February 2022

iFLAME WORKING PAPER SERIES

2022-iFLAME-02

Education and Childlessness in India

Thomas Baudin
IESEG School of Management,(LEM UMR 9221) and IRES, Université Catholique de Louvain

Koyel Sarkar
Center for Demographic Research, Université Catholique de Louvain

IÉSEG School of Management Lille Catholic University 3, rue de la Digue F-59000 Lille Tel: 33(0)3 20 54 58 92
www.ieseg.fr

IESEG IFlame Working Paper Series-2022-IFlame-02

Education and Childlessness in India*

Thomas Baudin† and Koyel Sarkar‡

January 03, 2021

Abstract

In a developing setting like India, women have started their long way to emancipation both at the family and societal levels. In this context, we study what may be perceived as a key sign of emancipation regarding marriage and motherhood: childlessness. Using micro-level regressions, we show that the probability of a woman ending her reproductive life without children exhibits a U-shaped relationship with her educational attainment. This is indicative of the fact that poverty and sterility are not the sole determinants of childlessness, but that better economic opportunities and empowerment within couples also matter. This result is robust to the introduction of important control variables such as the development level of the state where women live, the husband's education, age at marriage, religion, and caste. India seems to be joining a list of countries where adjustments to childlessness are much more than simple responses to boom-and-bust poverty.

Keywords: Childlessness, education, poverty, sterility, development.

*This research has received scientific and financial support from the ARC Project 15/19-063 on "Family Transformations - Incentives and Norms". Thomas Baudin also benefited from financial support from the French National Research Agency through the project MALYNES (ANR-18-CE26-0002). We would like to thank Dudley Poston, Ester Rizzi, David de la Croix, Philippe Bocquier, Mikko Myrskyla, Akansha Singh, Li Ma, Malgorzata Mikckuka, and Robert Stelter for their helpful remarks. We have also significantly benefited from presentations in seminars and conferences in Rostock, Louvain la Neuve, Prague, Durbuy, and Cape Town. All errors and imprecisions remain solely ours. **This document is a pre-publication version of an article to appear in *Population*.**

†IÉSEG School of Management (LEM UMR 9221), and IRES, Université catholique de Louvain, E-Mail: t.baudin@ieseg.fr. Phone number: +331.55.91.10.10. Corresponding author.

‡Center for Demographic Research, Université Catholique de Louvain, E-Mail: koyel.sarkar@uclouvain.be

Introduction

We propose a new interpretation of the dynamics of childlessness in India over the last decades. Based on the recent decomposition of childlessness by Baudin *et al.* (2015), we use micro-level data to show that a significant part of childlessness among Indian couples may be explained by the emergence of better educational and economic opportunities for women. This phenomenon exhibits a clear geographical heterogeneity, without contradicting the fact that childlessness among many Indian women is due to primary infertility related to sterility, sexually transmitted diseases, and poverty.

Given the strong demographic pressure faced by India,[1] the question of when and how the Indian fertility transition will end, as well as questions about family planning programs have monopolized the attention of demographers. As a result, hardly any space is left to study the other important dimension of fertility - childlessness. We contribute in filling this gap in literature by focusing on definitive childlessness among married Indian women. Childlessness is defined by the absence of any living birth in a woman's life; when a woman remains childless after the age of 40, it is usually called definitive childlessness.[2] The latest census, conducted in 2011, recorded the highest ever definitive childlessness rate in India: 7.89 percent among women above 40 years of age. Although the rate is only around half of that measured in richer countries like the US for the year 2014,[3] it is nevertheless far from natural sterility rates (Leridon, 2008). Childlessness in India is clearly not a marginal issue.

The definition of childlessness, whether voluntary ("childfree") or involuntary ("childless"), has been vigorously debated for many decades among demographers and sociologists. As discussed by Baudin *et al.* (2015), Toulemon (1996) and Poston and Cruz (2017) defining voluntary and involuntary childlessness in a non-disputable way is almost impossible. It may lead to potentially weak statistical analysis in contexts where the availability of attitudinal data is limited. Instead, we use the concepts of poverty and opportunity-driven childlessness, a phenomenon which has been studied by Baudin *et al.* (2015, 2019b). We decompose the group of definitively childless women into three categories: women suffering from an innate inability to reproduce (natural sterility),

[1] See the projections from the United Nations in the World Population Prospects 2019: Highlights.

[2] The 40-year-old threshold may be disputable, especially when considering rich countries where the postponement of the first birth is constantly increasing. In the case of a developing country, it is less disputable as explained in Baudin *et al.* (2019b). Also, see Poston and Cruz (2017) for alternative types of childlessness at lower ages.

[3] According to the estimation proposed by Pew-Research-Center (2015).

women who are childless because of factors related to poverty (poverty-driven childlessness), and women who are childless because economic opportunities led them to make decisions leading to childlessness (opportunity-driven childlessness). As poverty and economic opportunities can be proxied by educational attainment, it gives these concepts of childlessness an empirical counterpart.

One of the main predictions of our theory is that if opportunity-driven childlessness exists in India, the relationship between the probability of a woman ending her reproductive life childless and her level of education is U-shaped. Using micro-level regressions, we validate the existence of such a U-shaped relationship. This is indicative that poverty (proxied by low educational attainment) and sterility are not the sole gradients of childlessness; better economic opportunities and empowerment within couples (proxied by high educational attainment) also determine the probability of being childless. We show that this result is robust to the introduction of important control variables and potential confounders, such as the development level of the state where women live, the husband's education, age at marriage, religion, and caste. We do not limit our analysis to simple associations between women's educational attainment and the probability of being childless; we identify causal relationships using some unique features of the District-Level Household and Facility Survey. India seems to be joining a list of countries where adjustments to childlessness are much more than simple responses to boom-and-bust poverty.

1 Theoretical Hypothesis

We first discuss why we rely on the decomposition of childlessness proposed by Baudin *et al.* (2015) rather than on the more usual distinction between voluntary and involuntary childlessness. Then, we develop a theoretical framework adapted to the specific context of India. From this will emerge a set of hypotheses which we will test in the subsequent sections. These hypotheses appear in bold in Subsection 1.2.

1.1 Decomposing Childlessness

The literature about childlessness has extensively discussed the issue of voluntariness and involuntariness. Exploring National Survey of Family Growth (NSFG) data on American women, Poston and Cruz (2017) discuss the alternative methods to separate

childless women between those who are voluntarily and involuntarily childless. They explain why none of these methods are perfect; they also point to the need for intensive personal interviews to determine the underlying motivations for childlessness. Such interviews should provide details about women's reproductive health, their personal aspirations, and their entire matrimonial history. This need has also been highlighted by Veevers (1972) who, at the time, mentioned the absence of accurate, representative datasets. To the best of our knowledge, such datasets still do not exist either in the specific case of India or in any other country.

The absence of accurate data leads to the impossibility of decomposing childlessness into its voluntary and involuntary components in India and, more importantly, it does not allow to evaluate the socio-economic gradient of the latter. To circumvent this difficulty and to be able to understand the determinants of definitive childlessness in India, we use the decomposition proposed by Baudin *et al.* (2015). They decompose childlessness into three modalities: natural sterility, poverty-driven childlessness, and opportunity-driven childlessness. Natural sterility refers to the innate inability to give birth, and is uniformly distributed in the population, see for instance Leridon (2008).

Poverty-driven childlessness refers to women who were not a priori sterile, but failed to have children because of their poverty. As discussed for instance by Romaniuk (1980), McFalls (1979) and Frank (1983), one main cause of definitive childlessness among the poor in less developed countries lies in their higher degree of exposure to venereal diseases and malnutrition. Similar regularities are reported by Retel-Laurentin (1974), Poston *et al.* (1985), Ombelet *et al.* (2008), and Wolowyna (1977). Recent epidemiological research (Weiss *et al.*, 2008) has confirmed that married women in India are at higher risk for getting infected by chlamydia, gonorrhea (which are potential infertility causing STD's) and HIV (Solomon *et al.*, 2009), if they belong to socially disadvantaged groups, have experienced spousal sexual violence and domestic violence. In addition to poverty, Madhivanan *et al.* (2009) has documented the risk for adverse pregnancy outcomes caused by Trichomonas vaginalis, is higher among women who have been exposed to sexual intercourse at an early age through marriage. With reference to malnutrition and fecundity, even if recent evidence may seem less salient than in the past, at least Nanda and Garden (2009) document a positive association between low fertility and stunting among women in parts of India. Additionally, and importantly in the present day, inegalitarian, poor societies do not offer universal access to assisted reproduction techniques (ART), which reinforces the socio-economic gradient of poverty-driven childlessness: the rich can afford expensive ART, while the poor are excluded from this

technology. In the case of India, the low access of poor people to ART is documented by Rasool and Akhtar (2018), Allahbadia (2013), and Malpani and Malpani (1992).

Opportunity-driven childlessness refers to women who have delayed motherhood to a point where having children became either unfeasible or undesirable.[4] One main reason behind this postponement is the time cost of having children. In the vast majority of societies, as is the case in India, having children requires, for women much more than for men, investing time in child-rearing activities to the detriment of labor market activities. The renunciation of labor market activities is part of the opportunity cost of having children (Becker, 1981). Another dimension of this opportunity cost lies in the need to abandon other personal aspirations, whose importance is assumed to be positively associated with education. This fact is documented by Ghosh (2015) in the case of Kolkata's second demographic transition, while Surkyn and Lesthaeghe (2004) provide a more general discussion.

Let us notice that the decomposition of Baudin *et al.* (2015) does not ascribe a reason for being childless to each childless woman; from that point of view, it does not do better than the classical decomposition into voluntary and involuntary childlessness. Nevertheless, in contrast to the classical decomposition, Baudin *et al.* (2015) get rid of attitudinal concepts and the psychological roots of childlessness[5] to rely on the measurable socio-economic and biological factors of childlessness. Thanks to this conceptual framework, in the next section, we formulate a theory of childlessness in India. This theory will lead to micro predictions which can be tested on usual datasets, like census or survey data. We will be able to determine to what extent childlessness in India is mainly, if not completely, due to natural sterility and poverty, or whether a new kind of childlessness is emerging, a childlessness due to improvements in the status of women within society.

[4]This shows that opportunity-driven childlessness cannot be identified as voluntary childlessness. Indeed, trying but failing to have a first child at 38 does not mean that the person is voluntarily childless but that the economic and other kinds of opportunities she enjoyed made her entering late.

[5]Like for instance wanting or never wanting to have children. This said, one can wonder whether, beyond numbers and identification strategies, (attitudinal) signs of modernization, which are compatible with modern forms of childlessness, can be detected in Indian society. We claim that it is the case. Recent works have pointed out the practice of modern forms of marriage (similar to the British) among ethnic groups in the Darjeeling hills of India (Allendorf, 2013; Allendorf and Pandian, 2016 and Allendorf and Thornton, 2015) as examples of developmental idealism. Other works also show that single-child families are an emerging fertility trend in the country (Basu and Desai, 2016). Also, recent qualitative projects have described how certain metropolitan (Mumbai, Chennai, Vadodara, and Pune) Indian working women are giving their careers and personal aspirations higher priority than merely being mothers (Bhambhani and Inbanathan, 2018).

1.2 A Theory of Childlessness for India

In line with Baudin *et al.* (2015), we argue that beyond natural sterility, childlessness, at the individual level, is the result of an opposition between poverty and economic opportunities.

1.2.1 Poverty, opportunities and education

One challenge when exploring poverty versus economic opportunities induced childlessness resides in the measurement of poverty and economic opportunities. Most of the time, reliable and precise data are difficult to find. In their recent survey, Baudin *et al.* (2019a) argue that these two phenomena are advantageously proxied by the educational attainment of women. In this section, we analyse how this argument can be transposed to the Indian context and what are its implications.

Measuring economic opportunities offered to women is a challenge per se. If most of the datasets offer information about the profession of women, such an information has deep limitations. Indeed, the job of a woman of age 40+ measures her realized opportunities if she works at the time of the interview but not the set of opportunities she has been offered during her fertile years. This becomes an issue in the case of women who have chosen (or ended up by) not working. Not working does not necessarily mean that the woman did not enjoy good economic opportunities, it may also mean that she is not enjoying these latter at the time of the interview. The low female participation rate to the labor force in India[6] seems to make current occupational status a weak proxy of economic opportunities. For this reason, the level of education of a woman might appear as a better measurement of the economic opportunities she was offered when young.

Regarding poverty, reliable measures exist in the case of India. These measures may be relative or absolute. When relative, they classify people in percentiles in function of their wealth. Even if useful, they do not offer a stable measurement of the severity of poverty among successive cohorts as the general level of wealth may increase or decrease along time. Absolute measures are more indicated when cohort comparisons are to be proposed, as is the case in this paper; nevertheless, their connection to reproductive poverty remains limited. On top of this argument, temporality remains a major issue like for economic opportunities: poverty at age 40+ is only a proxy for the

[6]Documented for instance by Chatterjee *et al.* (2018).

poverty suffered during reproductive life. Again, educational attainment circumvents this temporal issue as educational attainment is decided before or just at the beginning of fertile life.

If education seems to be a good proxy for both economic opportunities and poverty in general, it has to be the case for India. Tilak (2002) shows how educational poverty, as measured by low levels of educational attainment, is one of the main, if not the sole cause of income poverty. Even if income poverty can in turn amplify educational poverty, the positive association between them is indisputable. Educational poverty is also a main factor leading to capability poverty in the sense of Sen and Nussbaum (1993). In the specific case of India, Duraisamy (2002) shows that the educational premium is rather weak for low levels of education (primary), while it is significantly strong for higher levels (secondary and tertiary). They document a decrease in the wage premium for primary education and an increase for higher levels during the period 1983-1994. This movement amplified after the economic reform of 1991, and income/wage inequality has rocketed in India since then. Kijima (2006) shows that this movement is mainly due to the increase in the returns to skills and the associated increase in the demand for skilled labor. These results are also confirmed by Chakraborty and Bakshi (2016) who show how learning English leads to higher wages. They estimate that, on average, not learning English during primary grades reduces weekly wages by 68 percent. Using different datasets and alternative measurements, Tilak (2007) reaches the same results. Based on this rich literature, we assume, **in the Indian context, that lacking education is associated to poverty but that higher education opens the set of economic opportunities**.

Education is the main engine of poverty and economic opportunities, but the intensities of these two phenomena oppositely evolve when education attainment increases. At low levels of educational attainment, poverty is severe, so that an increase in education strongly reduces the burden of poverty. On the other hand, as explained in the previous paragraph, it does not increase economic opportunities that much. To increase economic opportunities significantly, an increase in education has to occur in a context where the person is educated enough. For such individuals, the burden of poverty is weak. Our argument finds strong support in the enlightening paper of Chatterjee *et al.* (2018), which evidence the existence of a U-shape relationship between women's education and participation to the labor force in India. This empirical regularity is mainly due to the fact that among women with moderate educational levels, an increase in education allows to marry into wealthier families what allows them to leave the labor force. Above

a certain educational threshold, this effect becomes weak and is dominated by the increase in the set of economic opportunities.

As a consequence, among women who have low levels of education, an increase in educational attainment tends to reduce the probability of being childless because of poverty, without significantly increasing the probability of being childless because of better economic opportunities. On the other hand, among women who already have a high level of education, the risk of deprivations leading to infertility is already minimal, so that an increase in educational attainment only translates into better economic opportunities, and thus a higher probability of being childless. Finally, women who have an intermediate level of education are those who are likely to face the lowest probability of ending their reproductive lives childless: they are protected against extreme poverty but do not enjoy the largest sets of economic opportunities. From this part of our theory and echoing Baudin *et al.* (2015, 2019b), **we assume that the education of women has a U-shape incidence on the probability of being childless**.

1.2.2 Complementary Mechanisms

What are the main factors which could mitigate the U-shape relationship between female education and the probability of being childless? We identify at least four factors: male education, the Indian caste system, the geographical and institutional diversity of India, and religion.

The husband's education does not only reduce poverty, it also shapes the way a couple values the economic opportunities offered to the wife. Indeed, the husband's income allows to reduce the relative opportunity cost of child-rearing activities for women, as the couple has relatively less to lose when the husband also enjoys a high salary.[7] We therefore assume that for highly educated women, the husband's education reduces the incentive to postpone births, which in turn reduces the probability of remaining childless. On the whole, we then reach the global assumption that **male education reduces a woman's probability of being childless**.

The relative level of education of the husband and the wife may also be determinant. Let us point out that higher education levels are linked to higher degrees of personal

[7]Said differently, let us assume that to raise a child, a woman needs to spend one year out of the labor force. Let us also assume that a woman earns 100,000 rupees per year. Having a child would cost 100,000 rupees, whatever the wage of the husband, but it would represent 10% of the household income if the husband earns 900,000 rupees per year, while it would represent 90% of the household income if the husband earns 11,111 rupees per year.

aspirations, as explained by Surkyn and Lesthaeghe (2004). As women get more education and economic opportunities, their status within their couple improves: they enjoy stronger negotiation power. Following Chiappori (1988), this means that the family's objectives are more aligned with the wife's aspirations. Related to childlessness, it means that the wife's fertility preferences will become more important when deciding to have children or not. Nevertheless, the direction of the effect remains ambiguous: it is not given that an Indian woman systematically wants fewer children than her husband, or that she more often wants to postpone her first birth. Symmetrically, if a man is less inclined toward Westernized ways of life than his wife, getting more education reinforces his negotiation power, which reduces the probability of the wife not having children, while the reverse is also true. We know that an effect transiting through negotiation power exists, but cannot formally identify its sign. Let's notice that the economic literature on negotiation power inside couples (Chiappori and Donni, 2011, Baudin *et al.*, 2015, etc.) always identifies these effects as second-order effects.

India is a large, culturally and institutionally heterogenous country. India's diversity lies in its 29 official languages,[8] its caste system, and its being home to all the major religions in the world. Studies have documented a north-south divide in the country with respect to (*i*) openness to fertility change (Dommaraju, 2009), (*ii*) how education level and religious affiliation lead to different fertility outcomes (Kulkarni and Alagarajan, 2005), and (*iii*) how caste differences lead to a differential utilization of maternal health care (Kumar and Gupta, 2015). Even the implementation of national policies differs by state; this is the case for instance for education policies. In a recent paper, Chakraborty and Bakshi (2016) document how West Bengal has forbidden English classes in primary schools and how it affects the well-being of children. This state diversity is also driven by bio-geographic factors like climate, the intensity of pollution, the types of natural resources, etc.

Based on these documented empirical regularities, we identify three ways in which state-level diversities may influence the probability of being childless. First, **state specificities contribute to the formation of reproductive norms regarding the ideal size of families and about childlessness and the status of childless women**.[9] Second, **state specificities and institutions directly influence the reproductive conditions which women face, like malnutrition, sanitation,**

[8]Our dataset does not offer variables about language spoken at home.

[9]In fact, this is also true for districts and villages/cities, nevertheless, we do not have access to these geographical scales.

protection against venereal diseases, delivery conditions, access to modern ART, etc. Third, **states have the power to change some aspects of educational policies.**

Indian economic development and the improvement of economic opportunities offered to women cannot hide the fact that India is a patrilineal country where women still have tremendous pressure to bear a child soon after marriage. In some states, cultural barriers may annihilate the positive impact of better economic opportunities on the probability of remaining childless at the end of reproductive life.

Religious affiliation is another potential factor influencing the probability of remaining childless at the end of reproductive life. Among others, Koropeckyj-Cox and Pendell (2007) document an effect of religious affiliation on attitudes and intentions toward childlessness. Ram (2005) documents an impact of some religious affiliations on definitive childlessness rates in India. These effects are discussed in alternative contexts like Europe (Sobotka, 2017) and the United States (Abma and Martinez, 2006) also. We then hypothesize that **religious groups may have specific attitudes toward childlessness.** Using the theoretical framework of Goldscheider and Uhlenberg (1969), we assume that **the minority status and pro-natalist values of religious groups like Catholics and Muslims may drive religious differentials regarding fertility and childlessness.**

Another important dimension linking culture and social structures in India pertains to the division of society into castes. In post-independence India, the constitution established scheduled castes (SC), scheduled tribes (ST), other backward castes (OBC), and Others, with the aim of acknowledging and uplifting the marginalized sections of society. The SCs and the STs (lowest in the caste hierarchy) are caste groups who were historically the most deprived of certain basic human rights, lived in extreme poverty, malnutrition and were socially excluded. The OBCs are also historically backward castes while the general castes comprise of all the other upper castes. Several studies have shown that caste remains a strong factor in Indian society and women from the SC and ST groups often experience the highest burden of social exclusion including educational exclusion, poverty, lowest maternal health care utilization (Kumar and Gupta, 2015), lack of occupational mobility across generations (Banerji, 2012), high fertility outcomes (Ramesh, 2014), etc. We then make the assumption that **belonging to a caste has a direct impact on the educational attainment of men and women, as well as on their poverty status since for a low educational level, the burden of poverty is stronger on the lowest castes**, a phenomenon which

disappears for high educational attainment.

From the theory developed above, we can build the causal diagram in Figure 1. It provides a complete picture of our reasoning.

2 Data and descriptive statistics

2.1 DLHS data

We use secondary data from all three rounds of the District-Level Household and Facility Survey (DLHS, 1998-99, 2004-05, and 2007-08). The DLHS provides cross-sectional, micro-level data that covers all districts and states in the country. The survey was conducted by the Indian Institute for Population Studies (IIPS) Mumbai, funded by the Ministry of Health and Family Welfare (MOHFW). DLHS data guarantee a large sample size and wide coverage of the Indian population on long time span. In addition, the dataset also gives sufficient information about childlessness, marriage, fertility, and other socio-demographic characteristics for all categories of respondents.[10] Because the profession has discussed some issues about the quality of DLHS data and point to a potential superiority of DHS data over DLHS, we also use the DHS 4 data (2015-16), to test the external validity of our findings in section 5.

After deleting missing values, input errors, grouping women into birth cohorts, selecting only age groups 40 to 49 years old, we have a final sample size of 158,112 currently-married women born between 1953 and 1968, among whom 4,725 are childless.[11]

2.2 Patterns of Fertility and Childlessness in India

The literature has widely documented a negative relationship between the fertility of mothers and income, and thus a positive relationship between the degree of poverty and fertility (Birdsall *et al.*, 2001). Thus, if childlessness was mainly due to poverty or infertility, one should find either an absence of correlation between childlessness

[10]Let us notice that single women were not included in all the rounds, and when included, they were not asked about their number of children.

[11]A total of 529,817 households and 474,463 ever-married women were covered by the DLHS in the first round, 620,107 households and 507,622 ever-married women in the second round, and 720,320 households and 643,944 ever-married women in the third round of the survey. The age group selection allows preventing selection bias due to cohort-based mortality after 50.

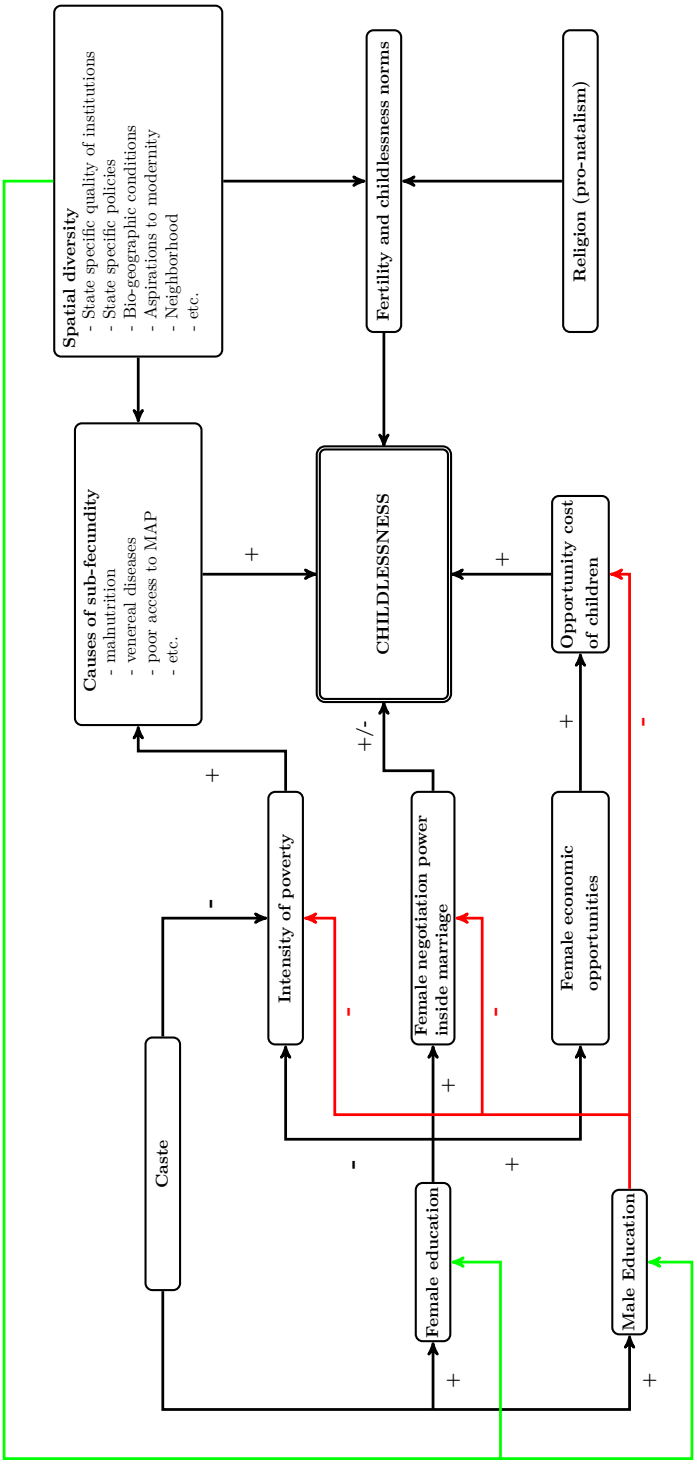

Figure 1: Causal diagram

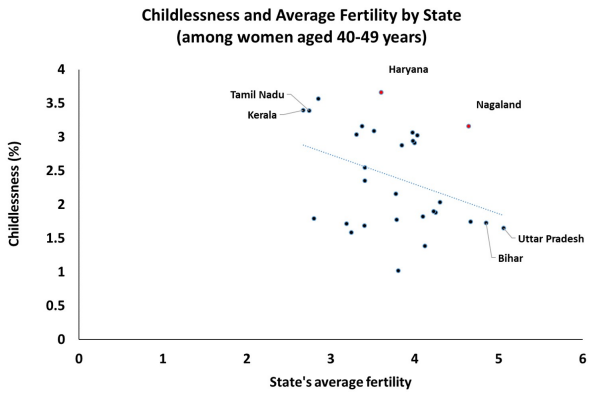

Figure 2: Correlation between childlessness rates and average fertility of mothers at the state-level among birth cohorts of women aged 40 - 49 years (right, Source: DLHS 1998-99, 2002-04 and 2007-08)

rates and the average fertility of women in Indian states (only sterility matters in that case), or a positive correlation between the two (poorer states should have both higher childlessness and higher fertility). Using state-level data (Figure 2 right), we show that the correlation between childlessness rates and the average fertility of women is strongly negative. This goes against the idea that childlessness in India is only due to infertility or poverty reasons, and is indicative of the possible existence of opportunity-driven childlessness. However, this does not mean that childlessness is only opportunity driven in India; outliers like Haryana and Nagaland indicate that both childlessness and the average fertility of women can be high, which is indicative of a potentially high prevalence of poverty-driven childlessness as well.[12]

In the sub-sample we have selected, the childlessness rate equals 3% against the 7% measured in the Indian Census. A first reason why childlessness rates may differ between our sample and census data comes from the age of respondents: in our sample, women are aged between 40 and 49 while census data consider all women above 40. The literature has pointed to associations between parity attainment and life expectancy of women, see for instance Doblhammer (2000) in the European context. Furthermore, one should keep in mind that in India, childless marriages are subject to much higher divorce rates than fertile marriages with at least a son. For this reason, our results should be seen as a lower bound of the reality.

[12]Even if we dispense with attitudinal concepts in our theory and analysis, it is interesting to notice within DLHS data, attitudinal signs confirm the existence of childlessness not due to medical reasons. See Appendix A.1.

2.3 The Education Gradient

In our sub-sample, 29 percent of women never went to school, while 24.8 percent received primary education, 38.6 received secondary education, and less than 10 percent of women have a university degree.[13] As indicated in Figure 3, at the country level, childlessness exhibits a J-shaped relationship with years of schooling. This shape indicates that in India, as in many other countries (Baudin *et al.*, 2019b), above an education threshold (9 or 11 years of schooling), with an increase in years of education, childlessness among women tends to increase. This fact is more salient when focusing on the youngest cohorts as shown in the right panel of Figure 3.

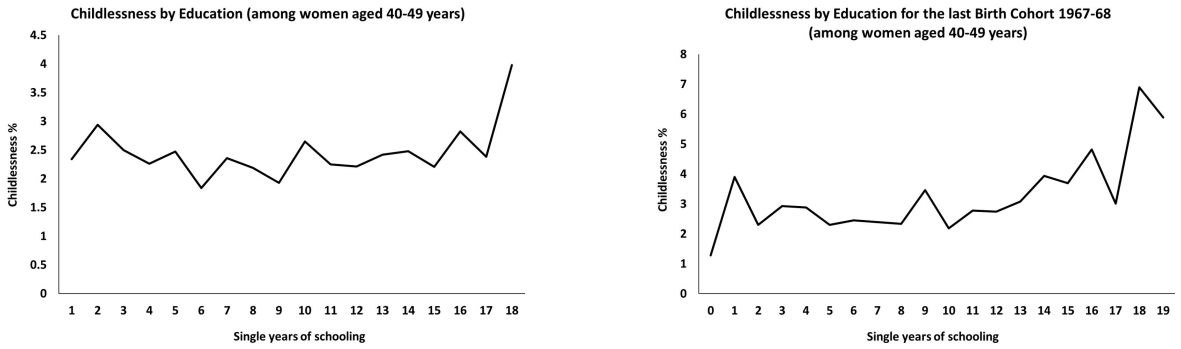

Figure 3: Childlessness rates by years of schooling among women aged 40-49. Source: DLHS 1998-99, 2002-04, and 2007-08

In a cross-state perspective (Figure 4), childlessness rates exhibit only a weak correlation with average education when all states are considered. Nevertheless, when we consider only the major states in India, this correlation becomes clearly positive. This suggests that the states with high levels of education (supposedly the most developed ones) are also those with higher childlessness, while the states with low education levels are those with the lower childlessness rates. This supports our hypothesis that opportunity-driven childlessness does exist in India. Nevertheless, our correlation charts are populated with outliers like Jharkhand (where average education is low and childlessness is high) and Haryana (where childlessness is low is spite of high education).

[13]Having primary education or secondary education does not mean here that a woman completed primary or secondary education, but that she had at least some years of this educational cycle.

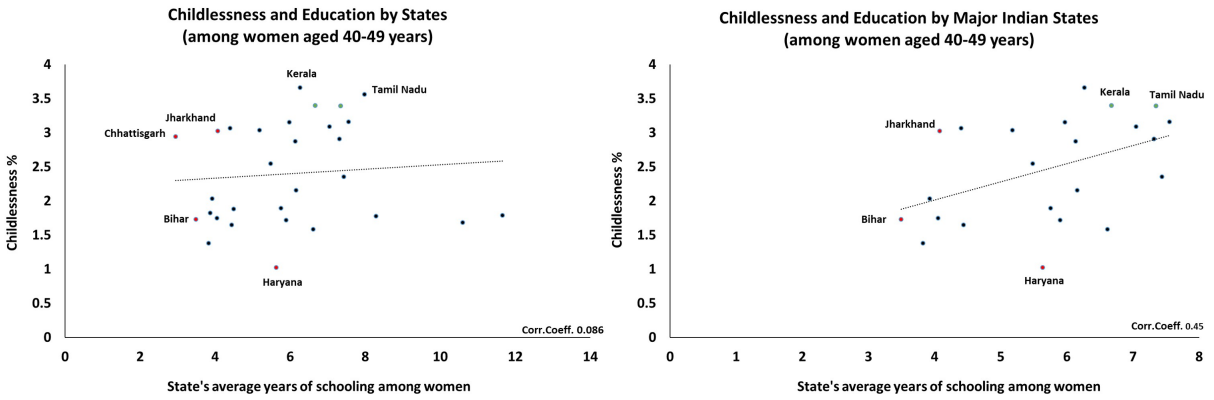

Figure 4: Correlation between childlessness rates and average education across states: including all states (left) and including only the major Indian states (right)

As shown in Figure 5, a similar pattern can be noticed in the district-level maps, which show that childlessness is higher among women who are highly educated (graduate and above) than among women with less education in all districts of India. The map at the center of the figure is very similar to the one proposed in the enlightening paper of Singh *et al.* (2017). The educational gradient that we evidence in the left and right maps refines the findings of Singh *et al.* (2017), it indicates that the spatial diversity of childlessness in India may find some of its roots in the spatial diversity of educational attainment. Indeed, childlessness rates are higher but less heterogenous among highly educated women than among lowly educated ones.

Though the existence of opportunity-driven childlessness can be expected from the above descriptive findings, it is still not clear whether the effect of education is robust to other socio-economic effects among childless women. This doubt will be ruled out in the next section using multivariate regression models.

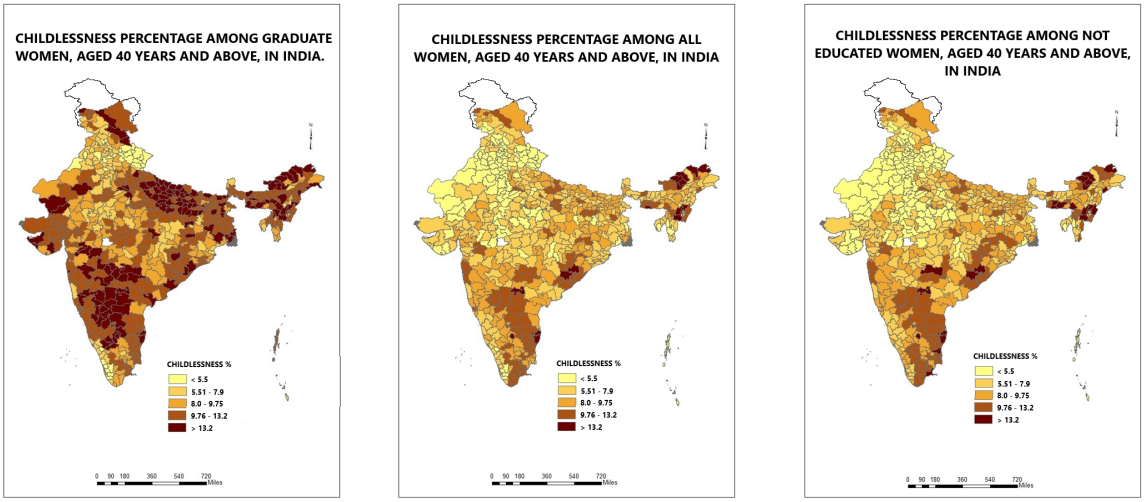

Figure 5: Childlessness rates among highly educated women (left), overall population (center), and not educated women (right) at the district level. Source: Census of India 2011

3 Regressions

In this section, we go beyond descriptive statistics, and identify the main determinants of the probability of a woman ending her reproductive life childless at the individual level.

3.1 Methodology

In our regression models, we study the determinants of the probability of a woman ending her reproductive life childless, for which we use information about completed fertility to build the dichotomous variable 'childlessness'. It takes value 1 if the respondent has no children and 0 otherwise. We use a logistic regression specification. From our causal diagram, it appears that the relationship between education and the causes of childlessness may be confounded by at least the caste system, spatial and cohort diversity. Following the methodology of Wunsch (2007), we have to control for these elements when estimating the relationship between education and childlessness.

We consider two kinds of fixed effects, the first is a cohort fixed effect and the second is a state fixed effect. Eight cohorts were compiled for the study, the oldest being women born in 1953-54 and the youngest, women born in 1967-68. All 35 states and union

territories in India were considered under the state fixed effects in the models.[14] It means that all the interpretations we propose are valid in a specific state and a specific cohort.

In order to take into account potential confounding factors and control variables in a comprehensive way, we have introduced independent variables stepwise. In the first step, we consider the respondent's education level grouped in four categories; no education, primary, secondary, and higher. In the second step, we add the husband's education level (no education, primary, secondary, and higher), child marriage (if the respondent got married before the age of 18, which is the legal age for marriage for women in India), and place of residence (rural or urban). In the third step, we add cultural variables like religious denominations (Hindu, Muslim, Christian, Sikh, Buddhist, and other) and caste categories (scheduled caste, scheduled tribe, other backward caste, general, and other). In the fourth step, we add the variable 'development level of the state', i.e. states are categorized into most developed, least developed, and intermediately developed, based on their average years of schooling among women, using the India census of 2011 estimates. States with an average of more than 7.4 years of education among women are categorized as developed, states with an average of less than 5.1 years of education among women are categorized as least developed, and other states are categorized as intermediate states.

It is important to notice that in our main regressions, observations are not weighted using the sample weights offered by DLHS waves. This decision comes from the difficulty we faced when trying to gather information about the way weights were computed in the first two waves of the DLHS. For this reason, we suspect that the comparability of data between waves is not guaranteed when using weights.[15]

3.2 Results

In Model 1, we regress childlessness with the education level of women. As hypothesized, we find a U-shaped curve after controlling for cohort and state fixed effects. This result

[14]The terminology fixed effect has to be understood here as the use of dummy variables controlling for the cohort of birth and the state of residence of the respondent.

[15]Nevertheless, we have tested all our regression models considering weighted data instead of unweighted data. The only significant change occurs when looking at the impact of being a Christian compared to being a Hindu. It has no significance when using weighted data, while it does when using unweighted data. For this reason, it is reasonable to assert that weighting issues are minor in our study. All results are available upon request.

still holds after controlling for background and marital characteristics, and cultural and ecological variables in subsequent models. These successive results confirm the main prediction of our theory, stating that the probability of women finishing their reproductive lives childless has a U-shape relationship with their level of education.[16]

In the second model, we show that the higher the husband's education level, the lower the probability of the woman remaining childless. The gradient is significantly and strongly negative. This means that the husband's education clearly plays the role of insurance against poverty-driven childlessness among low educated women, while it reduces the opportunity cost of having children for highly educated women. Interestingly enough, the non-linear shape of the probability estimations coming from our logistic underlying distribution delivers the following result: among low educated women, increasing education reduces the probability of being childless, and this marginal effect is reinforced by the husband's education. Conversely, for highly educated women, increasing their education increases their probability of being childless, but this marginal effect is smaller when the husband's education is high.[17]

Place of residence does not seem to have any effect on childlessness. Child marriage takes the value one if the respondent got married before the age of 18. In the case of a marriage during childhood or teenage years, the chance not to be childless is twice as high as that for a woman who married later. Beyond the fact that marrying before 18 prevents from a love marriage, it provokes a greater exposure time to the risk of conception. Being involved in a traditional arranged marriage may also increase the incentive or family pressure to conform to the traditional family system, in which having children is compulsory. Interestingly enough, we find that the correlation between the respondent's education and age at marriage is positive, but this does not prevent education from exerting a U-shape influence on the probability of remaining childless.

One may wonder why we have not divided age at marriage into more groups to capture late entry into marriage. This may be quite important knowing that women who do

[16]Goodness of fit measured by adjusted count R2 is low at first sight. It comes from the fact that being childless is a very rare event in India, which puts the logistic model in a bad position. As methods like penalized likelihood cannot be used with the high number of observations we have, we propose an alternative exercise in Appendix B.2. In this exercise, we draw a limited number of non-childless women randomly in order to diminish the size of their group and thus increase the prevalence of childlessness. We show that all our results hold in that situation, while adjusted count R2 increases significantly. This kind of issue with very rare events, as well as the impossibility of using penalized likelihood in our case, is documented for instance by King and Zeng (2001).

[17]Mathematically speaking, denoting female and male education e^f and e^m respectively and denoting $\mathcal{P}(\chi_i = 1|X_i)$ as the probability to be childless for a woman, one can verify that for low values of e^f, $\frac{\partial^2 \mathcal{P}(\chi_i=1|X_i)}{\partial e^f \partial e^m} > 0$, while for high values of e^f, $\frac{\partial^2 \mathcal{P}(\chi_i=1|X_i)}{\partial e^f \partial e^m} < 0$.

Table 1: Determinants of the probability to end reproductive life childlessness

VARIABLES	Model 1	Model 2	Model 3	Model 4
Education				
No education	1.36***	1.22***	1.21***	1.12*
Primary	0.937	0.97	0.97	0.94
Secondary Education	Ref.	Ref.	Ref.	Ref.
Higher	1.399***	1.38***	1.38***	1.35***
Husband's education				
No education		Ref.	Ref.	Ref.
Primary		0.73***	0.74***	0.78***
Secondary Education		0.66***	0.66***	0.66***
Higher		0.54***	0.54***	0.54***
Child marriage				
No		Ref.	Ref.	Ref.
Yes		0.54***	0.54***	0.58***
Place of residence				
Rural		Ref.	Ref.	Ref.
Urban		1	1	1.04
Religion				
Hindu			Ref.	Ref.
Muslim			0.89*	0.805***
Christian			0.74***	0.821***
Sikh			1.02	0.68***
Buddhist			0.8	0.77*
Others			1.04	1.12
Caste				
General			Ref.	Ref.
SC			1.05	1.08
ST			1.13*	0.98
OBC			0.98	0.89*
Others			1.11	1.14
State development level				
Developed States				Ref.
Least developed States				0.76***
Intermediate States				0.87***
Fixed effects				
Cohort FE	YES	YES	YES	YES
State FE	YES	YES	YES	NO
Pseudo R2	0.0123	0.0207	0.0213	0.0122
BIC	33585.197	33365.756	33454.349	33374.480
Number of obs.	158112	158112	158112	158112
Count (adj)	0	0	0	0

Notes: Odds-ratio reported. *** p<0.01, ** p<0.05, * p<0.1

not want to have children may also delay marriage as much as possible. We test this alternative in Appendix B.1, and show how our results improve in terms of goodness of fit. Indeed, women who marry later are likelier to remain childless. Let us point out that including this finer categorization for the age of entry into marriage makes the U-shape relationship between education and childlessness disappear. It is replaced by a decreasing relationship; see Appendix B.1. This comes from the fact that those women who marry late are these enjoying the best economic opportunities as they also are the more educated ones. Nevertheless, all the coefficients and measures of fit of this alternative model are subject to endogeneity issues and are strongly suspected of being spurious. Indeed, not wanting children (like not wanting to have a husband) determines both education decisions (and thus the history on the labor market) and the probability of remaining childless. This double causality issue may lead the effect of age at marriage to confound that of education.

This last criticism does not apply when only considering child marriage, as that kind of marriage has been decided by the respondent's family and not the respondent herself.[18] The possibility remains that reverse causality between education level and childlessness exists, as women who did not want children may have focused on studying - we are aware of this. Nevertheless, even if the causality is reversed, both mechanisms refer to opportunity-driven childlessness: if a woman did not want children, she was able to avoid having them by focusing on education and job opportunities.

In the third model, we see that caste has only a limited effect on the probability of remaining childless in our main regression, as only the ST have a higher probability of remaining childless compared to the General Caste. Said differently, everything else being equal, the SC, OBC, and Other castes seem not to suffer from any kind of excess childlessness compared to the General Caste. If this result is surprising at first sight, it is not once we recall that we control for state fixed effects. Indeed, discrimination against the lower castes has decreased over time and differs in space, which is captured by our fixed effects. Said differently, the ST seem to suffer from an extra risk of definitive childlessness at any time in any state.[19]

Turning to religion, we find that both Christians and Muslims are less likely to remain

[18]Notice here that we do not pretend that all marriages over 18 are not arranged marriages but that below 18, they are for sure.

[19]Interestingly enough, when we suppress state and cohort fixed effects, we find that women who belong to the General Caste are significantly less childless than women who belong to any of the other castes. The geographical and temporal aspects of the discrimination against lower castes is then included in the caste variable fully.

childless than Hindus. As explained in our theoretical section, this could be due to the pro-natalist aspects of these religions, as well as to their minority status in the country.

In the fourth model, we add the development level of the state, which is the state-specific average years of schooling among women. In order to not introduce multicollinearity issues, we have suppressed the state fixed effect from that model. The finding is in line with the macro model of Poston and Trent (1982), as well as with those of Baudin *et al.* (2019b). It shows that education at the micro level has an impact in itself on the probability of being childless, and this effect is reinforced by a macro effect of development, proxied by average education at the state level, on the individual probability of being childless.

From Model 3, we learn that compared to a woman who has completed secondary education, a woman who never went to school is 1.21 times likelier to finish her fertile life childless. A woman who has some years of college is 1.38 times likelier to be childless than a woman who has a secondary education level.[20]

4 Identification Checks

Following Wunsch (2007) but also the econometric literature on causality (see for instance Heckman, 2008), we know that measuring educational attainment and caste, and taking into account state dummies allows us to detect the potential existence of opportunity-driven and poverty-driven childlessness. Nevertheless, the signs of the effect of education on the probability of remaining childless do not constitute an irrefutable proof that poverty explains the decreasing part of the relationship between childlessness and education, or that better economic opportunities explain the increasing part. Indeed, some underlying factors may influence both education and childlessness in a way not taken into account in our theory. To explore this possibility and progress toward a causality analysis, we propose some identification checks using some unique features of the third wave of DLHS.

[20] As an alternative to state fixed effects, we have tested models with state-level ecological variables like the average childlessness rate and average fertility. We find that the higher the childlessness rate in a state, the higher the individual probability of remaining childless. The opposite is true for average fertility. This may reflect the existence of norms about family size and childlessness, but also factors related to state specificities in terms of fertility. Said differently, state fixed effects may control effectively for cultural norms about reproduction and other kinds of ecological differences, like the prevalence of venereal diseases or other factors leading to sub-fecundity. Results are available upon request.

Table 2: Determinants of the probability to end reproductive life childlessness

VARIABLES	Model A	Model B	Model C	Model D	Model E
Education					
No education	-	1.002	-	-	1.002
Primary	-	0.977	-	-	0.991
Secondary Education	-	Ref.	-	-	Ref.
Higher	-	1.567***	-	-	1.488***
Wealth Index					
Poor	1.254*	1.280*	-	1.266*	1.286***
Middle	Ref.		-	Ref.	Ref.
Rich	0.877*	0.818**	-	0.839*	0.803***
Occupation					
No occupation	-		Ref.	Ref.	Ref.
Labourer	-		1.144	1.086	1.111
Lowly skilled	-		0.969	0.883	0.901
Medium skilled	-		1.212	1.220	1.170
Highly skilled	-		1.386***	1.413***	1.191
Child marriage	YES	YES	YES	YES	YES
Religion	YES	YES	YES	YES	YES
Caste	YES	YES	YES	YES	YES
Fixed effects					
Cohort FE	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES
Number of obs.	51,709	51,709	51,709	51,709	51,709

Notes: Odds-ratio reported. *** p<0.01, ** p<0.05, * p<0.1. Wave 3 of DLHS.

The ideal dataset would offer a precise measure of both poverty and economic opportunities offered to women for each wave of observations. This is not the case of the DLHS as all the waves do not offer these two measures, but Wave 3 does. We then have to focus on this latter and restrict our analysis to 51,709 women who have answered to the question about their economic activity during the year preceding the survey.[21] We start with Model A where education is absent while we introduce a measure of wealth/poverty that is a relative measure in quintile recoded into three groups.[22] We obtain that poor women have a much higher probability of being childless than other women. Wealth clearly exerts a negative effect on the probability to remain childless. Interestingly enough, once we re-introduce the education of the respondent into this model, poverty continues to have a positive effect on the probability to remain childless but education does not recover its negative effect for low educational attainment. It is a sign that low educational attainment is really capturing the effect that poverty exerts on childlessness in our main regressions.

In Model C, we inspect the effect that the economic opportunities enjoyed by the respondent may have on the probability to remain childless. Wave 3 of DLHS offers a unique variable, 'Occupation,' which accounts for about 97 occupation categories. However, we divide this variable into five categories: no occupation, laborers, low skilled, medium skilled, and high skilled.[23] Before commenting on our results, let us recall that occupation is measured at the time of the survey, it is then only a proxy for the economic opportunities which have been offered to women all along their life. It is then true that birth history may have influenced the professional history of these women. Nevertheless, this does not contradict the fact that present economic occupation is strongly linked to economic opportunities enjoyed in the past. The same remark is true for our poverty measure.

We obtain that those women who enjoyed very good opportunities are more childless than the others. Again, this first result tends to confirm our theoretical model. More interestingly, in Model D, we confront directly the effect of poverty to the effect of economic opportunities. Both variables have a positive effect on childlessness as predicted by our theory. The poorer a person the more chances she has to remain childless while

[21] In all the models of Table 3, we use all the control variables of our main model developed in the previous section except the husband's education to avoid a too strong collinearity with the additional variables we introduce. All the qualitative results we reach remain valid when we introduce male education but significance levels may change. All these results are available upon request.

[22] Technically speaking, we have recoded the Wealth Index into three categories: Poor which gathers Poor and Second, Middle which corresponds to Middle, and Rich which gathers Fourth and Richest.

[23] Classification details available upon request.

the more she enjoyed favorable economic opportunities the more chances to remain childless too.

In Model E, we reintroduce educational attainment of the respondent. We get that poverty continues to exert a negative effect on the probability to remain childless while education has a pure positive effect on this latter. Reversely, this pure positive effect of education annihilates the positive effects of economic opportunities on childlessness. Said differently, high education levels and highly skilled occupations capture the same effect.

5 External validity

In this section, we take advantage of the recent publication of the fourth wave of Demographic and Health Survey (DHS) in India. We use this dataset both because it is a recent one and because, for the first time, it covers all the districts in India while providing information on important indicators like education, marriage, family, health and others. The dataset is given by the International Institute for Population Sciences (IIPS), Mumbai, under the stewardship of the Ministry of Health and Family Welfare. In the next paragraphs, we evidence how DHS4 data confirm both the validity of our theory and the results obtained in our identification checks.

From DHS4, we have extracted information which are comparable to these we used in previous sections: childlessness is measured in the same way as well as state of residence. Education is grouped in six levels: no education, incomplete primary, complete primary, incomplete secondary, complete secondary, higher. We have built a cohort variable gathering respondent by 2 years of birth from 1967 to 1975. DHS4 proposes a wealth measure in 5 quintiles while the occupation of women is divided in the same way as in Section 4 except that a category unemployed has been added. It concerns persons who declared to be unemployed new job seekers: no occupation, unemployed, laborers, low skilled, medium skilled and high skilled. Our results are again controlled for the caste, child marriage and religion.

Models a and b both test the main prediction of our theory: education exerts a U-shape influence on the probability to ends reproductive life childless among Indian women. In Model a, it is remarkable to see how the U-shape relationship does appear among the most recent generations in a set-up where the number of educational groups is extended to 5 instead of 3. In Model b, we limit the analysis to the persons for

Table 3: Determinants of the probability to end reproductive life childlessness - DHS4

VARIABLES	Model a	Model b	Model c	Model d
Education				
No education	1.197**	1.425*	1.288	1.303
Incomplete primary	1.244**	1.217	1.173	1.177
Complete primary	Ref.	Ref.	Ref.	Ref.
Incomplete Secondary	1.106	1.248	1.330	1.312
Complete secondary	1.374***	1.246	1.423	1.365
Higher	1.667***	1.635**	1.947***	1.759**
Wealth Index	-	-	0.871***	0.856***
Occupation				
No occupation	-	-	-	Ref.
Unemployed	-	-	-	0.757
Labourer	-	-	-	1.359
Lowly skilled	-	-	-	0.738
Medium skilled	-	-	-	0.827
Highly skilled	-	-	-	1.413
Teen marriage	YES	YES	YES	YES
Religion	YES	YES	YES	YES
Caste	YES	YES	YES	YES
Fixed effects				
Cohort FE	YES	YES	YES	YES
State FE	YES	YES	YES	YES
Number of obs.	126,402	21,855	21,855	21,855

Notes: Odds-ratio reported. *** p<0.01, ** p<0.05, * p<0.1

who the information about occupation is available. It reduces severely the number of observations but it allows to perform identification checks equivalent to those proposed in Section 4. Results of model b are qualitatively the same as model a, the reduction in the number of observations nevertheless makes only the categories no school and higher education significant.

Interestingly enough, once we introduce the relative wealth index proposed by DHS, the negative impact of education on the probability to remain childless among less educated women looses all its significance. This result echoes quite well the one obtained with the DLHS data in the previous section. Furthermore, once we introduce the occupational status of the respondent, this variable has no effect on the probability to remain childless while the positive effect of education among highly educated persons persists. Again, this result echoes remarkably well the one obtained with DLHS data.

6 Conclusion

At around 7 percent, the Indian childlessness rate is not the highest in the world; this being said, childlessness concerns more than 12 million women above 40 years of age. This is clearly an issue in India, and yet it remains underrated and not explored enough. We have extended the theoretical framework developed by Baudin *et al.* (2015) to the Indian context. We show that our main hypothesis holds true: once controlling for micro and macro factors, the relationship between the probability of remaining childless and a woman's educational attainment is U-shaped. As confirmed by identification analysis, this U comes from the opposite effect that education has on poverty faced by women and the economic opportunities they may enjoy. Indian data are sometimes criticized regarding their reliability, especially DLHS data and its first wave. Reassuringly enough, we have shown that our main hypothesis are validated using the last wave of DHS data.

One could argue that very few women are highly educated in India, but this is not accurate. In highly developed states like Kerala, more than 60% of women aged between 40 and 50 have at least some years of high school, and 7.5% spent some years at university.[24] In the state of Goa, we find that more than 75% of women between 40 and 50 have at least completed high school, while more than 15% have a university degree or some years at university. These states prefigure the future of education for Indian women. The democratization of education can already be diagnosed comparing the

[24]The data come from our sample.

educational attainment of alternative age groups in our sample. For India as a whole, 6.35% of women aged between 40 and 50 went to university, while they are respectively 7.83% and 8.97% in the 30-40 and 25-30 age groups. This movement can be observed despite some variance in all the states of India. Our results prefigure the future of childlessness in India; it will be less and less poverty and sterility related and more and more opportunity related.

While this paper is part of a recent literature showing how Indian families are changing rapidly, it leaves many questions unanswered. We believe that two of them are key: is celibacy a way to avoid childbirth in modern India? What are the economic, social, and psychological consequences of remaining childless in India?

References

Abma, J. and Martinez, G. (2006). Childlessness among older women in the united states: Trends and profiles. *Journal of Marriage and Family*, **68**, 1045–1056.

Allahbadia, G. N. (2013). Ivf in developing economies and low resource countries: An overviewl. *The Journal of Obstetrics and Gynecology of India*, **63**, 291–294.

Allendorf, K. (2013). Going nuclear? family structure and youg women's health in india, 1992-2006. *Demography*, **50** (3), 853–880.

— and Pandian, R. K. (2016). The decline of arranged marriage? marital change and continuity in india. *Population and Development Review*, **42** (3), 435–464.

— and Thornton, A. (2015). Caste and choice: the influence of developmental idealism on marriage behavior. *American Journal of Sociology*, **121** (1), 243–287.

Banerji, M. (2012). Fertility as mobility in india: Salience of caste, education and employment opportunities, department of Sociology, University of Maryland.

Basu, A. M. and Desai, S. (2016). Hopes, dreams and anxieties: India's one-child families. *Asian Population Studies*, **12** (1), 4–27.

Baudin, T., de la Croix, D. and Gobbi, P. E. (2015). Fertility and childlessness in the us. *American Economic Review*, **105** (6), 1852–1882.

—, — and — (2019a). *Childlessness and Economic Development: A Survey*, in **Human Capital and Economic Growth: The Impact of Health, Education and Demographic Change**, A. Bucci, A. Prskawetz, and K. Prettner (Eds.) Palgrave - MacMillan.

—, — and — (2019b). Endogenous childlessness and the stages of development. *Journal of the European Economic Association*, **Forthcoming**.

Becker, G. S. . (1981). *A treatise on the family*. Cambridge, MA, 30: Harvard University Press.

Bhambhani, C. and Inbanathan, A. (2018). Not a mother, yet a woman: Exploring experiences of women opting out of motherhood in india. *Asian Journal of Women's Studies*, **24** (2), 159–182.

Birdsall, N., Kelley, A. and Sinding, S. (2001). *Population matters: demographic change, economic growth, and poverty in the developing world*. Oxford University Press.

Chakraborty, T. and Bakshi, S.-K. (2016). English language premium: Evidence from a policy experiment in india. *Economics of Education Review*, **50**, 1–16.

Chatterjee, E., Desai, S. and Vanneman, R. (2018). Indian paradox: rising education, declining women's employement. *Demographic research*, **38** (31), 855–878.

Chiappori, P. A. (1988). Rational household labor supply. *Econometrica*, **56** (1), 63–90.

Chiappori, P.-A. and Donni, O. (2011). Non-unitary models of household behavior: A survey of the literature. In A. Molina (ed.), *Household Economic Behaviors*, Springer, New York, pp. 1–40.

Doblhammer, G. (2000). Reproductive history and mortality later in life: a comparative study of england and wales and austria. *Population studies*, **54** (2), 169–176.

Dommaraju, P. (2009). India's north-south divide and theories of fertility change. *Journal of Population Research*, **26**.

Duraisamy, P. (2002). Changes in returns to education in india, 1983–94: by gender, age-cohort and location. *Economics of Education Review*, **21** (6), 609–622.

Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, **80** (1), 27–38.

Frank, O. (1983). Infertility in sub-saharan africa: estimates and implications. *Population and Development Review*, **9**, 137–144.

Ghosh, S. (2015). Second demographic transition or aspirations in transition: An exploratory analysis of lowest low fertility in kolkata, india. *Asia Research Centre working paper 68*, pp. 1–42.

Goldscheider, C. and Uhlenberg, P. R. (1969). Minority group status and fertility. *American Journal of Sociology*, **74** (4), 361–372.

Heckman, J. (2008). Econometric causality. *International statistical review*, **76** (1), 1–27.

Kijima, Y. (2006). Why did wage inequality increase? evidence from urban india 1983-99. *Journal of Development Economics*, **81** (1), 97–117.

King, G. and Zeng, L. (2001). Logistic regression in rare events data. *Political Analysis*, **9**, 137–163.

Koropeckyj-Cox, T. and Pendell, G. (2007). Attitudes about childlessness in the united states: Correlates of positive, neutral, and negative responses. *Journal of Family Issues*, **28** (8), 1054–1082.

Kulkarni, P. and Alagarajan, M. (2005). Population growth, fertility, and religion in india. *Economic and Political Weekly*, **40**.

Kumar, P. and Gupta, A. (2015). Determinants of inter and intra caste differences in utilization of maternal health care services in india: Evidence from dlhs-3 survey. *International Research Journal of Social Sciences*, **4**.

Leridon, H. (2008). A new estimate of permanent sterility by age: sterility defined as the inability to conceive. *Population Studies*, **62** (1), 15–24.

Long, J. S. and Freese, J. (2006). *Regression models for categorical dependent variables using Stata.* Stata Press.

Madhivanan, P., Bartman, M., Pasutti, L., Krupp, K., Arun, A., Reingold, A. and Klausner (2009). Prevalence of trichomonas vaginalis infection among young reproductive age women in india: implications for treatment and prevention. *Sexual Health*, **6**, 339–344.

Malpani, A. and Malpani, A. (1992). Simplifying assisted conception techniques to make them universally available-a view from india. *Human Reproduction*, **7**, 49–50.

McFalls, J. A. (1979). Frustrated fertility: A population paradox. *Population Bulletin*, **34** (2), 3–43.

Mehta, C. R. and Patel, N. R. (1995). Exact logistic regression: theory and examples. *Statistics in medicine*, **14** (19), 2143–2160.

Nanda, S. and Garden, P. (2009). Effect of women's nutritional status on fertility: study of a non-industrial population in orissa. *Man In India*, **89** (4), 525–534.

Ombelet, W., Cooke, I., Dyer, S., Serour, G. and Devroey, P. (2008). Infertility and the provision of infertility medical services in developing countries. *Human Reproduction Update*, **14** (6), 605–621.

Pew-Research-Center (2015). Childlessness falls, family size grows among highly educated women.

Poston, D. L., Briody, J. E., Trent, K. and Browning, H. L. (1985). Modernization and childlessness in the states of mexico. *Economic Development and Cultural Change*, **34** (3), 503–519.

— and Cruz, C. E. (2017). Voluntary, involuntary and temporary childlessness in the united states. *Revue Quetelet/Quetelet Journal*, **4.1**.

— and Trent, K. (1982). International variability in childlessness: A descriptive and analytical study. *Journal of Family Issues*, **3** (4), 473–491.

Ram, U. (2005). Childlessness in time, space and social groups and its linkages with fertility in india, paper presented at the EPC Conference 2006.

Ramesh, P. (2014). An analysis of fertility differentials among caste groups in andhra pradesh, working paper number - 1350, Gokhale Institute of Population and Economics.

Rasool, S. and Akhtar, O. S. (2018). The huge burden of infertility in india: Are we crumbling underneath? *Global Journal of Reproductive Medicine*, **5**.

Retel-Laurentin, A. (1974). *Infécondité en Afrique noire : maladies et conséquences sociales*. Masson.

Romaniuk, A. (1980). Increase in natural fertility during the early stages of modernization: Evidence from an african case study, zaire. *Population Studies*, **34** (2), 293–310.

Sen, A. K. and Nussbaum, M. (1993). *The Quality of Life*, Clarendon Press, chap. Capability and Well-Being, pp. 30–53.

Singh, A., Kumar, K., Chauhan, P. and Banerjee, A. (2017). Spatial patterns and determinants of fertility in india. *Population*, **72** (3), 505–526.

Sobotka, T. (2017). *Childlessness in Europe: Reconstructing long-term trends among women born in 1900-1972*. Springer.

Solomon, S., Subbaraman, R., Solomon, S. S., Srikrishnan, A. K., Johnson, S. C., Vasudevan, C., Anand, S., Ganesh, A. K. and Celentano, D. D. (2009). Domestic violence and forced sex among the urban poor in south india: Implications for hiv prevention. *Violence Against Women*, **15** (7), 753–773.

Sujata Ganguly, S. U. (2010). Trends of infertility and childlessness in india: Findings from nfhs data. *Facts, Views and Vision in ObGyn*, **2**, 131–138.

Surkyn, J. and Lesthaeghe, R. (2004). Value orientations and the second demographic transition (sdt) in northern, western and southern europe: An update. *Demographic research*, **3**, 45–86.

Tilak, J. (2002). Education and poverty. *Journal of Human Development*, **3** (2), 191–207.

— (2007). Post-elementary education, poverty and development in india. *International Journal of Educational Development*, **27** (4), 435–445.

Toulemon, L. (1996). Very few couples remain voluntarily childless. *Population*, **1-27**.

Veevers, J. E. (1972). Factors in the incidence of childlessness in canada: An analysis of census data. *Social Biology*, **19**, 266–274.

Weiss, H. A., Patel, V., West, B., Peeling, R., Kirkwood, B. and Mabey, D. (2008). Spousal sexual violence and poverty are risk factors for sexually transmitted infections in women: a longitudinal study of women in goa, india. *Sexually Transmitted Infections*, **84**, 133–139.

Wolowyna, J. E. (1977). Income and childlessness in canada: A further examination. *Social Biology*, **24** (4), 326–331.

Wunsch, G. (2007). Confounding and control. *Demographic research*, **16**, 97–120.

A Data

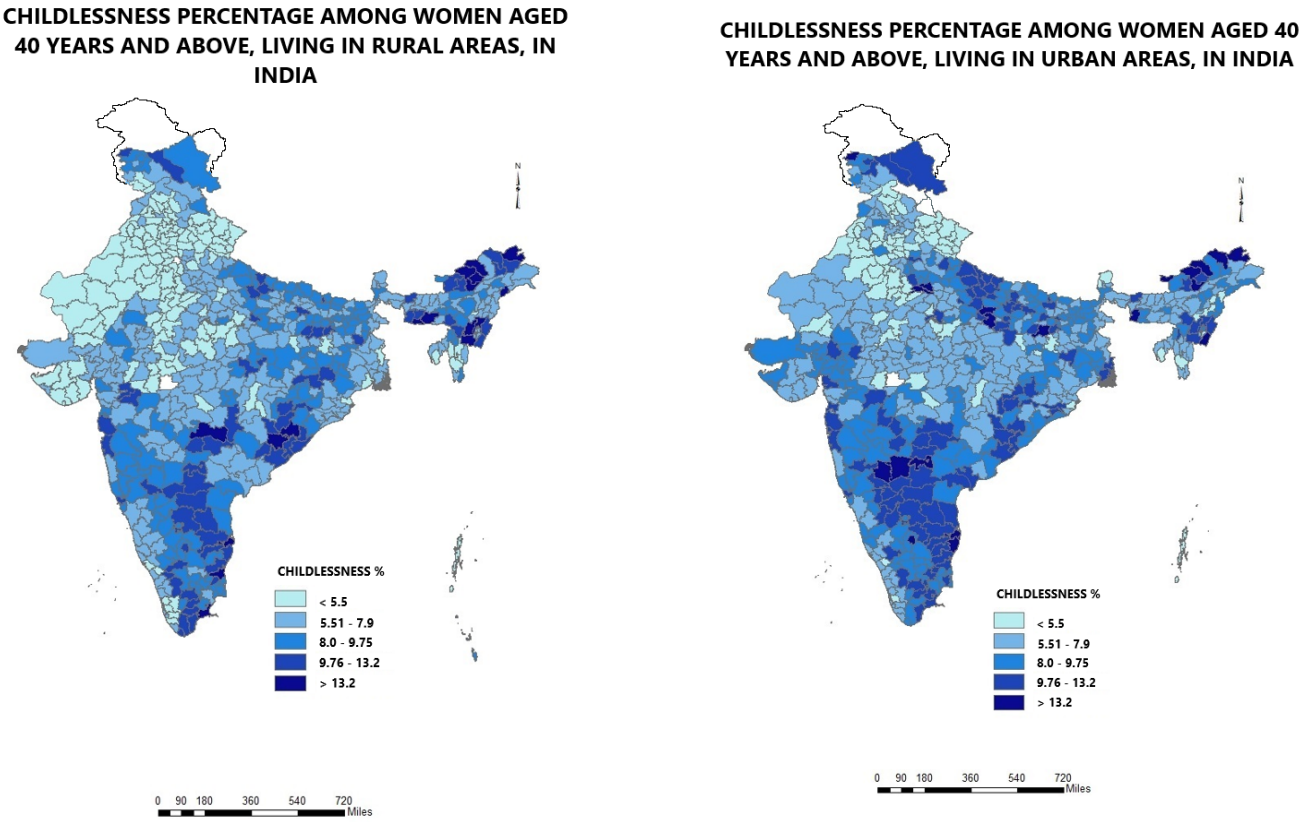

Figure 6: Definitive Childlessness in rural areas (left) and urban areas (right) in Indian Districts.

A.1 The Desire for No Children

Even if we dispense with attitudinal concepts in our theory and analysis, some attitudinal data exist in the DLHS. We explore them to check whether women were able to express reasons for their childlessness which were not related to reproductive issues, poverty, or violence. We conduct this analysis to again highlight how causes for childlessness in India cannot be limited to sterility and poverty. These data echo the facts we discuss at the end of Subsection 1.1.

We focus here on the variable 'desire for no children' among women with zero fertility in India. Figure 7 (left) shows that even if the percentage is low (2.34 percent, 1,157 women), the desire for no children exists in India. At the state level, this desire for no children is negatively correlated with a fertility problem (center panel) and positively correlated with definitive childlessness (right panel). This is indicative of the fact that the whole of Indian childlessness, as has been conceptualized until now (Ram, 2005 and Sujata Ganguly, 2010), may not be driven solely by infertility issues.[25]

[25]In the literature, Ram, 2005 and Sujata Ganguly, 2010 are among the very few who study infertility

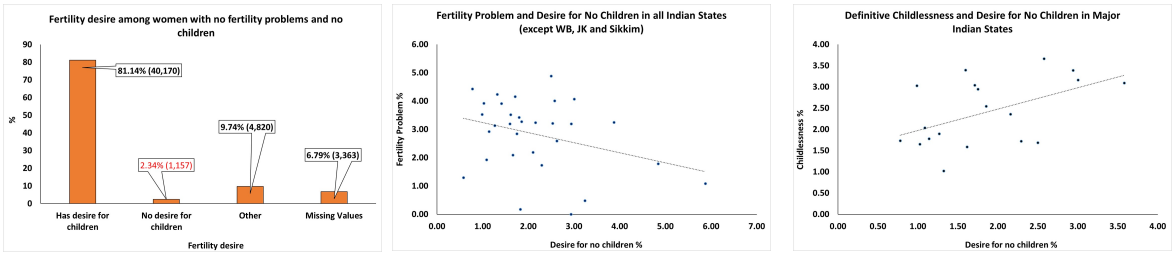

Figure 7: Fertility desire among women (left), Correlation between fertility problem and desire for no children in all Indian states (center), and Correlation between definitive childlessness and desire for no children in major Indian states (right). Source: DLHS 1998-99, 2002-04, and 2007-08

B Regressions

B.1 Age of entry into marriage

In Table 4, we use the age of entry into marriage instead of only teen-marriage.

and childlessness and give a comprehensive idea of the Indian context in this regard. However, an exploration of whether childlessness is solely driven by infertility, sterility, poverty, or opportunity seems to have been overlooked by Indian demographers.

Table 4: Determinants of childlessness - Using age of entry into marriage

Variable	Model 1	Model 2	Model 3	Model 4
Education				
Secondary Education	Ref.	Ref.	Ref.	Ref.
No education	1.40***	1.41***	1.42***	1.42***
Primary	0.92	1.06	1.06	1.06
Higher	1.38***	1.02	1.02	1.02
Husband's Education				
No education	Ref.	Ref.	Ref.	Ref.
Primary		0.74***	0.74***	0.74***
Secondary Education		0.72***	0.72***	0.72***
Higher		0.59***	0.58***	0.58***
Teen Marriage				
Less than 18 years	Ref.	Ref.	Ref.	Ref.
18 to 29 years		1.68***	1.68***	1.68***
30 years and above		25.05***	25.12***	25.13***
Place of Residence				
Rural	Ref.	Ref.	Ref.	Ref.
Urban		0.97	0.98	0.98
Religion				
Hindu	Ref.	Ref.	Ref.	Ref.
Muslim			0.82**	0.82**
Christian			0.85	0.85
Sikh			1.04	1.04
Buddhist			0.81	0.81
Others			1.2	1.2
Caste				
General	Ref.	Ref.	Ref.	Ref.
SC			1	1
ST			1.11	1.11
OBC			1.01	1.01
Others			1.13	1.13
State Development Level				
Developed States	Ref.	Ref.	Ref.	Ref.
Least developed States				0.83
Intermediate States				0.33***
Fixed Effects				
Cohort FE	YES	YES	YES	YES
State FE	YES	YES	YES	YES
Pseudo R2	0.0126	0.0706	0.0711	0.0711

B.2 Artificial samples and goodness of fit

As explained in the core of the paper, the low childlessness rates prevailing in our sub-sample explain the low values of our measures of goodness-of-fit. This problem is known as the rare event phenomenon in case of small samples, see King and Zeng (2001) and the blog of Paul Allison for a discussion. In our case, the total sample size is not small but as we use both state and cohort fixed effects, our sample is supposed to explain the probability of being childlessness for a given cohort in a given state, what reduces the size of each population tremendously and makes the number of events (women being childless) much smaller. Alternative methods of estimation exist, they are supposed to fix this issue but they have their own problems. For instance, the penalized likelihood estimation of a logistic model proposed by Firth (1993) suffers over-estimation bias of the coefficient of the regressions. In a paper like ours, it means that we could attribute meaning to meaningless variables. The exact-logistic regression model proposed by Mehta and Patel (1995) works only when the number of observations is below 200 because it is too demanding in terms of computing power.

As an alternative, we propose here to build artificial datasets in which we systematically keep all childless women while we select the non-childless persons randomly. We then show that the quality of our fit becomes much more satisfying. To build our three alternative datasets, we have generated a random variable following a standard normal distribution. In sample 1, we have kept all the persons who have drawn a value between -0.5 and +0.5 reducing the sample size to 58603 observations without reducing the number of childless women. In sample 2 and 4 respectively, we have kept women who have drawn a value between -0.1 and +0.1 and -0.05 and +0.05; sample sizes then become 9024 and 6141.

	Benchmark regressions			With age at marriage		
Measure of Fit	Sample 1	Sample 2	Sample 3	Sample 1	Sample 2	Sample 3
Age at marriage	NO	NO	NO	YES	YES	YES
M^c Fadden R^2	0.024	0.034	0.036	0.077	0.084	0.080
Count R^2	0.944	0.645	0.590	0.946	0.692	0.616
Adj-Count R^2	0.000	0.023	0.121	0.028	0.153	0.175

Note: Variables of regressions omitted for sake of space and clarity.

If we see that the improvements in Mc Fadden R^2 are real but limited, the main change comes from count-R^2. The count-R^2 measures the number of well-predicted cases over the total number of observations. It is reputed to be spuriously high when the event to be predicted is very rare because a lack of variance. To fix this bias, one can use the number of events which are well predicted beyond the largest marginal (the most common event which is not being childless). Long and Freese (2006) define this adjusted measure of fit as follows: "*The adjusted count R2 is the proportion of correct guesses beyond the number that would be correctly guessed by choosing the largest marginal.* One can notice that once we reduce our sample size, our model provide satisfying performances.