Approximate factor models in mortality modeling

Richard D. MacMinn
Illinois State University and National Chengchi University
Ko-Lun Kung
National Chengchi University
Weiyu Kuo
National Chengchi University
Chenghsien Tsai*
National Chengchi University

April 30, 2014

Abstract

The classical factor analysis may no longer be consistent when number of variable is large. Unfortunately when estimating a mortality model, one usually finds the number of variable is large. The problem is particular severe when we consider multi-population modeling. In addition, mortality rates are correlated and heteroskedastic. This implies that the assumption of classic method is violated; therefore may not even suits the need of mortality rates analysis. We introduce a robust method called asymptotic principal component analysis (APCA) to mortality modeling. It allows correlation and heteroskedasticity, in both cross-sectional and time series dimension. It is compatible to current mortality model such as Lee and Carter (1992). Using mortality data from Human Mortality Database, our (preliminary) result show that it is an efficient and accurate method and does capture the cross-sectional heteroskedasticity and time series heteroskedasticity in many populations.

1 Introduction

Principal component analysis (PCA) is widely used in the mortality modeling. The application of PCA in mortality modeling goes back to Bell and Monsell (1991). The famous Lee and Carter (1992) can be seen as a mortality model with one component, as described in Girosi and King (2007). Renshaw and Haberman (2003b) extended Lee-Carter model to two sets of component. Hyndman et al. (2007) applied functional PCA to mortality model. Yang et al. (2010) developed

*Corresponding author. Address for correspondence: 64, Section 2, Chihnan Road, Taipei, Taiwan 11623, ROC; Tel: +886-2-2936-9647; Fax: +886-2-2939-3864; e-mail: ctsai@nccu.edu.tw.

a mortality model using PCA as an exploratory tool. Recently Hunt and Blake (2014) developed a general framework around principal component analysis, also use it as a tool to identify the appropriate function of age and period effect in a mortality model.

A concept that deeply relates to PCA is factor model. Factor models of random variables (here mortality rates) decompose variables into common factors and idiosyncratic errors. These factors often will be used in subsequent analysis or forecasting. The factors are associated with factor loading, which is variable's sensitivity toward factor. Both factor and factor loading are unobservable, so without other assumption the factor model alone cannot gives us any information. To uniquely (up to column sign changes) identify a factor model, one needs to assume some empirical restrictions to factor loading and factor. After restrictions are set, the method used in PCA can also be used to estimate the factor model. Common methods used include singular value decomposition (SVD) and spectral decomposition. As an example, the original paper of Lee and Carter (1992) used SVD to estimate the model, and restricts the sum of factor loading equals to 1 and sum of factor equals to 0 to uniquely identify the model.

Brouhns et al. (2002) and Renshaw and Haberman (2006) developed a Newton-Raphson type estimation procedure based on maximum likelihood estimation (MLE), with the same identification restriction in Lee-Carter model. Brouhns et al. (2002) relaxed the assumption of normality as well as cross-sectional heteroskedasticity.However, these estimation procedures did not consider the time series heteroskedasticity.

In practice, one often encounters the situation where the population of interest has larger number of variable (mortality rates) than number of observation. This is particularly true in a multi-population context, since the number of mortality rates increase proportionally with number of population. We introduce a better way to estimate the common factor under such circumstance. The new methodology, proposed by Connor and Korajczyk (1986), is called asymptotic principal component analysis (APCA). The main advantage of APCA is in its wide range of applications. It can be applied to wherever a factor model can be applied to. The method accounts for both cross-sectional and serial correlation, as well as heteroskedasticity. Particularly, We show that it is fully compatible with the traditional Lee-Carter model but render better estimation results in many cases without apparent negative effect.

In addition, the asymptotic property of APCA makes it very attractive when we are interested

in estimating a multi-population model, where the number of variable is large. Lin et al. (2013) pointed out that correlation among populations exists and significantly affect the pricing of mortality security, where APCA is robust against correlation. We plan to apply the APCA to modified version of Li and Lee (2005) to demonstrate the performance of approximate factor model.

We also consider the heteroskedastic version of APCA, namely heteroskedasticity factor analysis (HFA) proposed by Jones (2001). Jones (2001) relaxed the assumption of time series heteroskedasticity of APCA, which it is non-trivial when we are interested in mortality data with a long history. HFA accounts for time series heteroskedasticity in idiosyncratic error by using an iterative estimator based on MLE to update the cross-product matrix until convergence. The idea is analogous to the estimation of generalized least square method (GLM), which is also used in Renshaw and Haberman (2003a) for cross-sectional estimation.

Using mortality panel data of every available population from Human Mortality Database (HMD), we compare the efficiency of different estimators. The cross-sectional and time series correlation and heteroskedasticity in mortality rates is well documented in the literature, which motivates the use of robust PCA method such as APCA and HFA. Our preliminary result suggests that APCA and HFA are more accurate and more robust than traditional Lee-Carter model.

2 The Factor Model

Given a data matrix X with N variables, the factor model of X can be written as:

$$X - \bar{X} = \beta F + \epsilon, \tag{1}$$

where $\bar{X}$ is the vector of average of X_n, a $1 \times T$ vector of the nth variable $n = 1, 2, \ldots, N$. The factor $F = (F_1, F_2, \ldots, F_m)^\top$ is a $m \times T$ matrix, each row represent a latent random variable that is (linearly) uncorrelated to each other. The term β is called factor loading, which describe the sensitivity of each variable to these factors. Since β and F are both latent variable, one has to impose further restrictions to uniquely (up to the sign of variables) identify them. The standard restriction are (1) $E(F) = 0$. (2) $Cov(F) = E(FF^\top) = I_m$, where I_m is a $m \times m$ identity matrix. (3) $E(\epsilon) = 0$. (4) $Cov(\epsilon) = diag(\sigma_1^2, \sigma_2^2, \ldots, \sigma_N^2) = D$. And (5) F and ϵ are independent. Based on the

assumption, the factor loading β can estimated from covariance matrix via spectrum decomposition using the following relation:

$$\Omega = Cov(X) = \beta\beta^\top + D. \tag{2}$$

The factor estimate $\widehat{F}$ can obtained from regress $X - \bar{X}$ on $\widehat{\beta}$

$$\widehat{F} = (\widehat{\beta}^\top\widehat{\beta})^{-1}\widehat{\beta}^\top(X - \bar{X}). \tag{3}$$

3 APCA and HFA

The estimation of β and F is not trivial. In practice, often the situation dictates that the available data has less observation in year than the number of age we are interested in. In other word, it is quite often that $N > T$. The dimensionality hence leads to inefficient estimate of the factor model, especially when we estimate the common factor with multiple populations. The classical factor analysis does not allow serial or cross-sectional correlation, which make application to mortality rate less appealing. Chamberlain and Rothschild (1983) first developed the "approximate" factor model to allow for weakly serial or cross-sectional correlation. Connor and Korajczyk (1986) proposed the idea of asymptotic principal component analysis (APCA), which is designed specifically to estimate the factor model efficiently when $N >> T$, and at the same time allows for cross-sectional correlation and serial correlation.

Connor and Korajczyk (1986) and Connor and Korajczyk (1988) showed that the factor model can estimated from the $T \times T$ cross-product matrix $\widehat{\Omega_T}$, i.e.,

$$\widehat{\Omega_T} = \frac{1}{N-1}\sum_{n=1}^{N}(X_n - \bar{X}_n)(X_n - \bar{X}_n)^\top. \tag{4}$$

Specifically, Connor and Korajczyk (1986) showed that the first K eigenvectors of the $T \times T$ cross-product matrix would converge to the true pervasive factor given that residuals were not serially correlated and homoskedastic. The first K factor estimates, $\widehat{F}$, are equal to the first K eigenvector of $\widehat{\Omega_T}$.

The result of Connor and Korajczyk (1986) suggests that when we are interested in estimating a $N \times T$ factor model, if $T >> N$ then the traditional spectrum decomposition to the $N \times N$

covariance matrix, or equivalently, the singular value decomposition (SVD) used in Lee-Carter model is more efficient; if $N >> T$ then the APCA estimate is more efficient, which is often true in our application. Our empirical estimates at the end of section show that, in most of the application, APCA produce more accurate estimate than Lee-Carter model. Connor and Korajczyk (1988) improved estimation procedure by considering conditional heteroskedasticity.

The estimate algorithm of APCA is as follows:

1. Compute $\widehat{\Omega_T}$.
2. Obtain first K eigenvectors.
3. For each $n = 1, 2, \ldots, N$, regress first K eigenvector on $X_n - \bar{X}_n$, and compute residual variance $\widehat{\sigma_n}^2$.
4. Scale X as $X\widehat{D}^{-1/2}$, where $\widehat{D} = \text{diag}(\widehat{\sigma_1}^2, \widehat{\sigma_2}^2, \ldots, \widehat{\sigma_n}^2)$
5. Compute the new cross-product from $X\widehat{D}^{-1/2}$.
6. Repeat step 1 - 3 to obtain first K eigenvector as factor estimate $\widehat{F}$, and regression coefficient as $\widehat{\beta}$.

Jones (2001) proposes heteroskedasticity factor analysis (HFA) that further relaxes the homoskedasticity assumption on the idiosyncratic errors. In Connor and Korajczyk (1986) they assume that there exists an average residual variance and it is constant through time. In Jones (2001) the average residual variance in each period is allowed to be varying. Using the Strong Law of Larger Number, the off-diagonal elements of residual covariance matrix converge to 0 almost surely. The diagonal elements of residual covariance matrix converge to the he average idiosyncratic variances in period. The estimation of HFA is based on Jöreskog (1967). Jöreskog (1967) provide an iterative procedure based Maximum Likelihood Estimation (MLE) to estimate the factor. The consistency of MLE ensure convergence to the true Ω. The procedure is as follows:

1. Compute $\widehat{\Omega_T}$.
2. Set $\widehat{D_0}$ as the initial guess of diagonal residual covariance matrix D.
3. Obtain first K eigenvectors for $\widehat{D_0}^{-1/2}\widehat{\Omega_T}\widehat{D_0}^{-1/2}$.

4. Let V be be the eigenvector matrix, where the kth column is the eigenvector associated with kth eigenvalue. Let Λ be a diagonal matrix of eigenvalue, with descending order. Compute $\widehat{F}$ by $\widehat{D_0}^{-1/2} V(\Lambda - I)^{1/2}$

5. Update the estimate of D by $\widehat{D} = \widehat{\Omega_T} - \widehat{F}^\top \widehat{F}$.

6. Iterate step 3 - 5 until convergence criterion is achieved.

7. (Optional) To compare with APCA estimate of factor, the factor estimate $\widehat{F}$ is orthonormalized to $\widehat{F}^N$. The first column of $\widehat{F}^N$, $\widehat{F}_1^N$ is the normalized first eigenvector. Every other eigenvectors is regressed on $\widehat{F}_1^N$ without intercept. The residual of the kth regression is the new kth orthonormalized factor $\widehat{F}_k^N$. This step yields the factor estimate $\widehat{F}^N$ such that $\widehat{F}^{N\top} \widehat{F}^N = I$

4 Preliminary Results

To conclude the section, Table 1 and 2 shows a comparison of estimation performance among Lee-Carter, APCA, and HFA with 1 factor and 2 factors, respectively. The data is obtained from Human Mortality Database. We gather mortality data of every male population from age 1 to age 100. The only exception is Belarus, because of the missing data problem, possibly due to WWI. We use the percentage of variance explained of the factor as a measure of estimation performance. Variance explained is equal to the ratio of first K eigenvalue to the sum of all eigenvalue. In Table 1 we set the number of factor K to 1 and Table 2 to $K = 2$. First column lists the code name of populations. Second to fourth column report the percentage of variance explained for SVD, APCA, and HFA, respectively. The last column lists the length of observation. Table 3 is for abbreviation used in Table 1 and 2.

[Insert Table 1 near here]

[Insert Table 2 near here]

[Insert Table 3 near here]

The result shows that APCA and HFA provide more accurate estimate in almost every population. Second, for those population which one or two factor does not explained the mortality well, APCA and HFA often provide better estimate, sometimes twice of as much boost to accuracy, for example in Israel (ISR) and Luxembourg (LUX). These populations are relatively smaller, hence cross-sectional variation is larger than the average. This suggests APCA is particularly suitable small population. Third, HFA tends to outperform the other two for populations with long historical data, for example in France (FRACNP, and FRATNP), UK (GBR_NIR, GBR_NP, GBR_SCO, GBRCENW, and GBRTENW), Italy (ITA), Norway (NOR), and Spain (ESP). These empirical evidences suggest APCA and HFA provide robust and accurate estimation to mortality model.

5 Future Research

We would like to apply our method to multiple populations. Since the number of variable increases when the number of population does, it might be good application to illustrate APCA and HFA's advantage when N is large. Our method has the potential to improve estimation of common factors among multiple populations.

Other plans include conducting the cross-sectional comparison in more detail. We plan to investigate the effect of cross-sectional and time heteroskadasticity on hedge effectiveness. Our second plan is to determine the number of factor via statistical method such as Bai and Ng (2002) and Onatski (2010). Also we would like to derive the MLE estimator under the assumption that death count is a heterogeneous Poisson distribution.

References

Bai, J., and S. Ng. 2002. Determining the number of factors in approximate factor models. *Econometrica* 70:191–221.

Bell, W., and B. Monsell. 1991. Using principal components in time series modelling and forecasting of age specific mortality rates. Paper presented at the American Statistical Association 1991 Proceedings of the Social Statistics Section.

Brouhns, N., M. Denuit, and J. K. Vermunt. 2002. A Poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance: Mathematics and Economics* 31:373–393.

Chamberlain, G., and M. Rothschild. 1983. Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets. *Econometrica* 51:1281–304.

Connor, G., and R. A. Korajczyk. 1986. Performance measurement with the arbitrage pricing theory: A new framework for analysis. *Journal of Financial Economics* 15:373–394.

Connor, G., and R. A. Korajczyk. 1988. Risk and return in an equilibrium APT: Application of a new test methodology. *Journal of Financial Economics* 21:255–289.

Girosi, F., and G. King. 2007. Understanding the Lee-Carter Mortality Forecasting Method. Working paper.

Hunt, A., and D. Blake. 2014. A general procedure for constructing mortality models. *North American Actuarial Journal* 18:116–138.

Hyndman, R. J., S. Ullah, et al. 2007. Robust forecasting of mortality and fertility rates: a functional data approach. *Computational Statistics & Data Analysis* 51:4942–4956.

Jones, C. S. 2001. Extracting factors from heteroskedastic asset returns. *Journal of Financial Economics* 62:293–325.

Jöreskog, K. G. 1967. Some contributions to maximum likelihood factor analysis. *Psychometrika* 32:443–482.

Lee, R. D., and L. R. Carter. 1992. Modeling and forecasting US mortality. *Journal of the American Statistical Association* 87:659–671.

Li, N., and R. Lee. 2005. Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method. *Demography* 42:575–594.

Lin, Y., S. Liu, and J. Yu. 2013. Pricing mortality securities with correlated mortality indexes. *Journal of Risk and Insurance* 80:921–948.

Onatski, A. 2010. Determining the number of factors from empirical distribution of eigenvalues. *The Review of Economics and Statistics* 92:1004–1016.

Renshaw, A., and S. Haberman. 2003a. Lee–Carter mortality forecasting: a parallel generalized linear modelling approach for England and Wales mortality projections. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 52:119–137.

Renshaw, A. E., and S. Haberman. 2003b. Lee–Carter mortality forecasting with age-specific enhancement. *Insurance: Mathematics and Economics* 33:255–272.

Renshaw, A. E., and S. Haberman. 2006. A cohort-based extension to the Lee–Carter model for mortality reduction factors. *Insurance: Mathematics and Economics* 38:556–570.

Yang, S. S., J. C. Yue, and H.-C. Huang. 2010. Modeling longevity risks using a principal component approach: A comparison with existing stochastic mortality models. *Insurance: Mathematics and Economics* 46:254–270.

Table 1: Percentage of variance explained by SVD, APCA, and HFA in the 1-factor model. The number in parentheses indicates the ranking of accuracy among three methods. 1 = best, 2 = second, 3 = worst. NA for population with missing data.

Population Code	LC	APCA	HFA	Span
AUS	89.18 (3)	90.17 (1)	90.02 (2)	1921-2009
AUT	80.93 (3)	89.39 (1)	82.37 (2)	1947-2010
BEL	NA (NA)	NA (NA)	NA (NA)	NA
BGR	71.01 (1)	67.98 (2)	66.72 (3)	1947-2010
BLR	70.80 (3)	84.63 (1)	73.40 (2)	1959-2010
CAN	93.99 (3)	94.00 (2)	94.38 (1)	1921-2009
CHE	90.09 (3)	95.80 (1)	91.58 (2)	1876-2011
CHL	48.56 (3)	67.87 (1)	52.00 (2)	1992-2005
CZE	74.45 (2)	77.70 (1)	73.64 (3)	1950-2011
DEUTE	80.58 (2)	86.53 (1)	80.56 (3)	1956-2011
DEUTNP	90.11 (3)	93.94 (1)	91.29 (2)	1990-2011
DEUTW	94.99 (3)	95.82 (1)	95.22 (2)	1956-2011
DNK	91.57 (3)	93.84 (1)	93.45 (2)	1835-2011
ESP	95.94 (2)	95.91 (3)	96.99 (1)	1908-2009
EST	36.70 (1)	36.05 (2)	31.38 (3)	1959-2011
FIN	82.48 (3)	88.34 (1)	88.11 (2)	1878-2009
FRACNP	95.31 (3)	95.41 (2)	97.19 (1)	1816-2010
FRATNP	93.32 (3)	93.76 (2)	97.26 (1)	1816-2010
GBR_NIR	76.08 (3)	77.70 (2)	93.16 (1)	1922-2011
GBR_NP	92.69 (3)	95.44 (2)	97.29 (1)	1922-2011
GBR_SCO	91.79 (3)	93.20 (2)	96.63 (1)	1855-2011
GBRCENW	96.60 (3)	97.05 (2)	98.25 (1)	1841-2011
GBRTENW	94.92 (3)	95.41 (2)	98.27 (1)	1841-2011
HUN	63.67 (2)	66.16 (1)	63.15 (3)	1950-2009
IRL	61.52 (2)	75.66 (1)	60.24 (3)	1950-2009
ISL	75.13 (3)	77.84 (1)	77.16 (2)	1838-2010
ISR	47.75 (3)	73.62 (1)	48.77 (2)	1983-2009
ITA	94.09 (3)	95.17 (2)	96.78 (1)	1872-2009
JPN	95.84 (3)	97.14 (1)	96.86 (2)	1947-2009
LTU	45.28 (3)	48.35 (1)	46.55 (2)	1959-2011
LUX	21.98 (2)	54.60 (1)	21.94 (3)	1960-2009
LVA	32.13 (3)	33.08 (2)	36.92 (1)	1959-2011
NLD	93.60 (3)	96.70 (1)	95.85 (2)	1850-2009
NOR	93.69 (2)	93.41 (3)	94.52 (1)	1846-2009
NZL_MA	28.85 (3)	32.68 (1)	30.15 (2)	1948-2008
NZL_NM	73.58 (3)	79.81 (1)	76.80 (2)	1901-2008
NZL_NP	64.97 (3)	81.98 (1)	65.56 (2)	1948-2008
POL	76.66 (3)	80.47 (1)	79.48 (2)	1958-2009
PRT	87.02 (3)	89.47 (1)	87.65 (2)	1940-2009
RUS	75.43 (3)	84.66 (1)	84.17 (2)	1959-2010
SVK	53.26 (3)	58.61 (1)	54.59 (2)	1950-2009
SVN	31.85 (3)	74.06 (1)	38.98 (2)	1983-2009
SWE	92.35 (3)	95.54 (1)	94.96 (2)	1751-2011
TWN	81.93 (3)	87.18 (1)	84.83 (2)	1970-2010
UKR	81.40 (3)	90.09 (1)	84.61 (2)	1959-2009
USA	94.27 (3)	96.18 (1)	95.54 (2)	1933-2010
Average Rank	2.7	1.38	1.91	

Table 2: Percentage of variance explained by SVD, APCA, and HFA in the 1-factor model. The number in parentheses indicates the ranking of accuracy among three methods. 1 = best, 2 = second, 3 = worst. NA for population with missing data.

Population Code	LC	APCA	HFA	Span
AUS	94.23 (3)	97.61 (1)	95.89 (2)	1921-2009
AUT	85.18 (3)	91.58 (1)	85.23 (2)	1947-2010
BEL	NA (NA)	NA (NA)	NA (NA)	NA
BGR	85.18 (2)	83.68 (3)	88.90 (1)	1947-2010
BLR	79.33 (3)	87.87 (1)	80.60 (2)	1959-2010
CAN	96.74 (3)	97.04 (1)	96.80 (2)	1921-2009
CHE	91.98 (3)	96.99 (1)	92.67 (2)	1876-2011
CHL	66.07 (3)	76.21 (1)	66.90 (2)	1992-2005
CZE	81.00 (2)	87.87 (1)	80.55 (3)	1950-2011
DEUTE	86.63 (2)	90.26 (1)	85.93 (3)	1956-2011
DEUTNP	92.42 (3)	95.99 (1)	93.25 (2)	1990-2011
DEUTW	96.57 (3)	97.23 (1)	96.62 (2)	1956-2011
DNK	93.65 (3)	94.82 (2)	95.29 (1)	1835-2011
ESP	97.87 (3)	98.04 (1)	98.02 (2)	1908-2009
EST	46.26 (2)	56.02 (1)	45.32 (3)	1959-2011
FIN	86.10 (3)	94.38 (1)	89.60 (2)	1878-2009
FRACNP	97.43 (3)	97.80 (2)	98.58 (1)	1816-2010
FRATNP	97.10 (3)	98.06 (2)	98.65 (1)	1816-2010
GBR_NIR	80.74 (3)	87.67 (2)	94.39 (1)	1922-2011
GBR_NP	97.73 (3)	98.24 (2)	98.26 (1)	1922-2011
GBR_SCO	94.52 (3)	96.50 (2)	97.60 (1)	1855-2011
GBRCENW	98.39 (2)	98.37 (3)	99.05 (1)	1841-2011
GBRTENW	97.83 (3)	98.24 (2)	99.02 (1)	1841-2011
HUN	79.97 (3)	86.25 (1)	81.02 (2)	1950-2009
IRL	68.37 (2)	81.47 (1)	67.88 (3)	1950-2009
ISL	77.73 (3)	80.33 (1)	80.20 (2)	1838-2010
ISR	54.44 (3)	76.49 (1)	56.64 (2)	1983-2009
ITA	97.05 (3)	97.96 (2)	98.33 (1)	1872-2009
JPN	98.14 (2)	98.88 (1)	98.09 (3)	1947-2009
LTU	60.04 (3)	69.23 (1)	60.39 (2)	1959-2011
LUX	29.10 (3)	56.69 (1)	29.49 (2)	1960-2009
LVA	56.66 (3)	75.73 (1)	62.99 (2)	1959-2011
NLD	95.06 (3)	96.86 (1)	96.74 (2)	1850-2009
NOR	95.65 (3)	96.76 (1)	96.34 (2)	1846-2009
NZL_MA	34.40 (3)	37.45 (1)	35.81 (2)	1948-2008
NZL_NM	78.17 (3)	88.21 (1)	81.63 (2)	1901-2008
NZL_NP	69.69 (3)	86.82 (1)	69.74 (2)	1948-2008
POL	88.17 (3)	90.29 (1)	89.49 (2)	1958-2009
PRT	91.71 (3)	93.16 (2)	93.86 (1)	1940-2009
RUS	89.16 (3)	92.13 (1)	91.21 (2)	1959-2010
SVK	64.68 (2)	70.00 (1)	64.57 (3)	1950-2009
SVN	43.44 (3)	76.23 (1)	49.49 (2)	1983-2009
SWE	94.21 (3)	97.35 (1)	95.98 (2)	1751-2011
TWN	88.90 (2)	90.55 (1)	88.82 (3)	1970-2010
UKR	88.13 (3)	91.90 (1)	89.88 (2)	1959-2009
USA	97.62 (3)	97.83 (2)	98.68 (1)	1933-2010
Average Rank	2.74	1.34	1.91	

Table 3: Code table for populations in Human Mortality Database.

Population code	Population name
AUS	Australia
AUT	Austria
BEL	Belgium
BGR	Bulgaria
BLR	Belarus
CAN	Canada
CHE	Switzerland
CHL	Chile
CZE	Czech Republic
DEUTE	East Germany
DEUTNP	Germany
DEUTW	West Germany
DNK	Denmark
ESP	Spain
EST	Estonia
FIN	Finland
FRACNP	France, Civilian Population
FRATNP	France, Total Population
GBR_NIR	Northern Ireland
GBR_NP	United Kingdom
GBR_SCO	Scotland
GBRCENW	England and Wales, Civilian National Population
GBRTENW	England and Wales, Total Population
HUN	Hungary
IRL	Ireland
ISL	Iceland
ISR	Israel, Total Population
ITA	Italy
JPN	Japan
LTU	Lithuania
LUX	Luxembourg
LVA	Latvia
NLD	Netherlands
NOR	Norway
NZL_MA	New Zealand – Maori
NZL_NM	New Zealand – Non-Maori
NZL_NP	New Zealand
POL	Poland
PRT	Portugal
RUS	Russia
SVK	Slovakia
SVN	Slovenia
SWE	Sweden
TWN	Taiwan
UKR	Ukraine
USA	The United States of America