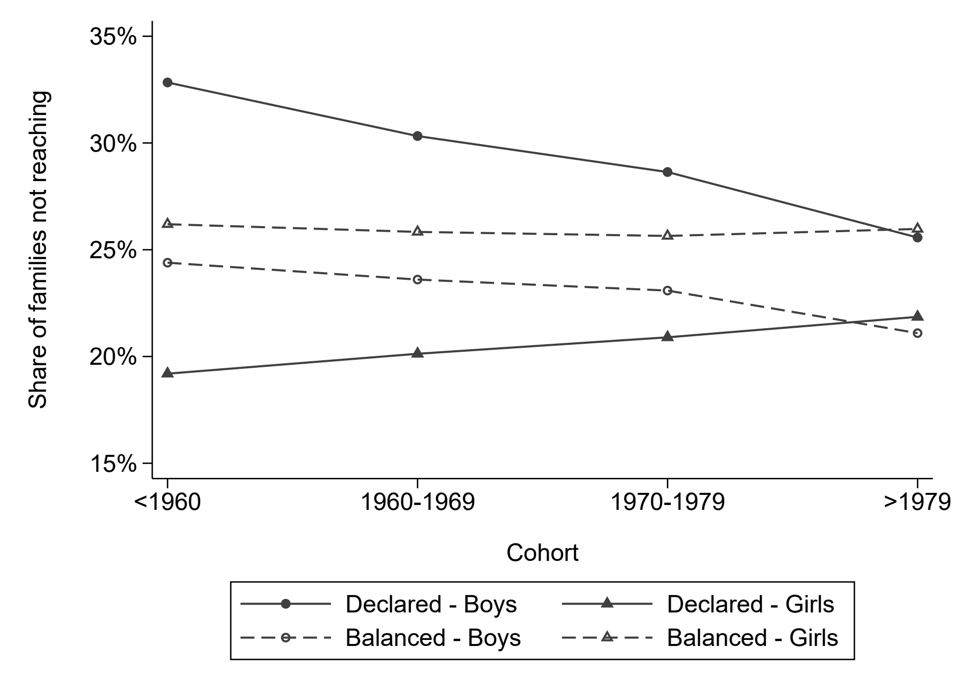
Reading: Under the balanced - girls to boys counterfactual, the intensity mechanism increases excess fertility by 19.66%, whereas the interaction effect decreases it by 3.59% and the composition component has a marginal decreasing effect of 0.55%.

Fig. 4 Evolution of the gender differential in intensity of fertility preferences



Reading: Mothers born between 1960 and 1969 will have on average 0.07 additional births if they miss a girl at their ideal family size. These mothers will have on average 0.63 more births if they miss a boy compared to if they miss a girl when reaching the ideal family size.

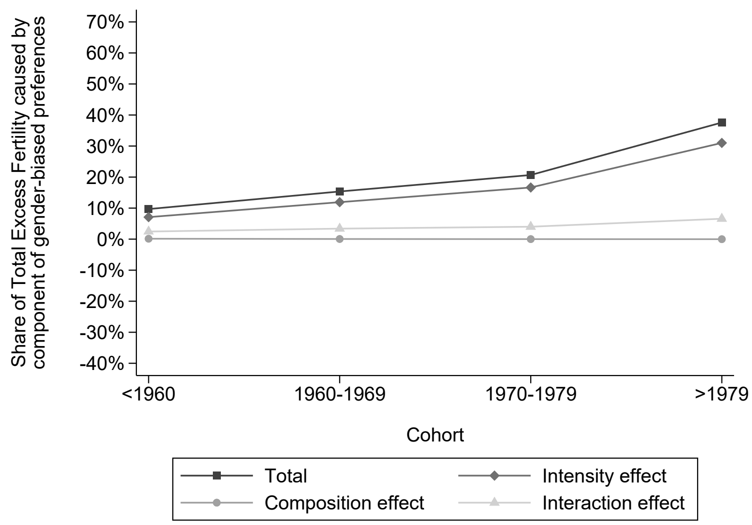
Fig. 5 Evolution of the distribution of families at desired family size



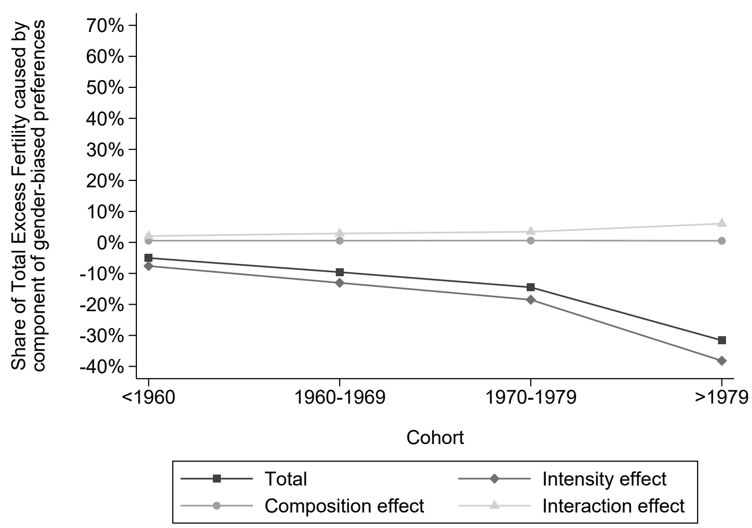
Reading: 20.13% of mothers born between 1960 and 1969 will not reach their ideal number of girls when reaching their ideal family size. 30.32% will not reach their ideal number of boys. Under a balanced composition, these numbers would be 25.84% and 23.60% respectively.

Fig. 6 Decomposition of excess fertility caused by gender-biased preferences over time, by intensity counterfactual (composition counterfactual: balanced)

(A) Counterfactual:
Boys to girls



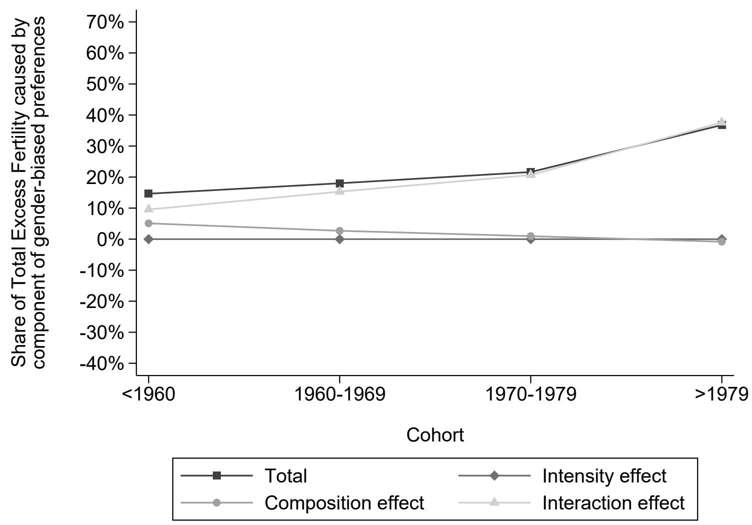
(B) Counterfactual:
Girls to boys



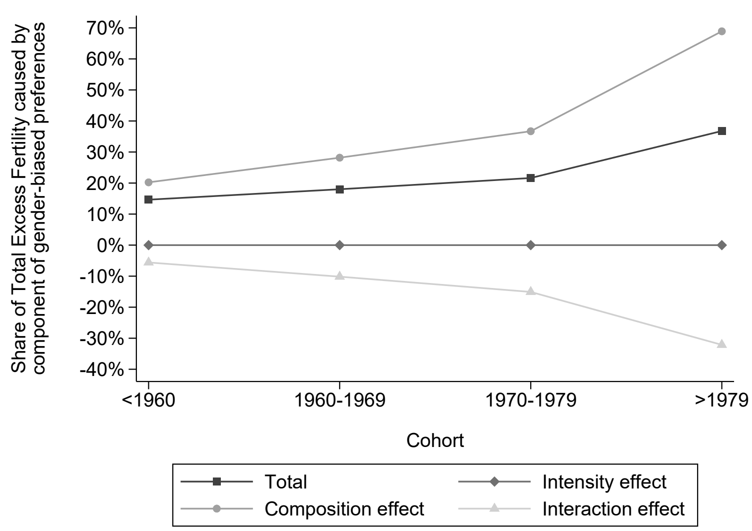
Reading: In the first cohort, under the balanced - girls to boys counterfactual, the intensity mechanism increases excess fertility by 7.62%, whereas the interaction effect decreases it by 2.04% and the composition component has a marginal decreasing effect of 0.56%.

Fig. 7 Decomposition of excess fertility caused by gender-biased preferences over time, by intensity counterfactual (composition counterfactual: indifferent)

(C) Counterfactual:
Boys to girls

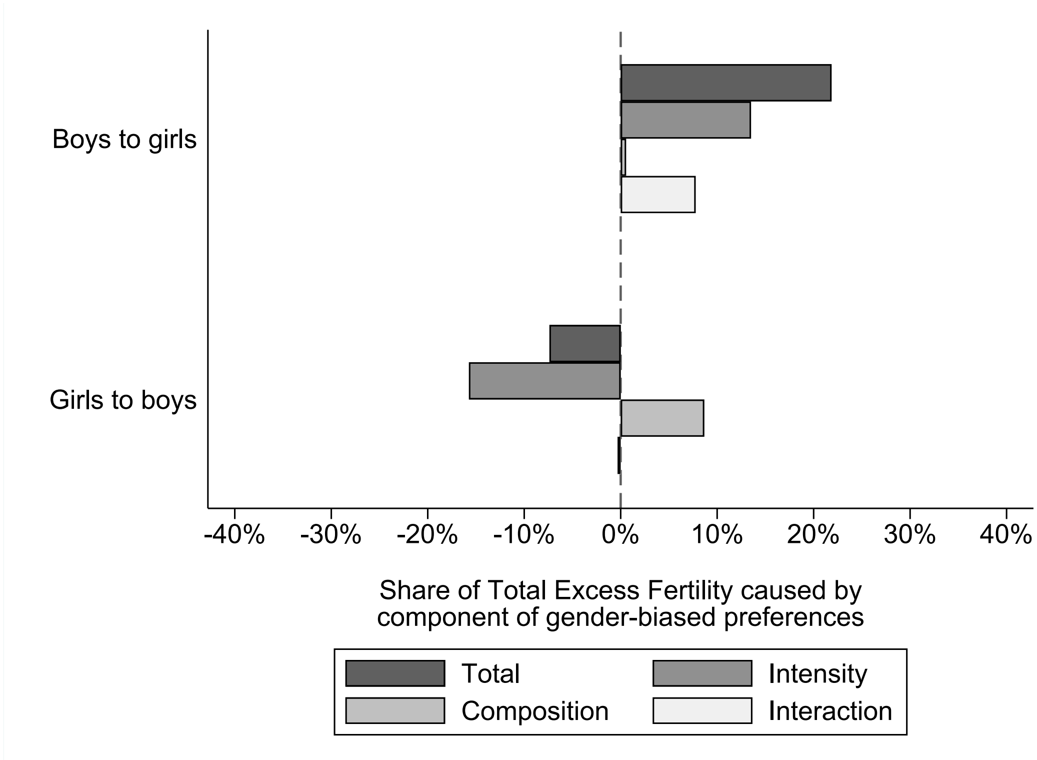


(D) Counterfactual:
Girls to boys



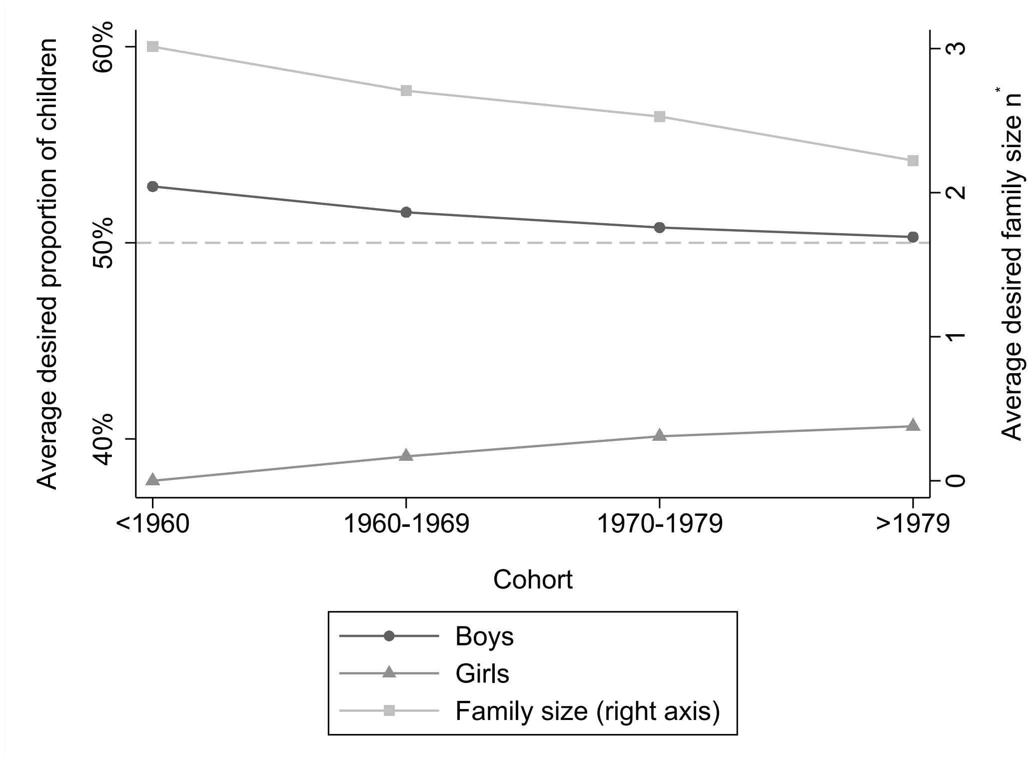
Reading: In the first cohort, under the indifferent - girls to boys counterfactual, the intensity mechanism does not impact excess fertility, whereas the interaction effect increases it by 5.59% and the composition component has a decreasing effect of 20.23%.

Fig. 8 Decomposition of excess fertility caused by gender-biased preferences, using an alternative definition of gender-balanced preferences



Reading: Under the balanced - girls to boys counterfactual, the intensity mechanism increases excess fertility by 15.74%, the interaction effect increases it by 0.33% and the composition component has a decreasing effect of 8.67%.

Fig. 9 Evolution of fertility preferences over time



Reading: In the first cohort, the average desired family size is 3.01 children, the share of desired boys is 52.87% and the share of desired girls is 37.87%.

Tables

Table 1
Descriptive statistics

	Mean	sd	Min	Max
Total children ever born	3.15	1.82	1.00	17.00
Desired number of children	2.47	1.13	0.00	30.00
Desired number of boys	1.26	0.81	0.00	12.00
Desired number of girls	0.99	0.66	0.00	10.00
Desired number of children of either gender	0.23	0.67	0.00	30.00
Reached desired family size	0.87	0.33	0.00	1.00
Observations	254,922			

Table 2
Share of families not reaching their ideal gender composition by composition of fertility preferences

	Not reaching desired # boys	Not reaching desired # girls
Gender-biased composition	27.31%	20.60%
Gender-balanced composition	22.03%	25.21%
Gender-indifferent composition	0%	0%

Table 3
Gender differential in intensity (OLS)

	nb of subs. Births
Birth missing	0.045^{***}
	(0.007)
Birth missing × Boy is missing	0.649^{***}
	(0.008)
Constant	1.024^{***}
	(0.014)
Observations	351942

Standard errors in parentheses
Standard errors are clustered at the primary sample unit
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4
Summary of the different mechanisms

	Boys to girls	Girls to boys
Intensity effect	$\hat{\gamma} * B_u$	$-\hat{\gamma} * G_u$
Composition effect	$\hat{\beta} * (B_b - B_u + G_b - G_u)$	$(\hat{\beta} + \hat{\gamma}) * (B_b - B_u + G_b - G_u)$
Interaction effect	$\hat{\gamma} * (B_b - B_u)$	$-\hat{\gamma} * (G_b - G_u)$

Table 5
Gender differential in intensity over time (OLS)

	(1) Before 1960	(2) 1960-1969	(3) 1970-1979	(4) After 1979
Birth missing	0.173***	0.067***	0.018	-0.008
	(0.026)	(0.015)	(0.009)	(0.011)
Birth missing × Boy is missing	0.514***	0.633***	0.673***	0.700***
	(0.028)	(0.018)	(0.012)	(0.014)
Constant	1.808***	1.143***	0.840***	0.683***
	(0.031)	(0.021)	(0.023)	(0.026)
Observations	47658	88233	141755	74287

Standard errors in parentheses
Standard errors are clustered at the primary sample unit
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$