A Mixed Approach to the Work-Motherhood Relation: An Application of fuzzy set Qualitative Comparative Analysis and Generalized Linear Models

Final version, forthcoming in *Comparative Sociology*

Giulia Maria Dotti Sani
University of Trento
giulia.dottisani@unitn.it

Mario Quaranta
Istituto Italiano di Scienze Umane
mario.quaranta@sumitalia.it

Abstract

Female labor force participation and total fertility rates have been negatively correlated until the early '80s. By the end of that decade, however, the relationship changed sign. Scholars have suggested a close link between individual reproductive behavior, labor market participation and institutional contexts, but we still lack clear evidence of the underlying micro-level mechanism. We propose the use of complementary techniques, fuzzy set Qualitative Comparative Analysis and Generalized Linear Models, as the different assumptions underlying the two, combinatory vs. additive, may lead to new insights on how the combination of institutional features can produce different outcomes in terms of the work-motherhood relationship.

Key words: work-motherhood, mixed methods, fuzzy set Qualitative Comparative Analysis, cross-country comparison, clustering.

Words count: 10.440.

1 Introduction

During the 20th century, two important albeit inverse trends have taken place in developed countries: while female labor market participation has steadily increased, there has been a substantial decline in fertility rates (Ahn and Mira 2002). Both trends have been all but uniform, as different countries have experienced different paces and varying levels of intensity in the change of the two phenomena.

Becker's Economic Theory of the Family (1991) has often been used to interpret the negative relationship between fertility and women's labor market participation. As women increasingly engage in higher education and paid employment, their earning potentials and consequently the opportunity-costs they face in childbearing increase greatly. This theoretically sound argument, however, fails to explain how the macro-level negative correlation has reversed since the last quarter of the 20th century. Currently, in fact, countries with higher levels of female employment also show higher fertility. Conversely, a more pronounced decline in total fertility rates is found in countries with lower female labor force participation rates (Engelhardt and Prskawetz 2004).

Researchers study the work-motherhood relation for several reasons. One is economic: working women buffer themselves and their families from poverty. Previous research has shown, in fact, that dual earner households face smaller poverty risks compared to single earner households (Barbieri, Cutuli and Tosi 2012). Another reason is social: couples in contemporary societies are having less children than they wish. A further reason is demographic: with fertility below the replacement rate[1], the working age population is becoming increasingly smaller and will have to sustain a growing number of older cohorts. It is therefore not surprising that the work-family conflict has received so much attention from the scientific community, and that the relationship between fertility rates, labor market institutions and welfare regimes has been vastly studied (Gauthier and Hatzius 1997; Castles 2003; D'Addio and D'Ercole 2005; Del Boca, Pasqua and Pronzato 2009). Despite so much research, the results have been far from enlightening. One of the reasons for this is, in our view, the difficulty in understanding the macro-level associations and the micro-level patterns of behavior in comparative perspective. Macro-level analysis, in fact, suffers greatly from the use of techniques that are sub-optimal when it comes to medium-N data and often fail to disentangle the contextual and the individual effects. On the other hand, contextual effects can be difficult to incorporate in the analysis and to understand when using cross-national micro-level data.

[1]The replacement rate, or substitution rate, of a population is around 2.1 children per woman (Castles 2003), and represents the number of children couples need to have to effectively "replace" themselves and guarantee stability in the population's size.

We attempt an innovative approach by using fuzzy-set Qualitative Comparative Analysis and Generalized Linear Models, in the framework of a mixed methodological strategy (Tarrow 2004). We develop a procedure to cluster a number of OECD countries according to relevant characteristics using fsQCA, and we then use the results to model micro-level data on fertility, introducing the country-clusters in probit regression models. The advantage of this strategy is that we are not just including country-dummies in the individual-level model: since the country-clusters are empirically built and tested using a method which forces the researcher to have a deep understanding of the cases (Ragin 2008), we are describing contexts with specific and known characteristics. Furthermore, rather than applying typologies that were developed in previous literature for other purposes – e.g. Esping-Andersen's (1990) welfare regime typology – we use fsQCA to test the presence of types and cluster our cases according to the empirical data (Kvist 1999, 2007).

We do not attempt a comparison of GLM and fsQCA – i.e. of two ontologies (Mahoney and Goertz 2006) – as this has been done in previous studies and has not provided enlightening results (Seawright 2005). Rather, we wish to build on the complementarity of the two techniques (Schneider and Wagemann 2010).

The paper is organized as follows. In the next section we review the main theories and empirical findings concerning the relationship between context, fertility and labor market participation, while in the third section we present our analytical strategy. In section four we introduce the data, the conditions and the calibration, while section five presents the results from the fuzzy set analysis. In the following sections we present the clusters of countries and the results from the micro-level analysis. The last section discusses and concludes.

2 Fertility rates and institutional characteristics

In the last quarter of the 20th century, in many OECD countries, women have massively entered paid employment and, at the same time, have greatly modified their reproductive behavior (Esping-Andersen 1999; Ahn and Mira 2002; Boongarts and Sobotka 2012). These trends have not been uniform across countries. While women's participation in the labor market was accompanied by a decrease and subsequent increase in fertility rates in Scandinavian countries, Southern European countries are currently featuring women's entrance in the labor market, while the fall in fertility rates has reversed only in very recent years (Boongarts and Sobotka 2012). Many other OECD countries are displaying in-between-patterns.

From a micro-level economic perspective, women's participation in the labor market is accompanied by declines in fertility rates because as

women gain more from paid labor they have greater opportunity costs in childbearing (Becker 1981). This micro-level theoretically sound argument is in contrast with the observed macro-level high fertility-high employment correlation that is today characterizing many OECD countries (Ahn and Mira 2002).

The childbearing-employment puzzle is important for a number of reasons. In industrial and even more so in post-industrial societies, women's employment is essential to guarantee their own independence, to protect their household from poverty, and to enhance the economic welfare of a country. Further, the fertility decline has been so steep that many countries are now well below the replacement rate. This means that in a near future a declining working age population will have to sustain an increasing number of elderly.

Given the large country-level variability, scholars have begun focusing on macro-level determinants that might be related to fertility and female labor market participation. Among these, the effect of family policies has been vastly investigated. The underlying idea is that family policies may reduce the opportunity costs of having a child for working women either by providing services – such as child-care – directly, or by financially sustaining the cost of private services. The effectiveness of such policies has been found to have varying degrees of success. Especially for what concerns fertility rates, Gauthier (2007) finds little evidence of a relation between policies and fertility. Nonetheless, the so-called social-democratic welfare regimes – known to provide universalistic support to households (Esping-Andersen 1990, 1999; Gauthier 1998) – are characterized by high levels of fertility.

In this context, we wish to draw the attention to two macro-level determinants that have received less attention in empirical research on this topic: the level of gender equality and of employment protection legislation. These two variables in combination might allow a better understanding of the female participation-fertility puzzle.

Gender equality is strictly linked to fertility rates because as gender equality increases, women gain control over their reproductive functions, and with contraceptive methods they avoid unwanted children. As equality increases, therefore, fertility rates are expected to decline. This idea was very clearly expressed by McDonald (2000:434-437): "[t]he transition from high fertility to fertility around replacement level is accompanied by an increase in gender equity within the institution of the family [...] When gender equity rises to high levels in individual-oriented institutions while remaining low in family-oriented institutions, fertility will fall to very low levels". However, the relation is non linear. When gender equality reaches a certain threshold in both individual-oriented institutions and family-oriented institutions, i.e. the work place and the family, then higher fertility can be reached. This is because, for instance, the opportunity costs

of having children are no longer on the shoulders of the mother but spread on the household, or because emancipated women have greater bargaining power among the couple.

Employment protection legislation is another macro-level variable that has changed over the past quarter of the century and might allow a better understanding of the relationship under study. High levels of employment protection, in fact, might create a dual labor market that excludes women from actively taking part of it. On the one hand, mothers may not wish to engage in long term, long hours contracts that might interfere with their family duties (Hakim 2000), and on the other hand employers might be discouraged from hiring individuals that are more likely to be on paid leave for maternity or sickness, such as women (Estévez-Abe 2005). Low levels of protection, on the contrary, can ease entry and exit from paid employment according to the needs of working mothers and are generally associated to more flexible working hours and part-time opportunities. They could, therefore, enhance childbearing (Adserà 2005). The backfiring argument to this is that low employment protection might impede access to a stable form of employment and income, thus hindering family formation first and childbearing afterwards (Del Boca 2003).

At the light of these considerations, we examine the fertility-employment relation by considering not only the association with family polices but also the implications tied to gender equality and labor market flexibility.

3 Analytical strategy

In the previous section we have introduced what are the potentially problematic issues tied to an unbalance between female labor force participation and childbearing. With such a picture, researchers have greatly stressed the importance of institutional factors in shaping women's participation in the labor market and their levels of fertility. In this section we present a combined methodological strategy which aims at reducing the weaknesses of quantitative or qualitative approaches, while improving the respective strengths (Tarrow 2004).

All the macro-level studies which have been interested in the work-motherhood relation encountered some difficulties due to the lack of data for an acceptable number of countries. The techniques which have been adopted vary greatly from one study to another, from pooled ordinary least squares regressions to time series analysis, and have been found to have varying degrees of success[2]. The question that arises, however, is how suitable regression analysis can be on medium-N data since the technique

[2]For a review see Gauthier (2007).

presents evident weaknesses. The problems with aggregate statistical analysis with a small number of cases are known (Berk 2005). Among others, the influential role that outliers assume on small samples and the problems that arise when analyzing few cases and many variables (Lijphart 1971).

To analyze the childbearing-working dilemma the most suitable techniques are, undoubtedly, those that use longitudinal, or to the least individual-level data. The research that has gone in this direction has yielded quite different results. As far as family policies and fertility are concerned: "disentangling the impact of social and welfare benefits on demographic behavior [...] is therefore a difficult exercise, and not surprisingly, one that has led to contradictory findings" (Gauthier 2007:339). Family policies can also have varying effects on women's labor market participation: in an influential study, Del Boca, Pasqua and Pronzato (2009) find that the availability of child-care significantly increases the probability of working, whereas long maternal leave and family allowances have the opposite effect.

Furthermore, the collection of longitudinal individual data is highly costly in terms of both time and financial resources and very often such data is not available for enough countries to allow satisfying cross-national comparisons. It is therefore not surprising that many studies have attempted to find macro-level relations between fertility, family policies and female labor market participation, even with the aforementioned limitations.

Macro-level data can prove very informative if analyzed in the correct way. In comparative micro-level analysis, in fact, the question that often arises is how to take into account clustering within the data. Individuals are nested, for example, in countries, and are influenced by the characteristics of the surrounding context, other than their own. How to use the information yielded by the context is an issue that social scientists have to face when attempting any study that involves different levels of analysis. It is not our aim here to discuss the advantages and drawback of several types of hierarchical models. It should be sufficient to point out that classic multi-level models require a relatively large number of level-two units, such as houses in counties, pupils in schools, etc. (Snijders and Bosker 1999, Maas and Hox 2005). Given that often this is not the case – and especially is not the case when comparing a handful of developed countries – the most commonly adopted solution is to introduce the level two units as dummy variables. The "country" effect, however, is not always easily interpretable. Moreover, this approach impedes introducing other macro-level information because of collinearity issues. Further, in our case this solution would confound the analysis, given the high number of dummies and relative – eventual – interaction terms. A more informative way of using country dummies is to group them according to a theoretically

and empirically driven clustering. The clustering of countries reduces the number of dummy variables and interaction terms without losing important information about the association between context and individual.

The question then is how to build an informative clustering of countries. Important scholars over the past decades have introduced several possibilities to cluster countries. Among these, we have been guided in our work by Esping-Andersen's (1990) welfare regime typology, including Ferrera's (1996) extension of the southern European cluster, and by Hall and Soskice's varieties of capitalism (2001). Public policies and family policies have been also used to form groups of countries, as in the work of Gauthier's (1998) and, more recently, of Thévenon (2011). Korpi (2000) by introducing gender as a source of inequality in the welfare state, produced a typology which included policies supporting not only families but also policies more specifically aimed at supporting working women and mothers. All the aforementioned typologies convey an important amount of information but do not explicitly address the work-motherhood relationship. As argued by Estévez-Abe (2005), we believe that for the study of reproductive behavior it is necessary to include additional information.

To achieve this, we use fsQCA to define groups of countries (Kvist 1999, 2007; Schneider and Wagemann 2010) that will be applied to the analysis of micro-level data. fsQCA is generally used to test theories and hypothesis (Schneider and Wagemann 2010), mainly to deal with "multi-conjunctural causation" (Ragin 2000; Braumoeller 2003; Ragin 2008; Berg-Schlosser et al. 2008; Aus 2009). But this is not the only possible application. In fact, fsQCA has been used to assess the conformity of cases to ideal-types (Kvist 1999, 2007), while we use it as a data reduction technique since it is a very strong tool to identify cross-case patterns (Ragin 2005, 2008). Variable-oriented data reduction techniques, such as cluster analysis or factor analysis, are not reliable when dealing with a low or medium number of observations, despite their extensive use (MacCallum et al. 1999). We propose to use fsQCA since it is an appropriate technique for intermediate-N macro-comparative analysis (Ragin 2000; Wagemann 2007; Ragin 2008; Schneider and Wagemann 2010). The application of fsQCA as a data reduction technique to define clusters of countries to be used as contextual information in a micro-level analysis could help deepen the understanding of the cases since "[f]amiliarity with the cases is a requirement before, during, and after the analytical moment of a QCA" (Schneider and Wagemann 2010:400). To build the clusters mentioned above, we select four conditions, introduced in section 2, which have been used less frequently in the study of the relation between employment and childbearing. Specifically, we include measures of female labor force participation, gender empowerment, generosity of parental leave schemes – as proxy of family policies – and employment protection legislation in the construction of the clusters.

After selecting and calibrating the aforementioned country-level characteristics – i.e. the conditions – we use fsQCA to derive the paths leading to high and not-high fertility rates. We then build clusters of countries based on the conditions we selected for the analysis. The clusters, which will represent a more complete scenario compared, for instance, to Esping-Andersen's and Gauthier's typologies, are then used in probit regression models as dummy variables.

The models will attempt to describe micro-level behavior and especially the relation between being in paid employment and having a high number of children. After estimating a pooled model including all countries, we introduce the country groups to control for contextual effects and then add interaction variables between the country groups and an individual-level variable, being employed or not. The use of interaction terms shall shed light on whether and how an individual-level variable varies in different groups of countries.

4 Data

4.1 Outcome and conditions

The aim of the following analysis is to examine country-specific fertility rates and their relation with gender equality, female labor market participation, labor market flexibility, and public policies in 20 OECD countries[3].

The selected countries are all highly developed, but provide a variegate scenario, as they represent different ways of organizing a welfare state (Esping-Andersen 1990; Gauthier 1998). Further, for this set of countries we have complete data for both the macro and the micro-level.

As far as the outcome is concerned, we follow a common procedure in demography to study the fertility of a country and use a period measure of fertility, the total fertility rate (TFR). The TFR is the average number of children per women in a given year. There has been much debate in the literature on the differences between period and cohort measures of fertility (Ni Bhrolchain 1992; Goldstein, Sobotka, and Jasilioniene 2009; Ni Bhrolchain 2011; Boongarts and Sobotka 2012). Cohort measures are in some cases preferred because they are not affected by the so-called tempo distortions. Tempo distortions, or tempo effects, occur when the age at childbearing of cohort(s) changes rapidly. In case of such changes, period measures such as the TFR might be distorted, and not measure the "true" fertility rate. We chose a measure of period fertility rather than cohort fertility, or tempo-adjusted fertility rates, because we are using the

[3]Countries are: Austria (AT), Australia (AUS), Belgium (BE), Canada (CA), Denmark (DK), Finland (FI), France (FR), Germany (DE), Greece (GR), Ireland (IE), Italy (IT), Japan (JP), Netherlands (NL), Norway (NO), New Zealand (NZ), Portugal (PT), Spain (ES), Sweden (SE), United Kingdom (UK), United States (US).

measure to establish "a population's current reproductive performance" (Ni Bhrolchain 1992:600) and not to predict long term population changes. The TFR is therefore appropriate for our analysis because we are interested in taking a "snapshot" of contemporary societies.

The first condition is female labor force participation (FLFP), which is the percentage of women between 15 and 64 years old in paid work (OECD 2009, United Nations 2009). We prefer a measure of female rather than maternal labor force participation because high female labor force participation should translate into more opportunities of outsourcing domestic work and especially childcare for working mothers. In other words, mothers may enter/remain in the labor market because other working women, not necessarily mothers, are working directly for them or indirectly providing services for them. Furthermore, we use a general measure of employment rather than distinguish between part-time and full-time because part-time regulations, in terms of both hours of work and working conditions, vary greatly between countries, making comparisons little informative.

The Gender Empowerment Measure (GEM), our second condition, is an indicator of gender equality and provides information on how much women count in important spheres of society. The GEM is built on three dimensions: 1) political participation and decision-making; 2) economic participation and decision-making and 3) power over economic resources. The first dimension is measured using the female and male shares of parliamentary seats, while the second derives from a combination of female and male shares of positions as legislators, senior officials and managers and female and male shares of professional and technical positions. Finally, power over economic resources considers female and male estimated earned income. Compared to other gender equality measures like the Gender Development Index, it is used to capture women's economic independence rather than the disparity between men and women[4] (UNDP 2007).

The third condition is a measure of parental leave generosity (PLG), which we use as proxy for the extent of family oriented public policies. The effect of parental leaves depends not only on their duration but also on the extent to which they are paid (Gornick et al. 1997). For this reason we build a composite index that accounts for duration and compensation. This is achieved by considering a) the total number of weeks of maternity leave times the cash benefits paid during maternity leave, plus b) the total number of weeks of parental leave times the cash benefits paid during parental leave, plus c) the total number of weeks of child-care leave times the cash benefits paid during child-care leave. The cash benefits are mea-

[4]Data are as of 31 May 2007 for most countries. Exceptions are detailed in the cited report.

sured as a percentage of female wages in manufacturing earnings received during the leave. The final value is then rescaled to range from 0 to 1. This measure, as well as TFR and FLFP, is derived from the Comparative Family Benefits Database 1960-2008 (Gauthier 2010). For sake of comparability, we select data from 2005, that is the most recent year for which values are available for all countries on all indicators.

The last condition is employment protection legislation (EPL), which we account for by using the data provided by OECD's Labor Force Series, Overall Strictness of EPL, referring to 2004. EPL is a three-dimensional index tapping three main areas of employment regulation: protection against individual dismissal, regulation of fixed term forms of employment, and regulation of collective dismissals (OECD 2004). Thus, the overall EPL provides a measure of average fluidity-rigidity of the labor market.

4.2 Calibration

In this section we discuss the criteria applied to set the cross-over points and the upper and lower thresholds used to calibrate the outcome and the conditions applying the direct method (Ragin 2008).

Our outcome of interest is the Total Fertility Rate of a country. In order to calibrate TFR, we compare its scores across a selection of developed countries. Considering that the population replacement rate is about 2.1 children per woman (Castles 2003) we set the upper threshold to 1.8 because only a few developed countries score higher. The lower bound is set to 1.4, as countries with fertility rates below this value belong to the so-called lowest-low fertility group (Kohler, Billari and Ortega 2002). The cross-over point is set to 1.6. This value was chosen by a direct observation of the selected sample of countries, in that this score creates two distinct groups.

The first condition is Female Labor Force Participation (OECD 2009; United Nations 2009). We set the thresholds by considering not only the current levels of female labor force participation but also the growth of the phenomenon from 1970 to 2005. Ideally, the high Female Labor Force Participation set should include those countries where women were actively part of the workforce already in the '70s. This was achieved by fixing the upper threshold at 73%. The lower threshold was instead set at 55%. Countries below this bound are those displaying low levels of FLFP nowadays as well as forty years ago. We decided to set the cross-over at 63%. The calibration of this condition also roughly reflects a classification made by the OECD (2009), which separates countries with high, medium and low Female Labor Force Participation. The thresholds also resemble those summarized by Aaberge et al. (2005) based on the work by Ahn and Mira (2002) and Engelhardt, Kogel and Prskawetz (2001) that analyze FLFP in

21 OECD countries and divide them into three groups: "The high participation group, in which the participation rate (FLP) is higher than 60 per cent, includes the United States, Canada, the United Kingdom, Sweden, Norway, Denmark, Finland and Switzerland. The medium participation group includes countries where the participation rate is in the 50-60 per cent range. The low participation countries are where the female participation rate is less than 50 per cent (Italy, Spain and Greece)" (Aaberge et al. 2005:128).

The second condition included in the analysis is the Gender Empowerment Measure (UNDP 2007, 2011). We are interested in defining a set which includes countries with high levels of gender equality, given that high human development does not automatically translate into gender equality. The GEM ranges from 0 to 1. A score higher than 0.9 means that women have important roles in social, political and economic spheres of society. The only countries overcoming this threshold – and widely recognized as those with the highest gender equality (Estévez-Abe 2005) – are Sweden and Norway. Hence, we set the upper threshold to 0.9. To gain a better understanding of the lower part of the GEM distribution, we observed the decomposed measure. Given that the only country which has low scores on each component is Japan, we set the lower threshold to 0.6. This allows separating Japan, which is completely out of the set, from the other countries which instead have higher scores on at least one component of the index. The cross-over point is set at 0.84 and distinguishes the countries with a high level of GEM from those which are still striving to achieve gender equality.

The third condition is an overall measure of the parental leave generosity. The countries scoring highest on the index were France and the Nordic countries. Given that these welfare states are renewed for their generosity in terms of family friendly measures, we set the higher threshold to 0.63, a value that neatly separates the mentioned countries from the others. We set the lower bound to 0.15, assuming that any measure under this level would not significantly affect the outcome. Finally, the cross-over is set at 0.31. The last condition is Employment Protection Legislation (OECD 2004). The EPL index is composed by three underlying dimensions, as shown in the previous paragraph. To set the upper threshold, we identified the countries which scored high on all the three dimensions and then set the upper threshold to 2. We applied the same procedure for the lower threshold. The score used to define the countries completely out of the set is 1, since all the countries scoring less than this value have low scores on all three dimensions. The cross-over point is set to 1.45. The choice of this value is guided by the weighting adopted by the ELP index.

Table 1 summarizes the thresholds used for calibrating the conditions and the outcome.

TABLE 1 ABOUT HERE

5 Fuzzy set analysis

The aim of the analysis is to provide an explanation for both high and low levels of fertility. To achieve this we repeat each step of the analysis twice: once for high TFR and once for not-high TFR (∼TFR). For both outcomes we present only the conservative and the most parsimonious solution because "the researcher is free to choose which formula(s) to put into the center of attention. Most likely, not all three solution formulas are extensively used for substantive interpretation" (Schneider and Wagemann 2010:408). In fact, the intermediate solution is not substantively useful for our purpose since both the conservative and the most parsimonious solutions will provide us with enough information to cluster the cases[5].

5.1 TFR

Although the "standards of good practice" suggest to begin the fsQCA with the analysis of necessary conditions (Schneider and Wagemann 2010), we analyze only sufficient conditions since we use fsQCA instrumentally to identify cross-case patterns which are helpful to produce clusters of countries[6].

The truth table, reported in table 2, conveys important information for the analysis of sufficient conditions leading to high TFR. The first is that 11 of the possible 16 combinations describe the cases included in the analysis. The 5 configurations of conditions for which we have no cases are removed. The second information is that 5 combinations of cases have a positive outcome. Third, the table indicates the consistency thresholds which define the presence of the outcome. We set this threshold to 0.8, so all the combinations which are greater to or equal than 0.8 get a 1 on the column of TFR, indicating the presence of a combination of conditions sufficient for the outcome.

TABLE 2 ABOUT HERE

Moving on to the analysis of sufficient conditions for high TFR, the conservative solution showed that there are three paths to high TFR:

[5]The intermediate solutions are provided in the Appendix.

[6]Tests for necessary conditions were not among the aims of this article; nonetheless, we performed them and we report the results in the Appendix.

$$\overbrace{\sim GEM * \sim EPL * \sim PLG}^{(a)} + \overbrace{GEM * FLFP * EPL}^{(b)} + \overbrace{GEM * FLFP * \sim PLG}^{(c)} \rightarrow TFR \quad (1)$$

The first path (1a) means that where the levels of gender equality, labor protection and parental leave are low, TFR is high. The second path (1b), instead, tells us a very different story: high gender equality, a high number of women who are active in the labor market and high levels of employment protection lead to high TFR. The third, (1c), says that high levels of gender equality, women participating in the labor market and low degree of parental leave benefits lead to high TFR. The solution has a good level of consistency (0.87) and has a coverage of 0.70. Practically, there are three equifinal paths to TFR. However, they should be distinguished, as they do not have the same importance. Their coverage measures allow us to differentiate them, as (1a) has a raw coverage of 0.32, while (1b) of 0.44 and (1c) of 0.31. It can therefore be argued that the second path (1b) is more relevant since it covers more cases[7].

Looking at the results more in detail, Ireland, New Zealand, the United Kingdom and the United States are members of path (1a). Denmark, Finland, the Netherlands, Norway and Sweden are instead members of path (1b), while Australia and the Netherlands belong to (1c). The interesting finding is that the outcome is the product of three completely different scenarios. In the first, parental leave schemes are not generous, the labor market is highly fluid and gender equality is still far from being reached. In the second scenario, instead, women are highly empowered and are a stable part of a well-regulated work force. The third scenario describes a situation which is similar to the previous with the exception of parental leave. We also see that the Netherlands is member of two paths which means that the route for this case to high TFR is twofold.

The second solution, the most parsimonious one, presents two paths to high TFR:

$$\overbrace{GEM * FLFP}^{(a)} + \overbrace{\sim EPL * \sim PLG}^{(b)} \rightarrow TFR \quad (2)$$

Gender equality and female labor force participation or low labor protection and low parental leave are sufficient for TFR. The solution consistency is 0.86, while the coverage is 0.80. The two paths show different levels of consistency and coverage. Path (2a) has a good consistency (0.85)

[7]Unique coverage: (1a) 0.14, (1b) 0.25, (1c) 0.02.

and a coverage of 0.59. Conversely, ∼EPL ∗ ∼PLG seems to be a better path to TFR. In fact, it has a high level of consistency (0.94) despite covering fewer cases (0.44). Cases with membership in (2a) are Denmark, Finland, the Netherlands, Norway and Sweden, while cases with membership in (2b) are Australia, Ireland, New Zealand, the United Kingdom and the United States.

5.2 ∼TFR

We now focus our attention on not-high TFR. Since set-theoretic methods are non-symmetrical, unlike correlational methods, the results will not be opposite to the ones from the analysis of sufficient conditions leading to high TFR (Schneider and Wagemann 2010).

Table 3 illustrates the truth table for not-high TFR. There are 11 configurations of conditions with empirical cases. As before, the configurations which do not have empirical instances have been removed. There are 3 configurations of conditions which can be sufficient for the outcome of interest and five cases show these patterns: Canada, Japan, Greece, Spain and Italy. We decided to relax the cut-off threshold for consistency in order to have Italy as an instance of ∼TFR. We know, in fact, that in the considered year this country had a very low fertility rate and it had also experienced a long period of fertility decline (Kohler, Billari and Ortega 2002). Therefore, the decision of lowering the cut-off point to 0.7, allowed in specific cases (Schneider and Wagemann 2010), was taken according to the literature.

TABLE 3 ABOUT HERE

The conservative solution leading to ∼TFR is the following:

$$\overbrace{\sim GEM * \sim FLFP * EPL}^{(a)} + \overbrace{\sim GEM * FLFP * \sim EPL * PLG}^{(b)} \rightarrow \sim TFR \tag{3}$$

Two equifinal paths lead to a not-high level of TFR. In contexts with low gender equality, low female employment, and rigid labor markets, fertility is not-high. At the same time, TFR is not-high in contexts where women are active in the labor market but have low levels of empowerment, and – while employment protection is low – parental leave generosity is high.

The solution's consistency is 0.81 and it has a coverage of 0.56. However, the two paths do not have the same relevance: they have the same consistency but path (3a) has a coverage of 0.46 while the coverage of path (3b) is 0.17. Further, the unique coverage of path (3a) is 0.38 while it only reaches 0.10 in path (3b). Path (3a), therefore, appears to be more relevant

in explaining ∼TFR. This solution defines two groups of cases with different characteristics. Cases members of (3a) – Greece, Italy and Spain – are clearly instances of the so-called familistic welfare regime (Ferrera 1996), and it is no surprise to find them clustered together. On the contrary, the Canadian-Japanese solution (path 3b) deserves some comments. While Canada could be, by and large, grouped together with the Anglo-Saxon countries, such as the US or Australia, for what concerns certain features (such as being a large albeit low populated country, with vast rural areas, high female employment, relatively low gender equity) it differs greatly in terms of fertility rates. Japan, on the other hand, could be imagined close to Southern European countries for its low levels of female empowerment and low fertility rates, but the high proportion of working women and especially its weak employment protection legislation clearly differentiate it from this cluster. In the light of this, the possibility of grouping together Canada and Japan could provide important insights in the study of fertility.

The next solution is the most parsimonious one:

$$\overbrace{\sim EPL * PLG}^{(a)} + \overbrace{\sim GEM * FLFP * EPL}^{(b)} \rightarrow \sim TFR \tag{4}$$

This solution does not achieve the consistency value of 0.8, but it provides similar results to the previous one. Given the consistency (0.77) and the coverage (0.59), it could be argued that this is not a good sufficient solution for not-high TFR. However, since the aim of our fsQCAs is the definition of groups of countries which will be used in the analysis of individual-level data, we briefly discuss it. The solution presents two paths. Path (4a) tells us that in countries where parenting leaves are generous but employment protection is low, fertility rates are also not-high. Canada and Japan are members of this configuration (consistency 0.71; raw coverage 0.21; unique coverage: 0.12). The second path, (4b) says that low levels of gender equality and low female labor force participation, combined with high employment protection, lead to not-high fertility. Instances of this configuration are Greece, Italy and Spain (consistency 0.81; raw coverage 0.46; unique coverage: 0.38).

6 Clustering the cases

In the previous section we have used fsQCA to empirically test the presence of clusters of countries to include in a GLM. In the present section, we define and name these clusters. To define the clusters, rather than aprioristically choosing one of the solutions that emerged, we use both solutions of the two analyses. This way, we allow the configurations of conditions leading to the outcomes to emerge more clearly, we avoid

overlapping clusters and we highlight cross-case patterns. To name the clusters, we use, when possible, the labels chosen by Gauthier (1998) in her public policy typology, and by Esping-Andersen (1990) in his welfare regime typology.

We can notice that the same cases belong to solutions (1b) and (2a), Denmark, Finland, the Netherlands, Norway and Sweden. We use solution (1b) because it presents a higher number of conditions which, by consequence, provide more information about the cluster itself. We name this group of countries, which is characterized by gender equality, female labor force participation and employment protection legislation, "Egalitarian". In solution (2b) we have five cases, Australia, Ireland, New Zealand, the United Kingdom and the United States, which also belong to solutions (1a) and (1c). To improve the parsimony of the classification we keep solution (2b) to define a cluster that we name "Non interventionist", characterized by not-high levels of EPL and not-high parental leave generosity. As far as the analysis of not-high TFR is concerned, the conservative solution is the best one. Compared to the most parsimonious one, it shows a better fit and it allows us to define two clear clusters, that we name "Traditional" and "Unequal". The "Traditional" group is characterized by not-high levels of gender equality, not-high levels of female labor force participation and high levels of employment protection legislation. The "Unequal" group, instead, has not-high levels of gender equality and of employment protection legislation, but on the contrary it has high levels of female labor force participation and of parental leave generosity. Despite the poor fit of the most parsimonious solution, it confirms the results of the previous one.

To sum up, following the results of the fsQCA, we have outlined the country groups in table 4.

TABLE 4 ABOUT HERE

While (a) and (b) are high fertility clusters of countries, (c) and (d) are those displaying not-high levels of fertility. We define a fifth group of countries comprising Austria, Belgium, Germany, France and Portugal for which no solution emerged and which can be considered a residual or remainder category. All countries of the group – with the exception of Portugal – belong to the same geographical area and are part of Esping-Andersen's "Continental" welfare regime. Although no solution emerged from the analysis of sufficient conditions, the Pseudo-Venn diagram (Schneider and Grofman 2006) in table 5 shows that the countries cluster along similar configurations of conditions. Since most of the non-explained countries present medium rather than high or low fertility rates, the absence of solutions for this group may be due to the way the analysis for sufficient conditions has been specified. We shall use this group as reference category in the micro-level analysis.

TABLE 5 ABOUT HERE

7 Micro-level analysis

In the previous paragraph we outlined the clusters that emerged from the fsQCA and we shall now use them in a micro-level analysis. We propose here a probit regression model where the outcome is higher order fertility. The data is derived by a harmonization of the World Values Survey (2009) and European Values Study (2011). The analysis is restricted to women aged 20-45, as standard procedure when studying fertility. Our dependent variable measures whether the respondent has had zero or one child, versus two or more. The scope of the analysis is to verify if (i) certain institutional contexts favor higher order births, and (ii) if the work-motherhood relation is affected by the institutional context. We know from the literature that specific contexts have higher fertility, therefore we expect a positive association between the probability of having two or more children and living in the "Egalitarian" and "Non interventionist" groups. To the contrary, we expect the outcome to be negatively associated with living in the "Traditional" and "Unequal" groups. Further, while being in paid work should be negatively associated to childbearing, given the opportunity-costs argument, in some contexts female labor force participation is positively associated to fertility rates. Therefore, we test whether the micro-level effect of working varies between country groups. The first model – the so-called pooled model – includes only a set of individual independent variables, namely: age, age squared, educational level (medium and high compared to low) and whether the respondent is working or not. Ideally, one would want to include other explanatory variables, for example educational level and employment status of the (eventual) partner, household income, religiosity and attitudes towards gender roles. In cross-national research, however, it is often difficult to obtain data that reaches acceptable standards of comparability on a large number of informative variables. Therefore, the researcher is often forced to sacrifice either the number of cases or the number of variables. In this case, we prefer maximizing the number of cases, although this means reducing the detail in the micro-level analysis. Model two includes country groups as dummies while model three adds the interaction between being in paid work and the country groups[8]. In both cases we use the countries belonging to the residual group as reference category. The interaction approach allows verifying if the effect of a variable changes conditional on another variable. In this case, it allows us to test if working status has different

[8]See table 11 in the Appendix for the micro variables descriptive statistics by country clusters.

effects across contexts[9].

For this analysis it results quite fruitful to have groups of countries in place of countries because the latter would have yielded a greater number of parameters, leading to an overparametrization of the models. Results are reported in table 6.

TABLE 6 ABOUT HERE

Model one – the pooled model – shows that, as expected, higher education and working negatively affect probability of having two or more children. Model two adds the country-groups. While the coefficients for education and working status remain roughly unchanged, a positive and statistically significant association emerges between living in the "Egalitarian", the "Non interventionist" and the "Unequal" group. Given the very high fertility rates of the "Non interventionist" cluster, it is not surprising to find the largest coefficient in these countries. The second largest effect is found in the "Egalitarian" countries, followed by the "Unequal" group. The "Traditional" context is the only one to be negatively associated to the probability of having two or more children. Given that we know the characteristics that define the "Traditional" cluster thanks to the fsQCA, we can interpret the negative effect on having a higher order birth as the result of living in a country with a rigid labor market, scarce female labor force participation and low levels of gender equality. This, by and large, confirms the hypothesis of McDonald (2000) regarding the relation between gender equality and childbearing. Notably, the "Unequal" group has a positive, albeit small, effect, meaning that the Canadian-Japanese combination favors fertility compared to the reference group. This could be due to the fact that in this group the flexibility of the labor market does not hinder family formation and fertility, but on the contrary allows a better reconciliation of motherhood and work duties (Adserà 2005). Lastly, model three includes the interaction effects. The main effects of the previous model remain unchanged in sign, although there is some loss in magnitude and in statistical significance. In particular, once the interaction between country groups and working status is included, the positive main association between belonging to the "Egalitarian" and the "Unequal" group and the outcome becomes not significant. Living in the "Non interventionist" and the "Traditional" clusters instead remains, respectively, positively and negatively associated to the outcome in a statistically significant way. The coefficient for working status is negative, suggesting that in the "Continental" cluster working is negatively associated to having a higher order birth. Working status is also negatively associated to the outcome in the

[9]Although we often use the term "effect" when referring to the association between variables, we are aware of the fact that due to the cross-sectional nature of our data it is not possible to speak of causal effect, only of associations.

"Non interventionist" context, while in "Egalitarian" countries being in paid work is positively related to the probability of having a second child. We gain a better understanding of the interaction terms by plotting predicted probabilities as in figure 1. The graph plots the predicted probabilities of working/not working and having two or more children by country group[10].

While there is an overall negative effect of working in all countries, in "Egalitarian" countries this effect is mitigated, meaning that in this context there appears to be little difference between women in or out of paid employment in terms of number of children. The effect is actually positive, albeit statistically not significant. Therefore, women in paid employment are partially sheltered by the institutional context and manage to have children and be in paid work better than in other contexts. The association between being in paid work and the probability of having two or more children is negative in all other country groups, but the negative effect of working is much more pronounced in the "Non interventionist" cluster.

FIGURE 1 ABOUT HERE

It could be argued that modeling the relationship between working status and high order fertility using probit models could be too simplistic. For this reason, in addition to the previous models, we propose the use of a recursive simultaneous bivariate probit model to estimate the joint probability of having children and working. Bivariate probit models are particularly useful when studying two outcomes that may be correlated with each other and have some common explanatory variables. To control for this possible source of endogeneity between working status and higher order fertility we estimate a model that has two equations with two dependent variables: y_1 denotes whether the respondent has had zero or one child versus two or more; y_2 denotes whether the respondent is working or not. This model is particularly useful for our purpose because it allows estimating:

a) the joint probability of being in different combinations of working status and presence of children;

b) the association between working and having a second child and its relation to the context.

The model takes the following form:

$$Prob[y_1 = 1, y_2 = 1|x_1, x_2] = \Phi_2(\boldsymbol{b_1'x_1} + \gamma y_2 + \boldsymbol{b_2'x_2}, \rho) \quad (5)$$

[10]Probabilities are calculated at the variable means.

Where Φ_2 represents the bivariate normal cumulative distribution function, and ρ the correlation between the two equations error terms[11]. The dependent variables of the first equation are: age in years; age in years squared; educational level in categories (medium and high compared to low); working status; country groups dummy variables ("Egalitarian", "Non interventionist", "Traditional", "Unequal", reference category is "Continental"); interaction terms between country groups dummies and working status.

The dependent variables in the second equation are age in years; age in years squared; educational level in categories and country groups dummy variables as in equation one. Note that y_2 (working status) is included as a predictor in the first equation, in order to obtain an estimate for the association between working status and presence of children[12].

As before, we run three nested models: a pooled model, a model with group dummies and a model with the interaction terms. Results are reported in table 7.

The dependent variables of the first equation are: age in years; age in years squared; educational level in categories (medium and high compared to low); working status; country groups dummy variables ("Egalitarian", "Non interventionist", "Traditional", "Unequal", reference category is "Continental"); interaction terms between country groups dummies and working status.

TABLE 7 ABOUT HERE

The bivariate probit estimates confirm the results from the probit models. Starting from the main effects, we find that education is positively associated to working status, as could be expected, and is negatively associated to the probability of having two or more children. Furthermore, the associations between country-group and higher order birth are the same that emerged from the probit models. As far as the endogenous variable – working status – is concerned, we find that is it negatively associated to the outcome, and that the negative association is stronger in the "Non interventionist" group. On the contrary, the interaction effect is positive in the "Egalitarian" countries. Therefore, even when we take into account that having additional children and working are likely to be joint decisions, we see that the association between the independent variables and our dependent variable of interest does not change. This can be seen as a

[11]For details about the estimation procedure and computational issues see Green (2000:849-856).

[12]Regarding the presence of a variable on both sides of the equations, Green (2000:852-853) points out that: "We can ignore the simultaneity in this model and we cannot in the linear regression model because, in this instance, we are maximizing the log-likelihood, whereas in the linear regression case, we are manipulating certain sample moments that do not converge to the necessary population parameters in the presence of simultaneity."

further confirmation of the results robustness and of the way institutions and contextual features relate to individual-level behavior.

8 Conclusion and discussion

"Quae non prosunt singula multa iuvant"[13]. Perhaps this quotation summarizes best the intention of this paper. The point we make is that while macro-level comparative analysis can provide important contextual information, it fails to enlighten micro-level mechanisms. Micro-level analysis, on the contrary, provides important insights on individual-level behavior, but may be less informative as regards the effects of context. This difficulty can be solved by using hierarchical models, but the number of level-two units must be sufficiently large to guarantee reliable estimates. The alternative solution to capture the contextual effects in individual-level analysis is the inclusion of dummy variables, usually as countries or clusters of countries. The latter is generally the most appropriate solution, as it reduces the number of parameters in the equation. The question, therefore, is how to cluster the countries in a meaningful, theoretically and empirically sound way.

What we suggested in this article was to a) identify a number of conditions associated with the outcome of interest, b) perform a fsQCA to obtain clusters and c) use the clusters in a GLM to account for contextual effects.

We investigate the motherhood fertility dilemma because the two phenomena that are implied, female labor force participation and fertility rates, have important implications for the well-being of populations and are currently being studied by a growing number of scholars. Thus, our work hopes to contribute to an existing and expanding field of research. Furthermore, this topic clearly presents a micro and a macro-level of analysis that overlap and must be kept into account and is therefore ideal for the two-step strategy that we use[14].

With TFR as outcome, we chose four conditions – female labor force participation, gender empowerment, generosity of parental leaves and employment protection legislation – that according to previous literature are

[13] What alone is not useful helps when combined. Ovid, Remedia amoris.

[14] Other authors propose the use of "two-step" approaches in the general framework of QCA (Schneider and Wagemann 2006; Schneider 2009; Mannewitz 2011). While these authors distinguish between "remote" and "proximate" conditions leading to an outcome, we use fsQCA to identify "contextual" conditions that interact with "individual" variables associated with the dependent variable of interest. Our approach, similarly to Schneider and Wagemann's, implies two analytical steps. Differently from these authors, however, we do not use fsQCA in the second step, but we combine it with a quantitative technique. As the second step of the analysis deals with individual-level behavior using survey data, we believe that applying large-N techniques is a more appropriate approach.

associated to fertility rates. After calibrating the conditions and outcome, we performed a fsQCA and obtained clusters of countries. At this point we ran probit and bivariate probit models estimating the likelihood of having two children or more and included the clusters in the equation, while controlling for relevant individual-level information. Having performed the fsQCA, we may now, albeit with caution, provide some empirically justified interpretation of the "group effect", because we are partially modeling country-specific characteristics. As a final step, we allow for an interaction between context and being in paid employment, and once again can attempt an empirically based explanation of why the individual effect varies between groups of countries. What can be said, in the end, about the relation between working status and childbearing? The fsQCA results showed that at the macro-level high fertility is reached in two different contexts. In one case, it is associated to sky-high levels of FLFP while in others to medium levels of female labor force participation. As far as the micro-level is concerned, the individual-level analysis provided two important pieces of information. One regards the baseline probability of having two or more children, that is much higher in the "Non interventionist" context than elsewhere. This means that, regardless of working status, women in this group have higher fertility. Notwithstanding, it is not the objective of the present analysis to investigate why this is the case, as there is a longstanding debate dealing with the topic. Among others, McDonald and Moyle (2010) in their article by the interrogative title "Why do English-speaking countries have relatively high fertility?", highlight a number of possible explanations, ranging from demographic to compositional, from access to contraception to value orientation. Again, our objective was not to identify the sources of high fertility, rather to investigate the relationship between work and motherhood. To this end, a second relevant piece of information emerges from figure 1: in the "Egalitarian" context, employment is positively associated to the presence of children, while in the "Non interventionist" context, the relationship is negative. The interpretation we suggest is that the same outcome – high fertility – is reached in two different contexts because of the way different conditions are combined. In "Egalitarian" countries, at the micro-level, working does not hinder childbearing, women do not opt out of the labor market when they have children, and therefore country level fertility and FLFP are high. Employment protection legislation and gender empowerment might be critical conditions to ensure women and mothers the same chances of being in the labor market. Instead, in "Non interventionist" countries, at the micro-level, being employed negatively affects childbearing, women exit the labor market when they have children, and country level fertility rates will be high while FLFP will be lower. Contextual traits such as the non generous nature of the parental leave schemes and the low levels of employment protection may not favor mothers' presence in the work force.

In other words, while most women have children, those who work are less likely to have many children. A note of caution on this interpretation is due: we do not make any causal claims on the relation between employment status and presence of children, nor on the nature of the relationship between macro-level circumstances and individual situations. What we suggest is that the macro-level characteristics we consider are associated to the "working side" of the work-motherhood relation. In other words, certain configurations of contextual traits can reduce the incompatibility of working and having children, but do not, according to our results, have an effect on the overall probability of having children. To conclude, we wish to stress the relevance of both levels of analysis, macro and micro, even when methodological and technical drawbacks may be hard to overcome, because they allow achieving a deeper understanding of the phenomena under investigation.

References

Aaberge, Rolf, Ugo Colombino, Daniela Del Boca, John Ermish, Marco Francesconi, Silvia Pasqua, and Steinar Strom. 2005. "Labor Supply and Fertility in Europe and the US." In Women and Work. An Econonomic Perspective, edited by Boeri Tito, Daniela Del Boca, and Christopher Pissarides pp. 125–153. Oxford University Press.

Adserà, Alìcia. 2005. "Where Are the Babies? Labor Market Conditions and Fertility in Europe." IZA Discussion Papers.

Ahn, Namkee and Pedro Mira. 2002. "A Note on the Changing Relationship between Fertility and Female Employment Rates in Developed Countries." Journal of Population Economics 15:667–682.

Aus, Jonathan P. 2009. "Conjunctural Causation in Comparative Case-Oriented Research." Quality and Quantity 43:173–183.

Barbieri Paolo, Giorgio Cutuli, and Marco Tosi. 2012. "Poor but Handsome." FamIne Working Paper Series 2.

Becker, Gary S. 1981. "Altruism in the Family and Selfishness in the Market Place." Economica, New Series 48:1–15.

Becker, Gary S. 1991. A Treatise on the Family: Enlarged Edition. Harvard: Harvard University Press.

Berg-Schlosser, Dirk, Gisèle De Meur, Benôıt Rihoux, and Charles C. Ragin. 2008. "Qualitative Comparative Analysis (QCA) as an Approach." In Configurational Comparative Methods. Qualitative Comparative Analysis (QCA) and Related Techniques, edited by B. Rihoux and C. C. Ragin, pp. 1–18. London: Sage.

Berk, Richard A. 2005. Regression Analysis: A Constructive Critique. Thousand Oaks: Sage.

Braumoeller, Bear F. 2003. "Causal Complexity and the Study of Politics." Political Analysis 11:209–233.

Bongaarts, John and Tómas Sobotka. 2012. "A Demographic Explanation for the Recent Rise in European Fertility." Population and Development Review 38:83–120.

Castles, Francis G. 2003. "The World Turned Upside Down: Below Replacement Fertility, Changing Preferences and Family-Friendly Public Policy in 21 OECD Countries." Journal of European Social Policy 13:209–227.

D'Addio, Anna Cristina and Marco Mira D'Ercole. 2005. "Trends and Determinants of Fertility Rates in OECD Countries: The Role of Policies." OECD Social, Employment and Migration Working Papers.

Del Boca, Daniela. 2003. "Why Are Fertility and Participation Rates so Low in Italy (and Southern Europe)?" Paper prepared for presentation at the Italian Academy at Columbia University, New York.

Del Boca, Daniela, Silvia Pasqua, and Chiara Pronzato. 2009. "Motherhood and Market Work Decisions in Institutional Context: a European

Perspective." Oxford Economic Papers 61:147–171.

Engelhardt, Henriette, Tomas Kogel, and Alexia Prskawetz. 2001. "Fertility and Female Employment Reconsidered." MPIDR Working Paper WP 2001-021. Max Planck Institute for Demographic Research.

Engelhardt, Henriette and Alexia Prskawetz. 2004. "On the Changing Correlation Between Fertility and Female Employment over Space and Time." European Journal of Population 20:35–62.

Esping-Andersen, Gosta. 1990. The Three Worlds of Welfare Capitalism. Princeton: Princeton University Press.

Esping-Andersen, Gosta. 1999. Social Foundations of Postindustrial Economies. Oxford: Oxford University Press.

Estévez-Abe, Margarita. 2005. "Gender Bias in Skills and Social Policies: The Varieties of Capitalism Perspective on Sex Segregation." Social Politics: International Studies in Gender, State & Society 12:180–215.

European Values Study. 2011. European Values Study 1981-2008, Longitudinal Data File. GESIS Data Archive, Cologne, Germany, ZA4804 Data File Version 1.0.0 (2011-04-30) DOI:10.4232/1.14804.

Ferrera, Maurizio. 1996. "The "Southern Model" of Welfare in Social Europe." Journal of European Social Policy 6:17–37.

Gauthier, Anne H. 1998. The State and the Family: A Comparative Analysis of Family Policies in Industrialized Countries. Oxford University Press.

Gauthier, Anne H. 2007. "The Impact of Family Policies on Fertility in Industrialized Countries: a Review of the Literature." Population Research and Policy Review 26:323–346.

Gauthier, Anne H. 2010. "Comparative Family Policy Database, Version 3 [computer file]." Technical report, Netherlands Interdisciplinary Demographic Institute and Max Planck Institute for Demographic Research (distributors).

Gauthier, Anne H. and Jan Hatzius. 1997. "Family Benefits and Fertility: An Econometric Analysis." Population Studies 51:295–306.

Goldstein, Joshua R., Tomá? Sobotka, and Aiva Jasilioniene. 2009. "The end of lowest-low fertility?." Population and Development Review 35:663–700.

Gornick, Janet, Marcia Meyers, and Katherine Ross. 1997. "Supporting the Employment of Mothers: Policy Variation across Fourteen Welfare States." Journal of European Social Policy 7:45–70.

Greene, William 2000. Econometric Analysis. Upper Saddle River: Prentice Hall. 4th edition.

Hakim, Catherine. 2000. Work-Lifestyle Choices in the 21st Century: Preference Theory. Oxford: Oxford University Press.

Hall, Peter A. and David Soskice. 2001. "Varieties of Capitalism: an Introduction." In Varieties of Capitalism: The Institutional Foundations

of Comparative Advantage, edited by Peter Hall and David Soskice, pp. 1–68. New York: Oxford University Press.

Kohler, Hans Peter, Francesco Billari, and José Antonio Ortega. 2002. "The Emergence of Lowest-Low Fertility in Europe During the 1990s." Population and Development Review 28:641–680.

Korpi, W. (2000). Faces of Inequality: Gender, Class, and Patterns of Inequalities in Different Types of Welfare States. Social Politics 7:127-91.

Kvist, Jon. 1999. "Welfare Reform in the Nordic Countries in the 1990s: Using Fuzzy-Set Theory to Assess Conformity to Ideal Types." Journal of European Social Policy 9:231–252.

Kvist, Jon. 2007. "Fuzzy Set Ideal Type Analysis." Journal of Business Research 60:474–481.

Lijphart, Arendt. 1971. "Comparative Politics and the Comparative Method." American Political Science Review 65:682–693.

Maas, Cora J. M., and Joop J. Hox. 2005. "Sufficient Sample Sizes for Multilevel Modeling." Methodology 1:86-92.

MacCallum, Robert C., Keith F. Widaman, Shaobo Zhang, and Sehee Hong. 1999. "Sample Size in Factor Analysis." Psychological Methods 4:84–99.

Mahoney, James and Gary Goertz. 2006. "A Tale of Two Cultures: Contrasting Quantitative and Qualitative Research." Political Analysis 14:227–249.

Mannewitz, Tom 2011. "Two-Level Theories in QCA: A Discussion of Schneider and Wagemann's Two-Step Approach." Compasss WP 2011-64.

McDonald, Peter. 2000. "Gender Equity in Theories of Fertility Transition." Population and Development Review 26:427–439.

McDonald, Peter and Helen Moyle. 2010. "Why do English-speaking countries have relatively high fertility?" Journal of Population Research 27:247–273.

Ni Bhrolchain, Marie. 1992. "Period Paramount? A Critique of the Cohort Approach to Fertility." Population and Development Review 18:599-629.

Ni Bhrolchain, Marie. 2011. "Tempo and the TFR." Demography 48:841–61.

OECD. 2004. "Employment Outlook." (Accessed July 2011).

OECD. 2009. "OECD StatExtracts. Labor - Labor Force Statistics – Annual Labour Force Statistics – ALFS Summary Tables." (Accessed July 2009).

Ragin, Charles C. 2000. Fuzzy-Set Social Science. Chicago and London: The University of Chicago Press.

Ragin, Charles C. 2005. "Core Versus Tangential Assumptions in Comparative Research." Studies in Comparative International Development 40:33–38.

Ragin, Charles C. 2008. Redesigning Social Inquiry: Fuzzy Sets and Beyond. Chicago: The University of Chicago Press.

Schneider, Carsten Q. 2009. The Consolidation of Democracy. Comparing Europe and Latin America. London and New York: Routledge.

Schneider, Carstern Q. and Bernard Grofman. 2006. "It Might Look like a Regression Equation But It's Not! An Intuitive Approach to the Representation of QCA and FS/QCA Results." Compasss: WP 2006–39.

Schneider, Carsten Q. and Claudius Wagemann 2006. "Reducing Complexity in Qualitative Comparative Analysis (QCA): Remote and Proximate Factors and the Consolidation of Democracy." European Journal of Political Research 45:751–86.

Schneider, Carsten Q. and Claudius Wagemann. 2010. "Standards of Good Practice in Qualitative Comparative Analysis (QCA) and Fuzzy-Sets." Comparative Sociology 9:397–418.

Seawright, Jason. 2005. "Qualitative Comparative Analysis Vis-A-Vis Regression." Studies in Comparative International Development 40:3–26.

Snijders, Tom A. B. and Roel Bosker. 1999. Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. London: Sage.

Tarrow, Sidney. 2004. "Bridging the Quantitative-Qualitative Divide." In Rethinking Social Inquiry: Diverse Tools, Shared Standards, edited by H. Brady and D. Collier, pp. 171–180. Lanham: Rowman & Littlefield.

Thévenon, Olivier. 2011. "Family Policies in OECD Countries: A Comparative Analysis." Population and Development Review 37:57–87.

UNDP. 2007. Human Development Report 2007/2008. Houndmills and New York: Palgrave Macmillan.

Wagemann, Claudius. 2007. "QCA e 'fuzzy set analysis'. Che cosa è e che cosa non è." Rivista Italiana di Scienza Politica 37:385–409.

World Value Survey. 2009. World Value Survey 1981-2008 official aggregate v.20090902, 2009. World Values Survey Association (www.worldvaluessurvey.org). Aggregate File Producer: ASEP/JDS Data Archive, Madrid, Spain.

Tables and Figure

Table 1: Thresholds for calibration.

	TFR	GEM	FLFP	EPL	PLG
Upper threshold	1.8	0.9	73%	2	0.63
Cross-over	1.6	0.84	63%	1.45	0.31
Lower threshold	1.4	0.6	55%	1	0.15

Table 2: Truth table for high Total Fertility Rate.

Conditions					Outcome		
GEM	FLFP	EPL	PLG	*N*	TFR	Consistency	Cases
1	1	0	0	1	1	0.981	AUS
0	1	0	0	3	1	0.936	NZ, UK, US
1	1	1	0	1	1	0.841	NL
0	0	0	0	1	1	0.836	IE
1	1	1	1	4	1	0.816	DK, FI, NO, SE
1	0	1	0	1	0	0.642	BE
0	1	1	0	1	0	0.566	PT
0	1	0	1	2	0	0.534	CA, JP
0	1	1	1	3	0	0.478	AT, DE, FR
0	0	1	1	1	0	0.419	IT
0	0	1	0	2	0	0.337	GR, ES

Table 3: Truth table for not-high Total Fertility Rate.

Conditions					Outcome		
GEM	FLFP	EPL	PLG	*N*	∼TFR	Consistency	Cases
0	0	1	0	2	1	0.834	GR, ES
0	1	0	1	2	1	0.812	CA, JP
0	0	1	1	1	1	0.743	IT
0	1	1	1	3	0	0.686	AT, DE, FR
0	1	1	0	1	0	0.669	PT
1	0	1	0	1	0	0.627	BE
0	0	0	0	1	0	0.395	IE
1	1	1	0	1	0	0.388	NL
1	1	1	1	4	0	0.306	DK, FI, NO, SE
1	1	0	0	1	0	0.252	AUS
0	1	0	0	3	0	0.246	NZ, UK, US

Table 4: Country clusters.

	Group	Path	Countries
(a)	Emancipative-Egalitarian	EPL $*$ FLFP $*$ GEM	DK, FI, NL, NO, SE
(b)	Liberal-Non inteventionist	$\sim$PLG $*$ $\sim$EPL	AUS, IE, NZ, UK, US
(c)	Traditional	$\sim$GEM $*$ $\sim$FLFP $*$ EPL	GR, IT, ES
(d)	Unequal	PLG $*$ $\sim$EPL $*$ $\sim$GEM	CA, JP
(e)	Continental	No solution	AT, BE, FR, DE, PT

Table 5: Pseudo-Venn diagram.

	GEM $*$ FLFP	GEM $*$ $\sim$FLFP	$\sim$GEM $*$ FLFP	$\sim$GEM $*$ $\sim$FLFP
EPL $*$ PLG	**DK, FI, NO, SE**		AT, DE, FR	*IT*
EPL $*$ $\sim$PLG	**NL**	BE	PT	*GR, ES*
$\sim$EPL $*$ PLG				*CA, JP*
$\sim$EPL $*$ $\sim$PLG	**AUS**		**NZ, UK, US**	IE

Bold cells represent positive outcome (TFR)
Italic cells represent negative outcome ($\sim$TFR)

Table 6: Probit models. Dependent variable: having zero or one child vs. two or more children.

	Model 1	Model 3	Model 3
	Coef./(s.e.)	Coef./(s.e.)	Coef./(s.e.)
Age	0.244***	0.268***	0.261***
	(0.027)	(0.028)	(0.028)
Age sq./100	-0.254***	-0.283***	-0.272***
	(0.040)	(0.041)	(0.041)
Not married (ref. cat.)			
Married	0.671***	0.716***	0.719***
	(0.037)	(0.038)	(0.038)
Low Ed. (ref. cat.)			
Medium Ed.	-0.305***	-0.268***	-0.251***
	(0.049)	(0.051)	(0.051)
High Ed.	-0.555***	-0.605***	-0.591***
	(0.051)	(0.053)	(0.053)
Not working (ref. cat.)			
Working	-0.249***	-0.268***	-0.306***
	(0.041)	(0.042)	(0.075)
Continental (ref. cat.)			
Egalitarian		0.367***	0.071
		(0.108)	(0.108)
Non interventionist		0.729***	0.938***
		(0.099)	(0.099)
Traditional		-0.237***	-0.296**
		(0.099)	(0.099)
Unequal		0.193**	0.112
		(0.060)	(0.108)
Working × Egalitarian			0.377**
			(0.122)
Working × Non interventionist			-0.302*
			(0.117)
Working v Traditional			0.091
			(0.120)
Working × Unequal			0.117
			(0.130)
Constant	-5.211***	-5.901***	-5.756***
	(0.445)	(0.462)	(0.466)
N		6191	

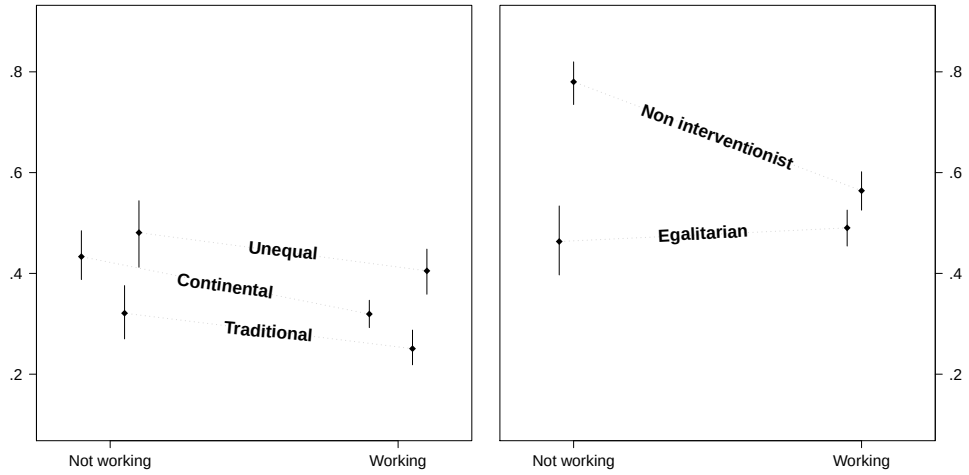
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Recursive simultaneous bivariate probit models. Dependent variables: y_1 = having zero or one child vs. two or more children; y_1 = working vs. not working.

	Model 1		Model 2		Model 3	
	Coef./(s.e.)		Coef./(s.e.)		Coef./(s.e.)	
	Two or more children	Working	Two or more children	Working	Two or more children	Working
Age	0.265***	0.140***	0.288***	0.139***	0.283***	0.141***
	(0.027)	(0.023)	(0.027)	(0.023)	(0.027)	(0.023)
Age sq./100	-0.280***	-0.162***	-0.310***	-0.162***	-0.300***	-0.165***
	(0.040)	(0.076)	(0.038)	(0.035)	(0.039)	(0.034)
Not Married (ref. cat.)						
Married	0.634***	-0.076*	0.636***	-0.061	0.667***	-0.064
	(0.049)	(0.038)	(0.051)	(0.038)	(0.050)	(0.038)
Low Ed. (ref. cat.)						
Medium Ed.	-0.230***	0.291***	-0.149*	0.268***	-0.166**	0.269***
	(0.066)	(0.045)	(0.060)	(0.045)	(0.062)	(0.045)
High Ed.	-0.395***	0.729***	-0.340***	0.695***	-0.397***	0.694***
	(0.109)	(0.047)	(0.093)	(0.048)	(0.100)	(0.048)
Not Working (ref. cat.)						
Working	-0.816**		-1.144***		-1.002***	
	(0.296)		(0.198)		(0.263)	
Continental (ref. cat.)						
Egalitarian			0.361***	0.096	0.114	0.086
			(0.051)	(0.052)	(0.102)	(0.052)
Non interventionist			0.634***	-0.122*	0.846***	-0.116*
			(0.065)	(0.050)	(0.110)	(0.050)
Traditional			-0.280***	-0.215***	-0.344***	-0.216***
			(0.055)	(0.051)	(0.094)	(0.052)
Unequal			0.149*	-0.101	0.075	-0.102
			(0.059)	(0.057)	(0.103)	(0.057)
Working × Egalitarian					0.326**	
					(0.117)	
Working × Non interventionist					-0.259*	
					(0.113)	
Working × Traditional					0.107	
					(0.114)	
Working × Unequal					0.127	
					(0.122)	
Constant	-5.280***	-2.609***	-5.709***	-2.524***	-5.735***	-2.546***
	(0.438)	(0.366)	(0.479)	(0.368)	(0.471)	(0.368)
N	6191					

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 1: Predicted probabilities with 95% confidence intervals of having two or more children for not working and working women in the five clusters of countries.



Appendix

Test of necessary conditions for TFR.

	Consistency	Coverage
GEM	0.627204	0.830000
∼GEM	0.455919	0.493636
FLFP	0.823678	0.714494
∼FLFP	0.230059	0.437002
EPL	0.585223	0.524849
∼EPL	0.450042	0.797619
PLG	0.481108	0.573000
∼PLG	0.588581	0.701000

Test of necessary conditions for ∼TFR.

	Consistency	Coverage
GEM	0.311496	0.280000
∼GEM	0.810878	0.596364
FLFP	0.563659	0.332119
∼FLFP	0.515451	0.665072
EPL	0.831891	0.506777
∼EPL	0.220025	0.264881
PLG	0.630408	0.510000
∼PLG	0.472188	0.382000

Intermediate solution for TFR.

	Raw coverage	Unique coverage	Consistency
∼PLG ∗ ∼EPL	0.436608	0.327456	0.940325
EPL ∗ FLFP ∗ GEM	0.447523	0.338371	0.843354
Solution coverage	0.774979		
Solution consistency	0.879048		

Assumptions: FLFP (present), GEM (present)

Intermediate solution for ∼TFR.

	Raw coverage	Unique coverage	Consistency
PLG ∗ ∼EPL ∗ ∼GEM	0.185414	0.100124	0.819672
EPL ∗ ∼FLFP ∗ ∼GEM	0.462299	0.377009	0.811280
Solution coverage	0.562423		
Solution consistency	0.813954		

Assumptions: ∼FLFP (present), ∼GEM (present)

Descriptive statistics by country cluster.

	More kids	Working	Age	Age sq.	Married	Low Ed.	Medium Ed.	High Ed.	N
	%	%			%	%	%	%	
Continental	40	70	34	1206	50	18	50	32	1834
Egalitarian	49	77	34	1230	52	10	32	57	1294
Non interventionist	61	66	33	1164	51	28	34	38	1214
Traditional	34	62	33	1149	54	24	42	35	1051
Unequal	50	67	34	1207	65	16	53	31	798
Total	47	68	34	1191	55	19	42	39	6191

Raw and fuzzy values for the outcome and the conditions.

	Raw values					Fuzzy values				
Country	TFR	GEM	FLFP	EPL	PLG	TFR	GEM	FLFP	EPL	PLG
Australia	1.79	0.85	69	1.15	0.00	0.95	0.5	0.86	0.11	0
Austria	1.41	0.79	65.8	1.93	0.57	0.05	0.33	0.7	0.93	0.92
Belgium	1.76	0.85	59.8	2.18	0.29	0.92	0.5	0.23	0.98	0.39
Canada	1.54	0.82	73	0.75	0.44	0.29	0.41	0.95	0.01	0.75
Denmark	1.8	0.88	76	1.5	0.33	0.95	0.86	0.98	0.53	0.53
Finland	1.8	0.89	73.2	1.96	0.66	0.95	0.92	0.96	0.94	0.96
France	1.92	0.72	65.1	3.05	1.00	0.99	0.17	0.65	1	1
Germany	1.34	0.83	67.4	2.12	0.47	0.02	0.44	0.79	0.98	0.81
Greece	1.34	0.62	53.4	2.73	0.13	0.02	0.06	0.03	1	0.03
Ireland	1.88	0.7	60.8	1.11	0.21	0.99	0.14	0.3	0.09	0.12
Italy	1.32	0.69	50.7	1.89	0.48	0.01	0.13	0.01	0.91	0.83
Japan	1.26	0.56	65.5	1.43	0.46	0.01	0.03	0.68	0.4	0.8
Netherlands	1.71	0.86	70.1	1.95	0.25	0.84	0.65	0.89	0.94	0.23
New Zealand	1.97	0.81	71.1	1.4	0.22	1	0.38	0.92	0.35	0.14
Norway	1.84	0.91	75.4	2.69	0.68	0.97	0.97	0.98	1	0.97
Portugal	1.41	0.69	71.7	3.15	0.27	0.05	0.13	0.93	1	0.3
Spain	1.35	0.79	58.4	2.98	0.25	0.02	0.33	0.15	1	0.23
Sweden	1.77	0.91	76	1.87	0.72	0.93	0.97	0.98	0.9	0.98
United Kingdom	1.79	0.78	68.9	0.75	0.09	0.95	0.3	0.85	0.01	0.01
United States	2.05	0.76	69.9	0.21	0.00	1	0.25	0.89	0	0