CESifo WORKING PAPERS

10287
2023

February 2023

Persistence in UK Historical Data on Life Expectancy

Guglielmo Maria Caporale, Juan Infante, Marta del Rio, Luis A. Gil-Alana

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 10287*

Persistence in UK Historical Data on Life Expectancy

Abstract

This paper provides estimates of persistence in historical UK data on life expectancy applying fractional integration methods to both an annual series from 1842 to 2019 and a 5-year average from 1543 to 2019. The results indicate that the former exhibits an upward trend and is persistent but mean reverting; the same holds for the latter, though its degree of persistence is higher. Similar results are obtained for the logged values. On the whole, this evidence suggests that the effects of shocks to the series are transitory though persistent, which is useful information for policy makers whose task is to take appropriate measures to increase life expectancy.

JEL-Codes: C220, C400, D600.

Keywords: life expectancy, long memory, fractional integration.

*Guglielmo Maria Caporale**
Department of Economics and Finance
Brunel University London
United Kingdom – Uxbridge, UB8 3PH
Guglielmo-Maria.Caporale@brunel.ac.uk
https://orcid.org/0000-0002-0144-4135

Juan Infante
University Villanueva
Madrid / Spain
juan.infante@villanueva.edu

Marta del Rio
University Villanueva
Madrid / Spain
marta.delrio@villanueva.edu

Luis A. Gil-Alana
University of Navarra
Pamplona / Spain
alana@unav.es

*corresponding author

February 2023
Luis A. Gil-Alana gratefully acknowledges financial support from the MINEIC-AEI-FEDER ECO2017-85503-R project from ‘Ministerio de Economía, Industria y Competitividad’ (MINEIC), `Agencia Estatal de Investigación' (AEI) Spain and `Fondo Europeo de Desarrollo Regional' (FEDER). An internal Project from the Universidad Francisco de Vitoria is also acknowledged.

1. Introduction

Life expectancy is a useful indicator of a population's health (Roser et al., 2013). Historically, it was extremely low (about 30 years) throughout the globe prior to the Age of Enlightenment, it then started to climb in the early 1800s in the countries where the industrial revolution took place but not elsewhere, which resulted in a huge disparity in health conditions between affluent and poor countries. Since the beginning of the 20th century, the average lifespan in the world has increased sharply (at an annual rate of a quarter of a year over a 160-year span, according to Oeppen and Waupel, 2002 - see also Riley, 2001) as a result of better living conditions and medical care (Katz et al., 1983) – it reached 50 years in the period between 1990 and 2000 (World Bank Group 2018), the long-term enhancement of the welfare of entire populations being achieved for the first time (Riley, 2005). and is currently above 70 years. However, there are still considerable differences between countries in terms of life expectancy (Cutler et al., 2006): in 2019, the Central African Republic had the shortest one (53 years) and Japan the longest (three decades more). Differences also remain between individuals in the same country, although they have declined over time; these can be measured using the Gini coefficient in the same way as income inequality (Peltzman, 2009). Also, life expectancy is positively linked to GDP per person, a relationship known as the “Preston curve” (Preston, 1975), which can be estimated rather accurately using a Poisson common factor model (Li, 2013).

An interesting issue not much investigated in the literature is the persistence of life expectancy. One of the few studies on this topic is due to Caporale and Gil-Alana (2016), who analysed it in the case of sub-Saharan Africa using fractional integration methods and obtained rather mixed results. A similar approach was followed by Gil-Alana et al. (2017) for a wider set of 37 countries, again with mixed findings. This modelling framework has the advantage of being more general and flexible than the standard one based on the stationary I(0) versus nonstationary I(1) dichotomy since it allows the differencing parameter to take any real value, including fractional ones, as opposed to integers only; as a result, it encompasses a much wider range of stochastic processes, and provides evidence on whether or not those are mean-reverting, how fast the adjustment process is and how long-lived the effects of shocks are, the estimated fractional differencing parameter measuring persistence in this context. The present study also uses fractional integration methods as the two mentioned above; however, it applies it to UK historical data with a much longer time span.

The layout of the paper is the following: Section 2 describes the data and the methodology; Section 3 discusses the empirical findings; Section 5 offers some concluding remarks.

2. Data and Methodology

The series used for the analysis are UK life expectancy from 1842 to 2019 at an annual frequency and its 5-year average from 1543 to 2018. They have been constructed by 'OurWorldinData', which is a project of the Global Change Data Lab, a non-profit organisation based in the UK (Registered Charity Number 1186433), and are available from the website https://ourworldindata.org/life-expectancy. The annual figures are the average number of years a newborn would live if age-specific mortality rates in the current year were to stay the same throughout its life, whilst the 5-year average is calculated in each case as the arithmetic mean over 5 years.

FIGURES 1 AND 2 ABOUT HERE

Figures 1 and 2 display the two series of interest, namely the annual one (1842 – 2019) and the 5-year average (1543 – 2018) respectively. Despite the effects of some negative shocks (corresponding to World War I and II in Figure 1, and also to the Anglo-Spanish Wars in the XVI and XVIII centuries in Figure 2), an upward trend is clearly visible; this reflects improved health care, sanitation, immunizations, access to clean running water and better nutrition, as well as better treatment of chronic diseases such as cerebrovascular disease, chronic lower respiratory disease, and gastrointestinal cancers (stomach, liver and esophageal cancer).

The degree of persistence of these series is examined by estimating the following regression model:

$$y_t = \beta_0 + \beta_1 t + x_t, \qquad t = 1, 2, \ldots, \qquad (1)$$

where y_t stands for the series of interest, β_0 and β_1 denote the intercept and the coefficient on a linear trend respectively, and x_t is the error term, which is assumed to be integrated of order d:

$$(1 - B)^d x_t = u_t, \qquad t = 1, 2, \ldots, \qquad (2)$$

Using a Binomial expansion, the left-hand side of (2) can be expanded, with B being the backshift operator, such that $B^k x_t = x_{t-k}$, and u_t is I(0) (as in Granger and Joyeux, 1980 and Hosking, 1981), to obtain the following expression:

$$(1-B)^d = \sum_{j=0}^{d} \binom{d}{j} (-1)^j B^j = 1 - dB + \frac{d(d-1)}{2} - \ldots \qquad (3)$$

where the parameter d is a measure of persistence and sheds light on the properties of the process being modelled. Specifically, if d=0 this exhibits short memory, whilst d>0 implies long memory; if d<0.5, it is covariance stationary and mean reverting; if $0.5 \leq d < 1$ it is nonstationary but mean reversion still occurs; if $d \geq 1$, the process is explosive.

We employ the testing procedure proposed by Robinson (1994) that is based on the Lagrange Multiplier (LM) principle and includes a version of the Whittle function in the frequency domain, where the null is the following:

$$H_o : \; d \; = \; d_o \, , \tag{4}$$

Note that in equations (1) and (2) d can be any real number, including decimals from the nonstationary range ($d \geq 0.5$), but the limit distribution of the test statistc is standard N(0, 1) (for its functional form see Gil-Alana and Robinson, 1997).

3. Empirical Results

Tables 1 – 8 report the estimated values of d alongside their 95% confidence intervals for three different model specifications: i) without deterministic terms, ii) with an intercept only, and iii) with an intercept as well as a linear time trend. The coefficients in bold are those from the models selected on the basis of the statistical significane of the regressors.

TABLES 1 – 4 ABOUT HERE

Tables 1 – 4 display the results for the original data, while Tables 5 – 8 show the corresponding ones for the logged transformed series, in both cases under the alternative assumptions of white noise and autocorrelated residuals; the latter are modelled using the exponential spectral approach of Bloomfield (1973) that approximates well AR structures in the context of Robinson's (1994) tests (see Gil-Alana, 2004). Concerning the raw data, with white noise residuals the time trend is significant and positive for both series (Table 2); as for d, the estimated values are 0.65 and 0.73 respectively for the annual and 5-year average series (Table 1). Moreover, the confidence intervals do not include 1, which supports the hypothesis of mean reversion. When allowing for autocorrelated residuals (Table 3 and 4) the time trend is again statistically significant and positive for both series, and the estimates of d are now 0.82 and 1.00 (i.e., they are much higher than in the previous case) and mean reversion ($d < 1$) is only found in the case of the annual series.

TABLES 5 – 8 ABOUT HERE

Concerning the logged values (Tables 5-8), the time trend is significant for both series regardless of the error term specification, and the estimated values of d are now 0.61 (annual) and 0.57 (5-years) with white noise errors, and 0.77 and 0.78 with autocorrelation, whilst the hypothesis of mean reversion cannot be rejected in any single case, which implies that the effects of shocks will gradually die away.

4. Conclusions

This paper provides estimates of persistence in historical UK data on life expectancy applying fractional integration methods to both an annual series from 1842 to 2019 and a 5-year average from 1543 to 2019. The results indicate that the former exhibits an upward trend and is persistent but mean reverting; the same holds for the latter, though its degree of persistence is higher. Similar results are obtained for the logged values. On the whole, this evidence suggests that the effects of shocks to the series are transitory though persistent, which is useful information for policy makers whose task is to take appropriate measures to increase life expectancy.

Future work could investigate possible nonlinearities, for example by including in the model non-linear deterministic components such as Chebyshev polynomials in time (Cuestas and Gil-Alana, 2016), Fourier functions in time (Caporale et al., 2022) or neural networks approximations (Yaya et al., 2021), all within a fractional integration framework.

References

Bloomfield, P. (1973). An exponential model in the spectrum of a scalar time series. *Biometrika*, 60, 217-226.

Caporale, G. M. and Gil-Alana, L. A. (2016).Estimating Time Trends in Life Expectancy in Sub-Saharan Africa. *Population Review,* 55, 2, Project MUSE, doi:10.1353/prv.2016.0004.

Caporale, G.M., Gil-Alana, L.A., and Yaya, O.S. (2022). Modelling persistence and non-linearities in the US Treasury 10-year bond yields. *CESifo Working Paper No. 9554. Economics Bulletin, forthcoming.*

Cuestas, J.C., and Gil-Alana, L.A. (2016).A nonlinear approach with long range dependence based on Chebyshev polynomials. *Studies in Nonlinear Dynamics and Econometrics*, 20, 57-94.

Cutler, D., Deaton, A., and Lleras-Muney, A. (2006).The Determinants of Mortality.*Journal of Economic Perspectives*, 20(3), 97-120.

Gil-Alana, L.A. (2004).The use of the Bloomfield model as an approximation to ARMA process in the context of fractional integration. *Mathematical and Computer Modelling*, 39(4), 429-436.

Gil-Alana, L.A., and Robinson , P.M. (1997). Testing of unit roots and other nonstationary hypotheses in macroeconomic time series. *Journal of Econometrics*, 80, 241-268.

Gil-Alana, L.A., Cunado, J. and Gupta, R. (2017). Persistence, Mean-Reversion and Non-linearities in Infant Mortality Rates. *Soc Indic Res*,131, 393–405. https://doi.org/10.1007/s11205-016-1253-1

Granger, C.W.J., and Joyeux, R. (1980). An introduction to long memory time series and fractional differencing. *Journal of Time Series Analysis*, 1, 15-29.

Hosking, J.R.M. (1981). Fractional Differencing. *Biometrica*, 68, 168-176.

Katz, S., Branch, L.G., Branson, M.H., Papsidero, J.A., Beck, J.C., and Greer, D.S. (1983). Active Life Expectancy. *The New England Journal of Medicine,* 1983; 309:1218-1224. DOI: 10.1056/NEJM198311173092005

Li, J. (2013). A Poisson common factor model for projecting mortality and life expectancy jointly for females and males. *A Journal of Demography*, 67, Issue 1.

Oeppen, J., and Waupel, J.W. (2002). Broken Limits to Life Expectancy. *Science*, 296, 5570, 1029-1031. DOI 10.1126/science 1069675

Our world in data
https://ourworldindata.org/organization

Peltzman, S. (2009). Mortality Inequality. *Journal of Economic Perspectives*, 23(4), 175-19.

Preston, S. H (1975). The Changing Relation between Mortality and Level of Economic Development. *Population Studies*, 29 (2), 231–248. doi:10.2307/2173509.

Riley, J. C. (2001). Rising Life Expectancy: A Global History. *Cambridge University Press.*

Riley, J.C. (2005). Estimates of Regional and Global Life Expectancy, 1800–2001. *Population and Development Review*, 31, 3, 537–543.

Robinson, P.M. (1994). Efficient tests of nonstationary hypotheses. *Journal of the American Statistical Association*, 89, 1420-1437.

Roser, M., Ortiz-Ospina, E., and Ritchie, H. (2013). Life Expectancy. *Published online at OurWorldInData.org. Retrieved from*: 'https://ourworldindata.org/life-expectancy' [Online Resource]

Yaya, O., Ogbonna, A.E., Furuoka, F., and Gil-Alana, L.A. (2021). A New Unit Root Test for Unemployment Hysteresis Based on the Autoregressive Neural Network. *Oxford Bulletin of Economics and Statistics, forthcoming.*

World Bank Group (2018). https://data.worldbank.org/indicator/SP.DYN.LE00.IN?locations=ZG. Accessed 17.03.2018.

Figure 1

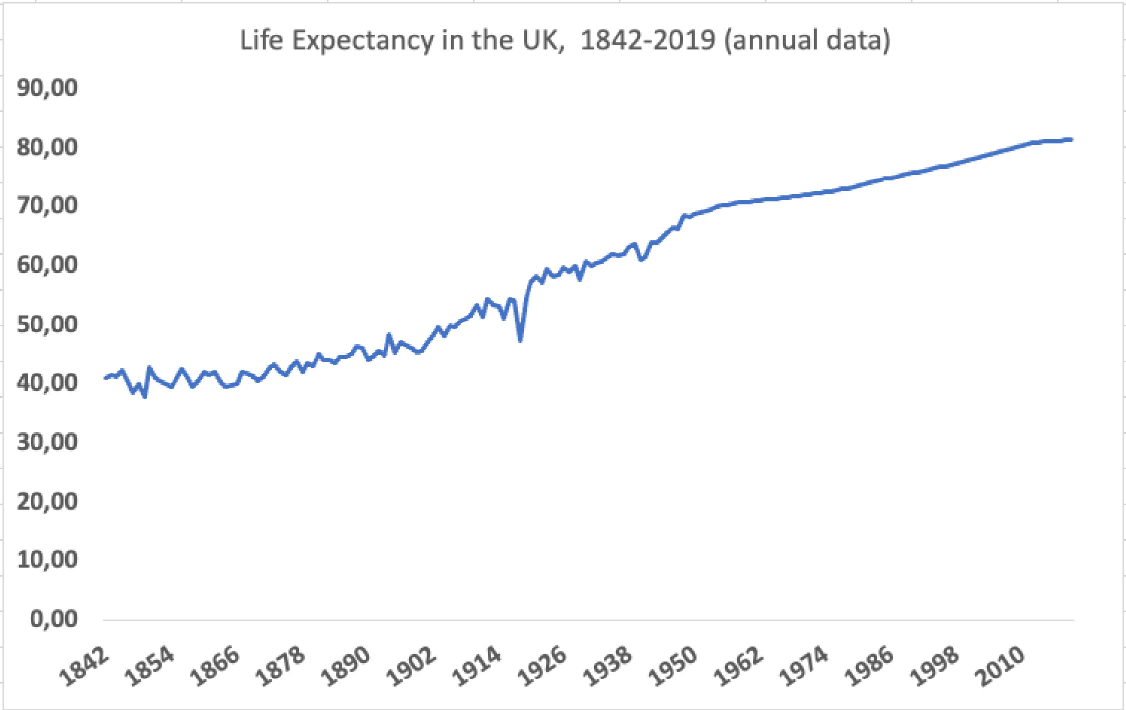

Figure 2

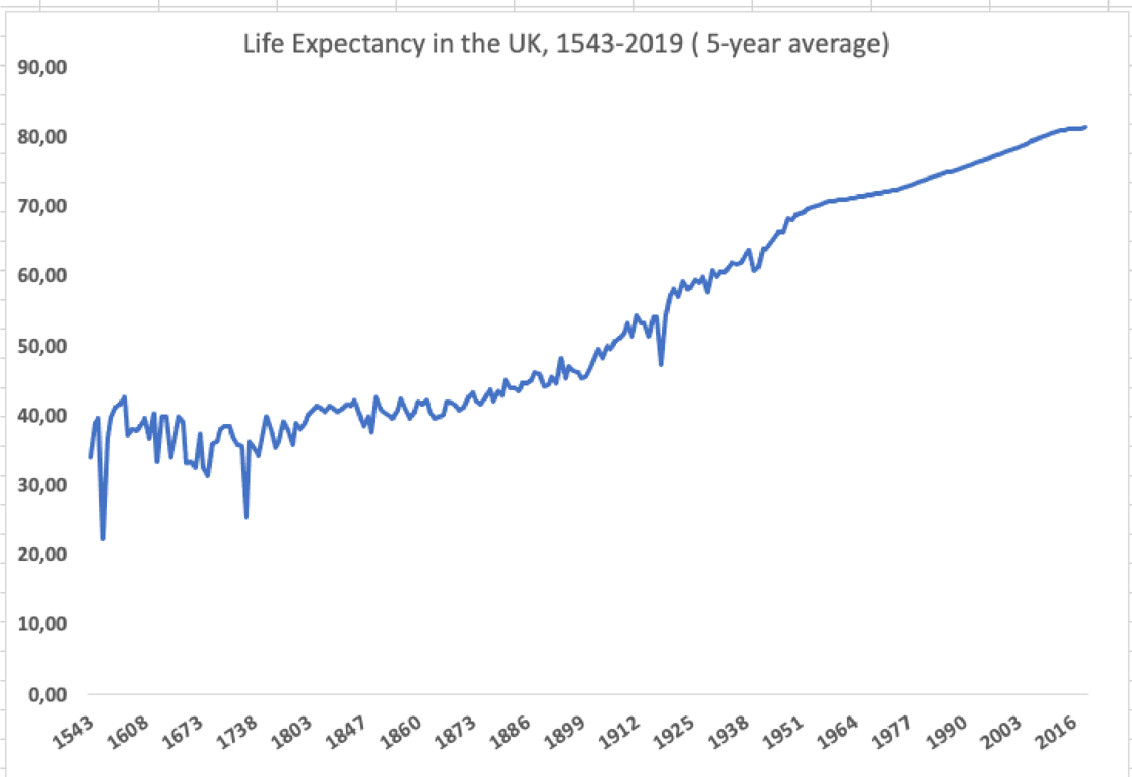

Table 1: Estimates of the differencing parameter. White noise errors

Series	No regressors	An intercept	A linear time trend
Annual	0.91 (0.81, 1.04)	0.75 (0.71, 0.80)	**0.65 (0.58, 0.73)**
5-years	0.86 (0.76, 1.00)	0.75 (0.69, 0.84)	**0.73 (0.66, 0.83)**

The values inside the parentheses are the 95% confidence intervals for the non-rejection values of d, while the coefficients from the selected models are in bold.

Table 2: Estimated coefficients from the selected model in Table 1

Series	No regressors	Intercept (t-value)	Time trend (t-value)
Annual	0.65 (0.58, 0.73)	**39.681 (37.88)**	**0.242 (13.54)**
5-years	0.73 (0.66, 0.83)	**33.243 (9.94)**	**0.448 (3.67)**

In bold in columns 3 and 4 the t-values of the corresponding coefficients.

Table 3: Estimates of the differencing parameter. Autocorrelated errors

Series	No regressors	An intercept	A linear time trend
Annual	0.85 (0.66, 1.08)	0.88 (0.82, 0.97)	**0.82 (0.72, 0.96)**
5-years	0.93 (0.76, 1.12)	1.00 (0.87, 1.15)	**1.00 (0.85, 1.17)**

The values inside the parentheses are the 95% confidence intervals for the non-rejection values of d, while the coefficients from the selected models are in bold.

Table 4: Estimated coefficients from the selected model in Table 3

Series	No regressors	Intercept (t-value)	Time trend (t-value)
Annual	0.82 (0.72, 0.96)	**40.480 (34.29)**	**0.236 (6.09)**
5-years	1.00 (0.85, 1.17)	**33.940 (9.96)**	**---**

In bold in columns 3 and 4 the t-values of the corresponding coefficients.

Table 5: Estimates of the differencing parameter. White noise errors

Series in logs	No regressors	An intercept	A linear time trend
Annual	0.97 (0.88, 1.10)	0.72 (0.68, 0.77)	**0.61 (0.55, 0.69)**
5-years	0.95 (0.82, 1.13)	0.62 (0.55, 0.70)	**0.57 (0.49, 0.67)**

The values inside the parentheses are the 95% confidence intervals for the non-rejection values of d, while the coefficients from the selected models are in bold.

Table 6: Estimated coefficients from the selected model in Table 5

Series in logs	No regressors	Intercept (t-value)	Time trend (t-value)
Annual	0.61 (0.55, 0.69)	**3.689 (179.98)**	**0.00422 (13.66)**
5-years	0.57 (0.49, 0.67)	**3.498 (43.65)**	**0.00809 (4.29)**

In bold in columns 3 and 4 the t-values of the corresponding coefficients.

Table 7: Estimates of the differencing parameter. Autocorrelated errors

Series in logs	No regressors	An intercept	A linear time trend
Annual	0.93 (0.78, 1.14)	0.84 (0.77, 0.93)	**0.77 (0.67, 0.90)**
5-years	0.89 (0.67, 1.15)	0.81 (0.69, 0.97)	**0.78 (0.64, 0.96)**

The values inside the parentheses are the 95% confidence intervals for the non-rejection values of d, while the coefficients from the selected models are in bold.

Table 8: Estimated coefficients from the selected model in Table 7

Series	No regressors	Intercept (t-value)	Time trend (t-value)
Annual	0.77 (0.67, 0.90)	**3.703 (155.23)**	**0.00411 (6.43)**
5-years	0.78 (0.64, 0.96)	**3.520 (37.59)**	**0.00859 (2.12)**

In bold in columns 3 and 4 the t-values of the corresponding coefficients.