Ca' Foscari University of Venice

Joint Degree programme
in quantitative methods (QEM)
laurea magistrale

Final Thesis

Parenthood decision making

The Spanish case

Supervisor
Ch. Prof. Carlo Santagiustina

Graduand
María Gallego Tomás 884870

Academic Year
2020 / 2021

Introduction

An essay on the Principle of Population (1872) written by Malthus exposed that economic and population cycles were linked, so the movements of fertility are only explained by its correction means. Across the first demographic transition the fertility rate changed from this cyclical to a declining pattern, and during the second one the birth rate started to follow a constant path. Several studies have proposed theories to explain the reasons for this change.

In Mulder (1998) three different hypotheses for demographic transition were discussed. The first one (Smith and Fretwell, 1974 and Kaplan, 1996) suggests that the number of offspring per year declined because of the competitive environment in which they are raised, where high economic investment would be needed to provide them a prosperous future. The second hypothesis (Boyd and Richerson, 1988), called indirect *bias*, suggests that people tend to imitate the parenthood decisions that most successful individuals in a community have made. The third one (Pérusse, 1993) states that the decreasing of fertility rates is caused by the non-adaptiveness of social status with reproductive and mating data, thus with newborns figures. This previous hypothesis presents some difficulties, so some psychological decision-making process must be considered.

Becker (1987) argued that education investment is negatively linked with fertility figures. Due to the necessity of making such an investment, women have to invest more time and effort in their professional careers, leaving no time for having children. This is explained by the rising of the opportunity cost of motherhood, see Becker (1960).

From a social perspective let us mention the innovation theory and the adjustment theory (Carlsson, 1966). The author gives more emphasis to the adjustment theory, highlighting the importance of the motivation and social situation that affects the parenthood decision. In contrast, the diffusion perspective (Cleland and Wilson; Casterline, 2001) supports that fertility decline was caused by the spread of new ideas and behaviours as for the spread of knowledge about contraceptives. This theory is more in line with the innovation theory than the adjustment theory, where socio-economic changes are considered more important (Carlsson, 1966).

Social interactions are crucial when a parenthood decision is being taken (Bongaarts, & Watkins, 1996; Montgomery and Casterline, 1996; Bernardi et al, 2007; and Balbo and Barban, 2014). The widespread diffusion of social media gives the opportunity to search for opinions and behaviours from people that we consider successful in a rapid and efficient way. This way of sharing information is sometimes even more efficient than personal interactions. New ways of retrieving information from the web, such as Google searches (Billari et al., 2016 and Wilde et al 2020) and twitter data (Sulis, 2016) have been already used to predict fertility rates and study the behaviour towards parenthood, respectively.

Southern European countries are among the ones that in the last 50 years showed the lowest fertility levels (Matysiak et al, 2021). None of the previous studies Spain, which will be the focus of this project. The first demographic transition has already finished in Spain, like in other occidental countries. Spanish fertility had been declining throughout the 20th century, but during Franco's dictatorship the figures followed a constant pattern.

However, it must be mentioned that a baby boom happened in this political period, from the 50' until late 60' (Fernández Cordón, 1986; Rowland, 2007). A timid growth in the natality was again experimented in Spain during the first years of the 21st century. This tendency was then interrupted by a natality decline due to the 2009 economic crisis (Miret, 2019; Puig-Barrachina et al, 2020). Nowadays, the future of Spain's fertility rate is more uncertain because of the COVID-19 crisis, among other issues. Some recent studies are proposing for other EU countries, like Italy, similar approaches to the one developed in this work (Luppi, Arpino and Rosina, 2021).

The aim of this study is to describe how different variables that may affect the parenthood decision-making in Spain can be proxied using Google Trends data, macroeconomic data, and the CIS survey from 2004 to 2020.

Recently new types of data sources, such as Google Trends, have been used to proxy preferences and socio-cultural factors that, according to literature, could play a role in parenthood decisions. These variables can capture time varying characteristics of the Spanish population that are not captured by macroeconomic variables. For this reason, in this work both types of variables are going to be used.

The rest of the paper is organized as follows: In section 2, literature review of decision-making patterns and previous fertility models. In section 3, description of the data strategy and of the variables employed in the models. Section 4 presents the models, the regression analysis and the main findings. Finally in section 5, the results are summarised, contextualized and discussed in relation to existing literature.

Literature review

Previous studies have given some light to parenthood decision-making. Some of them focus on who participates in the decision. Is it a decision made by the couple? Do the two members have similar decisional roles and power, or one of them has more to say than the other?

Let us present some models where the decision power does not depend neither on the gender nor on the social status. The golden mean rule (Jansen and Liefbroer, 2006) and the social drift rule (Bauer and Kenip, 2014; Jansen and Liefbroer, 2006) proposed that when the partners agree the decision output is clear, however when this situation does not occur the one that opposites the changes with respect to the status quo is the one who wins the negotiation. Also, the joint utility model states that the path followed is the one who reports a higher utility than costs, so gender-economic variables do not play a role. This proposal it is in line with the higher opportunity costs that affect women when having children, it is rational to think that the one that is more affected by the decision is the one with more decision-making power, However, a more gender egalitarian society may lead to an even power distribution (Duvander et al, 2019). Other models stated that these personal characteristics do affect parenthood decision-making power. On one hand, the patriarchal rule (Bauer and Kenip, 2014) gives more power to the one that has a better economical position (many times it is the men). On the other hand, the matriarchal rule (Jansen and Liefbroer, 2006) states that women have more power when it comes to the childbearing decision.

When asking childless women and men the reasons that influenced their decision, it comes out that men's decisions are more individualistic than women`s. Women take into account the quality of life that the children would have, whereas men think more on their self's quality of life after the child is born (Blackstone and Stewart, 2016; Park, 2005). This goes in line with the results reported on Agrillo e Nelini (2008) that show that men take parenthood decisions in a more efficient way. The individualistic character of men also plays a role in the way of deciding, as women emphasize more often than men that it is a relationship decision (Blackstone and Stewart, 2016).

Other studies focus on the decision-making process, is it a one-day decision or is it a far-sighted and continuous process? On one hand Blackstone and Stewart (2016) suggested that only one conversation may be needed to decide, but this can happen because the partners have previously made the decision individually and later on have discussed their decisions in just one conversation. On the other hand, some studies have reported that remaining childless is a more conscious decision, as it requires the use of contraceptives to actively prevent pregnancy (Blackstone and Stewart, 2016), whereas for having children no preventive behaviour is needed.

As for the timing of the decision process, some studies (Blackstone and Stewart, 2016; Settle and Brumley, 2014) suggest that it is a decision revisited every so often, as the couple's situation changes over the time, as so for, the variables that affect the discussion.

There are some variables that are known in demographic research to be good predictors of the fertility rate. It is a well-known result from the literature that the education level is negatively correlated with the number of new-borns (Campisi et al, 2020; Cleland and Wilson, 1987), this result has also been proposed for the Spanish case (Requena and Salazar, 2014; Hicks and Martinez-Aguado, 1987). This may be explained by increases of the cost of having children (Caldwell, 1980) or by changes of parenthood timing preferences due to changing studying and working opportunities (Kulu and Vikat, 2007; Sánchez-Barricarte (2019).

On one hand, the unemployment rate has a well-known negative effect on the fertility rate (Adsera, 2004; Campisi et al, 2020, Puig-Barrachina et al, 2020). This relationship is also true in Spain when studying the employment rate which has a positive effect on the number of new-born, because it increases the individual's purchasing power so deciding to become a parent is more likely (Sánchez-Barricarte, 2019). The duration of the unemployment period also plays a role on the fertility rate, i.e., long term unemployment has a higher effect (Busetta, Mendola and Vignoli, 2019). In the same way the expectation on the labour market affects the variable of study (Gerson, 1993).

On the other hand, parenthood has also an effect on the individual's working situation, as finding an employment is more challenging for them (Adsera, 2011) as well as re-entering on the full-time labour force market (Rønsen, M., & Sundström, M., 2002). Google Trends data about the number of searches of unemployment related topics and words have been used to proxy its influence on the fertility rate (Wilde, Chen and Lohmann, 2020).

The economic and social uncertainty about the future is also an important factor that affects the fertility rate (Comolli et al, 2021; Goldstein et al, 2013). Individuals reported that one of the reasons why they do not have children is an unexpected labour market instability (Gerson, 1993). The uncertainty also affects the relationship's duration, increasing the number of dissolutions, reducing the likelihood of having children (Jansen and Liefbroer, 2006). Another study determined that parenthood is postponed during periods with high uncertainty (Campisi et al, 2020).

Some social extremist social movements that may influence the population's way of thinking may also affect the parenthood decision-making process, by defending or questioning the traditional family model, and may hence also affect the number of new-borns. Let us focus on feminism and far-right movements. On one hand, the feminism movement increased the number of women in the labour market, this diminished the likelihood of having children, due to the lack of time (Gillespie, 2013; Ireland, 1993). Social movements that give women more freedom to break with their traditional family role will have a negative effect on the fertility rate (Meggiolaro and Ongaro, 2007). On the other hand, the increase of the far-right movement's popularity are often linked to periods of uncertainty (Campisi et al, 2020). In Norway, this type of government has implemented incentives to increment the fertility (Aassve and Lappegård, 2009) so it is reasonable to think that the popularity decrease of the conservative parties may lead to a decrease of the new-borns (Campisi et al, 2020).

Lastly, the popularity of the contraceptives methods comes from the first half of the 20th in the Spanish case (Nicolau-Nous, 1991). However, it has been proposed as an underlying explanation of the fertility decline produced during the first demographic transition (Coale and Watkins, 1986). Contraceptives searches in google have been to explain the future fertility rate (Wilde, Chen and Lohmann, 2020).

The models used in previous demographic researches are vector error correction models (VECM) (Hodryiannis and Papapetrou, 2002; Narayan and Peng, 2006; Herter et al, 2014; Frini and Miller, 2012), dynamic ordinary least squares (DOLS) (Hodryiannis, 2010; Sánchez Barricarte, 2017; Halner and Mayer-Foulkes, 2013; Herzer et al, 2014), Fully modified ordinary least squares (FMOLS) (Sánchez Barricarte, 2017; Hartani et al 2015; Bakar et al, 2014), random effects spatial panel regression models (Campisi et al 2020). The studies that used Google Trends employed the following models: ordinary least squares with variable and fixed effects (OLS) (Wide, Chen and Lohnman, 2020) and autoregressive and moving average model (ARMA) (Billari, D'Amuri and Marcucci, 2013). In the present study the last one mentioned is going to be estimated.

Table 1

Variables	Effect on the fertility rate
Unemployment	Negative effect
Uncertainty	Negative effect
Feminism	Negative effect
Far-right movement	Positive effect
Contraceptives	Negative effect

The present research paper's purpose is to evaluate the effects of the different variables which are hypothesized to influence the fertility rate. Due to the lack of macroeconomic data for some of them we are going to ad-hoc time series constructed using use Google Trends (Ballarin, D'Amiuri and Marcucci, 2013; Wilde, Chen and Lohmann, 2020), in other cases well-known macroeconomic variables and the CIS-survey data will be used.

Data

The variables that are going to be used are the following: contraceptive, ovulation test, morning after pill, pregnancy test and in vitro, pregnancy, feminism, extreme right and fecundation are proxied through search engine query frequencies from Google Trends. Also the economic and policy uncertainty (Baker, Bloom and Davis, 2015) and the unemployment rate are based on official statistical data from EPU webpage and OCDE respectively. From the CIS survey variables that capture the worry about the pension's problem, education's problem, actual economic situation, actual political situation, expectations on the expectations on the economy future situation, expectations on the political situation and the far right/left sentiments. Lastly, the fertility rate data comes from the Instituto Nacional de Estadística (INE). For each of these variables we have a time-series of the monthly frequency, for the time interval ranging from 2014M12 to 2020M06. The following the variables are going to be explained more in depth as well as the transformations required to do the statistical analysis.

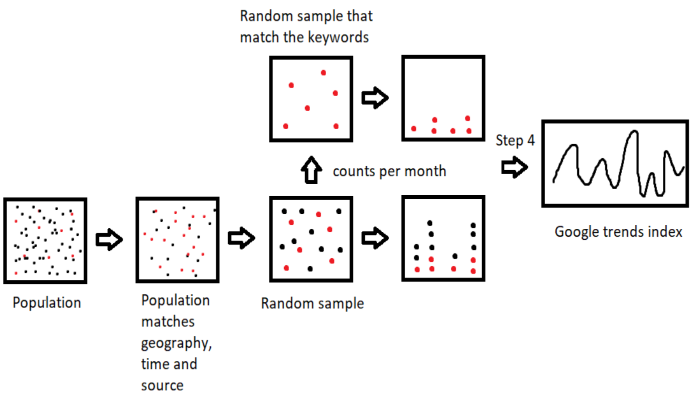
Google data

Keyword monthly search frequency in Spain from Google Trends, is going to be used to proxy the population interest that may affect the parenthood decision-making. The data provided from Google trends is normalized with a value equal to 100 for the month with higher number of searches.

Google Trends data is going to be used as a proxy of individual's consideration in several topics when making the parenthood decision. This index is constructed following the next steps:

1. Data that matches the area of interest (in this case Spain), the sources where the searches were done (in this case all: google, youtube and news) and the time framework.
2. A random selection of these matches is performed.
3. In this step the keywords findings are searched in the previously constructed population.

Figure 1



4. Samples generated by 2 and 3 are counted by month. Afterwards the time series is normalized. To do so, firstly divide elementwise 3 by 2 . Secondly, subtract the global minimum. Thirdly, divide by the global maximum. Fourthly, the result is multiplied by 100. Finally, the outcome is rounded to an integer.

The list of keywords and expressions have been selected by brainstorming and by using patterns similar to the ones employed in literature (Ballarin, D'Amiuri and Marcucci, 2013; Wilde, Chen and Lohmann, 2020). The variables are going to be divided as follows.

1. Medical variables (and keywords): contraceptives ("anticonceptivos"), ovulation test ("test de ovulación"), in vitro fecundation ("fecundación in vitro"), morning after pill ("píldora del día después"), pregnancy test ("test de embarazo"), breastfeeding ("lactancia"), pregnancy interruption ("aborto") and preparation for delivery ("preparación parto").

2. Social movements variables (and keywords): feminism ("feminism") and far right ("extrema derecha")

In order to employ these variables in a regression analysis, there are some changes that need to be done to the timeseries. By the Augmented Dickey-Fuller test (Novales, 1993) we computed the unit root test. The variables "contraceptives", "morning after pill", "pregnancy test" and "breastfeeding", rejected the stationarity test, so first-differences will be taken. Furthermore, logarithms are going to be taken for all the variables to reduce the heteroskedasticity In order to do these transformations null values were substituted by 0.00001.

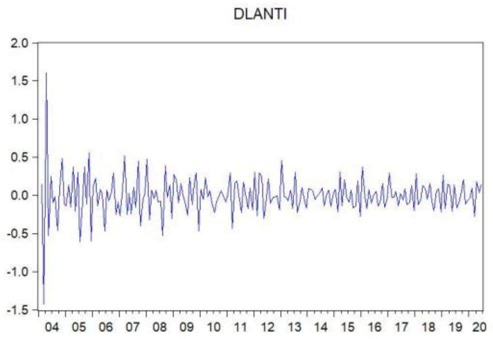
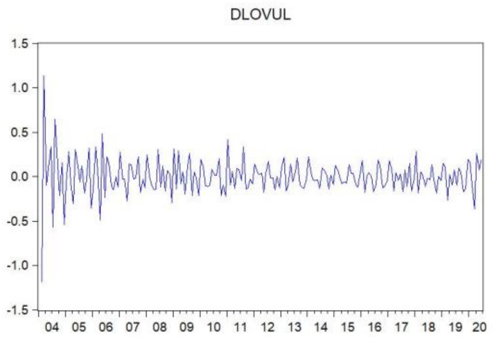
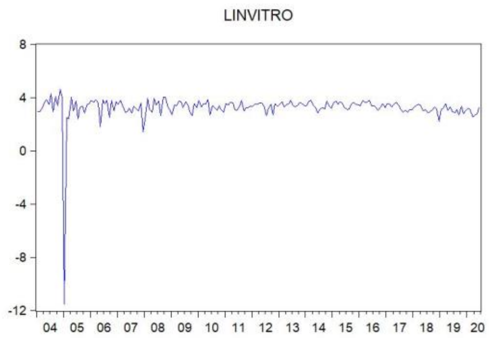
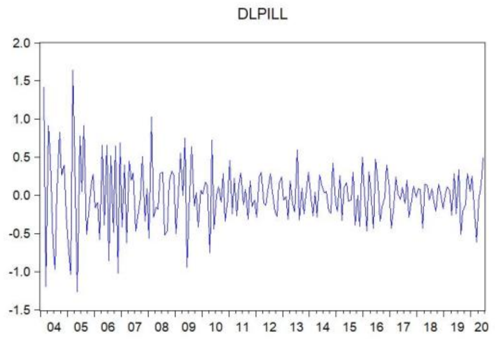
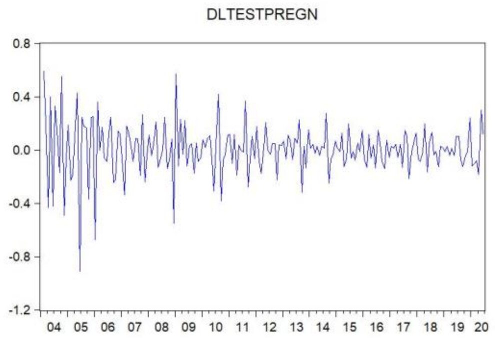
Figure 2 	The contraceptives time series (DLANTI) presents a constant trend due to its stationarity, and a higher volatility during the oldest stages of the time series, due to the lower quality data during this time.
Figure 3 	The ovulation test data (DLOVUL) presents a constant trend due to its stationarity, furthermore, as in the previously mentioned time series, the quality of data provokes a higher volatility during the first stages of the series.
Figure 4 	In vitro fecundation time series (LINVITRO) presents a consistent pattern, as for its stationarity. It is important to notice the outliers due to the lack of data during the month of issue.

Figure 5



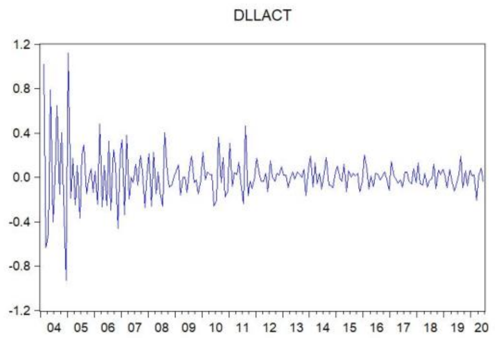
Morning after pill series (DLPILL), as the previous ones present a constant trend and a higher volatility in the oldest data due to its quality.

Figure 6



Pregnancy test time series (DLTESTPREGN) , as before, present a constant trend and a higher volatility caused by the quality of data during the first years of the time series.

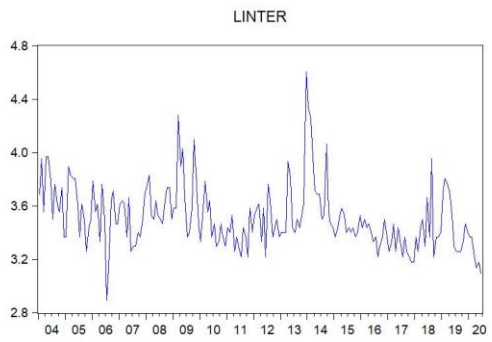
Figure 7



Breastfeeding series (DLLACT), present a constant pattern and a noticeable higher volatility caused by the lack of quality data during the first years of the timeseries.

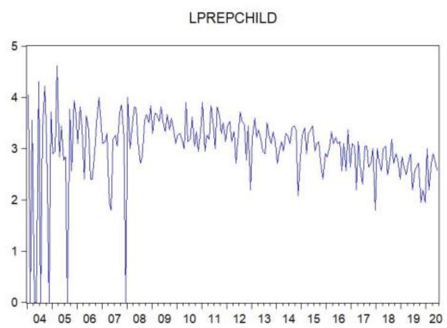
<table>
<tr>
<td>

Figure 8



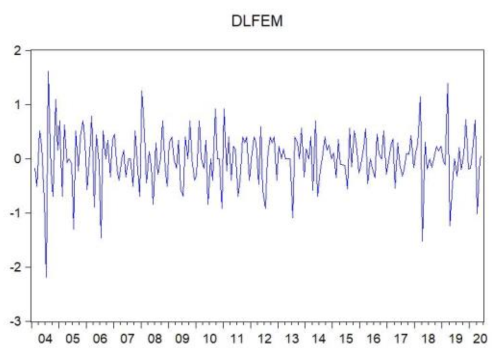
</td>
<td>Pregnancy interruption (LINTER), this time series presents a slightly more noticeable sub-trend, however the whole time series presents constant trends for its stationarity. The volatility during the first years is higher, as before, due to the quality of data. The first upward group of outliers (around 2010) were provoked by the abortion law enforcement. The second upward outlier was caused by the modification of the mentioned law.</td>
</tr>
<tr>
<td>

Figure 9



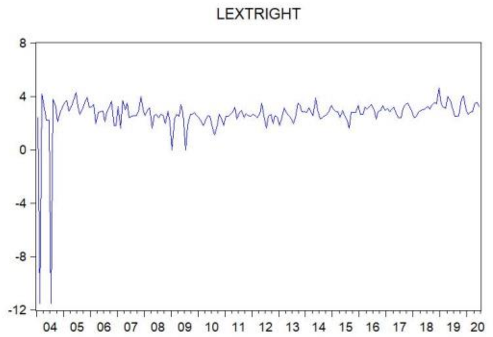
</td>
<td>Delivery preparation time series (LPREPCHILD) presents a slightly downwards trend during the last three years, but, overall the series has a constant pattern. As before, there is a bigger volatility at the beginning of the series due to the lack of quality data.</td>
</tr>
<tr>
<td>

Figure 10



</td>
<td>Feminism time series (DLFEM) presents a constant trend with a similar volatility throughout the data. A peak can be seen in late 2017 may provoked by the “me too” movement against sexual abuse. The peaks in 2019 and 2020 seem to be caused by Spanish women’s day.</td>
</tr>
</table>

Figure 11

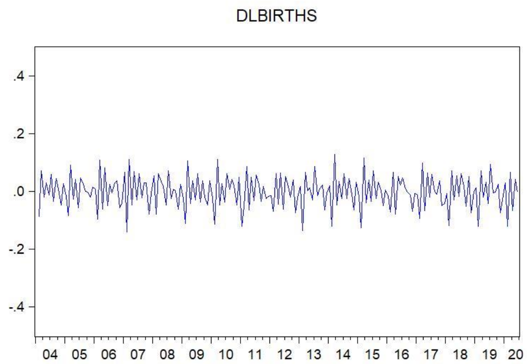


Far-right movement time series (LEXTRIGHT) presents a high volatility at the beginning, due to the lack of quality data. As for the trend it is constant due to the times series stationarity.

Fertility rate

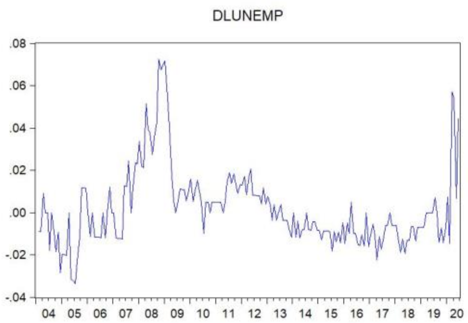
Measures the number monthly of new-borns in Spain. This time series has also been modified to make it stationary, so the first logarithm differences were computed. As it can be seen below the time series is stationary, thus showing a constant trend with no remarkable fluctuation.

Figure 12



Unemployment

Measures the number of working age individuals that take some actions to find a job, but do not have it. As the following variables, it has been modified to be stationary and to reduce its variance by taking first differences and logs to the original figures. The figure presents an overall constant trend, however, there are two groups of high value

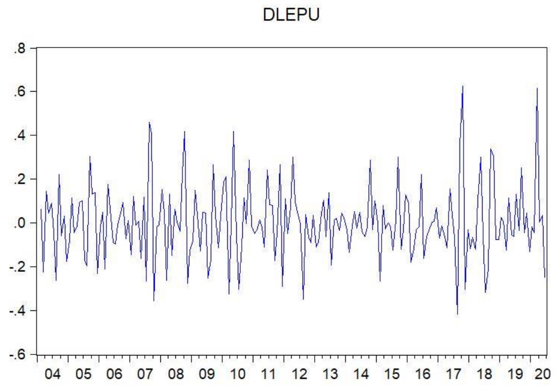


outliers, the first one during the Great Recession of 2008, and the second one during the COVID-19 pandemic.

Economic and political uncertainty

This time series created by Baker, Bloom and Davis (2015) measures the number of articles in the Spain's top journals that contain at least one word referring to uncertainty, policy, and economics. This time series has been modified by computing its logs and first differences to make it stationary. The time series' figures present a constant trend and a high value outliers' group during the 2008 great recession, as well as, around 2017 and 2018 due to the Catalan independentist issue. The last positive outlier may be due to the beginning of the current pandemic.

Figure 14

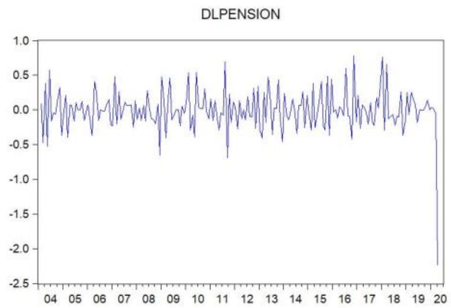
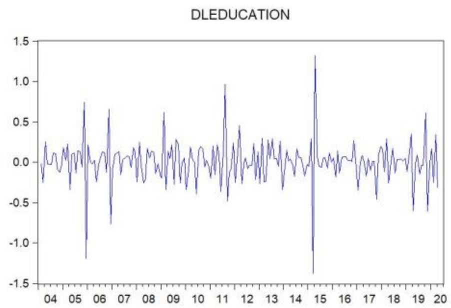


CIS survey data

Several variables have been created with data reported by the monthly survey made by the "Centro de investigaciones sociologicas" (CIS). Interpolation has been done to cover the lack of observations made for the month of August. The maximum number of participants in the survey was 17650 and the minimum 2425.

1. **Perspectives about the pensions and the education problem**

Data from a question that says as follows: "¿Cuál es, a su juicio, el principal problema que existe actualmente en España?" (What is, in your opinion, the major problem that exists nowadays in Spain?). The question gives several options and reports the percentage of participants that answered each one. In this report we are going to use the ones related to education and pensions. Both variables were stationarized by taking first differences and logarithms to reduce the volatility.

Figure 15  	Pensions as a national problem variable (DLPENSION) shows a constant trend due to its stationarity. A massive drop maybe due to the increase in thoughts given to the pandemic situation and the consequently importance drop of the pension problem.
Figure 16  	Education as a national problem variable (DLEDUCATION) shows a constant trend with a peak caused by a new education law enforcement.

2. **Opinion about the actual situation of economy and politics in Spain.**

Data was collected from two different questions that say as follows: "Refiriéndonos a la situación económica general de España actualmente, ¿cómo la calificaría Ud.: muy buena, buena, mala o muy mala?" (Referring to today's general Spanish economic situation, how would you rate it: very good, good, bad or very bad?) and "Refiriéndonos a la situación política general de España actualmente, ¿cómo la calificaría Ud.: muy buena, buena, mala o muy mala?" (Referring to today's Spanish general political situation, how would you rate it: very good, good, bad or very bad?). To simplify the indicator the options "very bad" and "bad", as well as "very good" and "good" were respectively summed up.

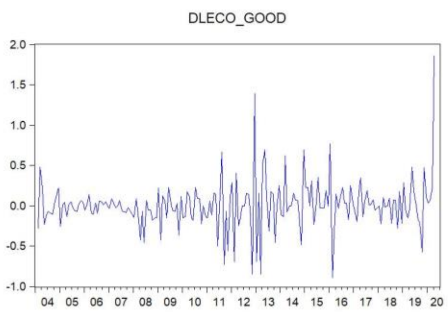
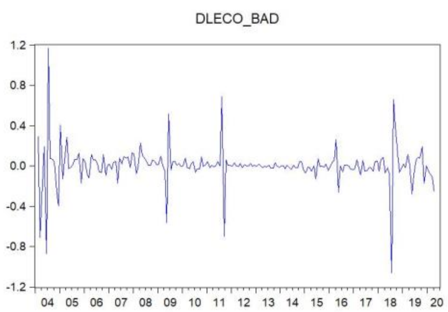
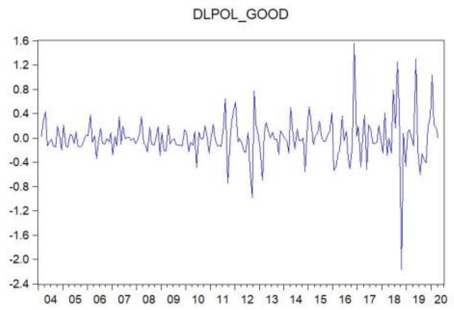
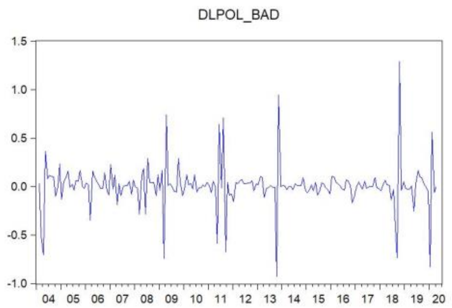
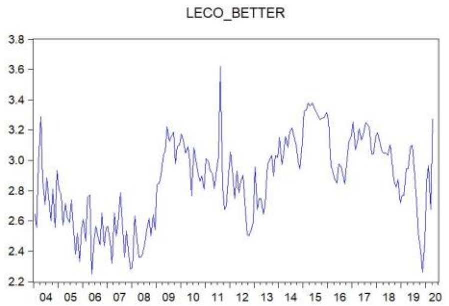
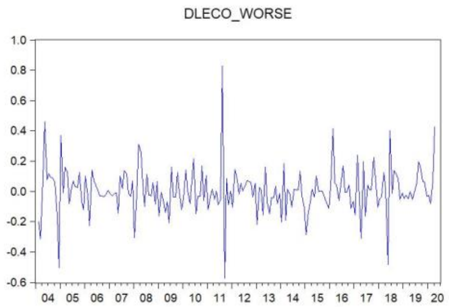
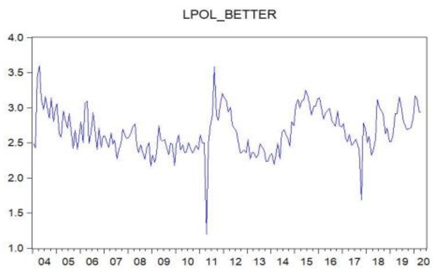
Figure 17 	Good economic situation series (DLECO_GOOD) shows a constant trend as for its stationarity, and a bigger variability around 2012-2013.
Figure 18 	Bad economic situation series (DLECO_BED) present less variability than its good version, however there are some outliers, during the economic recession of 2008, and during the Catalonian political issues.

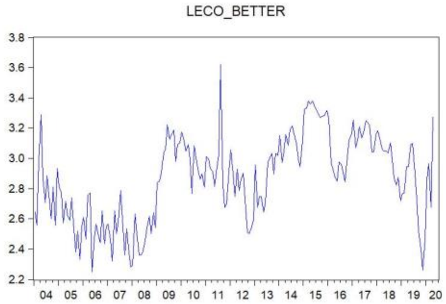
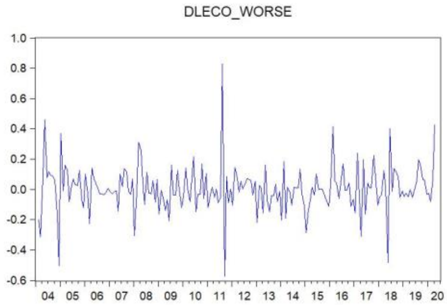
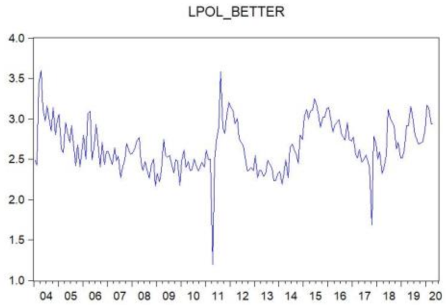
Figure 19 	Good political situation series (DLPOL_GOOD) presents a constant trend due to its stationarity. The end of the bipartidism structure of politics in Spain may have provoked an increasement in the volatility throughout the time
Figure 20 	Bad political situation (DLPOL_BAD) shows a constant trend as for its stationarity, and few outliers can be explained by the mentioned Great Recession, the early elections in 2011, the political corruption exposed in 2013 and the Catalonian independence movement.

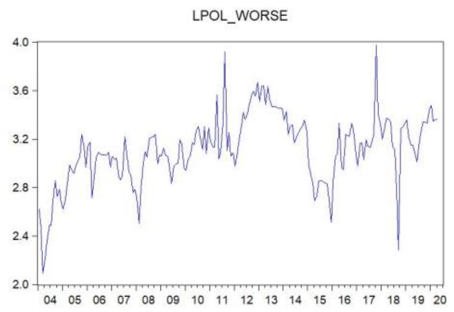
Some computations have been done to the time series to reach stationarity and a small variance. All the variables were stationarized by taking first differences and logarithms to reduce the volatility.

3. Predictions about the future economic and political situation in Spain.

Data was collected through two different questions that say as follows: “Y, ¿cree Ud. que dentro de un año la situación económica del país será mejor, igual o peor que ahora?” (Do you think that after a year, the Spanish economic situation will be better, equal or worse than now?) and “Y, ¿cree Ud. que dentro de un año la situación política del país será mejor, igual o peor que ahora?” (Do you think that after a year the Spanish political situation will be better, equal, or worse than now?). The variables that recovered a bad economic prospective were stationarized by taking first differences, the rest of variables were already stationary at levels. All the variables were transformed to logarithmic measures to reduce the volatility.

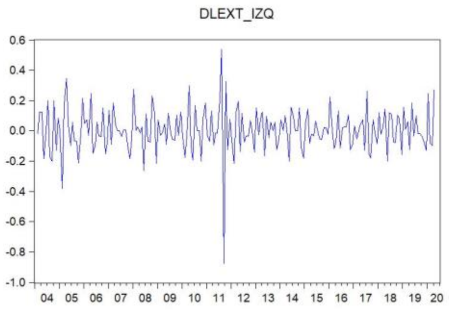
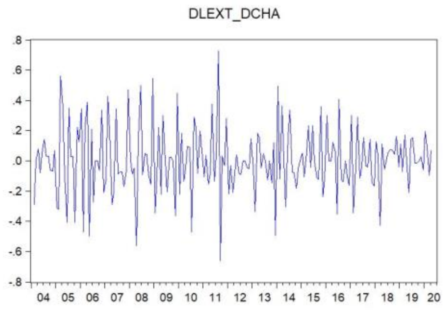
Figure 21 	Better future economic situation (LECO_BETTER) presents a constant trend divided by a structural change around 2009. As for the large fluctuations between months it can be due to the fact that this is a long memory process.
Figure 22 	Worse future economic situation (DLECO_WORSE) shows a constant trend and a high volatility peak during the 2011 elections.
Figure 23 	Better future political situation (LPOL_BETTER) presents an overall constant trend, however there are slightly upper and downwards sub-trends that always return to the constant mean.

<table>
<tr><td>Figure 24

</td><td>Better future economic situation (LECO_BETTER) presents a constant trend divided by a structural change around 2009. As for the large fluctuations between months it can be due to the fact that this is a long memory process.</td></tr>
<tr><td>Figure 25

</td><td>Worse future economic situation (DLECO_WORSE) shows a constant trend and a high volatility peak during the 2011 elections.</td></tr>
<tr><td>Figure 26

</td><td>Better future political situation (LPOL_BETTER) presents an overall constant trend, however there are slightly upper and downwards sub-trends that always return to the constant mean.</td></tr>
</table>

<table>
<tr>
<td>Figure 27
</td>
<td>Worse future political situation (LPOL_WORSE) shows an overall constant trend with downwards and upwards that always return to the mean.</td>
</tr>
</table>

4. Sympathy for the extremes: far right and far left ideologies.

Data was collected through a question that says as follows: “Cuando se habla de política se utilizan normalmente las expresiones izquierda y derecha. Situándonos en una escala de 10 casillas, como un termómetro, que van del 1 al 10, en la que 1 significa 'lo más a la izquierda' y 10 'lo más a la derecha', ¿en qué casilla se colocaría Ud.?” (When taking politics, the most common division is left and right, in a ten-box scale, like in a thermometer, from 1 to 10, where 1 represents the farthest left and 10 the farthest right, where would you place yourself?). To simplify the indicator the options “9” and “10”, as well as “1” and “2” were respectively summed up. Both variables were rendered stationary by taking first differences and logarithms.

Figure 28 	Far left movement from CIS data (DLEXT_IZQ) presents a constant trend, with a noticeable outlier in 2011, maybe for the general elections previously mentioned that pushed people to engage themself with a far left ideology.
Figure 29 	Far right movement from CIS data (DLEXT_DCHA) present a constant trend because its stationarity and a higher overall volatility than the opposite political movement.

Models

ARMA models are going to be used along this report. For estimating them a general to particular method is used, i.e., starting from the most general model and eliminating no statistically significant lags (always taking into account that the Akaike selection criteria goes down afterwards). During the model's discussion a 10% critical value is going to be used.

Ca' Foscari University of Venice

Model 1: Before pregnancy model

This model will be use as predictors for birth rates (DLBIRTHS) all variables proxied with Google Trends that are hypothesized to play a role when deciding if having children. For instance, the use of contraceptives is proxied with contraceptive information searches (DLANTI), attention towards the ovulation period is proxied by ovulation test searches (LTESTOVUL) and interest towards in vitro fecundation techniques is proxied through in vitro fecundation searches (LINVITRO). In the extensive form model lags from -14 to -9 were used, as this allows us to capture the time framework when these variables may affect individuals' parenthood decision.

As mentioned in the literature review, contraceptives usage reduces the number of new-borns (Coale and Watkins, 1986). The main objective of this model is to check if this negative effect continues in the Spanish case, approximating the usage of contraceptives by the interest on them by the Google Trends search data. A more general value "contraceptives" and a more specific one "ovulation test" were introduced in the model. This last-mentioned variable was chosen because it can be used as a contraceptive or as a way of checking the best days to become pregnant. "Invitro fecundation" was also included as it is an obvious indicator of the willingness of becoming parents. The variables used in this model recover actions that must be performed before pregnancy, as for this the lags chosen were the ones before pregnancy when the decision is made.

$$DLBIRTHS(i) = DLANTI(i-10) + DLANTI(i-12) + DLANTI(i-13) + LINVITRO(i-13) + LINVITRO(i-14) + C$$

Equation 1

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/15/21 Time: 15:55
Sample (adjusted): 2005M03 2020M06
Included observations: 184 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLANTI(-10)	-0.044667	0.019389	-2.303695	0.0224
DLANTI(-12)	-0.028778	0.018243	-1.577483	0.1165
DLANTI(-13)	-0.039406	0.018206	-2.164449	0.0318
LINVITRO(-13)	0.006288	0.003499	1.797033	0.0740
LINVITRO(-14)	-0.005124	0.003471	-1.476370	0.1416
C	-0.005678	0.016266	-0.349061	0.7275

R-squared	0.088091	Mean dependent var	-0.001314
Adjusted R-squared	0.062476	S.D. dependent var	0.056458
S.E. of regression	0.054666	Akaike info criterion	-2.943081
Sum squared resid	0.531932	Schwarz criterion	-2.838246
Log likelihood	276.7634	Hannan-Quinn criter.	-2.900590
F-statistic	3.438998	Durbin-Watson stat	3.171891
Prob(F-statistic)	0.005435		

The contraceptives search significatively diminishes the new-borns after 10 and 13 months of the search. This effect can be explained in two different ways, firstly the search of this kind of information may happen after making the decision use this type of products and by the fact that people may be encouraged to use them after looking for information in regard to this product.

As for the in vitro fecundation, we have a mixed effect. However, only the positive effect in the new-born figures is statistically significant at the 10% significance level, so, when this pregnancy method is searched the individual has already made the decision or the information read on google encourages them to take that choice.

Model 2: When may be pregnant.

This model will use as predictors for the birth rate (DLBIRTHS) all the variables from Google trends that may play a role when you may not know if you are pregnant, for instance morning after pill data searches (DLPILL) and test pregnancy searches (DLTESTPREGN). At the beginning lags from -9 to -7 were used as the time framework between conception and delivery.

Previous research has used these Google Trends' variables to predict the fertility rate, but without putting attention on its effect (Wilde, Chen and Lohmann, 2020). This model pretends to check the significance of these variables and their effect on fertility rates for the Spanish case. This model has been constructed with the variables that may play a role when individuals do not know if the conception has been successful or not. It is hypothesized that a higher interest on the morning after pill may produce a declining in the future deliveries as for its nature of emergency contraceptive. For pregnancy test searches the sign may be ambiguous because it depends on a posterior decision of interrupting or not the pregnancy if the test turns at to be positive. The lags were chosen to capture the earlier stages of the potential pregnancy when this kind of medical products were used.

$$DLBIRTHS(i) = DLTESTPREGNANCY(i-7) + DLPILL(i-8) + C$$

Equation 2

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/23/21 Time: 18:19
Sample (adjusted): 2004M10 2020M06
Included observations: 189 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLTESTPREGN(-7)	0.033416	0.021571	1.549150	0.1230
DLPILL(-8)	-0.020038	0.009784	-2.048016	0.0420
C	-0.001943	0.004047	-0.479940	0.6318

R-squared	0.029586	Mean dependent var	-0.001884
Adjusted R-squared	0.019152	S.D. dependent var	0.056182
S.E. of regression	0.055641	Akaike info criterion	-2.924049
Sum squared resid	0.575840	Schwarz criterion	-2.872593
Log likelihood	279.3227	Hannan-Quinn criter.	-2.903203
F-statistic	2.835429	Durbin-Watson stat	3.113819
Prob(F-statistic)	0.061233		

The morning after pill searches produces a negatively significant at the 5% significance level decline in the birth rate after 8 months, this may be due to the fact that individuals that search for information about this emergency contraceptive finally made the decision of use it, or they made have previously taken the decision, and because this contraceptive is taken in a potential stressful situation when the quicker source of information is Google

Model 3: When parenthood is not a decision but a fact.

As mentioned before, this kind of Google searches has already been used to predict the fertility rate (Wilde, Chen and Lohmann, 2020) but without mentioning the direction of the effect. This model tries to solve this question for the Spanish case, specifically, taking into account variables that play a role after a positive conception is confirmed. The variables used are going to be: "breastfeeding" (DLACT), it is hypothesized that this will have a strong positive effect in the number of new-born since people will look for information about this topic only after being informed of their pregnancy state. "Delivery preparation" (LPREPCHILD) information should only interest people that are expecting a child and "pregnancy interruption" (LINTER) information will interest people that may not be happy with the parenthood decision after knowing their pregnancy state. The lags (9 to 5) used in this model are the ones that recover information about the first half of the pregnancy period when decisions related with abortion can be taken.

$$DLBIRTHS(i) = DLLACT(i-7) + DLPREPCHILD(i-6) +$$

$$+ DLPRERPCHILD(i-5) \ + C$$

Equation 3

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/15/21 Time: 16:31
Sample (adjusted): 2004M09 2020M06
Included observations: 190 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLLACT(-7)	0.043479	0.018356	2.368653	0.0189
LPREPCHILD(-6)	0.007614	0.005623	1.354171	0.1773
LPREPCHILD(-5)	-0.007951	0.005536	-1.436126	0.1526
C	-0.000984	0.023505	-0.041850	0.9667

R-squared	0.055268	Mean dependent var	-0.001635
Adjusted R-squared	0.040030	S.D. dependent var	0.056138
S.E. of regression	0.055003	Akaike info criterion	-2.942038
Sum squared resid	0.562707	Schwarz criterion	-2.873680
Log likelihood	283.4936	Hannan-Quinn criter.	-2.914347
F-statistic	3.627066	Durbin-Watson stat	3.174470
Prob(F-statistic)	0.014078		

After running this model, the only variable that reported a significant variable at the 5% significance level was the breast-feeding searches, this leads to an increment in the new-born figures after 7 months. This positive influence on the birth rate appears because usually people that search for information about breastfeeding it is due to the fact that they are planning to breastfeed their children, if they are not already pregnant this way of thinking would not have sense.

Model 4: Social movements

Previous studies have proposed that far-right parties, as for their values, tend to apply policies that incentivizes new-borns increasements (Campisi et al, 2020 and Aassve and Lappegård, 2009). There are also studies that evidence the negative effect in the increment of the feminism movement popularity on the fertility rate (Gillespie, 2013; Ireland, 1993). There is no literature evidence about the possible far-left effects in the fertility figures, however, due to lack of quality data on Goggle Trend it could not be included in the regression. The purpose of this model is to check if the direction of this effect can also be found using Google Trends data from Spain. The lags taken into account (from 14 to 8) are the ones that enables the model to recover the period when the parenthood decision is taken and the variables of interest: far right searches (LEXTRIGHT) and feminism searches (DLFEM) may play a role.

$$\begin{aligned} DLBIRTHS(i) = {} & LEXTRIGHT(i-11) + DLFEM(i-8) + DLFEM(i-9) \\ & + DLFEM(i-10) + DLFEM(i-11) + DLFEM(i-12) + DLFEM(i \\ & - 13) + C \end{aligned}$$

Ca' Foscari University of Venice

Equation 4

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/15/21 Time: 18:10
Sample (adjusted): 2005M03 2020M06
Included observations: 184 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LEXTRIGHT(-11)	0.007743	0.003580	2.162994	0.0319
DLFEM(-8)	-0.023592	0.008070	-2.923573	0.0039
DLFEM(-9)	-0.015609	0.008002	-1.950478	0.0527
DLFEM(-11)	-0.030195	0.008940	-3.377486	0.0009
DLFEM(-12)	-0.022029	0.008933	-2.465952	0.0146
DLFEM(-13)	-0.037283	0.008225	-4.532763	0.0000
C	-0.021889	0.010316	-2.121867	0.0352

R-squared	0.151052	Mean dependent var	-0.001314
Adjusted R-squared	0.122274	S.D. dependent var	0.056458
S.E. of regression	0.052894	Akaike info criterion	-3.003753
Sum squared resid	0.495206	Schwarz criterion	-2.881446
Log likelihood	283.3453	Hannan-Quinn criter.	-2.954180
F-statistic	5.248881	Durbin-Watson stat	3.101997
Prob(F-statistic)	0.000053		

In one hand, the far-right movement searches in Google shows a significant positive at 5% significance level effect on the fertility rate after 11 months of the search as expected. On the other hand, the feminism searches show a significant negative effect, as expected, after 8, 9, 11, 12 and 13 months in the fertility rate.

Model 5: Other extremes political movements model approach

As mentioned before, no good-quality variables were available to approach the far-left social movement from google data searches. For this reason, CIS survey variables were used to approach the political movements, in order to figure out the possible effect of far-left movements The same lags as before are going to be taken into account for each separate regression.

First, since the far-right movement has already been approached by the Google trends data, we are going to use it as a control variable.

$$DLBIRTHS(i) = DLEXT\_DCHA(i-9) \ + \ C$$

Equation 5

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/16/21 Time: 10:49
Sample (adjusted): 2004M11 2020M06
Included observations: 188 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLEXT DCHA(-9)	0.034985	0.018220	1.920154	0.0564
C	-0.002076	0.004080	-0.508760	0.6115

R-squared	0.019437	Mean dependent var	-0.001935
Adjusted R-squared	0.014165	S.D. dependent var	0.056327
S.E. of regression	0.055927	Akaike info criterion	-2.918967
Sum squared resid	0.581771	Schwarz criterion	-2.884537
Log likelihood	276.3829	Hannan-Quinn criter.	-2.905017
F-statistic	3.686990	Durbin-Watson stat	3.138324
Prob(F-statistic)	0.056368		

The variable created to proxy the collective attention to the far-right movement by the CIS Survey (DLEXT-DCHA) has reported also a positive and significative effect on the birth rate after six months. However, this new measure is coetaneous to the conception month, while the far-right movement by google trends variable that produces an effect on the new-borns figure is obtained 2 months after the conception.

Secondly, the following model was constructed to check the hypothesis of a negative effect that may be produced by the far-left movements. In the following model a negative and significant effect of the far-left variable measured by the CIS survey on the birth rate after 10 Months, as it was hypothesized before.

$$DLBIRTHS(i) = DLEXT\_RIGHT(i-10) + C$$

Equation 6

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/16/21 Time: 10:52
Sample (adjusted): 2004M12 2020M06
Included observations: 187 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLEXT_IZQ(-10)	-0.083768	0.028301	-2.959906	0.0035
C	-0.001350	0.004041	-0.333952	0.7388

R-squared	0.045216	Mean dependent var	-0.001694
Adjusted R-squared	0.040055	S.D. dependent var	0.056381
S.E. of regression	0.055240	Akaike info criterion	-2.943625
Sum squared resid	0.564517	Schwarz criterion	-2.909068
Log likelihood	277.2290	Hannan-Quinn criter.	-2.929623
F-statistic	8.761043	Durbin-Watson stat	3.173076
Prob(F-statistic)	0.003480		

Model 6: Economical variables

Having children implies a great monetary inversion, so a situation without monetary stability and great unemployment it is not ideal to rising a kid. Many studies have pointed out the negative influence of unemployment in the parenthood decision (Sánchez-Barricarte, 2019; Busetta, Mendola and Vignoli, 2019; Gerson, 1993). Also, Wilde, Chen and Lohmann, 2020 used Google Trends data related to unemployment topics in order to predict the fertility rate.

At this point it is obvious to think that the uncertainty about the future may provoke a drop in the individuals that decide to be parents, as a possibility of worse personal finances may affect their capacity of taking care of the children. Several studies have explained this effect (Comolli et al, 2021; Goldstein et al, 2013; Gerson, 1993; Jansen and Liefbroer, 2006; Campisi et al, 2020).

These two variables may be taken into consideration when making the parenthood decision, so the lags considered (from 9 to 12) will recovered data from this time period.

$$DLBIRTHS(i) = DLUNEMP(i-9) + DLUNEMP(i-10) + DLUNEMP(i-11) + DLUNEMP(-12) + C$$

Equation 7

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/16/21 Time: 09:58
Sample (adjusted): 2005M02 2020M06
Included observations: 185 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLUNEMP(-9)	0.735936	0.459991	1.599893	0.1114
DLUNEMP(-10)	-1.726506	0.559960	-3.083269	0.0024
DLUNEMP(-11)	2.089078	0.557366	3.748125	0.0002
DLUNEMP(-12)	-1.227030	0.459696	-2.669220	0.0083
C	-0.001558	0.003984	-0.391091	0.6962

R-squared	0.109678	Mean dependent var	-0.001762
Adjusted R-squared	0.089893	S.D. dependent var	0.056634
S.E. of regression	0.054029	Akaike info criterion	-2.971945
Sum squared resid	0.525439	Schwarz criterion	-2.884909
Log likelihood	279.9049	Hannan-Quinn criter.	-2.936671
F-statistic	5.543516	Durbin-Watson stat	3.136319
Prob(F-statistic)	0.000313		

Model 7: Another economic and uncertainty model approach

In the previous model no significant results for the uncertainty variable were obtained. So, other variables are going to be used to try to figure out the sign of the uncertainty effect in the fertility rate. The same lags as before are going to be used (from 9 to 12) to capture the moment when the decision is taken, and if those variables play a roll.

The model 8 it shows that the influence of a better economical situation provokes an ambiguous effect, lags 9 and 11 shows a positive sign and lags 10 a negative sign, further study will be needed to determine the nature of this result.

$$DLBIRTHS(i) = LECO\_BETTER(i-9) + LECO\_BETTER(i-10) + LECO\_BETTER(i-11) + C$$

Equation 8

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/16/21 Time: 10:07
Sample (adjusted): 2004M12 2020M06
Included observations: 187 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LECO_BETTER(-9)	0.057550	0.026009	2.212705	0.0282
LECO_BETTER(-10)	-0.118347	0.031375	-3.771966	0.0002
LECO_BETTER(-11)	0.056110	0.026093	2.150366	0.0328
C	0.011813	0.044653	0.264547	0.7917

R-squared	0.072170	Mean dependent var	-0.001694
Adjusted R-squared	0.056960	S.D. dependent var	0.056381
S.E. of regression	0.054751	Akaike info criterion	-2.950872
Sum squared resid	0.548581	Schwarz criterion	-2.881757
Log likelihood	279.9065	Hannan-Quinn criter.	-2.922866
F-statistic	4.744795	Durbin-Watson stat	3.164624
Prob(F-statistic)	0.003274		

The model 9 relates a proxy of the currently economic and political situation with the future fertility rate. This proxies are going to be studied then the parenthood decision is been discussed by the individuals, for this propose the lags from 9 to 12 are being used to create the general model that later will be reduced with the mentioned procedure. As seen in literature the expected outcome is that a good economic situation affects positively in the number of new-borns (Sánchez-Barricarte, 2019), as for the political situation no hypothesis has been proposed previously, but it is expected the same effect.

$$DLBIRTHS(i) = DLECO\_GOOD(i-12) + DLPOL\_GOOD(i-12) + C$$

Equation 9

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/16/21 Time: 10:12
Sample (adjusted): 2005M02 2020M06
Included observations: 185 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLECO_GOOD(-12)	0.034297	0.015606	2.197647	0.0292
DLPOL_GOOD(-12)	-0.010840	0.012361	-0.876890	0.3817
C	-0.001597	0.004132	-0.386570	0.6995

R-squared	0.026524	Mean dependent var	-0.001762
Adjusted R-squared	0.015826	S.D. dependent var	0.056634
S.E. of regression	0.056184	Akaike info criterion	-2.904277
Sum squared resid	0.574514	Schwarz criterion	-2.852055
Log likelihood	271.6456	Hannan-Quinn criter.	-2.883112
F-statistic	2.479425	Durbin-Watson stat	3.152885
Prob(F-statistic)	0.086618		

The proxy for individual's perception of a currently good economic and politic situation shows a positive and significative effect (at a 5% significance level) on a year after fertility rate, as expected. However, a good political situation does not present a significant effect on the fertility rate.

Not all the variables studied reported significative results. On one hand, the Google Trends' proxies to the attention towards "pregnancy interruption", the economic and polity uncertainty and the data from the CIS survey regarding a bad economic and political situation, the variables that capture a worse economic and political situation and the variable that measures the population worry about the current education issues (model 10 in the appendix), were supposed to show a negative effect on the fertility rate. On the other hand, variables such as the Google Trends' proxies for the attention towards "delivery preparation" and the data from the CIS survey regarding a current good political situation, a future better political situation, and the population worries about the pension problem (model 10 in the appendix), were expected to report a positive lagged effect on the new-born's figures.

Conclusion

This research scope is to study the different variables obtained by Google Trends that approach individuals' interest on different topics and keywords in relation with the fertility rate. Several studies have used Google Trends as predictors for the fertility rate, but none of them approaches the Spanish case (Wide, Chen and Lohnman, 2020; Billari, D'Amuri and Marcucci, 2013). For this reason, during this research the Spanish case was studied to find variables that influence the fertility rate in order to make a more detailed prediction in a future research. Furthermore, some variables related with social movements approached by Google Trends data and by the monthly CIS survey that have not been studied before in this framwork.

After modelling several ARMA regressions some interesting results came out. Some of the variables obtained by Google Trends that reported significant effects: contraceptives and feminism searches provoked a decrease in the fertility rate, whereas attention towards in-vitro fecundation, breast-feeding and far right political movements affect the variable of interest positively. The variables from the CIS survey related with the present and future situation of the policy did not affect the fertility rate, just the positive perceptions of the economic present and future situation reported a positive affect and an ambiguous effect respectively, also the far left and far right political movements popularity affect in a negative and a positive way the fertility rate, respectively.

The main limitation of this research is the usage of aggregated variables to analyse people's intentions, because, as in the case of the Google Trends data it cannot be controlled if an individual has searched more than one term or the same term throughout different days. This data issue may cause correlations between the independent variables, provoking a lack of model accuracy. Also, the data quality was not optimal as during the first years of the Google Trends many data figures were empty, and interpolation was needed. Furthermore, some variables of interest for this project did not fulfil the minimum quality required.

The are several topics that can be studied in future research projects- Firstly, a study about the abortion decision making process using web-based indicators as Google Trends. Secondly, in order to solve the problems produced by the usage of aggregated data previously mentioned, a longitudinal study that recovers couples Google searching behaviour through the time, as well as their number of children and their birth date. Secondly, this study leaves unexplained the ambiguity of some of the variables studied as the unemployment and the perspective of a better economic situation, so further research will be needed. Lastly, this study opens several lines of research related with the abortion decision process. Also, a cross-country and cross-regional comparison of the Google searching patterns of usage at its influence on the parenthood decision making process.

Ca' Foscari University of Venice

Appendix

Equation 1 general form

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/27/21 Time: 13:25
Sample (adjusted): 2005M04 2020M06
Included observations: 183 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLANTI(-9)	-0.038102	0.025531	-1.492361	0.1375
DLANTI(-10)	-0.068332	0.029852	-2.288996	0.0234
DLANTI(-11)	-0.031612	0.031545	-1.002119	0.3178
DLANTI(-12)	-0.035689	0.030587	-1.166784	0.2450
DLANTI(-13)	-0.037423	0.028884	-1.295630	0.1969
DLANTI(-14)	0.011112	0.022424	0.495545	0.6209
LTESTOVUL(-9)	0.006398	0.007025	0.910658	0.3638
LTESTOVUL(-10)	-0.008201	0.007448	-1.101075	0.2725
LTESTOVUL(-11)	-0.002991	0.007474	-0.400201	0.6895
LTESTOVUL(-12)	0.004705	0.006925	0.679514	0.4978
LTESTOVUL(-13)	-0.000504	0.006816	-0.073954	0.9411
LTESTOVUL(-14)	-0.000957	0.006186	-0.154692	0.8773
LINVITRO(-9)	-2.49E-05	0.004221	-0.005893	0.9953
LINVITRO(-10)	0.000977	0.004292	0.227540	0.8203
LINVITRO(-11)	-0.002887	0.004209	-0.685804	0.4938
LINVITRO(-12)	-0.002789	0.004201	-0.664008	0.5076
LINVITRO(-13)	0.009543	0.004370	2.183595	0.0304
LINVITRO(-14)	-0.007041	0.004342	-1.621693	0.1068
C	0.010388	0.032072	0.323894	0.7464

R-squared	0.108490	Mean dependent var	-0.001817
Adjusted R-squared	0.010641	S.D. dependent var	0.056197
S.E. of regression	0.055898	Akaike info criterion	-2.832561
Sum squared resid	0.512424	Schwarz criterion	-2.499336
Log likelihood	278.1794	Hannan-Quinn criter.	-2.697489
F-statistic	1.108754	Durbin-Watson stat	3.216293
Prob(F-statistic)	0.347960		

Equation 2 general model

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/27/21 Time: 13:33
Sample (adjusted): 2004M11 2020M06
Included observations: 188 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLPILL(-7)	0.017686	0.011496	1.538516	0.1257
DLPILL(-8)	-0.012808	0.012409	-1.032096	0.3034
DLPILL(-9)	0.004859	0.011146	0.435964	0.6634
DLTESTPREGN(-7)	0.036858	0.026505	1.390599	0.1661
DLTESTPREGN(-8)	0.010656	0.029285	0.363882	0.7164
DLTESTPREGN(-9)	-0.017449	0.025368	-0.687841	0.4924
C	-0.001987	0.004076	-0.487476	0.6265

R-squared	0.049239	Mean dependent var	-0.001935
Adjusted R-squared	0.017722	S.D. dependent var	0.056327
S.E. of regression	0.055826	Akaike info criterion	-2.896640
Sum squared resid	0.564089	Schwarz criterion	-2.776134
Log likelihood	279.2841	Hannan-Quinn criter.	-2.847815
F-statistic	1.562302	Durbin-Watson stat	3.074783
Prob(F-statistic)	0.160558		

Equation 3 general model

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/27/21 Time: 13:37
Sample (adjusted): 2004M11 2020M06
Included observations: 188 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLPILL(-5)	0.008751	0.013912	0.628995	0.5302
DLPILL(-6)	-0.026231	0.016555	-1.584425	0.1149
DLPILL(-7)	0.007198	0.016525	0.435574	0.6637
DLPILL(-8)	-0.015216	0.016760	-0.907867	0.3652
DLPILL(-9)	0.003910	0.013113	0.298198	0.7659
DLLACT(-5)	0.002863	0.028086	0.101932	0.9189
DLLACT(-6)	0.019344	0.033277	0.581286	0.5618
DLLACT(-7)	0.085014	0.035718	2.380119	0.0184
DLLACT(-8)	0.030907	0.031151	0.992161	0.3225
DLLACT(-9)	0.020252	0.025030	0.809095	0.4195
C	-0.003082	0.004038	-0.763287	0.4463

R-squared	0.116917	Mean dependent var	-0.001935
Adjusted R-squared	0.067026	S.D. dependent var	0.056327
S.E. of regression	0.054407	Akaike info criterion	-2.927930
Sum squared resid	0.523935	Schwarz criterion	-2.738564
Log likelihood	286.2255	Hannan-Quinn criter.	-2.851206
F-statistic	2.343424	Durbin-Watson stat	3.150685
Prob(F-statistic)	0.012763		

Equation 4 general model

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/27/21 Time: 14:04
Sample (adjusted): 2005M04 2020M06
Included observations: 183 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LEXTRIGHT(-8)	-0.009794	0.006633	-1.476693	0.1416
LEXTRIGHT(-9)	0.006905	0.004495	1.536153	0.1264
LEXTRIGHT(-10)	-0.003638	0.003615	-1.006345	0.3157
LEXTRIGHT(-11)	0.009558	0.003671	2.603627	0.0101
LEXTRIGHT(-12)	-0.002249	0.003612	-0.622610	0.5344
LEXTRIGHT(-13)	-0.000146	0.003581	-0.040835	0.9675
LEXTRIGHT(-14)	-0.007874	0.003280	-2.400837	0.0174
DLFEM(-8)	-0.020697	0.009363	-2.210458	0.0284
DLFEM(-9)	-0.011212	0.010104	-1.109685	0.2687
DLFEM(-10)	0.004073	0.010261	0.396954	0.6919
DLFEM(-11)	-0.027822	0.010085	-2.758791	0.0064
DLFEM(-12)	-0.016425	0.010241	-1.603921	0.1106
DLFEM(-13)	-0.034040	0.010038	-3.391268	0.0009
DLFEM(-14)	0.018889	0.008861	2.131722	0.0345
C	0.018106	0.024092	0.751531	0.4534

R-squared	0.196628	Mean dependent var	-0.001817
Adjusted R-squared	0.129680	S.D. dependent var	0.056197
S.E. of regression	0.052427	Akaike info criterion	-2.980376
Sum squared resid	0.461764	Schwarz criterion	-2.717303
Log likelihood	287.7044	Hannan-Quinn criter.	-2.873739
F-statistic	2.937035	Durbin-Watson stat	3.028834
Prob(F-statistic)	0.000509		

Equation 5 general model

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/27/21 Time: 19:37
Sample (adjusted): 2005M04 2020M06
Included observations: 183 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLEXT_DCHA(-8)	-0.034681	0.018634	-1.861117	0.0644
DLEXT_DCHA(-10)	-0.019160	0.021156	-0.905692	0.3663
DLEXT_DCHA(-11)	-0.018064	0.023534	-0.767566	0.4438
DLEXT_DCHA(-12)	0.024671	0.025102	0.982806	0.3271
DLEXT_DCHA(-13)	-0.015544	0.023590	-0.658906	0.5108
DLEXT_DCHA(-14)	0.030648	0.020961	1.462099	0.1455
C	-0.001628	0.004116	-0.395450	0.6930

R-squared	0.059725	Mean dependent var	-0.001817
Adjusted R-squared	0.027670	S.D. dependent var	0.056197
S.E. of regression	0.055414	Akaike info criterion	-2.910453
Sum squared resid	0.540453	Schwarz criterion	-2.787686
Log likelihood	273.3065	Hannan-Quinn criter.	-2.860690
F-statistic	1.863220	Durbin-Watson stat	3.130014
Prob(F-statistic)	0.089599		

Equation 6 general model

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/27/21 Time: 19:45
Sample (adjusted): 2005M04 2020M06
Included observations: 183 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLEXT_IZQ(-8)	-0.026710	0.033643	-0.793915	0.4283
DLEXT_IZQ(-9)	0.007951	0.038392	0.207089	0.8362
DLEXT_IZQ(-10)	-0.074953	0.041599	-1.801811	0.0733
DLEXT_IZQ(-11)	-0.015139	0.042049	-0.360036	0.7193
DLEXT_IZQ(-12)	0.035229	0.041710	0.844613	0.3995
DLEXT_IZQ(-13)	-0.027485	0.038206	-0.719377	0.4729
DLEXT_IZQ(-14)	0.048437	0.033510	1.445437	0.1501
C	-0.001581	0.004103	-0.385368	0.7004

R-squared	0.084217	Mean dependent var	-0.001817
Adjusted R-squared	0.047586	S.D. dependent var	0.056197
S.E. of regression	0.054844	Akaike info criterion	-2.925917
Sum squared resid	0.526376	Schwarz criterion	-2.785612
Log likelihood	275.7214	Hannan-Quinn criter.	-2.869044
F-statistic	2.299038	Durbin-Watson stat	3.173232
Prob(F-statistic)	0.028896		

Equation 7 general model

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/12/21 Time: 15:40
Sample (adjusted): 2005M02 2020M06
Included observations: 185 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLEPU(-9)	0.027488	0.027246	1.008874	0.3144
DLEPU(-10)	0.005136	0.027576	0.186256	0.8525
DLEPU(-11)	0.037005	0.027092	1.365887	0.1737
DLEPU(-12)	-0.024344	0.027075	-0.899122	0.3698
DLUNEMP(-9)	0.738233	0.464538	1.589176	0.1138
DLUNEMP(-10)	-1.716931	0.570689	-3.008522	0.0030
DLUNEMP(-11)	1.940597	0.565532	3.431454	0.0007
DLUNEMP(-12)	-1.099342	0.471290	-2.332625	0.0208
C	-0.001691	0.003988	-0.423976	0.6721

R-squared	0.129298	Mean dependent var	-0.001762
Adjusted R-squared	0.089720	S.D. dependent var	0.056634
S.E. of regression	0.054034	Akaike info criterion	-2.950985
Sum squared resid	0.513860	Schwarz criterion	-2.794319
Log likelihood	281.9661	Hannan-Quinn criter.	-2.887492
F-statistic	3.266961	Durbin-Watson stat	3.133599
Prob(F-statistic)	0.001696		

Equation 8 general model

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/27/21 Time: 20:16
Sample (adjusted): 2005M01 2020M06
Included observations: 186 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LECO_BETTER(-9)	0.060398	0.032382	1.865166	0.0638
LECO_BETTER(-10)	-0.109744	0.037712	-2.910059	0.0041
LECO_BETTER(-11)	0.027844	0.037603	0.740492	0.4600
LECO_BETTER(-12)	0.018878	0.032303	0.584403	0.5597
LPOL_BETTER(-9)	-0.003858	0.021149	-0.182443	0.8554
LPOL_BETTER(-10)	-0.013053	0.023116	-0.564689	0.5730
LPOL_BETTER(-11)	0.020139	0.023040	0.874112	0.3832
LPOL_BETTER(-12)	-0.001705	0.020976	-0.081277	0.9353
C	0.001893	0.057558	0.032892	0.9738

R-squared	0.077274	Mean dependent var	-0.001848
Adjusted R-squared	0.035568	S.D. dependent var	0.056493
S.E. of regression	0.055479	Akaike info criterion	-2.898436
Sum squared resid	0.544798	Schwarz criterion	-2.742352
Log likelihood	278.5546	Hannan-Quinn criter.	-2.835185
F-statistic	1.852853	Durbin-Watson stat	3.164442
Prob(F-statistic)	0.070233		

Equation 9 general model

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/27/21 Time: 20:20
Sample (adjusted): 2005M02 2020M06
Included observations: 185 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLECO_GOOD(-9)	0.000665	0.017872	0.037225	0.9703
DLECO_GOOD(-10)	-0.001465	0.018873	-0.077621	0.9382
DLECO_GOOD(-11)	-0.009132	0.018836	-0.484818	0.6284
DLECO_GOOD(-12)	0.029516	0.017398	1.696567	0.0915
DLPOL_GOOD(-9)	-0.004403	0.012910	-0.341079	0.7335
DLPOL_GOOD(-10)	-0.011019	0.012952	-0.850790	0.3960
DLPOL_GOOD(-11)	0.010673	0.013007	0.820522	0.4130
DLPOL_GOOD(-12)	-0.011436	0.013297	-0.860060	0.3909
C	-0.001778	0.004201	-0.423368	0.6725

R-squared	0.036525	Mean dependent var	-0.001762
Adjusted R-squared	-0.007269	S.D. dependent var	0.056634
S.E. of regression	0.056840	Akaike info criterion	-2.849739
Sum squared resid	0.568612	Schwarz criterion	-2.693073
Log likelihood	272.6008	Hannan-Quinn criter.	-2.786246
F-statistic	0.834017	Durbin-Watson stat	3.156991
Prob(F-statistic)	0.573771		

Equation 10 general model

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/12/21 Time: 16:28
Sample (adjusted): 2005M02 2020M06
Included observations: 185 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLECO_BAD(-9)	0.018923	0.025563	0.740250	0.4601
DLECO_BAD(-10)	-0.012951	0.027850	-0.465035	0.6425
DLECO_BAD(-11)	0.014253	0.027931	0.510289	0.6105
DLECO_BAD(-12)	-0.009495	0.025022	-0.379476	0.7048
DLPOL_BAD(-9)	0.006336	0.024141	0.262445	0.7933
DLPOL_BAD(-10)	-0.001202	0.026048	-0.046145	0.9632
DLPOL_BAD(-11)	0.011029	0.024624	0.447896	0.6548
DLPOL_BAD(-12)	-0.005226	0.022552	-0.231727	0.8170
C	-0.001915	0.004235	-0.452058	0.6518

R-squared	0.023143	Mean dependent var	-0.001762
Adjusted R-squared	-0.021260	S.D. dependent var	0.056634
S.E. of regression	0.057233	Akaike info criterion	-2.835945
Sum squared resid	0.576510	Schwarz criterion	-2.679279
Log likelihood	271.3249	Hannan-Quinn criter.	-2.772452
F-statistic	0.521203	Durbin-Watson stat	3.140124
Prob(F-statistic)	0.839514		

Equation 10 *reduce* model

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/16/21 Time: 10:22
Sample (adjusted): 2005M01 2020M06
Included observations: 186 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLECO_BAD(-9)	0.030073	0.021804	1.379212	0.1695
DLECO_BAD(-11)	0.025678	0.020903	1.228418	0.2209
C	-0.002143	0.004131	-0.518698	0.6046

R-squared	0.017587	Mean dependent var	-0.001848
Adjusted R-squared	0.006850	S.D. dependent var	0.056493
S.E. of regression	0.056299	Akaike info criterion	-2.900274
Sum squared resid	0.580038	Schwarz criterion	-2.848246
Log likelihood	272.7255	Hannan-Quinn criter.	-2.879190
F-statistic	1.638042	Durbin-Watson stat	3.137401
Prob(F-statistic)	0.197197		

Equation 11 general model

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/12/21 Time: 16:21
Sample (adjusted): 2005M02 2020M06
Included observations: 185 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLECO_WORSE(-9)	0.031800	0.035030	0.907785	0.3652
DLECO_WORSE(-10...	0.043234	0.036835	1.173723	0.2421
DLECO_WORSE(-11...	0.022895	0.035381	0.647112	0.5184
DLECO_WORSE(-12...	0.017036	0.028711	0.593350	0.5537
LPOL_WORSE(-9)	0.005740	0.028717	0.199878	0.8418
LPOL_WORSE(-10)	-0.037788	0.032795	-1.152242	0.2508
LPOL_WORSE(-11)	0.038196	0.032852	1.162683	0.2465
LPOL_WORSE(-12)	0.003361	0.029252	0.114916	0.9086
C	-0.031553	0.055584	-0.567662	0.5710

R-squared	0.020711	Mean dependent var	-0.001762
Adjusted R-squared	-0.023802	S.D. dependent var	0.056634
S.E. of regression	0.057304	Akaike info criterion	-2.833458
Sum squared resid	0.577945	Schwarz criterion	-2.676792
Log likelihood	271.0949	Hannan-Quinn criter.	-2.769965
F-statistic	0.465279	Durbin-Watson stat	3.162447
Prob(F-statistic)	0.879299		

Equation 11 reduce model

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/16/21 Time: 10:40
Sample (adjusted): 2004M12 2020M06
Included observations: 187 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LPOL_WORSE(-11)	0.007043	0.014316	0.491956	0.6233
C	-0.023504	0.044526	-0.527871	0.5982

R-squared	0.001307	Mean dependent var	-0.001694
Adjusted R-squared	-0.004092	S.D. dependent var	0.056381
S.E. of regression	0.056496	Akaike info criterion	-2.898663
Sum squared resid	0.590479	Schwarz criterion	-2.864105
Log likelihood	273.0250	Hannan-Quinn criter.	-2.884660
F-statistic	0.242020	Durbin-Watson stat	3.149491
Prob(F-statistic)	0.623334		

Equation 12 general model

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/12/21 Time: 17:16
Sample (adjusted): 2005M02 2020M06
Included observations: 185 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLEDUCATION(-9)	-0.008613	0.018878	-0.456232	0.6488
DLEDUCATION(-10)	0.015299	0.022034	0.694356	0.4884
DLEDUCATION(-11)	0.015073	0.021973	0.685997	0.4936
DLEDUCATION(-12)	-0.013841	0.018786	-0.736786	0.4622
DLPENSION(-9)	0.019436	0.018085	1.074678	0.2840
DLPENSION(-10)	-0.003110	0.019942	-0.155945	0.8763
DLPENSION(-11)	0.002380	0.019909	0.119556	0.9050
DLPENSION(-12)	0.002177	0.017997	0.120945	0.9039
C	-0.001871	0.004208	-0.444536	0.6572

R-squared	0.027048	Mean dependent var	-0.001762
Adjusted R-squared	-0.017177	S.D. dependent var	0.056634
S.E. of regression	0.057119	Akaike info criterion	-2.839951
Sum squared resid	0.574205	Schwarz criterion	-2.683285
Log likelihood	271.6954	Hannan-Quinn criter.	-2.776458
F-statistic	0.611607	Durbin-Watson stat	3.135018
Prob(F-statistic)	0.767387		

Equation 12 reduce model

Dependent Variable: DLBIRTHS
Method: Least Squares
Date: 06/16/21 Time: 11:12
Sample (adjusted): 2005M01 2020M06
Included observations: 186 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLEDUCATION(-11)	0.015535	0.015517	1.001152	0.3181
C	-0.001881	0.004142	-0.454158	0.6503

R-squared	0.005418	Mean dependent var	-0.001848
Adjusted R-squared	0.000012	S.D. dependent var	0.056493
S.E. of regression	0.056493	Akaike info criterion	-2.898715
Sum squared resid	0.587223	Schwarz criterion	-2.864030
Log likelihood	271.5805	Hannan-Quinn criter.	-2.884659
F-statistic	1.002306	Durbin-Watson stat	3.157182
Prob(F-statistic)	0.318068		

Ca' Foscari University of Venice

Bibliography

Adsera, A. (2004). Changing fertility rates in developed countries. The impact of labor market institutions. *Journal of population economics*, *17*(1), 17-43.

Adsera, A. (2011). Where are the babies? Labor market conditions and fertility in Europe. *European Journal of Population/Revue européenne de Démographie*, *27*(1), 1-32.

Agrillo, C., & Nelini, C. (2008). Childfree by choice: A review. *Journal of cultural geography*, *25*(3), 347-363.

Arpino, B., Luppi, F., & Rosina, A. (2021). Changes in fertility plans during the COVID-19 pandemic in Italy: the role of occupation and income vulnerability.

Balbo, N., & Barban, N. (2014). Does fertility behavior spread among friends?. *American Sociological Review*, *79*(3), 412-431.

Becker, G. S., & de Grado, C. P. (1987). *Tratado sobre la familia* (Vol. 1). Madrid: Alianza editorial.

Bernardi, L., Keim, S., & Von der Lippe, H. (2007). Social influences on fertility: A comparative mixed methods study in Eastern and Western Germany. *Journal of mixed methods research*, *1*(1), 23-47.

Billari, F., D'Amuri, F., & Marcucci, J. (2016, October). Forecasting births using Google. In *CARMA 2016: 1st International Conference on Advanced Research Methods in Analytics* (pp. 119-119). Editorial Universitat Politècnica de València.

Blackstone, A., & Stewart, M. D. (2016). “There’s More Thinking to Decide” How the Childfree Decide Not to Parent. *The Family Journal*, *24*(3), 296-303.

Bloom, N., Davis, S. J., & Baker, S. R. (n.d.). Economic Policy Uncertainty Index. Economic Policy Uncertainty. Retrieved 12 June 2021, from https://www.policyuncertainty.com/

Bongaarts, J., & Watkins, S. C. (1996). Social interactions and contemporary fertility transitions. *Population and development review*, 639-682.

Boyd, R., & Richerson, P. J. (1988). *Culture and the evolutionary process*. University of Chicago press.

Busetta, A., Mendola, D., & Vignoli, D. (2019). Persistent joblessness and fertility intentions. *Demographic Research*, *40*, 185-218.

Caldwell, J. C. (1980). Mass education as a determinant of the timing of fertility decline. *Population and development review*, 225-255.

Campisi, N., Kulu, H., Mikolai, J., Klüsener, S., & Myrskylä, M. (2020). Spatial variation in fertility across Europe: Patterns and determinants. *Population, Space and Place*, *26*(4), e2308.

Carlson, E. A. (1966). The gene: a critical history. *The gene: a critical history.*

Casterline, J. B., & National Research Council (US) Committee on Population. (2001). Diffusion processes and fertility transition: Introduction. In *Diffusion processes and fertility transition: Selected perspectives*. National Academies Press (US).

Cleland, J., & Wilson, C. (1987). Demand theories of the fertility transition: An iconoclastic view. *Population studies*, *41*(1), 5-30.

Comolli, C. L., Neyer, G., Andersson, G., Dommermuth, L., Fallesen, P., Jalovaara, M., ... & Lappegård, T. (2021). Beyond the economic gaze: Childbearing during and after recessions in the Nordic countries. *European Journal of Population*, *37*(2), 473-520.

Duvander, A. Z., Lappegård, T., Andersen, S. N., Garðarsdóttir, Ó., Neyer, G., & Viklund, I. (2019). Parental leave policies and continued childbearing in Iceland, Norway, and Sweden. *Demographic Research*, *40*, 1501-1528.

Fernández Cordón, J. A. (1986). Análisis longitudinal de la fecundidad en España. *Tendencias demográficas y planificación económica. Madrid: Ministerio de Economía y Hacienda*, 49-75.

Gerson, K. (1993). No man's land: Men's changing commitments to family and work.

Gillespie, R. (2003). Childfree and feminine: Understanding the gender identity of voluntarily childless women. *Gender & Society*, *17*(1), 122-136.

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

Hicks, W. W., & Martínez-Aguado, T. (1987). Las determinantes de la fecundidad dentro del matrimonio en España. *Reis*, (39), 195-212.

Ireland, M. S. (1993). *Reconceiving women: Separating motherhood from female identity*. Guilford Press.

Jansen, M., & Liefbroer, A. C. (2006). Couples' attitudes, childbirth, and the division of labor. *Journal of family issues*, *27*(11), 1487-1511.

Kaplan, H. (1996). A theory of fertility and parental investment in traditional and modern human societies. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists*, *101*(S23), 91-135.

Kulu, H., & Vikat, A. (2007). Fertility differences by housing type: The effect of housing conditions or of selective moves?. *Demographic research*, *17*, 775-802.

Malthus, T. R. (1872). *An Essay on the Principle of Population.*.

Matysiak, A., Sobotka, T., & Vignoli, D. (2021). The Great Recession and fertility in Europe: A sub-national analysis. *European Journal of Population*, *37*(1), 29-64.

Miret, P. (2019). ¿ Reacias a la maternidad? Primofecundidad en España a principios del siglo XXI. *Encrucijadas: Revista Crítica de Ciencias Sociales*, (18), 7.

Montgomery, M. R., & Casterline, J. B. (1996). Social learning, social influence, and new models of fertility. *Population and development review*, *22*, 151-175.

Mulder, M. B. (1998). The demographic transition: are we any closer to an evolutionary explanation?. *Trends in ecology & evolution*, *13*(7), 266-270.

Novales, A. (1993). Econometría (2ª edición). *McGraw-Hil*.

Park, K. (2005). Choosing childlessness: Weber's typology of action and motives of the voluntarily childless. *Sociological inquiry*, *75*(3), 372-402.

Pérusse, D. (1993). Cultural and reproductive success in industrial societies: Testing the relationship at the proximate. *Behavioral and Brain Sciences*, *16*, 2.

Puig-Barrachina, V., Giró, P., Artazcoz, L., Bartoll, X., Cortés-Franch, I., Fernández, A., ... & Borrell, C. (2020). The impact of active labour market policies on health outcomes: a scoping review. *European journal of public health*, *30*(1), 36-42.

Requena, M., & Salazar, L. (2014). Education, marriage, and fertility: The Spanish case. *Journal of Family History*, *39*(3), 283-302.

Rønsen, M., & Sundström, M. (2002). Family policy and after-birth employment among new mothers–A comparison of Finland, Norway and Sweden. *European Journal of Population/Revue europeenne de demographie*, *18*(2), 121-152.

Rowland, D. T. (2007). Historical trends in childlessness. *Journal of family Issues*, *28*(10), 1311-1337.

Sánchez-Barricarte, J. J. (2017). The long-term determinants of marital fertility in the developed world (19th and 20th centuries) The role of welfare policies. *Demographic Research*, *36*, 1255-1298.

Sánchez-Barricarte, J. J. (2019). Factores determinantes del descenso histórico de la fecundidad marital en España. *Revista Internacional de Sociología*, *77*(3), 133.

Settle, B., & Brumley, K. (2014). 'It’s the Choices You Make That Get You There': Decision-Making Pathways of Childfree Women. *Michigan Family Review*, *18*(1).

Smith, C. C., & Fretwell, S. D. (1974). The optimal balance between size and number of offspring. *The American Naturalist*, *108*(962), 499-506.

Sulis, E., Bosco, C., Patti, V., Lai, M., Farías, D. I. H., Mencarini, L., ... & Vignoli, D. (2016). Subjective well-being and social media: A semantically annotated Twitter corpus on fertility and parenthood. In *3rd Italian Conference on Computational Linguistics, CLiC-it 2016 and 5th Evaluation Campaign of Natural Language Processing and Speech Tools for Italian, EVALITA 2016* (Vol. 1749, pp. 1-6). CEUR-WS.

Wilde, J., Chen, W., & Lohmann, S. (2020). *COVID-19 and the future of US fertility: what can we learn from Google?* (No. 13776). IZA Discussion Papers.