Birth and Death number variation

Projects in Bioinformatics

Spring Semester 2018/2019

Name: Tomas Sladecek

Supervisor: Palle Villesen Fredsted

1 Introduction

Births along with deaths are two life events inevitable in life of every man. These events mark the two major timestamps and neither one can be completely influenced by our actions or the decisions of others (not regarding murders, since their numbers are negligible in our analyses). In this study we focused on studying birth and death number patterns between different countries and also the yearly birth numbers in Denmark, with the intention to connect famous events with deviations from expected behaviour.

2 Yearly Birth Variation: 1901-2017 (Denmark)

The twentieth century is a century that feels rather polemic. On one hand, there were two world wars, several financial crises, cold war, nuclear disasters, etc. and at the same time, it was the period of many well-known scientific discoveries, inventions and cultural novelties.

Figure 1 shows the pattern of number of births in the period from 1901 to 2017. Several

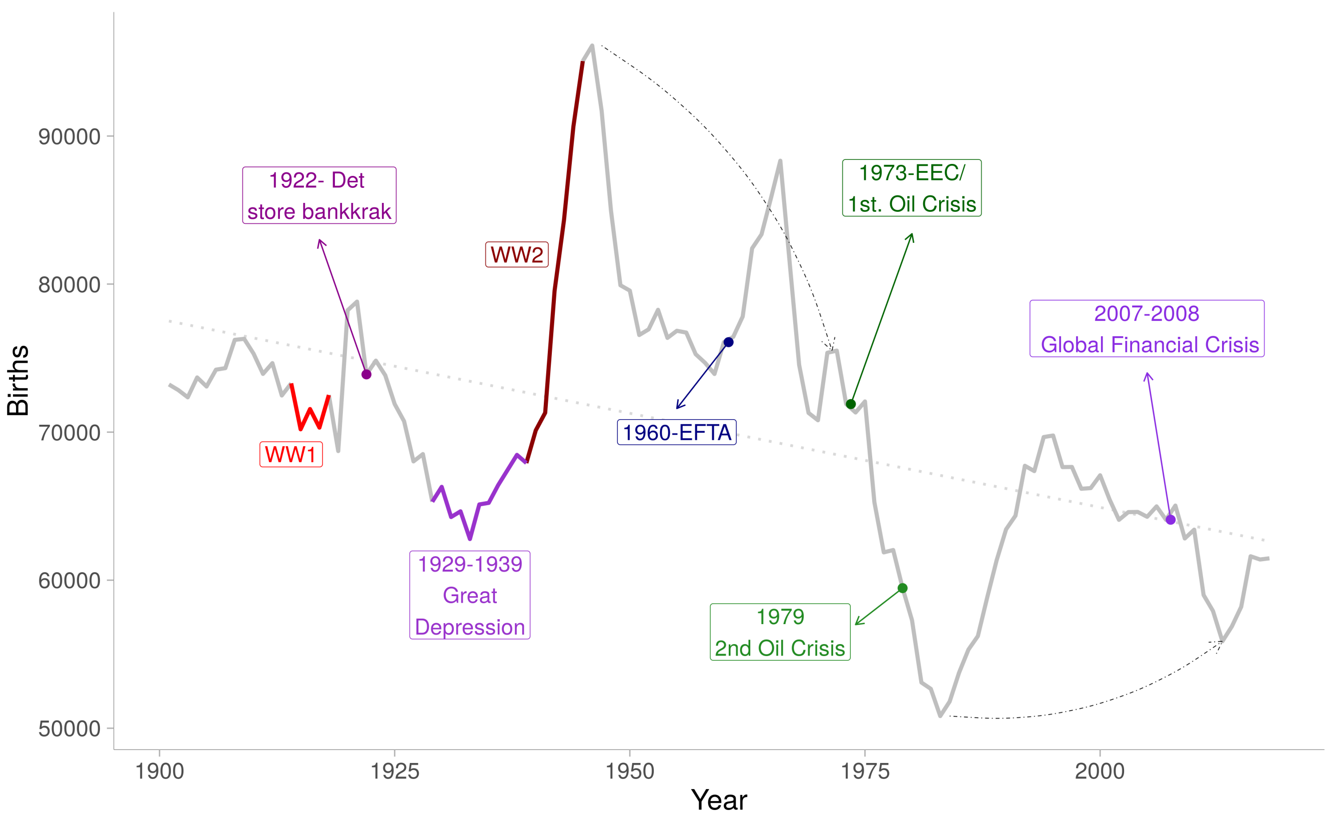

Figure 1: *Yearly variation of live births in Denmark*

events that might have influenced peoples behaviour are shown too. Yet, all of them are only hypotheses and we can not tell with complete certainty that any one of the events was causal for some observed patterns. Nevertheless, we can at least try to describe them and lay foundations to further research.

The expectation is, that after an event that affects the economic and/or existential status of society, there will be some deviation from the overall birth pattern.

For Denmark it seems like the number of births has been decreasing since 1901 (faint dotted line). However, the mortality of children during the first 5 years of their life, shows similar pattern until approx. 1980, and then it stabilizes [1]. With better healthcare and living conditions, the number of births decreased. This is a well known pattern for many countries, where the number of babies per woman and income are inversely correlated [2]. A similar pattern can be seen with child mortality when plotted against income [3].

During the first world war (WW1) there is not any strong pattern or deviation from other years in terms of birth numbers. And this is, in a way, not unexpected since, as some say, "World War One was a war that Denmark and the Danes “missed”, that the war remained a rather distant "southern thunder"" [4]. The most influential event for Denmark came in 1917, when US Army banned imports of goods to Denmark. This led to development of new trade system between Scandinavian countries, but Denmark remained still highly dependent on imports from Germany, at least until September 1918 [4].

Few years after WW1, there was a slight increase in births numbers but from around 1922 it had begun to decrease steadily. One of the events known to Danish people from this period is the so called 'det store bankkrak', where Danish government had lost about 500 million kroner, due to money circles introduced by Emil Glückstadt, the chairman of the involved Farmer's bank [5].

Birth numbers continued to decrease until 1933. This period is known as the Great Depression, which was "the worst economic downturn in the history of the industrialized world" [6], and which impact on economic and social status of people helped the rise of extremist political movements, such as Hitler's Nazi regime.

Between 1933 and 1946 birth numbers had been rising quite steadily, even during the infamous world war 2. This time, Denmark was more involved because of the occupation by Germany in 1940. In order to remain in as peaceful conditions as possible, it began cooperating with its southern neighbour and agreed on supplying food and raw materials. This agreement meant that the life standards of many people did not change drastically [7, 8]. Nevertheless, the pattern of live births is very interesting, added that child mortality does not have such a big peak around this time period [1].

After the second world war, birth numbers began to decrease until circa 1960, which is the year of the enter of Denmark in the European Free Trade Association (EFTA). This, along with the EEC (European Economic Community) entry in 1973 helped the economy of Denmark, because of their great reliance on export [9]. In the same year members of OPEC (Organization of Petroleum Exporting Countries) decided to stop the export of oil to USA what resulted in almost fourfold increase in oil prices [10]. The second strike came in 1979, where sudden shortages of oil supply occurred as a consequence of Iranian revolution in 1978-1979. Governments realized, that relying only on oil is not sustainable for future development. This led to development of new alternatives, such as power plants [11].

Between the first oil crisis and present, we see 2 notable dips/valleys in birth numbers. The first might be a result of the two oil crises, while the one in 2013 is very likely a consequence of that in 1983, since the average age of mothers giving birth in 2013 was 30.9 years. However, this pattern does not hold for the peak 1966, where we would assume that the average age of mothers to be circa 20 years, since the previous peak was in 1946. It was 26.6, which might suggest that something else caused the rapid increase in births around that period. Anyway, high birth numbers during WW2 at least manifested in higher numbers in 1972, what can be seen by one of the arrows in Figure 1.

In 2008 struck one of the worst crises since great depression - The global financial crisis. If there was any effect is hard to tell, because of the correlation between dips in 2013 and 1983.

2.1 Comparison with Finland and Sweden

If we compared yearly birth number variation in Denmark with other countries such as Sweden and Finland, we would find several similar patterns. For example the valley in 1933 and subsequent peak rising through WW2 can be seen in all three countries, while steady, almost non-changing birth numbers during WW1 are only seen in Denmark. After the second world war, these behaviours begin to diverge, accounting for different development of these countries.

3 Monthly Birth and Death Variation

As we saw in the previous section, the number of born children varies non-randomly. But is also present during the year? Is there any significant variation, or is it possible to explain the changes by a simple constant curve? And is this variation distinct for different countries around the globe?

3.1 Comparison of variation for different countries

For many countries, Denmark included, birth numbers follow some monthly pattern which seems to resume every year. Figure 2 shows this behaviour and how well it is conserved in 10 years range (2007-2017). There are already some very obvious differences between our groups, like the almost machine like repetition of behaviour in Sweden and Norway and rather random pattern in Australia or Ireland.

After plotting normalized mean monthly birth numbers (Figure 3) we can see the discrepancies

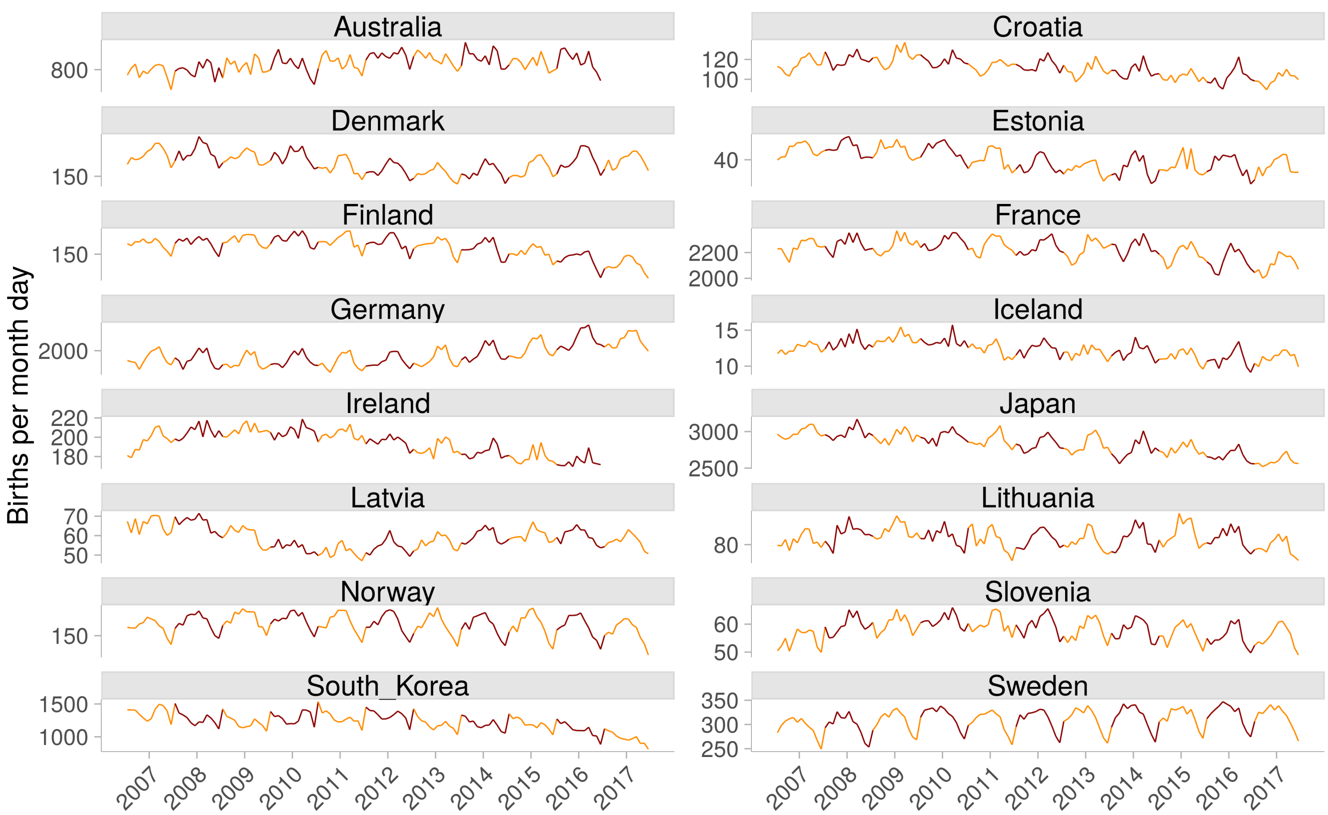

Figure 2: *Monthly variation of live birth numbers in different countries. To address the inequality of number of days in different months, we divided the number of births in each month by the number of days in that particular month.*

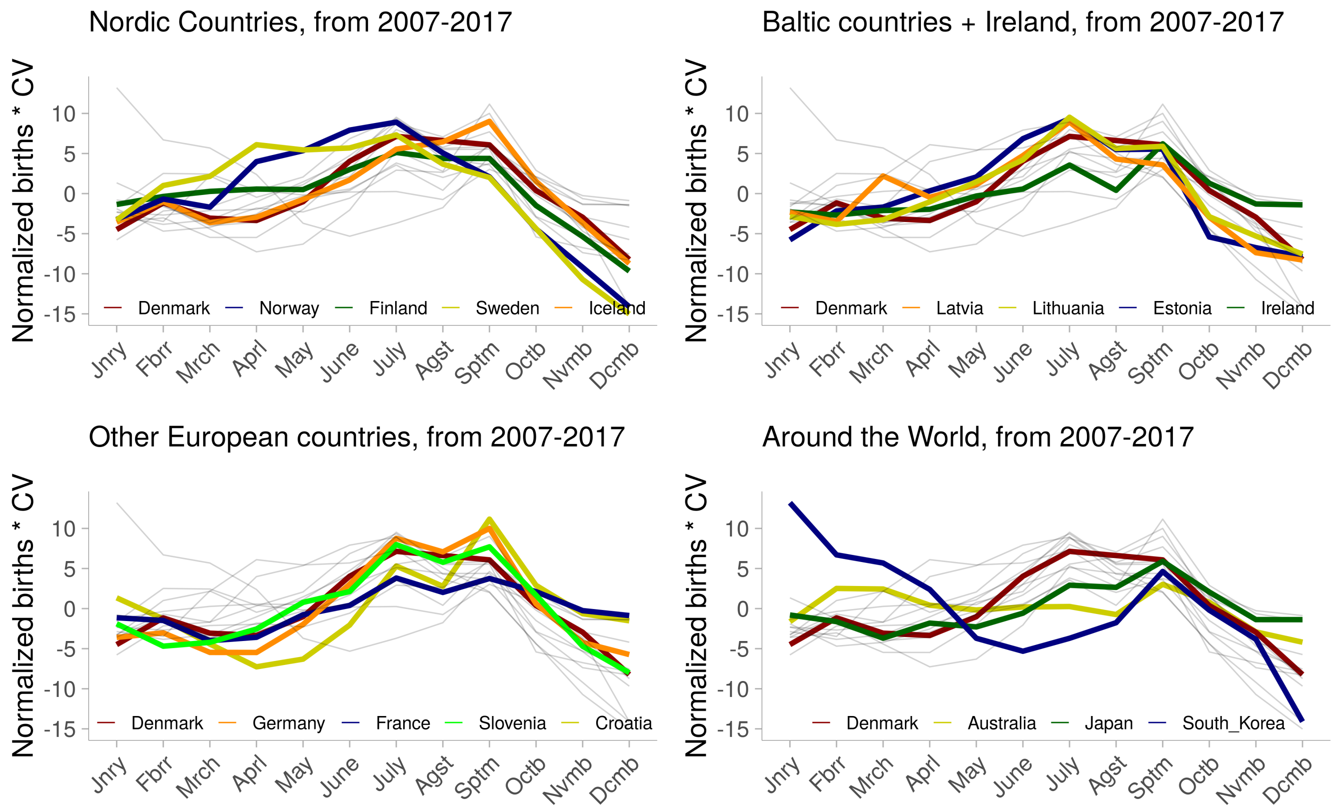

Figure 3: *Comparison of birth number variation in different countries. In order to compare different countries, birth numbers are normalized (corrected for unequal month day numbers). To retain the information about the magnitude of variation, we multiplied these normalized births by the coefficient of variation.*

between different groups of countries. The first three of those are composed of European countries, which seem to follow very similar pattern, meaning that more babies are born in summer months and less during winter. The situation is more difficult when comparing Denmark (which was taken as the representative of European countries) and the studied non-European countries, since the behaviour of births does not seem to follow a strict cyclical pattern[12–27].

3.2 Modelling the Variation

Before we can start finding causes underlying the variation, we have to prove that it is not just a random noise. Figure 2 and Figure 3 suggest, that the variation might be well explained by a sinusoidal curve. Our null hypothesis is that birth numbers do not change much throughout the year, and the observed deviations are seen only due to statistical noise.

Figure 4 depicts our two hypotheses and how the data might have looked like if it was randomly generated according to certain criteria. We fitted a constant (orange) and sinusoidal (blue) curve

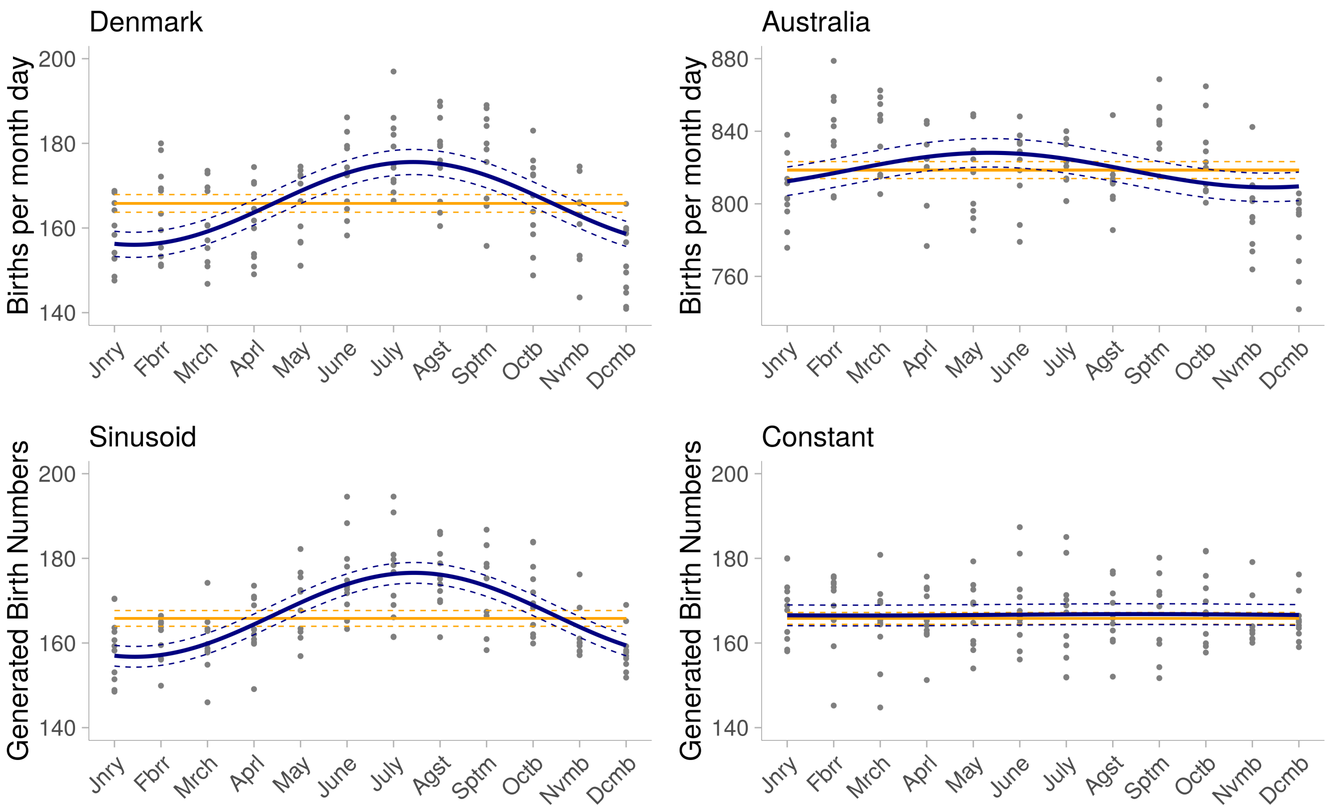

Figure 4: *Comparison of real and modelled sinusoidal and constant birth number variation. Top row shows real data, while the second shows modelled variation, with parameters (intercept, standard deviation, stretch, phase shift) estimated from real data.*

to both real and modelled data, along with their confidence intervals (in this case it is just 2 times the standard error). This was done to visually compare randomly generated noise with real data and how the fitting part would perform there.

The sinusoidal fit seems to capture the behaviour of birth numbers quite well for Denmark, while the opposite is true for Australia. To prove this statistically, we performed the analysis of variance, where we compared these two models for each country.

The results are shown in Figure 5. For most of the countries the sinusoidal fit was significantly better than a constant one, even after correcting for the number of carried out tests. The only two countries that fell behind were Australia and South Korea. In a way, we expected this for Australia, since the variation does not seem to have any particular behaviour. However, for South Korea the situation is more difficult to address, because there is strong suspicion that the data provided are not entirely correct. One can see that in Figure 2 because every year there is a huge spike somewhere in December/January. We did not find any explanation for this phenomenon,

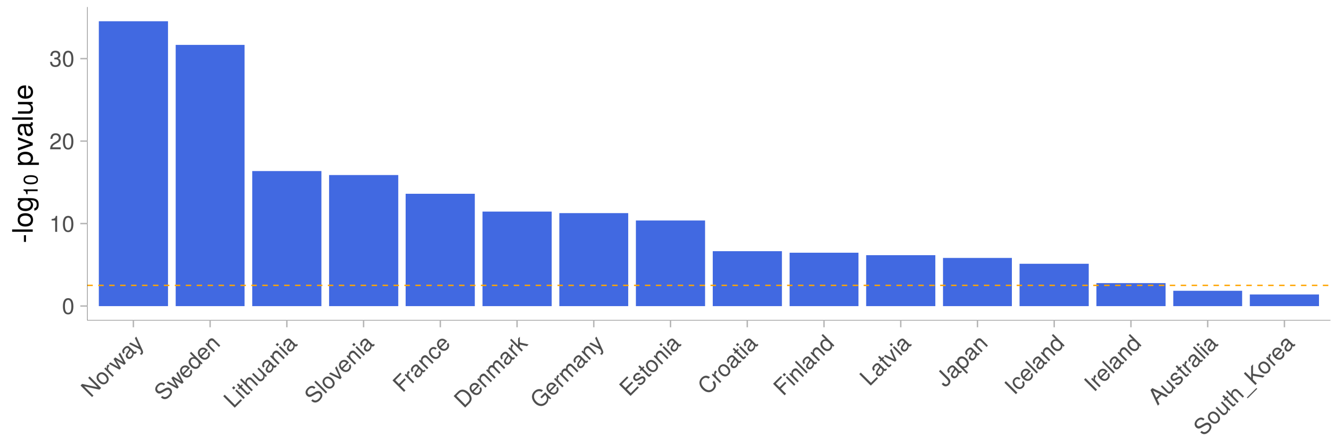

Figure 5: *p-values on logarithmic scale from comparisons of sinusoidal and constant fits. The orange line represents corrected significance level (Bonferroni correction).*

however one speculation might be that births that could not be assigned to other days in a year, are automatically assigned to the 1. of January, as default.

Even though these probabilities look very promising, we are actually interested in the amount of explained variation by the sinusoidal fit and also how strong the effect is. Here, by effect we understand the magnitude of variation with respect to the overall mean of birth numbers, both of which are shown in the top row of Figure 6. Together with the results from Figure 5 we can see a pattern, where the Nordic countries (except of Finland) seem to have very conservative periodical fluctuation of birth numbers. The explained variation decreases as we move farther away from Norway (~70%), where most for most of the European countries the sinusoidal curve explained around 20-40% of the variation. The lowest score was recorded by Ireland (~9%) followed by Iceland (~15%). The non-European countries scored very low - 3%, 5% and 18% for South Korea, Australia and Japan respectively.

The strength of the effect is correlated positively with the explained variation ($r \approx 0.78$), which is not surprising for countries scoring low in the $Adj.R^2$ metrics. But if the relationship is linear, this would imply that if the variation is well explained by sinusoid, than it is very likely that the amount stretch divided by intercept rises accordingly.

3.3 Births and Deaths

Similarly to previous discussion about birth number variation, we looked at other vital event - deaths. Again, we assumed that the numbers of deaths per month are uniformly distributed

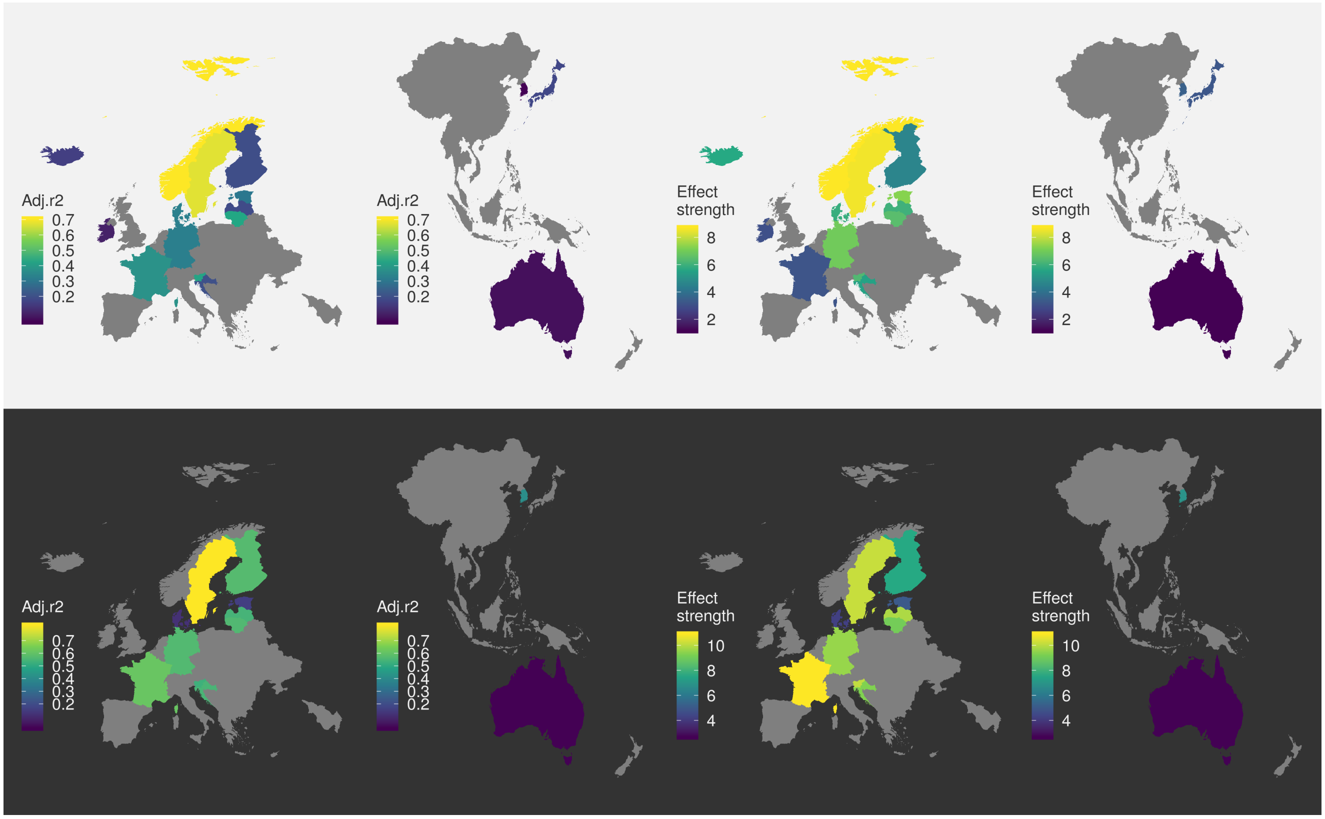

Figure 6: *Maps of countries used in birth number analyses. Colour fill is based on either the explained variation by sinusoidal fit (Adjusted R squared) or the effect size (stretch of the sinusoid divided by its intercept represented in percents). The plot at the top (with a lighter background) represents births, while the one with the dark background represents deaths.*

throughout the year and as an alternative we offered the sinusoidal model. Even though that both of these events are strictly biological, births can be planned to a certain extent. It makes sense that people prefer births after winter, because in the first few month babies are very weak and fragile and people are more prone to diseases in winter, or more precisely in cold weather [28]. On the other hand, it is not possible to plain natural death, which makes the results in Figure 7 and Figure 6 (bottom row with dark background) very appealing.

For most of the countries, the peak is observed in winter months and valley somewhere in summer. Swedish people, again, seem to be very "behaved" and the patterns for each year do not deviate much from other years. France showed more cyclical death variation than birth and the strongest effect size of all studied countries. In Denmark it looks like that the pattern was well conserved until around 2012/2013 where the death numbers began to wiggle in, as it seems like, random fashion. This means that the sinusoidal pattern was broken, and it makes sense why in Figure 6 it is coloured with such dark shade of purple. Australia once again scored very low,

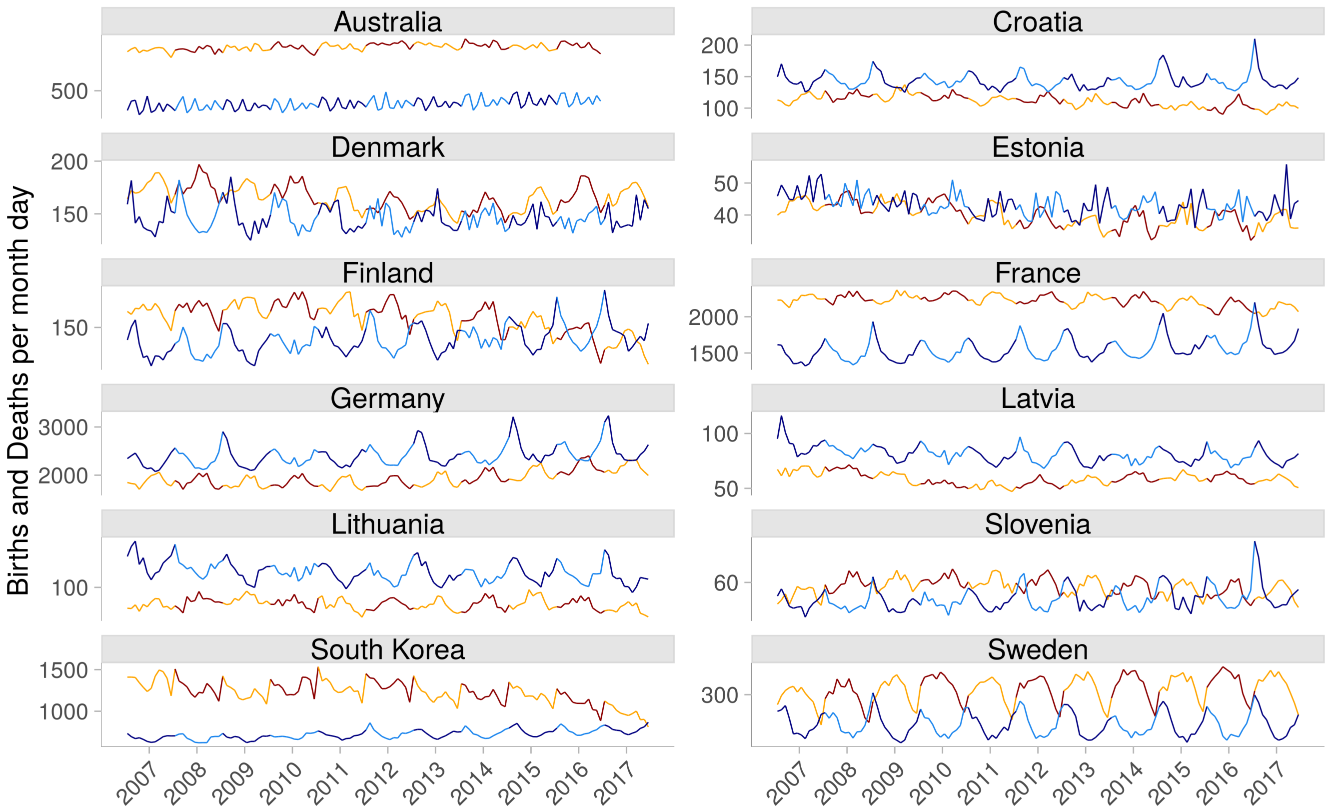

Figure 7: *Monthly birth (orange/red curve) and death (blue/dark-blue curve) variation.*

meaning that the pattern is either constant or more complicated than the proposed sinusoidal one [12–15, 17, 19–22, 25–27, 29–32].

One of the possible explanations to the question of why more people die in winter than in summer, is the higher occurrence of ”common” diseases (such as flu). There are many suggestions why this is happening. First there is less sunlight during winter, meaning that people fail to acquire the required dose of vitamin D, which strengthens our immune system. Than, since the weather is usually wetter and colder, people spend more time inside with others, what helps the spread of viruses. But probably the main reason is the dry air. Even though winter months are usually rainier/snowier than other, the air is actually drier which, as it turns out is a great condition for the spread of viruses [28]. All this likely means, that people with already weak immunity might have harder times in cold winter months, what probably manifests in higher death numbers.

4 Conclusion

Each season in the year brings along with its unique natural and climatic changes also some downsides. According to some results we have a suspicion that warmer seasons are considered by many as better choices to bring newborns into the world. Cold weather brings with it many common diseases which probably result in an increase in death numbers.

Looking at birth number from higher (yearly) perspective, we observed that it would seem very unlikely to witness such deviations as those shown in Figure 1 just by chance. There were definitely events that influenced choice of people to conceive children and it might be interesting to compare these patterns in different countries and find possible causes for common deviations.

References

[1] Statistic Denmark. *Deaths during first five years of life by age, sex and time*. 2018. URL: http://www.statbank.dk/statbank5a/Graphics/MakeGraph.asp?menu=y&maintable=DOD2&pxfile=2019416185918245042952DOD2.px&gr_type=0&PLanguage=1.

[2] Gapminder. *Correlation between Income and the number of babies per woman. FREE TO USE! CC-BY GAPMINDER.ORG*. 2018. URL: https://www.gapminder.org/tools/#$state$time$value=2018;&marker$axis_x$domainMin:null&domainMax:null&zoomedMin:null&zoomedMax:null&scaleType=genericLog&spaceRef:null;&axis_y$which=children_per_woman_total_fertility&domainMin:null&domainMax:null&zoomedMin:null&zoomedMax:null&spaceRef:null;&size$which=_default&use=constant&domainMin:null&domainMax:null&scaleType=ordinal&extent@:0&:0.08;&spaceRef:null;;;&chart-type=bubbles.

[3] Gapminder. *Correlation between Income and the child mortality. FREE TO USE! CC-BY GAPMINDER.ORG*. 2018. URL: https://www.gapminder.org/tools/#$state$time$value=2018;&marker$axis_x$domainMin:null&domainMax:null&zoomedMin:null&zoomedMax:null&scaleType=genericLog&spaceRef:null;&axis_y$which=child_mortality_0_5_year_olds_dying_per_1000_born&domainMin:null&domainMax:null&zoomedMin:null&zoomedMax:null&scaleType=log&spaceRef:null;&size$which=_default&use=constant&domainMin:null&domainMax:null&scaleType=ordinal&extent@:0&:0.08;&spaceRef:null;;;&chart-type=bubbles.

[4] Sørensen Nils Arne. *Denmark*. 2014. URL: https://encyclopedia.1914-1918-online.net/article/denmark.

[5] CPH POST. *The biggest financial disaster of the 20th century ... in Denmark*. NA. URL: http://cphpost.dk/history/the-biggest-financial-disaster-of-the-20th-century.html.

[6] History.com Editors. *Great Depression History*. 2019. URL: https://www.history.com/topics/great-depression/great-depression-history.

[7] National Museum of Denmark. *German occupation (1940-1945)*. NA. URL: https://en.natmus.dk/historical-knowledge/denmark/german-occupation-1940-1945/.

[8] Simkin John. *Denmark in the Second World War*. 2014. URL: https://spartacus-educational.com/2WWdenmark.htm.

[9] Henriksen Ingrid. *An Economic History of Denmark*. 2006. URL: https://eh.net/encyclopedia/an-economic-history-of-denmark/.

[10] Amadeo Kimberly. *OPEC Oil Embargo, Its Causes, and the Effects of the Crisis*. 2019. URL: https://www.thebalance.com/opec-oil-embargo-causes-and-effects-of-the-crisis-3305806.

[11] Chen James. *1979 Energy Crisis*. 2018. URL: https://www.investopedia.com/terms/1/1979-energy-crisis.asp.

[12] ABS.Stat. *Australian Bureau of Statistics*. 2017. URL: http://stat.data.abs.gov.au/Index.aspx.

[13] Statistics Denmark. . . URL: https://www.statbank.dk/statbank5a/default.asp?w=1920.

[14] Statistics Finland's PX-Web databases. *003 – Live births by month in 1900 to 2018*. . URL: http://pxnet2.stat.fi/PXWeb/pxweb/en/StatFin/StatFin__vrm__synt/statfin_synt_pxt_003.px/?rxid=b542950a-7b71-48c1-b7b6-6102837fa9ef.

[15] Statistiches Bundesamt. . . URL: https://www-genesis.destatis.de/genesis/online/data?operation=sprachwechsel&language=en.

[16] Central Statistics Office. *VSA19: Births by Sex of Child, Date of Occurrence, Month of Occurrence, Year and Statistic*. . URL: https://www.cso.ie/px/pxeirestat/Statire/SelectVarVal/Define.asp?maintable=VSA19&PLanguage=0.

[17] Statistics Latvia. *IE040m. Live births and deaths by sex and by month*. . URL: http://data1.csb.gov.lv/pxweb/en/iedz/iedz__iedzskaits__isterm/IE040m.px/?rxid=d8284c56-0641-451c-8b70-b6297b58f464.

[18] Statistics Norway. *Births*. . URL: https://www.ssb.no/en/statbank/table/05531/.

[19] Korean Statistical Information Service. . . URL: http://kosis.kr/eng/statisticsList/statisticsListIndex.do?menuId=M_01_01&vwcd=MT_ETITLE&parmTabId=M_01_01&statId=1962001&themaId=#A11.3.

[20] Croatian Bureau of Statistics. . . URL: https://www.dzs.hr/default_e.htm.

[21] Statistics Estonia. *PO10: LIVE BIRTHS BY MONTH OF BIRTH.* . URL: http://pub.stat.ee/px-web.2001/Dialog/varval.asp?ma=PO10&ti=LIVE+BIRTHS+BY+MONTH+OF+BIRTH&path=../I_Databas/Population/03Vital_events/04Births/&lang=1.

[22] L'Institut national de la statistique et des études économiques. *Demography - Number of live births - France (including Mayotte since 2014) Identifier 001641601.* . URL: https://www.insee.fr/en/statistiques/serie/001641601.

[23] Statistics Iceland. *Births by months 1853-2018.* . URL: http://px.hagstofa.is/pxen/pxweb/en/Ibuar/Ibuar__Faeddirdanir__faeddir__faedingar/MAN05302.px/?rxid=86314c20-a8ee-4efd-a5bd-ad8d3664349d.

[24] Statistics Japan. . . URL: https://dashboard.e-stat.go.jp/en/dataSearch.

[25] Statistics Lithuania. . . URL: https://osp.stat.gov.lt/statistiniu-rodikliu-analize?theme=all#/.

[26] Statistical Office RS. *Live births by months, Slovenia, 1960 and from 1970 on, annually.* . URL: https://pxweb.stat.si/pxweb/Dialog/varval.asp?ma=05J1030E&ti=&path=../Database/Demographics/05_population/30_Fertility/05_05J10_Fertility_SL/&lang=1.

[27] SCB Statistics Sweden. *Births and deaths per month by sex. Year 1851 - 2018.* . URL: http://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__BE__BE0101__BE0101G/ManadFoddDod/?rxid=ca97c1f1-b891-4f5e-b825-d0520f4d8f16.

[28] Robson David. *The real reason germs spread.* 2015. URL: http://www.bbc.com/future/story/20151016-the-real-reason-germs-spread-in-the-winter.

[29] Statistics Finland's PX-Web databases. *003 – Deaths by month in 1945 to 2018.* . URL: http://pxnet2.stat.fi/PXWeb/pxweb/en/StatFin/StatFin__vrm__kuol/statfin_kuol_pxt_003.px/?rxid=b542950a-7b71-48c1-b7b6-6102837fa9ef.

[30] Statistics Estonia. *PO40: DEATHS BY MONTH OF DEATH.* . URL: http://pub.stat.ee/px-web.2001/Dialog/varval.asp?ma=PO40&ti=DEATHS+BY+MONTH+OF+DEATH&path=../I_Databas/Population/03Vital_events/06Deaths/&lang=1.

[31] L'Institut national de la statistique et des études économiques. *Demography - Number of deaths - France (including Mayotte since 2014) Identifier 001641603.* . URL: https://www.insee.fr/en/statistiques/serie/001641603#Telechargement.

[32] Statistical Office RS. *Deaths by months, Slovenia 1960 and from 1977 on, annually.* . URL: https://pxweb.stat.si/pxweb/Dialog/varval.asp?ma=05L1008E&ti=&path=../Database/Demographics/05_population/32_Mortality/05_05L10_Mortality_SL/&lang=1.