LUND UNIVERSITY

Master Programme in Economic Development and Growth

Female Labour Supply and Fertility in Spain.
A Regional Analysis of Interdependencies

Alexandra Bussler
alexandra.bussler.971@student.lu.se

Abstract: In light of the pronounced regression in total fertility rates in Spain over the last 40 years and the permanence well below replacement fertility levels up until the present day, this thesis aims to add evidence to the on-going debate about the determinants of reproductive behaviour in this specific country. Given the presence of a significant heterogeneity in the pace of changes in and levels of fertility across the 19 autonomous regions, which can be related to differences in the value systems around the role of women in society, it seems of utmost importance to engage in an analysis that includes regional realities when trying to explain fertility behaviour. Drawing on cross-sectional information from the national census of 2011, including demographic, educational, socio-economic and geographic information of a nationally representative sample of Spanish women, this paper engages in a micro-level analysis of the relationship between personal traits and both the decision to enter motherhood as well as to progress to higher order births with a special focus on the regional context.

Keywords: *Female labour force participation, fertility, demography, gender equality, Spain, motherhood, two-step discrete choice models, lowest-low fertility*

EKHM52
Master Thesis, second year (15 credits ECTS)
June 2016
Supervisor: Faustine Perrin
Co-supervisor: Maria Fibaek
Examiner: Kerstin Enflo
Wordcount: 16,907

Note: Poster representing the mother and Spain's children by the nationalist humanitarian organisation 'Auxilio Social' under the General Franco during the Civil war.[1]

[1] Image source: http://losojosdehipatia.com.es/wp-content/uploads/cartelas.jpg

Acknowledgements

I would like to thank all the faculty, staff and students of the Lund University and the Economic History Department whom I have had the pleasure of meeting in the past year. Specifically, a big thank you to my thesis supervisor, co-supervisor and the readers for all their help.

First, to Faustine Perrin for her guidance, patience, insightful comments, critiques and enlightening discussions.

Another huge thank you to Maria Fibaek for her helpful comments, feedback and guidance in the process of writing.

Finally, many thanks to my family who made this all possible, to my partner David Martin for his endless support, and to Lena Pfeifer for her great help and support during the final weeks of this thesis.

Table of contents

List of figures **5**

List of tables **5**

Introduction **6**

I. Historic background: women and the spanish society **9**

A. CATHOLIC VALUES, TRADITIONALISM AND THE BASIS OF A FAMILISTIC-PATRIARCHAL SOCIETY IN MODERN SPAIN 9

B. THE 2ND REPUBLIC AND THE FRANCO REGIME: BACK AND FORTH OF ATTAINMENTS 10

C. TRANSITION TO DEMOCRACY AND CONTEMPORARY SPAIN 11

i. Female labour force participation, educational attainment and gender equality *12*

ii. Compatibility of work and motherhood on a rigid labour market *16*

II. Regional heterogeneity in spain **21**

A. FERTILITY BETWEEN 1977 AND 2014 22

B. FEMALE LABOUR FORCE PARTICIPATION BETWEEN 1977 AND 2014 23

C. VALUES AND CULTURAL FACTORS 25

III. Theoretical framework: women's economic activity and fertility **28**

A. LOWEST-LOW FERTILITY 28

B. LITERATURE REVIEW AND DIMENSIONS OF INTERACTION BETWEEN FLFP AND TFR 29

i. Value systems *31*

ii. Institutional dimension *31*

iii. Economic dimension *32*

iv. Personal dimension *33*

IV. Empirical strategy **34**

A. A TWO-STEP MODEL OF THE FERTILITY DECISION 34

B. DATA AND VARIABLES 35

i. Data *35*

ii. Explanatory variables *36*

iii. Dependent variable *38*

C. METHODOLOGY 38

i. First step: probability to become a mother *39*

ii. Second step: probability of subsequent births, conditional on already being mother *40*

D. DESCRIPTIVE STATISTICS 41

V. Empirical results **43**

A. FIRST STEP: PROBIT 43

i. Probit regression *43*

ii. Marginal effects *47*

B. SECOND STEP: ZERO-TRUNCATED POISSON (ZTP) 48

i. Ztp regression *48*

ii. Marginal effects *51*

VI. Conclusion **53**

References **56**

Annexes **61**

Appendices **71**

LIST OF FIGURES

Figure 1 FLFP per age group 13
Figure 2 Female population aged 25-64 by educational attainment 13
Figure 3 Fertility and mean age at first birth 18
Figure 4 Age specific fertility 19
Figure 5 Autonomous and macro regions Spain 2016 21
Figure 6 Fertility in 1977 and 2014 22
Figure 7 Female Labour Force Participation ages 25-54 in 1977 and 201425 24
Figure 8 Dimensions of interaction between FLFP and TFR 30
Figure 9 Average fertility per birth cohorts 41

LIST OF TABLES

Table 1 Occupied per type of occupation in 2014 14
Table 2 Unemployed by educational level in 2015 15
Table 3 Part time employment 2012 16
Table 4 Sample regional distribution of average number of children (< age 5) 42
Table 5 Probit regressions on binary outcome 44
Table 6 Marginal effects Probit 47
Table 7 Ztp regressions on positive fertility 50
Table 8 Marginal effects Ztp 52

Introduction

Spanish women have never been more successful outside of the household sphere than today. Having already trespassed men in the educational attainment, in quantity as well as quality, Spanish women have made an impressive entry in the labour market since the transition to democracy in the 1970s. Overall female labour force participation (FLFP) has nearly doubled between 1976 and 2014, namely from 28.33 to 53.6 per cent (Figure 1). Even if this number is still relatively low in comparison to the European average and to men's activity levels, one can clearly see an increasing inherent trend: in 2015, women aged 25-34 are the group with the highest FLFP, followed in decreasing manner by the 35-39 year old and the 25-29 year old. The lowest participation can be found among women aged 55-64 and 15-19 year old (Annex 1).[2]

At the same time, fertility levels in Spain have experienced a sharp decline into a "regime of low fertility" (White et al., 2007, p. 3). Between the mid-1970s and 1996, fertility dropped from 2.77 to 1.15 children per women, the historically lowest fertility level of Spain, meaning that women have reduced their fertility by more than one and a half children in a timespan of only 20 years (Annex 13). The TFR has since then not reached to recuperate levels closer to replacement fertility: while there was a slight recovery of fertility over the 2000s, the aftermaths of the financial crisis have pushed down fertility again below "*lowest-low* levels"[3]. These levels well below replacement fertility (2.1 children per women) follow a trend experienced by a large number of industrialised countries over the last decades (Matysiak, 2011).[4]

Persisting extremely low fertility levels are assessed to lead to, for instance, population aging and a decreasing share of active population, both putting pressure on the functioning of the social welfare state (Bloom and Sousa-Poza, 2010). In light of these risks, it is important to understand what influences fertility to stay that low, and especially the reasons for regional differences, in order to revert *lowest-low* fertility trends.

Spain's persisting low fertility remains, till date, puzzling to the existing demographic literature. In 2015, Spanish women gave birth to slightly over 1.3 children on average. However, Esping-Andersen (2013) found out that the desired number of children per person in Spain is 2, for which we can suspect that there exists a discrepancy between the amount of children a women (or a couple) has, and the one she (or they) *would like* to have. One factor that has been put into relation with the latter is the evolvement of a sort of incompatibility between the increasing female labour activity on one side, and the desire of motherhood on the other side during the last decades (see Ibañez, 2010).

---

[2] The decreasing FLFP for age group 15-24 might due to the increasing commitment of women in tertiary education over the past 30 years. Students enter university after completing post-obligatory education taking place from age 16-18. Tertiary studies can then take on average 5 years.

[3] *Lowest-low* fertility is the concept of the TFR dropping below the by Kohler et al. (2002) arbitrarily chosen threshold of 1.3 children per woman. With a TFR of 1.3, a population would half in 45 years, all other things equal, such as migration. With a TFR of 1.6, it would take 90 years. The lower the TFR, the stronger the pressure on population dynamics and social welfare systems (Goldstein et al., 2009).

[4] In 2014, 20 countries had TFRs below 1.5, and 11 below 1.3 (Statista, 2016).

This raises the question whether there is evidence for a negative relationship between women's activity on the labour market and their decision to be mothers in Spain, a question that will be addressed in this thesis by trying to identify the underlying mechanisms.

However, by only looking at national figures one misses the large variability of the phenomenon on a more disaggregated level. Spain is composed of nineteen autonomous regions that vary considerably in their fertility and FLFP pattern and in how these numbers changed over time. Analysing the regional behaviour during the transition to very low fertility levels and higher FLFP can help to better understand the development of the total fertility rate (TFR). Special importance is thus given to the specific reality of each autonomous region. Spain has undergone various changes (e.g. economic, cultural, institutional) since the end of the Franco dictatorship in the 1970s; but the speed and direction of them vary greatly across the regions (Arpino and Tavares, 2013).

Among many possible approaches that try to explain the recent fertility behaviour of Spanish women, the present thesis will engage in a socio-economic and regional analysis of the probability to enter motherhood and progress in giving birth, with a special regard on the underlying cultural factors.

In the first part, we will analyse the dimensions of interdependence between TFR and FLFP. For this purpose, we will use aggregate data at national and regional level from the national statistical Institute from Spain (INE). The current regional *level* differences of the variables of interest furthermore call for an analysis addressing the regional specificities that might trigger the former.

Therefore, the second part of this paper engages in a micro-level analysis of the factors that influence the fertility choices of Spanish women in the actuality with a specific regional focus. Various studies have already tried to carve out the determinants of the choice of motherhood in Spain. However, the regional component has never been a central point in the literature so far. Given the regional heterogeneity of *values* (Arpino and Tavares, 2013), including a regional component in the analysis of the choice of motherhood is one of the contributions of this paper. This analysis draws upon individual-level information from the national census of 2011, including information on residence, economic status, education, number of children and the partners' information. The research question we pursued has two inherent dimensions:

i. What influences women to become a mother in the first place (in contrast to not having children at all)?

ii. What are the determinants of a mother's decision to continue giving birth to more children once she is already mother of one child?

These questions will be addressed with a two-step discrete choice model instead of standard statistical procedures, which represents another great virtue of this thesis.

The structure of the thesis is as follows. *Section I* introduces the history of the position of women in Spanish society, with a special focus on the last century and the current situation. *Section II* has a look upon the strongly pronounced regional variability of TFR and FLFP rates in the present, analysing the regional diversity in selected spheres, such as culture and traditions. *Section III* reviews the theoretical concepts that link fertility and female labour force participation. It includes a proposition of the different dimensions of how TFR and FLFP affect each other mutually as well as a literature review. This is followed by *section IV,* the empirical analysis presenting the data used and econometric model applied in the two-step analysis of the probability of motherhood with a special regional focus. *Section V* presents the empirical results. In *section VI* we discuss, conclude and open up ways for future investigation.

I. Historic background: Women and the Spanish society

A. Catholic values, traditionalism and the basis of a familistic-patriarchal society in modern Spain

In order to be able to understand the Spanish society, especially the position of women inside it, as well as inherent gender inequalities, one has to look back quite far in history. Many norms, traditions and cultural values related to gender that we can find in the Spain of today have developed over many centuries under the strong influence of traditional Catholicism.

First evidences for the separation of gender roles in the occidental society have been associated with the adoption of plough-agriculture in Western rural Europe. Due to men's advantage in applying upper-body force, they got specialised in the work with plough animals on the fields. Consequently, women had to specialise around the household activities, such as milling the grains and cooking, spinning yarn, feeding the children or washing clothes – their main duty being procreation to supply workforce for the family farm (Boserup, 1970; Andersen et al, 2016). Thus, occidental societies saw a development of the role of women inside the household that persisted even into modernity, when activities shifted away from agriculture to other productive sectors (Giuliano, 2014).

The separation of gender roles in the case of Spain has been reinforced during the modern époque within a Catholicism-dominated society. The completion of the Christian *Reconquista*[5] of the Spanish peninsula from the Islamic kingdom in 1492 represents the end of the medieval and the beginning of the modern époque for Spain. In Christianity, Eve, the representation of women in the Bible, stands for the committing of the ever-lasting sin, while the concept of the virgin Maria represents just the opposite: purity and love, which should serve as a role model for women in the value system of the traditional Catholic Church. These two opposed figures have served in Spanish Catholicism to create a societal role of women that condemned them to a life dedicated to the private sphere and motherhood, excluded from all public activities that remained reserved for men (Muñoz, 2006; Aixelà, 2003).

In the 17th century, the Catholic Church expanded its dominion via its influence in education, literature and from the pulpit. Spain remained a dominantly rural, poorly educated and poor society until the arrival of the Industrial Revolution at the end of the 19th century. In this setting, the Catholic institutions in Spain are claimed to have influenced the role division between men and women in form of the promotion of a patriarchal *familistic*[6] system (Cantero, 2007).

[5] The *Reconquista* refers to the nearly 800 years in which the Christian Iberian kingdoms reconquered the Spanish territory that had been conquered by the Islamic crown at the end in 718AC. The fall of the last Islamic bastion, Granada, under the recently united Spanish kingdom in 1492 marks the official ending of the *Reconquista.*

[6] A *familistic* system refers to a cultural system with a tight support between generations, such as that parents' influence on children, and their dependence (financially, emotionally) on them, is prolonged. Leaving the parental home is postponed and grandparents support parents in childcare (Ibañez, 2010).

In the patriarchal *familistic* system, women were not only excluded from public activities, but also from scholarly instruction until the beginning of 1800, when primary education opened up for girls. Yet, apart from literacy, girl's education aimed at teaching basic skills "in accordance to their gender", related to the household sphere (Fernández, 2006). It is not until the last quarter of the 19th century that women were allowed access to secondary education and even in 1910 for them to join universities – however in absurdly low numbers in the latter. Yet, over the years these numbers started to increase gradually (Flecha, 2006).

To conclude, in post-1492 Spain women remained in an inferior position compared to men until far into the advanced 20th century, strongly influenced by Catholic traditions. Furthermore, the Civil Code of 1889 created a basis of profound inequality between men and women via discriminatory propositions based on the supposed weakness of women in need of protection. Women's rights were compared to the ones of minors or incapacitated persons (Torralbo, 2011). Only from the early 1900 feminist movements started to question the position of women in Spain's society.

Nevertheless, the remains from a *familistic* and patriarchal societal system can still be perceived in the Spain of today, which the following section will try to explain.

B. The 2nd Republic and the Franco Regime: back and forth of attainments

From 1939, the almost 40 years long lasting Franco regime abolished many of the rights and achievements towards gender equality gained during the 2nd Republic of the early 1930s. One could claim that women during this period were pushed even further backwards than ever before in history.

The often referred to as *revolutionary* years of the 2nd Republic lasted from 1931 until the Civil war in 1936[7], after which the general Franco took over the state power by means of a military coup in 1939. The socialist and left wing orientated democratic government of the 2nd Republic embraced ambitious steps towards a modernization of the Catholicism-dominated traditionalist Spanish society with many far-reaching policies and legal reforms. It broke for instance for the first time in modern history the relationship between the state and the Catholic Church, with the 3rd Article of the Constitution of 1931 claiming that Spain had no official religion.[8] Furthermore, women reached various steps towards equality. They obtained the right to vote in 1931 and participated for the first time in the legislative elections of 1933, thanks to the decisive protagonism of one of the first two female members of parliament, Clara Campoamor. The participation of female workers in syndicates increased significantly (Núñez, 1993). In addition –

[7] The Spanish Civil War lasted until 1939, dividing country and its society into two ferociously opposed sides (the *old* and *modern* values of Spain), provoking around half a million deaths.
[8] The Constitution of the II. Republic of 1931 can be accessed at the wikisource internet portal.

and quite anachronistic – abortions and divorces for women became completely legalized under the 2nd Republic. In this vein, civilian marriage was introduced in 1932 as an alternative to the religious one (Vadillo, 2014; Lafuente, 2015).

Yet, with the beginning of the dictatorship women were pushed back into the role of wife and mother: article II.1 of the first Franco-law *Fuero del Trabajo* proclaimed in 1938 tied married women back to the household, 'liberating' them from all outside working duties in the fabrics or workshops introduced during Republican times.[9] Fertility became strongly prioritized and traditional national-catholic values that would discriminate and exclude women from the working place became again centre in Spanish society, divorces and abortions illegalised. Basic schooling around the inherent values of 'femininity' were designed for women, but once the domestic market opened up in the end 1950s, it was needed that women start to become more educated – but just enough so that they could satisfy the increasing demand for a higher-educated workforce in the nascent industries. From that point on, basic female enrolment rates increased continuously, although the higher educational spheres still remained an above all male-dominated world until the end of the dictatorship, as shown in the educational enrolment statistics of the historical statistical yearbooks.[10]

Between 1939 and 1975, women were again forced into a role of subordination, procreation and service to the society (De Lecea, 2006). The highest active female labour force participation during the Franco regime was reached towards its end, with a national average of participation around 29 per cent.[11]
After fertility had decreased to 2.5 by the mid-20th century, it raised again up to 2.8[12] children in 1975, the final year of the dictatorship. This can be explained by the return of traditional-catholic values as discussed above.

C. Transition to Democracy and Contemporary Spain

In present-day Spain, we find a stark contrast between the old traditional-Catholic heritage and the new and modern influences from the outside. With the transition to democracy after Franco's death in 1975, Spain experienced a period of rapid changes. One of the first steps under the new democratic government was the equalization of the rights between women and men and the formal abolition of discriminatory laws. Yet, it was only in 2007 that a Spanish government (PSOE[13]) took the initiative to pass a law aiming to promote the equality between men and women with a special focus on the labour market, recognizing the goal to enhance the compatibility of work and motherhood.[14]

[9] The complete legal text of the *Fuero del Trabajo* is accesseble at the wikisource internet portal.
[10] Those are retrievable on the webpage of the INE (National Statistical Institute).
[11] For tables consult COMISIONES OBRERAS (2004) p. 30 and INE-EPA for that year.
[12] Synthetic Fertility Index calculated with demographic data from the National Institute of Statistics (INE).
[13] PSOE is the acronym for *Partido Socialista Obrero de España* (Socialist Worker's Party of Spain), the Spanish centre-left party.
[14] Reference is made to the *Ley Orgánica 3/2007 de 22 de marzo, para la igualdad efectiva de mujeres y hombres* (Organic Law for the effective equality between women and men).

Turning away from a backwards orientated, Catholic-traditionalist society, Spain has managed to adopt ever more values which the European Union stands for: "Respect for human dignity, liberty, democracy, equality, the rule of law and respect for human rights, including the rights of persons belonging to minorities".[15] The influence of the Catholic Church on the public sphere has been incrementally constrained, for instance in the sphere of civil unions and divorces (under Franco impossible for women) or the legalization of abortions in 1985.[16] Yet, only in 2010 *voluntary* abortions become legalized.[17]

i. Female Labour Force Participation, Educational Attainment and Gender Equality

Keeping in mind the discriminatory treatment of women and their position in society under the Franco regime until as recently as 1975, it becomes evident that the transition to an egalitarian and modern society has not been an easy endeavour. Especially for women, that in some cases experienced up to three types of completely opposed periods (the 2nd Republic, the Franco Regime and the recent period of democracy), the process of adaptation remains complicated.
For a woman whose 'natural' role was to be a housewife and mother for most of her life, raised under a cultural regime where the discrimination of women was a common practice, probably unable to study anything relevant outside of the household sphere, it might have resulted quite difficult to enter the labour market in the democratic and fast-changing post-Franco Spain.

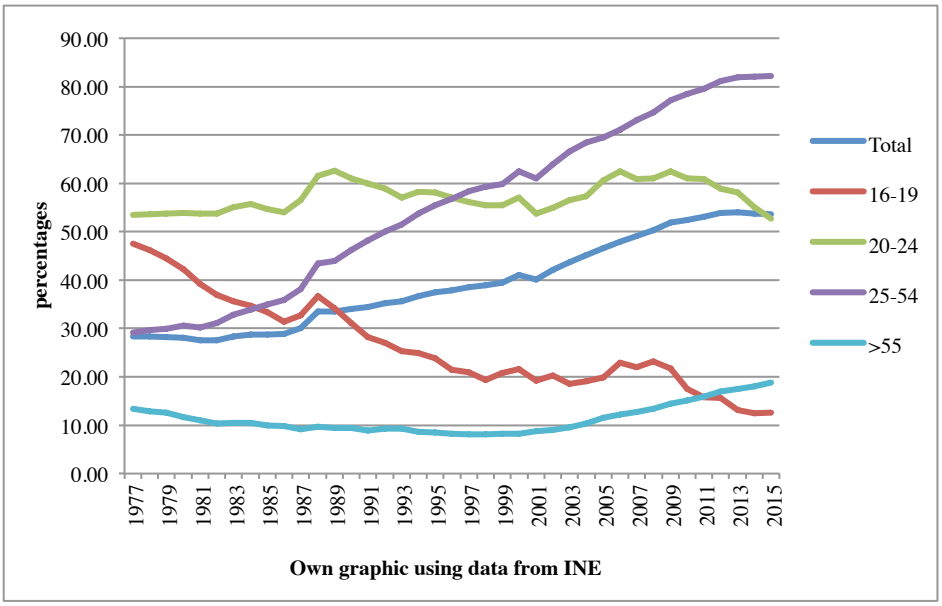
This reasoning can help to explain why the FLFP for the older female cohorts (Figure 1) remains low in 2015: only about 19 per cent of the women born before 1960 are economically active. During the first 30 years of democracy, the FLFP of women aged over 55 did not even reach 10 per cent. For those women born under Franco but who experienced their youth and early adulthood in democracy, FLFP is today already more than 30 per cent higher with respect to the former group.[18]

Interestingly, at the end of the dictatorship in 1975, the very young women (age 16-19, 20-24) had participation rates nearly double the national average. Of the 'main trunk' of women, aged 25 to 54, 82.14 per cent are currently economically active. This group has made a stunning development since the end of the dictatorship: a more than 50 percentage points increase from initially 30 per cent in 1977. One could conclude that the nowadays still quite low total FLFP in Spain is mainly driven by the low participation of the older cohorts and the decreasing FLFP of the 16-19 year-old, who most certainly are still enrolled in the educational system. Furthermore, the activity of the 20-24 old is today somewhat above 50 per cent and shows a decreasing pattern, which can be explained by the increasing female presence in the higher educational system.

---

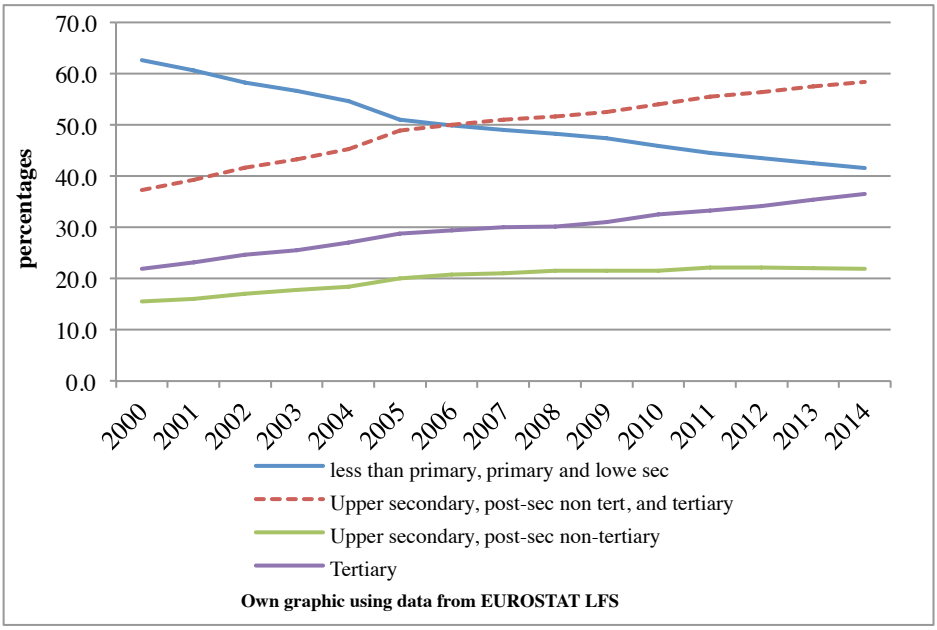
[15] Article I-2 of the 2004 Constitutional Treaty of the European Union.
[16] Reference is made to the *Ley Orgánica 9/1985* (Organic Law 9/1985). However, under this law abortions were only legal in specific medical or psychological cases.
[17] Reference is made to the *Ley Orgánica 2/2010 de salud sexual y reproductiva y de la interrupción voluntaria del embarazo* (Organic Law on Sexual and Reproductive Health and Voluntary Termination of Pregnancy).
[18] See Annex 1 for FLFP disaggregated by age using data from the EUROSTAT monthly labour force survey, which draws on representative extracts of the population. The levels differ thus from the national statistics. We trust the official INE information.

Figure 1: FLFP per age group



The latter can be seen in Figure 2. Since the turn of the century, the share of women with no or very low levels of educational attainment has decreased drastically, while all other levels have increased. Most notably is the tertiary education, which increased from 21.9 to 36.5 per cent between 2000 and 2014. The overall category *upper secondary, post-secondary non-tertiary and tertiary attainment* represents the majority of Spanish women in 2014, namely 58.4 per cent.[19] Its increase seems to be mainly driven by the tertiary attainments.

Figure 2: Female population aged 25-64 by educational attainment



[19] The blue, green and purple lines add up to 100, while the red dotted line is the sum of the green and purple lines.

If we furthermore compare the educational levels of men and women (Annex 2), we can see that in 2005 women have trespassed men in both upper secondary and post-secondary as well as tertiary educational levels. The trend seems to be further increasing.

Although the presence of women in the labour market is enhancing and their educational attainments higher than ever before, we are far from being able to state that they have attained a level of equal treatment and positioning to men in Spanish society. For instance, women are still overrepresented in certain types of occupation not requiring high levels of qualification. This can be seen in the gender ratio[20] greater than 1 in occupation 5 and 9 (Table 1). Also, more than twice as many women than men work in typical white-collar jobs (category 4). On the other hand, in the category *Technicians and scientific or intellectual professionals*, representing the second-highest qualification level, women are with 21.1 per cent much more present than men.
However, women are highly underrepresented in the highest category, *directors and managers*. The latter insinuates that in nowadays Spain there still exists discrimination of women in the labour market when it comes to occupational choice – or *occupational attainment*.

Table 1. Occupied per type of occupation in 2014

Type of occupation	*Men**	*Women**	*Gender ratio⁺*
1. Directors and managers	5.7	3	0.52
2. Technicians and scientific or intellectual professionals	14.5	21.1	1.44
3. Technicians and assistant professionals	12	8.9	0.7
4. Bookkeeping employees, administrative staff and other office jobs	6.3	14.7	2.35
5. Restoration, personal services or protection service workers or sales staff	17.7	29.8	1.67
6. Qualified workers in the agricultural, farming, forestry or fishery sectors	3.8	1	0.27
7. Artisans and qualified workers in the manufacturing and construction (except Installation and machinery operators)	18.4	1.9	0.1
8. Installation and machinery operators and assemblers	11.8	2	0.16
9. Elementary occupations	8.8	17.3	1.85
0. Military occupations	0.9	0.2	0.2
Total	*100*	*100*	*0.95*

Using data from INEbase

* in percentages, ⁺ *(# women per occupation / total FLFP)/(# men per occupation / total MLFP)*

Furthermore, looking at the average earnings women obtain in fulltime employment, we can find a quite significant wage-gap of 14.5 per cent in 2013. The average gross income of women is even only 76 per cent of the men's. Interestingly, the wage gap of older women is most pronounced with 21 per cent for women aged 45-54 compared to 12.2 per cent for the 25-34 year-old (INE,

[20] The gender ratio (values 0 to 1) refers to the ratio of women working in the specific sector with respect to the entire female active population, to the ratio of men working in this specific sector with respect to the entire male active population. Value 1 means an equal distribution of women and men in this sector. Values exceeding 1 indicate that relatively more women than men work in this sector (controlled for the gender-respective labour force participation).

2016). The latter raises the assumption that women are not only disadvantaged by their gender but also age.
This can be linked to one of the theories that have been used to explain why women earn less than men, the motherhood penalty: based on the fact that women become mothers at some point in time and have to leave their jobs, they lose work experience, which will affect their future wages. This has to do with the idea of working productivity in the basic Human Capital theory. The rationale is that the wage-gap grows with age, since the more time women have already spent out of the job (due to motherhood), the 'more losses in productivity' they have accumulated (Budig and England, 2001).

Table 2 contains information about gender-specific unemployment levels per educational attainment. Here again, we can see that women tend to be discriminated the higher their skills level is: at any of the three higher educational levels (5.-7.) the gender ratio of unemployment[21] is higher than 1. The same counts for the lowest educational category (1.). This means that with no education and higher educational attainment, there are relatively more unemployed women than men. Especially with tertiary education there are nearly double as many women as men unemployed (ratio 1.71). At lower-to-medium educational levels, however, it seems that men are suffering relatively more from unemployment than women (categories 2.-4.). Yet, the gender ratio of unemployment is higher for women on average (1.16).

Table 2. Unemployed by educational level in 2015

Type of Education	*Men**	*Women**	*Number of men°*	*Number of women°*	*Gender ratio⁺*
1. Analphabets	0.5	1.0	13	23	2.04
2. Primary education incompleted	3.4	2.1	80.6	50	0.72
3. Primary education completed	13.3	9.0	318.4	215.1	0.78
4. Lower secondary education	42.8	36.4	1 021.3	870	0.98
5. Higher secondary education	12.0	13.1	287.6	313.5	1.26
6. Secondary, post-secondary and non-tertiary education	9.7	11.4	230.6	273.3	1.37
7. Tertiary education	18.3	27.1	435.9	647.1	1.71
			2387.6	*2391.9*	*1.16*

Using data from INEbase

** in percentages, ° in thousands of persons, ⁺ (# women per edu. Level / total FLFP)/(# men per edu. Level / total MLFP)*

All the latter information leads us to conclude that women are still suffering from inequalities and discrimination in the present labour market. There are more unemployed women than men, women are less represented in occupations that require higher skill levels (even though women are today, on average, more educated than men) and receive significantly less pay than men.

[21] The gender ratio here follows the same logic and can be interpreted as explained in the previous footnote.

The fact that that there are more unemployed women than men can also be linked to an important feature of the Spanish labour market: low flexibility and the difficulty to combine parenthood and work. With this, we mean the ability for workers to reduce from full- to part time employment due to, for instance, family obligations such as caring for elderly or children. In 2012, a total of 1.975 million women were employed part time, which corresponds to only about 18.5 per cent of the female active population (Table 3). For men, it is even only about 5.2 per cent of the actives. If we compare these numbers to Sweden, one of the countries which seems currently to be one of the most equal countries with respect to the labour market, we can see that in 2013 more than 30 per cent of women and 11 per cent of men worked part time (SCB, 2014).

Furthermore, the reason for women to choose part time due to caring is in more than half of the cases (51.8 per cent) that childcare is either not available, or too expensive. In fact, Ibañez (2010) discussed that there exists a significant problem in the public childcare between age 0-3: there are not enough public places, and the private ones are simply too expensive for the majority of families. This is confirmed by the statistics in Table 3.

Hence, if the public childcare service is insufficient or inaccessible, women seem to majorly be the ones who enter part time work, which results in lower current earnings and future pension payments. Only 8,200 men in contrast to 263,000 women were part time employed due to caring for children or elderly. Some women even quit their jobs: 24,400 unemployed women in 2013 left their last job due to "childcare or elderly care", in contrast to only 2800 men (INE, 2016).

Table 3. Part time employment 2012

Reason for part time employment		
	*Women**	*Men**
Still studying or in professional formation	84.3	55.3
Own illness or incapacity	20.7	10.7
Childcare or care for elderly, incapacitated or sick	**263.1**	**8.2**
Other familiar obligations	118.6	7.4
Could not find a full time job	1,160.3	414.1
Do not want to work full time	140.5	30.5
Other motivations	183.5	110.7
Do not know the motivation	5.0	3.2
Total	***1,975.9***	***640.1***

Reason for part time employment due to care		
	% Women	*% Men*
Childcare non available or not affordable	51.8	38.8
Elderly or sick people's care non available or not affordable	4	7.8
Both previous reasons	3.2	-
Other reasons	40.3	53.4
Do not know	0.7	-
Total	***100***	***100***

Using data from INM

*in thousands of persons

Form these information we can induce that it still seems to be seen more of a women's duty to care for children in Spanish society. Thus, there are two major factors that still impede women's equality to men on the labour market: the impossibility of reconciliation of work and motherhood, and the cultural reality of unequal parenthood leave.

The latter two are related to institutional as well as cultural factors. The Spanish government is among those in the EU who spends least on family and children aid programs (Ibañez, 2010). The legislation concerning the conciliation of parenthood and work actually still seems to be driven by a discriminatory conception of motherhood, since it foresees the mother as principal carer of the children.

Paid maternity leave accounts for 16 weeks, of which ten weeks can be shared with the father. However, only in 1.8 per cent of the cases women shared parental leave with their partners in 2014. Furthermore, only since 2007 the *Ley por la Igualdad* gives fathers the right to 13 unpaid days of independent paternity leave. In 2014, only 235,678 men enjoyed paternity leave for 427,595 children born – which might be related to the unpaid character of the leave.
Interestingly, the government has been constantly postponing the implementation of a law to increase paternity leave to one month that was ratified in 2009 (El Diario, 2016).

During the paid maternity/paternity leave, the parent receives 100 per cent of the basis of assessment of his or her respective last salary. The fact that those who take the leave are predominantly women can therefore not be attributed to the economic rationale of the man having a higher salary (as seen above), for which it would be logic that the woman leaves the job to maximize household income. It is rather explainable by a persisting cultural conception of the *women-carer* in Spain.

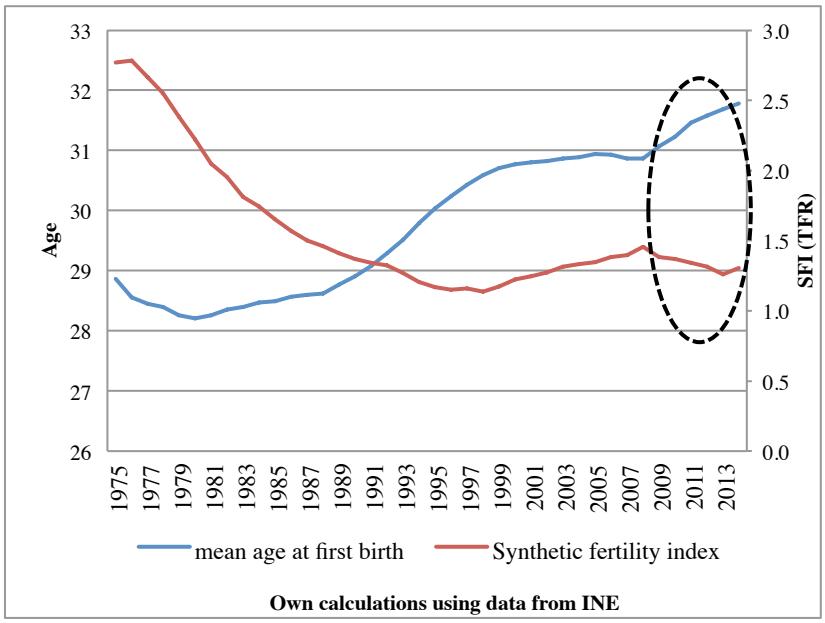
Additionally, in order to receive the motherhood benefits, a woman must have previously been contributing to the Social Security. If this prerequisite is not met, she is in 2016 entitled to receive a monthly subsidy of 532.51€ during the immediate 6 weeks after giving birth. It would actually result more convenient to be registered as unemployed during motherhood, since unemployment benefits are not limited to 42 days. This can help to explain the higher entries into unemployment due to parenthood for mothers.

As we can infer from the facts of this section, it seems that there still exists a cultural and institutional conception and persistence of gender discrimination. Especially when it comes to the sensitive case of parenthood, mothers are still the ones that majorly suffer a loss in order to care for the children: institutional fails such as insufficient public childcare or discriminatory and dis-incentivising parenthood-legislation make it harder for women to be able to combine motherhood and work.

Given their continuing advancement in educational attainment and presence on the labour market, the younger generations of women are not willing anymore to follow their mothers or grandmothers in going back to being housewives and exiting the labour market when they

become mothers. The current hurdles for women to combine motherhood and work, and the lacking/insufficient participation of men in the task of childcare, have resulted in a striking development of fertility in Spain over the last decades: TFR has decreased to a *lowest-low* level, and has until date not managed to increase again towards replacement fertility. We calculated that the synthetic fertility index (SFI[22]) for Spain was 1.31 children per women in 2014, meaning that a woman in 2014 was expected to give birth to 1.31 children during her whole lifetime (Figure 3).

Figure 3. Fertility and mean age at first birth



While at the end of the Franco dictatorship women were expected to have about 2.8 children during their lifetime, in the beginning of the nineties this number had dropped to less than 1.3 children.

After nearly ten years on *lowest-low* levels, the TFR increased again during the 2000s to reach its 21st century high of 1.46 in the year the financial crisis hit Europe (2008), after which it decreased again below 1.3. Today, the TFR has recovered again above 1.3 – even though it is questionable to call a TFR that is more than 0.7 percentage points below replacement fertility[23] a 'recovery'. It is expected to further increase towards 1.6 until 2019 (Annex 4).

[22] The synthetic fertility index (SFI), or total fertility rate (TFR), is a measure based on synthetic fertility: it adds up age-specific fertility for women aged 15-49 at a certain point in time. It indicates how many children a 'synthetic women' in a specific year is expected to have in her whole reproductive lifetime, including the information of all cohort fertilities.

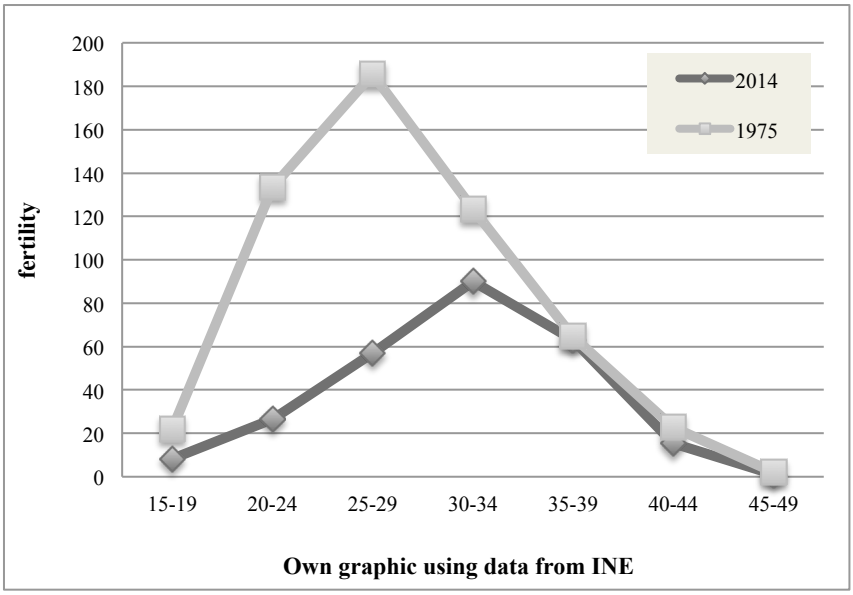
[23] *Replacement fertility* refers to a TFR of less than 2.1 children per woman. The continued fall of developing countries under replacement fertility has increased the concept's presence in demographic literature (Smallwood and Chamberlain, 2005).

Another development that has been assessed to influence a temporary decrease of period fertility rates is the increase of mean age at first birth, particularly in the case of Spain. The mean age increased from around 28.5 to nearly 31 years during the same time that the TFR dropped to *lowest-low* levels. Bongaarts and Feeney (1998) argue that these *lowest-low* levels will only remain until the postponement period of age at first birth has levelled off. In fact, once the 'postponement of first birth' effect had finished in the 2000s, the TFR began to increase again. Since 2008, we can see anew a period of increase in the mean age at first birth paired with decreasing TFR (the dashed oval in Figure 3).

However, the synthetic fertility measure applied in this paper with the demographic data from INE represents a synthetic but "consistent projection of the level, timing and distribution of the completed fertility" (Kohler and Ortega, 2002, p.126). The renewed decrease in TFR might be thus, at least partially, triggered by a decrease in the amount of children women have, and not the postponement of birth.

A look at the age specific fertility rates in Spain for the years 1975 and 2014 yields the insight that also the cohort-fertility behaviour has changed drastically during the transition to very low fertility (Figure 4).

Figure 4. Age specific fertility



While in 1975 fertility was highest among women aged 25-29, followed by 20-24, in 2014 it was highest for the 30-34 year old, followed by 35-39. This represents an important shift of fertility towards higher ages. What we can further see is a drop in the level of cohort fertility: the highest cohort fertility in 2014 was 90.07, representing half of the highest cohort fertility in 1975 (185.59).

In the context of an increasing number of highly educated women that want to – and do already – participate on the labour market, entering motherhood seems to have become less attractive in

the last decades. Women on average have chosen to have fewer children, and have them at ever-higher ages. Yet, this does not mean they do no *desire* to have more (Esping-Andersen, 2013).

Yet, from this last section we can learn that there seems to be a negative relationship between working and women's choice to enter motherhood, or have more children (this difference cannot be deduced from the calculated TFR). In fact, 58 per cent of women polled in 2006 thought that motherhood hinders a woman's career (CIS, 2006 cited in Ibañez, 2010). Furthermore, the factor that influences this incompatibility seems to be the cultural and institutional conception that the mother is responsible for the childcare.

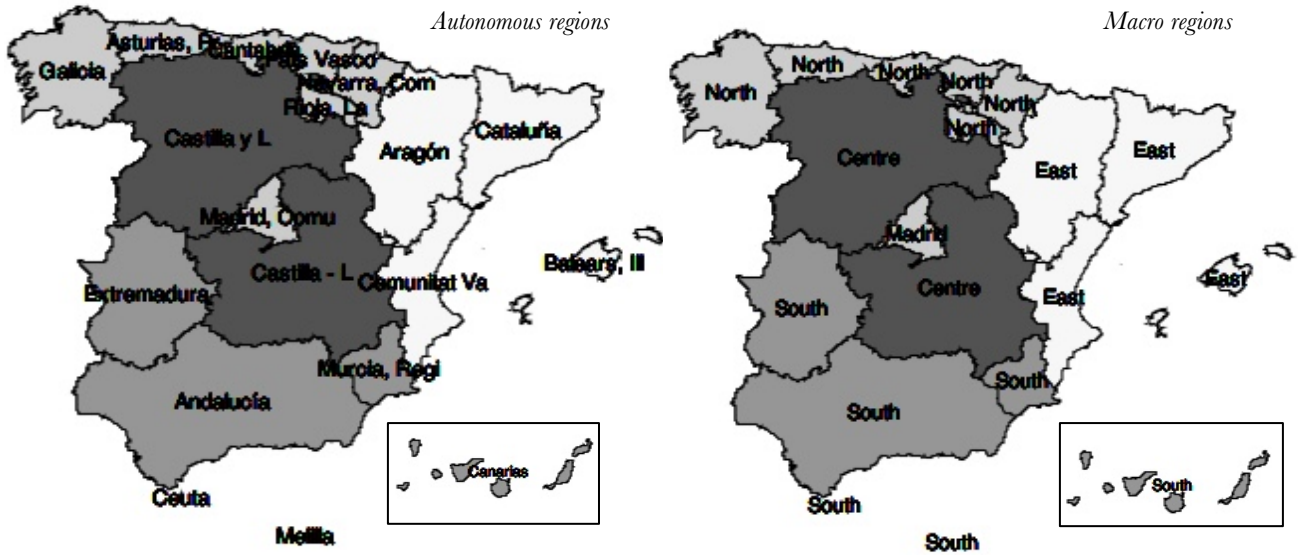
Nevertheless, we have for now only been looking at national aggregate figures. This can give us an idea of how the overall situation for Spanish women looks like, but it hides heterogeneities that can be found on a more disaggregated level, such as the 19 Spanish *comunidades autónomas*. The next section presents the regional realities of FLFP and TFR, with a revision of cultural and traditional heritage that still have an influence in present-day Spain.

II. Regional heterogeneity in Spain

When thinking of the country of Spain, it always appears as a little Europe itself stretched over 865 km from North to South: the North and Northeast seem culturally, and geographically, very close to Central Europe. The capital, Madrid, and the second-most important city, Barcelona, are comparable to any other modern European capital. The regions in the Centre of Spain surrounding Madrid, however, are a rather desolate place (dry, desert-like, scarcely populated and traditionalist). The East, with its orientation towards summer tourism on the coastline, is densely populated in constant exchange with the exterior. Further South we come into a completely different world: for instance, Andalucía, with its Arabic heritage in architecture and culture, its incredibly hot climate, and on the other hand traditional, poor and agricultural society outside of the touristic places. Then, there are Ceuta and Melilla, the two autonomous cities and former Spanish colonies on the north coast of Africa, as well as the isolated but touristic Canary Islands and Balearics.

These perceived differences between the Spanish regions can be found in many other indicators that we will look upon in this section. Seizing the region-specific cultural heritage is important if we want to understand why there is a pronounced regional heterogeneity in women's fertility and economic activity (Kertzer, 1995, 1997 cited in White et al., 2007).

Figure 5. Autonomous and macro regions Spain 2016



Note: Maps using data from INE

For the regional analysis, the 19 autonomous regions of Spain were grouped into six macro-regions combining the NUTS1 classification of the European Union with the paper's own research interest.[24] Figure 5 shows both the autonomous regions (left) and their grouping into macro regions (East, South, North, Centre and Madrid) (right). We can see that the autonomous

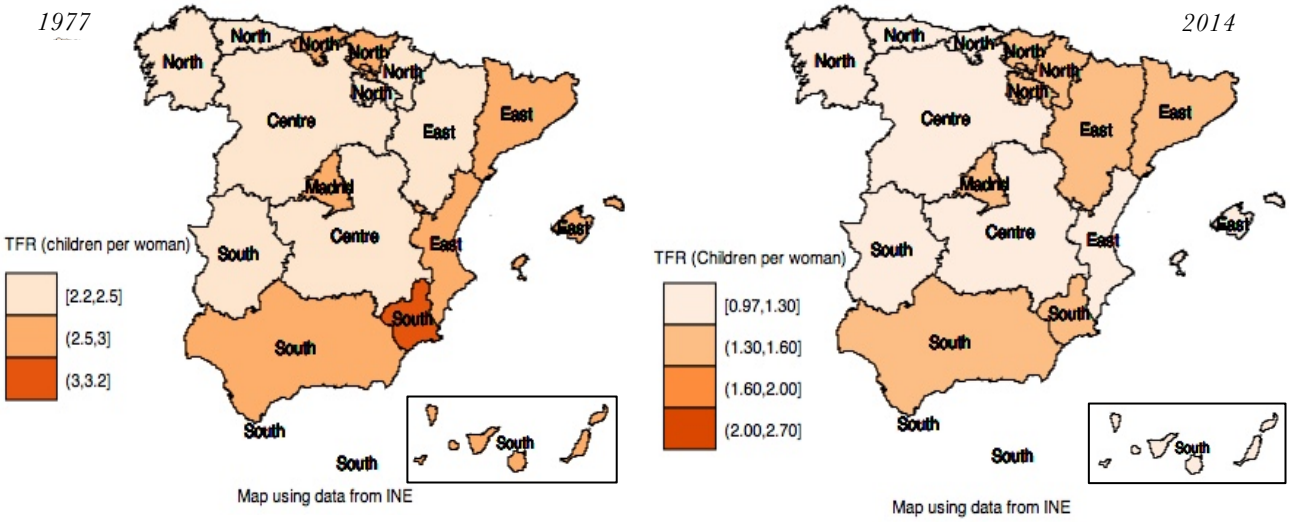
[24] The distribution of the regions into macro regions can be found in Annex 5.

cities Ceuta and Melilla,[25] as well as the Canary Islands and the Balearics are not on Spanish ground.

A. Fertility between 1977 and 2014

We see in Figure 6 that average fertility has declined in all Spanish regions form the end-70s until 2014.[26] Over 40 years, the regions followed commonly a u-shaped like pattern, sharply declining until the 2000s, followed by a common increase until the financial crisis, with an anew decline after 2008. Ceuta and Melilla diverged from the rest of Spain in timing and levels, especially since their TFRs are incessantly increasing since the 2000s.[27] While in 1977 the TFR of a majority of regions was around 2.2 to 2.5 children per woman, in 2014 the most frequent TFR was 1 to 1.3. The difference between lowest and highest fertility on the mainland in 1977 was 1 child per woman, today it is 0.6 (excluding Ceuta and Melilla).

Figure 6. Fertility in 1977 and 2014



Note: The data for Ceuta and Melilla are not visible in the graph. TFR for Melilla was in 1977: 2.22, in 2014: 2.7. For Ceuta, in 1977: 2.8, in 2014: 2.0

Only 7 *comunidades* have TFRs above *lowest-low* levels today. As much as six regions, majorly of the North, have actually never again experienced fertility levels above *lowest-low* levels.[28] The TFR differentials of some regions have reverted between 1977 and 2014. The South, known for its economic underdevelopment, had highest fertility levels in 1977, but has lower or lowest levels today.

[25] The *shapefile* of the Spanish autonomous regions does not include Ceuta and Melilla, since they are on the African continent. Their specific numbers/values can be found in the footnotes below the graphs.
[26] 1977 and 2014 were the first and last year of comparable data at regional level available at the National Statistical Institute of Spain (INE).
[27] Annex 8 contains the regional development of TFR.
[28] Asturias, Galicia, Cantabria, Castilla y León, País Vasco and Canarias.

The highest present-day TFR on the mainland is shared between some specific *comunidades autónomas* of the North, East and South, among which we find the most and least developed regions of Spain: Murcia (1.54), followed by Navarra (1.42), Cataluña and Andalucía (both 1.37). Interestingly, the highest fertility of Spain cannot be found on the mainland, but in Ceuta and Melilla, the old colonies bordering North-Morocco, which are utterly diverse from Spain in terms of culture, geography, ethnic composition and economic development.

If we have a look at the timing, the TFR in some Northern regions[29] dropped below 1.3 as early as in 1985, when the South and Centre still had replacement fertility levels. The South[30] reached a TFR below 1.3 only as late as in 1995.

Regarding the mean age at first birth (Annex 9), the regions followed a common convergence towards higher ages during the observed period but present divergent levels today: Compared to 31.8 years, the national average in 2014, women are much older in the Northern regions (e.g. nearly one year older in the País Vasco), while they are younger in the South (e.g. 31.1 years in Murcia).

B. Female Labour Force Participation between 1977 and 2014

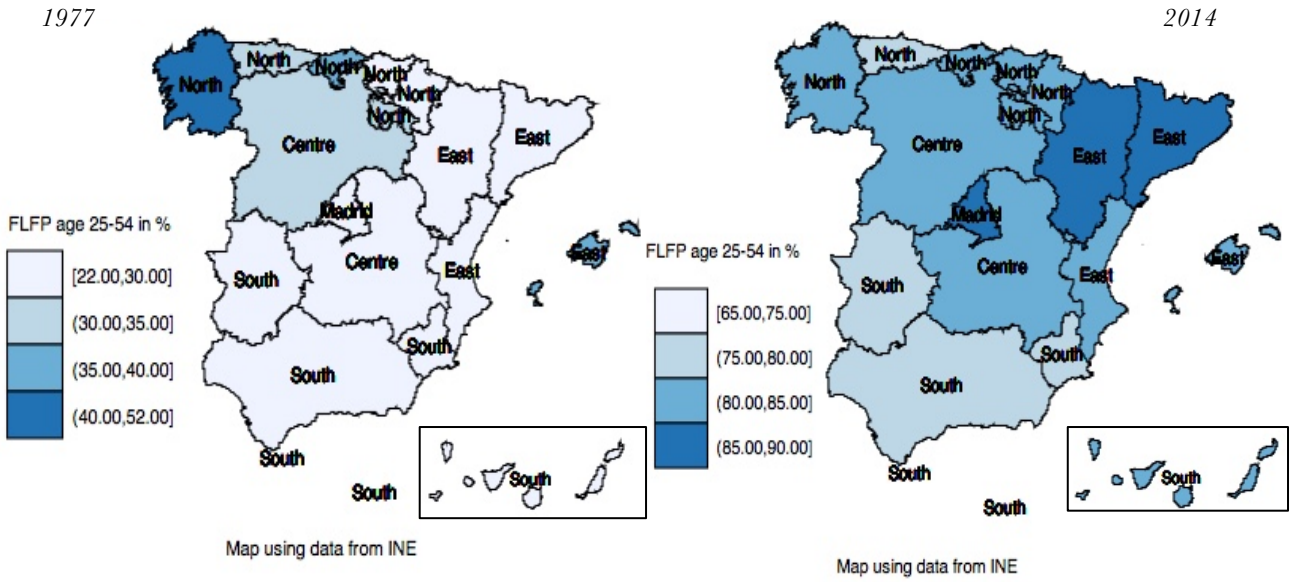
Female labour force participation (FLFP) offers similar regional heterogeneity to fertility. Altogether, the regions have converged towards higher FLFP between 1977 and 2014, while the gap between highest and lowest FLFP has decreased. It is worthwhile noting that some regions that had highest female activity levels in 2014, such as Madrid and Cataluña, where among the less participative ones in 1977. Some regions have even reversed completely from being among the most participative ones in 1977 to the bottom of the distribution today.[31]

In Figure 7 we see that women aged 25-54 in the Northern half of Spain[32], especially in the North-East, are more likely to be working than women in the South. In 1977, in contrast, it was more likely that a woman was in the labour force when she was living in the Northwest. The *forerunners* by far in 1977 were Galician women, while today it is women from the regions comprising the three economically most important cities of Spain: Barcelona, Madrid and Bilbao (dark blue).

The overall increase in total female activity has been important (from 28.3 to 53.6 per cent)[33]. Yet, some regions of the South have made astonishing advances, for example Andalucía with an increase from 22.6 to 52.3 per cent, while others, especially in the North, have experienced less progress, such as Asturias, Galicia or Cantabria.

[29] Asturias and País Vasco.
[30] Andalucía, Murcia, Extremadura, as well as central Castilla la Mancha.
[31] Annex 10 contains the regional development of FLFP.
[32] Canarias, Madrid, Cataluña, Baleares, Murcia, Valencia, la Rioja, Navarra, Aragón, Castilla la Mancha.
[33] See for this the map in Annex 11.

Figure 7. Female Labour Force Participation ages 25-54 in 1977 and 2014



Note: The data for Ceuta and Melilla are not visible in the graph. Furthermore, for 1977 no data is available. The FLFP in Melilla in 2014 was 65.88, in Ceuta, it was 65.2 (the lowest category of light blue in the right map).

Over the observed period, the composition of the female working force has also changed notably. Nearly in all *comunidades autónomas*, the participation rates of women in their main working age (24 to 54) is over 75 per cent – with the exception of Melilla and Ceuta who have today lowest levels. In 1977, the highest FLFP was reached by women aged 16 to 24 (on average 49 per cent), while the participation of women aged 25-65 was on average only 22 per cent.[34] According to Guner et al. (2014), this change was half triggered by changes in the composition of the working force (education, maternity, marital status), and half by the behaviour of the group of *married* women. Interestingly, the increase in FLFP happened independently from the number of children per women: for instance, the activity for women with two children increased from 18 to 49 per cent between 1977 and 2007.

However, today's figures show a clear difference between the regions where women are increasingly present in the labour market, the darker blue regions in Figure 7, and the regions that are falling behind: majorly the South. It seems that among the EU-15 only Italy, known for its persistent low FLFP, has lower female activity rates than Andalucía (Fundación Centro de Estudios Andaluces, 2007, p.117; White et al., 2007).

[34] The low FLFP can be explained by the fact that at that time marriage was followed soon by motherhood, a status that excluded women from economic activity.

C. Values and cultural factors

An important factor showing high regional variation across Spain is the value system change linked to the Second Demographic Transition (SDT[35]) and in turn to gender norms, which influences both women's economic activity and fertility decisions.

In the past decades, Spain has experienced cultural value changes in the sense of the SDT towards a weakening of the family and marriage as institutions: marriages (especially religious ones) are decreasing since the 2000s, cohabitation before marriage as well as divorces and births out of the wedlock are increasing (Arpino and Tavares, 2013). The speed at which these changes are taking place, as well as the level differences, are quite substantial across Spanish regions.

As a matter of fact, some regions have experienced very quick and important advances of various dimensions: in terms of individual autonomy, individualism (of relationships and towards children) as well as gender equality (in the household sphere and on the labour market). For instance, the South and Canarias are falling behind the rest of Spain with respect to individualism in relationships, while the North and Northeast are clearly leading. All regions but the Northeast and Centre became less individualistic towards children, that is, more inclined towards starting a family. The East fell behind in advances of gender equality on the labour market, and was situated, together with the Northwest, below the European average in 2008. The Centre and Madrid have made the biggest advances in this direction, while the South is falling behind. Concerning gender equality in the household, Madrid, the South, the Canaries and Northwest became more equal, while the Centre has gotten less equal. The leader is clearly Madrid. This insinuates that there might persist a significant regional variation in the perception of gender norms today (Arpino and Tavares, 2013).

This can further be related to the process of secularization and the influence of traditions, in Spain notably the Catholic traditionalism, on gender norms. According to a survey on religious conviction carried out by CIS in 2012, there is a substantial regional variation in the percentage of people who perceive themselves as practicing Catholics, reaching from 58.6 per cent in the País Vasco to 85 per cent in Murcia. It is worth mentioning that the regions with the lowest percentage of Catholics[36] are among those with the highest fertility levels on the mainland – with the exception of the South (Murcia and Andalucía), who has both, high fertility and high religious foothold. Women aged 25-54 of those two regions are actually among the least economic active ones.

However, we have to keep in mind that the 'highest' TFRs that we are talking about here refer to solely 1.54 children per women in Murcia, 1.42 in Navarra, 1.37 in Andalucía and Cataluña, 1.36 in the País Vasco and 1.33 in La Rioja and Madrid – we can see they are still very close to the *lowest-low* threshold.

[35] This concept is explained in section III.
[36] Madrid, Cataluña, País Vasco and Navarra.

Another important feature is that Ceuta and Melilla stand out from the other Spanish regions: these two autonomous cities in North-Africa are a cultural melting pot due to their mixed ethnic and religious population composition. In Ceuta, approximately half of the population is Spanish and Christian, while around 49 per cent is Arab-Berber and Muslim, the rest being Jewish, Hindu or other minor ethnic groups. The Arab population in Ceuta is known for its high level of illiteracy, poverty, youth unemployment and extremely diverging fertility patterns from the Spanish population, namely over 4 children per woman (Roa, 2006). Melilla is also host for four major religions: Christianity, Islam, Judaism and Hinduism, while the Christian population seems to be decreasing in recent times. Furthermore, being a major attraction point for people wanting to enter the European Union via Spain, Melilla is also experiencing increasing numbers of refugees from poorer countries of Africa, and faces similar ethnic fragmentation problems as Ceuta.

We can see that these two autonomous cities belonging to the Spanish territory are highly divergent in cultural traditions, norms and ethnic composition, which might play a role in their previously observed divergent fertility and female activity behaviour. In fact, these two autonomous cities have the lowest female active population aged 25-54 of the Spanish territory as well as the highest fertility rate per woman.

Another indicator of secularization is the evolution of the importance of religious versus civil unions. All over Spain, civil marriages have started to outnumber religious ones since 2009. In 2013, the number of civil unions was nearly double of the religious ones. Interestingly, religious marriages have started to lose importance only in 2000, but have since then more than halved. The decreasing importance has been related to the fact that marriages seem to have turned into something more like a ritual for younger couples; the religious aspect has lost its importance. However, the decrease has been unequal across the Spanish regions: it has been greatest in Madrid, Cataluña, the Comunidad Valenciana, the Baleares and Canarias. In contrast to those regions, the Centre and South (mostly rural and still influenced by Catholic traditions) and some regions of the North have shown much more modest declines in religious marriages. The South is the region that still presents a high percentage of religious marriages, for instance in Andalucía in 2013 the number of civil and religious unions were equal. Barcelona and Madrid are the regions were the ratio civil to religious unions is by far highest in Spain. At the same time, civil marriages have increased significantly for all Spanish regions together, but the increase was most pronounced in all provinces around Madrid and the Northeast (La Información, 2013; Allón and Díaz, 2010). Spanish scholars have claimed that these disparities have to do with the process of urbanization that goes hand in hand with higher degrees of individualization, as well as the higher presence of immigrants (Allón and Díaz, 2010).

It seems like during the process of convergence towards very low fertility levels of Spanish regions, the regional disparities and level differences have not disappeared. We argue that the regional realities, among others cultural differences and the process of secularization, can account for parts of the regional heterogeneity in the fertility behaviour of Spanish women. Thus, a regional

component in the multivariate analysis of the probability to enter motherhood will help to better understand what influences women in their decision to enter motherhood.

III. Theoretical framework: women's economic activity and fertility

A. Lowest-low fertility

The current persistence of very low fertility rates below replacement level in a significant number of European as well as other developed countries around the world is till date a rather odd phenomenon for demographers. This section tries to give an oversight of main currents trying to explain the drop in fertility in occidental societies, to arrive at the theoretical background and previous research that underlie the research question of the present thesis.

The first significant decline of fertility behaviour across the Western world took place around the end of the 19th century, and was explained by the theory of First Demographic Transition (FDT). It was related to decreasing necessity to give birth to a great number of children due to declining mortality and increasing longevity thanks to medical advances, when societies moved from a traditional to a more modernised state (Coale, 1989).

During the 20th century, TFRs in the developed world experience again a sharp decline towards below replacement levels. Some countries even reached *lowest-low* fertility levels starting in the 1990s (Kohler et al., 2002). It became necessary to explain these recent and odd fertility developments, especially why TFR remained at such low levels. The theory of the Second Demographic Transition (SDT) tries to explain the second historic drop in fertility from a macro point of view. It attributes it to changes in the value system of modernising and urbanising societies towards more individualistic behaviour. Parenthood and the institutions *family* and *marriage* started to become less central in people's lives, which in turn influenced marital behaviour[37] (Laesthaeghe and van de Kaa, 1986; Van de Kaa, 1987, 2004). One could speak of an emerging new view on parenthood, away from the urge to have many children towards having few 'quality children' (Ibañez, 2010). The latter happened parallel to a process of secularization, that is, the decreasing influence of religious and traditional institutions on Western societies (Lesthaeghe and Surkyn, 1998, cited in White et al., 2007).

The problem with the SDT theory is that it would predict that today the most individualistic societies should present lowest levels of fertility. Yet, today's countries with the lowest TFRs are known for its still quite present influence of Catholic-traditionalism and inherent *familistic* cultural systems. What is worth noting is that those countries with lowest fertility levels today were also found to have lowest female economic activity rates (Del Boca, 2002; McDonald 2000). In contrast to that, the Protestant North, more egalitarian, secularised and less *familistic*, experienced increasing or at least on average stable fertility levels at nearly-replacement levels.[38]

The importance to understand the phenomenon of below replacement fertility is related to its implications on long-run population dynamics, such as: active labour-force aging and -decrease, pressure on the welfare state and public finance, the advent of 'cultural nationalism' as well as the

[37] Such as increasing age at first marriage and first birth, divorces, cohabitation and extra-marital fertility.
[38] Excluding Germany and Switzerland, with TFRs around 1.5 since 1990.

cleavage between the increasing older generations and the youth (World Bank, 1994; Galasso and Profeta, 2004; Cremer et al., 2000; Caldwell et al., 2002).

However, the literature is still divided as to whether very low fertility actually poses a serious problem in the long run, or not. The transition paradigm, on the one hand, assumed that once the postponement of births to higher ages finished, fertility would automatically go back and stabilise at replacement levels (Bongaarts and Feeney, 1998). The 'Easterlin Hypothesis' claimed further that there exists an inverse relation between fertility and cohort sizes: under restrictive immigration[39], smaller cohort sizes would lead to higher fertility in the subsequent generations (Easterlin, 1980, cited in Kohler et al., 2006).

Yet, Spain is among the countries with highest positive net migration in Europe over the last three decades – while fertility has continued to decline, even if immigrants' fertility behaviour was, on average, higher than for the native population (Roig and Castro, 2007, cited in Castro and Rosero, 2011). In this vein, and contrary to the transition paradigm, the concern about a downward spiral towards a 'low-fertility trap' emerged, since various countries show persistent low fertility levels even after the transition to older-age parenthood finished (Lutz et al., 2006).

It rather becomes evident that the development of fertility is not explainable by a mere aggregate-level analysis. As we can learn from the theory of SDT, culture and values have an important influence on the micro-behaviour of woman and their partners in the decision of fertility.

Even if hard to disentangle conceptually, there exist different micro theories around the complex phenomenon of fertility, depending on which socio-cultural approach is chosen. According to Ibañez (2010), one can apply an institutional or economic focus. Fertility can also be approached from a cultural point of view, such as the influence of culture and values/traditions on micro fertility behaviour. Going further in this direction, some scholars have recently focused on the geographical influence (via culture and norms) on the choice of motherhood (White et al., 2010).

B. Literature review and dimensions of interaction between FLFP and TFR

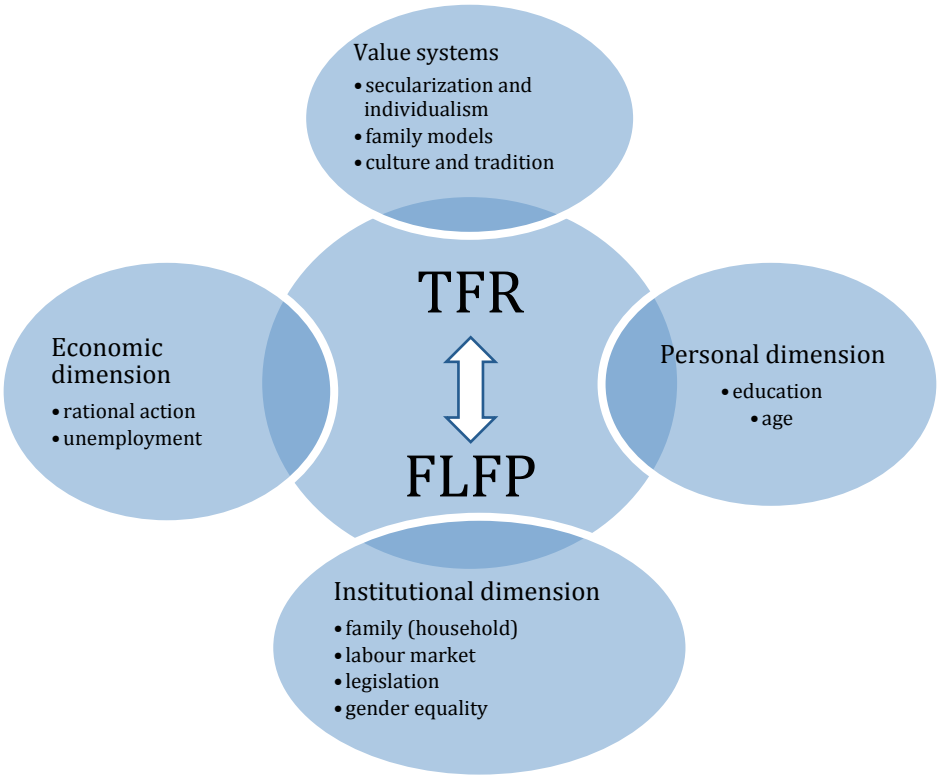
One aspect of the micro-behaviour around fertility choices has been widely looked upon and dominates in the literature: the relation between female labour force participation and fertility. The main idea in this strand of literature is that there exists a sort of *incompatibility* between being a mother and working. In light of the increasing opportunities on the labour market for ever better educated women, it might seem plausible that they would substitute away from being a housewife and mother and work instead.

[39] The foreign population in Spain has increased from 0.52 per cent in 1981 to 10.7 per cent in 2014. Since 2007, it fluctuates around 10-12 per cent (INE, 2015a). After the financial crisis in 2007-08, net migration has turned negative (World Bank, 2016).

The macro view on this matter has been inconclusive, but Mishra et al (2010a,b) have recently found an aggregate inverse long-run relationship between TFR and FLFP for the G7 countries and the OECD[40] over the past 40 years. These findings sustain the hypothesis of *incompatibility* between work and motherhood. Such an aggregate analysis between FLFP and TFR for the case of Spain might however not be that fruitful, since it would solely give an idea as to which of the two variables 'triggered' the other in the long run, but not analyse which factors underlie this relationship.

Figure 8 shows the different channels through which fertility (TFR) and female labour force participation (FLFP) are connected in a micro-behaviour framework, such as applied in this thesis' analysis. In the following sub-sections, the different channels will be explained more explicitly.

Figure 8. Dimensions of interaction between FLFP and TFR



[40] The OECD paper included Spain.

i. Value Systems

As discussed above, the SDT has entailed significant changes in parenthood, together with a movement towards a more individualistic and secularised society. Parenthood is not anymore central in people's lives. As a matter of fact, motherhood is not determining younger women's lives anymore, since birth control and termination have become available and legal, and women's prominence in the public sphere has increased.

However, the degree to which value systems have changed in the peer group surrounding a woman at the moment of her decision to enter motherhood might still be quite influential. Such factors are for instance, the degree to which religion, especially marriage, is still seen as a necessary condition for family founding, or cohabitation and out-of-the wedlock births are socially accepted.

The type of prevailing family model might influence a woman's fertility behaviour. Pfau-Effinger (2004, cited in Ibañez, 2010) defines five possible models in western society: the family economic gender model, the male breadwinner/female home carer model, the male breadwinner/female part-time carer model, the dual breadwinner/state carer model, and the dual breadwinner/dual carer model. The traditional family can be modernised in two directions: the male breadwinner system can be paired with part-time work popularisation for women (Germany, UK, Holland), or move towards a dual breadwinner and external care model (France, Sweden). In these two modernised models childcare is covered. If however the family model does not reach to develop into one of those categories, such as when the labour market becomes increasingly equal but the woman has still the single role of the carer, this can lead to negative effects on fertility behaviour (McDonald, 2000).

In a micro-behaviour analysis of regional fertility behaviour for Italian women, White et al. (2007) have accounted for such cultural value factors. For instance, they included the degree of family-friendliness of the public sector (childcare availability) that stands for the family model, and the process of secularization and other regional factors (religiosity, importance of marriage). Contextual and structural factors seem to the authors important in explaining fertility behaviour.

ii. Institutional dimension

A strand of the literature about gender equality claims that inequalities between men and women can be made accountable for the persistence of fertility levels below the desired level, more particularly below replacement. If gender equality has achieved different levels in the central institutions of a country, such as the professional sphere (labour market, educational institutions) or the household sphere (sharing childcare and housekeeping duties), then it becomes increasingly difficult for women to match their roles as mothers and workers. This, in turn, results in persisting sub-replacement fertility until corresponding gender equality levels are reached (McDonald, 2000).

In this line, if the TFR and FLFP evolve in opposite directions, such as that TFR decreases with increasing FLFP, it is possible to argue that in this country economic activity reduces the fertility behaviour of women. Supporting McDonald's theory, Arpino and Tavares (2013) found in their analysis of the role of value changes in determining fertility that for Italy and Spain higher gender equality on the labour market had led to lower TFR because gender equality within the family remained unchanged.

The analysis of the involvement of the state in providing an environment of role compatibility further follows an institutional approach. For instance, public childcare, maternity and paternity leave legislations or tax benefits can help to reduce the negative effect of FLFP on TFR (Bernhardt, 1993, cited in White et al., 2010). However, Del Boca (2002) discussed that childcare as means to improve compatibility of motherhood and work has its limits. For instance structural failures, such as untimely and too limited opening hours, miss the point of childcare and render it futile.

It is possible to analyse the latter with a behavioural framework of the impact of economic activity – and related factors like the possibility to work part-time (representing the flexibility of the institution labour-market) – on the decision of motherhood. Ibañez (2010) was interested in this specific relationship for today's Spanish women, and looked besides the influence of domestic culture also at the institutions: labour market and public policies (possibility of reconciliation). Her results on the probability to become a mother (and progress to more children) were however inconclusive, carving out that both working (in combination with higher educational attainments) as well as being unemployed impact this decision importantly. Alonso-Antón et al. (2015) were as well interested in how FLFP and TFR affected each other mutually and performed a one-step analysis of the probability-of-motherhood for Spanish women, paired with how the fertility decisions affect women's professional decisions. They found that the probability to have a child is influenced negatively by a woman's economic activity, fulltime work and the occupation requiring higher educational attainments. They did however not explicitly analyse the progression in childbearing and performed a national analysis.

iii. Economic dimension

Another dimension that links TFR and FLFP can be explained through rational action linked to the two decisions to fist be a mother, and then have subsequent births. According to Bernardo and Requena (2003, cited in Ibañez, 2010), the first step is rather influenced by the fact that there exists a tradition of needing to be completely independent from the parents before couples commit to parenthood in Spain. For instance, young adults are among the ones that emancipate latest in Europe (a Mediterranean phenomenon), which can be explained by the high level of economic uncertainty due to, for instance, extremely high youth unemployment. The level of income and type of occupation are thus factors often accounted for in micro-studies of fertility (Ibañez, 2010; Alonso-Antón et al., 2015). For the second decision to progress in parenthood, five aspects should be taken into account: direct costs, opportunity costs, benefits, uncertainty and

gender equality.[41] The latter are, however, quite difficult to capture in order to control for them in an empirical work, since they often represent personal perceptions, but help to understand the underlying process of decision-making.

An important insight comes further from Baizán (2006, cited in Ibañez, 2010): unemployment seems to have overall a negative impact on the number of children a woman has. To the contrary, Ibañez (2010) found in Spain an unemployed woman is more likely to progress to have more children.

iv. Personal dimension

Lastly, we can think of various personal traits of a woman that might influence her decision to become a mother, and in some cases through the channel of FLFP.

Age is certainly one of these traits: at relatively high ages, when approaching the natural frontier of fertility, it becomes less and less probable that a woman will get pregnant anew. However, given the developments explained by the theory of SDT and women's increasing participation in higher education, pregnancy sets in at ever older age (Figure 3) since women postpone entering the workforce to later stages, once education is completed. Immediately after entering the active labour force, the probability to decide for motherhood might then be rather low, for which older ages might be related to higher probability of motherhood.

Connected to the latter, Ibañez (2010) discussed that a mother's level of education is a decisive factor in the choice to have *more* children in Spain: women with low educational level are more likely to be housewives, or even unemployed, than women with higher educational attainment, due to a high fragmentation of the Spanish labour market for women. Furthermore, it is most likely for women to be in a stable employment with a university degree, which in turn will influence fertility. Our reference works altogether account for the latter factors and found similar results (Ibañez, 2010; Alonso-Antón et al., 2015; White et al., 2007).

[41] For more details see Ibañez (2010, p. 9-10).

IV. Empirical Strategy

The empirical part of this thesis consists of a first sub-section introducing the model that will be analysed, a second one introducing the dataset and all variables used in the analysis, a third one explaining the econometric approach and a last one presenting some descriptive statistics.

A. A two-step model of the fertility decision

We are interested in a micro-level analysis of the socio-economic and regional factors that determine the likelihood to enter motherhood and the transition to subsequent births in Spain. We are especially interested in the linkage between labour force activity and fertility, for which key explanatory variables will be related to economic activity.

The advantage of using a *two-step* model in contrast to a *one-step* model is that we account for the fact that there exist two independent choices concerning the fertility behaviour of women: whether she decides to become a mother in the first place, and whether she progresses in giving birth to more children.

We assume that these two processes are driven by different factors. One can imagine that being married (that is, in a stable relationship), having a job and age are quite important when a woman considers to get pregnant for the first time. The degree of religious influence might be responsible of how important marriage is for her in this decision.

On the other hand, once a woman is already mother of a child, some factors might become more relevant for her to progress to subsequent births, and others less. The possibility to reduce work to part-time, if she is working or unemployed, if her partner is economically active and can support the family, if somebody can care for the children (grand-parents etc.) or how many children she already has, might get more important for this decision. Other variables, such as being married, might become less relevant.

B. Data and variables

i. Data

The data set used for this analysis is a cross-section of the Spanish population from the year 2011 (Spanish national census), accessible through the webpage of IPUMS international. It contains individual-level data for every person of the polled households, such as demographic, educational, socio-economic and geographic information.

One great advantage of censuses is the availability of information on a large set of persons, here 4,107,465, and the wide geographical coverage, here all of the 19 Spanish regions (*comunidades autonómas)*. Therefore, we assume it to be a representative cross section of the Spanish population.

Yet, a shortcoming of this census is the lack of detailed information about employment and occupations unlike the national Labour Force Surveys[42] for which the present analysis was limited in the inclusion of some important variables.[43] Furthermore, it lacks a time dimension that would be feasible with panel data. Hence, no dynamic changes and event-historic components can be analysed, such as for instance yearly varying regional peer-group information (FLFP, TFR).

The final dataset contains information of 406,244 women of Spanish nationality (and their spouses, if applicable) living in Spain at the time of the census and are aged 19 to 45. This age interval is based on Ibañez (2010), defining the general period in which decision about reproduction is taken in Spain. Childbirth in women below age 19 and over age 45 is extremely low.[44] However, the age range applied in the literature varies. Alonso-Antón et al. (2015) chose age 20-44 for Spain, White et al. (2007) age 18-49 for the Italian case.

Women of other nationality than Spanish do not form part of this analysis, since their fertility behaviours were assessed to be quite different compared to native Spanish women, especially the timing and number of children for women from the Caribbean (Castro and Rosero, 2011).

[42] Ibañez (2010), Alonso-Antón et al. (2015) and White et al. (2007) used National Labour Force Surveys.
[43] Micro-level data of the trimestral Spanish Labour Force Survey (ENA) has a significant economic cost (332€), for which the present analysis was limited to the freely available census from IPUMS international.
[44] Annex 12 contains the distribution of fertility by age groups.

ii. Explanatory variables

In order to analyse factors influencing women in their fertility choice, we chose explanatory variables following the approaches of Ibañez (2010) and Alonso-Antón et al. (2015) for personal socio-economic factors and White et al. (2007) for the regional components. Annex 15 explains the variables, corresponding name in STATA as well as basic variable statistics.[45]

Economic activity is the key explanatory variable in this analysis and has a binary outcome: 1 if the woman is economically active (employed, unemployed), and 0 if she is inactive (e.g. retired). According to the literature and previous findings, we expect activity to have a positive influence on the decision to enter motherhood in the first place. However, for higher order births, activity might turn out to have a negative effect, since there are still many hurdles to combine motherhood and work in nowadays' Spain (IPF, 2016). We included further *interaction terms of activity and regions*, for which we control for the effect of being active in specific cultural circumstances.

Unemployment is a factor that is expected to influence women in their fertility decision: the benefits could be added to the maternity-leave-pay, for which we could think it influences motherhood positively. We could also argue to the contrary that as long as a woman is unemployed the needed level of economic security is not met, and hence motherhood will be postponed.

It is further important to analyse if the reduction of work to part-time has a significant influence on women's choice of fertility. For first births, activity might primarily be the decisive factor. For the decision to have more than one child, it might however become relevant if the mother is able to reduce her working time to part-time in order to be able to care for the additional toddler. Furthermore, including *interaction variables of full-time and the specific regions* accounts for geography-specific effects of work-flexibility.

Marital status is another decisive factor in the choice of motherhood. This dummy variable is 1 if a woman is currently married or has been so before, and 0 if she has never been married or is divorced. Marriage is expected to positively influence fertility, especially in more traditional regions. Given the increasing cohabitation and out-of-the-wedlock births in Spain, we are interested in how marital and extra-marital fertility vary. We also included *interaction terms of the different regions with married,* in order to capture for instance how 'being married in the South' affects fertility. There is no information about the type of union (civil/religious/cohabitation) in the census, which limits the impact analysis of union-security.

We control for *age* and its *square*[46]. Unfortunately, age at marriage is not available in the dataset, for which we cannot determine how long a woman is in a stable union. Given the renewed transition to higher mean age at first birth (Figure 3), older ages are expected to be positively

[45] We used STATA (Version 12) for our econometric analysis.
[46] We control for a possible u-shaped influence on fertility: low probability at very young and very old ages.

related to giving birth to the first child. However, age could have a negative effect on subsequent births, since women approaching the natural limit of fertility are less probable to progress in giving birth.

In order to control for the cohort-specific fertility behaviour, we include *cohorts* in five-year intervals in the models, starting with the 1965-69 cohort and ending with 1990-94. It is expected that for younger cohorts the probability to enter motherhood and progress to higher births should be lower, consistent with the sharply decreasing fertility observable at the aggregate national level (see Figure 9).

A further important explanatory variable is *education*. We established a 5-category variable according to the Spanish educational system: 0 unknown, 1 basic primary, 2 secondary and higher non-tertiary, 3 higher technical education, 4 undergraduate and graduate studies, 5 PhD. A Dummy for *university degree* (1 if education is level 5, 0 otherwise) captures the effect of very high education on fertility decisions. Both the latter are expected to positively influence on fertility.

As explained in section II, four Dummy variables for the Spanish *macro regions* were created out of the 19 autonomous regions: South, North, East and Centre, taking the value 1 if the region lies within the respective macro-region. As reference category we chose the region Madrid, since it has today the highest FLFP and is among the seven regions with above *lowest-low* TFR; it functions as a sort of forerunner in value change in Spain. In order to control for the divergent behaviour of Ceuta and Melilla, we constructed a second South region without those two regions, and added another Dummy for them separately.

Another control is *household size*. We expect that in households with more family members the task of caring becomes easier, since for instance grandparents could care for a toddler if its parents have to work. In the second-step decision, we control for number of children a woman already has (dummy variables for *one, two, or multiple children*). It is expected that the more children she already has, the lower the probability she will be pregnant anew (Ibañez, 2010). As further controls, we include the *spouse's educational level* and *activity* status (if applicable), to check for possible influence of the partner's characteristics on her fertility decision.

Unfortunately, we cannot deduce whether the polled individuals live in a rural or urban area, which would have shed further light on the influence of traditions on TFR. Ideally, we also would have liked to include regional context variables, such as for instance regional FLFP, TFR or civil marriage ratios, in order to explore the pressure women receive from their peer-groups (see White et al., 2007). However, since the data is cross-sectional and for a binary outcome probability model the variables need to contain varying information, we could not include this information in the model.

iii. Dependent variable

The dependent variable in this analysis is not completed fertility (the total number of children a women has had over her fertile period), but 'recent net fertility'*:* the number of own children under the age of 5 that are present in the women's household, according to Dribe et al. (2014). Also Alonso-Antón et al. (2015) state that analysing the probability of going into motherhood is only relevant when those women who are *currently* being a mother are chosen as a counterfactual. Put differently, a woman's total number of children might no be affected by her current economic activity, age or marital status, if her children are already over twenty years old and emancipated. A possible effect of labour activity on giving birth is only relevant for those women that currently have one or more children in an age of needing care. Hence, we selected those women for whom the total number of children corresponds to the number of children below age 5. In fact, the 2011 census only includes the latter two information concerning women's fertility behaviour.

We are thus left with the fertility of women who are currently mothers and have no other children aged 0 to 4, and the fertility of those that have never been mother before and are currently not having children. The dependent variable *fertility* can take values of 0 and greater and states the number of children aged 0 to 4 that currently live in the household of the mother.

C. Methodology

The dependent variable *fertility* (the number of children under age five living with the mother) is a restricted discrete non-negative variable that can take the values zero, one, two, three or four in the analysed sample. The distribution of the dependant variable can be observed in the Annex 16. The variable suffers from the problem of excess zeros, which is common to count data models (Winkelmann and Zimmermann, 1995). The mode of the dependent variable is 0, its variance is 0.26 and its mean 0.22, which further indicates overdispersion.[47]

Data in cross-sectional format and a dependent variable with the given characteristics justify the usage of a *Hurdle model*, also called a two-step decision model, to pursue the research question about the progression of fertility. The econometric discussion of Winkelmann and Zimmermann (1995) has served as a guideline for this approach. The Hurdle model addresses the problem of excess zeros thanks to the assumption of two underlying data generating processes, one for (motherhood/no motherhood), and another for (one child/more children).[48]

The residuals of an ordinary-least-square (OLS) regression for this type of dependant variable would violate the assumption of homoskedasticity and normality of the errors of OLS regressions. This would imply invalid standard errors and significance levels, erroneous hypothesis tests (see Scott Long, 1997, p. 38-40).

[47] Since variance > mean.

[48] First econometric work using this methodology was Mullahy (1986).

The first step of the hurdle model consists of an ordinary Probability process looking at the probability to be a mother or not; we used a Probit model. The second step works on the positive integers of the dependent variable only. For robustness issues, negative binomial, ordinary Poisson and zero-truncated Poisson regressions where run in order to see which predicts the actual observations best.[49] We chose the zero-truncated Poisson model over all.

The next sub-sections explain the two steps of the hurdle model in detail.

i. First step: Probability to become a mother

The first step of the hurdle model is a binary probability model with the dichotomous outcome of Y = 0 for no children at all, and Y = 1 for one or more children. Moffitt (1984) claims that if the outcome of a process is a dichotomous event (as in the case of fertility), a Probit model is appropriate, although there is no convention about the preference of a Probit or Logit model (Gujarati and Porter, 2009). Alonso-Antón et al. (2015) used for instance a one-step Probit model in their analysis of the decision to motherhood for Spanish women. We tested Probit against Logit and found no significant difference in the outcomes, while Probit did slightly better[50] than Logit.

Whether a woman decides in favour of becoming mother or against it is a qualitative process with a binary outcome. The decision of the i_{th} woman depends on an unobservable latent variable I_i, which is determined by the set of explanatory variables (X_i) introduced above. Following Gujarati and Porter (2009, chapter 15.9), we can represent this decision as follows:

$$I_i^* = \beta_1 + \beta_2 * X_i \qquad (1)$$

The change to enter motherhood or not depends on an unobservable threshold I^*_i. However, what we *can* observe is a binary outcome I_i in the sample we analyse:

$$I_i = \begin{cases} 1 & if \quad I_i^* > 0 \\ 0 & otherwise \end{cases} \qquad (2)$$

For I_i=1 a woman has children, for I_i=0 not. Under the assumption of a normal distribution with same mean and variance, the probability conditional on X that I_i^* is less or equal than I_i can be written as:

$$P_i = \mathrm{P}(I_i^* \leq I_i) = \mathrm{F}(I_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\beta_1 + \beta_2 * X_i} e^{-z^2/2}\, dz \qquad (3)$$

[49] Due to non-concavity of positive fertility, negative binomial regressions could not be perfomed by STATA: likelihood calculations did not converge.

[50] It did better according to the AIC and BIC information criteria. We also tested the Baseline Probit against a negative binomial and ordinary Poisson: the Probit fit much better (Appendix 2, 3).

where F is a standard normal cumulative distribution function (CDF). P represents the probability that the event (fertility) occurs, given that the underlying utility for the woman to have a child is higher than to not have a child.

The results of a Probit estimation only become interpretable if we calculate the marginal effects. Then we can identify the marginal effect of a unit change in X on the probability that Y = 1 (a woman is mother), when e.g. all other variables are at their means.

ii. Second step: Probability of subsequent births, conditional on already being mother

The second step of the hurdle model represents the decision of a mother to have subsequent births when she already has one child, independent from the first decision about entering motherhood.
We used a zero truncated Poisson model, according to Winkelmann and Zimmermann (1995) who applied it for a similar data generating process: the probability of a job change of German immigrant workers. The zero truncated Poisson only takes into account the count data outcome of *positive* fertility (1, 2, 3 or 4 children).

Usually, one shall use "the zero-truncated form of a standard discrete distribution such as the Poisson or negative binomial distribution" for the second step of the hurdle model (Ridout et al., 1998, p. 4).

The first and second step of the hurdle model are specified as:

a) The probability to enter fertility, π_o
b) A distribution (defined on positive values of the dependent variable) for the number of children a woman has, once she is mother.

The probability to 'clear the hurdle', that is, to have non-zero outcomes, is $\pi_+ = 1 - \pi_o$.
For the process b), the probability distribution of the outcome Y can be written as follows:

$$\Pr(Y = y) = \begin{cases} \pi_o & y = 0 \\ \frac{(1-\pi_o)\, e^{-\lambda}\, \lambda^y}{(1-e^{\lambda})\, y!} & y > 0 \end{cases} \qquad (4)$$

If the same covariates will affect π_o and λ, then a standard Poisson model applies to the distribution (Ridout et al., 1998).

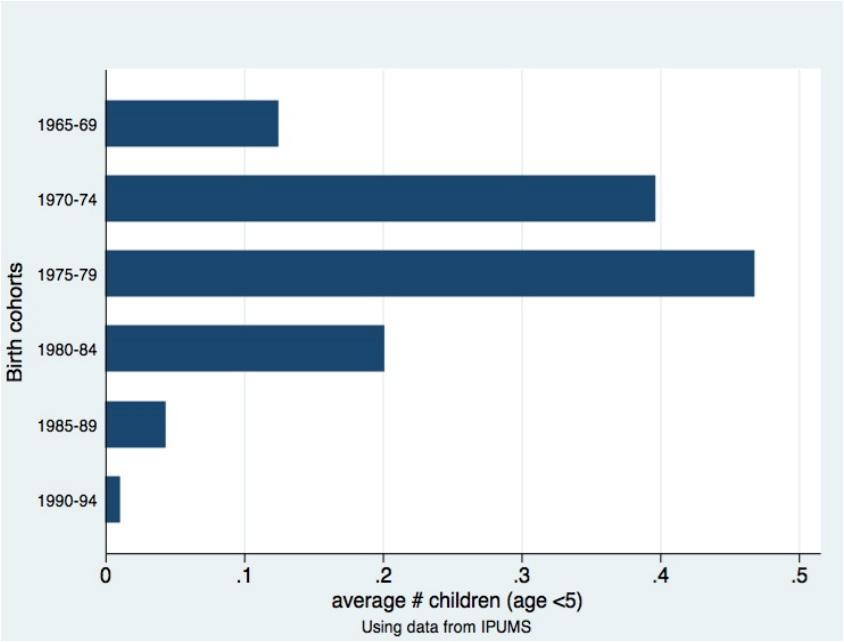
The models were estimated with maximum likelihood methods. In order to interpret the results, we again calculate the marginal effects of a unit change in X on the probability that a woman is mother of more than one child.

D. Descriptive statistics

Figure 9 shows the mean of the dependent variable distributed over the birth cohorts. We can see that younger cohorts (born 1980-94) have much smaller average fertility than the older women (born 1970-79). The older cohorts have either on average more children, or more women in these cohorts have children.

This can be put into relation with the development of mean age at first birth that increased between 1980 and 2004 (Figure 3), and is increasing again since the financial crisis of 2007. Women of the younger cohorts (1980-94) could actually be postponing anew their first child to even later ages than age 32.

Figure 9. Average fertility per birth cohorts



The mean fertility, excluding those women for whom *fertility* is 0, is highest in the 1970-74 cohort, followed by women born 1975-79 (see Annex 14). Polled mothers aged 19-31 in 2011 had on average not much more than 1.1 children. However, they still have not completed their lifetime fertility period and could have children at later stages. In this vein, the highest average number of children in the observed sample can be found at quite high ages (37-41).

It is also worth having a look at the sample regional distribution of the average number of children below age 5 (Table 4). There seems to be a slight regional unbalance. The macro-regions South, East and North are overrepresented compared to the Centre and Madrid, which is however due to the way the regional Dummies are created.[51] It seems like in the East women have on average the highest fertility, whilst in the Centre by far the lowest. Furthermore, the

[51] North, South and East contain more *comuninades* than the rest.

capital Madrid has the highest relative incidence of women with more than one child (that is, two, three or four). Women from the East sample have more unique children than in the other regions of Spain at the moment of the census.

Table 4. Sample regional distribution of average number of children (< age 5)

Region	*Mean number of children (including 0)*	*Percentage of 0 children*	*Percentage of 1 child*	*Percentage of 2 children*	*Percentage of 3 children*	*Percentage of 4 children*	*Number of obs*
South	0.224	82.01	13.78	4.06	0.15	0.01	107,234
North	0.217	82.45	13.53	3.86	0.15	0.00	68,135
East	0.243	80.51	14.93	4.40	0.16	0.00	118,409
Centre	0.199	84.10	12.04	3.37	0.14	none	59,914
Madrid (Reference)	0.224	82.55	12.7	4.51	0.23	0.01	52,552
National	*0.225*	*82.02*	*13.68*	*4.14*	*0.16*	*0.00*	*406,244*

Table using data from IPUMS

An interesting pattern is that the percentage of women that have progressed in motherhood in the four subsequent years after bearing their first child is extremely low: less than 4.5 per cent for Spain as a whole. 82.02 per cent of all women of the sample aged 19 to 45 that have not been mothers before have actually not given birth between 2007 and 2011.

If we have a look at descriptive statistics of all variables in the Annex 15, we can see that more than 82 per cent of the sample women are economically active. From the ones for which we have information about full-time/part-time work and unemployment, only 23 per cent work part-time, and around 31 per cent are unemployed. Of the husbands for whom we have information, 93.2 per cent were economically active, and of those only 5.9 per cent worked part-time in 2011, following the national trend.

Interestingly, only around 30 per cent of the sampled women aged 19-45 were married in 2011. The average educational attainment of those women (for whom we know the education of their husbands) is higher than for their spouses: 2.26 in contrast to 1.92. In fact, for the whole sample women are on average higher educated than men. Of all women, stunning 22 per cent had a university degree. The average household size is 2.89, meaning that the nuclear family consists more or less on average of two parents and one child.

Another finding is that there well exist out-of-the-wedlock births: 16,106 out of 275,940 unmarried women had children in 2011. However, mothers are much more frequent among the group of married women (54,695 out of 116,626). On average, significantly more women had an only child compared to two or more children.

V. Empirical results

The empirical analysis aims at answering the question of the influence of some socio-economic factors on the probability that woman a) enters motherhood in the first place, and b) that she proceeds to subsequent births, once she is already mother of one child. A special focus lies on the interrelation between economic activity and fertility
The two following sub-sections present the regression results of the econometric models described in section IV. C. i and ii.

A. First step: Probit

i. Probit regression

Table 5 contains the estimated coefficients for the first step Probit regression on motherhood participation. We estimated six different models including different sets of explanatory variables on the probability to become a mother.[52] The correlation matrix between the variables used in the models can be found in Appendix 1. Due to multicollinearity[53], some variables cannot be included in the model simultaneously but have to be introduced one by one (e.g. work fulltime and activity are highly correlated).

The reported pseudo R-squareds are not interpretable on their own, but only make sense when comparing models that use the same data set to predict the same outcome, like in our case (UCLA: What are pseudo R-squareds?). According to this measure, Model 5 fits best. Yet, the AIC and BIC information criteria, as well as the evaluation of the predicted outcomes across the models (Appendix 4), point commonly towards the Baseline and Model 6 as best fitting relatively to the others, and Model 5 as the second-*worst* fitting.

The main findings from the Probit regressions (Table 5) can be summarized as follows:
First of all, nearly all calculated regression coefficients are significant, with the exception of unemployment.[54]

We can learn from the models having economic activity as main explanatory variable that being economically active influences the probability that a woman decides to become a mother positively (highly significant). This, as expected, can be explained via the channel of economic security, which leads to the emancipation from the parents and is a prerequisite to enter parenthood.

[52] We corrected regression with the given personal sampling weights (p-weights), as advised by IPUMS international.
[53] Multicollinearity refers to highly significant correlation (>0.7) between two explanatory variables in a multivariate analysis. It indicates that the variables can predict each other mutually and contain very similar information. One should not include them simultaneously in a regression analysis, this could lead to biased coefficients of the independent variables.
[54] A high number of coefficients with high significance is common in Ibañez (2010) and Alonso-Antón et al. (2015); yet, we conducted robustness tests. The very significant p-values of the likelihood ratio chi-square tell us that our models are as a whole statistically significant, compared to a model without predictors.

Table 5. Probit regressions on binary outcome

	Baseline	Model 2	Model 3	Model 4	Model 5	Model 6
VARIABLE	*p*	*p2*	*p3*	*p4*	*p5*	*p6*
activity		0.572***	0.058***		0.354***	
		(0.014)	(0.019)		(0.016)	
age	0.030***	0.029***	0.007***		0.653***	0.021***
	(0.001)	(0.001)	(0.001)		(0.009)	(0.001)
married	1.343***	1.318***	0.681***	1.328***	1.193***	0.750***
	(0.010)	(0.008)	(0.010)	(0.008)	(0.008)	(0.012)
edu	0.020***	-0.010***	-0.018***		-0.051***	0.014***
	(0.004)	(0.003)	(0.004)		(0.003)	(0.005)
hhsize	0.359***	0.263***		0.289***	0.315***	
	(0.004)	(0.004)		(0.003)	(0.005)	
S	0.030**	0.065***	0.103***	0.044***	0.072***	0.048***
	(0.014)	(0.011)	(0.014)	(0.012)	(0.012)	(0.017)
N	-0.038***	-0.023*	0.002	-0.030**	-0.026**	-0.014
	(0.014)	(0.012)	(0.015)	(0.012)	(0.012)	(0.017)
E	0.102***	0.099***	0.040***	0.095***	0.103***	0.039***
	(0.013)	(0.011)	(0.013)	(0.011)	(0.011)	(0.015)
C	-0.052***	-0.034***	-0.012	-0.046***	-0.028**	-0.037**
	(0.015)	(0.013)	(0.016)	(0.013)	(0.013)	(0.019)
fullt	-0.209***					-0.410***
	(0.010)					(0.012)
edu_s			0.008**			0.014***
			(0.004)			(0.005)
activity_s			0.733***			0.754***
			(0.021)			(0.028)
unemp				0.013		
				(0.008)		
cohort				-0.133***		
				(0.003)		
uni				-0.059***		
				(0.008)		
agesq					-0.010***	
					(0.000)	
Constant	-3.393***	-3.713***	-1.678***	-1.990***	-13.439***	-1.959***
	(0.035)	(0.030)	(0.042)	(0.016)	(0.152)	(0.050)
Observations	222,119	392,563	143,055	322,956	392,563	100,240
Pseudo R2	0.268	0.263	0.0595	0.245	0.309	0.0849

Dependent variable: fert_3 (0 if # children below age 5 is 0, and 1 otherwise). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Pseudo R_2 are Mc Fadden's mirrors.

In Model 4, we see that being unemployed has also positive influence on motherhood. Yet, it seems to be unrelated to this decision since it is not statistically significant – in line with our expectations that this factor becomes important when it comes to subsequent births.

When a woman works full-time, the picture changes: it is significantly less probable that she will decide to become a mother. This goes in line with the analysis of the inflexible Spanish labour market that implies an incompatibility between full-time economic activity and motherhood.

In general, we can see that the older a woman is, the more probable it is that she will be mother. Yet, Model 5 shows that this effect is not linear: the coefficient for *agesq* is negative and significant. Hence, if women get older, the positive effect of age decreases. The negative coefficient of cohort in Model 4 delivers further evidence for the fact that the younger a woman is, the less probable it is she will be having a child under age five, as expected.

Living in a bigger household where many family members such as the grandparents live together with the core family influences the probability of motherhood positively, as we expected.

The coefficient for marriage is, in line with expectations, large and positive in all models: married women are still more likely to be mothers in present-day Spain. Underlying traditional values still seem to be largely influencing the maternity choices.

The variables of the educational attainment of women are not conclusive. One the one hand, there is a negative influence of higher education on the probability of motherhood in Model 2, 3 and 5 using *economic activity* as key explanatory. The completion of a university degree also seems to be negatively related to this decision. This adds to the hypothesis of incompatibility of work and motherhood, especially for those women with the highest level of education – which is however contrary to our expectations. On the other hand, when controlling for *fulltime* instead of *activity*, the educational coefficient changes to positive (Model 2, 6). This would go in line with the fact that at higher educational levels fulltime-working women are more likely to have a stable economic situation to affront the cost of becoming parents (Ibañez, 2010).

Model 3 looks at the spouse's education and economic activity. Both have, as expected, a positive and significant influence on the probability of the wife becoming a mother, since higher education and economic activity of the husband add to household economic security (as already explained). The educational attainment coefficient is however relatively small.

The regional controls are not always significant, but living in the South or East of Spain makes it significantly more probable to be a mother than living in Madrid. In contrast, coming from North or Central Spain makes it less probable with respect to Madrid. We would have expected that both the East (with Barcelona) and Madrid would be the regions where it was most probable that women have a child aged 0-4, since they stand for the most developed and advanced regions in Spain, where the cultural change of the SDT is most advanced and motherhood and work should be combinable best. In fact, the East has the highest coefficient of all regions. The South's positive coefficient can be explained by the cultural traditionalism with a strong centrality of the motherhood role for women.

The predicted constants are all significant and negative and tell us the predicted z-score of the probability if all other controls are zero. Yet, interpreting them is not very insightful.

We are aware that the models suffer from Omitted Variable Bias, since we lack important information that might correlate with the choice of becoming a mother (e.g. income or type of occupation). We should then assume that the calculated coefficients are biased due to distributional misspecification (since the error is not normally distributed), but still consistent.[55] Yet, the bias converges towards zero (Williams, 2015).
Additionally, we controlled that multicollinearity is not a problem in our regressions.[56]

In order to control for problems that can arise due to model-misspecification and to check for the robustness of the results, we performed three 'robust models' (Appendix 5). Again, all coefficients are highly significant.
In Robust 1, we used S2 (South excluding Ceuta and Melilla) and CeuMel (Dummy for the autonomous cities Ceuta and Melilla) instead of S. The coefficient of N becomes insignificant. The signs of all variables remain unchanged, while the magnitudes vary somewhat from the Baseline model (e.g. S2 (versus S), age and activity). CeuMel is positive but insignificant, for which we can conclude that CeuMel and the South can be treated as one homogeneous region.

In Robust 2 and 3, we used regional interaction variables in order to see whether there are not only regional level differences, but also slope differences for different explanatory variables.[57] From these models, we can conclude that in all regions being married makes it more likely to be mother than in Madrid, most so in the Centre. Having higher education influences motherhood highest in Madrid, and least in the Centre and South. From Robust 2, we infer that in the South and East being economically active influences the likelihood to be a mother more than in Madrid – most in the South, against our expectations. In Robust 3, we see that full-time work reduces the probability of motherhood in all regions with respect to Madrid (quite importantly), most so in the Centre and least in the South.
Madrid seems to be the leader in the combinability of full-time work and motherhood and the place where cultural traditionalism matters least, as expected.

However, a drawback from interaction terms in nonlinear models is that the interpretation of their marginal effects does not follow the same logic as for linear models. They cannot be interpreted straightforwardly, which is why we did not calculate them (for more details see Norton et al., 2004, p. 154).

The interpretation of the coefficients of a Probit regression is limited. Their signs inform about the direction of influence on the predicted probability of motherhood if the predictor increases. Yet, the effect varies depending on the level evaluation of the explanatory variable.

[55] For example, the income from work might increase the probability to become pregnant: the bias caused by not including this variable into the regression might be upwards.
[56] We used for this the Stata "vif" (variance inflation factors) command.
[57] Due to multicollinearity, one should not regress interaction terms alongside the single interacted variables.

ii. Marginal effects

A common practice in this case is the calculation of marginal effects, which are an approximation of how much the dependent variable is expected to increase or decrease for a unit change in an explanatory variable. Hence, the effect has to be in interpreted on an additive scale (Buis, 2010). Table 6 contains the marginal effects of the Baseline Model and Model 6.

Table 6. Marginal effects Probit

VARIABLE	*Baseline*	*Model 6*
fullt	-0.048***	-0.157***
	(0.002)	(0.005)
age	0.007***	0.008***
	(0.000)	(0.000)
married	0.311***	0.287***
	(0.002)	(0.004)
edu	0.005***	0.005***
	(0.001)	(0.002)
S	0.007**	0.018***
	(0.003)	(0.006)
N	-0.009***	-0.005
	(0.003)	(0.006)
E	0.024***	0.015***
	(0.003)	(0.006)
C	-0.012***	-0.014**
	(0.004)	(0.007)
hhsize	0.083***	
	(0.001)	
edu_s		0.005***
		(0.002)
activity_s		0.289***
		(0.011)
Observations	222,119	100,240

Dependent variable: fert_3 (0 if # children below age 5 is 0, and 1 otherwise).
Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Note: This table contains the partial marginal effects at the means of the vector of the other explanatory variables.

We notice that nearly all marginal effects are highly significant. It is positive that they do not vary in sign and only little in magnitude across the two models, when variables are introduced or left out.

The increase in age by one year increases the probability to be a mother of a child aged 0-4 by 0.007 (0.008). A one-unit increase in the household size increases the probability by 0.083. The marginal effect for education is now clearly positive: with a one-level increase in educational attainment the probability of motherhood increases by 0.005. Being married, versus not being in a formal union, has the highest marginal effect: it increases the probability by 0.311 (0.287). Working full-time, versus part-time, decreases the probability of motherhood by 0.048, and even

by 0.157 when we control additionally for the spouse's education and activity, indicating that in higher-educated economically active partnerships full-time-working women are less prone to have children. However, the spouses' isolated influence of activity and higher education *increase* the motherhood probability by 0.005 and 0.289 respectively.

Concerning the regional controls, we can now see that being a women from the South or East increases the probability of being a mother by 0.07 (0.018) and 0.024 (0.015), with respect to Madrid. Coming from the North versus Madrid decreases the probability. Motherhood is the least probable in the very traditional, rural and scarcely populated Centre. Yet, the regional coefficients are not as overwhelmingly large as we would have wished.

B. Second Step: Zero-truncated Poisson (Ztp)

i. Ztp regression

We remind the reader that the second step of the hurdle model is interested in which factors influence the probability to have more than one child. The regressions have as dependent variable the positive values of fertility. As expected, the influence of some explanatory variables became insignificant, while other factors now turn out to be significant. No longer all controls are significant in this second step.

Table 7 contains the three best-fitting Ztp regressions according to the AIC and BIC information criteria.[58] A good sign is that the size and sign of the coefficients do not vary importantly throughout the models.

Appendix 7 displays the pairwise correlation matrix between the model's predicted outcomes and the observed values of the dependent variable. In Appendix 9 all other regressions (including robustness checks) can be examined.

The main findings from Table 7 can be summarized as follows:
Age does not influence the probability to progress in childbearing. Being married is positively related to the latter, but the coefficient is very small (0.002) and even insignificant in Model 6; we can think of the traditionalist social pressure on couples vanishing after the first child is born.

[58] We checked Ztp is versus ordinary Poisson regressions. Appendix 6 contains the model fit statistics indicating that Ztp is preferable.

Table 7. Ztp regressions on positive fertility

VARIABLES	Model 1 *(ztp1)*	Model 6 *(ztp6)*	Model 9 *(ztp9)*
age	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
onech	6.983***	6.983***	6.983***
	(0.359)	(0.359)	(0.359)
twoch	7.490***	7.490***	7.491***
	(0.380)	(0.380)	(0.379)
hhsize	0.058***	0.059***	0.059***
	(0.020)	(0.020)	(0.020)
fullt	-0.002	-0.002	
	(0.002)	(0.002)	
married	0.002*	0.002	
	(0.001)	(0.001)	
edu		0.002**	
		(0.001)	
uni	0.005**		
	(0.002)		
S	-0.001		
	(0.003)		
N	-0.003	-0.003	
	(0.003)	(0.003)	
E	-0.002	-0.002	
	(0.002)	(0.002)	
C	-0.002	-0.002	
	(0.002)	(0.002)	
S2		-0.001	
		(0.003)	
CeuMel		-0.008	
		(0.006)	
Sxm			-0.003
			(0.002)
Cxm			-0.002
			(0.002)
Exm			-0.000
			(0.002)
Nxm			-0.002
			(0.002)
eduxC			0.000
			(0.001)
eduxE			0.000
			(0.001)
eduxN			0.002
			(0.002)

Table 7. Ztp regressions on positive fertility (continued)

VARIABLES	Model (1)	Model (6)	Model (9)
eduxS			0.002
			(0.001)
fulltxC			-0.000
			(0.001)
fulltxE			0.000
			(0.002)
fulltxN			-0.009
			(0.006)
fulltxS			-0.001
			(0.002)
Constant	-6.755***	-6.757***	-6.758***
	(0.281)	(0.280)	(0.280)
Observations	47,978	47,978	47,978

Dependent variable: positive values of fert_3. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Education is positively and significantly influencing the probability to become a mother of more than one child; the coefficient of having a university degree is even 2.5 times the size of *edu*, indicating that when considering working women, it is more likely that the highly educated ones progress in motherhood (Model 1, 6), which confirms Ibañez' (2010) findings about the special role of education for Spanish women's progression in childbearing.

The effect of household size is positive. We can imagine that for working couples with more than one child is becomes especially tricky to satisfy adequate childcare needs, for which the family's support is crucial.

What was not expected is that when looking at working women, having a fulltime job is not influencing the probability to progress in motherhood significantly (even if the sign is negative as thought). On the other hand, being economically active has a positive influence (Model 3, Appendix 9).
Yet, it seems that it is rather unemployment (a sub-category of economic activity), and not work, that is significantly influencing the probability of motherhood: once we control for education and activity of the husband, being unemployed has a quite large and highly significant positive influence on the probability to progress in giving birth (Model 5, Appendix 9). Furthermore, fulltime work becomes negatively related to the probability of birth progression (Model 4, Appendix 9). This tells us that unemployed women are more likely to progress in motherhood if the husband works and the higher his education is. In contrast to that, fulltime working women with a working husband are less expected to progress in childbearing with his education level raising. These findings are similar to Ibañez (2010) and can be explained by the segmented nature of the Spanish labour market.

Already being mother of one or two children does significantly increment the probability of being pregnant again. The coefficient of *twoch* is actually larger than for *onech*, which is puzzling. We would have expected that, given the scarcity (in reality and in the sample) of multiple mothers, the fact that one has already got a child would reduce the probability to progress in giving birth.

With respect to the regional Dummies, we see that for all regions the probability of progression is lower than in Madrid – but the effect is unexpectedly insignificant. The same occurs if we further look at the regional interaction effects (Model 9).

In the robustness regressions (Appendix 7), we see that the sizes and signs do not vary greatly when leaving out some crucial variables or introducing other robustness check variables. The significance does however change: some regional controls and age become significant in Model 3, 4, 5 and 7.

ii. Marginal effects

One convenient way to interpret the Ztp coefficients is, as in the Probit regressions above, to calculate the marginal effects. They can be observed for Model 1 and 6 in Table 8.[59]

The previous amount of children a mother already has as well as the household size have the only significant coefficients. The fact of having already two children is increasing the probability of another birth by 0.04934. This is 0.003 higher than if having one child (0.046). Furthermore, one person more in the household increases the probability of repeated motherhood by around 0.00038. Overall, the sizes of the effects are quite low.

The signs of the regional controls indicate that in every region it is less likely that a mother progresses to higher births than in Madrid. This would confirm the assumption that the capital is the place where it is more likely that women get anew pregnant after having already one child. Yet, the regional coefficients are not significant, against our expectations.

[59] Marginal effects for the interaction effects were not calculated, as explained above.

Table 8. Marginal effects Ztp

	Model (1)	Model (6)
VARIABLES	*marg_ztp1*	*marg_ztp6*
age	0.00000	0.00000
	(0.00000)	(0.00000)
onech	0.046***	0.046***
	(0.00980)	(0.00980)
twoch	0.04934***	0.04934***
	(0.01055)	(0.01055)
hhsize	0.00038*	0.00039*
	(0.00023)	(0.00023)
fullt	-0.00001	-0.00001
	(0.00001)	(0.00001)
married	0.00001	0.00001
	(0.00001)	(0.00001)
edu		0.00001
		(0.00001)
uni	0.00003	
	(0.00002)	
S	-0.00000	
	(0.00002)	
N	-0.00002	-0.00002
	(0.00002)	(0.00002)
E	-0.00001	-0.00001
	(0.00002)	(0.00002)
C	-0.00001	-0.00002
	(0.00002)	(0.00002)
S2		-0.00000
		(0.00002)
CeuMel		-0.00005
		(0.00004)
Observations	47,978	47,978

Dependent variable: positive values of fert_3. Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Note: This table contains the partial marginal effects at the means of the vector of the other explanatory variables.

VI. Conclusion

The persistent low fertility in Spain has till date remained a puzzling phenomenon. In this thesis, we tried to improve our understanding of this topic by analysing fertility from a regional perspective. We were interested to see if underlying regional patterns drive the national fertility level. More specifically, we analysed how economic activity and related factors, such as fulltime work or unemployment, affect the motherhood decisions of contemporary Spanish women. The first section introduced the current social position of women in Spanish society in a historic and regional context. We then analysed the national as well as regional patterns of fertility and female labour force participation, and discussed a theoretic framework of the interdependencies of fertility and economic activity for women. Our literature review presented some important work that has already dealt with this topic. In our empirical analysis, we addressed the question of which socio-economic and regional factors determine the likelihood to enter motherhood and the transition to subsequent births in Spain. We applied a micro-level two-step discrete choice model for the two independent decisions of fertility: becoming mother in the first place, and proceeding to give birth to more children.

The regional analysis shows that between 1970 and 1990 the TFR of all 19 Spanish regions decreased to *lowest-low* levels following a similar trend. During the same period of time, the FLFP in all regions increased significantly. However, these developments took place at different paces and the levels of TFR and FLFP are today quite heterogeneous across Spain. Some regions have experienced unexpected trend reversals in fertility: as a matter of fact, those regions with traditionally very high fertility have today the lowest TFRs.

We presented a variety of approaches to provide explanations of the drop to and persistence in sub-replacement fertility as well as how economic activity affects fertility. We established a theoretical framework including the economic dimensions, value systems, personal traits and institutional dimensions. We integrated these dimension in our empirical approach. Through our introductory historic analysis, we saw that women all over Spain have increased their influence in the public life (labour market, educational attainments, presence in higher education) at a stunning pace since the transition to democracy, but are nowadays confronted by the reality of hardly being able to combine their new role of being an educated working woman with motherhood. The persisting cultural reality of gender inequalities, both in the household and professional sphere, can help to explain this *role incompatibility*. The same can be said of institutional failures and labour market rigidities.

This *role incompatibility* has – besides other factors attributable to the Second Demographic Transition – led women to choose work prior to motherhood. Nowadays' women have the lowest average number of children in history, since only around 30 per cent of the sample progress over bearing one child, and even only 661 over two. In contrast to that, 82.4 per cent of the polled women are economically active. In this context, we developed a micro-level multivariate probability model on the factors that influence the independent decision to enter, or not, motherhood, and to progress in giving birth to more children. Given the cross-sectional and limited nature of the data, we were not able to test for the influence of gender equalities in the

household as well as in the professional sphere; also institutional failures and the effect of timing of birth could not be addressed. We are thus aware that our results represent a modest alternative approach to the phenomenon of fertility, while a better underlying dataset could have improved the explanatory power of our analysis.

For the empirical work, we followed a similar approach to Ibañez (2010), Alonso-Antón et al. (2015) and White et al. (2007) and could confirm some of the findings of these authors as well as add new evidence to the phenomenon of persisting sub-replacement fertility in Spain.

Regarding the first question addressed in this thesis on the determinants a woman's choice to enter motherhood, we can conclude the following: the economic activity seems to have a positive impact on the choice of motherhood, confirming the economic channel. Yet, fulltime work represents a hindering factor, which seems to confirm our suspicions of the *incompatibility* between full working activity and motherhood. Catholic traditionalism still seems to have a significant influence on fertility, shown by the high positive effect of being married on motherhood. The significance of the household size suggests that there exists an institutional failure of sufficient/appropriate childcare so that the family has to support the parents in this duty. Alternatively, we could as well conclude that the *familistic* cultural system is still quite present in the Spanish society. We confirmed Ibañez (2010) in that education plays a significant role in the first fertility decision. The findings on the regional fertility behaviour confirming the influence of culture are an important contribution of this paper: in Madrid, the capital, backwards-orientated traditions matter least, education most and work and motherhood are combinable best, with respect to the rest of Spain. We found evidence for Madrid being a *forerunner* region with respect to fertility, adding evidence to Arpino and Tavares' (2013) findings about Madrid's precursor role in value change.

Concerning the second research question on what determines a mother's decision to go on in motherhood, we were able to show that other factors than those influencing the decision to have children matter. In fact, solely the household size, the mother's educational attainment and the number of children she already has are significant. Hence, the availability of other family members for childcare positively affects the decision to have more children. The same holds for the higher educated a mother is (confirming Ibañez' (2010) discussion on the high fragmentation of the Spanish female labour market, where stable employment only is available for highly educated women). A finding that needs further clarification is that having already one or two children makes it more likely to proceed to subsequent births. This stands in contrast to Ibañez' (2010), who found that having one or more children reduces the probability to have subsequent births. Also, it goes against the logic that follows the reality of very few mothers of several children, who are outnumbered by single-child mothers in today's Spain. We could interpret this finding as that for all women who already have taken the step to be a mother, becoming pregnant another time is less of a hurdle. This, in turn, would mean that the low fertility rate of Spain might rather be triggered by the amount of women who do not have children at all – which in fact represents the great majority of the polled women. Furthermore, the regional effects are explained away by other factors: across all regions, the same factors influence the progression in

giving birth with the same magnitude as in Madrid. Thus, there is no evidence for the capital being the leading place where mothers are most prone to have more than one child, against our initial assumptions. Women in Madrid are thus more likely to be mothers, but women are equally likely to progress to subsequent births as in Madrid.

Analysing fertility choices with a two-step decision model represents an alternative, but in our opinion more adequate, methodological framework. Additionally, a major contribution of this paper is to show that cultures and traditions do significantly matter when it comes to the decision of becoming a mother. Each region of Spain has its very own specificities and historical background, which still today seems to influence major aspects of society, such as the position of women and fertility. On the other hand, we presented some evidence that for the progression of childbearing it is rather the institutional surrounding which seems to be decisive: as Ibañez (2010) pointed out – and as often reiterated in surveys – public and private childcare are far from covering the specific needs of women to be able to combine work and motherhood in modern Spain.

After all, it seems that for the analysed sample of actual mothers in Spain, only a very small number progress to have more children in the subsequent four years after they gave birth. Furthermore, only a very small percentage of the, in 2011, women aged 19-45 had actually become mothers between 2007 and 2011. This timespan is quite peculiar, since it includes the aftermaths of the financial crisis that had a far-reaching impact on the Spanish society (sky-rocketing (youth) unemployment, bankruptcies, austerity policies etc.). This opens up space for a discussion of the possibly temporary nature of the downward trend in fertility triggered by the financial crisis and the following economic recession (see for this Sobotka et al., 2011; Goldstein et al., 2013; Larañaga et al., 2014; Örsal and Goldstein, 2010).

For future research, it might thus be of interest to further carve out how exactly cultural heritage and institutional surrounding influence fertility behaviour. It would furthermore represent an important advancement to investigate the reciprocal influence between FLFP and TFR in a multi-step decisional choice model (Antonio-Antón et al. (2015) already did this for a simple one-step decision model), in order to control for the possible endogeneity of FLFP.

References

Alonso, F. G. (2011). Los estudios sobre el descenso histórico de la fecundidad en España y sus pautas territoriales: un estado de la cuestión. *Biblio 3w: revista bibliográfica de geografía y ciencias sociales*, *16*.

Alonso-Antón, A., Fernández-Sainz, A., and Rincón-Diez, V. (2015). Análisis de la actividad femenina y la fecundidad en España mediante modelos de elección discreta. *Lecturas de Economía*, *82*, 127-157.

Andersen, T., Jensen, P. and Skovsgaard, C. (2016). The heavy plough and the agricultural revolution in medieval Europe. Journal of Development Economics Vol. 118, pp: 133-149.

Arpino, B., and Tavares, L. P. (2013). Fertility and values in Italy and Spain: A look at regional differences within the European context. *Population Review*, *52*(1).

Bloom, D. E., and Sousa-Poza, A. (2010). Introduction to Special Issue of the European Journal of Population:'Economic Consequences of Low Fertility in Europe'. *European Journal of Population/Revue européenne de Démographie, 26(2), 127-139.*

Bongaarts, J., and Feeney, G. (1998). On the quantum and tempo of fertility. *Population and development review*, 271-291.

Boserup, E. (1970). Woman's role in Economic Development. Earthscan, Trowbridge.

Budig, M. J., and England, P. (2001). The wage penalty for motherhood. *American sociological review*, 204-225.

Buis, M. L. (2010). Stata tip 87: Interpretation of interactions in non-linear models. *The stata journal*, *10*(2), 305-308.

Cantero Rosales, M. Á. (2007). De "perfecta casada" a "ángel del hogar" o la construcción del arquetipo femenino en el XIX. *Tonos Digital*, *14*(0).

Castro-Martín, T., and Rosero-Bixby, L. (2011). Maternidades y fronteras. La fecundidad de las mujeres inmigrantes en España. *Revista internacional de sociología*, *69(M1), 105-138.*

Coale, A. J. (1989). *Demographic transition* (pp. 16-23). *Palgrave Macmillan UK.*

Del Boca, D. (2002). The effect of child care and part time opportunities on participation and fertility decisions in Italy. *Journal of Population Economics*, *15*(3), 549-573.

Dribe, M., Hacker, J. D., and Scalone, F. (2014). The impact of socio-economic status on net fertility during the historical fertility decline: A comparative analysis of Canada, Iceland, Sweden, Norway, and the USA. *Population studies*, *68(2), 135-149.*

Esping-Andersen, G. (Ed.). (2013). El déficit de natalidad en Europa: la singularidad del caso español. *Obra Social" La Caixa".*

Fernández Valencia, A. (2006). La educación de las niñas: ideas, proyectos y realidades. *Historia de las mujeres en España y América Latina, 3, 427-53.*

Flecha, C. (2006). Mujeres en institutos y universidades. *Historia de las mujeres en España y América Latina. Del siglo xix a los umbrales del xx, Madrid*, 3, 455-503.

Giuliano, P. (2014). The Role of Women in Society: from Preindustrial to Modern Times. *CESifo Economic Studies.*

Goldstein, J., Kreyenfeld, M., Jasilioniene, A., and Örsal, D. D. K. (2013). Fertility reactions to the" Great Recession" in Europe: Recent evidence from order-specific data. *Demographic Research*, *29*, 85-104.

Gujarati, D. N., and Porter, D. (2009). *Basic Econometrics Mc Graw-Hill International Edition.*

Guner, N., Kaya, E., and Sánchez-Marcos, V. (2014). Gender gaps in Spain: policies and outcomes over the last three decades. *SERIES, 5(1), 61-103.*

Ibañez, M. (2010). Fertility, Female Participation in Employment and Reconciliation Policies in Spain. *REC-WP 13/2010 Working Papers on the Reconciliation of Work and Welfare in Europe RECWOWE Publication, Dissemination and Dialogue Centre, Edinburgh.*

Lesthaeghe, R., and van de Kaa, D. J. (1986). Twee demografische transities. *Bevolking: groei en krimp, 9-24.*

Lutz, W., Skirbekk, V., & Testa, M. R. (2006). The low-fertility trap hypothesis: Forces that may lead to further postponement and fewer births in Europe. *Vienna yearbook of population research, 167-192.*

Kohler, H. P., Billari, F. C., and Ortega, J. A. (2002). The emergence of lowest-low fertility in Europe during the 1990s. *Population and development review, 28(4), 641-680.*

Kohler, H. P., and Ortega, J. A. (2002). Tempo-adjusted period parity progression measures, fertility postponement and completed cohort fertility. *Demographic research*, *6*, 91-144.

Matysiak, A. (2011). Interdependencies between fertility and women's labour supply. *Springer Science and Business Media.*

McDonald, P. (2000). Gender equity in theories of fertility transition. *Population and development review*, *26*(3), 427-439.

- (2006). Low fertility and the state: The efficacy of policy. *Population and development review*, *32*(3), 485-510.

Mishra, V., Nielsen, I., and Smyth, R. (2010a). On the relationship between female labour force participation and fertility in G7 countries: evidence from panel cointegration and Granger causality. *Empirical Economics*, *38*(2), 361-372

Mishra, V., and Smyth, R. (2010b). Female labor force participation and total fertility rates in the OECD: New evidence from panel cointegration and Granger causality testing. *Journal of Economics and Business*, *62*(1), 48-64.

Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of econometrics*, *33*(3), 341-365.

Norton, E. C., Wang, H., and Ai, C. (2004). Computing interaction effects and standard errors in logit and probit models. *Stata Journal*, *4*, 154-167.

Núñez-Pérez, M., (1993). Evolución de le situación laboral de las Mujeres en España durante la Segunda República (1931-36). *Cuadernos de Relaciones Laborales N°3. Edit. Complutense Madrid.*

Örsal, D. D., and Goldstein, J. R. (2010). The increasing importance of economic conditions on fertility. In *annual meetings of the Population Association of America. Dallas, Texas, April* (pp. 15-17).

Ridout, M., Demétrio, C. G., and Hinde, J. (1998). Models for count data with many zeros. In *Proceedings of the XIXth international biometric conference* (Vol. 19, pp. 179-192).

Roa-Venegas, J. M. (2006). Scholastic Achievement and the Diglossic Situation in a Sample of Primary-School Students in Ceuta. *Revista Electrónica de Investigación Educativa, 8(1).* Available Online: http://redie.ens.uabc.mx/vol8no1/contents-roa.html [Accessed 7th of May 2016]

De Lecea, T., (2006). Las mujeres y la iglesia, in Morant, I. (dir.) Historia de las mujeres en España y América Latina: Del siglo XX a los umbrales del XXI. *2nd Edition, volume IV, Ediciones Cátedra.*

Scott Long, J. (1997). Regression models for categorical and limited dependent variables. *Advanced quantitative techniques in the social sciences*, *7*.

Smallwood, S., and Chamberlain, J. (2005). Replacement fertility, what has it been and what does it mean?. *Population Trends*, *119*, 16-27.

Sobotka, T., Skirbekk, V., and Philipov, D. (2011). Economic recession and fertility in the developed world. *Population and development review*, *37*(2), 267-306.

Torralbo-Ruíz, Á. (2011). El rol de la mujer en el Código Civil: especial referencia a los efectos personales del matrimonio. *Master Thesis at the University of Salamanca.*

Van de Kaa, D.J. (1987). Europe's Second Demographic Transition. *Population Bulletin 42.1:1-57.*

- (2004). Is the Second Demographic Transition a Useful Research Concept: Questions and Answers. *Vienna Yearbook of Population Research 2004: 4-10.*

White, M. J., Gabrielli, G., Bernardi, L., Kertzer, D. I., and Perra, S. (2007). Regional context and fertility in contemporary Italy. In *Annual Meeting of the Population Association of America, New York, USA*.

Williams, R. (2015). Comparing Logit and Probit Coefficients Between Nested Models (Old Version). *University of Notre Dame.* Available Online: https://www3.nd.edu/~rwilliam/stats3/Nested03.pdf [Accessed 18th May 2016]

Winkelmann, R., and Zimmermann, K. F. (1995). Recent developments in count data modelling: theory and application. *Journal of economic surveys*, *9*(1), 1-24.

COMISIONES OBRERAS (2004). ¿Invisibles? Mujeres, trabajo y sindicalismo en España. *Edita la Confederación Sindical de Comisiones Obreras, Madrid.*

Web Sources:

Allón, D. and Díaz, P. (2010). Los matrimonios civiles desbancan a los religiosos, Público, 1 November, Available Online: http://www.publico.es/espana/matrimonios-civiles-desbancan-religiosos.html [Accessed 7th of May 2016]

CIS (Centro de Investigaciones Sociológicas) (2012). Barómetro autonómico de septiembre-octubre de 2012 [pdf] Available at: http://www.cis.es/cis/export/sites/default/-Archivos/Marginales/2940_2959/2956/Es2956_mapa_pdf.html [Accessed 7th of May 2016]

Constitución Española de 1931. Available Online: https://es.wikisource.org/wiki/Constituci%C3%B3n_espa%C3%B1ola_de_1931 [Accessed 20 April 2016]

El Diario (2016). El Gobierno sigue retrasando el permiso de paternidad de un mes hasta el 2017, El Diario, 4 August, Available Online: http://www.eldiario.es/economia/Gobierno-sigue-retrasando-permiso-paternidad_0_416458828.html [Accessed 2nd of May 2016]

EUROSTAT (2016). File: Total fertility rate, 1960–2014, 14 March, EUROSTAT, Available Online: http://ec.europa.eu/eurostat/statistics-explained/index.php/File:Total_fertility_rate,_1960%E2%80%932014_%28live_births_per_woman%29_YB16.png [Accessed May 8th 2016]

Fuero del Trabajo. Available Online: https://es.wikisource.org/wiki/Fuero_del_Trabajo [Accessed 1 May 2016]

Fundación Centro de Estudios Andaluces (2007). 25 Años de Autonomía en Andalucía. [e-book] Sevilla: Centro de Estudios Anadaluces. Available at: Google Books: books.google.com [Accessed May 6th 2016]

INE (2016). Mujeres y hombres en España. Avalaiblae online: http://www.ine.es/ss/Satellite?L=es_ES&c=INESeccion_C&cid=1259925408327&p=1254735110672&pagename=ProductosYServicios%2FPYSLayout&param1=PYSDetalle&param3=1259924822888. /[Accessed 23 April 2016]

INE (2015a). Fenómenos demográficos, 4 December, Available Online: http://www.ine.es/jaxi/menu.do?type=pcaxis&path=%2Ft20%2Fp277&file=inebase&L=0 [Accessed 1 May 2016]

INE (2015b). Cuidado de personas dependientes (niños o adultos enfermos, incapacitados o mayores). Available Online: http://www.ine.es/ss/Satellite?L=es_ES&c=INESeccion_C&cid=1259925472720&p=1254735110672&pagename=ProductosYServicios%2FPYSLayout&param3=1259926137287 [Accessed 2 May 2016]

INM (2016). Derechos de conciliación de la vida familiar y laboral. Instituto de la Mujer, Available Online: http://www.inmujer.gob.es/conoceDerechos/empleo/conciliacion/home.htm [retrieved 2 May 2016]

IPF (Instituto de Política Familiar) (2016). Informe de Evolución de la Familia en España 2016, 11 May, Available Online: http://www.ipfe.org/Espa%C3%B1a/Documentos/IPF [retrieved 9 May 2016]

Lafuente, P. (2015). 82° aniversario del voto femenino. Clara Campoamor y las mujeres del S. XX, El Diario, 15 November, Available Online: http://www.eldiario.es/aragon/elprismatico/aniversario-femenino-Clara-Campoamor-SXX_6_452564747.html [Accessed 7 May 2016]

La Información (2013). En España se registran casi el doble de uniones civiles que religiosas, 2 February, La Informacíon, Available Online: http://noticias.lainformacion.com/asuntos-sociales/familia/en-espana-se-registran-casi-el-doble-de-uniones-civiles-que-religiosas_OiUREcp6wLUG4QdA5n2Dw5/ [Accessed 8 May 2016]

Ley Orgánica 9/1985, de 5 de julio, de reforma del artículo 417 bis del Código Penal. Available Online: http://www.boe.es/buscar/doc.php?id=BOE-A-1985-14138 [Accessed 23 April 2016]

Ley Orgánica 3/2007, de 22 de marzo, para la igualdad efectiva de mujeres y hombres. Available Online: http://www.boe.es/buscar/doc.php?id=BOE-A-2007-6115 [Accessed 23 April 2016]

Ley Orgánica 2/2010 de salud sexual y reproductiva y de la interrupción voluntaria del embarazo. Available Online: http://www.boe.es/buscar/doc.php?id=BOE-A-2010-3514 [Accessed 23 April 2016]

SCB (Statistiska Centralbyrån) (2014). Women and men in Sweden 2014: 30 years of statistics on gender equality, SCB, 3 July, Available Online: http://www.scb.se/en_/Finding-statistics/Statistics-by-subject-area/Living-conditions/Gender-statistics/Gender-statistics/Aktuell-Pong/12259/Behallare-for-Press/375659/ [Accessed 29 April 2016]

Statista (2016). The 20 countries with the lowest fertility rates in 2014, January, http://www.statista.com/statistics/268083/countries-with-the-lowest-fertility-rates/ [Accessed 8 May 2016]

UCLA: Statistical Consulting Group. What are pseudo R-squareds? Available Online: http://www.ats.ucla.edu/stat/mult_pkg/faq/general/Psuedo_RSquareds.htm [Accessed 18 May 2016]

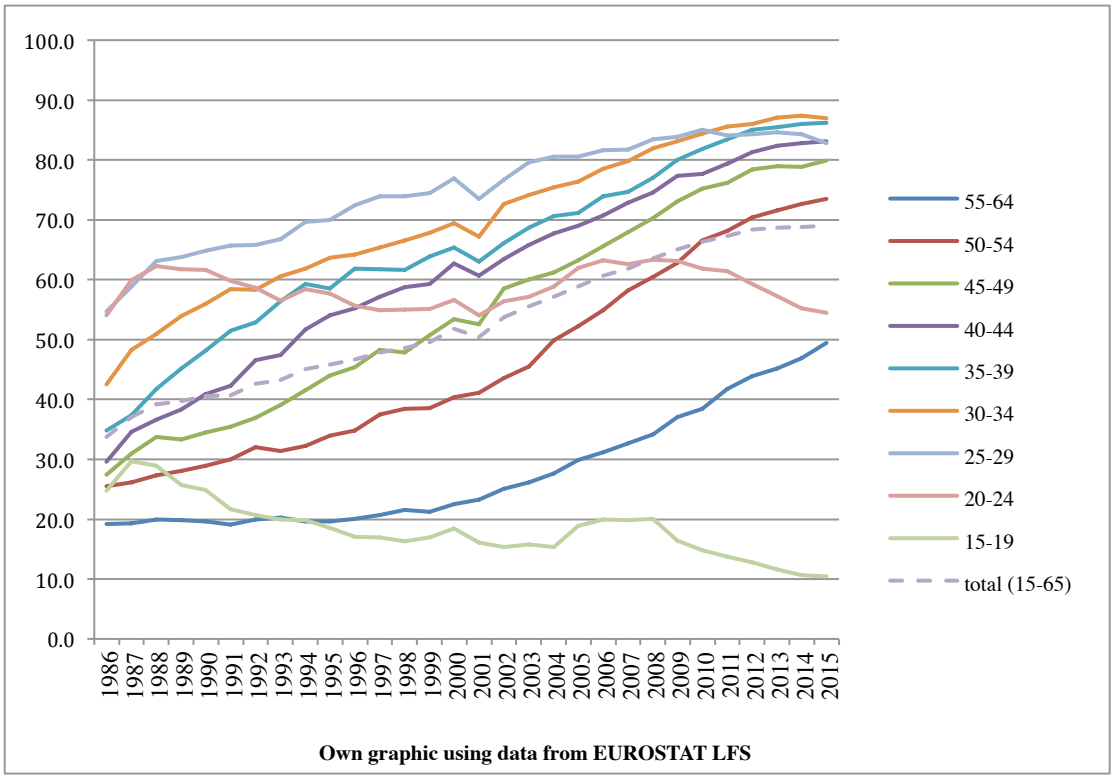
Vadillo, J. (2014). Conquista y esperanzas de la II República española. La historia de la II Republica es la historia de una ruptura con el periodo anterior y de avances sociales sin precedentes en España. Diagonal, 14 April, Available Online: https://www.diagonalperiodico.net/saberes/22582-conquista-y-esperanzas-la-segunda-republica-espanola.html [Accessed April 26th 2016]

World Bank (1994). Averting the old age crisis : policies to protect the old and promote growth. Washington DC; World Bank. Available Online: http://documents.worldbank.org/curated/en/1994/09/698030/averting-old-age-crisis-policies-protect-old-promote-growth [Accessed 1 May 2016]

World Bank (2016). Net Migration, Available Online: http://databank.bancomundial.org/data/reports.aspx?source=2&country=&series=SM.POP.NETM&period= [Accessed 8 May 2016]

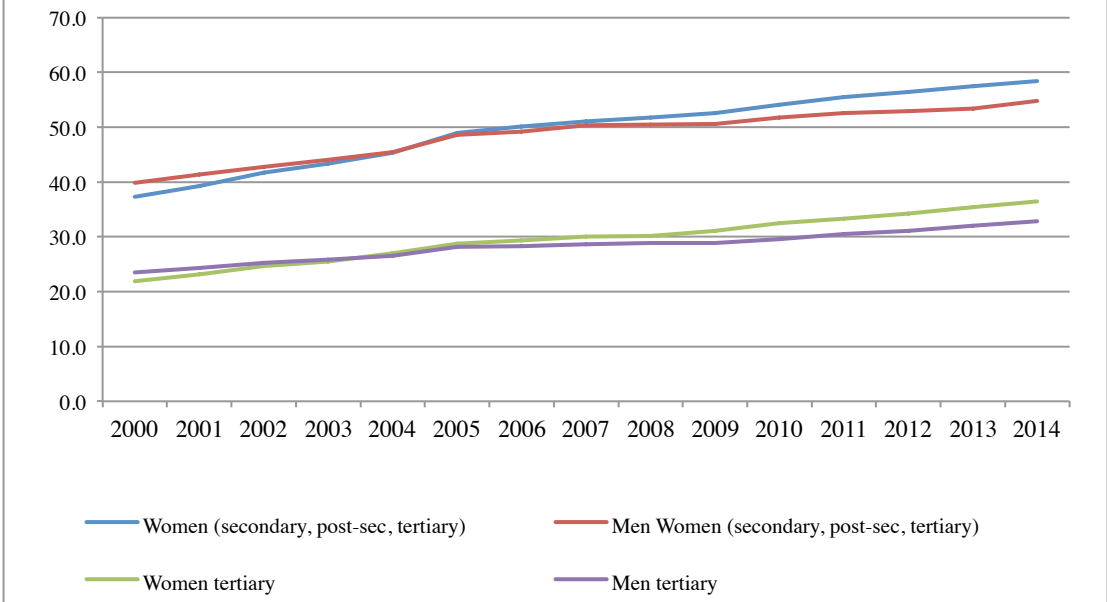
ANNEXES

Annex 1. Female labour force participation per age-group



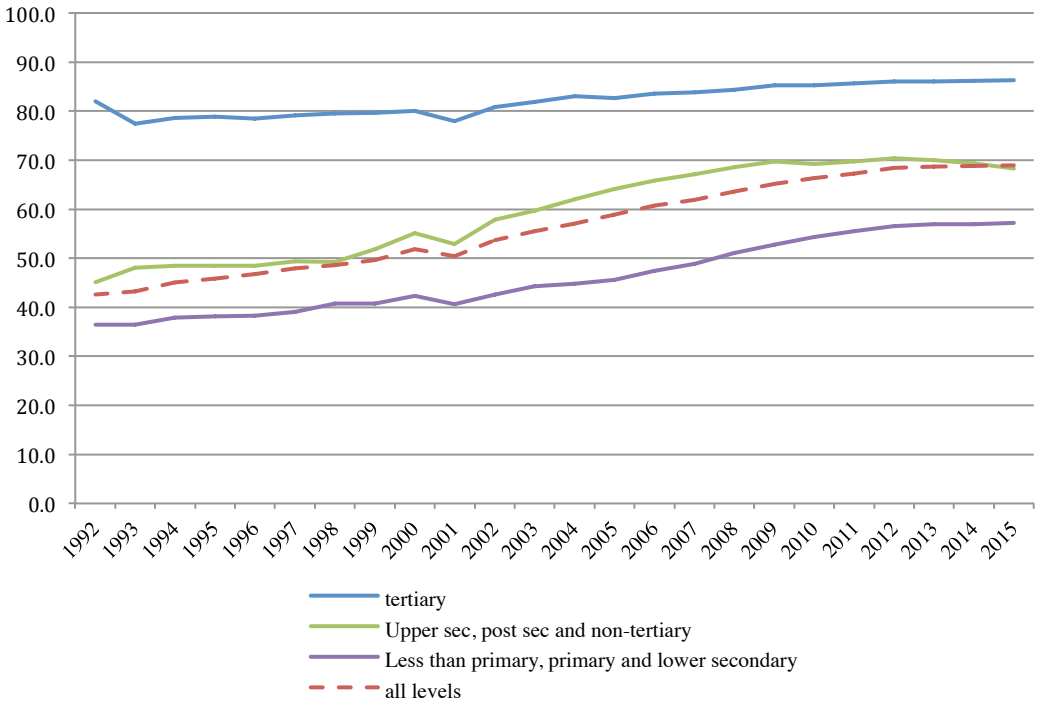
Own graphic using data from EUROSTAT LFS

Annex 2. Female and male population aged 25-64 by educational al attainment level



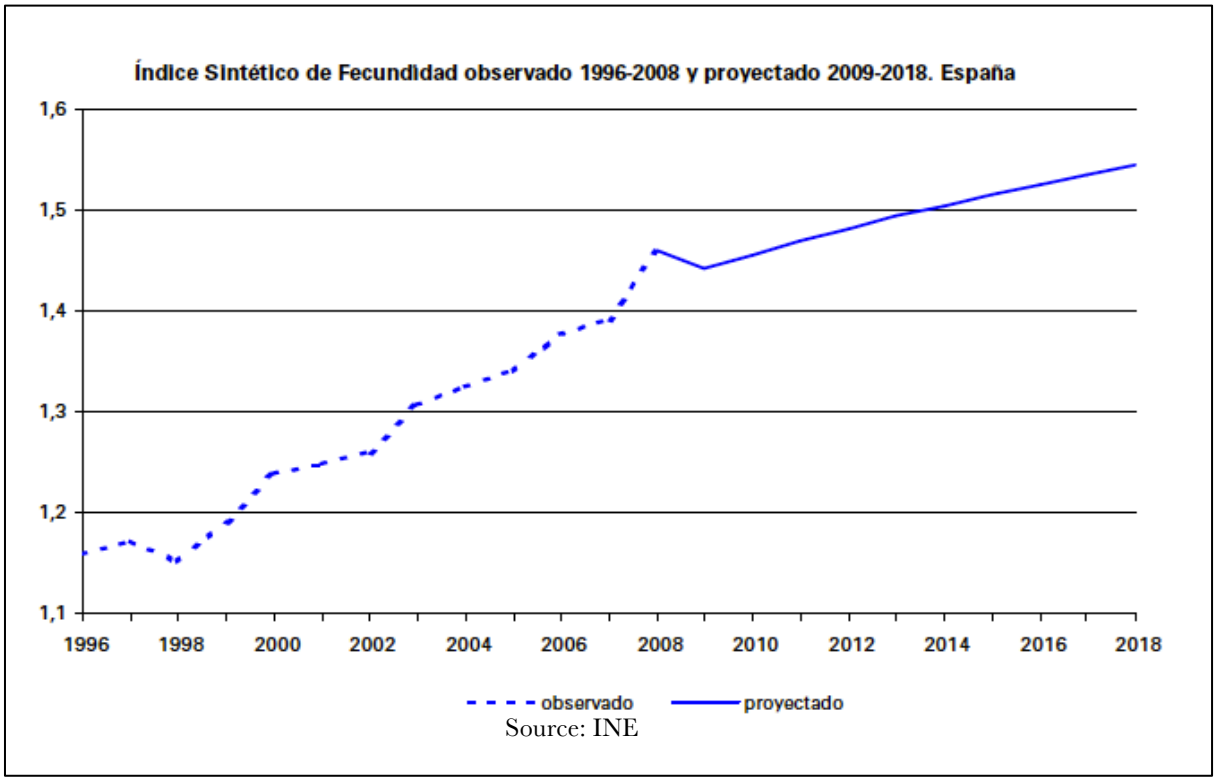
Own graphic using data from EUROSTAT LFS

Annex 3. FLFP by educational attainment level



Own graphic using data from EUROSTAT LFS

Annex 4. Synthetic Fertility Index (observed 1996-09 and projection 2009-18)



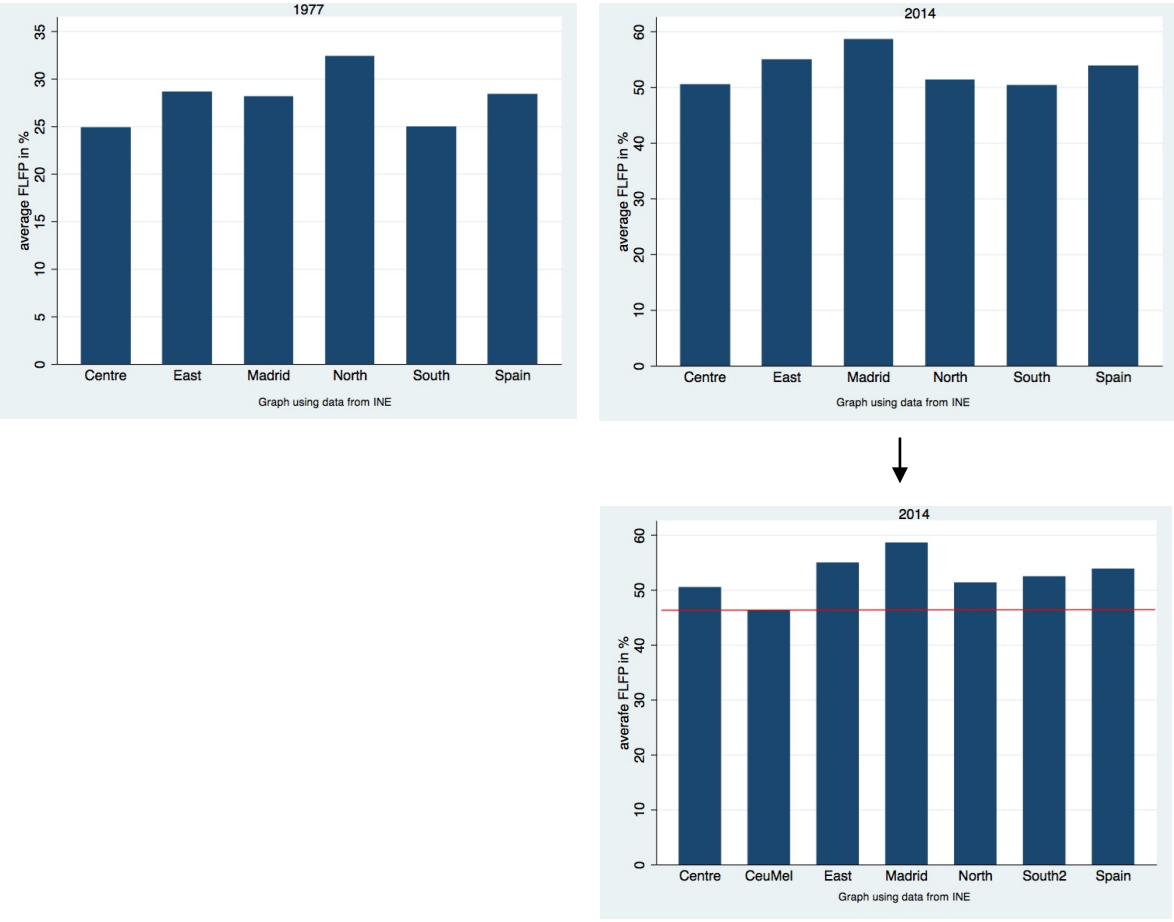
Source: INE

Annex 5. Macro-regions

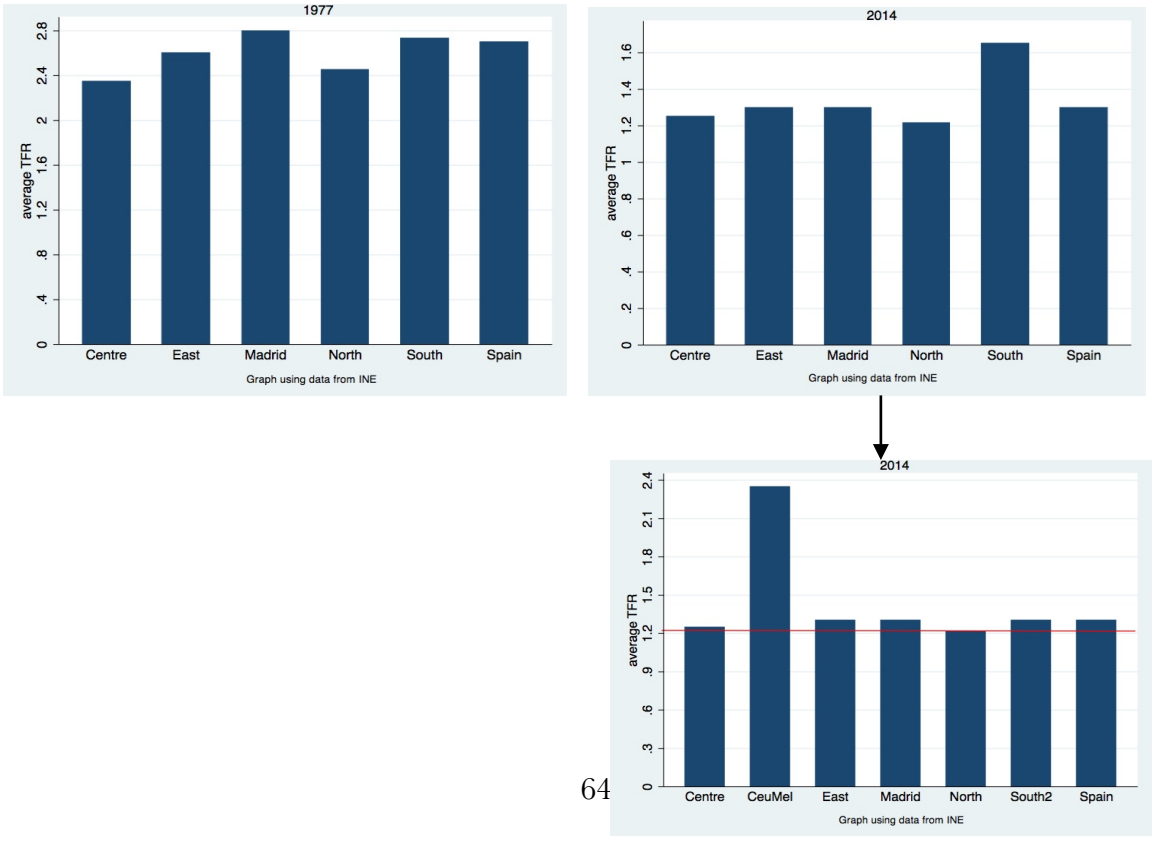
Constructed Macro-region	*Autonomous Region/City*
North	Galicia, Principado de Asturia, País Vasco, Comunidad Foral de Navarra, La Rioja
East	Aragón, Cataluña, Comunidad Valenciana, Illes Balears
Centre	Castilla y León, Castilla-La Mancha
South	Extremadura, Andalucía, Región de Murcia, Ciudad Autónoma de Ceuta, Ciudad Autónoma de Melilla, Canarias
South2	Canarias, Extremadura, Andalucía, Región de Murcia
CeuMel	Ciudad Autónoma de Ceuta, Ciudad Autónoma de Melilla
Madrid	Comunidad de Madrid (Reference category for Dummies)

Using information from NUTS1 (EUROSTAT)

Annex 6. Average FLFP per region in 1977 and 2014



Annex 7. Average TFR per region in 1977 and 2014



Annex 8. Regional development of TFR

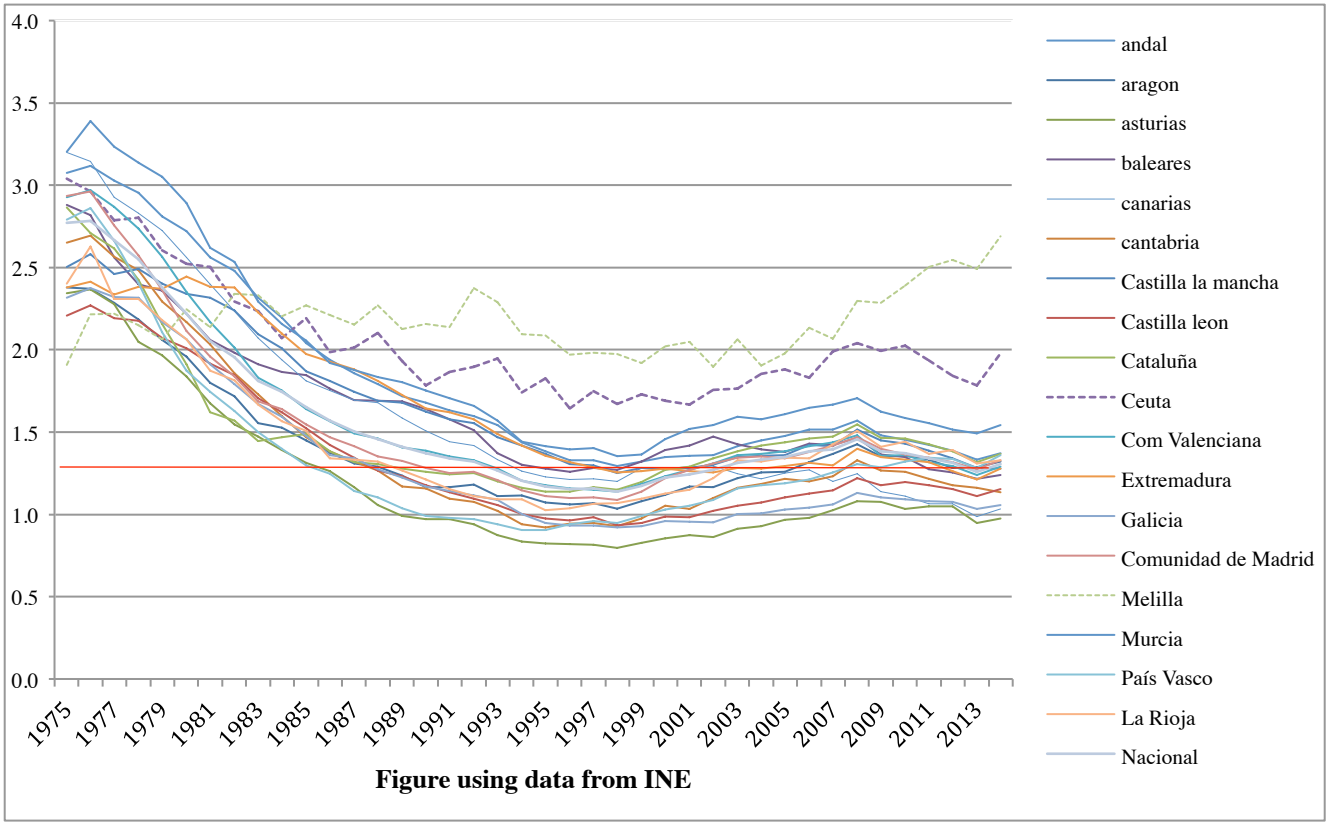


Figure using data from INE

Annex 9. Regional development mean age at first birth

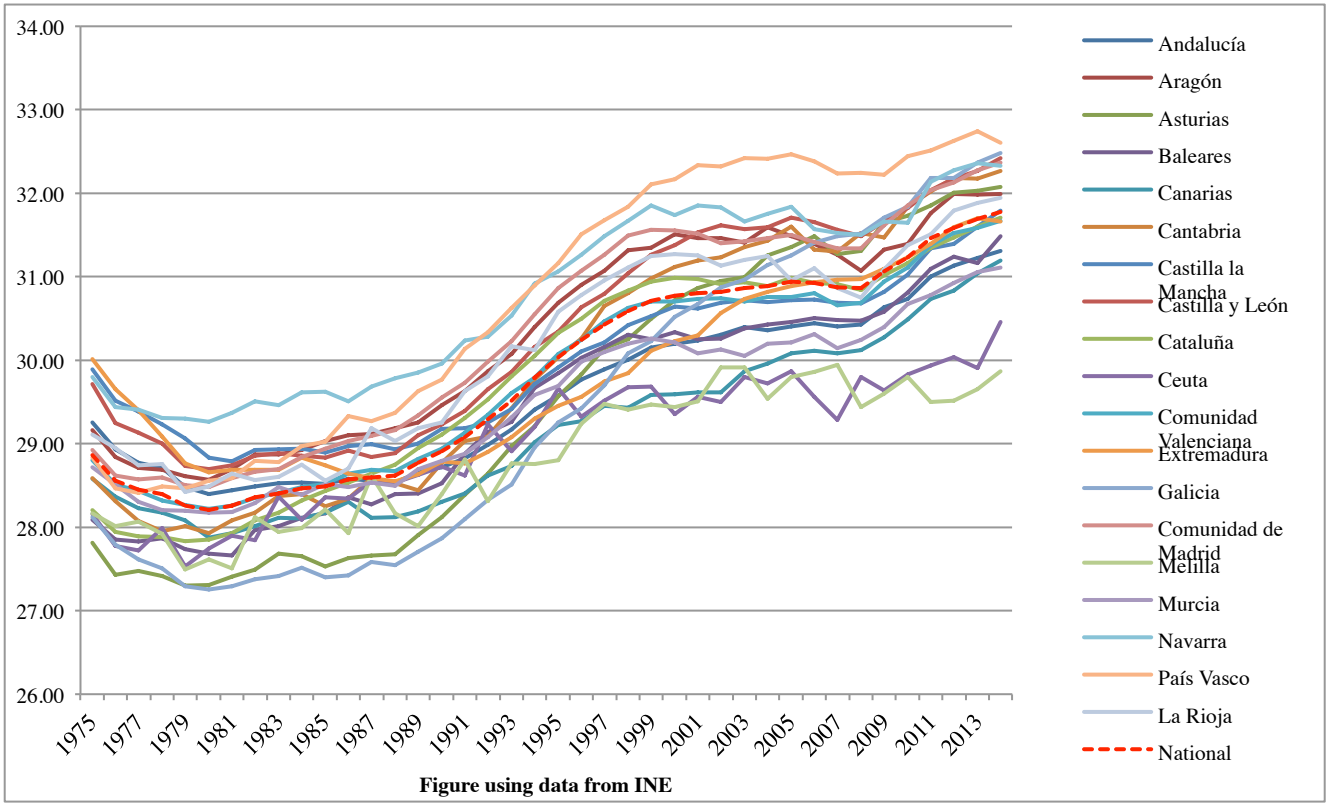


Figure using data from INE

Annex 10. Regional development of FLFP

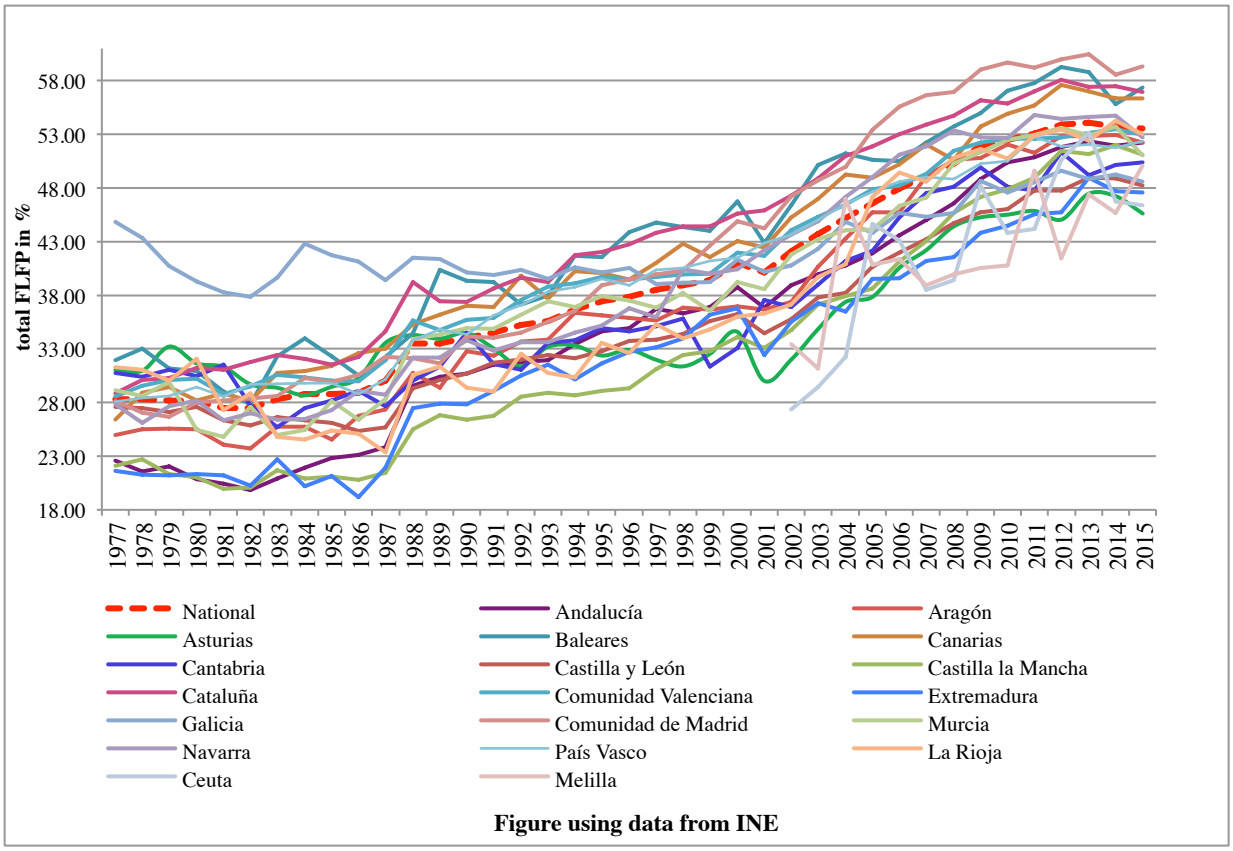
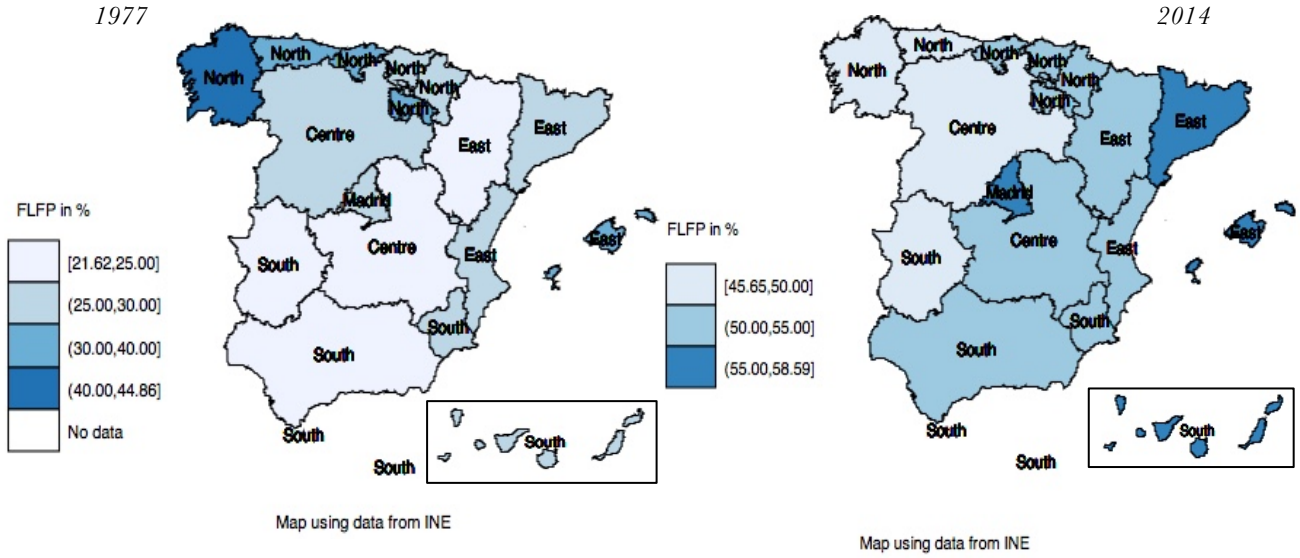


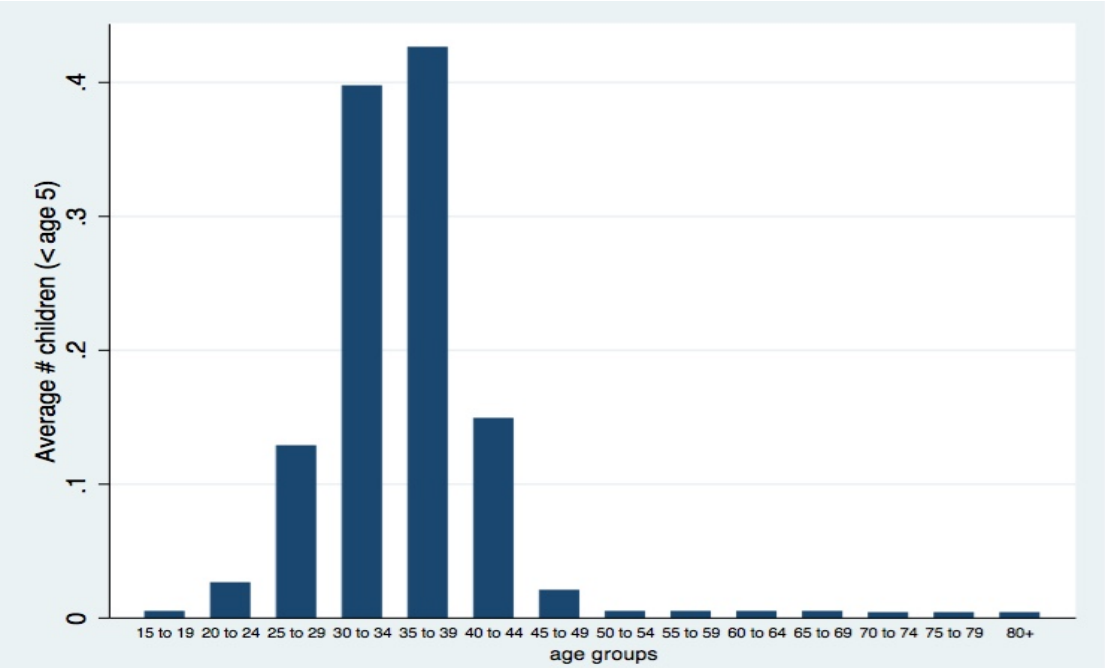
Figure using data from INE

Annex 11. Total FLFP in 1977 and 2014



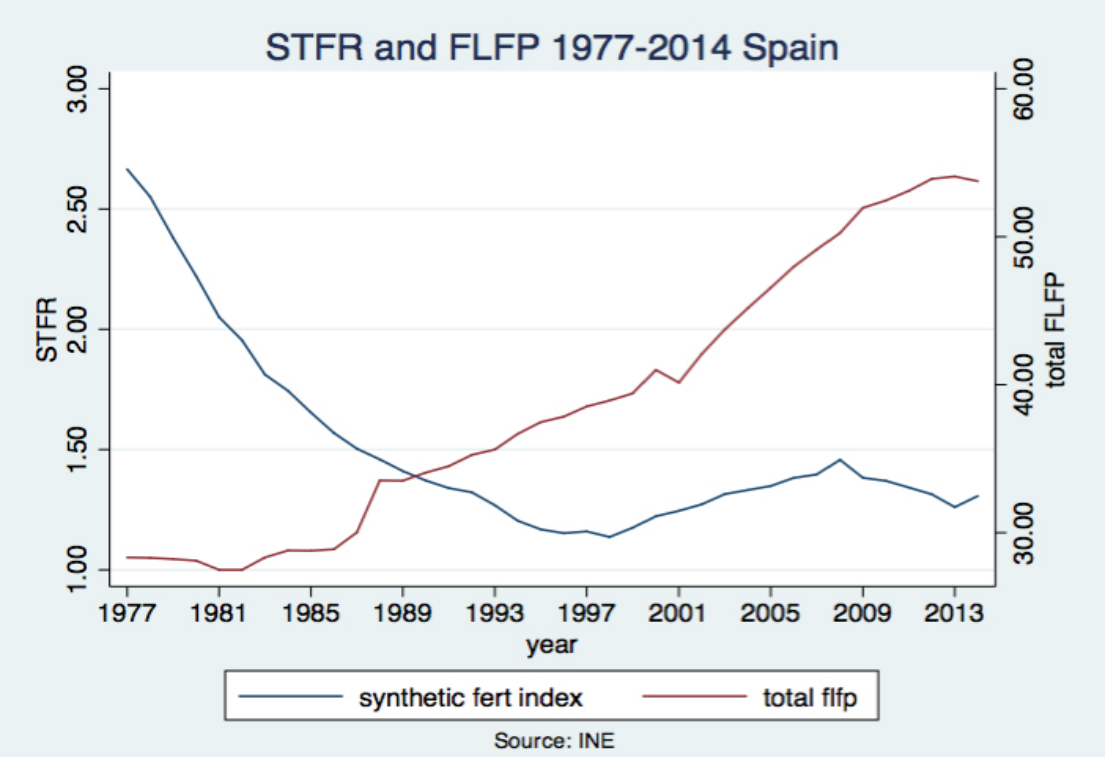
Note: The data for Ceuta and Melilla are not visible in the graph. Furthermore, for 1977 no data is available. The FLFP in Melilla in 2014 was 45.7. In Ceuta, it was 46.8 in 2014.

Annex 12. Average number of children under age 5

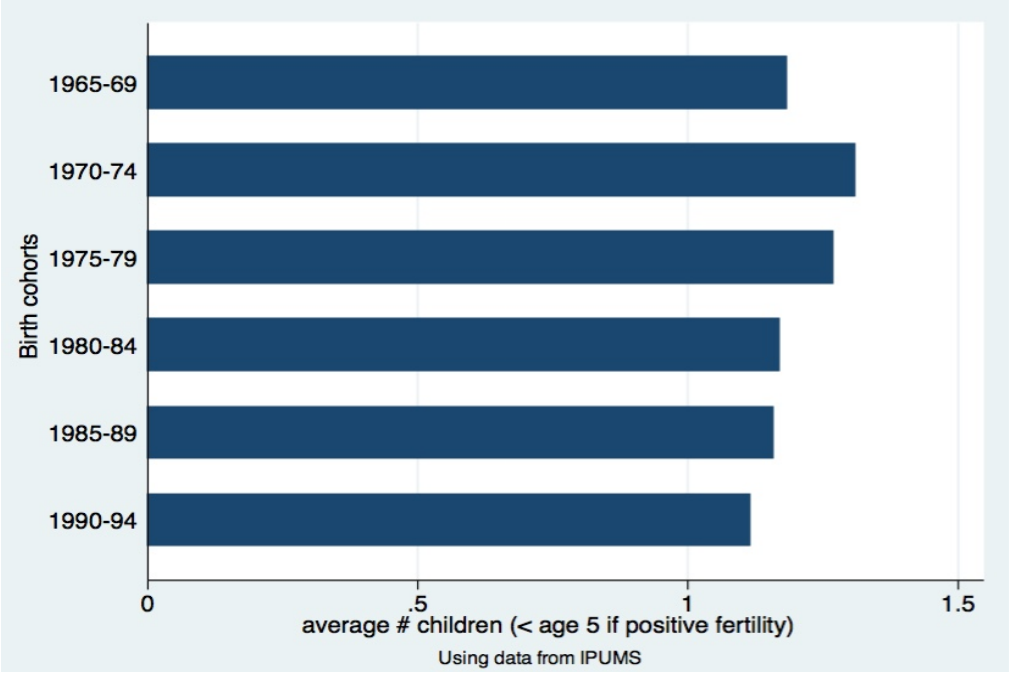


Note: Whole sample from the IPUMS International Census data (2001).

Annex 13. Fertility and FLFP transition 1977 to 2014



Annex 14. Average number of children (< age 5) if positive number of children



Annex 15. Descriptive statistics of variables

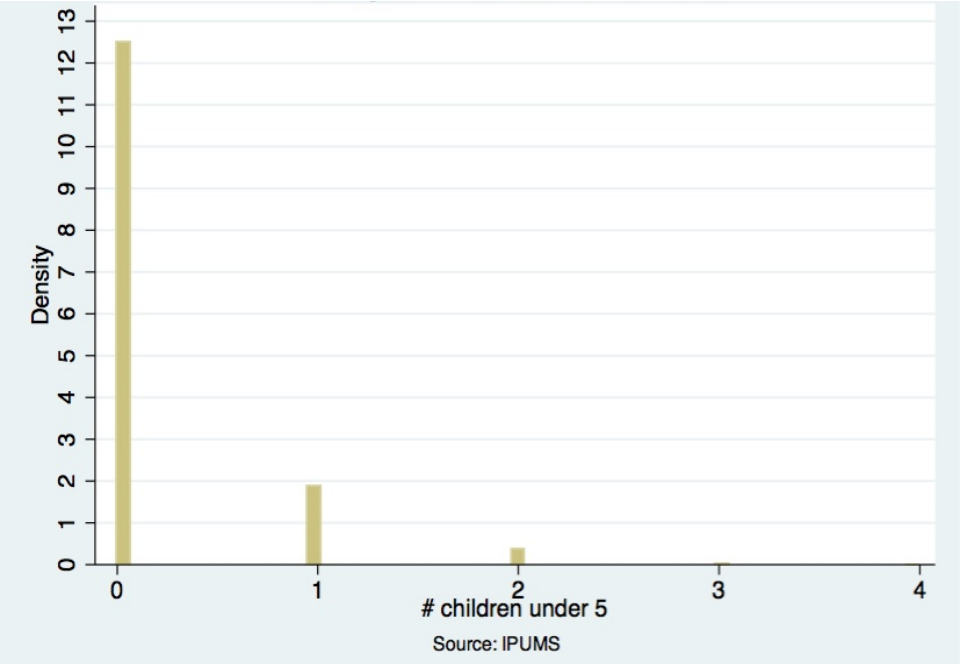
Variable	*Variable name in STATA*	*Mean*	*Standard Deviation*	*Min*	*Max*	*Number of observations*
Fertility (Dependent Var.)	*fert_3*	0.224	0.517	0	4	406,244
Personal traits:						
Economic Activity	*activity*	0.824	0.38	0	1	
Unemployment	*unemp*	0.311	0.463	0	1	334,760
Full-time	*fullt*	0.771	0.420	0	1	230,541
Marriage	*married*	0.297	0.457	0	1	392,566
Age	*age*	30.3	7.135	19	45	406,244
Age squared	*agesq*	970.2	445.818	361	2025	406,244
Birth cohorts (1960-65, …, 1990-94)	*cohort*	2.715	1.465	0	5	406,244
Education	*educ*	2.234	1.176	0	5	406,244
University degree	*uni*	0.217	0.413	0	1	406,244
One child	*onech*	0.0414	0.199	0	1	406,244
Two children	*twoch*	0.002	0.04	0	1	406,244
Multiple mother	*multmom*	0.002	0.04	0	1	406,244
Exogenous factors:						
Household size	*hhsize*	3.056	1.226	1	24	406,244
Education spouse	*edu_s*	1.917	1.156	0	5	146,511
Activity spouse	*activity_s*	0.932	0.252	0	1	146,511
Macro-regions:						
- South	*S*	0.264	0.441	0	1	406,244
- North	*N*	0.168	0.374	0	1	406,244
- East	*E*	0,291	0.454	0	1	406,244
- Centre	*C*	0.148	0.355	0	1	406,244
- South 2	*S2*	0.261	0.439	0	1	406,244
- Ceuta&Melilla	*CeuMel*	0.003	0.052	0	1	406,244
Interaction terms (i)						
- Married & South	*Sxm*	.078	.268	0	1	392,566
- Married & Centre	*Cxm*	.043	0.202	0	1	392,566
- Married & East	*Exm*	0.089	0.285	0	1	392,566
- Married & North	*Nxm*	0.05	0.219	0	1	392,566
- Married & South2	*S2xm*	0.078	0.268	0	1	392,566
- Married & Ceuta/Melilla	*CMxm*	0.001	0.028	0	1	392,566
- Activity & Centre	*actxC*	0.118	0.322	0	1	406,244
- Activity & Ceuta/Mellila	*actxCM*	0.002	0.045	0	1	406,244
- Activity & East	*actxE*	0.246	0.431	0	1	406,244
- Activity & North	*actxN*	0.139	0.346	0	1	406,244
- Activity & South	*actxS*	0.211	0.408	0	1	406,244
- Activity & South2	*actxS2*	0.209	0.407	0	1	406,244
- Education & Centre	*eduxC*	0.312	0.871	0	5	406,244
- Education & Ceuta/Melilla	*eduxCM*	0.005	0.121	0	5	406,244
- Education & East	*eduxE*	0.661	1.211	0	5	406,244
- Education & North	*eduxN*	0.391	0.994	0	5	406,244
- Education & South	*eduxS*	0.542	1.08	0	5	406,244

Variable	*Variable name in STATA*	*Mean*	*Standard Deviation*	*Min*	*Max*	*Number of observations*
Interaction terms (ii)						
- Education & South2	*eduxS2*	0.537	1.075	0	5	406,244
- Fulltime & Centre	*fulltxC*	0.107	0.309	0	1	406,244
- Fulltime & Ceuta/Melilla	*fulltxCM*	0.002	0.39	0	1	406,244
- Fulltime & East	*fulltxE*	0.232	0.429	0	1	406,244
- Fulltime & North	*fulltxN*	0.141	0.348	0	1	406,244
- Fulltime & South	*fulltxS*	0.157	0.365	0	1	406,244
- Fulltime & South2	*fulltxS2*	0.156	0.363	0	1	406,244

Sample statistics using data from IPUMS

Note: The Table Annex 15 is a summary of the variables used for the empirical analysis. If the number of observations is smaller than the total sample (406,224), it is either due to missing observations for specific variables (for instance if the definition of a variable only applies to certain individuals: the Dummy fulltime is only 1 or 0 if the women is active in the first place, and has a missing observation if she is not active), or that the original data delivered a non-classifiable information (such as for instance the label "Not in the universe" or "Other" that were classified as missing observations).

Annex 16. Histogram of the dependent variable (fert_3)



Appendices

Appendix 1. Correlation Matrix

	fert_3	age	agesq	cohort	married	activity	fullt	unemp	activity_s	edu	uni	edu_s	onech	twoch	multmom	S	N	E	C	CeuMel	S2
fert_3	1.0000																				
age	0.2285*	1.0000																			
agesq	0.2042*	0.9933*	1.0000																		
cohort	-0.2256*	-0.9806*	-0.9724*	1.0000																	
married	0.4677*	0.4188*	0.4016*	-0.4110*	1.0000																
activity	0.1418*	0.3539*	0.3165*	-0.3538*	0.1690*	1.0000															
fullt	-0.0520*	0.1591*	0.1458*	-0.1557*	0.0140*	.	1.0000														
unemp	-0.0246*	-0.1987*	-0.1804*	0.1965*	-0.0800*	.	.	1.0000													
activity_s	0.1245*	-0.1220*	-0.1357*	0.1172*	-0.0410*	0.1778*	0.0131*	-0.0501*	1.0000												
edu	0.0384*	0.1063*	0.0862*	-0.1079*	0.0010	0.1162*	0.1068*	-0.1852*	0.1069*	1.0000											
uni	0.0300*	0.1296*	0.1123*	-0.1306*	0.0120*	0.1186*	0.0898*	-0.1230*	0.0612*	0.8055*	1.0000										
edu_s	0.0730*	0.0852*	0.0691*	-0.0827*	0.0014	0.0701*	0.0812*	-0.1425*	0.1711*	0.4912*	0.4049*	1.0000									
onech	0.7141*	0.1254*	0.1137*	-0.1236*	0.2521*	0.0686*	-0.0279*	-0.0088*	0.0572*	0.0463*	0.0370*	0.0776*	1.0000								
twoch	0.2143*	0.0255*	0.0236*	-0.0250*	0.0519*	0.0097*	-0.0016	-0.0040*	0.0062*	0.0182*	0.0172*	0.0336*	-0.0083*	1.0000							
multmom	0.2188*	0.0256*	0.0237*	-0.0252*	0.0526*	0.0098*	-0.0017	-0.0040*	0.0066*	0.0186*	0.0178*	0.0343*	-0.0084*	0.9878*	1.0000						
S	-0.0009	-0.0683*	-0.0664*	0.0664*	-0.0013	-0.0374*	-0.0492*	0.1406*	-0.0252*	-0.0926*	-0.0710*	-0.0849*	-0.0023	-0.0020	-0.0017	1.0000					
N	-0.0063*	0.0512*	0.0498*	-0.0506*	0.0041*	0.0045*	0.0074*	-0.0515*	-0.0114*	0.0375*	0.0204*	0.0120*	-0.0062*	-0.0010	-0.0011	-0.2688*	1.0000				
E	0.0221*	0.0124*	0.0106*	-0.0121*	0.0158*	0.0334*	-0.0057*	-0.0560*	0.0127*	0.0188*	0.0141*	0.0134*	0.0086*	-0.0004	-0.0005	-0.3841*	-0.2879*	1.0000			
C	-0.0204*	-0.0152*	-0.0123*	0.0151*	-0.0090*	-0.0287*	0.0032	0.0180*	-0.0029	-0.0420*	-0.0437*	-0.0706*	-0.0085*	-0.0023	-0.0027	-0.2491*	-0.1867*	-0.2668*	1.0000		
CeuMel	-0.0050*	-0.0102*	-0.0096*	0.0097*	-0.0009	-0.0074*	-0.0026	0.0138*	-0.0088*	-0.0073*	-0.0091*	-0.0016	-0.0004	-0.0009	-0.0009	0.0862*	-0.0232*	-0.0331*	-0.0215*	1.0000	
S2	-0.0004	-0.0673*	-0.0655*	0.0655*	-0.0012	-0.0367*	-0.0491*	0.1394*	-0.0244*	-0.0920*	-0.0701*	-0.0850*	-0.0022	-0.0019	-0.0016	0.9932*	-0.2670*	-0.3815*	-0.2474*	-0.0307*	1.000

*Note: Pearson pairwise correlation coefficients displayed. * stands for p-value<0.05.*

Appendix 2. Probit versus Logit (Baseline model)

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
probit	392563	-2282895	-1760498	9	3521014	3521112
logit	392563	-2282895	-1766236	9	3532490	3532588

Note: N=Obs used in calculating BIC; see [R] BIC note

Appendix 3. Comparison Negative Binomial, Poisson and Probit (1st step)

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
Baseline	222119	-1450891	-1062515	10	2125050	2125153
nbreg	222119	-1803697	-1401314	11	2802651	2802764
poisson	222119	.	-1401359	10	2802737	2802840

Note: N=Obs used in calculating BIC; see [R] BIC note

Appendix 4. Comparison Probit models 1-6

i. Predictions

stats	p	p2	p3	p4	p5	p6	fert_3
N	222121	392566	143056	322958	392566	100240	406244
mean	.2200275	.1849092	.4222272	.2125765	.1834461	.4107477	.2244316
sd	.2157344	.2000719	.1341257	.2060166	.2111829	.1588793	.5165222
variance	.0465413	.0400288	.0179897	.0424428	.0445982	.0252426	.2667952
min	.0034198	.0015535	.0562927	.0066957	9.44e-06	.025798	0
max	.9999812	.9999996	.5970678	.9996915	1	.7469621	4

ii. AIC and BIC criteria

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
Baseline	222119	-1450891	-1062515	10	2125050	2125153
M2	392563	-2282895	-1682806	10	3365632	3365741
M3	143055	-1326727	-1247812	11	2495645	2495754
M3	143055	-1326727	-1247812	11	2495645	2495754
M5	392563	-2282895	-1576471	11	3152964	3153084
M6	100240	-931661.1	-852540.9	11	1705104	1705209

Note: N=Obs used in calculating BIC; see [R] BIC note

Appendix 5. Robust Probit Regressions (1-3)

VARIABLES	(1) *Robust 1*	(2) *Robust 2*	(3) *Robust 3*
age	0.551***	0.044***	0.041***
	(0.008)	(0.001)	(0.001)
hhsize		0.243***	0.352***
		(0.004)	(0.004)
activity	0.260***		
	(0.015)		
agesq	-0.008***		
	(0.000)		
married	1.170***		
	(0.007)		
edu	-0.064***		
	(0.003)		
S2	0.099***		
	(0.011)		
N	-0.020		
	(0.012)		
E	0.068***		
	(0.011)		
C	-0.037***		
	(0.013)		
CeuMel	0.013		
	(0.078)		
actxC		-0.058***	
		(0.020)	
actxE		0.092***	
		(0.014)	
actxN		-0.070***	
		(0.018)	
actxS		0.106***	
		(0.015)	
Sxm		1.167***	1.237***
		(0.014)	(0.018)
Cxm		1.309***	1.386***
		(0.019)	(0.023)
Exm		1.137***	1.209***
		(0.012)	(0.014)
Nxm		1.164***	1.208***
		(0.016)	(0.019)
eduxC		-0.108***	-0.064***
		(0.007)	(0.008)
eduxE		-0.070***	-0.020***
		(0.004)	(0.005)
eduxN		-0.062***	-0.016***
		(0.006)	(0.006)
eduxS		-0.102***	-0.062***
		(0.005)	(0.006)
fulltxC			-0.462***
			(0.023)
fulltxE			-0.302***

			(0.015)

Appendix 5. Continued

	(1)	(2)	(3)
VARIABLES	*Robust 1*	*Robust 2*	*Robust 3*
fulltxN			-0.465***
			(0.020)
fulltxS			-0.268***
			(0.019)
Constant	-10.644***	-3.338***	-3.358***
	(0.130)	(0.024)	(0.031)
Observations	392,563	392,563	222,119
Pseudo R2	0.264	0.222	0.235

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix 6. Comparison Zero-truncated Poisson and Poisson (2nd step)

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
poisson2	47978	.	-661099.6	12	1322223	1322328
ztp1	47978	.	-282267.2	12	564558.3	564663.7

Note: N=Obs used in calculating BIC; see [R] BIC note

Appendix 7. Pairwise correlation Matrix (Ztp 1-9 and fert_3)

Variable	ztp1	ztp2	ztp3	ztp4	ztp5	ztp6	ztp7	ztp8	ztp9	fert_3
ztp1	.	1.000	0.411	0.320	0.324	1.000	0.016	1.000	1.000	0.747
(p)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ztp2	1.000	1.000	0.379	0.320	0.313	1.000	0.016	1.000	1.000	0.748
(p)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ztp3	0.411	0.379	1.000	0.754	0.706	0.411	0.349	0.242	0.411	0.242
(p)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ztp4	0.320	0.320	0.754	.	0.989	0.320	0.073	0.320	0.320	0.242
(p)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ztp5	0.324	0.313	0.706	0.989	.	0.324	0.069	0.313	0.324	0.228
(p)	(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ztp6	1.000	1.000	0.411	0.320	0.324	.	0.016	1.000	1.000	0.747
(p)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)
ztp7	0.016	0.016	0.349	0.073	0.069	0.016	.	0.016	0.016	0.009
(p)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)
ztp8	1.000	1.000	0.242	0.320	0.313	1.000	0.016	.	1.000	0.749
(p)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)
ztp9	1.000	1.000	0.411	0.320	0.324	1.000	0.016	1.000	.	0.747
(p)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		(0.000)
fert_3	0.747	0.748	0.242	0.242	0.228	0.747	0.009	0.749	0.747	1.000
(p)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Note: Significance level in brackets (0.01)

Appendix 8. Estimate statistics for Ztp models 1-9

i. Predictions

stats	ztp1	ztp2	ztp3	ztp4	ztp5	ztp6	ztp7	ztp8	ztp9	fert_3
N	222121	322958	392566	102502	137617	222121	230541	392566	222121	406244
mean	.0863811	.0834624	.284759	.4059619	.4167785	.0863827	.3830983	.0726239	.0863952	.2244316
sd	.3671499	.3613702	.3407715	.1932833	.1877613	.3671575	5.246024	.3381448	.3672044	.5165222
variance	.134799	.1305884	.1161252	.0373584	.0352543	.1348046	27.52077	.1143419	.1348391	.2667952
min	.0012317	.0010502	.0420723	.1647421	.1800548	.0012251	.0392016	.0010871	.0012257	0
max	3.357831	3.215183	120.1438	4.632885	4.835293	3.358481	2143.643	3.087622	3.355772	4

ii. AIC and BIC information criteria

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
ztp1	47978	.	-153310.6	12	306645.2	306750.5
ztp2	67169	.	-215838	12	431699.9	431809.3
ztp3	70801	.	-457387	11	914795.9	914896.8
ztp4	44484	.	-313177	11	626375.9	626471.7
ztp5	62002	.	-441477.9	11	882977.8	883077.2
ztp6	47978	.	-153310.5	13	306647.1	306761.2
ztp7	49440	.	-291906.8	11	583835.6	583932.5
ztp8	70801	.	-227529.6	17	455093.3	455249.1
ztp9	47978	.	-153310.8	17	306655.5	306804.7

Note: N=Obs used in calculating BIC; see [R] BIC note

Appendix 9. Zero-truncated Poisson regressions (2nd step)

VARIABLE	Model (2) ztp2	Model (3) ztp3	Model (4) ztp4	Model (5) ztp5	(Model 7) ztp7	Model (8) ztp8
unemp	-0.00144*			0.155***		
	(0.000808)			(0.0176)		
activity		0.660***				
		(0.0829)				
fullt			-0.112***		-0.156***	
			(0.0202)		(0.0212)	
age	4.74e-05	0.0296***	0.0389***	0.0309***		5.69e-05
	(0.000143)	(0.00172)	(0.00224)	(0.00186)		(0.000143)
married	0.00185**	0.358***				
	(0.000822)	(0.0227)				
uni	0.00357**	0.230***	0.185***	0.152***	0.298***	
	(0.00157)	(0.0148)	(0.0206)	(0.0181)	(0.0410)	
onech	7.216***					7.224***
	(0.328)					(0.309)
twoch	7.749***					7.775***
	(0.339)					(0.315)
multmom		1.309***	1.785***	1.796***	0.697***	
		(0.0793)	(0.0175)	(0.0152)	(0.0858)	
hhsize	0.0344***	0.309***			0.672***	0.0184**
	(0.0111)	(0.0422)			(0.0438)	(0.00888)
S	-0.000918	-0.106***	-0.0132	-0.0230		
	(0.00217)	(0.0221)	(0.0275)	(0.0235)		
N	-0.00276	-0.160***	-0.149***	-0.166***	-0.133**	
	(0.00221)	(0.0238)	(0.0295)	(0.0263)	(0.0581)	
E	-0.00184	-0.0554***	-0.0900***	-0.0700***	-0.0231	
	(0.00180)	(0.0205)	(0.0254)	(0.0224)	(0.0565)	
C	-0.00249	-0.0250	-0.0212	-0.0145	0.0723	
	(0.00165)	(0.0255)	(0.0334)	(0.0284)	(0.0449)	
S2					-0.0418	
					(0.0571)	
CeuMel					0.0609	
					(0.238)	
activity_s			-0.0578	0.0252		
			(0.0597)	(0.0517)		
edu_s			0.0897***	0.0947***		
			(0.00802)	(0.00692)		
cohort					-0.134***	
					(0.0186)	
Sxm						-0.002
						(0.00150)
Cxm						0.0003
						(0.000697)
Exm						0.00142
						(0.000919)
Nxm						0.00274
						(0.00307)
actxC						-0.00289**
						(0.00127)
actxE						-0.00252
						(0.00169)
actxN						-0.0111
						(0.00685)
actxS						-0.000751
						(0.00159)
eduxC						0.000258
						(0.000192)
eduxE						9.22e-05
						(0.000503)
eduxN						0.00265
						(0.00164)
eduxS						0.000978
						(0.00103)
Constant	-6.890***	-3.880***	-2.314***	-2.197***	-2.953***	-6.833***
	(0.285)	(0.142)	(0.103)	(0.0868)	(0.168)	(0.290)
Observations	67,169	70,801	44,484	62,002	49,440	70,801

Dependent variable: positive values of fert_3. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1