Institution: City, University of London

Unit of Assessment: UoA 17 Business and Management

Title of case study: Detecting and correcting historic anomalies in national population and mortality data.

Period when the underpinning research was undertaken: 2014 – 2016

Details of staff conducting the underpinning research from the submitting unit:

Name(s):	Role(s) (e.g. job title):	Period(s) employed by submitting HEI:
Professor David Blake	Professor of Finance & Director of Pensions Institute	2004 – ongoing

Period when the claimed impact occurred: 2016 – ongoing

Is this case study continued from a case study submitted in 2014? N

1. Summary of the impact

Professor Blake and collaborators developed methods for identifying anomalies in national population data and (as a consequence) data on national mortality rates which were revealed in the wake of the 2011 UK Census. The work has changed practices in the UK Office for National Statistics (ONS), the UK Continuous Mortality investigation (CMI), the Human Mortality Database (HMD), as well as pension schemes and actuarial consultancies (e.g. USS, Milliman). Additionally, historical mortality tables data are being revised and future data will be collected and reported in a different way in order to avoid the identified anomalies being repeated. To illustrate the quantitative impact, the use of more accurate data, derived using the methods developed in the research, has contributed to the reduction in prices charged by insurers for the transfer of pension liabilities, saving UK pension funds up to £1bn over the period.

2. Underpinning research

The 2011 census revisions to England & Wales (E&W) population estimates by the ONS drew attention to the possibility that there were widespread errors in how population data were measured and reported in many countries, in particular the UK. Professor Blake in collaboration with Professor Andrew Cairns (Maxwell Institute) and Professor Kevin Down (Durham) investigated these anomalies, leading to the publication of the peer-reviewed research "Phantoms Never Die: Living with Unreliable Population Data" **[3.1].**

The research identified the cause behind the potential anomalies in E&W population data. The standard practice is to report the population size as the mid-year estimate, taken as the mid-point between the population estimated at the end of the given year and the population at the end of the previous year, with the former requiring an estimate to be made of births and deaths during the year, together with an assumption that births and deaths are **spread evenly throughout the year**. The research identified this practice as flawed due to the fact that in some years, the assumption of evenly distributed births is invalid. For example, in 1947, births were heavily concentrated in the early months of the year (as a result of service personnel returning home from the Second World War during 1946). This meant that the mid-year population for 1947 was underestimated. It would have been overestimated if the births had been concentrated in the second half of the year. Further, even if deaths are accurately measured (which is the case in countries like the UK), mortality rates (deaths divided by population) will be biased if mid-year population is mis-estimated. Since population estimates are almost never revised, the bias in mortality rates persists over time.

The analysis of national mortality trends is critically dependent on the quality of the population, exposures and deaths data that underpin death rates. Blake et al. developed methods for identifying anomalies and specific errors in national population and mortality data, using data from E&W by way of example. First, they proposed a set of graphical diagnostics that help to pinpoint anomalies. Second, they developed a Bayesian model that allows the researchers to quantify objectively the size of any anomalies. Finally, they introduced two-dimensional graphical

diagnostics and modelling techniques enabling the detection of small systematic errors in exposures and deaths of less than 1%. These are shown to significantly improve the researchers' ability to identify and quantify anomalies **[3.1].**

Apart from analysis of historical data, these methodologies can also be used as an early warning system for the detection of emerging anomalies in data. Analysis using these methodologies has shown that errors remain in ONS population data. For example, in the case of E&W, errors of more than 9% in the estimated size of some birth cohorts can be attributed to an uneven pattern of births. Blake et al. developed the cohort–births–deaths exposures methodology which can be used to explain many of the bigger errors. Collectively, these errors can make substantial differences, particularly in respect of cohorts that are still sufficiently large to have a significant financial effect on pension and annuity providers. An example would be the 1947 birth cohort, for which national mortality rates will be biased upwards.

The same sources of errors—with possible variants—will apply to other countries. Some countries are similar to E&W in that they derive their population data from periodic (typically decadal) censuses. An example of the practice is the data collection in the US, which shares similar characteristics and reveals similar anomalies to those for E&W.

Finally, Blake et al. considered the effect of anomalies on mortality forecasts and annuity values and found significant effects for some cohorts.

The implication is that definitions of published data need to be read carefully and published population figures that are derived from these (possibly by other agencies) should be used with caution **[3.1].**

3. References to the research

3.1 Cairns A., **Blake D**., Dowd K., Kessler A. (2016) "Phantoms Never Die: Living with Unreliable Population Data"*, Journal of the Royal Statistical Society Series A (Statistics in Society),* 179 (4), pp. 975–1005.

Research was published in a journal that applies stringent peer-review process prior to accepting articles for publication and has been consistently rated highly by the Charted Association of Business Schools.

4. Details of the impact

The methodology developed in the course of this research – hereafter referred to as Cairns-Blake-Dowd-Kessler (CBDK) after the four authors of **[3.1]** – is being used by organisations in the UK, France and the United States. Examples include the UK Office for National Statistics (ONS), the UK Continuous Mortality Investigation (CMI) of the Institute and Faculty of Actuaries, the US Human Mortality Database (HMD), as well as pension schemes and actuarial consultancies, such as USS (UK) and Milliman (FR).

ONS: 2016 Methodological Review & Impact in Planning for the Census

Due to the UK's reliance on 10-year censuses to measure population, the ONS has to work on assumptions rather than accurate data between censuses, which can lead to unreliable results as suggested by the underpinning research. In 2016, the ONS acknowledged the issue and conducted a methodological review of official high-age population estimates, referring specifically to the research as being one of two drivers for the review **[5.1, page 4].** The review acknowledged errors such as discrepancies between 2011 Census record of year of birth and the probable "true" year of birth. The impact is evident in the recommendations which include that (a) findings should feed into planning for the 2021 Census and (b) alternative methods for allowing for mortality improvement be investigated **[5.1, page 64].**

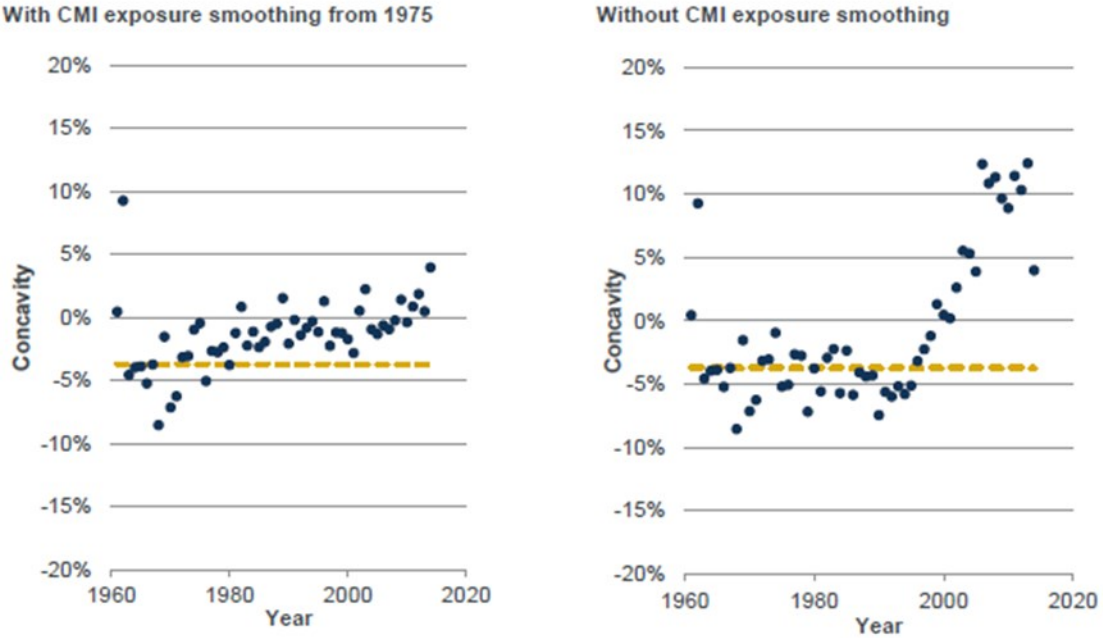
By challenging the established methodology, Blake et al. prompted the ONS to identify a range of weaknesses in the methodology and make recommendations impacting on the planning and execution of the 2021 Census. The conclusions of the evaluation also prompted calls for further investigation of the impact of birth cohorts, where births are not spread evenly over the year and further work investigating the over-count in the 2001 Census/mid-year estimates of the oldest old and in the 2002 to 2010 inter-censal years, both of which are in line with the issues the research has identified.

Continuous Mortality Investigation

The CMI, by the Institute and Faculty of Actuaries (IFoA), has a long history of providing

authoritative and independent mortality and sickness rate tables for UK life insurers and pension funds. It carries out research into mortality and morbidity experience and produces practical tools that are widely used by actuaries. In 2016, CMI's working papers tackled the issue of projecting mortality improvements which require use of a table of historical mortality rates **[5.2].** The results of the underpinning research highlighted to CMI the importance and potential impact of discrepancies in such tables. This led to a simplified method to adjust for anomalies with the CMI's flagship projection tool incorporating the choice of whether to use the adjusted or unadjusted population data, with use of the adjusted data being the default **[5.3].** Further impact was achieved in June 2017, when CMI adapted the approach to identify anomalies in their own data. Based on **[3.1]**, the CMI Mortality Projections Committee proposed an alternative approach to smoothing anomalies and have adopted this since the release of the CMI_2016 model. See Chart 2.6 from **[5.4]** which utilises the "convexity adjustment ratio", one of the CBDK diagnostic tools.

Chart 2.6: CBDK – concavity by cohort (1920 cohort) for ONS male data (based on KT method with trend, k = 5, m = 2, join age = 85, Lexis adjustments)



Note that the convexity adjustment ratio shown in Chart 2.6 is a property of the ONS input data only, and does not itself depend on the population estimates produced under alternative variants of the KT method. It simply serves as an informative benchmark to help interpret the concavity in the KT population estimates. For example, it can be seen that the concavity in the unsmoothed population estimates prior to the early 1990s (navy dots on the right-hand chart) is broadly consistent with the convexity adjustment ratio required to correct for the known distribution of births in the 1920 cohort (gold dashed line on the right-hand chart). In other words, the birth distribution anomalies for that cohort largely explain the observed concavity. After the early 1990s there is a further drift in the observed concavity, which may be due to other effects (Cairns et al suggest this may relate to the method used to roll forward the census estimates after 1991). The left-hand chart shows that, after applying the CMI's exposure smoothing, the observed concavity in both regions is reduced.

Human Mortality Database & Global Reach

The HMD, which was created to provide detailed mortality and population data to researchers, students, journalists, policy analysts, and others interested in the history of human longevity, has also modified its methodology in a way that is consistent with the research findings. The Methods Protocol for HMD (Last Revised: October 5, 2019) reflects the recommendations discussed in the underpinning research. According to the Associate Director of HMD: "*The systematic checks of official vital statistics (birth and death counts) and census data or population estimates which make the basis of all HMD calculations have highlighted a number of problems which present important challenges for the accurate measurement of mortality level and trends, ..[such as] reliability (under-registration of vital events or population; immortals and phantoms)…..We will present examples of such issues for HMD countries, explain which methods are currently implemented to take them into account, and open a discussion about those remaining unresolved*". **[5.5]**

According to the HMD records, its data have been used in 3490 journal articles, 744 books,

178 Dissertations and Thesis, 163 Reports and 577 Technical Papers – not including conference papers – a considerable amount of which are post 2018 when the methodology was updated. Thanks to the contribution of the HMD's popularity, the underpinning research and associated methodologies, have been exposed to a global audience and have informed hundreds of academic and industry projects. **[5.6]**

Universities Superannuation Scheme (USS) and Other UK Pension Scheme Liabilities

The CMI projections models, including its high-age methodology, are used extensively by life insurance and pensions actuaries in setting best-estimate mortality improvement assumptions, Specifically, the 2018 USS Actuarial Valuation used the CMI_2017 model which incorporates the CBDK smoothing assumptions **[5.7].** These assumptions impact on the valuation of £2trn of UK pension liabilities. The research has contributed to the reduction in prices charged by insurers for the transfer of pension liabilities, saving UK pension funds up to £1bn over the period **[5.8].** The research by Blake et al. has also been used in analysing socio-economic exposures and mortality data in USS **[5.9]**, where the data have had anomalies removed using the methodology in **[3.1].**

Milliman

Milliman (France) is a leading actuarial consultancy firm with more than 3,000 actuaries and consultants spread over 65 offices covering all geographic areas (Europe, US, Latin America, Asia Pacific, Africa, Middle East). Since 2017, they have used the underpinning research to confirm past conjectures about the 1919 cohort effect (caused by an uneven pattern of births following the First World War). New developments, carried out at Milliman, highlighted the universal nature of these apparently isolated (and, in the light of **[3.1],** now known to be false) cohort effects, which are unfortunately present in most period tables of the Human Mortality Database **[5.10].** This had two impacts. Primarily, it led to corrections to national mortality tables used by the insurance (and related securitisation) markets, with subsequent integration of these corrections coherently into their internal processes when assessing risks. It also led to HMD, not only making corrections to their tables as already discussed, but also looking for other anomalies, such as in fertility data **[5.11].** Milliman's Director of R&D is an ardent proponent of correcting existing mistakes and updating methodologies, as accurate population data is important for framing national pension policy (e.g., setting the state pension age), estimating the number of pensioners alive in a given future year, and the implications for the taxation needed to pay for state pensions. Milliman itself was prompted – by the research – to update its own mortality tables, several years before the HMD update, leading to a change in its Internal Model for calculating the regulatory capital that life companies need to hold under the Solvency II regime. **[5.12]**

5. Sources to corroborate the impact

5.1 Office for National Statistics: "Accuracy of Official High-age Population Estimates, in England and Wales: An evaluation" (12 December 2016); https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/ageing/methodologies/accuracyofofficialhighagepopulationestimatesinenglandandwalesanevaluation

5.2 CMI: Working Paper 91 (2016) "CMI Mortality Projections Model consultation: Technical paper", Mortality Projections Committee

5.3 CMI: "Mortality Projections Model CMI_2018 Briefing Note" (The Excel-based toolkit is available to CMI subscribers.)

5.4 CMI: "High Age Mortality Working Party Working Paper 100: A second report on high age mortality" (June 2017). https://www.actuaries.org.uk/system/files/field/document/CMI WP100 v01 2017-06-30 - HAMWP Second Report.pdf

5.5 Associate Director, Human Mortality Database: "Data Quality Issues and Adjustments in the Human Mortality Database", presented in Plenary Session III (Data Reliability Challenges and Timeliness) at the Longevity 12 Conference, Chicago (30 September 2016);

https://www.cass.city.ac.uk/__data/assets/pdf_file/0010/323929/L12-Programme-v9.pdf

5.6 HDM Publications Using the HMD as of November 2020;
https://www.mortality.org/Public/HMD-Publist.pdf

5.7 USS Actuarial Valuation - March 2018

5.8 Prudential Retirement – Statement on the effect of population data projections.

5.9 Cairns A., **Blake D**., Dowd K., Coughlan G., Jones, O., Rowney, J. "Old Professors Never Die, They Simply Fade Away: Analysis of Mortality in the Universities' Superannuation Scheme", *Pensions Institute Working Paper* (23 September 2018)

5.10 Boumezoued A., Devineau l., Hagstrom D, "Reliability Issues in the Construction of National Mortality Tables for the General Population – What you should know", *Milliman White Paper* (22 February 2017).
https://us.milliman.com/en/insight/reliability-issues-in-the-construction-of-national-mortality-tables-for-the-general-popula

5.11 Boumezoud A. "Improving HMD Mortality Estimates with HFD Fertility Data", presented at the Longevity 12 Conference, Chicago (29 September 2016);
https://www.cass.city.ac.uk/__data/assets/pdf_file/0020/334064/L12-18-BOUMEZOUED.pdf

5.12 Letter of support provided by R&D Director, Milliman, Paris. Accompanied by Fabrice Balland, Alexandre Boumezoued, Laurent Devineau, Marine Habart, Tom Popa. "Mortality Data Reliability in an Internal Model". 2018. ffhal-01719216f