Item 6 – Forecasting demographic components: fertility

Forecasting the number of births in Portugal

Antonio Caleiro¹, University of Évora, Portugal

ABSTRACT

Portugal is characterised by a noteworthy decline in fertility: a phenomenon that requires some intervention given the costs, namely economic and social, associated with it. Notwithstanding the downward trend in fertility, a careful observation of the number of births in Portugal indicates that there are months where the number of births is clearly higher, as well as others where it seems to be lower. Generally speaking, in analysing the seasonal patterns of births in different societies and over different time periods, the literature has considered two major kinds of explanations: (a) societal/cultural explanations, based upon social phenomena such as religious practice, marriage patterns and the timing of holidays, and (b) environmental/climatic explanations, based upon variations in temperature and photoperiod. The paper extends the first kind of explanation for birth seasonality by the consideration of socio-economic factors, namely marriage patterns and end-of-year (economic) expectations, in order to understand how these factors help improving the forecasts of births. This is done using ARIMA models with regression variables allowing for seasonal effects. In particular, ex-post prediction analyses are performed in order to assess the importance of those socio-economic factors when forecasting the number of births in Portugal.

1. INTRODUCTION AND MOTIVATION

The decline in fertility that characterizes many countries around the world is particularly evident in Portugal. This phenomenon requires some intervention given the costs, namely economic and social, associated with it. Despite the significant downward trend in fertility, a careful observation of the data on the number of births in Portugal indicates that there are months where the number of births is apparently much higher, as well as others where the number of births seems to be much lower (Caleiro, 2010).

The existence of peaks and valleys in the number of births (in Portugal) leads to the need of identification of the factors that apparently explain that kind of seasonality. This is important as these factors could then be used as a basis for demographic policies intended to increase fertility or, at least, to reduce the decrease in it.

Generally speaking, in analysing the seasonal patterns of births, the literature has considered two major kinds of explanations: (a) societal/cultural explanations, based upon social phenomena such as religious practice, marriage patterns and the timing of holidays, and (b) environmental/climatic explanations, based upon variations in temperature and photoperiod.

In fact, for some time, authors have been proposing climate/weather or, more specifically, temperature, as an explanatory factor, although other factors are generally considered to be necessary to an explanation of the observed seasonal patterns on births (see, among others, Seiver, 1989). The most consistent result seems to be that extreme temperatures, especially summer heat, according to Lam & Miron (1991, 1996), suppress fecundity. This may partly explain the September peak in births for some northern hemisphere countries, but, even after controlling for temperature, some peaks in births, such as the persistent spring peak in births in northern Europe, do not seem to be completely explainable by temperature.

¹ Departamento de Economia & CEFAGE–UE, e-mail: caileiro@uevora.pt
² As a matter of fact, the seasonality in the number of births is an issue that has long deserved some attention. In a seminal work, Huntington (1938) called attention to the fact that seasonal variations in the number of births could essentially be explained by reasons related to climate/weather.
Closely related is the importance of photoperiod on reproduction. Barber (2002) supports the hypothesis that human reproduction is indeed suppressed by short photoperiods and low temperatures. Mathers & Harris (1983) also seem to confirm that environmental factors are more important than socio-cultural ones given the significant geographical trend in the seasonality of births that characterise a single, but vast enough, country, such as Australia, where the more northern states present a February–March peak whereas the more southern states present a September–October peak.\(^3\)

From the previous studies one may conclude that environmental factors may be (more) important but also that they do not explain all the seasonal variations in the number of births. In general, several authors call attention to the additional (primordial or secondary) importance of socio-cultural factors, namely marriage patterns and holiday practices (related to religiosity).\(^4\)

For instance, Matsuda & Kahyo (1994), considering the case of Japan from 1974 to 1983, conclude that seasonal variations in marriage are relevant to the seasonality of first births, while the seasonality of subsequent births is essentially due to other features such as environmental ones. Chatterjee & Acharya (2000) conclude that the distribution of conceptions over calendar months in rural west Bengal is indeed negatively associated with the average monthly temperature but that the marriage pattern of the community also has a significant effect on the monthly distribution of births. Grech et al. (2003), considering the case of Malta, conclude that the seasonality of births is closely related to the seasonality of marriages, as is also concluded by Demoliates & Katsouyiannopoulos (1995) for Greece, 1956-80.

As a matter of fact, most of the authors acknowledging the importance of marriage seasonality also admit that this is not the most important factor affecting the seasonality of births (Trovet & Odynak, 1993). Other authors do not even find any relationship between the two – see Polašek et al. (2005) for the case of Croatia. Rather, they point to the increased number of conceptions during (religious) holidays, such as Easter and, most notably, Christmas, as the most important explanation for birth seasonality. This would explain the spring and September peaks in births (James, 1990; Polašek et al., 2005).

Concerning the September peak in births, it is particularly worth mentioning that this seasonality may be associated with a “Christmas effect”, although, in fact, this may not necessarily be of a strictly Christian religious nature given that the peak in births in that month is also a characteristic of non-Christian countries (Cesario, 2002). Some authors call attention to the importance of leisure time, which increases during holidays or vacations. Following this argument, the September peak in births could substantiate such a ‘holiday theory’. This is the explanation provided by Haandrikman & van Wissen (2008) considering the case of The Netherlands. This view could also be used to explain the (late) spring peak in births following increased leisure time during summer vacations, as generally happens in Europe (Cesario, 2002).

Despite some evidence supporting the diminishment of birth seasonality around the world (Seiver, 1985; Roenneberg, 2004; Cancho-Candela et al., 2007), in what concerns Portugal seasonality in births seems to exist, being more evident in the most recent years. This impression is confirmed by Caleiro (2010), based upon a time series analysis of the data, which show that, in general, May and September are, indeed, months where more births take place and that December and February are months where fewer births take place. In particular, the September peak in births is the most impressive result.

With regards to the main results of the literature, as surveyed above, the results in Caleiro (2010) are more in accordance with a predominance of socio-cultural factors as the main determinants of the seasonality of births in Portugal. In particular, the local peak in May seems to reflect both an increase in leisure time associated with summer vacations, as well as the surge in marriages that, traditionally, take place in August in Portugal. Regarding the birth peak on September, which may also reflect the local peak in weddings that take place on December, this seems to be closely associated with an aspect that the literature has largely ignored, i.e. the expectations that couples form at the end of the year, which depend upon confidence in the economic evolution of the country.

This paper builds upon the kind of explanation for birth seasonality in Portugal offered by Caleiro (2010), namely marriage patterns and end-of-year (economic) expectations, in order to understand how these factors help improving the forecasts of births. This is done using ARIMA models with regression variables allowing for seasonal effects (and eventual differencing of the data). In particular, \textit{ex-post} prediction analyses are performed in order to assess the importance of those socio-economic factors when forecasting the number of births in Portugal.

The remaining part of this paper is structured as follows. Section 2 describes the data. Section 3 introduces and applies the specific time series methodology in order to forecast the number of births. Section 4 concludes.

\(^3\) Nevertheless, for the Czech Republic, for instance, Bobak & Gjonca (2001) show that socio-demographic factors are more important than temperature or photoperiod.

\(^4\) As Polašek et al. (2005) point out, the seasonality of marriages could be related to climatic and/or cultural factors given that the appropriate period for weddings may depend on meteorologically suitable conditions and/or the existence of (religious) holidays.
2. THE DATA

The data for the number of births by month, which covers the period January 1977 to December 2008, is readily available at the Eurostat site. These observations, in terms of daily averages given the different durations of the months, are plotted in Figure 1.

![Figure 1: The daily average of births in Portugal](image1)

Roughly speaking, Figure 1 shows that from 1977 onwards there was a general decline in the number of births, which lasted up to around 1996, after which a tentative increase could be observed until 2001. This increase in the number of births is said to be due to a specific demographic policy aimed at family support, which coincided with a favourable economic situation in the country. Besides these characteristics, there is also marked seasonality in the data, in particular at the end of the period.

In what concerns marriages, the data for the number of marriages by month, which covers the period January 1977 to December 2008, is also readily available at the Eurostat site. These 384 observations, in terms of daily averages given the different durations of the months, are plotted in Figure 2.

![Figure 2: The daily average of marriages in Portugal](image2)

---

6 In fact, the available data (for Portugal) starts on January 1950 and finishes on December 2006. Given the beginning of data for births we consider only data after January of 1977. The data for 2007 and 2008 was obtained through the Estatísticas Demográficas from the Instituto Nacional de Estatística.
Figure 2 shows a tendency for a diminishment in the number of marriages and a clear seasonality. In fact, August is an evident peak and December is a peak of second order in marriages.

In what concerns confidence, the data covers the period June 1986 to December 2008, and is also readily available at the Eurostat site. Figure 3 plots the data. It shows that ure 3 shows the usual behaviour of confidence, i.e. some evident volatility, being also clear an abrupt decline around the end of 1991 and a significant increase at the end of 1993 until mid-1997, followed by another significant decline until the end of period, despite a slight recover during the period 2003-2006.

3. THE FORECAST OF THE NUMBER OF BIRTHS

Observing the pattern exhibited by the number of births brings about a crucial question: if there are months that are characterised by a significantly different number of births, what explains it? In particular, it seems important to understand what may lead couples to have more babies in certain months of the year, whether from planned or unplanned pregnancies. An explanation to this fact is provided in Caleiro (2010). Given that there are regularly peaks in the number of births in May and, even more notably, September we would like to test that explanation by the use of the well-known autoregressive moving average (ARMA) (with seasonal factors) methodology.

The ARMA($p,q$) is a model that considers $p$ autoregressive terms and $q$ moving average terms and is given by the following expression:

$$Y_t = \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{i=1}^{q} \beta_i \epsilon_{t-i} + \epsilon_t,$$

possibly adjusted to take into account seasonal factors.

However, regarding the series for the number of births, the application of that model appears to be a problem, given the need of considering stationary series. From Figure 1, is evident that the time series of births is likely to have a downward trend and seasonal spikes, which implies level non-stationarity. Therefore we decided to use the methodology behind the SEATS/TRAMO approach as it handles, in an accurate way, all the problems that the original series seems to present (Gómez & Maravall, 1996).

By the use of that methodology a one-year ex-post forecast of the number of births, i.e. until December 2008, is particularly appropriate as it enables to verify how the time series of births itself is able to forecast well the number of births.\(^7\) Figure 4 shows the results.

\(^7\) This methodology was applied through the module TramoSeat in gretl (Gnu Regression, Econometrics and Time-series Library), freely available at http://gretl.sourceforge.net.
Plainly, the *ex-post* prediction values are reasonably close to the actual ones, despite being evident also some underestimation in the forecasted number of births. This leads to a possible improvement of the forecasts by the consideration of explanatory factors (other than the series of births itself). Furthermore, of particular relevance is the notable peak of births in September. Although the local peak of births in May does not characterize 2008, Caleiro (2010) shows that it also relevant.

From the explanatory factors suggested in the literature, the peaks in births can be associated with socio-cultural practices, such as the marriage pattern, which in Portugal, shows a global peak in August and a local peak in December (Caleiro, 2008). As shown in Caleiro (2008), the degree of synchronisation between marriage seasonality and birth seasonality, where a nine month lag is particularly relevant, and the fact that marriages do explain births in Portugal, support the fact that the May peak in births may reflect the marriage practice in Portugal. In addition, most Portuguese choose to take their summer vacation in August.

Considering the number of marriages occurred nine months before as an explanatory factor of births, requires the use of autoregressive moving average models with exogenous variables (ARMAX). In this case ARMAX\((p,q,b)\) is a model that considers \(p\) autoregressive terms, \(q\) moving average terms and the last \(b\) terms of an exogenous time series \(X_t\). As such, is given by the following expression:

\[
Y_t = \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{i=1}^{q} \beta_i \epsilon_{t-i} + \sum_{i=1}^{b} \gamma_i X_{t-i} + \epsilon_t.
\]

By the use of the methodology behind SEATS/TRAMO, the ex-post prediction results for 2008 (as well as for 2009), with regression effects of marriage assigned to the seasonal component, are plotted in Figure 5.

---

*Unfortunately, due to the kind of data that is available, Caleiro (2008) does not distinguish first births from the rest of the parities. Still, if the seasonality of marriages does explain the seasonality of all births, it is to be expected that more/better would explain first births.*

*As a curiosity, when the author asked an obstetrician if summer vacations could help explain the birth peak in May, the answer was clearly in favour of an increased number of conceptions being due to the increase of leisure time, enjoyed together by couples, on summer holidays.*

*The results were obtained through the use of TSW (Tramo Seats for Windows), freely available at [http://www.bde.es/servicio/software/tswe.htm](http://www.bde.es/servicio/software/tswe.htm).*

---

Figure 4: The *ex-post* prediction of the number of births for 2008
Figure 5: The ex-post prediction of the number of births (with marriages)

In order to verify how the inclusion of marriages in the model changed the forecasts of births let us consider Figure 6, which shows the results of both ex-post forecasts exercises, as well as the real numbers for births in 2008.

Figure 6: The ex-post prediction of the number of births (with and without marriages)

Clearly, the consideration of marriages helps to improve the forecasts, even if by some a small amount, as it does not underestimate so much the real numbers. Figure 6 shows also that the September peak, despite being detected, assumed a value considerably higher than the forecasted one. As pointed out before, the most remarkable peak of September could be explained as some sort of holiday-related peak, or due to the increased number of conceptions allegedly taking place immediately after (or even before) marriages in December, (which register an increase in Portugal in this month). However, this does not appear to be the complete story.

Indeed, if it were only the result of those two factors, i.e. some sort of Christmas effect or mid-winter festivities, in conjunction with the increased number of marriages that take place in December, then one should obtain peaks in births in (almost) every September. In fact, as shown in Caleiro (2010), there are some years where that does not occur. Moreover, even in the years associated with the existence of a September peak in births these register somewhat distinct magnitudes/sizes. It is our hypothesis that this has to do with a factor that, to the best of our knowledge, has not yet been explicitly considered as an explanation for the September peak in births, i.e. the expectations that couples form at the end of the year about the situation they will face during the year ahead.

In order to proxy the level of expectations the most appropriate measure is the (consumer) confidence indicator, given its prospective nature and applicability to the generality of the population – see Caleiro (2006) to verify how this prospective indicator relates to some relevant macroeconomic variables. Given that the level of confidence is quite
sensitive to the evolution of the economy, we decided to consider this other explanatory variable but in differenced terms. Again using the methodology behind SEATS/TRAMO, the ex-post prediction results for 2008 (as well as for 2009), with regression effects of marriage and confidence assigned to the seasonal component, are to be obtained.

Due to the fact that data for confidence only starts in June 1986, these results are not directly comparable to those presented so far, as these consider the period 1977-2008. As a matter of fact, the data for births starts on March 1987 (data for marriages and variation of confidence start 9 months before) and finishes on December 2007, leaving the months of 2008 as the ex-post prediction interval. The results are plotted in figures 7 and 8.

Figure 7: Forecast without regressors

Figure 8: Forecast with regressors

Clearly, the results are similar, despite being also visible some change in the pattern of forecasts and the end of the year. Figure 9 clarifies these facts by plotting both kind of forecasts as well as the actual numbers of births that occurred in 2008.

Figure 9: A comparison of the results

The results clearly show that the inclusion of the regressors marginally increased the quality of results. As a matter of fact, this is already of significance given that when the ARIMA (without regressors) methodology is correctly applied (i.e. by experienced experts and/or recurring to well-founded automatic procedures, such as the one considered in SEATS/TRAMO), the inclusion of lagged values of regressors may not improve the forecasts of an ARIMA model, which has already exploited the history of the original time series.
4. CONCLUSIONS AND DIRECTIONS FOR FURTHER RESEARCH

This paper builds upon the existence of a month effect, understood as a seasonality effect, in the number of births in Portugal. This issue is assumed to be relevant because, as we know, Portugal has been characterised by a remarkable decline in fertility: roughly speaking, at the beginning of the 1970s about 500 babies were born every day, while this number is now around 300. This is, obviously, a serious problem that needs to be addressed, given the costs of this phenomenon.\footnote{These include some (well-known) consequences, such as the pressure on the social security system, as well as some other (not so well-known) consequences (not necessarily costs), such as the fact that the ageing population makes pensioners more important from an electoral point of view.}

The paper extends the usual explanations for birth seasonality by the consideration of socio-economic factors, namely marriage patterns and end-of-year (economic) expectations, in order to understand how these factors help improving the forecasts of births. This is done using ARIMA models with regression variables allowing for seasonal effects. In particular, ex-post prediction analyses are performed in order to assess the importance of those socio-economic factors when forecasting the number of births in Portugal.

The results show that those two factors help in forecasting the number of births despite being apparent a subsidiary importance. As a matter of fact, an improvement in forecasting by the use of the ARIMA methodology, when this is already being used in a most efficient way, should be always be non-ignorable.

As for further work, we would like to consider other data allowing us to better consider the possible link between births and marriages, which clearly is to exist (if so) essentially for first births. To robustify the results, the analysis of births that occur outside marriages is also a possibility. In what concerns confidence, it seems possible to consider other approaches rather than assuming that the variation over two months is a factor that may help to explain some variation in the number of births (Caleiro, 2006).

5. REFERENCES


