“Modeling multi-population life expectancy: a cointegration approach”

AUTHORS	A. Ntamjokouen S. Haberman G. Consigli

ARTICLE INFO	A. Ntamjokouen, S. Haberman and G. Consigli (2014). Modeling multi-population life expectancy: a cointegration approach. *Insurance Markets and Companies*, *5*(2)
RELEASED ON	Tuesday, 16 December 2014
JOURNAL	"Insurance Markets and Companies"
FOUNDER	LLC “Consulting Publishing Company “Business Perspectives”

NUMBER OF REFERENCES	NUMBER OF FIGURES	NUMBER OF TABLES
0	0	0

A. Ntamjokouen (Italy), S. Haberman (United Kingdom), G. Consigli (Italy)

Modeling multi-population life expectancy: a cointegration approach

Abstract

The continuous improvements in mortality rates and life expectancy of the last century have been given a great deal of attention by academics, life insurers, financial engineers, and pension planners, particularly in developed countries. Mortality-linked securities such as longevity bonds (EIB & BNP as well as the Swiss Re bond), survivor swaps, and mortality forward (q-forward) have appeared recently in the industry to help operators hedge such risks. A classic survivor bond has been proposed in the literature with coupon payment linked to the life time of the last survivor in an insurance reference portfolio. It appears therefore to be crucial to improve the accuracy of future life expectancy forecasts. In this paper, the authors investigate time-varying dependency associated with common trends that drive regional life expectancy within Canada. The aim is to compare three major models that have recently appeared in the literature, the autoregressive integrated moving average (ARIMA), the vector autoregressive model (VAR) and the vector error correction model (VECM), to analyze the common factors that have determined a progressive shift of life expectancy in specific Canadian regions. Results show that VECM performs better than VAR and ARIMA in terms of backtesting and its ability to capture the dynamics of common life expectancy. Findings from these analyses are useful for local insurers and demographers in their goal to project life expectancy improvements and also to forecast future trends.

Keywords: life expectancy at birth, VECM, VAR, ARIMA, confidence interval.

Introduction

The improvements in life expectancy and mortality rates have been investigated in many studies across the 20th century, especially in developed countries [(see Lee and Carter, 1992), (Russolillo and Haberman, 2005), (Tuljapurkar, 2007) and (Oeppen and Vaupel, 2002)]. These progressions in human life have been induced by higher life quality, typically associated with structural improvements of medical healthcare systems (Shaw et al., 2005), social advancements and economic development (Chen and Ching, 2000). Due to increasing levels of life expectancy, longevity risk is borne by insurance companies, pension funds, and social security. Furthermore, young active taxpayers are facing issues about their future pension payouts after retirement. Researchers are working to improve the accuracy of life expectancy computations with the goal of reducing the incidence of policy payouts as pension liability amounts are increasing. Two main methods of quantifying future life expectancy have been identified in the literature: the biological techniques and extrapolative methods (Whiteford, 2006). As for the first group, calculations are based on medical scenarios (Oeppen and Vaupel, 2002) to project life expectancy. However, these results are in underestimation of future life expectancy. Olshansky et al. (2005) explained that diseases such as obesity slow down human longevity especially in developed countries, and that this is one of the reasons for the underestimation of future life expectancy.

The main example of the extrapolative method is the Lee Carter model (1992), that forecasts life expectancy and mortality rates. It is based on the extrapolation of past mortality trends. This method has been adopted by the US Social Security. Another extrapolative method (Whitehouse, 2007) is based on three steps to generate a 50 year forecast. (Rusolillo and Haberman, 2005) improved the life expectancy forecasts by using ARIMA (see also (Torri, 2011)) which presents better results over the Lee Carter model. Several other papers, including (De Beer and Alders, 1999), (Keilman et al. 2001), (Maarten, 2007), (Booth, 2006), (Alho and Spencer, 2005) and (Denton, 2005) and (Torri, 2012) have explored this approach. In addition (Adekola, 2002) used a generalized model to explain life expectancy while (Raftery et al., 2013) projected life expectancy at birth for all the countries in the world using the Bayesian probabilistic model. Other forecasting approaches, including (Andreev and Vaupel, 2006) and (Lee, 2006) have explored methods based on the hypothesis of a non-stochastic component of the life expectancy variable.

In an international context, (Torri, 2011) used the cointegration approach methodology to examine future life expectancy in several countries. The results showed that the VAR model increases the accuracy of life expectancy predictions over ARIMA and VECM models for four countries, France, Italy, Norway and Sweden. Other research led by (Babel, 2007) has investigated life improvements in Australia, Europe, Japan and North America. In this paper, we implement the cointegration analysis, which takes into account the long run historical relationships across groups. This method examines the potential dependency between regions within a country with the aim of extrapolating their future life expectancies. We examine male life expectancy, but similar analysis can be conducted on females groups. We apply the cointegration method, which has proved to be successful in modeling time series (see commitee Nobel prize, 2003), to life expectancy

data from six heavily populated Canadian provinces by taking into account their correlation structure. In the literature some models, including VAR (Torri, 2011) and ARIMA have shown improvements in explaining the dynamics of life expectancy. The purpose of this paper is to explore how VECM performs on data. In addition, the use of the VAR model is related to previous literature, which has shown strong performance in explaining multiple time series. We will not a priori eliminate one model, but compare different models in order to show which is suitable to explain life expectancy time series among the Canadian provinces.

Life expectancy data were provided by the Canadian Human Mortality Database (CHMD) through the website www.bdlc.umontreal.ca/chmd, which is managed by the Department of Demography of the Université de Montréal in collaboration with the Max Planck Institute for Demographic Research and the Department of Demography at the University of California in Berkeley (CHMD). This database was created to provide information on human longevity in Canada to researchers, students, journalists and policy makers. It supplies the data used here, which has a frequency of 1 year spanning of 1921-2009, and covers the six Canadian provinces of Nova Scotia, New Brunswick, Quebec, Ontario, Alberta, and British Columbia. It also provides detailed information regarding births, population size, exposure-to-risk, death rates, and life expectancy at birth.

Life expectancy (Figure 1) shows an increasing trend for all of the provinces analyzed from 1921 to 2009. The historical pattern of life expectancy can be subdivided into two main periods. From 1921 to 1960 we observe a divergence in provincial life expectancy. However, after 1960, we can clearly observe a convergence as the six canadian provinces show common trends, as seen in Figure 1. British Columbia shows the highest life expectancy level for the sample period of 1960-2009, followed by Ontario. Quebec recovers from its low level during the period of 1921-1960 to become one of the provinces where residents live longest. New Brunswick, in contrast, is the province where people have lower life expectancy. We observe the common trends across the provinces to increase over the years. Accordingly, it can be deduced that life expectancy is converging to the same level in all parts of Canada. In the rest of this paper, we present the features of the three models (ARIMA, VAR and VECM) used in the analysis. The methodology of cointegration applied here includes several steps:

- the computation of the optimal value of lag of the vector autoregressive model;
- the Johansen cointegration test, which estimates the dynamic relationship among the regional life expectancies;
- the estimation of the VAR and VECM models and the forecasting of the derived model;
- we also present the backtesting output from the different models, and finally we generate the values of future life expectancy 50 years ahead, using VECM.

The procedure for VECM also involves determining the order of integration for each of the six life expectancy data sets by using the Augmented Dickey Fuller, Phillips-Perron, and KPSS tests. The steps of cointegration analysis are described by (Hamilton, 1994), (Juselius, 2007) and (Harris and Sollis, 2002). The empirical analysis will be done using R statistical software package developed by (Pfaff, 2008).

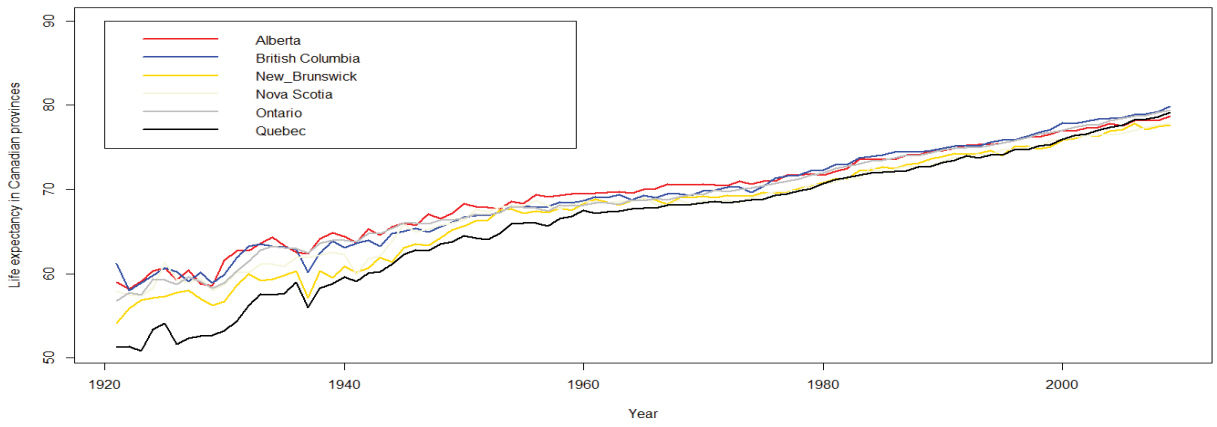

Fig. 1. Male Life expectancy in the six Canadian provinces

1. Explanation of various models

1.1. ARIMA model: description and fit. In the ARIMA model, life expectancy is modeled as a stochastic process. The methodology consists of three phases: identification, estimation, and diagnostics. These three steps are all described in Box and Jenkins (1976) and Hyndman and Athanasopoulos (2013) who explain that the process involves choosing

an appropriate ARIMA (p, d, q) when modeling a variable. The goal is to identify the correct model that will best fit the time series under study. Two options outlined by the literature may help to select the most appropriate model: selection of the model by the user, or the automatic ARIMA that will be used from here. In general, ARIMA is described as:

$$L_t = a_0 + a_1 L_t - 1 + \varepsilon_t, \quad (1)$$

where a_0 is the drift term, a_1 is a coefficient, L_{t-1} is the time series, and e_t is the error term distributed with $\varepsilon \sim (0, \sigma^2)$. The principal steps in selecting the best model as follows:

Identification of the model: consists of plotting data and identifying the pattern of the time series. As we can observe in Figure 1, life expectancy presents an increasing trend, with drift, for all the six provinces. The basic analysis also consists of differencing the data until they appear to be stationary. The unit root tests, including Augmented Dickey Fuller (ADF), Phillips-Perron (PP) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) are useful in determining the level of stationarity. Results obtained from the three unit root tests are used to determine the order of integration which corresponds to the value of the parameter d. The best model corresponds to the lowest Akaike Information criterion (AIC).

The first visualization of life expectancy from the six provinces (Figure 1) indicates that the variables are non-stationary. In order to confirm this, we compute the unit root test for life expectancy through the Augmented Dickey Fuller (see Dickey and Fuller, 1979), the Phillips Perron (see Phillips and Perron, 1988) and KPSS tests (Kwiatkowski et. al., 1992). The values for the KPSS are greater than the critical value of the test and KPSS also confirms the hypothesis of non stationarity. The order of integration analyzed through the ADF test (Table 10), PP test (Table 11), and KPSS test (see Table 12) shows that life expectancy time series are non-stationary at 5%. Under the criterion of constant, all the *p*-values are greater than the critical values. However, the analysis under the criterion of trend and constant, shows significance only for New Brunswick, Nova Scotia, Ontario, and Quebec, where the p-values are greater than the critical values (see the ADF results from the Table 10). Under the PP test (see Table 11) only Quebec, Ontario and Nova Scotia are significant. Last, the analysis of Alberta, New Brunswick, Nova Scotia, Ontario and Quebec with the KPSS shows that life expectancy are non-stationary (see results in Table 12 and the relative critical values in Table 13). Furthermore, the *p*-values from the three models measured on first difference data from each life expectancy are less than the critical values. Overall, the three tests accept the null hypothesis that life expectancy for each province is integrated of order 1 under the constant criterion and the criterion of trend.

Estimation of the order of the model: after derivation of the order of stationarity, it is necessary to experiment with various combinations of p, d, and q where p is the number of autoregressive parameters d is the drift, and q is the moving average parameter. A Box-Cox transformation may also be necessary to stabilize the variance.

It is recommended at this stage to examine the autocorrelation (ACF), the partial autocorrelation (PACF), and the diagnostics of residuals graph to choose the appropriate model. (Hyndman and Athanasopoulos, 2013) developed an automated algorithm which consists of the inclusion of a constant. (Box and Jenkins, 1976) advised relying on the AIC (Akaike) and SIC (Schwarz criterion) to choose the best model.

Model validation checks the diagnostics of residuals from the chosen models by plotting and conducting a Portmanteau test of the residuals. The residuals diagnostics are investigated to see whether there is white noise. The procedure is completed by computing the forecasts through the choice of the best fitting model. The best numerical results of the ARIMA are described in Table 1. The Portmanteau test (see Table 2) indicates non autocorrelation of residuals with 4, 10, 15, or 20 lags for each of the provinces life expectancy. These results suggest that ARIMA appears to behave well with white noise disturbances.

Table 1. The best ARIMA models from the analysis of life expectancy

models	Alberta	Columbia	Brunwick	Scotia	Ontaio	Qubec
ARIMA (p,d,q)	(1,1,1)	(0,1,2)	(0,1,1)	(1,1,2)	(0,1,0)	(0,1,1)
ar1	0.44	-0.35	-	-0.83	-	-
(se)	(0.14)	(0.11)	-	(0.10)	-	-
ma1	-0.79	-0.46	-0.42	0.58	-	-0.34
(se)	(0.09)	(0.14)	(0.10)	(0.16)	(0.10)	
ma2	-	-	-	-	-0.38	-
(se)	-	-	-	-	(0.12)	-
drift	0.22	0.23	0.25	0.22	0.24	0.31
(se)	(0.02)	(0.01)	(0.04)	(0.04)	(0.045)	(0.04)

Table 2. The *p*-values of the Portmanteau test from ARIMA models over the period of 1921-2009

lags	Alberta	Columbia	Brunwick	Scotia	Ontaio	Qubec
4	0.83	0.57	0.63	0.23	0.19	0.91
10	0.55	0.54	0.39	0.092	0.55	0.91
15	0.67	0.52	0.57	0.11	0.67	0.26
20	0.83	0.67	0.76	0.83	0.83	0.35

1.2. The vector autoregressive model theory. To forecast and explain the historical pattern and forecast of each variable as a function of others in the system, the vector autoregressive model of order p is used. The optimal lag length of the variables in the VAR model is derived by choosing the lag of order p that minimizes the value of information criteria models such as Akaike (AIC), HQ (Hannan-Quinn), Schwarz (SC), and the Final Prediction criteria (FPE). When these information criteria choose different values of p, (Lutkepohl, 2005) recommends considering only the lag chosen by the SC criterion. The VAR (p) models behave well with white noise in forecasting whether the residuals are normally distributed and non-autocorrelated. We start by determining whether they are non stationary (see results from the previous sections). We then derive the optimal lag order of these variables.

1.2.1. Optimal lag length. We analyze the optimal lag length of the VAR model. The information criteria shows contradictory results: AIC and FPE indicate three optimal lags while HQ indicates a lag order of two and finally SC indicates a lag order of only one. Since they differ, following Lutkepohl (2005), preference will be given to SC. Consequently, the lag length is 1.

1.2.2. Estimation of the VAR model. The VAR model is derived in (2):

$$L_t = b_0 + b_1 L_{t-1} - b_2 L_{t-2} + ...b_p L_{t-p} + \varepsilon_t, \qquad (2)$$

where $L_t = (L_{1t}, L_{2t}, ..., L_{kt})$ for $k = 1,……, K$ time series, $(b_0…..b_i)$ are the coefficients and ε_t is the error term distributed with $\varepsilon \sim (0,\sigma^2)$.

The following equations describe the VAR (p) of each of the variables included in the model (A = ALBERTA; BC = BRITISH COLUMBIA; NB = NEW BRUNSWCIK; NS = NOVA SCOTIA; O = ONTARIO; Q = QUEBEC):

$$\begin{bmatrix} L_{A,t} \\ L_{BC,t} \\ L_{S,t} \\ L_{O,t} \\ L_{Q,t} \end{bmatrix} = \begin{bmatrix} 21.98 \\ 9.80 \\ 10.16 \\ -0.73 \\ 4.43 \\ -4 \end{bmatrix} + \begin{bmatrix} .54 & -.16. & .07 & -0.6 & .13 & -.12 \\ .53 & .13 & .22 & -0.1 & -.29 & -.36 \\ .58 & -.13 & .36 & .36 & -.40 & .02 \\ .16 & -.05 & .22 & .45 & .29 & -.07 \\ .27 & 05 & -.03 & -0.8 & .80 & -.09 \\ .49 & -.19 & .13 & -0.14 & 0.15 & 0.59 \end{bmatrix} \begin{bmatrix} L_{A,t-1} \\ L_{BC,t-1} \\ L_{NB,t-1} \\ L_{S,t-1} \\ k_{O,t-1} \\ k_{Q,t-1} \end{bmatrix} + \lambda \begin{bmatrix} .06 \\ .07 \\ .06 \\ -.005 \\ .03 \\ .02 \end{bmatrix} \qquad (3)$$

Diagnostic tests of residuals are computed for both Portmanteau and Normality. The results in Table 1 show remaining autocorrelation (p = 0.0009), but normality on the residuals, as the p-value is equal to 0.23. These results can be expected since we use only a few parameters. However, for the purpose of forecasting, it is better to use as few lags as possible. The autocorrelation(ACF) and partial autocorrelation functions (PACF) are performed on residuals as shown in (Figure 2) which shows that the residuals for life expectancy in Alberta are an appropriate fit and do not present autocorrelation.

Table 3. The diagnostics tests of residuals under the VAR model

Type of test	Specific name	*p*-values
Autocorrelation	Portmanteau (4 lags)	0.0009
Normality	Both	0.23
	Kurtosis	0.195
	Skewness	0.36

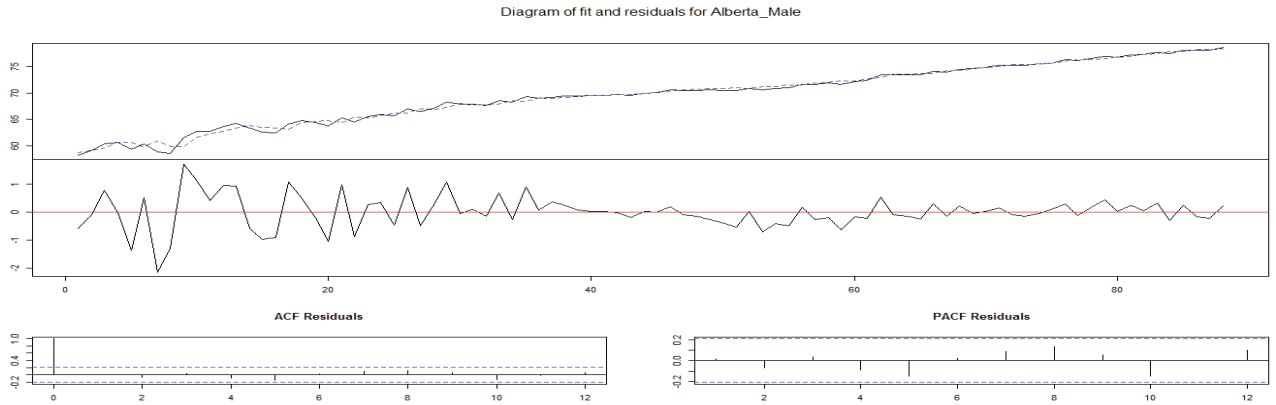

Fig. 2. Diagnostics of residuals with reference to Alberta

1.3. The vector error correction model. Once the test of unit root and the optimal lag are determined, the VECM is determined by conversion of the VAR (p). It can be described in two versions of this model: the short run and the long run version, where each variable in the vector system is explained by its own past values, lagged changes in other variables, and residuals. Each lagged difference of the L_{t-1} variable included must be stationary. The long version of VECM, which will be used here, is defined in (4) as follows:

$$\Delta L_t = \Gamma_1 \Delta L_{t-1} + \Gamma_2 \Delta L_{t-2} + ... + \Gamma_{p-1} \Delta L_{t-p+1} + \prod L_{t-p} + \\ + A_0 + e_t, \quad (4)$$

where $L_i = -(I - A_1 - ..... - Ai), i = 1, ..., (p-1)$ $\Pi = -(I - A_i, -... - A_p)$ is $N-$ dimensional time series, A_0 is the intercept term, and e_t is white noise.

The vector error correction model is used for forecasting and estimations, performed with the Johansen maximum likelihood methodology. It is used to determine the number of common trends (or cointegrated equations) derived from multiple data. The presence of cointegrated equations between variables is determined according to the three following hypotheses.

If $r = K$, the number of cointegrated variables, r, which is stationary, equals the rank (K) of Π, then the model will be estimated by using the standard statistical model.

If $r = 0$, this means that there are no cointegrated relationships between the variables. The variables are stationary if we take the differences of variables above.

If $0 < r < K$ there are two matrices, α and β, such that $\Gamma = \alpha\beta'$, and there will be r cointegrating relationships or $n - r$ common trends. The test of cointegration is reduced to the two following hypotheses:

The rank test is specified in the following form as in (5):

$$H_0 : rank(\prod) = r, \\ H_1 : rank(\prod) > r, \quad (5)$$

and the likelihood ratio statistic is described in (6) as:

$$LR(r) = -(T - p)\sum \ln(1 - \lambda_i), \quad (6)$$

where r represents the number of cointegrated relationships and A is the eigenvalue associated with the linear relationship.

The cointegration rank is determined in the trace test and the maximum eigenvalue test of (Johansen, 1988 and 1991). In addition, the test on the maximum eigenvalue test is specified as follows in (7):

$$H_0 : rank(\prod) = r \\ H_1 : rank(\prod) = m + 1, r = 0, 1, ..., n - 1. \quad (7)$$

The statistic value is written here in (7):

$$LR(r) = -(T - p)\sum \ln(1 - \lambda_{m+1}). \quad (8)$$

The eigenvalue statistic value tests the null hypothesis of m cointegrated relations against the alternative m+1. For example, the null hypothesis of five cointegrated relations is accepted against the alternative of six cointegrated relations.

1.3.1. Model fitting. The eigenvalue and trace test results from Johansen's procedure are reported in Tables 4 and 5. In the remaining sections of this paper, our computations will be given based on the trace test. Obviously, the same procedure can be accomplished with the eigenvalue test. But in order to save space we will illustrate the results obtained under the trace test. For any r, if the test value is less than the critical values then the corresponding r represents the number of cointegrated equations. For example, from $r = 0$ to $r = 2$ there are no cointegrated equations.

Table 4. The cointegrating relationship under eigen test

Cointegrating relationship	critical values	5%	1%
5	3.09	9.24	12.97
4	7.20	15.67	20.20
3	22.16	22.00	26.81
2	38.87	28.14	33.24
1	47.18	34.40	39.79
0	74.58	40.30	46.82

Table 5. The cointegrating relationship under trace test

Cointegrating relationship	critical values	5%	1%
5	3.09	9.24	12.97
4	10.29	19.96	24.60
3	32.45	34.91	41.07
2	71.31	53.12	60.16
1	118.49	76.07	84.45
0	193.08	102.14	111.01

The $r = 3$ test value equals 32.45 which is less than the critical value (34.91), therefore the number of cointegrated equations is three at a 5% significance level. We can say that the null hypothesis of three cointegrating relations is accepted against the alternative of two, while the null hypothesis of zero cointegrated relations is rejected. Consequently, according to these two tests, there are three cointegrated relations under the

trace test and four for the eigen test among the six groups of regional life expectancy data used in this study. The results of the fitted VECM are presented below as:

$$\begin{bmatrix}\Delta A\\ \Delta BC\\ \Delta NB\\ \Delta NS\\ \Delta ON\\ \Delta Q\end{bmatrix}=\begin{bmatrix}0.36 & 0.09 & 0.04 & -0.40 & 0.26 & -0.36\\ 0.28 & -0.08 & 0.04 & -0.14 & 0.28 & -0.13\\ 0.16 & -0.01 & -0.08 & -0.009 & 0.57 & -0.64\\ 0.36 & 0.19 & -0.34 & -0.20 & 0.53 & 0.03\\ 0.18 & 0.15 & 0.08 & -0.07 & -0.03 & -0.32\\ 0.08 & 0.41 & -0.10 & -0.20 & 0.34 & -0.48\end{bmatrix}\begin{bmatrix}\Delta A(-1)\\ \Delta BC(-1)\\ \Delta NB(-1)\\ \Delta NS(-1)\\ \Delta ON(-1)\\ \Delta Q(-1)\end{bmatrix}+$$
$$+\begin{bmatrix}-0.84 & 0.21 & -0.11 & 0.11 & 0.13 & 0.34\\ 0.21 & -0.43 & 0.14 & 0.06 & 0.30 & -0.22\\ 0.27 & 0.13 & -0.54 & 0.25 & -0.27 & 0.16\\ -0.27 & -0.23 & 0.43 & -0.46 & 0.53 & -0.05\\ 0.05 & 0.13 & -0.02 & -0.10 & -0.03 & -0.02\\ 0.23 & -0.32 & 0.21 & -0.11 & 0.41 & -0.34\end{bmatrix}\begin{bmatrix}A(-1)\\ BC(-1)\\ NB(-1)\\ NS(-1)\\ ON(-1)\\ Q(-1)\end{bmatrix}+\begin{bmatrix}12.10\\ -4.97\\ 1.12\\ 4.56\\ 0.10\\ -5.85\end{bmatrix}. \quad (9)$$

These equations measure the long run relationship between the six times series throughout the period of 1921 to 2009. Here $Z_{i,t}$ represents the stationary variable which quantifies the deviation from the equilibrium of the various life expectancies analyzed. Changes in provincial life expectancy are reflected in these three equations, which also involves change in trends of life expectancy.

The equation in matrix form for males representing the dynamic of life expectancy derived by the vector error correction model is given below. The equations explain the variations of the improvements in mortality by patterns observed from other provinces at the first lag level. The variation in Alberta is explained by the other provinces in their first lag (coefficients are -0.84, 0.21, -0.11 -0.11, 0.13 and 0.34) and also by the first difference in the mortality of each, as can be observed from the coefficients of the matrices (coefficients are 0.36, 0.09, 0.04, -0.40, 0.26 and -0.36). The same interpretation can be applied to the other provinces.

The three cointegrating relations with the criteria of the trace test are:

$$\begin{aligned}Z_{1t-1} &= A_{1t-1} - 12.21BC_{2t-1} + 13.74NB_{3t-1} - 6.08NS_{4t-1} + 12.97O_{5t-1} - 92.72 - 8.19Q_{6t-1}\\ Z_{2t-1} &= -12.43A_{1t-1} + BC_{2t-1} + 0.72NB_{3t-1} + 1.20NS_{4t-1} + 3.16O_{5t-1} + 163.05 - 4.25Q_{6t-1}.\\ Z_{3t-1} &= -0.17A_{1t-1} - 0.14BC_{2t-1} + NB_{3t-1} + 1.08NS_{4t-1} + 1.013O_{5t-1} + 0.07 - 0.61Q_{6t-1}\end{aligned} \quad (10)$$

Diagnostic tests of residuals are conducted for both Portmanteau and Normality tests. The results provided by Table 6 show remaining autocorrelation as the *p*-value is equal to 0.0018. However, they show evidence of normality on the residuals as *p*-value is equal to 0.0675. These results can be expected since we use only a few paramaters. Increasing the number of lags could improve the significance of autocorrelation test. However, for the purpose of forecasting, it is better to use as little lag as possible.

Table 6. The diagnostics tests of residuals of VECM

Type of test	Specific name	*p*-values
Autocorrelation	Portmanteau (4 lags)	0.0018
Normality	Both	0.0675
	Kurtosis	0.07
	Skewness	0.195

2. Forecasting procedure and backtesting of the various models

In this section, we fit data from six samples periods including 1921-2000, 1921-2001, 1921-2002, 1921-2003, 1921-2004, and 1921-2005 with the three models analyzed and forecast life expectancy for the remaining part of each sample up to 2009. In this backtesting phase, we compute the Mean Absolute Percentage Error (MAPE) of the three models in six different sample periods 2001-2009, 2002-2009, 2003-2009, 2004-2009, 2005-2009 and 2006-2009. The results are presented in Table 5 and show that VAR (0.31%, 0.40%, 0.26% and so on) and VECM (0.29%, 0.27%,0.24% and so on) are reliable in being a good fit for the data as the errors are low for each sample. The ARIMA model presents poor results with the highest error performance (35.73%, 35.92%, 36.28% and so on). This illustrates the fact that forecasts from VECM and VAR are much closer

to the historical data. In addition, we deduce that the VECM performed better than the VAR model in the quantification of residuals. This contrasts with the overall results obtained by Torri (2011). In the six regions as we can see (see Figures 3 to 8), the confidence intervals from the VECM performed better than the VAR and ARIMA models as we can see in Table 7. It allows one to account for more risk than other models. Consequently the VECM produces better results than the VAR model in terms of backtesting out-of-sample and quantification of future. The additional variables (the first difference of mortality index with Γ coefficients) included in the VECM provide improvements over the VAR model in terms of confidence interval as well as future life improvements since the VAR model only considers variables in terms of their levels. Accordingly, a new approach based on the VECM explains time varying long-run relationship dependence between the various life expectancies of the Canadian regions considered.

Table 7. The average MAPE for the ARIMA VAR and VECM models for the six provinces

Out-of-sample	VECM	VAR	ARIMA
h = 2001-2009	0.29%	0.31%	35.73%
h = 2002-2009	0.27%	0.40%	35.92%
h = 2003-2009	0.24%	0.26%	36.28%
h = 2004-2009	0.28%	0.44%	33.73%
h = 2005-2009	0.20%	0.23%	34.34%
h = 2006-2009	0.28%	0.37%	35.24%

Table 8. The confidence interval of the VAR, VECM and ARIMA models for the six provinces derived from predictions 50 years forecasts

Provinces	VECM	VAR	ARIMA
Alberta	(1.04-4.58)	(1.19-1.73)	(1.20-3.44)
British Columbia	(1.07-7.06)	(1.04-1.49)	1.34-2.32
New Brunswick	(1.05-6.52)	(1.18-2.20)	(1.36-5.65)
Nova Scotia	(1.11-6.73)	(1.27-2.09)	(1.32-6.21)
Ontario	(0.65-6.40)	(0.75-1.57)	(0.83-5.88)
Quebec	(1.08-6.33)	(1.24-2.64)	(1.30-6.07)

Table 8 reports the confidence intervals for each model presented in this work. Results from this analysis show that the VECM performs better than the other models. For example, Alberta's confidence interval length is 3.58, which is equal to six times that of VAR (0.54) and 1.5 times that of ARIMA (2.24). Observing, British Columbia province shows the greatest interval confidence length. However, the confidence interval associated with ARIMA is smaller for the provinces Nova Scotia, Ontario and Quebec. We observe overall that the VECM is better than the ARIMA and VAR models in capturing the increasing level of life expectancy as well as in fitting historical data.

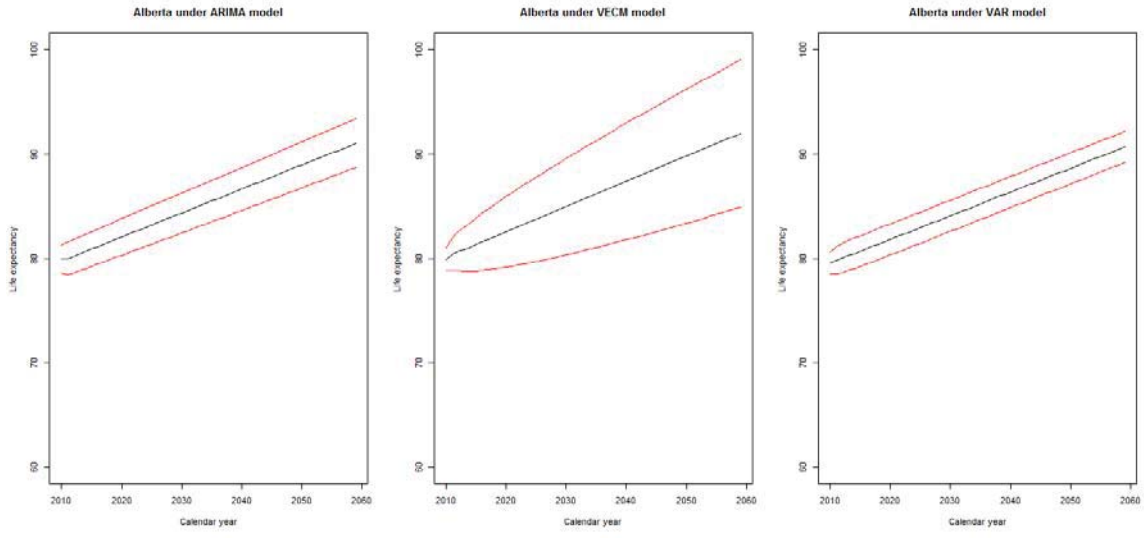

Fig. 3. Life expectancy of Alberta under the three models: the red lines represents the lower and upper forecasting, the black line represents the point forecast

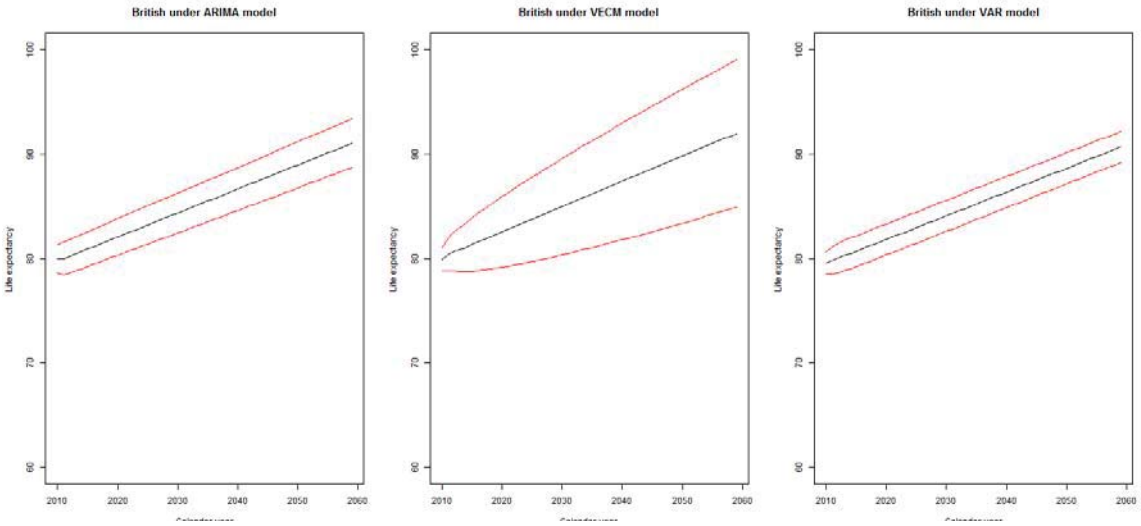

Fig. 4. Life expectancy of British Columbia under the three models: the red lines represent the lower and upper forecasting, the black line represents the point forecast

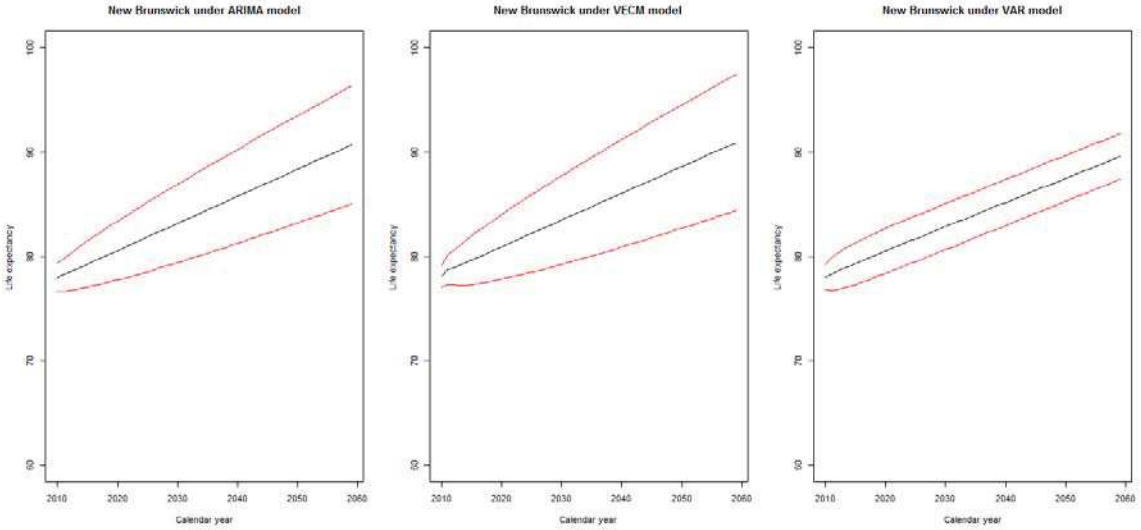

Fig. 5: Life expectancy of New Brunswick under three models: the red lines represent the lower and upper forecasting, the black line represents the point forecast

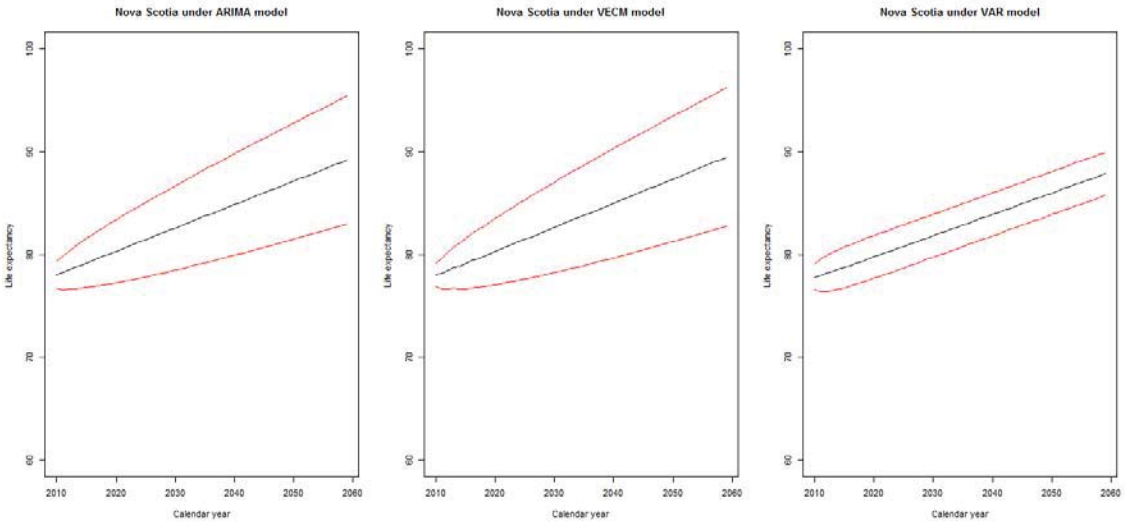

Fig. 6. Life expectancy of Nova Scotia under the three models: the red lines represent the lower and upper forecasting, the black line represents the point forecast

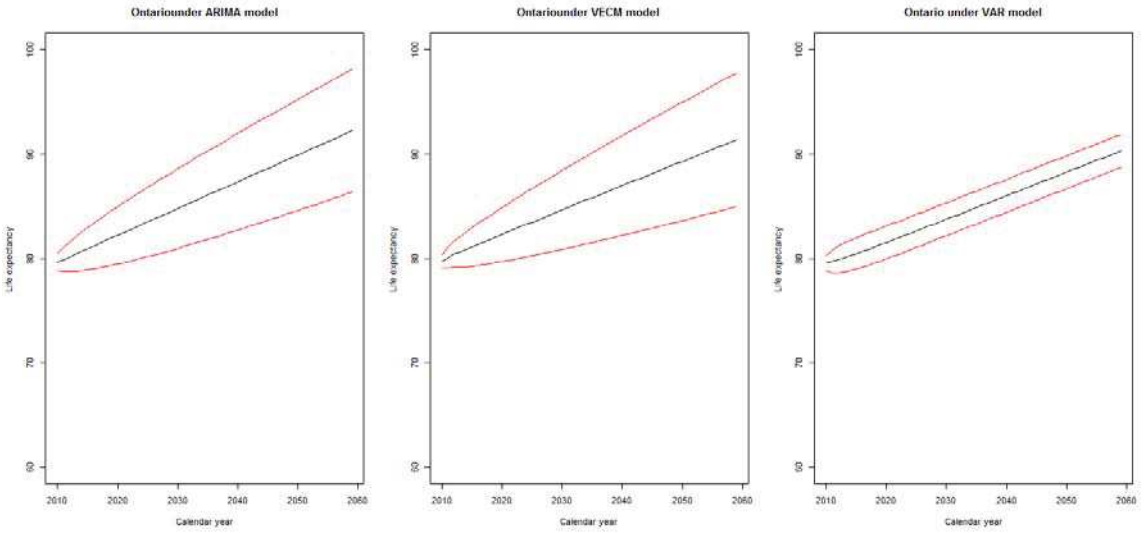

Fig. 7. Life expectancy of Ontario under the three models: the red lines represent the lower and upper forecasting, the black line represents the point forecast

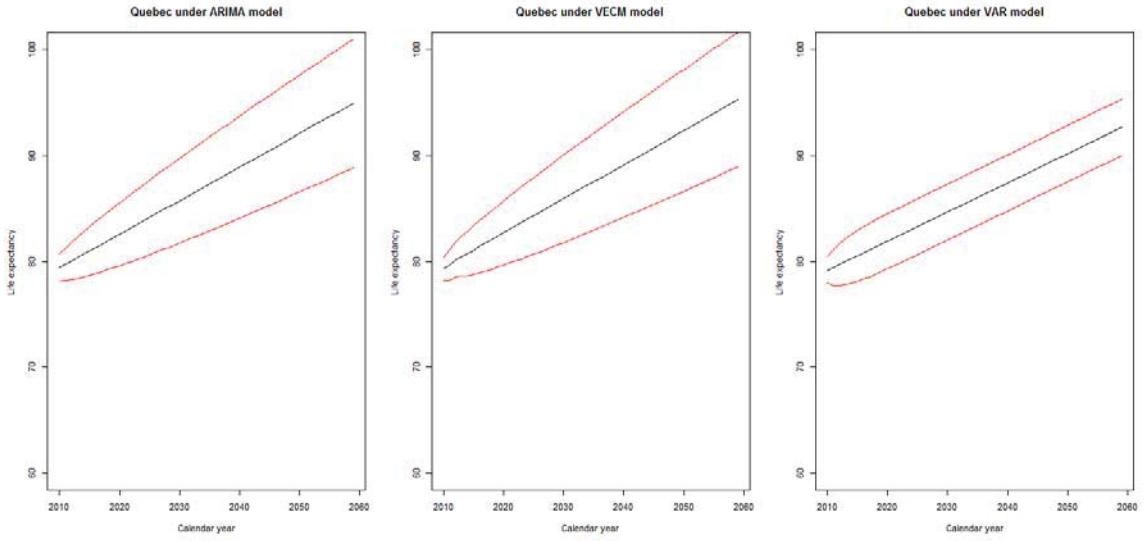

Fig. 8. Life expectancy of Quebec under the three models: the red lines represent the lower and upper forecasting, the black line represents the point forecast

Here we show the computations of life expectancy at birth derived from the VECM which has proven to be the best model of the three models investigated. The results are exposed in this framework with a frequency of 10 years as in Table 9, where we can see future life expectancy from 2010, 2020, 2030, 2040 and 2059. Data from annual frequency results are also available on request. We observe that life expectancy at birth from 2010 to 2059 is close to 90 years particularly in Alberta, British Columbia, New Brunswick, Ontario, and Quebec. Only Nova Scotia shows a life expectancy level below 90.

Table 9. Mean forecast of life expectancy with the VECM for the six provinces

Year	Alberta	British Columbia	New Brunswick	Nova Scotia	Ontario	Quebec
2010	79.28	78.12	78.02	79.74	79.74	79.30
2020	81.26	82.29	80.67	80.06	82.18	82.36
2030	83.57	84.71	83.23	82.41	84.72	85.59
2040	85.89	87.13	85.79	84.75	87.26	88.82
2050	88.21	89.55	88.35	87.10	89.79	92.05
2060	90.63	91.97	90.92	89.45	92.33	95.27

Conclusions

In this paper we have investigated the forecasting scenario of multi-population life expectancy for provinces in Canada. We have presented three principal econometric models ARIMA, VAR and the VECM which have appeared recently in the literature. The VECM presents a better performance than ARIMA and VAR models in terms of backtesting, goodness of fit, and future trend uncertainty quantification as shown by the confidence interval measured here. Furthermore, VECM highlights improvements in understanding the dynamics of life expectancy patterns over time as it captures common trends and also the correlation structure from the provinces monitored. We also illustrate the values of forecasts of life expectancy in the six provinces and found that it will surpass 90 years in the next 50 years except in Nova Scotia. The results from these analyses aim to help social security and insurance companies improve the quantification of future life expectancy and thus price pensions fairly.

Acknowledgement

The authors acknowledge financial support from the Fondazione Cariplo.

References

1. Adekola, O.A. (2002). A generalised life-expectancy model for a population, *Journal of the Operational Research Society*, vol. 53, pp. 919-921.
2. Alho, J.M. and Spencer, B.D. (2005). Statistical demography and forecasting. Springer.
3. Andreev, K.F., and Vaupel, J.W. (2006). Forecasts of cohort mortality after age 50, *Working paper*, *Max Planck Institute for Demographic Research*, Rostock, Germany.
4. De Beer, J. and Alders, M. (1999). Probabilistic population and household forecasts for the Netherland, *Working paper European Population Conference*, EPC99.
5. Babel, B., Bomsdorf, E., Schmidt, R. (2007). Future life expectancy in Australia, Europe, Japan and North America, *Journal of Population Research*, vol. 24(1), pp. 119-131.
6. Booth, H. (2006). Demographic forecasting: 1980 to 2005 in review, *International Journal of Forecasting*, vol. 22, pp. 547-581.
7. Box, G.E.P. and Jenkins, G.M. (1976) *Time Series Analysis for Forecasting and Control*. San Francisco: Holden-Day.
8. Canadian Human Mortality Database (2013). University of Montreal, available at: http://www.bdlc.umontreal.ca/chmd/.
9. Chen, Miranda and Ching, Michael (2000). A statistical analysis of life expectancy across countries using multiple regressions, Sys 302 Project.
10. Committee, Nobel Prize (2003). Time-series Econometrics: Cointegra-tion and Autoregressive Conditional Heteroskedasticity.
11. Denton T. Frank, C.H. Feaver, B.G. Spencer (2005). Time series analysis and stochastic forecasting: An econometric study of mortality and life expectancy, *Journal of Population Economics*, vol. 18, pp. 203-227
12. Dickey, D.A. and Fuller, W.A. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root, *Journal of the American Statistical Association*, vol. 74 (366): pp. 427-431.
13. Hamilton, J.D. (1994). Time Series Analysis Princeton University Press.
14. Harris, R. and Solis, D. (2002). Applied Time Series Modelling and Forecasting.
15. Hyndman, R.J. and Athanasopoulos, G. (2013). Forecasting: principles and practice, available at: http: otexts.com/fpp/ . Accessed on 05 September.
16. Juselius, K. (2007). *The Integrated VAR Model: Methodology and Applications. Advanced Texts in Econometrics*, New York: Oxford University Press.
17. Keilman, N., Pham, D.Q, and Hetland, A. (2001). Norways Uncertain Demographic future. Social and Economic Studies 105. Statistics Norway.
18. Maarten, A., Keilman, N. and Cruijsen, H. (2007). Assumptions for long-term stochastic population forecast in 18 European countries, *European Journal of Population*, 23: pp. 33-69
19. Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root, *Journal of Econometrics*, vol 54, pp. 159-178.
20. Lutkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*, Springer 2005, XXI.
21. Lee, R.D. and Carter, L.R. (1992). Modeling and Forecasting U.S. mortality, *Journal of the American Statistical Association*, vol. 87, pp. 659-675.
22. Lee, R. (2006). Mortality forecasts and linear life expectancy trends. In T. Bengtsson (Ed.), Perspectives on mortality forecasting. III. The linear rise in life expectancy: history and prospects. National Social Insurance Board, pp. 1940.
23. Maarten Alders, Nico Keilman and Harri Cruijsen (2007). Assumptions for long-term stochastic population forecast in 18 European countries, *European Journal of Population*, 23(1), pp. 33-69.
24. Oeppen, J. and Vaupel, J.W. (2002). Enhanced: Broken Limits to Life Expectancy. Science vol. 296 (4), pp. 1029-1031.
25. Olshansky, S.J., Passaro, D., Hershow, R., Layden, J., Carnes, B.A., Brody, J., Hayflick, L., Butler, R.N.; Allison, D.B., Ludwig, D.S. (2005). A Possible Decline in Life Expectancy in the United States in the 21st century, *New England Journal of Medicine*, vol. 352., pp. 1103-1110.

26. Pfaff, B. (2008). Var, SVAR and SVEC models: Implementation within R package vars, *Journal of Statistical Software*, Vol. 27 (4) p. 132.
27. Phillips, P.C.B., Perron, P. (1988). Testing for a unit root in time series regression, *Biometrika*, vol. 75, pp. 335-346.
28. Raftery, A.E, Chuun, J.L, Gerland, P., Sevcikva, H. (2013). Bayesian Probabilistic Projections of Life Expectancy for all Countries., *Demography Research*, vol. 30 (27), pp. 795-822, doi 10.1007/s13524-012-0193-x.
29. Russolillo, M. and Haberman, S. (2005). *Lee-Carter mortality forecasting: application to the Italian population*. London. Faculty of Actuarial Science and Statistics. Cass Business School.
30. Shaw, J.W., Horrace, C.W. and Vogel, R. (2005). The Determinants of Life Expectancy: An Analysis of the OECD Health Data, *Southern Economic Journal*, vol. 71 (4), pp. 768-783.
31. Torri, T. (2012). Building blocks for a mortality index: an international context, *European Actuarial Journal*, 1, pp. 127-141.
32. Tuljapurkar, S., Puleston, C.O., Gurven, M.D. (2007). *Why Men Matter: Mating Patterns Drive Evolution of Human Lifespan*. PLoS ONE 2 (8): e785. doi:10.1371.
33. Torri, T. and Vaupel, J.W. (2012). Forecasting life expectancy in an international context, *International Journal of Forecasting*, 28, pp. 519-531.
34. Whiteford, P. and Whitehouse, E.R. (2006). Pension Challenges and Pension Reforms in OECD Countries, *Oxford Review of Economic Policy*, 22 (1), pp. 78-94.

Appendix

Table 10. Unit root (Augmented Dickey Fuller) testing for nine Canadian provinces

Life expectancy	Constant	Lags	Dw stat	Constant and trend	Lags	Dw stat
Alberta	0.62	1	2.06	0.027	0	2.05
Δ*Alberta*	0	0	1.94	0	0	2.06
British Columbia	0.91	2	2.03	0	0	1.56
Δ*British Columbia*	0	0	1.87	0	1	2.04
New Brunswick	0.64	1	2.08	0.42	1	2.08
Δ*New Brunswick*	0.0063	4	1.94	0	0	2.08
Nova Scotia	0.83	0	2.42	0.07	0	2.18
Δ*Nova Scotia*	0	0	2	0	0	2.08
Ontario	0.67	0	2.11	0.1	0	2
Δ*Ontario*	0	0	1.93	0	0	1.93
Quebec	0.6	1	2.08	0.47	1	2.07
Δ*Quebec*	0	0	2.08	0	0	2.09

Table 11. Unit root (Phillips Perron) testing for nine Canadian provinces

Life expectancy	Constant	Lags	Dw stat	Constant and trend	Lags	Dw stat
Alberta	0.72	13	2.32	0.027	0	2.05
Δ*Alberta*	0	9	2.05	0	10	2.06
British Columbia	0.99	11	2.02	0	2	1.56
Δ*British Columbia*	0	7	2.05	0	7	2.05
New Brunswick	0.35	17	2.71	0.03	0	2.45
Δ*New Brunswick*	0	8	2.08	0	8	2.06
Nova Scotia	0.85	3	2.42	0.098	3	2.18
Δ*Nova Scotia*	0	3	2.08	0	3	2.08
Ontario	0.66	1	2.11	0.1	2	2
Δ*Ontario*	0	2	1.93	0	2	1.93
Quebec	0.53	10	2.52	0.44	2	2.38
Δ*Quebec*	0	5	2.08	0	6	2.09

Table 12. Unit root (KPSS) testing for nine Canadian provinces

Life expectancy	Constant	Lags	Dw stat	Constant and trend	Lags	Dw stat
Alberta	1.19	7	0.01	0.21	6	0.52
Δ*Alberta*	0.15	13	2.32	0.1	14	2.33
British Columbia	1.2	7	6.43	0.09	5	0.94
Δ*British Columbia*	0.16	11	2.02	0.07	11	2.03
New Brunswick	1.18	7	0.01	0.21	6	0.41
Δ*New Brunswick*	0.27	19	2.69	0.11	22	2.72
Nova Scotia	1.18	7	0.006	0.15	6	0.44
Δ*Nova Scotia*	0.059	4	2.43	0.053	4	2.43

Table 12 (cont.). Unit root (KPSS) testing for nine Canadian provinces

Life expectancy	Constant	Lags	Dw stat	Constant and trend	Lags	Dw stat
Ontario	1.19	7	0.006	0.15	6	0.25
Δ*Ontario*	0.12	0	2.1	0.08	1	2.11
Quebec	1.18	7	0.025	0.18	6	2.52
Δ*Quebec*	0.23	9	2.51	0.11	10	2.52

Table 13. Table of unit root (KPSS) of critical values at 5%

Critical values	Constant	*Constant-trend*
Levels	0.463	0.146
Different variable	0.463	0.146