Are the birth dates of our ancestors real?
Registration of false birth dates in 20th century Poland
(Extended abstract)

Jacek Cypryjański[(1)], Marta Hozer-Koćmiel[(2)]

[(1)] Department of Information Society and IT Applications, University of Szczecin, Poland
[(2)] Department of Econometrics and Statistics, University of Szczecin, Poland

Introduction

Falsification of birth dates, i.e. registration of a date other than the actual date of birth, was first observed in Italy in 1895/96 by R. Benini and then in many other countries in the world at the beginning of the 20th century (Gini and D'Addario 1931; Livi 1928). Two types of falsifications were observed: (*P1*) registering children born in December as being born in January of the following year and (*P2*) registering children as being born on specific days of the year, month or week. Today, when we study the seasonality of births over long periods of time (e.g. Doblhammer, Rodgers, and Rau 2000; Cassel 2002; Cancho-Candela, Andrés-de Llano, and Ardura-Fernández 2007; Herteliu et al. 2015; Björnsson and Zoega 2017; Dahlberg and Andersson 2018) this phenomenon is no longer just a historical curiosity, but can disrupt research results and lead to erroneous conclusions.
Falsification as a cause of the disproportion between the December and January birth rates is indicated by Takahashi (1964) analysing seasonality in Japan in the 1950s, Ruiu and Breschi (2017) in Italy in the first half of the 20th century, and Holzer in Poland in 1959–1960. Unnatural differences between December and January are also noticed by Anderson and Silver (1988) in the Soviet Union. The authors point to the flaws in the registration system as the cause, although it seems that it is also about the intentional behaviour of parents.
The second type of falsification was rarely described. We found publications documenting this phenomenon in Poland in the years 1927-1932 (Szulc 1931, 1936a, and Zaremba 1936) while Breschi, Ruiu, and Gonano (2019) quote confirming results from Italy. It is most likely that this phenomenon has been encountered by Herteliu et al. (2015), whose research results indicate higher daily birth rates in Romania on the 1st, 10th, 12th, 14th, 15th and 20th of each month. These regularities were not the focus of attention of the authors and were not interpreted, however, they might have distorted some of the results they presented.
In this paper we document the phenomenon of registering false birth dates in Poland in the 20th century. We try to estimate its scale and duration. To the best of our knowledge, this is the first comprehensive study of this phenomenon.

Data and methods

The study was conducted on the basis of birth dates of 44,187,032 Polish citizens born in the years 1900–1999, and registered in the General Electronic System for Registration of the Population (PESEL). PESEL contains only birth dates of those who were citizens and lived in the new, delimited after World War II, borders of the Polish state, and additionally, they lived to the time when PESEL was created, that is from 1974 to the early 90s. For this reason, the PESEL database does not contain all births (Fig. 1), but it is the only digital source of individual data for the entire analysed period, including wars (Statistics Poland, personal communication), allowing for a consistent study of the fluctuations in the daily number of births. In order to verify the exact extent to which the PESEL database reflects the examined regularities, the results obtained for 1932 were compared with the results calculated on the basis of the only found data on daily number of births published by the Central Statistical Office (Zaremba 1936). The comparisons were made using standard, statistical measures of central tendency, dispersion and association.

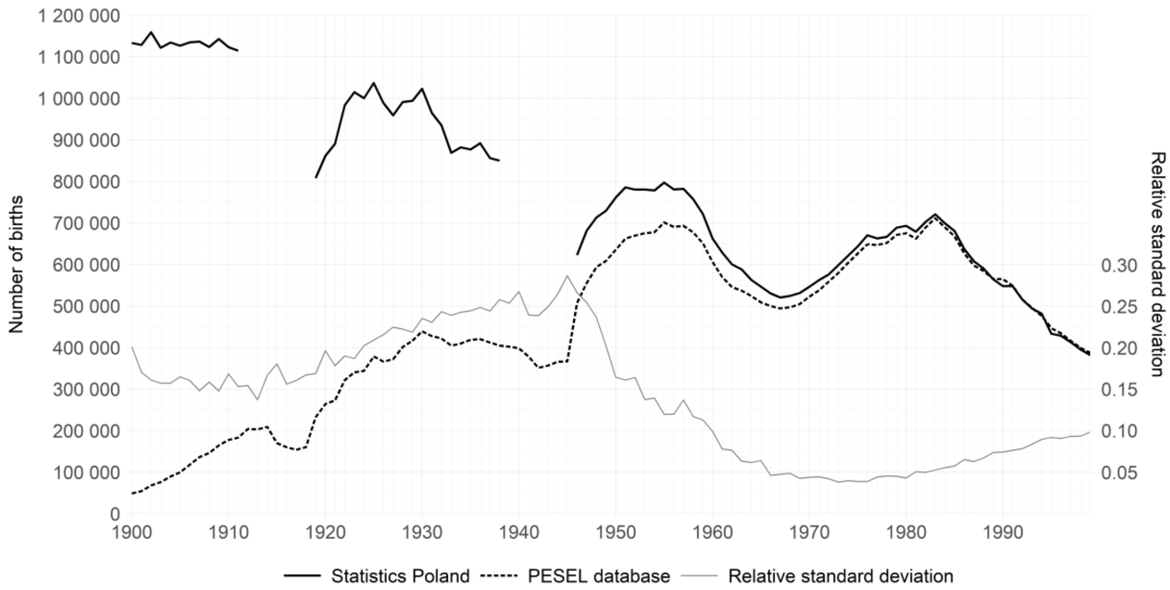

Fig. 1 Number of live births in the years 1900-1999 (obtained from Statistics Poland, Szulc 1936b, and Jelonek 1986) and its representation in PESEL database; Relative standard deviations of the daily number of births from the monthly average for the individual years

To analyse *P1* and *P2*, we transformed birth dates from the PESEL database into daily series of births, which we detrended by taking them to the daily average of the relevant year, and then presented graphically using heat maps. *P1* and *P2* overlapped to a large extent because January days with large deviations could probably have been the days to which the birth dates of December were shifted. Since it is impossible to determine to what extent this could have been the case and, on the other hand, not being certain that the December births were shifted only to the days of January with the largest deviations, *P1* and *P2* were estimated separately using different methods.

When estimating *P1*, we assumed that the observed over the entire period increase in the number of births in January in relation to December of the previous year should not exceed the average increase from the period when none of the above mentioned phenomena occurred or occurred to a small extent (Cypryjański 2019). To determine this period we used the relative standard deviation (RSD_y) calculated separately for each year. As shown in Figure 1, *RSD* had the lowest value (< 4%) between 1973 and 1976. Thus, the estimated number of shifted birth dates (p_{1y}) as a percentage of all births in a given year was calculated according to the formula:

$$p_{1y} = \frac{b_{Jan,y} - b_{Dec,y-1} * R}{(R+1) * b_y}$$

where: $b_{Jan,y}$ – the number of births in January of year *y*; $b_{Dec,y-1}$ – the number of births in December of a year *y–1*; b_y – the number of births in year *y*; *R* = 110.2% – average increase in the number of births in January in relation to December of the previous year in 1973–1976.

Estimating *P2*, we assumed, similarly to Breschi, Ruiu, and Gonano (2019), that if birth dates were not shifted to specific days, the births should be equally distributed between the days of the month. We interpreted equal distribution between the days of the month as the expected one. Taking the average daily number of births per month instead of per year as an expected value was intended to eliminate seasonal fluctuations. Therefore, in order to estimate the number of shifted birth dates (p_{2y}), we first compiled a list of days to which the dates of birth were shifted and then, for all days in the list, we calculated, separately for each year, the sum of the differences between the daily number of births and the daily average in the relevant month. The results were expressed as a percentage of the annual number of births.

Results

Below we present only selected results of the research. All the findings together with the discussion will be included in the final version of this paper.

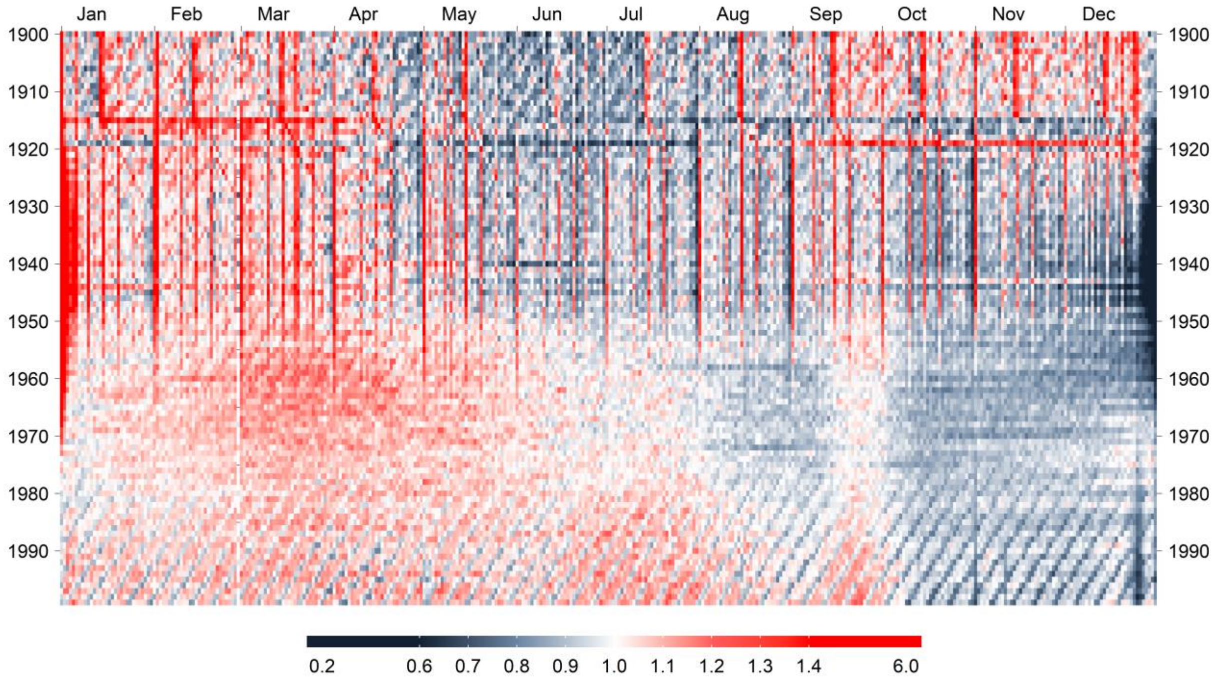

Fig. 2 Daily numbers of births in Poland in the years 1900 – 1999, as deviations from yearly average

The representation of daily deviations d_y by means of a heat map allows to observe a number of phenomena at the same time. Fig. 2 shows how over the period of 100 years the pattern of seasonality of births in Poland has changed (from the pattern with a clear well in summer months, through the pattern with peak in March and the local maximum in September to the current pattern with two peaks in July and September). We can also see in the form of horizontal lines how these seasonal patterns were disturbed by the wars taking place on Polish territory (World War I 1914 – 1918, Polish-Bolshevik war 1919 – 1921, and World War II 1939 – 1945).

Phenomenon of shifting birth dates from December to January of the following year (*P1)* is clearly visible in Fig. 2 in the form of two vertical stripes: intense red on the first days of January with simultaneous dark blue in the last days of December. This phenomenon is visible from the beginning of the studied period until the beginning of the 1970s, but it clearly intensifies after Poland regained its independence after World War I and reaches its peak during and just after the end of the World War II. In the years 1939 – 1948 deviations d_y on the 1st day of January ranged from 4.54 to 5.96, on January 2 from 3.40 to 3.99, and on three consecutive days from 1.30 to 1.99.

However, the most puzzling thing (regularity being a part of P2) is visible in the upper half of the heat map as vertical red lines indicating large positive deviations from the average on some days of the year. These lines appear with some regularity, especially on the 1st, 10th, 14th (to 1915), 15th and 20th day of each month, which can be observed by superimposing on each other the deviations from individual months ($\bar{d}_y$), as in Fig. 3a. Fig. 3a allows to observe one more, slightly weaker regularity in the form of a bright descending line from 1901 to 1931 (next regularity being a part of P2). It indicates positive deviations on days with dates in which the day number is the same as the year number. In turn, the heat map of deviations ($\bar{d}_y$) by days of the week (Fig. 3b) shows clear weekly patterns and their changes throughout the decades.

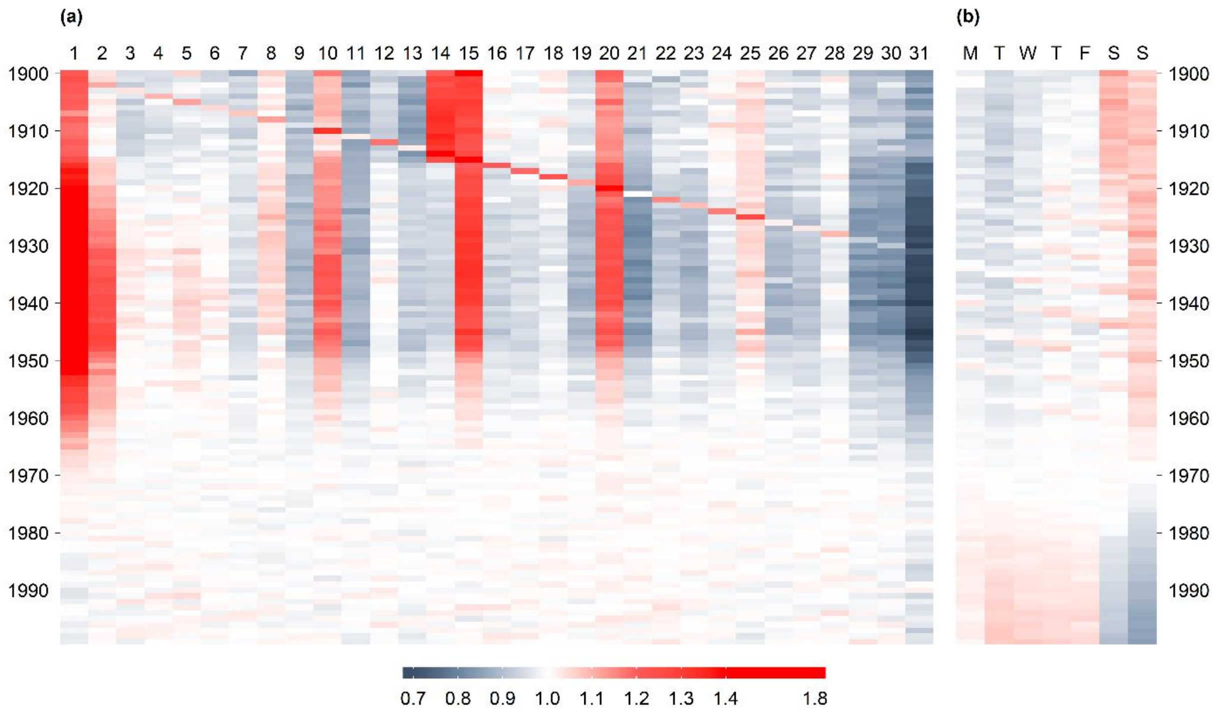

Fig. 3 Daily numbers of births in Poland in the years 1900 – 1999, as deviations from yearly average: (a) – averaged for the selected day of the month, (b) – averaged for the selected day of the week

References

Anderson, B.A. and Silver, B.D. (1988). The Effects of the Registration System on the Seasonality of Births: The Case of the Soviet Union. *Population Studies* 42(2): 303–320. DOI:10.1080/0032472031000143376

Björnsson, D.F. and Zoega, G. (2017). Seasonality of birth rates in agricultural Iceland. *Scandinavian Economic History Review* 65(3): 294–306. DOI: 10.1080/03585522.2017.1340333

Breschi, M., Ruiu, G., and Gonano, G. (2019). One day is not as good as another…at least not across the entire territory of Italy: an analysis of the differences in the daily distribution of births across Italian macro-regions, 1999–2016. *Genus* 75: 15. DOI:10.1186/s41118-019-0062-1

Cancho-Candela, R., Andrés-de Llano, J.M., and Ardura-Fernández, J. (2007). Decline and loss of birth seasonality in Spain: analysis of 33 421 731 births over 60 years. *Journal of Epidemiology & Community Health* 61(8): 713–718. DOI:10.1136/jech.2006.050211

Cassel, P-G. (2002). Changing seasonality of births in Sweden 1900–1999. In: Carling, J. (ed.). *Nordic demography: Trends and differentials. Scandinavian Population Studies, Volume 13*. Oslo: Unipub/Nordic Demographic Society: 97–110.

Cypryjański, J. (2019). Changes in seasonality of births in Poland in the years 1900–2009. *Demographic Research* 40: 1441–1454. DOI:10.4054/DemRes.2019.40.49

Dahlberg, J. and Andersson, G. (2018). Changing seasonal variation in births by sociodemographic factors: a population-based register study. *Human Reproduction Open* 2018(4): 1–8. DOI:10.1093/hropen/hoy015

Dahlberg, J. and Andersson, G. (2019). Fecundity and human birth seasonality in Sweden: a register-based study. Reproductive Health 16: 87. DOI:10.1186/s12978-019-0754-1

Doblhammer, G., Rodgers, J.L., and Rau, R. (2000). Seasonality of birth in nineteenth-and twentieth-century Austria. *Social Biology* 47(3–4): 201–217. DOI:10.1080/19485565.2000.9989019

Gini, C. and D'Addario, R. (1931). Intorno alla portata delle date di nascita differite per i nati alla fine dell'anno. *Annali di Statistica* serie VI, vol. XVIII.

Herteliu, C., Ileanu, B.V., Ausloos, M., and Rotundo, G. (2015). Effect of religious rules on time of conception in Romania from 1905 to 2001. *Human Reproduction* 30(9): 2202–2214. DOI:10.1093/humrep/dev129

Holzer, J. Z. (2003). *Demografia*. Warszawa: PWE.

Jelonek, A. (1986). Ruch naturalny ludności w Polsce w latach 1948–1984. Wrocław: Wydawnictwo Polskiej Akademii Nauk (Zeszyty Dokumentacji Geograficznej 6).

Livi, L. (1929). Sulle false dichiarazioni delle data di nascita per i nati alla fine dell'anno, e rettifica della distribuzione mensile delle nascite nel trennio 1923–1925. *Annalia di Statistica* serie VI, vol. III: 41–109.

Ruiu, G. and Breschi, M. (2017). Seasonality of livebirths and climatic factors in Italian regions (1863–1933). *Historical Life Course Studies* 5: 145–164.

Szulc, S. (1931). Fałszywe zgłaszanie dat urodzenia w Polsce. *Kwartalnik statystyczny*, 8(1), 40-44.

Szulc, S. (1936a). Dokładność rejestracji urodzeń i zgonów. *Statystyka Polski*, seria C, zeszyt 41, 133-158.

Szulc, S. (1936b). Ruch naturalny ludności w Polsce w latach 1895–1935. *Statystyka Polski* seria C, zeszyt 41: 1–132.

Takahashi, E. (1964). Seasonal Variation of Conception and Suicide. *Tohoku Journal of Experimental Medicine* 84: 215–27. DOI:10.1620/tjem.84.215

Zaremba, Z. (1936). Fałszywe zgłaszanie dat urodzenia w Polsce. *Statystyka Polski* seria C, zeszyt 41: 159–173.