

# HUMAN FERTILITY DATABASE DOCUMENTATION: SPAIN

## Authors:

**Daniel Devolder**

Centre d'Estudis Demogràfics, Universitat Autònoma de Barcelona

E-mail: [ddevolder@ced.uab.es](mailto:ddevolder@ced.uab.es)

**Elsa Ortiz**

Centre d'Estudis Demogràfics, Universitat Autònoma de Barcelona

**Kryštof Zeman**

Wittgenstein Centre, Vienna Institute of Demography

E-mail: [krystof.zeman@oeaw.ac.at](mailto:krystof.zeman@oeaw.ac.at)

29 August 2016

**WARNING: Due to lower quality of input data, age-specific indicators prior to 1975 should be used with caution. Quality issues are also frequent in data after 1975 (under-registration of births in the early 1980s, incorrect assignment of birth order in 1975–1979, 1986, and 2005–2009), indicating that data by birth order be used carefully for this period. All cohort fertility indicators for Spain, especially cohort childlessness, should be used cautiously as well. For further details on data quality issues, please see Section 4.2.**

## 1 General information

This report documents Spanish data collected for the Human Fertility Database project, obtained from the civil registration system and usually published in the successive volumes of the *Movimiento Natural de la Población* (Vital Statistics publications), or in the Statistical Yearbooks. The data have been available as individual birth records since 1975.

In general, the following types of birth count data are available for Spain, which were collected and published in increasing detail during different periods of time:

- (i) Birth statistics were published on a national basis as early as the 1858–1877 period, but these published data contained only aggregate numbers of baptisms provided by the Catholic Church.
- (ii) For the period 1878–1899 only the annual total number of live births is available.
- (iii) From 1900 onwards monthly total numbers of live births are available.
- (iv) Starting in 1922 births also became available by five-year age group of parents. We included only the information for mothers in the database.
- (v) In 1975 the last important reform of the civil registration system took place, pertaining to the collection of information regarding biological birth order and a change in the definition of a live birth to conform to international standards. The population covered by vital statistics was changed from the *de facto* population (i.e., *de hecho*, based on place of birth/death) to the *de jure* population (i.e., *de derecho*, based on permanent legal residence) (Glei et al., 2016). Another positive change involved the use of computers to process the information, which later made it possible to work directly with individual (micro) datasets from the civil registration system.

For the period 1900–1940, information regarding children born alive can be found in the annual publication of the Vital Statistics. For the period starting in 1941 this information is available on the website of the National Statistics Institute (INE); <http://www.ine.es>.

On its website INE publishes the provisional data six months after the end of the reference year, and the final data are then published between 12 and 18 months after that period has ended.

Throughout the period 1878–2014 there is no homogeneous coverage of these data. Details for each sub-period and the type of information that can be obtained are described in section 2. Also, it is important to note that before 1975 birth counts did not include deaths during delivery and the first 24 hours of life (“false stillbirths”), which were classified as “miscarriages”.

The Vital Statistics system was profoundly reorganised in 1975, adopting the international definition of live births, collecting increasingly detailed information (including, for example, the distribution of parents by single year of age, and of mothers by biological birth order), and using the new means of dissemination: computers (which explains why microdata files for births are available for the period starting in 1975).

More details about Spanish demographic statistics can be found in Nicolau (2005) or Reher and Valero (1995). For detailed, almost encyclopaedic, information on Vital Statistics for the period 1878–1974, see Cusidó (2008)

The HFD features age-specific data on births starting in 1922, whereas monthly births are provided from 1900 onwards. All the input data used for generating the HFD output data and indicators for Spain are specified in Appendix 1.

## **1.1 Territorial coverage**

The geographical area to which the birth count data published in the HFD pertain includes the Spanish part of the Iberian Peninsula and the two archipelagos (Balearic and Canary Islands). The data also include the two major exclaves of Ceuta and Melilla and surroundings, located on the north coast of Morocco.

There have been no territorial changes in Spain during the period covered by HFD data. There was, however, a change in population coverage in 1975 (as noted earlier), when the collection of vital statistics data changed from covering the de facto to the de jure population. In the calculation of population exposure estimates, this change is taken into account (see section 3.1 for details).

## **2 Birth Count Data**

The availability of fertility data for Spain is summarized in Table 1. Within the period 1878–1899, data on births were published only as annual totals, with no information concerning the month of childbirth or the age of the parents. From 1900 through 2014, live birth counts were also given by month of birth, whereas from 1922 to 1974 the data also covered five-year age groups of the mother (below 15, 15–19,..., 45–49, and 50 years and older). After 1975, the data are given by single years of age for the mother as well as by biological birth order (for the mother). All the information for this period was obtained from the official Vital Statistics publications (*Movimiento Natural de la Población*). For the HFD the true birth order was re-estimated, correcting for the wrong classification of births from multiple deliveries. Appendix 3 provides more detailed information on this procedure.

**Table 1** Availability of birth data by period, Spain, 1878–2014

Period	Yearly total	Monthly total	Age of mother	Birth order	Availability in HFD
1878–1899	X				
1900–1921	X	X			X
1922–1974	X	X	Five-year age groups		X
1975 onwards	X	X	Single years of age	X	X

### 3 Population Count Data

#### 3.1 Population data by age

Population exposures by single age group for Spain were estimated using population estimates and death counts from the Human Mortality Database (HMD; [www.mortality.org](http://www.mortality.org)). The data cover the time period since 1908. Although there have been no territorial changes in this period, there was a change in population coverage in 1975: the collection of vital statistics data shifted from covering the de facto to the de jure population. Although this is not a territorial change, we treat it as such in order to make the appropriate adjustments in the formulas that calculate the population estimates and birth rates (Glei et al., 2016).

#### 3.2 Population data by age, birth cohort and parity

The distribution of women by the number of children born alive was obtained from the Population Censuses of 1991 and 2011. For the HFD, 1991 census data were tabulated by birth cohort of the mother and her number of children born alive<sup>1</sup> from the 10% microdata sample. The 2011 Census was based on a combined approach: information from registers was available for the total population with a limited number of variables. In order to complement this information from registers, a 10% sample survey was carried out. The microdata file from this sample survey is used for the tabulation of women by birth cohort and parity for 2011. Both 1991 and 2011 microdata files can be downloaded at [http://www.ine.es/prodyser/micro\\_censopv.htm](http://www.ine.es/prodyser/micro_censopv.htm)

In several previous censuses, the fertility question was also asked of all women but, regrettably, the results were tabulated only for ever-married women. Also, note that there was no question on fertility in the 2001 Census.

### 4 Specific Details

#### 4.1 Definitions

**Born alive 1878–1974:** In this period, the birth statistics in Spain used the definition of a live born child specified in the Civil Code: “a foetus will only be considered a live-born child if it has a human figure and has lived for twenty-four hours entirely detached from the mother's womb” (Art. 30 of Spain's Civil Code), while the rest of the births, considered miscarriages, were classified as follows:

- (i) Deaths during delivery (DC, *Muertos al nacer*).
- (ii) Live births, which died during the first 24 hours of life (D24, *Nacidos vivos, muertos antes de 24 horas*).

<sup>1</sup> Exact wording of the census question and English translation: “Sólo para mujeres que han tenido hijos: ¿Cuántos hijos nacidos con vida ha tenido? Incluya los ya fallecidos.” [Only for women who had children. How many live-born children have you had? Include the ones who died.]

(iii) Stillbirths (ND, *Nacidos muertos*).

These two sub-groups—deaths during delivery and the births that did not survive past the first 24 hours of life—were in fact live births, or “false stillbirths”. These counts were neither included in registered births nor in the death statistics. Consequently, there was an underestimation of the total of births and deaths until the last important reform of the civil registration system in 1975. To correct this underestimation, deaths during delivery and deaths before the first 24 hours of life have to be subtracted from the statistics of miscarriages and stillbirths and added to the statistics of births (Cusidó, 2008). This correction has to be applied to annual counts as well as monthly ones and to births by age of mothers. As miscarriage data were not published in necessary detail, we had to perform estimates of early deaths (see details in Appendix 2).

**Born alive since 1975:** In 1975, the United Nations definition regarding the live-born concept was adopted: “Live birth is the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of pregnancy, which after such separation, breathes or shows any other evidence of life, such as beating of the heart, pulsation of the umbilical cord or definite movement of voluntary muscles, whether or not the umbilical cord has been cut or the placenta is attached; each product of such a birth is considered liveborn (all live-born infants should be registered and counted as such, irrespective of gestational age or whether alive or dead at the time of registration, and if they die at any time following birth they should also be registered and counted as deaths).” See

<http://unstats.un.org/unsd/demographic/sconcerns/natality/natmethods.htm>

**Miscarriage 1878–1974:** In this period, miscarriages included the stillborn, deaths during delivery, deaths during the first 24 hours of life, and all those declared foetal deaths regardless of the time of gestation.

**Miscarriage since 1975:** Since this year, miscarriages have been defined as the “expulsion or extraction from its mother of a product of conception that was not viable”.

## 4.2 Data Quality Issues

This section discusses the data quality issues as related to Spanish fertility data in general as well as the fertility data included in the HFD, namely:

- Data problems during the pre-1975 era (section 4.2.1 and Appendix 2).
- Under-registration of births in the province of Barcelona in the early 1980s (section 4.2.2).
- Incorrect classification of birth order, affecting exclusively first and second birth order, in various years and various provinces (section 4.2.3). This data quality issue has been corrected using the method described in Appendix 3.
- Incorrect classification of the birth order for multiple births, corrected using the information from micro-data (section 4.2.4).
- Excessive classification of births of first birth order (section 4.2.5).
- Incorrect classification of births by year of mother's birth in official statistics after 1975 (section 4.2.6).

### 4.2.1 Data problems during the pre-1975 era

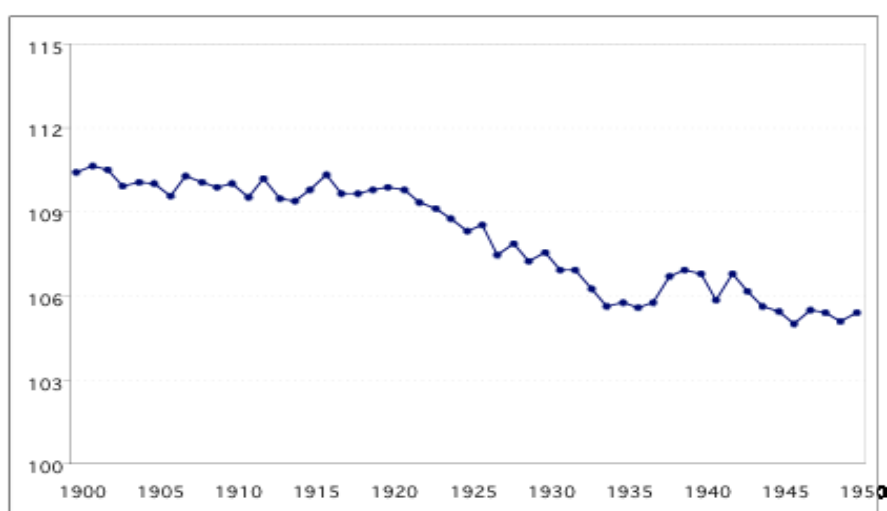
At the beginning of the Civil Register in the 19th century, the vital statistics did not cover a considerable amount of births (Ferrer, 1952), and until the beginning of the 20th century there were serious doubts by contemporaries as to the Register's reliability (Reher and Valero, 1995, p.85).

Due to the 1870 law which entrusted the management of the Civil Register to municipal judges, not to a group of specialists, registering problems soon emerged. The register suffered from under-

registration of events since the beginning. Even the authorities in charge of the official statistics in those years warned that the data from the Civil Register in its first years did not reflect the same number of events as those registered by the clergy (Instituto Geográfico y Estadístico, 1888: p. 11).

Leasure (1962) estimated the undercount of births in 1900 to be 11%. In 1905 the Geographical and Statistical Institute estimated that omitted births in the Civil Register made up approximately 6% of the total births (Instituto Geográfico y Estadístico, 1905). Livi-Bacci (1968 and 1993) and Reher and Valero (1995) also confirmed this under-registration of births in Spain until the beginning of the 20th century, based on the reported sex ratio at birth (see Figure 1). Despite certain improvements in the completeness of birth records during the 20th century, there were problems in data collection during the Civil War and the immediate post-war period (Villar, 1942; Consejo del Servicio Estadístico, 1943: Introducción). After these periods, the data relative to births became much more reliable (Livi-Bacci, 1968 and 1993; Nicolau, 2005; Reher and Valero, 1995).

**Figure 1** Evolution of the sex ratio (number of live born boys per live born girls) at birth in Spain during the period 1900–1950



Source: Author's calculations based on data from the Vital Statistics 1900–1950.

The percentage of “*No consta*” (Not stated) related to the mother's age presented a decline of 3.6% to 2.0% between 1922 and 1931, at which point it started increasing again to 4.4% in 1939. From 1940 to 1974 there was a drop in the share of births with “not stated” age of the mother from 3.9% to 0.9%. Since 1975 this category has no longer been included in the birth statistics.

On the other hand, underreporting of deaths during childbirth and on the first day must have been extensive during the first half of the 20th century. In the 1950s, the introduction of measures aiming to standardise the registration forms and questionnaires of the civil register resulted in a significant increase in reported early deaths (Blanes, 2007).

The re-estimation of numbers of live births from the numbers of “false stillbirths” is explained in Appendix 2.

#### **4.2.2 Under-registration of births in the early 1980s**

Even if the systematic under-registration of births seems to have almost disappeared in the 1950s, it is well known that there were cases of episodic sub-registration in the municipal birth series in some provinces, especially in the 1980s. After revising the birth series of the municipalities with

more than 10,000 residents in the period since 1975, one can come to the conclusion that this problem occurred in dozens of municipalities. Although under-registration was not general and was confined to only a few locations, its effects are notable, especially for the province of Barcelona during the first part of the 1980s and for the province of Guipúzcoa at the end of the 1970s.

A validation survey of the Vital Statistics registers carried out by the National Statistics Institute in 1987 (INE, 1987) gave clear proof of the birth coverage deficit across the country in the late 1970s and the early 1980s. This deficit was more sizeable for deaths than for births. The birth series indicate significant under-registration in selected provinces, especially in the early 1980s but also in the subsequent decades. The drop in under-registration typically occurs in short periods of one or two years.

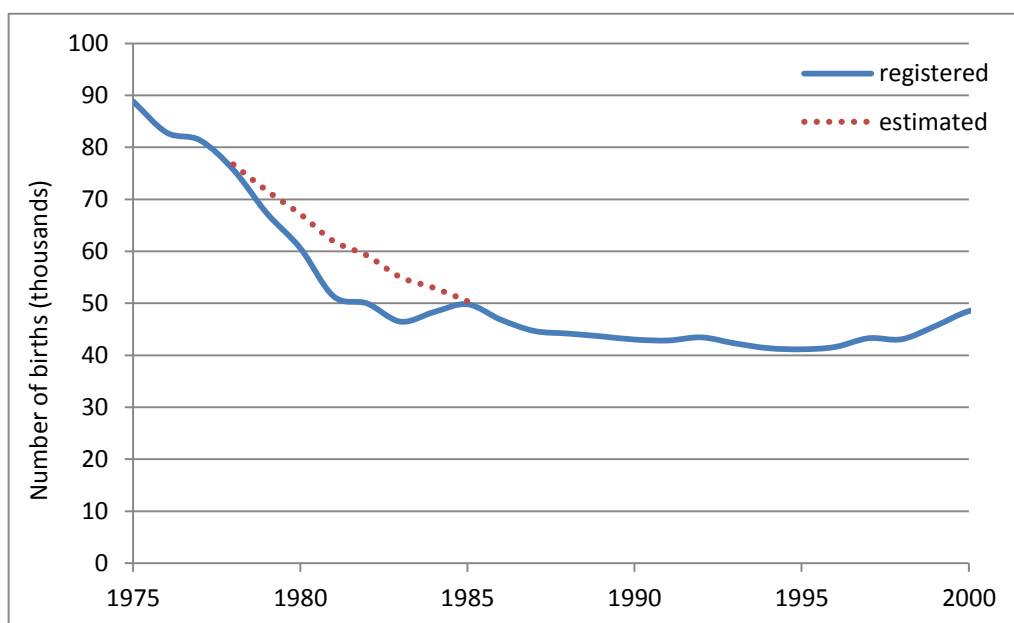
Although under-registration has been detected in many provinces, in only few of them a considerable number of municipalities is affected, including in particular Barcelona, Guipúzcoa and, to a smaller extent, La Coruña, Huelva, Madrid, and Santa Cruz de Tenerife.

The most serious problem is found in the province of Barcelona, where we identified 15 municipalities with more than 10,000 citizens that suffered sizeable under-registration in the 1980–1985 period. This problem notably affected the birth statistics at the province level. Overall, a general problem of the birth coverage deficit in Vital Statistics can be detected in the whole region of Catalonia. A 2007 study by the Centre d'Estudis Demogràfics (Devolder and Treviño, 2007) pointed out that this problem had large territorial variability and took place for a longer period of time than had been identified earlier in the study published by INE (1987).

Figure 2 shows the impact of the birth registration deficit in the province of Barcelona, 1981–1983, on the overall number of births and on the level of fertility indicators. The figure compares the number of registered births with the estimation based on the trend of proportions of births in Barcelona province among total births in Spain, which had been declining from 12.2% in 1976 to 10.6% in 1989. The estimated undercount in the Barcelona province was maximum 8–10 thousand births yearly from 1981–1983, which is around 2% of the total births in Spain and would mean an underestimation of TFR of 0.03–0.04.

No correction has been performed to adjust for the under-registration of births reported for the early 1980s. Users are therefore warned that they must use these data with caution. For further details regarding this issue, see Devolder and Treviño 2007.

**Figure 2** Births registered in province of Barcelona, 1975–2000, and estimated births accounting for under-registration, 1978–1985

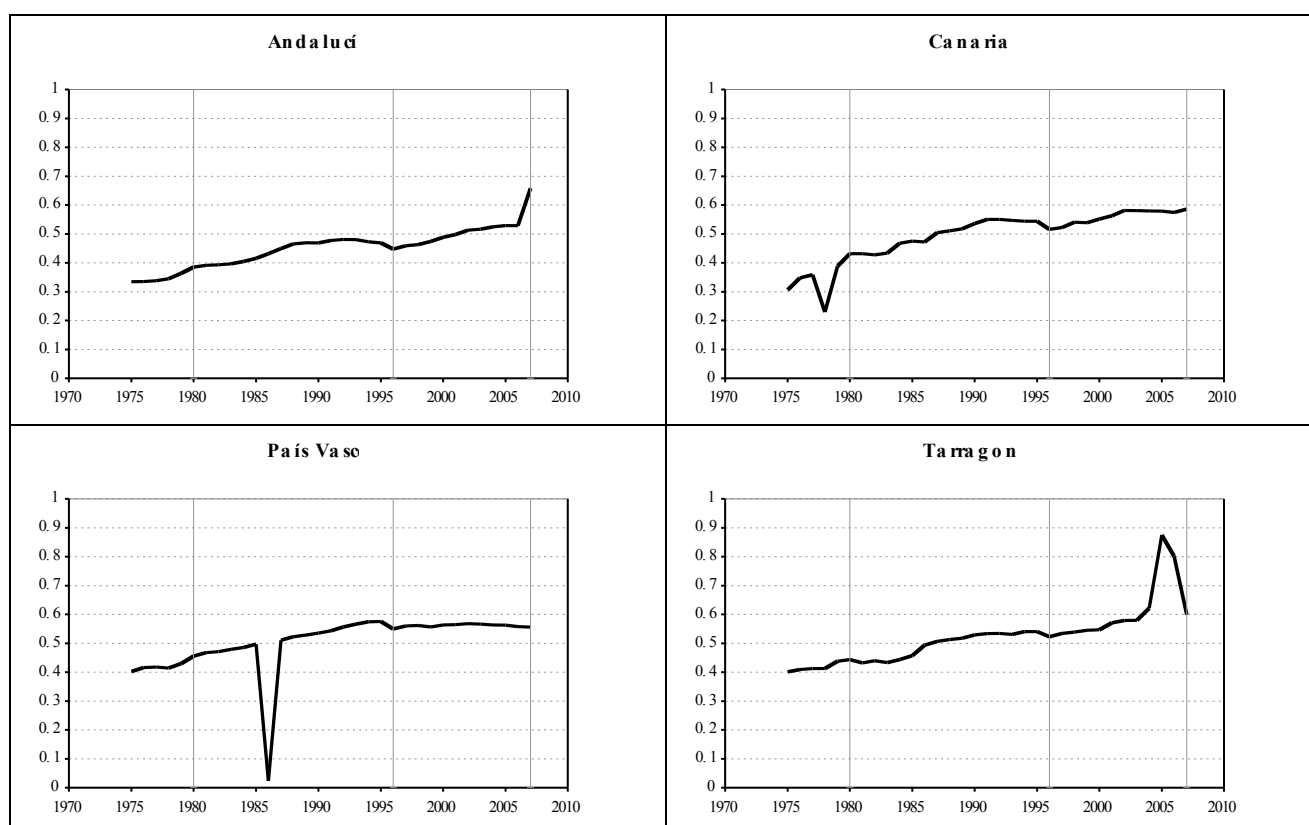


#### 4.2.3 *Incorrect classification of the birth order*

A simple graphic analysis (Figure 3) of the evolution of the birth series by birth order at the regional level helps detect many cases that appear as an incorrect allocation of the birth order, typically in short periods of one calendar year. These problems usually pertain to the misclassification of first births. There are two specific kinds of problems: incorrect transfers from the first to second birth order, which manifest themselves by a temporary downward shift in the proportion of births at order 1 (e.g. Canary Islands 1978; Basque Country 1986), and upward peaks in births at order 1 due to (incorrect) transfers from second birth order (e.g. Andalusia 2007; Tarragona 2005–2006). The problem is limited to the birth order classification, so the totals of births are not affected. For the years and provinces with clear misclassification of births of first and second birth order, the numbers were corrected using the method described in Appendix 3:

- Year 1978: both provinces of the Canary Islands
- Year 1979: Balearic Islands
- Year 1986: all three provinces of Basque Country
- Years 2005–2006: Tarragona
- Year 2007: all eight provinces of Andalusia

**Figure 3** Proportion of first births in selected regions (Andalusia, Canary Islands, Basque Country, Tarragona)



Note: The birth order classification has already been corrected for the incorrect allocation of birth order to multiple childbirths (see Section 4.2.3 and Appendix 3). The vertical lines signal the introduction of a new childbirth form: on January 1st of 1980, 1996, and 2007.

#### 4.2.4 Incorrect classification of the birth order for multiple births

Another relatively common problem in the microdata from the Vital Statistics for the period since 1975 pertains to the incorrect allocation of the birth order for multiple births. There is an easy and satisfactory solution to this problem if one exploits the information from the microdata files of the childbirths and carries out a new classification, using the correct procedure of giving a different order to each birth coming from the same delivery.

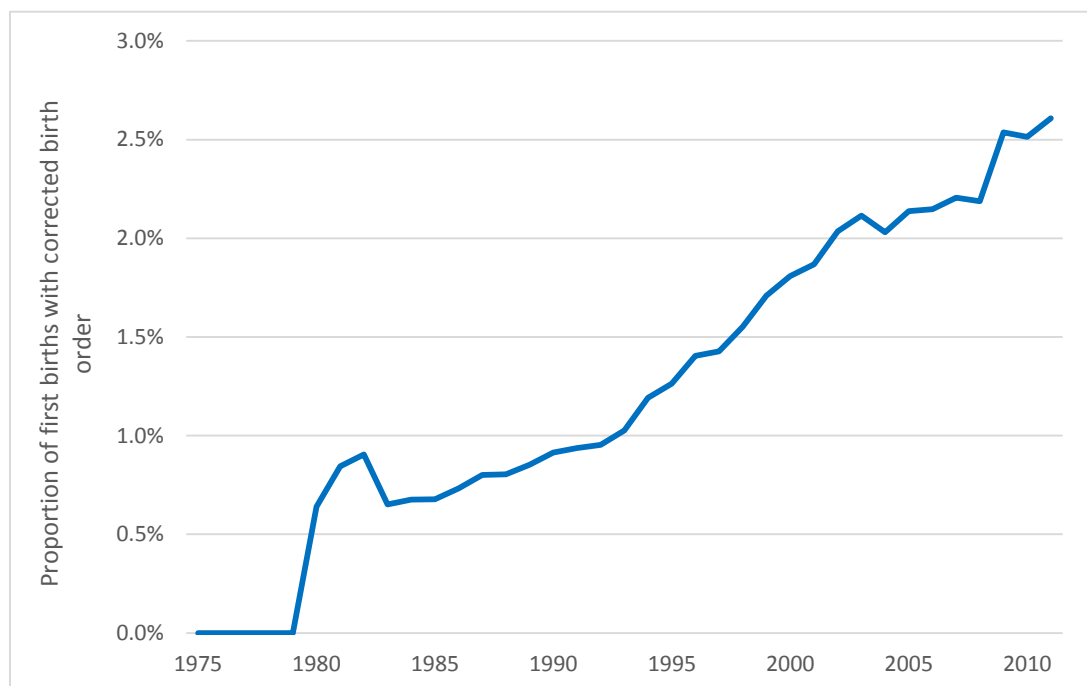
Until 1979, the registration form was filled out separately for each birth, and births from the same delivery were assigned the birth order subsequent to the count of previous live births. Since 1980 there is one registration form per delivery, and multiple births were assigned the order equal to the total number of live births, including those from the actual delivery. For completeness, we include the questions on birth order and their English translation in Appendix 4.

Another observation is that the proportion of births where the birth order needed to be corrected increases progressively over time (see Figure 4), mainly due to the use of assisted reproduction technologies, but also to the postponement of childbearing, which contributes to the rise in the number of multiple deliveries (as the natural risk of multiple births is higher after age 30).

As of 2007 INE has adopted a new procedure for the classification of multiple births by birth. Unfortunately, this change has not been documented, and the past series has not been modified. This partly explains the increase in the proportion of births with corrected birth order (Figure 4). Another cause of this increase is the misclassification problem observed in the region of Andalusia in 2007 (as documented in Figure 3).



**Figure 4** Proportion of first births where the birth order had to be corrected, Spain 1975–2011



#### **4.2.5 Excessive classification of births at first order**

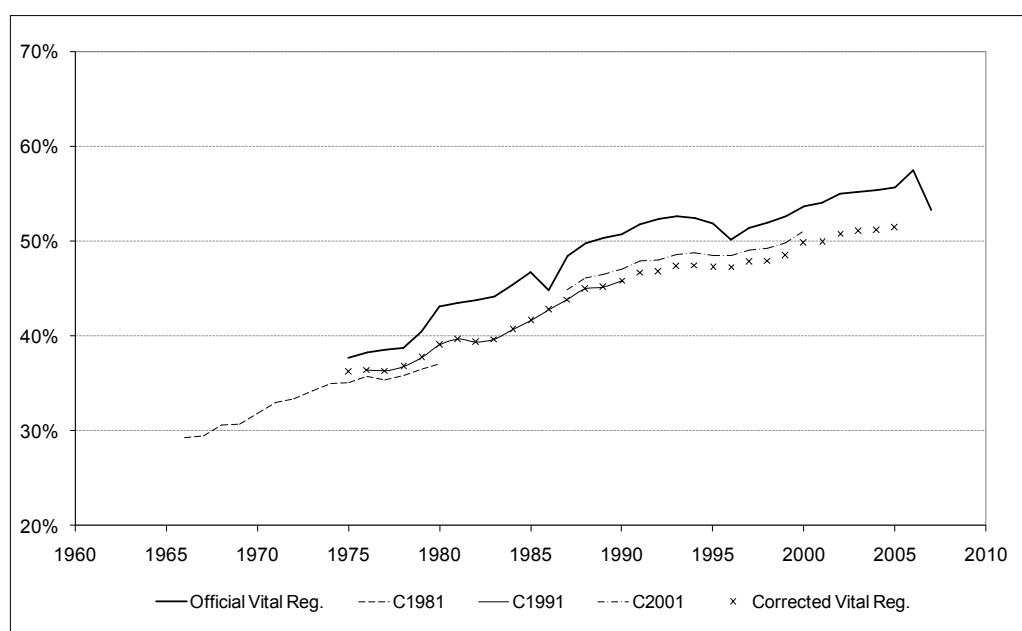
This problem has more serious impact on data quality. It concerns the incomplete or incorrect birth order on the childbirth forms. The excessive imputation of first births becomes apparent when cohort total fertility at birth order 1 is estimated for the Spanish regions. The problem is most acute for the region of Catalonia, where the estimated level of total fertility at birth order 1 is too high for the cohorts of women born in 1955–1958, with 98% to 101% of women estimated as having at least one child. The problem regarding inflated assignment of births of birth order 1 can also be observed for the cohorts born between 1953 and 1965 in the Balearic Islands, where the estimated levels of cohort first-birth rates also exceed an impossible value of more than one child per woman. The same pattern is observed for the region of Murcia for the cohorts born in 1953, 1958, 1959, and 1960; also in many birth cohorts of Ceuta and Melilla, or in Madrid in 1953 and 1954. A further examination of these estimates reveals that the completed fertility level for the first birth order reaches implausible values close to 1 in numerous cases.

In order to study the trend in time of the excess of births at birth order 1 in the Vital Statistics, we reconstructed birth histories of women from census data using the “own children method” (Devolder et al., 2008). We obtained the annual series of the distribution of births by birth order, which we used to estimate the degree of error of the same distribution obtained from the vital registration system. For example, Figure 5 compares the annual proportion of first births obtained from three censuses and from the vital registration system. The data series obtained from different census data are similar, and in all cases much lower than the corresponding values from the vital registration system. As we consider the census data to be of high quality, we used them to derive a “correct” proportion of first births, taking as a basis the proportion obtained from the 1991 Census, which we consider to be most comprehensive and accurate for this purpose. The “correct” proportion derived from the 1991 Census is then extended with the help of the 1981 and 2001 censuses allowing us to expand the estimated time series of the “corrected” proportion of first births over a longer period of time. We then extrapolate the time series up to 2006, assuming that the relative difference between the “correct” proportion of first births and the proportion observed

from vital statistics for the year 2000 remained constant thereafter. The results of this exercise can be summarised as follows:

- The excess of first births, due to transfers to birth order 1 of births of birth orders 2 or higher, is persistent over time, beginning with the reform of the Vital Statistics system in 1975.
- The bias was strongest during the period 1980–1995, when approximately 8–10% of births classified as birth order 1 were in fact births of birth order 2 or higher.
- Before 1980 and after 1995 the incorrectly classified “excess” births of birth order 1 represented around 5% of the total of births of this birth order.
- The introduction of a new childbirth form in 2007 represented an occasion to eliminate incorrect birth order reporting and reduce the excess of births of order 1. Unfortunately, the reported proportion of first births was clearly wrong again in that year, due to data problems in the Andalusia region. But the proportion of first births in year 2008 and from then onward suggests that birth order reporting had improved, as its level went down to that reported for 1998, and closer to the values we deduced from the census data.

**Figure 5** Proportion of first births among the total number of births: official and corrected values from the Vital Registration System and proportion calculated from the 1981, 1991, and 2001 Census data, by birth cohorts, for Spain



Note: The proportion of first births derived from the vital registration system statistics has been corrected to take account of multiple births (see Section 4.2.4).

Age-specific and total fertility rates by birth order are among the most important fertility indicators computed for the HFD. Thus it may be useful to discuss briefly what could be the consequences of the excess in the reported proportion of first births analysed above. The consequences are not the same for the different types of rates that can be derived from the birth data:

- The effects on the conventional indicator of Period Total Fertility Rate (TFR) are limited to the level of this index for each birth order. The biased reporting of birth order—resulting in a surplus of births of order 1, and a deficit of births at birth order 2 and higher—directly affects the estimated level of the TFR for each order.
- The level of mean age at childbearing is also affected, but only for birth order 1: as there is a transfer of births of birth order 2 and higher (which take place at later ages) to birth order 1, the estimated mean age at first birth will be higher than the correct value. The mean age

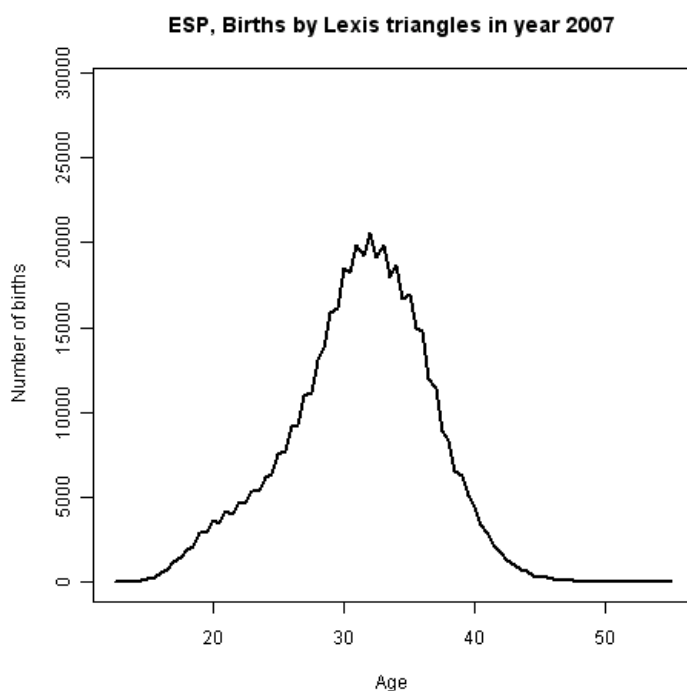
derived for higher birth orders should not be significantly affected, however, unless births wrongly classified as being of order 1 have a different distribution by age of mothers than the correctly classified ones.

- The effect of this problem on the Bongaarts and Feeney adjustment of the period TFR for tempo effect should be relatively limited. The misclassification problem is persistent over a long period of time and is therefore unlikely to cause sudden fluctuations in estimated mean ages at childbearing, to which the computation of tempo-adjusted TFR is particularly sensitive.
- The effects on the computation of fertility indexes based on occurrence–exposure rates (conditional rates) are more complex to analyse, as the birth order classification problem affects both the numerator and the denominator of the rates. Clearly, estimating the distribution of women by age and parity constitutes an additional problem, as the use of biased data with over-reported numbers of first births leads to implausibly low estimates of childlessness.

#### 4.2.6 *Incorrect classification of births by Lexis triangles*

In selected years, the age pattern of births by age and birth cohort of mother yields an implausible zig-zag pattern, suggesting that either age or cohort was not correctly stated for some births. This problem is especially evident at the beginning of the period covered by the birth order-specific data in the HFD (1975–1979) and in more recent years (2007–2012); see Figure 6. Therefore, we do not use the original classification by Lexis triangles for the entire period 1975–2014. Instead, we use only the distribution of births by age of mother, redistributing these into Lexis triangles using standard HFD methodology.

**Figure 6** Zig-zag pattern in births by age and cohort of mother, Spain 2007



## References

- Blanes Llorens, Amand (2007). *La mortalidad en la España del siglo XX. Análisis Demográfico y territorial.*- Tesis doctoral, dirigida por Anna Cabré Pla, leida en el Departamento de Geografía, Facultad de Filosofía i Letras, Universidad Autónoma de Barcelona.
- Consejo Del Servicio Estadístico (1943). *Movimiento Natural de la Población 1939: Introducción.* Dirección General de Estadística.
- Cusidó Vallverdú, Teresa (2008). Análisis crítica dels continguts del Movimiento Natural de la Población, 1858–1974: Història, normativa i conceptes relacionats. Trabajo de investigación dirigido por Albert Esteva i Palós, leida en el Departamento de Geografía, Facultad de Filosofía i Letras, Universidad Autónoma de Barcelona.
- Devolder, Daniel, and Treviño, Rocío (2007). *Informe sobre fuentes y metodología para el análisis de la fecundidad. La calidad de los datos en relación a la fecundidad: diagnóstico, documentación y propuestas de corrección.* Centre d'Estudis Demogràfics, Barcelona.
- Devolder, Daniel, Merino, Marta and Treviño, Rocío (2008). The use of Population Censuses to correct the distribution of births by order of the vital registration system. Application to Spanish data in the 1975–2005 period. EAPS European Population Conference, Barcelona.
- Ferrer Martín, Daniel (1952). *El Registro Civil y sus defectos. Compendio de los errores y defectos que pueden apreciarse y procedimiento para su subsanación.* Madrid.
- Glei, Dana A., Rosa Gómez Redondo, Antonio Argüeso, and Vladimir Canudas Romo (2016). *About Mortality Data for Spain.* Human Mortality Database Background and Documentation, last revised on 25 April 2016. <http://www.mortality.org/>
- Instituto Geográfico Y Estadístico (1888). *Reseña geográfica y estadística de España, 1888.* Dirección General del Instituto Geográfico y Estadístico.
- Instituto Geográfico Y Estadístico (1905). *Anuario Estadístico 1905.* Dirección General del Instituto Geográfico y Estadístico.
- Instituto Nacional De Estadística (1987). *Informe sobre la encuesta de validación del Movimiento Natural de la Población,* Madrid.
- Leasure, James W. (1962). *Factors involved in the decline of fertility in Spain, 1900–1950.* Michigan: Princeton University PhD, Ann Arbor Microfilms.
- Livi-Bacci, Massimo (1968). Fertility and nuptiality changes in Spain from the late 18th to the early 20th century, *Population Studies* 22(1): 83–102 and 22(2): 211–234.
- Livi-Bacci, Massimo (1993). *Introducción a la demografía.* Barcelona: Ariel (Ariel Historia).
- Nicolau, Roser (2005). Población, Salud y Actividad, in *Estadísticas Históricas de España*, Carreras, A. and Tafunell, X (eds), Bilbao, Fundación BBVA.
- Reher, David and Valero, Angeles (1995). *Fuentes de información demográfica en España*, Madrid, Centro de Investigaciones Sociológicas.
- Villar Salinas, Jesús (1942). *Repercusiones demográficas de la última guerra civil española, problemas que plantean y soluciones posibles.* Madrid: Memoria premiada por la Real Academia de Ciencias Morales y Políticas.

## APPENDIX 1

### DESCRIPTION OF DATA USED FOR LEXIS DATABASE

#### **BIRTHS**

Period	Type of data	Age range	Birth order	RefCode(s)
1922–1974	Annual number of live births by age of mother (Lexis squares)	–15, 15–19, ..., 45–49, 50+, UNK	–	1, 2
1975–2014	Annual number of live births by age of mother and birth order (Lexis squares)	12, ..., 55	1, 2, 3, 4, 5+	3, 6
1900–2014	Annual number of live births by month	–	–	1, 2, 3, 6

#### **FEMALE POPULATION: Distribution by age and parity**

Period	Type of data	Age range	Year of birth, range	Parity	RefCode	Notes
1.3.1991	Number of women by year of birth and parity	—	1870, ..., 1991	0, ..., 8+	4	Golden census; No unknown parity
1.11.2011	Number of women by year of birth and parity	—	1900, ..., 2011	0, ..., 8+	5	No unknown parity

#### **FEMALE POPULATION: Exposure by age and year of birth**

Female exposure population by calendar year, age, and year of birth (Lexis triangles) is estimated using data on population size and deaths from the Human Mortality Database, which is available at <http://www.mortality.org> or <http://www.humanmortality.de>.

## APPENDIX 2

### DESCRIPTION OF THE ESTIMATION PROCEDURES FOR THE PERIOD BEFORE 1975

The main complication with the data prior to 1975 is the fact that INE did not include infant deaths during childbirth and deaths during the first 24 hours of life into the count of live births. These infants' deaths were also not included with the other deaths, but instead were placed in a curious statistic category of 'miscarriages' that summed up stillbirths (ND), infant deaths during childbirth (DC), and live births who died during the first 24 hours of life (D24).

The dataset prepared for the HFD includes re-estimated data, obtained after estimating a number of infant deaths during childbirth and deaths during the first 24 hours of life (by age of mother; by month of birth), and adding them to the available number of live births.

**Table A1** Available and estimated data on live births wrongly classified as miscarriages in Spanish Vital Statistics during the period 1900–1974

Period	Total of miscarriages			Deaths during childbirth and deaths during the first 24 hours of life		
	Yearly total	Monthly total	Age of mother	Yearly total	Monthly total	Age of mother
1900–1921	X	X		E	E	
1922–1931	X	X	X	E	E	E
1932–1933	X	X	X	X	X	E
1934–1940	X	X	E	X	X	E
1941–1950	X	E	X	X	E	E
1951–1953	X	E	E	X	E	E
1954–1974	X	E	X	X	E	E

Note: X indicates that the information is available and E indicates that data have to be estimated in order to correct the birth counts.

The availability of data is summarised in Table A1. Data on miscarriages (which are the total of the three categories: DC, D24 and ND)<sup>2</sup> are available on a monthly basis for the time period 1900–1940, but for the 1941–1974 period only yearly totals are published. Miscarriages sums are also detailed by age of mothers, but curiously only for the time periods of 1922–1933, 1941–1950, and 1954–1974. For the two sub-groups of miscarriages that in fact are live births (DC and D24), the data published are less detailed, and we dispose of monthly data only for the time period 1932–1940, and of yearly data for the 1941–1974 period. No data were published regarding the distribution by age of mother. Therefore, we used different methods of estimating the deaths during childbirth and deaths during the first 24 hours of life for different periods.

### Estimation of deaths during childbirth and deaths during the first 24 hours of life by year and month

**Period 1900–1931:** The monthly adjustment factors  $f_m$  are calculated as the ratio of infant deaths during childbirth and deaths during the first 24 hours of life among the total of miscarriages for each month  $m$  during the 1932–1940 to total number of miscarriages in specific month  $A_{t,m}$ :

$$f_m = \frac{\sum_{t=1932}^{1940} DC_{t,m} + D24_{t,m}}{\sum_{t=1932}^{1940} DC_{t,m} + D24_{t,m} + ND_{t,m}} = \frac{\sum_{t=1932}^{1940} DC_{t,m} + D24_{t,m}}{\sum_{t=1932}^{1940} A_{t,m}}$$

<sup>2</sup> DC: *Muertos al nacer* / Deaths during childbirth

D24: *Nacidos vivos, muertos antes de 24 horas* / Live births, and deaths during the first 24 hours of life

ND: *Nacidos muertos* / Stillbirths

The estimated number of infant deaths during childbirth and deaths during the first 24 hours of life in 1900–1931 is then computed as  $DC_{t,m} + D24_{t,m} = f_m \cdot A_{t,m}$  where  $A$  is the known number of miscarriages in given month and year.

**Period 1941–1974:** Because for this period the number of miscarriages  $A$  is known only for the whole year, and not for individual months, the monthly weights are computed from observed proportions of DC + D24 during the time period 1932–1940:

$$w_m = \frac{\sum_{t=1932}^{1940} DC_{t,m} + D24_{t,m}}{\sum_{t=1932}^{1940} DC_t + D24_t}$$

The estimated number of infant deaths during childbirth and deaths during the first 24 hours of life in 1941–1974 is then computed as  $DC_{t,m} + D24_{t,m} = w_m \cdot f_m \cdot A_t$  where  $A_t$  is the known number of miscarriages in a given year.

### **Estimation of deaths during childbirth and deaths during the first 24 hours of life by age of mother**

**Periods 1922–1933, 1941–1950, 1954–1974:** The number of infant deaths during childbirth and of deaths during the first 24 hours of life is estimated using the known distribution of miscarriages by age  $x$  and known total number of infant deaths during childbirth and deaths during the first 24 hours in given year as:

$$DC_t(x) + D24_t(x) = \frac{A_t(x)}{A_t} \cdot [DC_t + D24_t]$$

**Periods 1934–1940, 1951–1953:** As in these periods the distribution of miscarriages by age was not known, we use the average of age distribution in neighbouring years, 1933 and 1941 for the former period and 1950 and 1954 for the latter one. The age distribution of miscarriages was very stable around these periods. The rest of the estimation is similar to other periods.

## APPENDIX 3

### DESCRIPTION OF DATA AND THE ESTIMATION PROCEDURES FOR 1975–2014

For the years and provinces with clear misclassification of births of 1st and 2nd birth order, as discussed in section 4.2.3, the numbers were corrected using following method. The list of affected years  $t$  and provinces  $p$  is as follows:

- Year 1978: both provinces of the Canary Islands
- Year 1979: Balearic Islands
- Year 1986: all three provinces of Basque Country
- Years 2005–2006: Tarragona
- Year 2007: all eight provinces of Andalusia

The number of first births is estimated using the total number of births, and the estimated proportion of first births, for each age  $a$ . This proportion of first births is estimated as the interpolation of proportion of first births from the years preceding and following the affected period (when the proportion is considered correct). The number of second births is then subtracted from the observed number of births in first plus second birth order, so that the original totals of first+second births remain constant:

$$B_1^*(a, t, p) = B(a, t, p) \cdot \left[ \frac{B_1(a, t-1, p)}{B(a, t-1, p)} + \frac{B_1(a, t+1, p)}{B(a, t+1, p)} \right] / 2$$

$$B_2^*(a, t, p) = B_{1+2}(a, t, p) - B_1^*(a, t, p)$$



## APPENDIX 4

### QUESTIONS ASKED ABOUT NUMBER OF BIRTHS ON SPANISH BIRTH AND CHILDBIRTH FORMS

Spanish	English
<b>Boletín de nacimiento 1960–74</b>	<b>Birth form for the period 1960–74</b>
<p>Fecundidad: ¿Cuántos hijos nacidos vivos ha tenido, contando este parto (aunque sólo vivieran unos instantes)?.... De ellos, ¿cuántos viven?....</p> <p>¿Cuántos se malograron antes de nacer, sin llegar a vivir un solo instante?....</p> <p>Parto: si fue múltiple, dígame cuántos nacidos vivieron 24 horas o más (si fueran dos o más se llenará un Boletín para cada uno, y si hubiera además, algún nacido muerto o que viviera menos de 24 horas, se utilizarán para ellos los correspondientes Boletines de Aborto.)</p>	<p>Fertility: How many live-born children have you ever had, including this childbirth (even if they lived for only a few moments)?... Of those, how many are still alive...</p> <p>How many died before birth, without living for even a moment?....</p> <p>Childbirth: if this was a multiple birth, how many children born lived 24 hours or more? (if there are two or more, fill out a form for each one, and if there was a stillbirth or children who lived less than 24 hours, fill out an miscarriage form for them.</p>
<b>Boletín de nacimiento 1975–79</b>	<b>Birth form for the period 1975–79</b>
<p>– Clase de parto: simple, doble, triple, cuádruple o más (en caso de parto múltiple, se rellenará para cada nacido un Boletín Estadístico de Nacimiento, y para cada aborto un Boletín Estadístico de Aborto).</p> <p>– Número de hijos nacidos con vida que ha tenido, contando este parto (incluidos también los que sólo hayan vivido unos instantes).</p> <p>– El hijo anterior a éste nacido con vida, nació el día / mes / año 19..</p>	<p>– Kind of childbirth: single / double / triple / quadruple or more (in case of a multiple birth, fill out a birth statistical form for each birth and an abortion statistical form for each abortion).</p> <p>– Number of live-born children you have ever had, including this childbirth (including also those who lived only for a few moments).</p> <p>– The previous live birth before the current one was on day / month / year 19..</p>
<b>Boletín de parto 1980–95</b>	<b>Childbirth form for the period 1980–95</b>
<p>– Número de hijos nacidos con vida que ha tenido, contando este parto</p> <p>– El hijo anterior a este parto nacido con vida, nació el día / mes / año 19..</p>	<p>– Number of live-born children you have ever had, including this childbirth.</p> <p>– The previous live birth before this childbirth was on day / month / year 19..</p>
<b>Boletín de parto 1996–2006</b>	<b>Childbirth form for the period 1996–2006</b>
<p>– Número de hijos que ha tenido contando este parto ...</p> <p>– ¿Cuántos de ellos nacieron con vida? ...</p> <p>– El hijo anterior a este parto nacido con vida, nació el día .. del mes .. del año ....</p>	<p>– Number of children you have ever had, including this childbirth ...</p> <p>– How many of those were alive at birth? ...</p> <p>– The previous live birth before this childbirth was on day / month / of year.</p>
<b>Boletín de parto de 2007</b>	<b>Childbirth form of year 2007</b>
<p>– ¿Cuántos hijos/hijas ha tenido a lo largo de su vida y contando este parto? (Por favor, si alguno de ellos nació sin vida, inclúyalo también) ...</p> <p>– ¿Ha tenido algún hijo/hija (nacido vivo) en partos anteriores: Sí / No ¿Cuántos?</p> <p>– Datos del hijo o hija nacido con vida anterior a este parto, si lo hubiera: ¿En qué fecha nació? Día:... Mes:... Año:.... ¿Dónde nació? En España (Indique la provincia y el municipio donde nació) / En el extranjero (Indique el país)</p>	<p>– How many children have you ever had during your life, including this childbirth? (if any were dead at birth, please include them as well) ...</p> <p>– Have you had any children (born alive) in previous childbirths: Yes / No (how many?).</p> <p>– Information on a child born alive before this childbirth. What is his/her birth date) Day:... Month:... Year:... Where was he/she born? In Spain (indicate the province and the municipality of birth) / In a foreign country (indicate the country).</p>