



**HAL**  
open science

# Sons as Widowhood Insurance: Evidence from Senegal

Sylvie Lambert, Pauline Rossi

► **To cite this version:**

Sylvie Lambert, Pauline Rossi. Sons as Widowhood Insurance: Evidence from Senegal. 2014. halshs-00948098v1

**HAL Id: halshs-00948098**

**<https://shs.hal.science/halshs-00948098v1>**

Preprint submitted on 17 Feb 2014 (v1), last revised 1 Oct 2015 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**PARIS SCHOOL OF ECONOMICS**  
ÉCOLE D'ÉCONOMIE DE PARIS

**WORKING PAPER N° 2014 – 05**

## **Sons as Widowhood Insurance: Evidence from Senegal**

**Sylvie Lambert**

**Pauline Rossi**

**JEL Codes : D13, I15, J13, O15**

**Keywords: Intra-household insurance, Gender, Fertility, Health, Senegal**



**PARIS-JOURDAN SCIENCES ÉCONOMIQUES**

48, Bd JOURDAN – E.N.S. – 75014 PARIS  
TÉL. : 33(0) 1 43 13 63 00 – FAX : 33 (0) 1 43 13 63 10  
[www.pse.ens.fr](http://www.pse.ens.fr)

# Sons as Widowhood Insurance :

## Evidence from Senegal.

Sylvie LAMBERT and Pauline ROSSI\*

February, 2014<sup>†</sup>

### Abstract

Exploiting original data from a Senegalese household survey, we provide evidence that fertility choices are partly driven by women's needs for widowhood insurance. We use a duration model of birth intervals to show that women most exposed to the risk of widowhood intensify their fertility until they get a son. Insurance through sons entails substantial health costs : short birth spacing raises maternal and infant mortality rates.

**Keywords** : Intra-household insurance, Gender, Fertility, Health, Senegal.

**JEL Codes** : D13, I15, J13, O15.

---

\*Sylvie Lambert: PSE - INRA (sylvie.lambert@ens.fr); Pauline Rossi: PSE and CREST (pauline.rossi@ensae.fr)

<sup>†</sup>Dominique van de Walle sensitized us to the situation of widows in Sub-Saharan Africa and qualitative field work in Senegal in her company proved very inspiring. The paper benefited from early discussions on son preference in Africa with Annamaria Milazzo. The authors are grateful for the comments of participants to seminars at LMI - CREST and Namur University, to conferences at Paris 1, ISS, NEUDC, and Nanterre University. For their comments we are also grateful to Xavier d'Haultfoeuille, Guilhem Cassan, Nicolas Pistoiesi, Maelys de la Rupelle, Eric Strobl and Dominique van de Walle. We thank IDRC (International Development Research Center), INRA Paris and CEPREMAP for funding data collection and the World Bank for funding qualitative field work.

# 1 Introduction

In many developing countries, the lack of insurance markets and social safety nets is mitigated by family. A large literature describes the use of family as insurance, mentioning for instance family income diversification strategy through migration (Rosenzweig, 1988), or children acting as old age insurance for their parents (Nugent, 1985; Hoddinott, 1992).

In this paper, we focus on one mechanism whereby women rely on sons as insurance in case of widowhood in Senegal. In this country, women are highly exposed to the risk of losing their husbands, in particular because of the age difference between spouses : 10 years on average, and often much more. The question of the means available to women to cope with the impact of widowhood on living standards is a very urgent one. Formal mechanisms are weak : finance and insurance markets are very imperfect, social safety nets are missing, and opportunities for self-support are limited due to labor market restrictions and constraints stemming from social norms. In this context, children, especially sons, turn out to be women's best claim to their late husbands' resources in case of widowhood. Indeed, a widow's access to those resources depends on the number and gender composition of her husband's children. In Senegal, where 95% of the population is Muslim, inheritance practices are ruled by Islamic and customary, patrilineal laws. Wives are *de facto* excluded from a bequest after the death of their husband ; the inheritance is to be shared among the husband's children, sons inheriting more frequently and quantitatively than daughters (Lambert, Ravallion and van de Walle, 2013). When the husband has children from other women, they compete for the man's inheritance. In particular, the number and gender composition of the children of the current wife, relative to other children the husband might have, determine the share of inheritance she might control *de facto* ; it influences, in turn, the possibility to carry on residing in the late husband's house. Women might therefore have incentives to be pregnant often in order to maximize the number of sons they have.

Demonstrating such a mechanism may matter for two main reasons. First, frequent pregnancies are known to cause adverse maternal and child health outcomes. In Senegal, maternal and infant mortality rates are a matter of concern : out of each 1,000 births, 4 mothers died as well as 47 children before turning a year old (among which, 29 passed away before 1 month) over the

period 2007-2011 (ANSD and ICF International, 2012). Reducing maternal and child mortality as well as implementing birth control are two key objectives of the Senegalese national health development plan (called PNDS-II, covering 2009-2018). On the supply side, the budget of the Health Ministry was multiplied by 3.5 between 2000 and 2010, reaching 108.4 billions FCFA (about 220 millions dollars), and now accounts for 10.4% of the State budget (ANSD and ICF International, 2012). A great deal of effort has been made to increase access to maternal health care services, in particular tetanus toxoid injections to pregnant mothers and child delivery assistance by skilled professionals. However, efficient supply-side interventions require understanding people's behaviors : why do frequent pregnancies persist at such a large scale ? Women's needs for widowhood insurance might be part of the answer.

Second, the costs of informal family insurance are rarely studied ; it is generally assumed that taking part in family insurance networks has no cost other than reciprocity. Recent papers have focused on the monetary costs for relatively high income members, and on the distortions generated by the "kin tax" in African economies (Jakiela and Ozier, 2012). The monetary costs are often considered negligible for low income members. Non-monetary costs have been totally overlooked, and yet they might be important from a welfare perspective. The cost of widowhood insurance through sons is potentially huge : health risks related to frequent pregnancies are well-documented in the medical literature on developing countries. Those risks are well-known : increasing the interval between births up to 24 months is a rationale for many family planning programs<sup>1</sup>. On the child side, short birth intervals are associated with an increased risk of adverse perinatal outcomes : lower birth weight, smaller infants for gestational age, higher risk of preterm delivery, stillbirth and neonatal death. A meta-analysis of medical studies on developed and developing countries concludes that the association between short birth spacing and adverse outcomes remains significant even after controlling for maternal characteristics and socioeconomic status (Conde-Agudelo et al., 2006). In Senegal, children born less than 2 years after a sibling are 2.4 times more likely to die than the others, whatever mortality rates we consider : neonatal, post-neonatal, infant, juvenile, infant-

---

1. Some programs such as the Optimal Birth Spacing Initiative, created under the auspices of USAID, even recommend an optimal birth spacing of 3 to 5 years.

juvenile (ANSD and ICF International, 2012)<sup>2</sup>. On the mother's side, Conde-Agudelo and Belizan (2000) studied data from 456,889 women delivering singleton infants in Latin America and the Caribbean between 1985 and 1997. After controlling for a wide range of maternal characteristics<sup>3</sup>, they found that birth intervals lower than 15 months are associated with an increased risk of adverse maternal outcomes : third trimester bleeding, premature rupture of membranes, puerperal endometritis and anaemia; mothers are also 2 times more likely to die. Another study following 21,244 pregnancies in Matlab, Bangladesh, between 1996 and 2002 finds very similar results : women with birth intervals lower than 15 months bear a significantly higher risk of maternal complications (DaVanzo et al., 2004). One medical explanation is the maternal depletion hypothesis, stating simply that short intervals do not allow the mother to recover from the physiological stresses imposed by the previous pregnancy. In the economic literature, Milazzo (2012 and 2013) finds that son preference induces short birth spacing, and might be a significant cause of female excess adult mortality in Nigeria and in India.

In this paper, we investigate whether fertility choices are partly driven by women's insurance strategies. We focus here on the sequential nature of male monogamous unions : rivalry for inheritance occurs between children of the current wife and children of ex-wives, whether divorced or deceased. Our hypothesis is that the existence of children of ex-wives (henceforth called rivals) should exacerbate son preference and intensify fertility. Rivalry may also arise in polygamous unions between children of co-wives, but we chose not to consider them for two reasons. First, our identification strategy fails to deal with simultaneity issues : in polygamous unions, women might adjust the number of their children to react to a co-wife's births, so that the rivals of a given wife cannot be considered as exogenous to her own fertility choices. Second, it can be argued that co-wives rely less on their children to stay in the husband's house after his death since they may agree to continue to live there together (see qualitative evidence in Lambert and van de Walle, 2012).

---

2. Ronsmans (1996) shows that in rural Senegal, short birth intervals also increase the odds of dying for the preceding child : if the mother delivers another child within 2 years, the risk of mortality for the index child is 4 times higher in the second year of life than if the next birth takes place more than 2 years after his birth. However, this correlation is not necessarily causal : short subsequent birth intervals may be either a cause (abrupt weaning) or a consequence (willingness to "replace" a dead child) of child mortality.

3. Maternal age, number of previous deliveries, history of miscarriage, stillbirth and early neonatal death, previous caesarean delivery, marital status, education, cigarette smoking, body mass index before pregnancy, trimester during which prenatal care was started, number of prenatal care visits, geographic area, hospital type, and year of delivery.

Exploiting original data from a nationally representative household survey, we develop a strategy that is close to a difference-in-difference framework. We analyze how the gender composition of the first children influences subsequent choices of women with and without rivals. Both groups of women might be different ; what matters to our identification strategy is that (i) the gender composition is exogenous, and (ii) in the absence of rivals, both groups would have reacted similarly to this exogenous shock : having one son vs. one daughter. Using such strategy to identify the existence of widowhood insurance motive for fertility is very similar in spirit to Nugent's (1985) suggestion : in order to evidence old-age security motive, one has to compare similar populations with different intensity of the motive. Here, among women with rivals, we compare those having a son and those having only daughters ; they are similar, but have different intensities. Then, we factor in women without rivals to cancel out any son preference prevailing in the population, but not related to widowhood insurance.

Building on demographers' methodology, we use non-, semi- and fully-parametric estimations of a duration model of birth intervals to show that women with rivals and no son do indeed tend to have another child more quickly. When there are rivals, having only girls versus having at least one son multiplies by 1.5 to 2 the probability of short birth spacing (intervals shorter than 24 months). In the absence of rivals, the gender composition of first births does not affect subsequent birth intervals. Pushing the logic further, we show that the effect is clearly evidenced for women who co-reside with their husband, in particular for the poorest half of the population, comforting the prior that the house is the main stake in the bequest process. Also, fertility choices are more strongly affected after the third child : one interpretation is that the vast majority of women want at least three children and would start feeling the pressure to have a son only after the third « missed try ».

All in all, we find that son preference only appears in the presence of rivals, especially when women have no independent housing. Women's needs for widowhood insurance seem therefore to play a role in fertility choices, leading to more frequent pregnancies for the most exposed women : those with rivals and no son. This suggests that the lack of formal widowhood insurance hampers birth control efforts, and even more worryingly, imposes a potentially heavy health cost on both women and children, whether widowhood occurs or not. Such non-monetary costs point to the fact

that, even if it plays its role in terms of access to resources in case of widowhood, such informal insurance system is a poor substitute to a formal insurance scheme. The benefits of implementing some public widowhood insurance scheme, or changing inheritance laws, would thus be far reaching.

The outline of the paper is as follows. Section 2 provides background on widowhood and inheritance practices in Senegal. Section 3 presents the data and some descriptive statistics. Section 4 discusses the empirical strategy, the identification assumptions and the methods of estimation. Section 5 reports the main results. And section 6 concludes with policy implications and directions for future research.

## 2 Widowhood and Bequest in Senegal

### 2.1 An ex-ante risk-coping strategy based on children

The general idea is that women's vulnerability in Sub-Saharan Africa is conducive to large family size<sup>4</sup>. Family law renders women dependent on spouses, or on adult children in case of death or abandonment by the husband. Women have therefore an interest in early marriage and high fertility. Anthropologists such as Bledsoe (1990) also endorse the idea that in all societies, people make efforts to construe marital status and parenthood to their advantages. She claims that in Africa, as long as children remain women's best claim to male resources, women will continue to want many children.

Demographers working on Sub-Saharan Africa have come to the same conclusion : among the key factors shaping the reproductive regime in this region, they mention inheritance patterns and women's subordinate status. In patrilineal societies, in which wives "belong" to the male kin, women's security and status critically depend on their ability to have sons (Lesthaeghe, 1989).

All this requires that fertility is an active choice of women. In Senegal, modern contraceptive methods are widely known, but not much used : over 92% of married women can name at least one modern method (sterilization, oral contraceptive, implants, IUD, condom) ; but only 12% - mostly urban, educated and wealthy women - report using one of them. However, only 7.5% of married

---

4. In Senegal, fertility is still high : 5 is both the average and the median number of children among women older than 45 years (cf. Figure 4 in Appendix A).



women report unmet needs for birth control to limit the number of children. This proportion is even smaller for younger women : around 4.5% for women between 15 and 39 years old. Unmet needs for birth control to increase intervals between births are higher, though (22%). Another way to assess whether women do indeed have some control over their fertility is to look at the percentage of unwanted or unplanned pregnancies. Only 4% of pregnancies that occurred during the period 2005-2010 are reported by the mothers as unwanted and 20% as unplanned (ANSD and ICF International, 2012). A majority of Senegalese women therefore claim to be choosing the number of children they have, and to a lesser extent the timing of births. One explanation is that birth spacing is partly determined by breastfeeding amenorrhea and abstinence, and breastfeeding is in the domain of the mother. So women may have considerable control over the spacing of births, even in the absence of modern contraceptive methods, as argued by some anthropologists (Mace and Sear, 1997). In Senegal, the median postpartum non-susceptible period (related to amenorrhea and/or abstinence) is 14 months, and the average 15 months (ANSD and ICF International, 2012).

## 2.2 The bequest process

As described in Lambert *et al.* (2013), the Islamic inheritance law is the dominant rule in Senegal, despite the existence of alternative possibility ; in particular, wills and testaments are rarely made, and even less implemented. *De jure*, the Islamic rule entitles wives to 1/8 of the total bequest, to be shared among all wives in case of polygamous union. In practice, women are generally excluded from the bequest. Further, this inheritance system splits the bequest between the children <sup>5</sup>, limiting the inheritance of daughters to half of that of sons. In addition, entrenched tradition favors sons for inheriting land <sup>6</sup>. Since daughters typically move to their husband's abode, they are supposedly compensated by their brothers with money or other forms of wealth, for what would have otherwise been their share of land inheritance ; the same holds with respect to inheriting the father's house. But in practice, the compensation does not always take place. In our data, among people who lost their father, we find that daughters are significantly less likely to report that the father left a

---

5. A man's extended family (uncles, cousins, nephews, etc.) is entitled to a share of the bequest only if he died childless.

6. With a small exception in Basse Casamance where there is slightly more access to land for women, due to strong tradition in women-dominated rice cultivation.

bequest, and among those who do, daughters are 2.5 times more likely to have not received anything (10% vs. 4%); they are also less likely to inherit land or a house, but not significantly more likely to inherit money, which contradicts the compensation story.

Another reason why daughters might be deprived of their rights is the non-implementation of the inheritance process. Typically, brothers and their own family may stay together in the father's house after his death; as long as they are willing to live together, they do not feel the need of an actual sharing to take place and may consider that the father left no bequest. Among people who reported no bequest, we find that sons are significantly more likely to live in the household of their late father. If we consider that in such a case, the house is transmitted *de facto*, we can estimate the proportion of people who inherited (part of) a house : 60% of sons vs. 49% of daughters. This might not seem such a large imbalance. Nevertheless, when one looks at women at the time of their own death rather than at the time of their parents' death, it appears that women are vastly less likely than men to bequeath a house : less than a quarter of adults whose mother died have inherited anything from her, and only 22% of those who did inherited a house. By contrast, 72% of adults whose father died have inherited something, and 71% of them inherited a house. This significant gap between men and women in the possibility to transfer property rights at the time of their own death suggests that, even if a large fraction of women claim to have inherited a house and hence appear to feel entitled to part of the house property, this might be only a temporary situation. In fact, by the time they die, they seem to have lost any access to it.

In total, men seem to be more successful than women in claiming their inheritance rights and in holding to their property; hence widows facing rivals need a son to insure some access to the bequest.

Further, having a son might not only raise the likelihood of remaining in the house, but could also provide an alternative in case of eviction. Indeed, widows are more likely to be hosted by a son than a daughter : among widows hosted by their children, 92% live with one son vs. 34% with one daughter. If we further consider widows living with exactly one child, 68% live with one son vs. 32% with one daughter, and both situations are very different : widows are hosted by their sons in 2/3 of the cases, while they host their daughters in 2/3 of the cases (being head of household, or

remarried and co-residing with the new husband).

### 2.3 The vulnerability of widows

So far, the economic literature on vulnerability has paid little attention to the situation of widows in Sub-Saharan Africa. Yet, a few papers highlight that women are extremely vulnerable to the loss of their husbands : households headed by widows are significantly poorer than male headed and other female headed households (Appleton, 1996 in Uganda ; Horrell and Krishnan, 2007 in Zimbabwe). In Mali, van de Walle (2011) finds that widowhood has a strong, lasting impact on women's living standards and welfare indicators : the detrimental effects persist over time even if women remarry, and are passed on to children. The vulnerability of widows stems from women's subordinate status regarding legal protection, individual endowments, social norms and access to and control over resources. Historians have also documented that women who lack male support, like widows, are more vulnerable in case of famine (see Vaughan, 1987 for the history of a famine in Southern Malawi in 1949).

Vulnerability is exacerbated when women have rivals for their husband's resources. The question of wives rivalry has only been studied in the context of polygamous households. Researchers working on Mali have documented that co-wife rivalry is responsible for resource-consuming struggles (Kazianga and Klonner, 2006) ; it even adversely affects child survival, especially that of sons since they remain to compete for land (Strassmann, 1997).

In this paper we are interested in the competition with ex-wives, rather than co-wives, through children. The existence of children of at least one ex-wife is a common feature in the Senegalese context where widowhood, divorce and remarriage are very frequent for both men and women. Antoine and Dial (2003) point to the lack of reliable data on divorce in Africa : cross-sectional studies often underestimate the frequency of divorce because remarriage happens quickly. Using a biographical approach, they find that one union in three ends with a divorce in Dakar ; 25% of divorces occur after less than 7 years of marriage, and 50% of divorced women are remarried 5 years later. To our knowledge, there are no studies about the impact of past marriage of the husband on current wife's well-being.

For current wives, the main stake in case of widowhood is probably the husband's house. First, notably for poor households, the house is likely to be the main asset. Second, according to qualitative evidence, women fear being thrown out of the house by the children of an ex-wife and having nowhere to go (Lambert and van de Walle, 2012). For instance, a childless widow reported how she was confronted with the children of her husband after his death : they offered her a room to rent in their father's house at a steep price that she could not pay, and she was forced to leave. Clearly, this fear is relevant only for women that were not housed independently from their husband. Finally, remaining in the house of the late husband can be a way either to avoid the obligation of remarrying, or, in case of remarriage, to keep a relatively high degree of autonomy in this new marriage.

These two dimensions are illustrated in Table 1 in which we provide some statistics on widows, broken down by the place of residence after the husband's death. Widows in our data were on average 46 years old when they lost their husband<sup>7</sup>, and children were born from this union in the vast majority of cases (88%). Two widows out of 3 were able to remain in the same household after the death of their husband, while the other third joined another household. Among those who could remain in the same household, 90% still reside there at the time of the survey : they became head of household, or stepped aside for one of their children. Only 20% got remarried, and among those who did, only 1/3 live with their new husband. The story is completely different for widows hosted by another household at the death of their husband<sup>8</sup>. They generally joined up with their parents, a child or a sibling. The host family provided only temporary accommodation : in 75% of the cases, widows had moved to another place at the time of the survey, after a stay shorter than 2 years for half of them. Consistently, the fraction of remarried women living with the new husband is much higher in this population : 48% remarry, and among those who do, 2/3 live with their new husband. Hence, women who remained in their late husband's house were more likely to have avoided remarriage, and if they remarried to be non-coresiding. For married women, non-coresidence seems to be a favorable situation. In Senegal, nearly a quarter of married women do not co-reside permanently with their husband. It is in particular the case for women who have

---

7. According to World Development Indicators, life expectancy at birth for women in Senegal is 62 years old.

8. Here, the decision to leave the husband's house may be either taken by the widow herself, or imposed on her.

their own house, that may be inherited from a previous husband now deceased. In case spouses do not co-reside, the husband has another dwelling and comes for regular visits. Women who do not co-reside seem to benefit from greater autonomy and enjoy relatively higher per-capita consumption levels (De Vreyer and Lambert, 2013). Therefore, even if other possibilities exist, remaining in the late husband's house is associated with a gain in autonomy for the rest of the life, even in the event of a remarriage.

It is important to note that the argument that women need a son to guaranty access to the husband resources is likely to be less relevant for women without rivals, even if the probability of termination of this union is clearly positive. In Senegal, there is no legal provision for alimony and maintenance allowance for children. Hence divorcees do not obtain anything from their ex-husband whatever the number of children. In addition, at the death of the ex-husband, not only are ex-wives not entitled to any share of the inheritance, but even if her children obtain the house, she would not return to live in the house of a man she divorced, or who repudiated her : first, she is likely to be remarried and would keep the living arrangements she has with her current husband, and second, in such a patrilineal society, it would not be conceivable that she returns as household head in the house of a former marriage that belongs to another lineage. Hence, a first wife who has no rivals, has no incentive to adjust her fertility to mitigate the impact of the termination of that union.

## 3 Data

### 3.1 The survey "Poverty and Family Structure"

The data used here come from an original survey entitled "Poverty and Family Structure" (PSF) conducted in Senegal in 2006-2007 (detailed description in DeVreyer et al., 2008)<sup>9</sup>. It is a nationally representative survey conducted on 1,800 households spread over 150 primary sampling units drawn randomly among the census districts. About 1,750 records can be exploited.

In addition to the usual information on individual characteristics, a comprehensive description

---

9. Momar Sylla and Matar Gueye of the Agence Nationale de la Statistique et de la Démographie of Senegal (ANSD) on the one hand and Philippe De Vreyer (University of Paris-Dauphine and IRD-DIAL) Sylvie Lambert (PSE) and Abba Safir (now with the World Bank) designed the survey. The data collection was conducted by the ANSD.

of the household structure is obtained. Of particular relevance for our purpose is the fact that the survey registers for each child how many half-siblings he has, counting separately siblings from the same father only, siblings from the same mother only and siblings from the same parents. Therefore, we are able to identify the children from previous unions on the mother's side, and the rivals on the father's side : since we consider monogamous unions only, all siblings having the same father but a different mother than the current wife, are children of an ex-wife. There is also information on all children below age 25 of household members, even if they do not live with their parents. Children who died are also reported but there is no information on the timing of births. As a result, a woman's complete birth history for surviving children is available only if all her children are less than 25 years old. Among the children registered, only 59% live with both their parents ; the figure is similar to the national average (ANSD and ICF International, 2012). When parents do not live together, children live with their mothers in 2/3 of the cases. That is why standard household surveys would largely underestimate the number of rivals, or even ignore their existence.

Moreover, detailed information is collected on all current spouses, whether they co-reside or not, and on past marital history of all members (age at first marriage, being or not in the first union). This can be relevant since having previous unions, or children from previous unions, might play a role in determining women's current fertility. Lastly, we use information on education, occupation and income, as additional controls.

### **3.2 Descriptive statistics on women**

The sample consists of 936 women below 39 years old, engaged in a monogamous union, with at least one child from the current union. We had to restrict the sample of interest to relatively young women to ensure that their complete birth history is known, as mentioned above. In our data, the lower bound for age at first birth is 13, so children of women below 39 years old are all younger than 25. Since birth cohort is orthogonal to other variables of interest, this selection will not induce any estimation bias. Further, we focus on a given woman's children with her current husband, because they are the ones at stake during the bequest process ; also, for a given woman, we have only limited information on the rivals of children from previous unions.

Table 2 gives an overview of important characteristics of these women. They have on average 2.7 living children with their current husband ; the number ranges from 1 to 9 (Figure 4 in Appendix A displays the distribution). Roughly 1 woman in 5 has lost at least one child, and 4% have children from previous unions. Of course, since they are in the middle of their reproductive lives, we do not observe their total fertility. The average age in the sample is 28, while the median ages<sup>10</sup> at birth 2 and 3 are respectively 25 and 28 years old. We will therefore focus on children of rank 1, 2 and 3; only a small fraction in our sample has already given birth to more children. Fertility varies significantly across places of residence, education levels and employment status; in our sample, 48% of women live in rural areas, 44% received no education at all, and 33% work. Regarding their marital lives, the vast majority of these women (87%) are in their first union ; the remainder has only one broken union. Women marry for the first time at 19 years old on average ; and they marry a much older husband : the age difference is 10 years on average, and up to 23 years in the top decile of the distribution of age differences. The threat of becoming a young widow is therefore probably a widespread matter of concern. Only 7% of husbands work in the public sector, so the public system of widowhood pensions covers a very small part of the population. The average annual income is 1.6 million FCFA (approx. 2,500 euros), but it is driven by the right tail of the distribution as 50% of the husbands earn less than 500 K FCFA annually (approx. 750 euros). Turning to statistics about rivals, marrying a man who already has children with ex-wives is not unusual : 17% of women have at least one rival, and 12% at least one male rival. The number of rivals in inheritance shares (1 share = 1 boy = 2 girls) is on average 1.7, but the variance is large with numbers ranging from 0.5 to 7.5 shares. 10% have strictly more rivals than own children at the time of the survey.

Women marrying a divorced man are described in Table 3. They have themselves more complex marital lives : they are older and more likely to have broken unions, in which children were born. The age difference with their husband is significantly higher. They are also more likely to live with their current husband, and to work. A first hint that the presence of rivals may have an impact on fertility is that women with rivals have on average more children with their current husband, and in quicker succession ; in line with shorter birth intervals is the increased child mortality rate.

---

10. The median age at birth 2 and 3 is derived from the sample of women above 45 years old, who have reached the end of their reproductive life.

A sub-sample of particular interest will be that of co-residing wives<sup>11</sup>. As explained in the preceding section, rivalry for inheritance seems to be essentially about the husband's house, hence the threat of being thrown out of the house by rivals only jeopardizes women who actually live in their husband's house. It is therefore natural to look at the potentially most exposed women to find evidence of insurance strategies. As shown in Table 4, only 3 married women in 4 live with their husband. Non co-residing wives are very similar to other women regarding socio-economic characteristics (age, area of residence, education, occupation), but they tend to have less children, both dead or alive, with their current husband. Such statistics are consistent with the idea that women who are housed independently from their husband feel less pressure to get insured through their children.

Note that we cannot exhibit any descriptive statistics on widows with rivals and no son, since, consistently with our claim, there are very few of them (11) in the sample. If we believe that the widowhood insurance mechanism is at play, such situation should indeed not exist.

## 4 Empirical Strategy

### 4.1 A duration model of birth intervals

The basic idea underlying our empirical strategy is to compare fertility choices and son preference of wives with and without rivals. We build on the literature on son preference to devise an empirical strategy based on duration models of birth intervals (Leung, 1988, 1991). The existence of son preference is highlighted when the length of the birth interval before the next child is longer for couples with more sons.

The advantage of duration models is their ability to deal properly with right-censored observations, i.e. families that are not yet complete by the time of the survey<sup>12</sup>. Indeed, they allow us to

---

11. One caveat is that co-residence status is observed at the time of the survey, and not when fertility choices were made. Consequently, some women are considered as co-residing, while they were housed independently when their children were born. If they anticipated that they would eventually join their husband's house, it is meaningful to include them in the most exposed women. Conversely, some women are considered as non co-residing, while they used to live with their husband. In our sample, out of 231 non co-residing women, 46 actually remained with their family-in-law after the husband left the household, generally for professional reasons. Those women clearly face the same risk of eviction as co-residing ones. We find that, indeed, they behave similarly : when we include them in the sub-sample of co-residing wives, the magnitude and significance level of our coefficients remain completely unchanged.

12. We have only limited information on the children of women with complete fertility : among women above 45



identify the distribution of a duration variable from potentially right-censored observations as long as the duration and the right-censoring variables are independent. This condition is very likely to be satisfied as the date of the survey is completely unrelated to the latest births. Another rationale for considering birth intervals is to account for health risks related to the spacing, and not only to the number, of births.

Using duration models, son preference has been tested and validated by an extensive literature focusing mainly on Asia : in China (Tu, 1991), in Bangladesh (Rahman and Da Vanzo, 1993), in the Chinese population of Malaysia (Pong, 1994), in Vietnam (Haughton and Haughton, 1995), in India (Arnold et al, 1998), in South Korea (Larsen et al., 1998), and in Taiwan (Tsay and Cyrus Chu, 2005).

Only a few papers examine this question in an African context<sup>13</sup> : Gangadharan and Maitra (2003) find evidence of son preference in South Africa, but only among the Indian community. To our knowledge, the only conclusive paper on African data using duration models was written by anthropologists working on Gabbra pastoralists in the North of Kenya ; they find that women with no son have significantly shorter subsequent birth intervals than women with at least one son (Mace and Sear, 1997). They relate son preference to the patrilineal and patrilocal family system prevailing among the Gabbra. Hence, evidence on son preference is quite limited in Sub-Saharan Africa.

## 4.2 The model with proportional hazard

To test whether son preference is related to the presence of rivals, we use a model with proportional hazard. Our variable of interest  $T$  is the duration between successive births, measured in months. Our coefficients of interest measure the impact of the gender composition of the first children on the subsequent birth interval, among wives with and without rivals. We consider intervals

---

years old who declare having at least one child, not a single child is found in our dataset in almost half of the cases ; all the children are found in only 7% of the cases. On average, a woman declares 5 children and only 1.3 is registered.

13. Some papers use duration models of birth intervals in Africa, but they are interested in the impact of socio-economic factors (e.g. mother's characteristics like birth cohort, age at first marriage and at first birth, residence, education in Ghilagaber and Gyimah, 2004), not in son preference. Other papers deal with son preference in Africa, but do not estimate duration models ; for instance, Milazzo (2012) estimates a linear model for the probability of short birth interval to convincingly underline the role of son preference in fertility choices of Nigerian women.

between births  $n$  and  $(n + 1)$  for  $n = \{1, 2, 3\}$ ; as a robustness check, we include all parities instead of limiting up to parity 3. We start by running a pooled regression of all intervals, and then three separate regressions to determine at which parity hazard rates diverge.

The key element of duration models is the hazard function : the instantaneous probability to have another child at date  $t$ . In the specification with proportional hazard, the hazard function is modeled as follows :

$$\lambda(t) = \lambda_0(t) \times \exp(X'_n\beta)$$

Where  $\lambda_0(t)$  is the baseline hazard function, common to all individuals, and  $X_n$  is a vector of individual covariates at birth  $n$  susceptible to influence the hazard function. Then the survival function - the probability not to have another child at least until  $t$  - is as follows :

$$S(t) = \Pr(T > t) = \exp\left(-\int_0^t \lambda_0(u)du \times \exp(X'_n\beta)\right)$$

In our case,  $X_n$  includes :

- *Girls* : a dummy equal to 1 if the first  $n$  children are all girls.
- *Rivals*<sup>14</sup> : a dummy equal to 1 if there is at least one rival. As robustness checks, we consider alternative definitions of rivals : there is at least one male rival and the number of rivals.
- *Girls*  $\times$  *Rivals* : the interaction term between *Girls* and *Rivals*.
- Additional characteristics of the woman related to fertility : age at birth  $n$ , area of residence, level of education, age at first marriage, whether in her first marriage, living or not with her husband, having at least one dead child<sup>15</sup>, having at least one child from previous unions.
- In the pooled regression, we also include  $n$  : the rank of preceding birth, to control for potential differences between parities.

The fact that we do not observe the birthdate of deceased children might lead us to overestimate the true duration between all successive births, and hence to underestimate son preference. If women

---

14. Since we consider women engaged in monogamous unions, rivals are children of ex-wives; therefore their existence and number are time-invariant.

15. This is the best we can do to mitigate the fact that we do not observe all successive births; results are qualitatively unchanged if the dummy is removed.

do indeed shorten birth spacing until they get a son, we would wrongly assign a large birth interval to those who lost a child as a consequence. Ideally, we would have regressed the intervals between all successive births on the gender composition of the children alive at the time of the subsequent birth. Since we only have information on children alive at the time of the survey, our coefficients can be viewed as a lower bound for son preference.

Also, one might worry about selective mortality : those women who intensify their fertility when they have only daughters are more at risk of maternal death, and therefore more likely to be under-represented in the sample. But this would again lead us to underestimate son preference. Moreover, we find no evidence of selective mortality among women with first born girls : in our data, girls are not more often maternal orphans than boys<sup>16</sup>.

### 4.3 Identification assumptions

Under the proportional hazard assumption, each covariate  $X_k$  has a multiplicative effect that is constant over time on the hazard rate.  $e^{\beta_k}$  might therefore be interpreted as the hazard ratio at any point in time between two individuals that only differ by one unit of  $X_k$ . The proportional hazard assumption will be tested when estimating the model.

In our case,  $\beta_1$ ,  $\beta_2$ , and  $\beta_3$  are the coefficients on *Girls*, *Rivals* and *Girls*  $\times$  *Rivals*, respectively.  $e^{\beta_1}$  measures the hazard ratio between women having only daughters and women having at least one son, among those without rivals.  $\beta_1$  identifies a causal impact if the gender composition of the first children is truly exogenous. This assumption is likely to hold insofar as there is no evidence of sex-selective abortion or infant mortality in Senegal (ANSD and ICF International, 2012). In our sample, sex ratios among children are perfectly balanced (cf. Figure 1). Moreover, we test that women having  $n$  daughters are indeed similar to women having at least one son for  $n = \{1, 2, 3\}$  : balancing tests presented in Appendix A, Table 13, provide some level of reinsurance that the *Girls* dummy is exogenous<sup>17</sup>. If we find that  $e^{\beta_1} > 1$ , in other words, having only daughters vs. at least

---

16. Among children younger than 15 years old (6,150 observations), 2.4% of boys and 2.5% of girls have lost their mothers. The difference between boys and girls is not significant (p-value=0.82).

17. We conduct univariate t-tests to compare covariate means across both subsamples, and we find no statistically significant difference at 5% when  $n = \{1, 2\}$  and only one (in 17 covariates) when  $n = 3$ , which is usually considered as a satisfactory degree of balance.

one son increases the hazard rate and hence decreases the expected birth interval, then we can infer the existence of son preference among women without rivals.

$e^{\beta_1+\beta_3}$  measures the same ratio among women with rivals. Therefore,  $e^{\beta_3}$  is a "ratio of ratio" that captures the difference in son preference between women with and without rivals. If our hypothesis is true, then we expect  $e^{\beta_3} > 1$ ; in other words, having only daughters vs. at least one son increases *more* the hazard rate and decreases *more* the expected birth interval when there are rivals.  $\beta_3$  would identify an insurance strategy through sons as long as both groups of women have the same "intrinsic" son preference, meaning outside of inheritance considerations.

Of course, wives with and without rivals are different; even after controlling for observable differences, some unobservable characteristics may drive both fertility choices and the decision to marry a divorced man or a widower<sup>18</sup>. However, what matters for our identification to hold, is that in the absence of rivals, both groups would have reacted similarly to an exogenous shock : having one son vs. one daughter. Under this assumption, if wives with rivals eventually display a stronger son preference than wives without rivals, the difference would be caused by the presence of rivals. In section 5, we run some placebo tests to ensure that (i) the common "intrinsic" son preference is plausible, and that (ii) an insurance-based interpretation explains the results better than potential alternative mechanisms.

One limit of our empirical strategy is that we are dealing with rather small samples. Table 12 in appendix A gives details about the number of observations for each regression, as well as the exact number of observations in each cell of interest. In particular we should keep in mind that there are very few women having 3 daughters and rivals when interpreting the results.

#### 4.4 Descriptive statistics on birth intervals

Table 5 presents some descriptive statistics of our duration variables. When we pool all durations together, durations after birth 1 account for almost half of the observations, durations after birth 2 for one third, and the remaining durations are observed after birth 3. A bit more than one third of the durations are right-censored, a proportion that logically increases in parity. Among non-

---

18. It is worth underlining that, given the general complex patterns of marital lives in Senegal, marrying a divorced man or a widower is likely to happen to most women at some point in their lives.

censored observations, the distribution is quite similar for the 3 parities, with a mean around 36 months, and a median around 30 months. One fourth of the births would be considered as risky with a birth spacing lower than 24 months. Note that duration between births 1 and 2 displays a larger variance due to "extreme" values<sup>19</sup>.

Before estimating the proportional hazard model, the first step to get a picture of the survival function is to compute the Kaplan-Meier estimator of  $S(t)$ (Kaplan and Meier, 1958). It is plotted in Figure 2 when durations are pooled together, and when they are separated by parity. All graphs display a common shape with three important features :

1.  $\widehat{S}_{KM}(t) = 1$  if  $t < 8$  : the shortest birth interval is equal to 8 months.
2. The exit rate is increasing until approximately 50 months and then decreasing.
3.  $\widehat{S}_{KM}(t) > 0$  even if  $t$  is large i.e. the distribution of  $T$  is defective ; in other words, some women will never have an additional child.

Also, note that we observe "jumps" in the survival function at 12, 24, 36 etc. months. They are due to measurement approximations : in roughly one third of the non-censored cases, we do not observe the exact dates of birth for the children  $n$  and  $(n + 1)$ , hence we approximate the birth interval by the age difference multiplied by 12, to get the duration in months. The same approximation is implemented on censored observations in 20% of the cases, when the exact birthdate of the latest child is missing.

As a consequence, when we observe  $t_n$ , the information we derive is more or less precise, depending on the type of observation<sup>20</sup> :

Not censored, precisely measured ; we infer that  $T_n \in [t_n - 1; t_n + 1]$ .

Not censored, imprecisely measured ; we infer that  $T_n \in [t_n - 12; t_n + 12]$ .

Censored, precisely measured ; we infer that  $T_n > t_n - 1$ .

Censored, imprecisely measured ; we infer that  $T_n > t_n$ .

---

19. Approximately 2% of birth intervals are larger than 8 years (mostly observed after birth 1) : in half of the cases, women have in fact lost 1 or 2 children in the meantime ; the remaining cases are mostly urban, educated, working women living in relatively rich households.

20. See Appendix B for further detail about the approximation and the number of observations of each type.

## 4.5 Estimation

### 4.5.1 Semi-parametric estimation

We start by considering a semi-parametric estimation of the Cox proportional hazard model (Cox, 1972). Using a partial likelihood, the method of estimation makes it possible to estimate the vector of coefficients  $\beta$  without imposing a functional form on  $\lambda_0(t)$ . The idea is to focus on the ordering of events rather than on the exact duration for each observation<sup>21</sup>. In the pooled specification, we use robust standard errors clustered at the woman level to account for the correlation between the error terms related to the different birth intervals for the same woman.

One can derive an estimate of the baseline survival function,  $\widehat{S}_0(t_{(i)}^*)$ , constant by pieces on the interval  $[t_{(i)}^*; t_{(i+1)}^*[,$  where  $t_{(1)}^* < \dots < t_{(k)}^*$  are the  $k$  ranked distinct values among non-censored observations (Box-Steffensmeier and Jones, 2004). We use it to compute the expected duration conditional on  $X$  for the women who eventually have another child<sup>22</sup> :

$$\mathbb{E}(T \mid X, T \leq t_{(k)}^*) = \left( \sum_{i=1}^{k-1} \widehat{S}_0(t_{(i)}^*)^{\exp(X'\hat{\beta})} \times (t_{(i+1)}^* - t_{(i)}^*) \right) - t_{(k)}^* \widehat{S}_0(t_{(k)}^*)^{\exp(X'\hat{\beta})}$$

Another quantity of interest is  $(1 - \widehat{S}(24 \mid X))$  : the probability of a short birth interval for women with characteristics  $X$ . We define "short" as inferior or equal to 24 months because it is the minimum length advocated by family planning programs and used in health statistics (e.g. DHS) below which a birth is considered risky.

In particular, we are able to compare these quantities for women having only daughters and/or rivals, holding everything else constant. We consider the median individual as the reference individual, by setting the values of the controls to the median of the sample<sup>23</sup>.

Nonetheless, the semi-parametric estimation has three main limits : first, it is not well-designed

---

21. In the three separate regressions, we use the exact marginal-likelihood method to handle ties among non-censored durations. It provides the most accurate and appropriate approximation when duration data is generated by a continuous time process and ties arise from imprecise measurement, as is the case here, because it takes into account all the possible orderings of tied failures. We check that estimates based on alternative approximations (Efron or Breslow method) are of very similar size and significance. In the pooled regression, we use Breslow as the exact marginal-likelihood method is not available when standard errors are clustered.

22. In the Cox model, some women will never have another child, so the global expected duration would be equal to infinity. That is why our quantity of interest is in fact the expected duration given that  $T \leq t_{(k)}^*$ .

23. We check that results are qualitatively unchanged if we consider the average individual (not shown).

to model individual-level unobserved heterogeneity, which would be useful here to capture the unobserved determinants of fertility across women (e.g. fecundity level). Second, the specific measurement approximations described above cannot be accounted for with the standard Cox estimation. Third, the estimation does not explicitly take into consideration the defective distribution of our durations.

#### 4.5.2 Fully parametric estimation

To account for (i) unobserved heterogeneity, (ii) measurement approximations, and (iii) a defective distribution, we estimate a fully parametric multispell model. The details of the specification are given in appendix C.

We start by defining a dummy  $R_n$  for long-term survivors at parity  $n$  *i.e.* women who stop having children after the  $n^{th}$  birth. The probability of stopping is likely to depend on the characteristics of women after birth  $n$  summarized in the vector  $X_n$ ; we estimate the probability with a logit :

$$Pr(R_n = 1 | X_n) = p(X_n) = \frac{\exp(X_n' \alpha)}{1 + \exp(X_n' \alpha)}$$

For the women who have another child, the hazard function is now given by :

$$\lambda(t | X_n, \nu) = \lambda_0(t) \times \exp(X_n' \beta + \nu)$$

Where :

- $\nu \sim \mathcal{N}(0, \sigma)$  :  $\nu$  has a normal distribution such that  $\mathbb{E}(\nu) = 0$  and  $Var(\nu) = \sigma^2$ .
- The baseline hazard function has a Weibull distribution with parameters  $\lambda$  and  $a$ .

$$\lambda_0(t) = \begin{cases} \lambda a (\lambda(t - 8))^{a-1} & \text{if } t \geq 8 \\ 0 & \text{if } t < 8 \end{cases}$$

This hazard function replicates two characteristics of the durations observed in our sample : (i) there is no exit before 8 months, and (ii) the inverted U-shape of the exit rate arises from the combination

of an increasing baseline hazard together with unobserved heterogeneity<sup>24</sup>. The survival function conditional on  $R_n$ ,  $X_n$  and  $\nu$  given in appendix C allows us to retrieve the aggregate survival function conditional on  $X_n$  and  $\nu$ .

Using this specification, we are able to disentangle the effect of covariates, in particular *Girls*, *Rivals* and *Girls*  $\times$  *Rivals*, on the number of births and on the spacing of births. The vector  $\alpha$  captures the impact on the probability of stopping having children, whereas the vector  $\beta$  captures the impact on durations. The contribution to the likelihood of each observation is specified in a way that takes into account measurement approximations; it uses a confidence interval that depends on how precisely the duration is measured, instead of the density and survival function.

Now,  $\theta \equiv \{\alpha, \beta, \lambda, a, \sigma\}$  is the vector of parameters to be estimated. We start by considering the same parameters for all parities and compare the estimates to those found in the pooled Cox regression. Then we allow the parameters<sup>25</sup> to vary depending on  $n$ , and compare the estimates to those found in the three separate Cox regressions. However, the last specification includes too many parameters relative to the number of observations; therefore we constrain the model and keep only the controls whose impact was significant in the estimation on all parities.

Last, we are able to compare across different categories of women, the probability of having another child,  $(1 - p(X))$ , and the expected birth interval conditional on  $X$  :

$$\mathbb{E}(T | X, \nu, \hat{\theta}) = \int_0^\infty S(u | X, \nu, \hat{\theta}) du$$

We also consider the probability of a birth interval being shorter than 24 months,  $(1 - S(24 | X, \nu, \hat{\theta}))$ . Again, we compute these quantities for the median individual (setting  $\nu = 0$ ) and compare them to the ones obtained with the Cox estimation.

---

24. Women with a high hazard rate exit at the beginning, and after some time, there remains only women with a low hazard rate. We do not investigate further whether the inverted U-shape results from the exit process *per se*, or from the variation in the composition of the population, because we are not mainly interested in duration dependence, here.

25. Except  $\sigma$ , the standard deviation of the unobserved heterogeneity. This is precisely the strength of the multispell model : random draws from the normal distribution are attributed once and for all to each individual, explaining simultaneously all durations.



## 5 Results

### 5.1 Preliminary evidence : non-parametric estimation

Using the Kaplan-Meier estimate of the survival function and its variance, it is possible to test for the equality of distributions between two sub-populations (log-rank tests)<sup>26</sup>. The log-rank tests are conclusive in two cases : first, when  $n = 3$ , we reject the null (at 11%) that women with 3 daughters and rivals have the same survival function than the rest of the population. Second, in the pooled specification, when we restrict the sample to co-residing wives, we also reject the null (at 3%) that women with only daughters and rivals have the same survival function as the others. As shown by Figure 3, the survival function of women having only daughters and rivals is systematically below the one of the others, which means shorter birth intervals.

The non-parametric estimation provides a first hint that women with the highest need for widowhood insurance also make different fertility choices. However, the non-parametric estimation does not allow us to control for other factors, and therefore does not permit us to infer anything else than a correlation at this point.

### 5.2 Semi-parametric estimation : Results

We first discuss the estimates of the pooled specification (Table 6), and then turn to the results disaggregated by parity (Table 7). When we consider the whole sample, there is no ratio significantly different from 1 ; we notice that in magnitude,  $e^{\beta_1}$  and  $e^{\beta_2}$  are indeed very close to 1, whereas  $e^{\beta_3}$  is higher, but imprecisely estimated.

Following the hint from the non-parametric estimation, we consider the sample of co-residing wives. For them, we have information about the husband's characteristics (income, sector of activity, age difference) that we include in the regression as additional controls ; we also add a dummy indicating whether the woman works.  $e^{\beta_1}$  remains insignificant, so we find no evidence of son preference among women without rivals.  $e^{\beta_3}$  is higher in this sub-sample, and now significantly different from 1 (at 5%). Among women with at least 1 rival, the hazard rate is multiplied by

---

26. We compare women with only daughters vs. those with at least a son ; women with rivals vs. those without any ; and finally women with both only daughters and rivals vs. others.

$e^{\beta_1+\beta_3} = 1.6$  for those having only daughters as compared to those having at least one son.

We further investigate whether the presence of rivals has a heterogenous impact depending on the socio-economic status of the household; to do so, we split the sample of co-residing wives on the median income<sup>27</sup>. In the poorest half of the population,  $e^{\beta_3}$  is very high and significant (at 1%) : having only daughters multiplies the hazard rate by 2.4 for women with rivals, while it has no significant impact for women without rivals. In the richest half of the population,  $e^{\beta_3}$  is not significantly different from 1<sup>28</sup>. At first sight, such a result might seem at odds with standard economic predictions stating that the larger the bequest, the stronger the incentives to take hold of it. However, what matters in our context is probably not the absolute value of the bequest as the relative value of the house in the bequest; and the house is the main asset in most of poor households. Moreover, women in those households have hardly any alternative in case of eviction because they cannot rely on personal resources or on their extended family to support them.

Another group of women particularly exposed to the risk of widowhood would be those married to an old husband. However, we would need more women married to an old enough husband to come up with satisfactory results<sup>29</sup>.

Also, note that  $e^{\beta_2}$  is never different from 1. Hence, after controlling for observable differences between women with and without rivals, both groups make similar fertility choices once they have a son.

Now we consider the three regressions corresponding to each parity; the only ratio significantly different from 1 is  $e^{\beta_3}$  when  $n = 3$ . Among women with at least 1 rival, the hazard rate is multiplied by 2.5 for those having 3 daughters as compared to those having at least one son.

The results disaggregated by parity suggest that the presence of rivals impacts fertility choices

---

27. Results are qualitatively unchanged if we take education instead of income as an indicator of socio-economic status :  $e^{\beta_3}$  is much higher for non-educated women than for educated ones.

28. If we interact a *Poor* dummy with our variables of interest instead of splitting the sample, the difference between rich and poor is confirmed : the coefficient on the triple interaction  $Girls \times Rivals \times Poor$  is equal to 2 and significant at 10%.

29. The distributions of husband's age at birth  $n$  are :  $Q_1 = 27, Q_2 = 31, Q_3 = 36$ , among women without rivals; and  $Q_1 = 30, Q_2 = 36, Q_3 = 42$ , among women with rivals. The impact of husband's age on women's choices is likely to be non-linear, but the relevant thresholds are difficult to figure out. As an example, we define "old" as above 42 to ensure there are enough women with rivals and an old husband. When we split the sample of poor co-residing wives between old and young husbands, we do find that  $e^{\beta_3}$  is larger for old husbands than for young ones, but both coefficients are not significantly different from each other.

particularly after the third child, and for the women who have not had a son yet. The average ideal family size among married Senegalese women is 5 children (ANSD and ICF International, 2012). Hence, the vast majority of women probably expect to have other "tries" for a son, when at first they have one or two daughters. The pressure to have a son would start increasing after the third "missed try", once some women have reached their ideal family size. We predict that the effect should keep growing at parities 4 and 5, but unfortunately, our sample size does not allow us to test this prediction<sup>30</sup>.

To better understand the consequences in terms of health, we estimate (i) the expected birth spacing and (ii) the probability of short birth spacing for the three populations in which we detected an effect : co-residing wives, the poorest half of co-residing wives, and women with at least 3 children (cf. Table 8).

Among co-residing wives, we find that, holding everything else constant, the expected birth spacing is quite similar for all women when there is no rival. But in presence of rivals, having only daughters vs. at least one son is predicted to reduce birth spacing by 10 months and to multiply the probability of short birth spacing by 1.5. These same figures are 11 and 2, respectively, when we consider the poorest half of the population<sup>31</sup>.

If we turn to the sample of women with at least 3 children, having 3 daughters vs. at least 1 son reduces the expected birth spacing by 13 months and doubles the probability of short birth spacing among women with rivals. In contrast, it has almost no impact among women without rivals, and if any, it is the other way around.

The probability of very short birth spacing (below 15 months)<sup>32</sup> is multiplied by 1.5 for co-residing wives and by 2.3 for women in relatively poor households, and for those having three daughters.

---

30. Jayachandran and Kuziemko (2011) have also put forward the hypothesis that gender is most predictive of subsequent fertility near ideal family size in India. They build a formal model of fertility choices under son preference, in which breastfeeding acts as a contraceptive. They predict that the probability of breastfeeding should be lowest for daughters without older brothers ; and the gender gap is expected to arise once ideal family size is reached.

31. Excluding extreme values of birth intervals (above 96 months) does not change the predicted probability of short birth spacing, but mechanically reduces the expected birth spacing and flattens out differences between categories : the gap between women having rivals and only daughters vs. at least one son amounts to 6 and 8 months, for co-residing wives and for the poorest half of co-residing wives, respectively.

32. Same computation as for the probability of birth spacing below 24 months (not shown).

Regarding the controls, hazard ratios are in line with expectations. The hazard rate decreases in parity : durations are longer between the latest children. It is higher for women in their first union and who married younger. Also, it is higher for co-residing wives, in particular when  $n = 3$  : the hazard rate doubles. The area of residence has a significant impact, especially when  $n = \{2, 3\}$  : the hazard rate is multiplied by approximately 1.5 for rural vs. urban mothers. Two covariates play a significant role only when  $n = 1$  : the hazard rate is halved for mothers who already have children from previous unions. And the hazard rate is 25% lower for women having one dead child, suggesting that our dummy variable mainly captures the death of children after the first child and before the second (living) child.

When we restrict the sample to co-residing wives, we find that women married to civil servants, and therefore entitled to a public system of widowhood pensions, have lower hazard rates (-30%) ; the effect is particularly strong for the richest half of the population. Age difference and employment status of the woman are never significant.

Last, we test the proportional hazard assumption following the procedure developed by Grambsch and Therneau (1994) and based on the Schoenfeld partial residuals (Schoenfeld, 1982). We fail to reject it in the pooled specifications, and when  $n = \{2, 3\}$ , which supports the choice of our model. But we reject the proportional hazard assumption when  $n = 1$ , both in the global test and in the tests for education and dead children variables. To further investigate whether the violation of this assumption introduces a strong bias in our coefficients of interest, we split the sample according to the median duration between births 1 and 2 (30 months) and run separate regressions on both subsamples. We find that, indeed the effect of the education and dead children variables are not constant in time<sup>33</sup>, but  $\beta_1$ ,  $\beta_2$  and  $\beta_3$  are not different for shorter and longer durations.

---

33. The lack of education has a positive impact on hazard rates (hence a negative impact on durations) only for shorter durations. And having one dead child has a negative impact on hazard rates (hence a positive impact on durations) only for larger durations : this seems consistent as the death of a child has a mechanical positive impact on the duration between births 1 and 2 only if it occurred between them ; it cannot be the case if the duration is short.

### 5.3 Fully parametric estimation : Results

The estimates of the fully parametric specification are reported in Tables 9 and 10 ; they confirm, strengthen and specify the results of the Cox model.

The first clarification is that, on the one hand, our variables of interest *Girls*, *Rivals* and *Girls*  $\times$  *Rivals* have no significant impact on the probability to stop having children, whatever sample we consider<sup>34</sup>. But on the other hand, the effect of *Girls*  $\times$  *Rivals* on durations is much higher and significant than estimated in the Cox model ;  $e^{\beta_3}$  turns out to be significant even for the whole sample. The ranking across sub-samples remains the same :  $e^{\beta_3}$  is still larger for co-residing wives, in particular for the poorest half, and for women with at least 3 children. Otherwise,  $e^{\beta_1}$  and  $e^{\beta_2}$  are similar in magnitude to the Cox estimates, and in the same way, never significantly different from 1.

Hence, estimations based on the explicit modeling of a defective distribution suggest that women's needs for a son influence the spacing, but not the number of births - at least up to the third child. Our results are in line with the literature on son preference : Jensen (2005) and Basu and De Jong (2010) point out that son preference does not necessarily lead to differences in sibship size when fertility is high ; in this case, they advocate looking at birth intervals to find evidence of son-preferring behavior.

Our quantities of interest capturing the magnitude of the effect display larger differences across groups than those derived from the Cox estimation : in the presence of rivals, having only daughters vs. at least one son multiplies the probability of short birth spacing by 2 to 4 times ; the reduction in the expected birth spacing is estimated between 11 and 21 months, depending on the population we consider. But just as in the Cox model, birth spacing is predicted to be quite similar for all women when there is no rival (cf. Table 11).

Further clarification is also obtained on the controls. Living in a rural area, being married for the first time, and co-residing with the husband, all make it significantly less likely to stop having children ; coefficients are particularly high when  $n = 3$ , as already noted in the Cox estimation.

Other controls influence birth spacing : the age at first marriage, the rank of the child and

---

34. The probability of having another child is estimated to be around 98% for the median individual, irrespective of the gender composition of the first children and/or the presence of rivals.

having children from previous unions increase durations ; whereas the age at the previous birth and a low level of education tend to shorten durations. Again, we find that having one dead child has a significant, positive impact on durations only between births 1 and 2. Co-residence is the sole variable influencing both the number of children and their spacing.

When we introduce controls related to the husband, marrying a civil servant has a mixed impact on fertility : it reduces the probability to stop, but it also lengthens durations. Age difference and women’s employment remain insignificant.

Regarding the baseline survival function, coefficients  $\lambda$  and  $a$  are always precisely estimated. The graphs corresponding to different populations of interest and to different parities, are plotted in Appendix D, Figure 5. When we consider all parities together, the baseline survival function of co-residing wives is always below the one of the whole sample, reflecting their increased fertility.

Last, in all specifications, estimates of the variance of the unobserved heterogeneity are significant at 1% and of the same order of magnitude<sup>35</sup>.

#### 5.4 Robustness and placebo tests

We further perform robustness and placebo tests reported in Appendix E. First, we check that hazard ratios in our main sample of interest - co-residing wives - are robust to alternative specifications : (i) including all parities instead of limiting up to parity 3 ; and (ii) adding controls for ethnic groups (Table 15). Coefficients are also unchanged when we control for previous durations in the regressions by parity (not shown).

Second, one can wonder whether the results of the pooled regression are actually only driven by durations after birth 3. It would be somewhat worrying due to the small number of women having three daughters and rivals. To remove any doubt, we focus on durations after birth 1 and consider the sample of co-residing wives, as a whole and next divided into poorest half and richest half<sup>36</sup>. As shown in Table 16, results of the pooled regression still hold when we consider only durations after birth 1 :  $e^{\beta_3}$  remains of the same order of magnitude and significance level. So the mechanism seems

---

35. We find  $\sigma = 0.68$  for the whole sample, 0.75 for co-residing wives, and 0.63 when coefficients are parity-specific.

36. We did the same exercise on durations after birth 2 : estimates are line with the pooled results, even though not significant, but we should be careful with the interpretation due to the small sample size when  $n = 2$ .

to be already at play after birth 1 but to a lesser extent, so that the impact of insurance strategies on fertility would be only detectable for the most exposed women. After birth 3, the pressure would be much stronger and the effect visible in the whole population.

Third, we check that our estimates do not capture preferences for diversity rather than son preference, by replacing the *Girls* dummy by a *Boys* one : we find that having only sons vs. at least one daughter has no significant impact on the hazard rate, either for women with or without rivals ; if anything, the ratio is always lower than one, which refutes the hypothesis of preferences for diversity (Table 17).

Fourth, we test the identification assumption of a common "intrinsic" son preference between women with and without rivals. The reasoning is the following : if one son is already enough to feel insured, women with rivals should display the same son preference as women without rivals from the moment they give birth to their first son. Restricting the sample to women of parity 2 with at least 1 son, we find indeed that the impact of having 2 sons vs. 1 son and 1 daughter is the same in the two groups ; women with 2 sons wait significantly longer before getting pregnant again, irrespective of the presence of rivals (Table 18).

Fifth, we investigate an alternative explanation for higher son preference in the presence of rivals, which emphasizes the role of the husband. One story is that husbands would be more likely to divorce, or repudiate, a first wife if she had not given birth to a son ; the next wife would then feel the pressure to have a son quickly to avoid divorce or repudiation. Such mechanism is very improbable in our case since there are slightly more boys than girls among rivals, and also slightly more ex-wives with at least one son than ex-wives with at least one daughter.

Another story assumes that men with a strong son preference would be more likely to take another wife in order to maximize the number of sons they have, and the arrival of a new wife could prompt the first one into divorcing ;  $e^{\beta_3} > 1$  would simply capture the son preference of the husband, not of the wife. If this were true, we should observe evidence of son preference among the rivals, since the husband would have put pressure on his first wife as well. We have no systematic, detailed information on the timing of births of the rivals so we cannot use a duration model of birth intervals. Instead, we test a standard prediction of son-preferring, differential stopping fertility

models : under son preference, an average girl is predicted to have more siblings than an average boy (Jensen, 2005, Basu and De Jong, 2010). It is not the case for the rivals in our sample : boys and girls have on average 2.5 siblings.

Last, we consider alternative definitions of rivals to capture potential effects of the gender composition and the number of rivals<sup>37</sup>. So far, our results indicate that the presence of one son plays a crucial part in the inheritance process. If we turn the reasoning around, rivals should represent a threat (i) only when there are male rivals ; and (ii) as soon as there is one male rival.

We begin with defining rivals as a dummy equal to 1 if there is at least one male rival, thus excluding women having only female rivals from the category. As shown in Tables 19 and 20, the same pattern of results observed in the specification "at least 1 rival" holds. Also if we compare the effect of having at least one male rival vs. one female rival in our sample of interest - co-residing wives -  $e^{\beta_3}$  is smaller and no longer significant in the female case. When we introduce both, the coefficient on  $Girls \times Rivals\_male$  is 35% higher than the coefficient on  $Girls \times Rivals\_female$ , even though both are not significant (Table 21).

Turning to the impact of the number of rivals, we create a variable equal to the number of rivals in inheritance parts (Tables 22 and 23). Again,  $e^{\beta_3}$  is particularly high for the poor and after the third child. Among co-residing wives, each additional inheritance part for the rivals increases the hazard ratio between women having only daughters and women having at least one son by 23%. However, the impact of the number of rivals is non-linear : most of the effect takes place between 0 and 1 rival ; each additional rival just slightly raises the ratio. Among women with exactly one rival, having only daughters doubles the hazard rate. Logically, the effect is entirely driven by women having one male rival ; it is zero for women having one female rival (Table 24).

One minor difference in the estimates using alternative definitions is that  $e^{\beta_2}$  is sometimes significantly lower than 1 for the poorest half of co-residing co-wives and for women with at least 3 children. But the ratio is imprecisely estimated, and in any case, of relatively small magnitude. It only suggests that women with more/male rivals tend to have larger birth intervals than other women as long as they have at least one son : one son seems enough to feel insured.

---

37. Note incidentally that the coefficients on the controls are robust to alternative definitions of rivals.



## 6 Conclusion

All in all, we find that wives with rivals for inheritance need to have one son, while there is no sign of son preference for wives without rivals. These findings support the idea that sons play a key role in insuring women in case of widowhood. However, the impact of insurance strategies on fertility only materializes for the most exposed women : those with no independent housing, with rivals and without a son among the first children. The cost of insurance through sons in terms of health is therefore borne by a small fraction of women : in cross-section, they represent only 7% of our sample<sup>38</sup>. But, from a longitudinal perspective, this fraction should be much higher given the high rates of divorce and remarriage in Senegal.

Still, for these women and their children, the health risk is huge. We showed that the probability of short birth spacing was multiplied by 1.5 to 2, and up to 4 for some sub-populations; and we learnt from the medical literature that short birth spacing multiplies the risk of infant death by 2.4 (ANSD and ICF International, 2012) and doubles the risk of maternal death (Conde-Agudelo and Belizan, 2000).

As a consequence, reducing women's reliance on children to eventually disassociate fertility choices from insurance considerations could help achieve the national health plan's objectives. The general consensus is that improving women's status is a precondition for the fertility transition. In a chapter entitled "Women's agency and social change", Sen argues that women's empowerment will bring about lower fertility rates and lower infant mortality (Sen, 1999). However, Sen maintains that the main drivers of fertility transition are (i) a change in women's aspirations and bargaining power, and (ii) the access to birth control methods. We argue that both conditions are necessary, but not sufficient : as long as women have no alternative insurance strategies, they will continue to have many children to mitigate downside risks, starting with widowhood.

One might think of three, non-mutually exclusive, ways ahead. First, reforming inheritance rules to grant wives, irrespective of the number of children, a significant percentage of the husband's wealth, or at least the usufruct of the house for some years. Second, expanding social benefits to protect a larger share of the population : in the Senegalese case, it would be broadening the

---

38. 75% of women currently live with their husband and among them, 18% have rivals and 50% have first born girl.

current system of pensions paid to widows of civil servants to all widows. Third, but in the longer run given the difficulties encountered by insurance companies in developing countries, promoting formal insurance and financial markets.

The next question on our research agenda is to investigate whether the same mechanism is at play in polygamous households. The situation is fundamentally different for two reasons : on the one hand, first wives might react to the arrival of a new co-wife by adjusting their number of children ; they might also anticipate the births of potential rivals, hence insuring themselves for the future. On the other hand, co-residing co-wives in polygamous unions have an alternative insurance strategy based on cooperation : they may agree to all stay in the house in case of husband's death. A theoretical approach is therefore necessary to describe fertility choices and family insurance strategies in a polygamous setting.

## References

- Agence Nationale de la Statistique et de la Démographie (ANSD) et ICF International (2012), 'Demographic and Health Survey - Multiple Indicator Cluster Survey (DHS-MICS) 2010-2011', Calverton, Maryland, USA : ANSD et ICF International.
- Antoine, P. and Dial F. B. (2003), 'Mariage, divorce et remariage à Dakar et Lomé.', Actes du colloque de l'AUF : Familles du Nord, familles du Sud, Marseille 23-26 juin 2003.
- Appleton, S. (1996), 'Women-Headed Households and Household Welfare : an Empirical Deconstruction for Uganda.', *World Development* 24(12), 1811-1827.
- Arnold, F., Choe, M. K. and Roy, T.K. (1998), 'Son Preference, the Family-Building Process and Child Mortality in India.', *Population Studies* 52(3), 301-315.
- Basu, D. and De Jong, R. (2010), 'Son targeting fertility behavior : some consequences and determinants.', *Demography* 47(2), 521-536.
- Bledsoe, C. (1990), 'Transformations in Sub-Saharan African Marriage and Fertility.', *Annals, AAPSS* 510, 115-125.
- Box-Steffensmeier, J. and Jones, B (2004), 'Timing and Political Change : Event History Modeling in Political Science.', Ann Arbor : University of Michigan Press.
- Conde-Agudelo, A. and Belizan, J.M. (2000), 'Maternal Morbidity and Mortality Associated with Interpregnancy Interval : Cross Sectional Study.', *British Medical Journal* 321, 1255-1259.

- Conde-Agudelo, A., Rosas-Bermudez, A. and Kafury-Goeta, A.C. (2006), 'Birth Spacing and Risk of Adverse Perinatal Outcomes. A Meta-analysis.', *The Journal of the American Medical Association*, 295(15), 1809-1823.
- Cox, D.R. (1972), 'Regression Models and Life-Tables (with discussion).', *Journal of the Royal Statistical Society Series B* 34, 187-220.
- DaVanzo, J., Razzaque, A., Rahman, M. and Hale L. (2004), 'The Effects of Birth Spacing on Infant and Child Mortality, Pregnancy Outcomes, and Maternal Morbidity and Mortality in Matlab, Bangladesh.', Rand Working Paper, WR-198, Santa Monica, California : Rand.
- DeVreyer, P. and Lambert, S. (2013), 'Household Risk Management in Senegal', background paper for the 2014 World Development Report, *Managing Risk for Development*, World Bank.
- DeVreyer, P., Lambert, S., Safir, A. and Sylla, M. B. (2008), 'Pauvreté et structure familiale, pourquoi une nouvelle enquête?', *Stateco* (102), 261-275.
- Gangadharan, L. and Maitra, P. (2003), 'Testing for Son Preference in South Africa.', *Journal of African Economies* 12(3), 371-416.
- Ghilagaber, G. and Gyimah, S. O. (2004), 'A Family of Flexible Parametric Duration Functions and their Applications to Modelling Child-Spacing in Sub-Saharan Africa.', *PSC Discussion Papers Series* 18(1).
- Grambsch, P.M. and Therneau, T.M. (1994), 'Proportional hazards tests and diagnostics based on weighted residuals.', *Biometrika* 81(3), 515-526.
- Haughton, J. and Haughton, D. (1995), 'Son Preference in Vietnam.', *Studies in Family Planning* 26(6), 325-337.
- Hoddinott, J. (1992), 'Rotten Kids or Manipulative Parents : Are Children Old Age Security in Western Kenya?', *Economic Development and Cultural Change* 40(3), University of Chicago Press, 545-65.
- Horrell, S. and Krishnan, P. (2007), 'Poverty and Productivity in Female-Headed Households in Zimbabwe.', *Journal of Development Studies* 43(8), 1351-1380.
- Jakiela, P. and Ozier, O. (2012), 'Does Africa Need a Rotten Kin Theorem? Experimental Evidence from Village Economies.', World Bank Policy Research Working Paper 6085.
- Jayachandran, S. and Kuziemko, I. (2011), 'Why Do Mothers Breastfeed Girls Less Than Boys? Evidence and Implications for Child Health in India.', *The Quarterly Journal of Economics* 126, 1485-1538.
- Jensen, R. (2005), 'Equal treatment, unequal outcomes? Generating sex inequality through fertility behavior.', Unpublished document. JFK School of Government, Harvard University.
- Kaplan, E.L. and Meier, P. (1958), 'Nonparametric estimation from incomplete observations.', *Journal of the American Statistical Association* 53, 457-481.

- Kazianga, H. and Klønner, S. (2009), 'The Intra-household Economics of Polygyny : Fertility and Child Mortality in Rural Mali.', MPRA Working Paper 12859.
- Lambert, S. and van de Walle, D. (2012), 'Transcription des entretiens menés avec des veuves et des divorcées au Sénégal en Novembre 2011.', unpublished document, World Bank.
- Lambert, S., Ravallion M. and van de Walle, D. (2013), 'Intergenerational Mobility and Interpersonal Inequality in an African Economy', Cepremap.
- Larsen, U., Chung, W. and Das Gupta, M. (1998), 'Fertility and son preference in Korea.', *Population Studies* 52, 317-325.
- Lesthaeghe, R. J., editor (1989), 'Reproduction and Social Organization in Sub-Saharan Africa.', Berkeley : University of California Press.
- Leung, S.F. (1988), 'On tests for sex preferences.', *Journal of Population Economics* 1, 95-114.
- Leung, S. F. (1991), 'A Stochastic Dynamic Analysis of Parental Sex Preferences and Fertility.', *Quarterly Journal of Economics* 106(4), 1063-1088.
- Mace, R. and Sear, R. (1997), 'Birth interval and the sex of children in a traditional African population : an evolutionary analysis.', *Journal of biosocial science* 29(4), 499-507.
- Milazzo, A. (2012), 'Son Preference, Fertility and Family Structure : Evidence from Reproductive Behavior among Nigerian Women', Job market Paper, Bocconi University, Milan.
- Milazzo, A. (2013), 'Why Are Adult Women Missing? Son Preference and Maternal Survival in India', mimeo, Bocconi University, Milan.
- Nugent, J. (1985), 'The Old-Age Security Motive for Fertility', *Population and Development Review* 11(1), 75-97.
- Pong, S.L. (1994), 'Sex preference and fertility in Peninsular Malaysia.', *Studies in Family Planning* 25(3), 137-48.
- Rahman, M. and DaVanzo, J. (1993), 'Gender Preference and Birth Spacing in Matlab, Bangladesh.', *Demography* 30(3), 315-332.
- Rosenzweig, M. R. (1988), 'Risk, Private Information, and the Family', *The American Economic Review*, 78(2), Papers and Proceedings of the One-Hundredth Annual Meeting of the American Economic Association (May, 1988), 245-250.
- Ronsmans, C. (1996), 'Birth spacing and child survival in rural Senegal.', *International Journal of Epidemiology* 25(5), 989-97.
- Schoenfeld, D. (1980), 'Chi-squared goodness-of-fit tests for the proportional hazards regression model.', *Biometrika* 67(1), 145-153.

Sen, A. (1999), *Development as Freedom.*, Oxford : Oxford University Press.

Strassmann, B. (1997), 'Polygyny as a Risk Factor to Child Mortality among the Dogon.', *Current Anthropology* (38), 688-695.

Tsay, W.J. and Cyrus Chu, C. Y. (2005), 'The pattern of birth spacing during Taiwan's demographic transition.', *Journal of Population Economics* 18, 323-336.

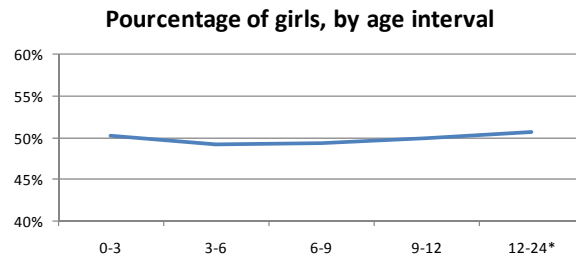
Tu, P. (1991), 'Birth spacing patterns and correlates in Shaanxi, China.', *Studies in Family Planning* 22(4), 255-63.

Van de Walle, D. (2011), 'Lasting Welfare Effects of Widowhood in Mali.', *World Development* 51, 1-19

Vaughan, M. (1987), *The Story of an African Famine : Gender and Famine in Twentieth-Century Malawi.*, Cambridge : Cambridge University Press.

## Figures and Tables

FIGURE 1: Sex ratio by age, among children from current union



Sample: All children from current union of the 936 women in our baseline sample.  
 Percentages are unchanged if we include the 72 children from previous unions.

\* The last age interval is larger to reach a sufficient number of observations (see below).

Age interval	0-3 years	3-6 years	6-9 years	9-12 years	12-24 years	Total
Observations	739	632	441	276	363	2451

FIGURE 2: Non-parametric estimation of survival functions (pooled and separated by parity)

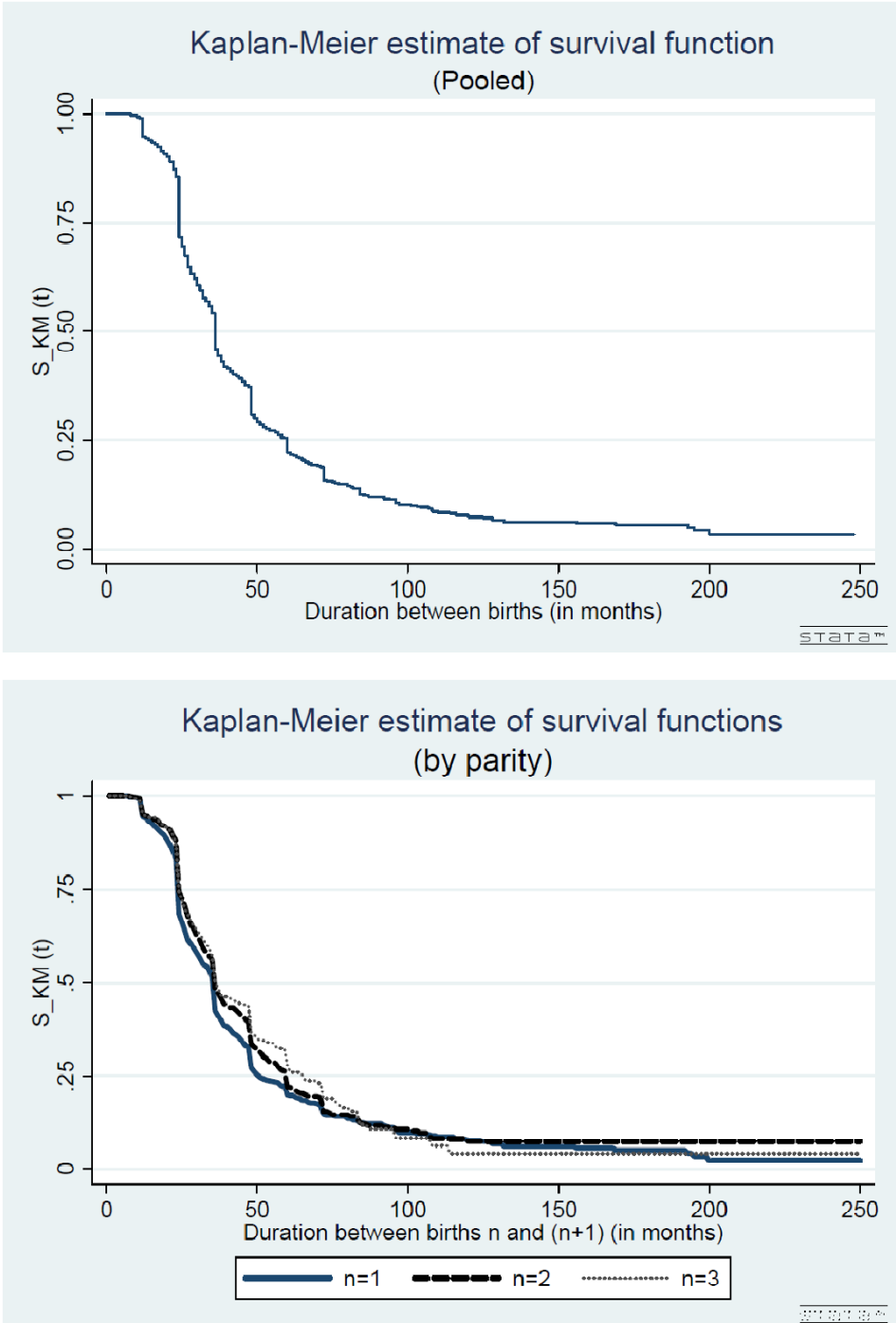


FIGURE 3: Non-parametric estimation of survival function by subpopulation

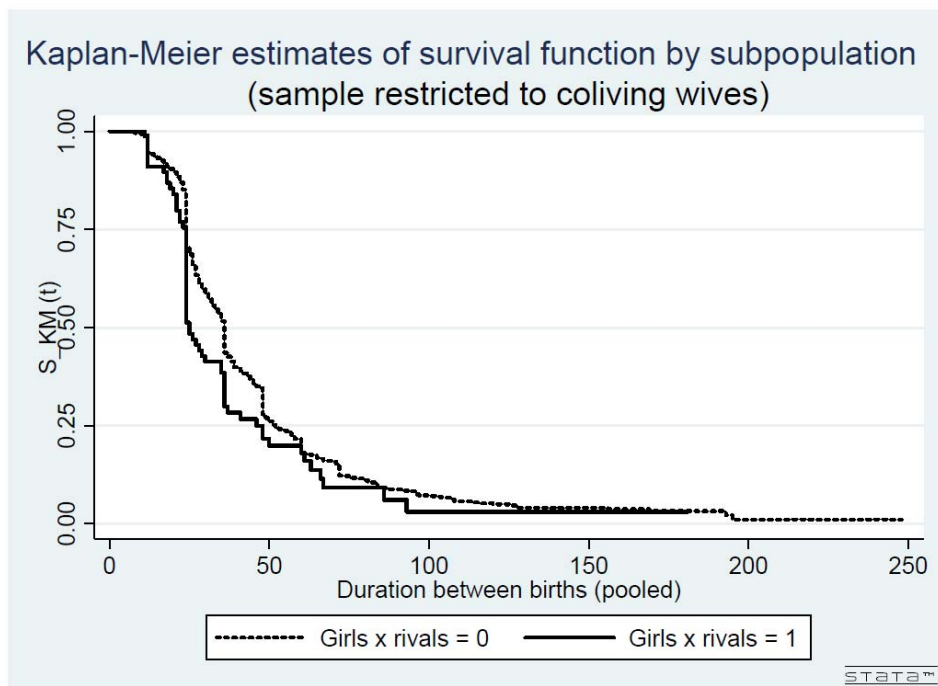
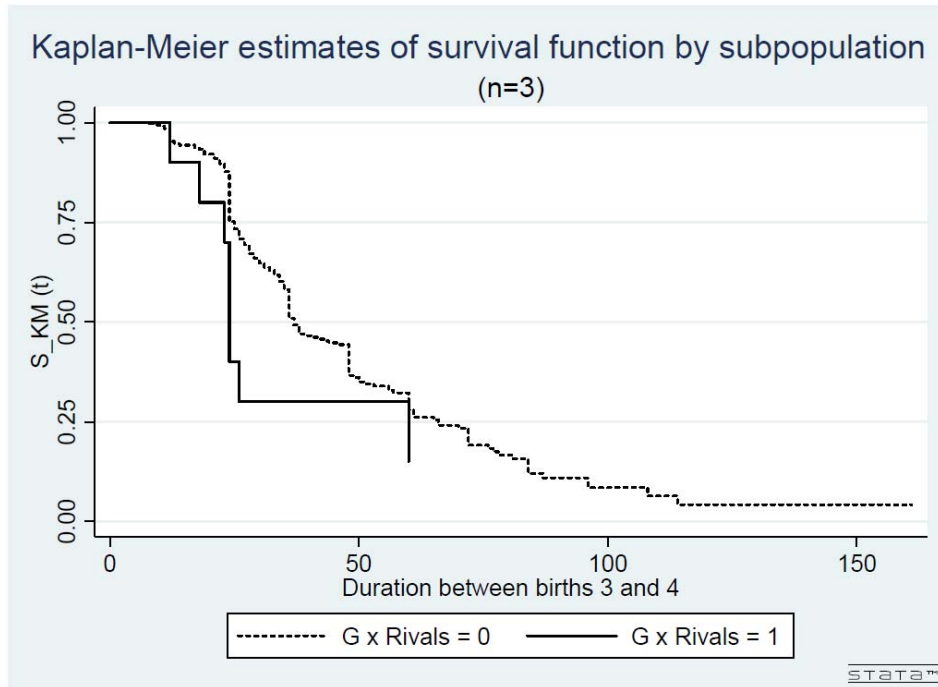




TABLE 1: Widows characteristics, by place of residence after the husband's death

	Remained in the same household	Hosted by another household
Observations	364	182
Still in the household (at the date of the survey)	90%	25%
If not : median length of stay (in years)	7	2
Household head	The widow herself	45%
	A child	30%
	A parent	4%
	A sibling	6%
	Other	15%
Remarried	20%	48%
If remarried, co-residing (with the new husband)	35%	67%

Sample : women whose latest broken union ended with the husband's death (statistics computed for the latest period of widowhood); only 20% of them have experienced more than one broken union.

TABLE 2: Descriptive statistics : Sample of women below 39 years old, engaged in a monogamous union, having at least one child from current union

Number of observations	936
Number of children from current union	2.7
At least one child from previous unions	4.1%
At least one dead child (from current union)	19.8%
Age	27.9
Rural	48.4%
No education	43.8%
Work	32.8%
First union	86.6%
If not : nb broken unions	1.1
Age at 1st marriage	18.9
Live with their husband	75.3%
If co-residing : Age difference	10.2
If co-residing : Public sector	7.3%
If co-residing : Annual income K FCFA (median)	500
At least 1 rival	17.1%
At least 1 male rival	11.9%
Average nb of rivals (inheritance parts)	1.7 (min 0.5 ; max 7.5)
Strictly more rivals than children (inheritance parts)	9.7%

The initial sample consisted of 966 women but we excluded 30 inconsistent cases, for which the number of children found in the database was different from the number reported by the women. The sample restricted to consistent observations is very similar to the initial sample. If anything, inconsistent cases tend to be slightly older, less educated and less likely to live with the husband.

TABLE 3: Descriptive statistics : Comparison between women with and without rivals

	At least one rival	No rival	p-value
Number of observations	159	770	
Number of children from current union	3.1	2.6	0.01***
At least one child from previous unions	10.7%	2.6%	0.01***
At least one dead child (from current union)	26.7%	18.5%	0.04**
Interval between births 1 and 2 (in months)	33.3	36.8	0.10*
Age	30.2	27.5	0.01***
Rural	44.7%	49.5%	0.27
No education	40.6%	44.6%	0.37
Work	40.5%	30.8%	0.02**
First union	78.5%	88.8%	0.01***
If not : nb broken unions	1.2	1.05	0.14
Age at 1st marriage	19.2	18.8	0.29
Live with their husband	81.2%	74.8%	0.07*
If co-residing : Age difference	13.7	9.5	0.01***
If co-residing : Public sector	8.7%	7.0%	0.52
If co-residing : Annual income K FCFA (median)	500	500	0.83

We conduct t-tests comparing the means of 14 covariates, and a nonparametric equality-of-medians test for the income : a low p-value indicates that both subsamples are statistically different with respect to the corresponding covariate. Significance levels : \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

The total number of observations is different from 936 because there are 7 missing values for the presence of rivals.

TABLE 4: Descriptive statistics : Comparison between wives co-residing and non co-residing

	Co-residing wives	Non co-residing wives	p-value
Nb observations	705	231	
Nb children (in current union)	2.8	2.1	0.01***
Children from previous unions $\geq 1$	3.4%	6.0%	0.12
Dead children (in current union) $\geq 1$	21.2%	15.2%	0.04**
Age	27.9	28.0	0.93
Rural	48.9%	46.8%	0.56
No education	44.0%	43.0%	0.79
Work	32.8%	32.9%	0.98
First union	92.9%	66.5%	0.01***
If not : nb broken unions	1.0	1.1	0.08*
Age at 1st marriage	18.7	19.4	0.04**
At least 1 rival	18.3%	13.4%	0.07*
At least 1 male rival	12.9%	8.9%	0.08*
Average nb rivals (in parts)	1.7	1.7	0.81

We conduct t-tests comparing the means of each covariate : a low p-value indicates that both subsamples are statistically different with respect to the corresponding covariate. Significance levels : \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

TABLE 5: Statistics on non-censored durations

	Pooled	$n = 1$	$n = 2$	$n = 3$
<i>Total number obs</i>	1939	904	644	391
Number non-censored obs	1250	634	395	221
Mean	36.2	36.4	35.9	36.2
Std dev.	21.3	23.3	18.5	19.8
Min	8	8	8	8
Q1	24	24	24	24
Q2	30	30	31	30
Q3	45	44	47	48
Max	200	200	120	114
Number obs $\geq 96$ months	29	20	5	4

Statistics computed on non-censored durations, expressed in months.

Approx. 2% of birth intervals are larger than 8 years : in half of the cases, women have in fact lost 1 or 2 children in the meantime. The remaining cases are mostly urban, educated, working women living in relatively rich households.

Results are qualitatively unchanged if we exclude extreme values (above 96 months) from the sample.

TABLE 6: Cox estimation (pooled)

<b>Hazard ratios = <math>e^{\beta_k}</math></b>	Whole sample (1)	Co-residing wives (2)	Poorest half (3)	Richest half (4)
<i>Girls</i>	1.00 (.069)	1.07 (.088)	1.08 (.118)	1.04 (.137)
<i>Rivals</i>	1.03 (.108)	0.94 (.112)	0.78 (.133)	1.01 (.164)
<i>Girls</i> $\times$ <i>Rivals</i>	1.23 (.217)	1.54** (.334)	2.26*** (.612)	1.21 (.395)
Baseline controls	Yes	Yes	Yes	Yes
Additional controls	No	Yes	Yes	Yes
Log-likelihood	-6569	-4607	-2111	-1956
Observations	1589	1117	569	548
Clusters	761	515	260	255
Obs in cell : <i>Girls</i> $\times$ <i>Rivals</i> = 1	93	64	36	28

Robust standard errors of  $e^{\beta_k}$  are in parentheses (clustered at the woman level) - Significance levels : \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . *Girls* : first children are only girls ; *Rivals* : there is at least one rival.

Column (2) : sample restricted to wives living with their husband ; this sample is then split on the median income into the poorest half in column (3) and the richest half in column (4).

Baseline controls mainly include characteristics of the woman : age at preceding birth, place of residence, level of education, age at first marriage, being or not in her first marriage, living or not with the husband, having at least one dead child, having at least one child from previous unions, rank of preceding birth.

Additional controls include : occupation of the wife, sector of activity of the husband, income of the husband, age difference with the husband.

Breslow method to handle ties among non-censored durations.

TABLE 7: Cox estimation (separated by parity, whole sample)

<b>Hazard ratios = <math>e^{\beta_k}</math></b>	$n = 1$	$n = 2$	$n = 3$
<i>G</i>	0.98 (.095)		
<i>Rivals</i>	0.92 (.158)		
<i>G</i> × <i>Rivals</i>	1.35 (.323)		
<i>GG</i>		1.07 (.149)	
<i>Rivals</i>		1.30 (.211)	
<i>GG</i> × <i>Rivals</i>		0.90 (.300)	
<i>GGG</i>			0.79 (.204)
<i>Rivals</i>			0.82 (.161)
<i>GGG</i> × <i>Rivals</i>			3.20** (1.548)
Controls	Yes	Yes	Yes
Log-likelihood	-1889	-1199	-645
Observations	737	528	324
Obs in cell : <i>Girls</i> × <i>Rivals</i> = 1	57	26	10

Standard errors of  $e^{\beta_k}$  are in parentheses - Significance levels : \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

*G*, *GG*, *GGG* : first  $n$  children are girls; *Rivals* : there is at least one rival.

Controls include characteristics of the mother : age at birth  $n$ , place of residence, level of education, age at first marriage, being or not in her first marriage, living or not with the husband, having at least one dead child, having at least one child from previous unions.

Exact marginal-likelihood method to handle ties among non-censored durations.

TABLE 8: Quantities of interest derived from Cox estimation

<b>Co-residing wives</b>	<i>Girls = 0</i> <i>Rivals = 0</i>	<i>Girls = 1</i> <i>Rivals = 0</i>	<i>Girls = 0</i> <i>Rivals = 1</i>	<i>Girls = 1</i> <i>Rivals = 1</i>
Expected birth spacing in months	38.2	36.7	39.5	29.8
Probability of birth spacing $\leq 24$ months	33.0%	35.9%	32.4%	47.6%
<b>Poorest half (of co-residing wives)</b>	<i>Girls = 0</i> <i>Rivals = 0</i>	<i>Girls = 1</i> <i>Rivals = 0</i>	<i>Girls = 0</i> <i>Rivals = 1</i>	<i>Girls = 1</i> <i>Rivals = 1</i>
Expected birth spacing in months	36.0	35.1	38.5	27.5
Probability of birth spacing $\leq 24$ months	33.8%	36.0%	27.5%	54.3%
<b><math>n = 3</math> (within whole sample)</b>	<i>GGG = 0</i> <i>Rivals = 0</i>	<i>GGG = 1</i> <i>Rivals = 0</i>	<i>GGG = 0</i> <i>Rivals = 1</i>	<i>GGG = 1</i> <i>Rivals = 1</i>
Expected birth spacing in months	35.8	39.0	38.5	25.4
Probability of birth spacing $\leq 24$ months	35.1%	28.9%	29.8%	59.2%

Expected birth spacing =  $\mathbb{E}(T \mid X, T \leq t_{(k)}^*)$

Probability of birth spacing  $\leq 24$  months =  $1 - \widehat{S}(24 \mid X)$

We use the baseline survival functions estimated in the Cox model (pooled specification for the first two tables, and  $n = 3$  specification for the last table).

We set the values of the controls to the median of the sample. We check that results are very similar if we set the values of the controls to the sample mean (not shown).

TABLE 9: Fully parametric estimation (same parameters for all parities)

<b>Hazard ratios = <math>e^{\beta_k}</math></b>	Whole sample (1)	Co-residing wives (2)	Poorest half (3)	Richest half (4)
<i>Girls</i>	1.01 (.106)	1.03 (.124)	1.19 (.195)	0.97 (.160)
<i>Rivals</i>	0.97 (.147)	1.02 (.179)	0.69 (.172)	1.08 (.288)
<i>Girls</i> × <i>Rivals</i>	1.60** (.387)	2.29*** (.653)	3.87*** (1.440)	1.79 (.863)
Baseline controls	Yes	Yes	Yes	Yes
Additional controls	No	Yes	Yes	Yes
Log-likelihood	-2950	-2134	-908	-1203
Observations	761	515	260	255

Standard errors of  $e^{\beta_k}$  are in parentheses - Significance levels : \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.

*Girls* : first children are only girls ; *Rivals* : there is at least one rival.

Column (2) : sample restricted to wives living with their husband ; this sample is then split on the median income into the poorest half in column (3) and the richest half in column (4).

Baseline controls mainly include characteristics of the woman : age at preceding birth, place of residence, level of education, age at first marriage, being or not in her first marriage, living or not with the husband, having at least one dead child, having at least one child from previous unions, rank of preceding birth.

Additional controls include : occupation of the wife, sector of activity of the husband, income of the husband, age difference with the husband.

TABLE 10: Fully parametric estimation (parameters specific to each parity, whole sample)

<b>Hazard ratios=</b> $e^{\beta_k}$	$n = 1$	$n = 2$	$n = 3$
<i>G</i>	1.02 (.130)		
<i>Rivals</i>	1.00 (.217)		
$G \times Rivals$	1.47 (.447)		
<i>GG</i>		1.03 (.179)	
<i>Rivals</i>		1.38 (.286)	
$GG \times Rivals$		0.82 (.344)	
<i>GGG</i>			0.86 (.263)
<i>Rivals</i>			0.85 (.207)
$GGG \times Rivals$			7.46*** (4.767)
Controls †		Yes	
Log-likelihood		-2935	
Observations		761	

Standard errors of  $e^{\beta_k}$  are in parentheses - Significance levels : \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

*G*, *GG*, *GGG* : first  $n$  children are girls; *Rivals* : there is at least one rival.

† In this specification, the model is constrained to ensure a manageable number of parameters. In the logit, we keep only the variables whose impact was significant in the estimation on all parities : place of residence, being or not in her first marriage, living or not with the husband. In the survival function, we keep our variables of interest and the controls with a significant impact : age at birth  $n$ , level of education, age at first marriage, living or not with the husband, having at least one dead child, having at least one child from previous unions.

TABLE 11: Quantities of interest derived from fully parametric estimation

<b>Co-residing wives</b>	<i>Girls = 0</i> <i>Rivals = 0</i>	<i>Girls = 1</i> <i>Rivals = 0</i>	<i>Girls = 0</i> <i>Rivals = 1</i>	<i>Girls = 1</i> <i>Rivals = 1</i>
Expected birth spacing in months	38.1	37.6	37.8	27.1
Probability of birth spacing $\leq 24$ months	20.5%	21.1%	20.8%	42.7%
<b>Poorest half (of co-residing wives)</b>	<i>Girls = 0</i> <i>Rivals = 0</i>	<i>Girls = 1</i> <i>Rivals = 0</i>	<i>Girls = 0</i> <i>Rivals = 1</i>	<i>Girls = 1</i> <i>Rivals = 1</i>
Expected birth spacing in months	35.5	33.3	40.8	23.6
Probability of birth spacing $\leq 24$ months	22.8%	26.4%	16.4%	56.1%
<b><math>n = 3</math> (within whole sample)</b>	<i>GGG = 0</i> <i>Rivals = 0</i>	<i>GGG = 1</i> <i>Rivals = 0</i>	<i>GGG = 0</i> <i>Rivals = 1</i>	<i>GGG = 1</i> <i>Rivals = 1</i>
Expected birth spacing in months	37.6	40.2	40.4	19.5
Probability of birth spacing $\leq 24$ months	23.5%	20.6%	20.4%	77.0%

Expected birth spacing =  $\mathbb{E}(T \mid X, \nu = 0, \hat{\theta})$

Probability of birth spacing  $\leq 24$  months =  $1 - S(24 \mid X, \nu = 0, \hat{\theta})$

We use the survival function and the parameters estimated in the fully parametric specification (same parameters for all parities in the first two tables, and parameters specific to  $n = 3$  in the last table).

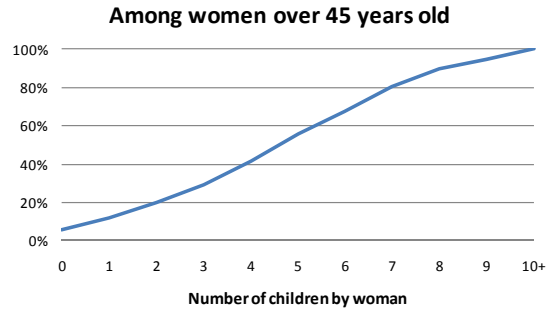
We set the values of the controls to the median of the sample. We check that results are very similar if we set the values of the controls to the sample mean (not shown).



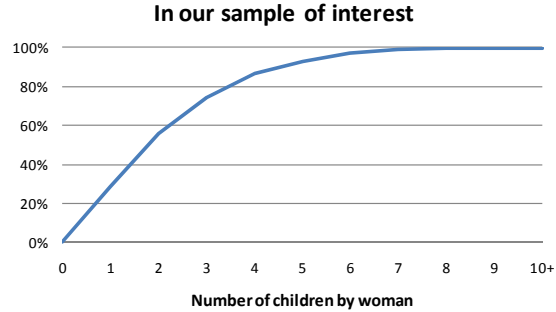
# Appendix

## Appendix A : Additional descriptive statistics

FIGURE 4: Cumulative distribution of the number of children by woman



Sample: Women over 45 years old in the PSF survey (1060 observations)  
Distribution in line with national statistics reported in DHS-MICS, 2010-11.



Sample: Women below 39 years old, engaged in a monogamous union, having at least one child from current union (936 observations).

TABLE 12: Number of observations : Breakdown by missing variables and cells of interest

	Pooled	$n = 1$	$n = 2$	$n = 3$
<b>Initial number obs</b>	<b>2083</b>	<b>966</b>	<b>686</b>	<b>431</b>
Inconsistency	67	30	20	17
Twins	27	9	7	11
Missing duration	50	23	15	12
<b>Number observations (non-parametric estimation)</b>	<b>1939</b>	<b>904</b>	<b>644</b>	<b>391</b>
At least 1 missing obs	350	167	116	67
sex children	2	2	0	0
rivals	14	7	4	3
previous children	0	0	0	0
age	0	0	0	0
Missing : rural	0	0	0	0
education	64	27	21	16
age first marriage	93	51	27	15
first union	124	55	45	24
co-residing	0	0	0	0
dead children	111	50	41	20
<b>Number observations (parametric estimation)</b>	<b>1589</b>	<b>737</b>	<b>528</b>	<b>324</b>
<i>Girls</i> = 1	505	365	140	43
<i>Rivals</i> = 1	270	112	91	67
Among which : <i>Girls</i> × <i>Rivals</i> =1	93	57	26	10

The "pooled" column brings together all birth intervals, which are then split by parity ; the 1589 observations correspond to 761 different women. But the duration between children 1 and 2 is missing for a few of them, that is why we have only 737 observations when  $n = 1$ .

Sample : Women below 39 years old, engaged in a monogamous union, having at least one child from current union. Inconsistency : women reporting a number of children from current union different from the number found in the database.

*Girls* : first  $n$  children are only girls, *Rivals* : there is at least 1 rival.

TABLE 13: Balancing tests : Having only daughters vs. having at least one son

	n=1		n=2		n=3		
	G=0	G=1	GG=0	GG=1	GGG=0	GGG=1	p-value
Nb observations	372	365	388	140	281	43	
Nb children (in current union)	2.6	2.7	3.3	3.5	4.1	4.4	0.13
Children from previous unions $\geq 1$	4.0%	3.8%	2.6%	5%	2.8%	2.3%	0.84
Dead children (in current union) $\geq 1$	18.8%	20.8%	25%	18.6%	26%	34.9%	0.25
Age	27.9	28.1	29.3	29.6	31.1	30.8	0.71
Rural	49.7%	44.3%	50.3%	50.7%	52.7%	58.1%	0.50
No education	45.7%	42.7%	45.6%	43.6%	46.3%	55.8%	0.25
First union	90.1%	86.6%	91.5%	85.7%	92.2%	97.7%	0.05**
If not : nb broken unions	1.1	1.05	1.1	1.05	1.1	1.0	0.08*
Age at 1st marriage	19.0	19.0	18.7	18.3	18.4	18.2	0.76
At least 1 rival	15.3%	16.7%	17.5%	18.6%	21%	25.5%	0.52
At least 1 male rival	9.1%	12.0%	11.1%	15%	12.8%	18.6%	0.36
Average nb rivals (in parts)	1.4	1.8	1.5	2.1	1.6	2.0	0.41
Work	32.8%	32.6%	32.9%	32.1%	33.3%	27.9%	0.47
Live with the husband	76.3%	79.4%	82.0%	85.0%	83.6%	88.4%	0.38
If co-residing : Age difference	10.0	10.4	10.1	10.8	10.2	12.2	0.07*
If co-residing : Husband in public sector	6.6%	9.5%	6.9%	8.6%	4.8%	8.1%	0.49
If co-residing : Annual income K FCFA (median)	480	600	500	600	500	310	0.43

We conduct t-tests comparing the means of 16 covariates, and a nonparametric equality-of-medians test for the income : a low p-value indicates that both subsamples are statistically different with respect to the corresponding covariate. Significance levels : \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.  $G$ ,  $GG$ ,  $GGG$  : first  $n$  children are girls.

## Appendix B : Measures of birth intervals

We have 4 types of observations. Table 14 below gives further detail about the number of observations of each type :

1. Not censored, precisely measured ; we observe :

$$t_n = (\text{birth year of child } (n + 1) - \text{birth year of child } n) \times 12 + \text{birth month of child } (n + 1) - \text{birth month of child } n$$

So depending on the birth days of both children, we infer that  $T_n \in [t_n - 1; t_n + 1]$ .

2. Not censored, imprecisely measured ; we observe :

$$t_n = (\text{age of child } n - \text{age of child } (n + 1)) \times 12$$

So depending on the exact birth dates of both children, we infer that  $T_n \in [t_n - 12; t_n + 12]$ <sup>39</sup>.

3. Censored, precisely measured ; we observe :

$$t_n = (\text{year of survey} - \text{birth year of child } n) \times 12 + \text{month of survey} - \text{birth month of child } n$$

We infer that  $T_n > t_n - 1$ .

4. Censored, imprecisely measured ; we observe :

$$t_n = (\text{age of child } n) \times 12$$

We infer that  $T_n > t_n$ .

---

39. E.g. if one child is 4 years old and the next one is 2 years old, we measure  $t = 24$  and the two extreme cases are given by (i) the eldest will turn 5 the day following the survey while the youngest turned 2 the day before : the real  $t$  is 36 ; and (ii) the youngest will turn 3 the day following the survey while the eldest turned 4 the day before : the real  $t$  is 12.

TABLE 14: Number and frequency of observations by censoring and measurement status

<b>Pooled</b>	Precisely measured	Imprecisely measured	Total
Non-censored	621 (39%)	413 (26%)	1034 (65%)
Censored	441 (28%)	114 (7%)	555 (35%)
Total	1062 (66%)	527 (34%)	1589 (100%)

<b>n=1</b>	Precisely measured	Imprecisely measured	Total
Non-censored	314 (43%)	203 (28%)	517 (70%)
Censored	176 (24%)	44 (6%)	220 (30%)
Total	490 (66%)	247 (34%)	737 (100%)

<b>n=2</b>	Precisely measured	Imprecisely measured	Total
Non-censored	198 (37%)	132 (25%)	330 (63%)
Censored	156 (30%)	42 (8%)	198 (38%)
Total	354 (67%)	174 (33%)	528 (100%)

<b>n=3</b>	Precisely measured	Imprecisely measured	Total
Non-censored	109(34%)	78 (24%)	187 (57%)
Censored	109(34%)	28 (9%)	137 (43%)
Total	218 (67%)	106 (33%)	324 (100%)

A duration is considered as imprecisely measured when it is derived from the age of the child instead of the exact birthdate.

## Appendix C : Fully parametric duration model

The survival function conditional on  $R_n$ ,  $X_n$  and  $\nu$  given in the appendix allows to retrieve the aggregate survival function conditional on  $X$  and  $\nu$ .

$$S_1(t | R_n = 1, X_n, \nu) = 1$$

$$S_1(t | R_n = 0, X_n, \nu) = S(t | X_n, \nu) = \begin{cases} \exp\left(-(\lambda(t-8))^a \exp(X_n' \beta + \nu)\right) & \text{if } t \geq 8 \\ 1 & \text{if } t < 8 \end{cases}$$

We derive the aggregate survival function conditional on  $X$  and  $\nu$  :

$$S_1(t | X_n, \nu) = p(X_n) + (1 - p(X_n)) \times S(t | X_n, \nu)$$

Therefore,  $S_1(t | X_n) \rightarrow p(X_n) > 0$  when  $t \rightarrow \infty$ . Using this specification, we are able to disentangle the effect of covariates, in particular *Girls*, *Rivals* and *Girls × Rivals*, on the number of births and on the spacing of births. The vector  $\alpha$  captures the impact on the probability to stop having children, whereas the vector  $\beta$  captures the impact on durations.

We estimate a multispell model in which the individual contribution to the likelihood  $L_i(\theta | t_i, X_i, \nu_i)$  is given by<sup>40</sup> :

1.  $p(X_{1,i}) + (1 - p(X_{1,i})) \times s(t_{1,i} | X_{1,i}, \nu_i)$  if woman  $i$  has exactly 1 child.
2.  $\left((1 - p(X_{1,i})) \times f(t_{1,i} | X_{1,i}, \nu_i)\right) \times \left(p(X_{2,i}) + (1 - p(X_{2,i})) \times s(t_{2,i} | X_{2,i}, \nu_i)\right)$  if woman  $i$  has exactly 2 children.
3.  $\left((1 - p(X_{1,i})) \times f(t_{1,i} | X_{1,i}, \nu_i)\right) \times \left((1 - p(X_{2,i})) \times f(t_{2,i} | X_{2,i}, \nu_i)\right) \times \left(p(X_{3,i}) + (1 - p(X_{3,i})) \times s(t_{3,i} | X_{3,i}, \nu_i)\right)$  if woman  $i$  has exactly 3 children.
4.  $\left((1 - p(X_{1,i})) \times f(t_{1,i} | X_{1,i}, \nu_i)\right) \times \left((1 - p(X_{2,i})) \times f(t_{2,i} | X_{2,i}, \nu_i)\right) \times \left((1 - p(X_{3,i})) \times f(t_{3,i} | X_{3,i}, \nu_i)\right)$  if woman  $i$  has 4 children and more.

Where  $t_{n,i}$  is the duration (potentially right-censored) between births  $n$  and  $(n + 1)$  for woman  $i$ ;  $f(\cdot)$  and  $s(\cdot)$  are defined to explicitly allow for measurement approximations :

$$f(t_{n,i}) = (S(t_{n,i} - 1) - S(t_{n,i} + 1))^{P_{n,i}} \times (S(t_{n,i} - 12) - S(t_{n,i} + 12))^{(1 - P_{n,i})}$$

---

40. The level of analysis is no longer the births as in the Cox model, but the women. Here, the number of observations is exactly equal to the number of clusters in the Cox model (=761).

$$s(t_{n,i}) = S(t_{n,i} - 1)^{P_{n,i}} \times S(t_{n,i})^{(1-P_{n,i})}$$

Where  $P_{n,i}$  is a dummy equal to 1 if the duration  $n$  is precisely measured for woman  $i$ . Hence, the contribution to the likelihood uses a confidence interval that depends on how precisely the duration is measured, instead of the density and survival functions.

Now,  $\theta \equiv \{\alpha, \beta, \lambda, a, \sigma\}$  is the vector of parameters to be estimated. We start by considering the same parameters for all parities and compare the estimates to those found in the pooled Cox regression. Then we allow the parameters<sup>41</sup> to vary depending on  $n$ , and compare the estimates to those found in the three separate Cox regressions. However, the last specification includes too many parameters with respect to the number of observations; therefore we constrain the model and keep only the controls whose impact was significant in the estimation on all parities.

To compute the standard errors of  $e^{\beta_k}$ , we use the following approximation :

$$se(e^{\beta_k}) = e^{\hat{\beta}_k} \times se(\beta_k)$$

Since there is no closed form expression for the expected likelihood, we use the technique of simulated maximum likelihood. We estimate  $\mathbb{E}(L_i(\theta | t_i, X_i))$  with :

$$\widetilde{L}_i(\theta | t_i, X_i) = \frac{1}{H} \sum_{h=1}^H L_i(\theta | t_i, X_i, \nu_{i,h})$$

Where  $H = 60$  is the number of random draws from a normal distribution for each individual.

Last, we are able to compare across different categories of women, the probability of having another child,  $(1 - p(X))$ , and the expected duration conditional on  $X$  :

$$\mathbb{E}(T | X, \nu, \hat{\theta}) = \int_0^{\infty} S(u | X, \nu, \hat{\theta}) du$$

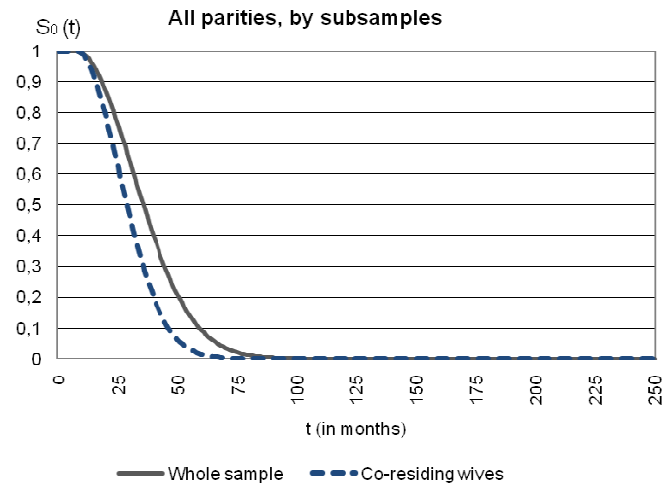
We also consider  $(1 - S(24 | X, \nu, \hat{\theta}))$ . Again, we compute those quantities for the median individual (setting  $\nu = 0$ ) and compare them to the ones obtained with the Cox estimation.

---

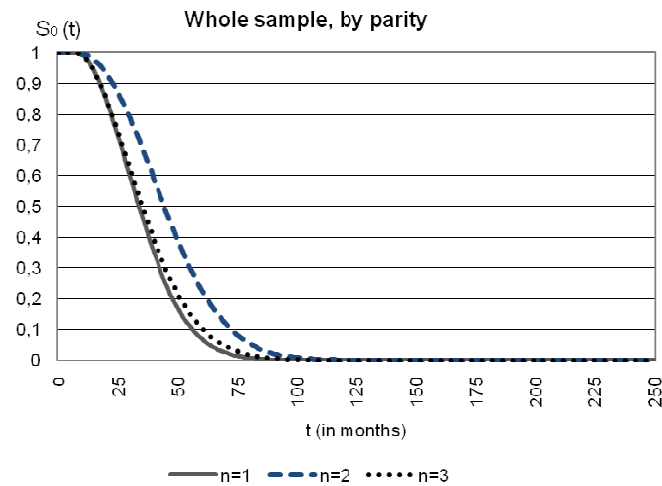
41. Except  $\sigma$ , the standard deviation of the unobserved heterogeneity. This is precisely the strength of the multispell model : random draws from the normal distribution are attributed once and for all to each individual, explaining simultaneously all durations.

## Appendix D : Additional estimates of the fully parametric specification

FIGURE 5: Fully parametric estimation of baseline survival function



Estimates	$\lambda$	$a$
<b>Whole sample</b>	0.03***	1.93***
<b>Co-residing wives</b>	0.04***	1.95***



Estimates	$\lambda$	$a$
<b>n=1</b>	0.03***	1.90***
<b>n=2</b>	0.02***	2.10***
<b>n=3</b>	0.03***	1.80***

Significance levels : \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.



## Appendix E : Robustness tests

TABLE 15: Robustness : Cox estimation (pooled, sample of co-residing wives)

Hazard ratios = $e^{\beta_k}$	Baseline specification	All parities	Ethnic groups
	(1)	(2)	(3)
<i>Girls</i>	1.07 (.088)	1.12 (.089)	1.08 (.088)
<i>Rivals</i>	0.94 (.112)	0.93 (.092)	0.94 (.114)
<i>Girls</i> $\times$ <i>Rivals</i>	1.54** (.334)	1.49** (.308)	1.55** (.334)
Parities	$n \leq 3$	all $n$	$n \leq 3$
Controls	Yes	Yes	Yes
Ethnic group	No	No	Yes
Log-likelihood	-4607	-5633	-4604
Observations	1117	1271	1117
Clusters	515	518	515
Obs in cell : <i>Girls</i> $\times$ <i>Rivals</i> = 1	64	68	64

Robust standard errors of  $e^{\beta_k}$  are in parentheses (clustered at the woman level) - Significance levels : \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . *Girls* : first children are only girls; *Rivals* : there is at least one rival.

Sample : co-residing wives; column (1) : baseline specification ( $n \leq 3$ , no control for ethnic groups); column (2) : we include all parities; column (3) : we control for ethnic groups.

Controls include characteristics of the woman : age at preceding birth, place of residence, level of education, age at first marriage, being or not in her first marriage, living or not with the husband, having at least one dead child, having at least one child from previous unions, occupation, sector of activity of the husband, income of the husband, age difference with the husband and rank of preceding birth.

Breslow method to handle ties among non-censored durations.

TABLE 16: Effect detectable after the first child : Cox estimation ( $n = 1$ , by sub-samples)

<b>Hazard ratios = <math>e^{\beta_k}</math></b>	Whole sample (1)	Co-residing wives (2)	Poorest half (3)	Richest half (4)
<i>G</i>	0.98 (.095)	1.03 (.119)	1.00 (.167)	1.13 (.195)
<i>Rivals</i>	0.92 (.158)	0.87 (.180)	0.52** (.169)	1.17 (.332)
<i>G</i> × <i>Rivals</i>	1.35 (.323)	1.72* (.491)	2.95*** (1.242)	1.22 (.503)
Baseline controls	Yes	Yes	Yes	Yes
Additional controls	No	Yes	Yes	Yes
Log-likelihood	-1889	-1295	-552	-624
Observations	737	500	252	248
Obs in cell : <i>Girls</i> × <i>Rivals</i> = 1	57	38	22	16

Standard errors of  $e^{\beta_k}$  are in parentheses - Significance levels : \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

*G* : first child is a girl; *Rivals* : there is at least one rival.

Column (2) : sample restricted to wives living with their husband; this sample is then split on the median income into the poorest half in column (3) and the richest half in column (4).

Baseline controls mainly include characteristics of the woman : age at birth 1, place of residence, level of education, age at first marriage, being or not in her first marriage, living or not with the husband, having at least one dead child, having at least one child from previous unions.

Additional controls include : occupation of the wife, sector of activity of the husband, income of the husband, age difference with the husband.

Exact marginal-likelihood method to handle ties among non-censored durations.

TABLE 17: No preference for diversity : Cox estimation (pooled, sample of co-residing wives)

<b>Hazard ratios= <math>e^{\beta_k}</math></b>	<b>Co-residing wives</b>
<i>Boys</i>	0.93 (.082)
<i>Rivals</i>	1.11 (.147)
<i>Boys</i> × <i>Rivals</i>	0.90 (.190)
Controls	Yes
Log-likelihood	-4610
Observations	1117

Robust standard errors of  $e^{\beta_k}$  are in parentheses (clustered at the woman level) - Significance levels : \* p<0.10, \*\* p<0.05, \*\*\* p<0.01. *Boys* : first children are only boys; *Rivals* : there is at least one rival.

Controls include characteristics of the woman : age at preceding birth, place of residence, level of education, age at first marriage, being or not in her first marriage, living or not with the husband, having at least one dead child, having at least one child from previous unions, occupation, sector of activity of the husband, income of the husband, age difference with the husband and rank of preceding birth.

Breslow method to handle ties among non-censored durations.

TABLE 18: Common "intrinsic" son preference : Cox estimation ( $n = 2$ , sample of women with at least one son)

<b>Hazard ratios= <math>e^{\beta_k}</math></b>	<b><math>n = 2</math></b>
<i>BB</i>	0.66*** (.107)
<i>Rivals</i>	1.39 (.295)
<i>BB</i> × <i>Rivals</i>	1.08 (.360)
Controls	Yes
Log-likelihood	-855
Observations	388

Standard errors of  $e^{\beta_k}$  are in parentheses - Significance levels : \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.

*BB* : first 2 children are boys; *Rivals* : there is at least one rival.

Controls include characteristics of the mother : age at birth 2, place of residence, level of education, age at first marriage, being or not in her first marriage, living or not with the husband, having at least one dead child, having at least one child from previous unions.

Exact marginal-likelihood method to handle ties among non-censored durations.

TABLE 19: *Rivals* = there is at least one male rival : Cox estimation (pooled)

<b>Hazard ratios = <math>e^{\beta_k}</math></b>	Whole sample (1)	Co-residing wives (2)	Poorest half (3)	Richest half (4)
<i>Girls</i>	1.00 (.067)	1.09 (.087)	1.13 (.121)	1.05 (.133)
<i>Rivals</i>	0.86 (.112)	0.81 (.117)	0.69* (.148)	0.89 (.185)
<i>Girls</i> $\times$ <i>Rivals</i>	1.37 (.286)	1.58* (.404)	2.24** (.731)	1.19 (.436)
Baseline controls	Yes	Yes	Yes	Yes
Additional controls	No	Yes	Yes	Yes
Log-likelihood	-6569	-4606	-2112	-1956
Observations	1589	1117	569	548
Clusters	761	515	260	255
Obs in cell : <i>Girls</i> $\times$ <i>Rivals</i> = 1	70	48	29	19

Robust standard errors of  $e^{\beta_k}$  are in parentheses (clustered at the woman level) - Significance levels : \* p<0.10, \*\* p<0.05, \*\*\* p<0.01. *Girls* : first children are only girls; *Rivals* : there is at least one male rival.

Column (2) : sample restricted to wives living with their husband; this sample is then split on the median income into the poorest half in column (3) and the richest half in column (4).

Baseline controls mainly include characteristics of the woman : age at preceding birth, place of residence, level of education, age at first marriage, being or not in her first marriage, living or not with the husband, having at least one dead child, having at least one child from previous unions, rank of preceding birth.

Additional controls include : occupation of the wife, sector of activity of the husband, income of the husband, age difference with the husband.

Breslow method to handle ties among non-censored durations.

TABLE 20: *Rivals* = there is at least one male rival : Cox estimation (separated by parity, whole sample)

<b>Hazard ratios= <math>e^{\beta_k}</math></b>	<i>n</i> = 1	<i>n</i> = 2	<i>n</i> = 3
<i>G</i>	0.98 (.092)		
<i>Rivals</i>	0.83 (.191)		
<i>G</i> × <i>Rivals</i>	1.58 (.473)		
<i>GG</i>		1.07 (.145)	
<i>Rivals</i>		1.06 (.214)	
<i>GG</i> × <i>Rivals</i>		0.86 (.325)	
<i>GGG</i>			0.83 (.201)
<i>Rivals</i>			0.61** (.147)
<i>GGG</i> × <i>Rivals</i>			4.24** (2.448)
Controls	Yes	Yes	Yes
Log-likelihood	-1889	-1201	-640
Observations	737	528	324
Obs in cell : <i>Girls</i> × <i>Rivals</i> = 1	42	21	7

Standard errors of  $e^{\beta_k}$  are in parentheses - Significance levels : \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

*G*, *GG*, *GGG* : first *n* children are girls; *Rivals* : there is at least one male rival.

Controls include characteristics of the mother : age at birth *n*, place of residence, level of education, age at first marriage, being or not in her first marriage, living or not with the husband, having at least one dead child, having at least one child from previous unions.

Exact marginal-likelihood method to handle ties among non-censored durations.

TABLE 21: At least one rival, by gender : Cox estimation (pooled, sample of co-residing wives)

<b>Hazard ratios</b> = $e^{\beta_k}$	At least 1 male rival (1)	At least 1 female rival (2)	Both (3)
<i>Girls</i>	1.09 (.087)	1.11 (.089)	1.09 (.088)
<i>Rivals_male</i>	0.81 (.117)		0.79 (.124)
<i>Rivals_female</i>		0.97 (.142)	1.07 (.161)
<i>Girls</i> × <i>Rivals_male</i>	1.58* (.404)		1.47 (.484)
<i>Girls</i> × <i>Rivals_female</i>		1.33 (.343)	1.09 (.360)
Controls	Yes	Yes	Yes
Log-likelihood	-4607	-4609	-4607
Observations	1117	1117	1117
Clusters	515	515	515
Obs in cell : <i>Girls</i> × <i>Rivals_male</i> = 1	48	na	48
Obs in cell : <i>Girls</i> × <i>Rivals_female</i> = 1	na	40	40

Robust standard errors of  $e^{\beta_k}$  are in parentheses (clustered at the woman level) - Significance levels : \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.

*Girls* : first children are only girls; *Rivals\_male* : there is at least one male rival; *Rivals\_female* : there is at least one female rival.

Controls include characteristics of the woman : age at preceding birth, place of residence, level of education, age at first marriage, being or not in her first marriage, living or not with the husband, having at least one dead child, having at least one child from previous unions, occupation, sector of activity of the husband, income of the husband, age difference with the husband and rank of preceding birth.

Breslow method to handle ties among non-censored durations.

TABLE 22: *Rivals* = number of rivals in inheritance parts : Cox estimation (pooled)

<b>Hazard ratios = <math>e^{\beta_k}</math></b>	Whole sample	Co-residing wives	Poorest half	Richest half
	(1)	(2)	(3)	(4)
<i>Girls</i>	1.02 (.067)	1.11 (.088)	1.09 (.119)	1.08 (.134)
<i>Rivals</i>	0.97 (.050)	0.94 (.056)	0.79** (.077)	1.03 (.043)
<i>Girls</i> $\times$ <i>Rivals</i>	1.06 (.064)	1.11 (.087)	1.59*** (.176)	0.97 (.056)
Baseline controls	Yes	Yes	Yes	Yes
Additional controls	No	Yes	Yes	Yes
Log-likelihood	-6570	-4608	-2109	-1956
Observations	1589	1117	569	548
Clusters	761	515	260	255
Obs in cell : <i>Girls</i> $\times$ <i>Rivals</i> > 0	93	64	36	28

Robust standard errors of  $e^{\beta_k}$  are in parentheses (clustered at the woman level) - Significance levels : \* p<0.10, \*\* p<0.05, \*\*\* p<0.01. *Girls* : first children are only girls; *Rivals* : number of rivals in inheritance parts.

Column (2) : sample restricted to wives living with their husband; this sample is then split on the median income into the poorest half in column (3) and the richest half in column (4).

Baseline controls mainly include characteristics of the woman : age at preceding birth, place of residence, level of education, age at first marriage, being or not in her first marriage, living or not with the husband, having at least one dead child, having at least one child from previous unions, rank of preceding birth.

Additional controls include : occupation of the wife, sector of activity of the husband, income of the husband, age difference with the husband.

Breslow method to handle ties among non-censored durations.

TABLE 23: *Rivals* = number of rivals in inheritance parts : Cox estimation (separated by parity, whole sample)

<b>Hazard ratios= <math>e^{\beta_k}</math></b>	<i>n</i> = 1	<i>n</i> = 2	<i>n</i> = 3
<i>G</i>	1.00 (.093)		
<i>Rivals</i>	0.92 (.077)		
<i>G</i> × <i>Rivals</i>	1.11 (.113)		
<i>GG</i>		1.06 (.141)	
<i>Rivals</i>		1.06 (.074)	
<i>GG</i> × <i>Rivals</i>		0.95 (.103)	
<i>GGG</i>			0.85 (.204)
<i>Rivals</i>			0.88 (.088)
<i>GGG</i> × <i>Rivals</i>			1.54** (.266)
Controls	Yes	Yes	Yes
Log-likelihood	-1890	-1200	-644
Observations	737	528	324
Obs in cell : <i>Girls</i> × <i>Rivals</i> > 0	57	26	10

Standard errors of  $e^{\beta_k}$  are in parentheses - Significance levels : \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.

*G*, *GG*, *GGG* : first *n* children are girls; *Rivals* : number of rivals in inheritance parts.

Controls include characteristics of the mother : age at birth *n*, place of residence, level of education, age at first marriage, being or not in her first marriage, living or not with the husband, having at least one dead child, having at least one child from previous unions.

Exact marginal-likelihood method to handle ties among non-censored durations.



TABLE 24: Non-linear impact of the number of rivals : Cox estimation (pooled, sample of co-residing wives)

<b>Hazard ratios = <math>e^{\beta_k}</math></b>	Nb rivals $\geq 1$	Nb rivals $\geq 2$	Nb rivals = 1	Nb rivals = 1, by gender
	(1)	(2)	(3)	(4)
<i>Girls</i>	1.11 (.088)	1.10 (.091)	1.05 (.087)	1.06 (.089)
<i>Nb_parts</i>	0.94 (.056)	0.96 (.056)	0.87 (.152)	
<i>Nb_male</i>				0.75 (.149)
<i>Nb_female</i>				1.30 (.208)
<i>Girls</i> $\times$ <i>Nb_parts</i>	1.11 (.087)	1.09 (.082)	1.90* (.697)	
<i>Girls</i> $\times$ <i>Nb_male</i>				2.20** (.866)
<i>Girls</i> $\times$ <i>Nb_female</i>				0.99 (.388)
Controls	Yes	Yes	Yes	Yes
Log-likelihood	-4608	-4039	-4097	-4095
Observations	1117	1005	1023	1023
Clusters	515	464	474	474
Obs in cell : <i>Girls</i> $\times$ <i>Nb_parts</i> $> 0$	64	38	26	na
Obs in cell : <i>Girls</i> $\times$ <i>Nb_male</i> $> 0$	na	na	na	21
Obs in cell : <i>Girls</i> $\times$ <i>Nb_female</i> $> 0$	na	na	na	5

Robust standard errors of  $e^{\beta_k}$  are in parentheses (clustered at the woman level) - Significance levels : \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

*Girls* : first children are only girls; *Nb\_parts* : number of rivals in inheritance parts; *Nb\_male* : number of male rivals; *Nb\_female* : number of female rivals.

The sample of co-residing wives in column (1) is split on the number of rivals : column (2) includes only women with 2 and more rivals whereas columns (3) and (4) include only women with exactly 1 rival.

Controls include characteristics of the woman : age at preceding birth, place of residence, level of education, age at first marriage, being or not in her first marriage, living or not with the husband, having at least one dead child, having at least one child from previous unions, occupation, sector of activity of the husband, income of the husband, age difference with the husband and rank of preceding birth.

Breslow method to handle ties among non-censored durations.