Distr.: General
30 March 2016

English only

Economic Commission for Europe

Conference of European Statisticians

Joint Eurostat/UNECE Work Session on Demographic Projections
Geneva, 18-20 April 2016
Item 3 of the provisional agenda
Sub-national projections

Subnational population projections for the Republic of Korea, 2013 - 2033

Prepared by Statistics Korea[1]

Summary

Statistics Korea has recently developed a systematic framework for producing population projections for sub-provinces. It enables local governments to conduct population projections for their sub-provinces of various population sizes for up to 20 years. Republic of Korea consists of 17 provinces and 226 sub-provinces. Statistics Korea has conducted national and provincial population projections. Low fertility has continued for more than three decades and the gap in aging population among regions have been widened over time. Because of these trends, the demand for the long-term, unified, sub-provincial level planning using population projections has increased. To meet this demand, Statistics Korea has developed a web based sub-provincial projection program based on cohort component method.

In this paper, we will outline how to conduct various kinds of such small area population projections using this projection program, including age specific fertility rates, life expectances and migration rates. These sub-provincial projection results are then controlled to the national and provincial level of projection results. It projects age-specific fertility rates for each sub-province using linear regression models based on the recent relationship between province and each sub-province. For mortality projection, it applies Li-Lee model with a region's common factor and each sub-region's specific residual factor in age specific mortality change. To account sub-provincial geographical mobility, Origin-Destination matrices using out-migration rates from each sub-province to other regions are applied.

For the implementation of the sub-province level of projections, Statistics Korea has developed a web-based program including a regionally customized software using R packages. This software provides statisticians in local government offices a tool for analyzing and conducting their sub-province projections. The projection results can be summarized and visualized in graphs and charts from the software. It provides user friendly graphic interfaces for an intuitive maneuvering.

[1] [Prepared by Sooyoung Kim, Jee S. Baek, Taejick Lee, Myeongjik Han, Subok Lee, and Ji -Youn Lee].

I. Introduction

1. Republic of Korea consists of 17 provinces and 226 sub-provinces. Once every five years, Statistics Korea conducts national population projections and that of the 17 provinces. Local autonomy has been settled for 65 years in Korea and the demand for unified local policies has been increased. Sub-provincial population projections which are crucial to reasonable policy making and assessment have been desired by the central government, the local governments, and researchers. Korea's demographic situation of the lowest low fertility and the population aging which needs region detailed strategies is another reason for small area population projections. As low fertility has continued for more than three decades and varies by region, the gap on aging population among regions has been widened over time such that "oldest" sub-province shows 37% of the total population falls in the 'old age' population, while the "youngest" shows 6% for 'old age' population.
2. Some local governments produce their own population estimates for sub-provinces for future planning of the region. There was a tendency to overestimate future population growths. For some local governments, their predictions of the future population size in their area are more optimistic than their actual population growth. This overestimating tendency may result in not only cost overruns but also inefficient regional budget allocation. Recently, the ministry which is in charge of control of city planning and distribution of budget to the local government modified guidelines on the city planning, that local governments' population estimates should be coherent with national population projections by Statistics Korea.
3. To meet the demand for reliable sub-provincial population projections, Statistics Korea has recently developed a systematic framework for producing population projections for 226 sub-provinces which incorporates methodologies and web-based program. It enables local governments conduct population projections for their sub-provinces.
4. This paper describes the methods used for sub-provincial population projections and the framework of projection program in Korea.

II. Methods for small area population projections

5. For the sub-provincial population projections, this study reviews three methodologies: Hamilton-Perry's survival ratio method, sub-provincial population ratio method, and cohort component method. Hamilton-Perry's survival ratio method is the most commonly used for small area population projections. It can be applied when population by age is available for more than two periods. But since it only considers survival ratios of the population by age, it has drawbacks upon regional application especially for those regions with a large yearly divergence associated with internal migration component (Hamilton and Perry, 1962).
6. Another method of sub-regional projections is the method of proportions. In this method, sub-provincial population could be projected by applying projected proportion to official

provincial population projections. Since this method also does not consider the components of population change, it is limited in producing reliable long-term population projections.

7. Next, cohort component method is commonly used for official population projections at national and regional levels (Smith and Keyfitz, 1977). However it is known to be difficult for applying to small area population projections because of the requirement of the detailed population component data which is usually unavailable and unstable in many small areas. However, we found that cohort component method is the most applicable among the three methods, in terms of Goodness of Fitness (GOF) between observed data and projected data. The fact that the cohort component method provides not only population statistics by age and sex but also the information of population components is an important advantage of the method and provides unique usefulness to this study. To apply cohort component method for sub-provincial projections despite of the insufficient data, we had to develop new methodologies for each component. We used them after the GOF test.

III. Cohort Component Method for Sub-Provincial Population Projections

a) Population of Sub-Provinces in Korea

8. The difficulty of making reliable sub-provincial projections in Korea attributes to two reasons: deviation of the size of regional population and high internal migration rates. The population sizes of sub-provinces in Korea vary from 10 thousands to 1,171 thousands. In 2010, there were 63 sub-provinces (27.4% of total number of sub-provinces) having more than 300 thousand people, and 20 sub-provinces (8.7% of total) having a population size of 30 thousands or less.

< Table 1>. Population size of Sub-provinces

thousands	2005 (A)		2010 (B)		Difference (C=B-A)	
	numbers	proportion	numbers	proportion	Numbers	proportion
less than 30	16	7.0	20	8.7	4	1.7
30 - 50	34	14.8	35	15.2	1	0.4
50 - 100	47	20.4	44	19.1	-3	-1.3
100 - 200	41	17.8	37	16.1	-4	-1.7
200 - 300	31	13.5	31	13.5	0	0.0
over 300	61	26.5	63	27.4	2	0.9

9. Also, many sub-provinces experience large fluctuations in the size of their population size due to internal migration. The number of sub-provinces having population change of 10% or larger during a five year period amounts 51 or 22% of total sub-provinces. The largest

population increase during the five year period happened in a sub-province in Gyeonggi province, showing 70% growth in population, from 290 thousands in 2005 to 490 thousands in 2010.

<Table 2>. Population Changes of Sub-Provinces

(Numbers, %)

Population change(2005~2010)					
	Numbers	Proportion		Accumulated numbers	Accumulated proportion
Less than 3%	60	26.1	Total	230	100.0
3%-5%	47	20.4	Over 3%	170	73.9
5%-10%	72	31.3	Over 5%	123	53.5
10%-20%	37	16.1	Over 10%	51	22.2
20%-30%	9	3.9	Over 20%	14	6.1
Over 30%	5	2.2	Over 30%	5	2.2

b) Methods for the projection of fertility, mortality, and internal migration

10. In this chapter, we describe how cohort component method is applied to the sub-provincial population projections despite of the large deviation in population size and rapid population change at the sub-province level. The methods for each component are briefly outlined in following chapters.

(1) Fertility

11. The analysis of fertility rates of sub-provinces in Korea showed that the gaps of fertility between a province and each sub-province are consistent over time even though fertility rates have been decreased for most sub-provinces. Having this in mind, we projects age-specific fertility rates (ASFR) using linear regression models based on the recent relationship between a province and its sub-provinces (Brass, 1979).
12. The model for fertility projections of sub-province is as follows.

$$f_{x,i} = \alpha_i + \beta_i * f_{x,k}$$

α : intercept of regression model between the k province and the i sub-province
β : slope of regression model between the k province and the i sub-province
$f_{x,i}$: ASFR of the i sub-province
$f_{x,k}$: ASFR of the k province

13. In this process, the study uses age-specific fertility rates in five-year age groups for each province and its sub-provinces which can be obtained from the Statistics Korea's web portal. Then the study extends them to one-year-group age-specific fertility rates. By applying parameters estimated from each sub-province's model to fertility projections for each province, projected fertility rates for each sub-province could be obtained.

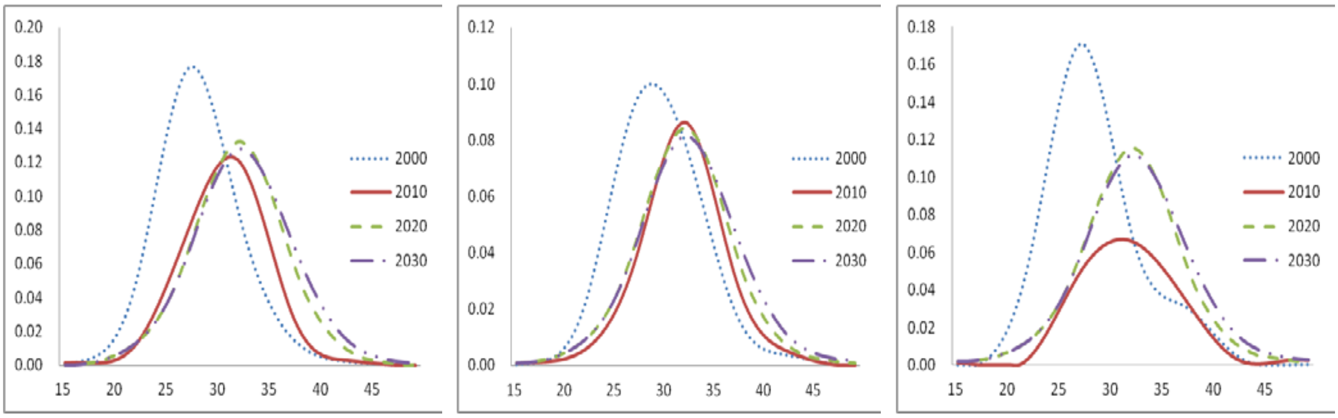
14. To assess the validity of models, we overlaid the estimated models to the observed ASFR for the period of 2000-2012 and compared the observed ASFR to the estimated fertility from models. Then the mean absolute percent errors (MAPE) of the forecasts have been examined. MAPE between the observed and the estimated by model appears to be less than 5% on average (<Table 3>). For the period of 2015-2030, the forecasted age specific fertility rates of the selected three sub-provinces with population sizes varying from 10 thousands to 1,171 thousands in 2015 are shown in < Figure 1 >.

<Table 3> Mean Absolute Percent Errors of Total Fertility Rates for Sub-Provinces, 2000~2012

Province	Average	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16
MAPE(%)	4.8	2.8	4.2	3.9	4.4	4.6	3.8	3.3	3.5	4.5	4.8	4.1	7.9	.5	6.9	5.7	3.6

< Figure 1> Forecasted Age Specific Fertility Rates of Sub-Provinces, 2015~2030

SP1 SP2 SP3



(2) Mortality

15. Mortality rates for each sub-province in Korea have been decreased over time but the gaps of mortality among sub-provinces have been narrowed. Reflecting these results, we applied Li-Lee model for mortality projections. This model is known to be useful for coherent group mortality forecasting by considering a region's common factor as well as each sub-region's specific residual factor of age specific mortality change (Li and Lee, 2005; Kim, 2011). The models are as follows.

$$\ln(m_{x,t,i}) = a_{x,i} + B_x * K_t + b_{x,i} * k_{t,i} + \varepsilon_{x,t,i}$$

$m_{x,t,i}$: death rates for sub-province i, year t, age x
$a_{x,i}$: age pattern of mortality for sub-province i, age x
B_x, K_t : common factors for a province: age pattern of mortality change and time varying index of the level of mortality
$b_{x,i}, k_{t,i}$: specific factors for each sub-province: age pattern of mortality change and time varying index of the level of mortality
$\varepsilon_{x,t,i}$: errors for sub-province i, year t, age x

16. To project death rates of sub-provinces for the period of 2013 -2033, we had to perform several advanced processes including interpolation, adjusting, and old age death rates extension (Lee et al., 2013). The details of each process are shown below;

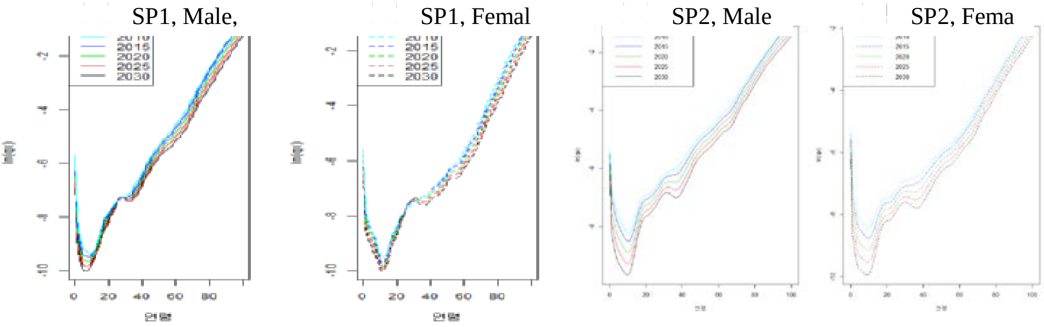
	1) **Interpolation**[1)]	2) **Adjusting**	**3) Old age death rates extension(80+ →100+)**[2)]	**4) Forecasting**
Data	The number of deaths for provinces and sub-provinces for the period of 2000~2012 (5 year age, age 0-79) Mid-year population	Death rates of sub-provinces for the period of 2000~2012 (single age, age 0-79)	Death rates by sub-province for the period of 2000~2012 (age 65~75)	Death rates by sub-province for the period of 2013~2033 (single age, age 0-100+)
Models	Spline interpolation	Brass-Logit Model	2 Parameters Logistic	Li-Lee model

1) Three year's moving average, extending death rates from five year age interval to one year age interval

2) In the old age death rate, the death rate of the oldest age group reflects an aggregate of the death rates of any age group overflowed above the limit. To overcome this limit and to get more accurate mortality forecasting including old age death rates, we extended the age boundary of the old age death rates from age 80 to age 100.

17. Figure 2 shows the forecasted mortality rates using above mentioned method for two sub-provinces for the period of 2010 to 2030. For a sub-province whose age specific death rates in 2010 were stable, the forecast result appears to be rather stable for forecasting period. While the forecasted mortality has some fluctuations for the other sub-province with an unsteady mortality, the fluctuation gets moderated in long-term projections.

< Figure 2 > Sub-provincial log-death rates, 2010~2030



18. The mean absolute percent errors of life expectancies between a province and each sub-province are less than 5% for the forecasting period. The coherent trends of observed death rates among province and their sub-provinces are reasonably reflected on the forecasted mortality.

< Table 4> Mean Absolute Percent Errors of life expectancies between province and sub-provinces, 2000~2030

	Males				**Females**			
year	**2000**	**2010**	**2020**	**2030**	**2000**	**2010**	**2020**	**2030**
P1	1.38	1.17	1.50	3.25	0.81	0.68	1.32	2.67
P2	1.21	1.12	1.00	1.69	0.94	1.06	0.89	1.84
P3	1.54	1.46	1.19	1.89	1.02	0.96	1.84	2.80
P4	1.13	0.96	1.00	0.83	1.57	1.37	1.62	1.40
P5	0.80	0.65	0.76	1.24	1.26	0.53	1.36	2.14
P6	1.74	1.64	2.02	4.28	0.54	0.46	2.40	4.40
P7	0.85	0.74	0.64	0.67	0.72	0.59	0.36	0.47
P8	1.90	1.59	1.30	1.96	1.11	1.09	0.86	1.49
P9	1.56	1.52	1.52	1.14	1.06	0.98	1.05	1.23
P10	1.62	1.01	1.07	1.08	0.90	1.32	0.88	1.08
P11	1.53	1.35	1.03	2.12	0.86	1.22	0.98	2.20
P12	1.39	1.45	1.04	1.80	0.94	0.89	1.44	2.36
P13	1.68	1.09	0.79	1.86	1.01	0.90	0.93	1.61
P14	1.40	1.19	0.98	1.13	1.08	1.15	1.13	1.36
P15	2.05	1.72	1.19	1.30	1.45	1.28	1.03	1.83
P16	5.58	3.66	2.46	3.03	3.25	2.93	3.56	3.71

(3) Migration

19. In 2015, inter sub-province migration rate is 10%. It means that 10 persons out of 100 Koreans move to different sub-provinces during a year. Migration is the most crucial component to the credibility of regional population projections. In the provincial population projections, a multi-regional model based on Origin-Destination migration matrix is used (Feeney, 1973). This method considers not only the size of migrant but also the direction of migration. The matrix for provincial projections is composed of 17 origin-province by 17 destine-province for two sexes and 101 single ages from age 0 to 100 and over. A problem of having too many data arrays arises upon constructing the Origin-Destination-age-sex dimension matrices for sub-provinces: two sexes and 101 ages for 226 sub-provinces of origin by 226 sub-provinces of destine. Therefore, the dimension of matrix of sub-provinces has been reduced to consider 1) migration of a sub-province to other sub-provinces within the province and 2)the total migration to all sub-provinces in other provinces. The model used in the study is shown below:

$$\mathrm{NM}_{\mathrm{i}} = (\mathrm{InM}_{\mathrm{Intra\ i}} + \mathrm{InM}_{\mathrm{Inter\ i}}) - (\mathrm{OutMI}_{\mathrm{Intra\ i}} + \mathrm{OutM}_{\mathrm{Iter\ i}})$$

NM_{i}: net-migration of sub-province i, i=sub-provinces, 1,2,…, J-1, J=otherwise
$\mathrm{InM}_{\mathrm{Intra\ i}}$: intra-province inmigration of sub-province i, $\sum_{j=1}^{J-1} \mathrm{InM}_{\mathrm{intra\ i}}\ (j)$
$\mathrm{InM}_{\mathrm{Inter\ i}}$: inter-province inmigration of sub-province i, $\mathrm{InM}_{\mathrm{Inter\ i}}(J)$
$\mathrm{OutM}_{\mathrm{Intra\ i}}$: intra-province outmigration of sub-province i, $\sum_{j=1}^{J-1} \mathrm{OutM}_{\mathrm{intra\ i}}\ (j)$
$\mathrm{OutM}_{\mathrm{Inter\ i}}$: inter-province outmigration of sub-province i, $\mathrm{OutM}_{\mathrm{inter\ i}}(J)$

20. In Korea, regional development programs such as city planning or civil construction evolve dynamically, having a great influence on population growth. To improve the validness of this projections, especially in short term, our framework includes an option of reflecting region’s development planning, either confirmed or in progress, to migration projections by using controlling weights.

IV. Statistics Korea Small Area Population Projection Program (KOSTAT-SPP)

21. To provide tools for the local governments to independently conduct the population projections for their sub-provinces, Statistics Korea developed a web-based program for sub-provincial population projections. It is constructed to apply cohort component method using component projection methods mentioned in the previous chapter.
22. This program is a turn-key solution providing functional modules required by the users, such as data input, components projection, tabulation of projection results, creating visual charts

and so on. It consists of five modules: data downloading, projecting cohort components, uploading projected data, compiling, and result and visualization. In data downloading, births, deaths, and migration data can be downloaded from Statistics Korea Data Base web portal (KOSIS). The module of projecting cohort components has a function of components projection of fertility, mortality and migration in a regionally customized program with free statistical package R. In a module of uploading, one can upload projected fertility, mortality, and migration. Through a module of compiling, projected fertility, mortality, migration, and base population data are combined for producing future population by age and sex. In a module of result and visualization, tables and visual charts are generated. Finally, these sub-provincial projection results are controlled to national and provincial level of population projection results. The sum of projected births, deaths, and migration for sub-provinces are made adjustments to those of the provincial projection results by sex and age. Then, the sum of population by sex and age for sub-provinces are regulated to those of the province, respectively. Through the final step, the major challenge of our sub-provincial projection for keeping internal consistency with national and provincial population projections is resolved.

<Table 5>. Modules of KOSTAT-SPP

Observed Data	Components Projection	Results from Cohort Component Method (consistency)		
Fertility rates Death Rates Migration Rates	Fertility Rate Death Rate Migration Rate	Base Population	Births Deaths Migrants	Projected Population
KOSIS (Statistics Korea Data Base Portal Site)	R program (Standardized and regionally customized)	Web-Based Program(KOSTAT-SPP)		

The menu of KOSTAT-SPP is as in <Table 6>.

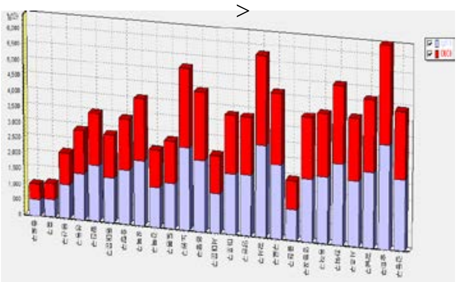
<Table 6>. Program Menu of KOSTAT-SPP

Data Uploading	Projecting Cohort Components	Projection Results	Notice
Base Population	Internal Migration Weight	Population by Sex and Age	Information
Sex Ratio at Birth	Demographic Balance Equation	Births, Deaths, Migrants	Q & A
Age-Specific Fertility Rate		Demographic Indicators	
Age Specific Survival Probability			
Origin-Destination Migration Matrix			

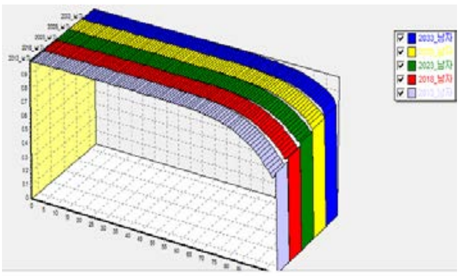
Through this Program, we also provide visualization module of various charts.

<Figure 3>. Charts from KOSTAT-SPP

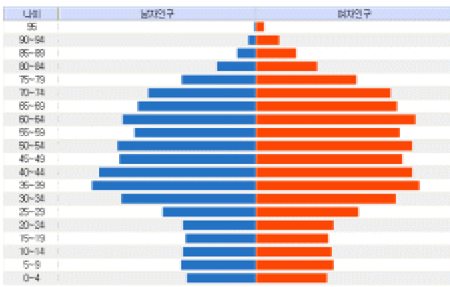
<Births>



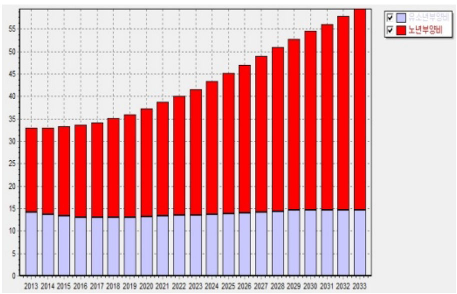
<Probability of Survival>



<Population Pyramids>



<Dependency Ratios>



V. Conclusion

23. Statistics Korea recently developed population projection framework for 226 sub-provinces. It incorporates components projection methodologies and web-based projection program, and comes with a user's manual. In this project, we found that cohort component method is available for small area population projection with large deviations of population size and rapid population changes, with the help of data adjusting techniques.

24. We expect this systematic framework for sub-provincial population projection could facilitate evidence based policy administration by providing reliable forecast of regional population. Through cooperation between national statistical office developing projection framework and local governments implementing population projections, Republic of Korea can successfully build the regional statistical infrastructure that is necessary for policy formation, implementation, and assessment. This projected sub-province populations may help national and local governments allocate resource efficiently, implement and monitor their programs effectively.

References

Baek, J. Jeong, M., Oh, Y., Lee, J., and Kim, S. 2013. Evaluation of Korean Mortality Forecasting Models. Eurostat/UNECE Work Session on Demographic Projection. October 2013.

Brass, W. 1971. On the Scale of Mortality. In W. Brass (ed.). Biological Aspects of Demography. London: Taylor and Francis.

_______. 1979. The relational Gompertz of fertility by age of woman. Asian Population Studies Series No 44.

Coale, A. and Kisker, E. 1990. Defects in Data on Old Age Mortality in the United States: New Procedures for Calculating Approximately Accurate Mortality Schedules and Life Tables at the Highest Ages. Asian and Pacific Population Forum, 4, 1-31.

Coelho, E., Magalhães, M. and Bravo, J. M. . 2007. Mortality and Longevity Projections for the Oldest–Old in Portugal. in Working Session on Demographic Projections, Bucharest, Romania, 10-12 October, 2007, EUROSTAT.

Feeney, G. 1973. Two models for multiregional population dynamics, Environment and Planning A, 5(1): 31-43.

Hamilton, C. H. and Perry, J. 1962. A short method for projecting population by age from one decennial census to another, Social Forces, 41(2): 163-170.

Kannistö, V. 1992. Development of Oldest-Old Mortality, 1950-1990: Evidence from 28 Developed countries. Odense University Press.

Kim, S. 2011. Mortality Forecasting for the Republic of Korea: the Coherent Lee-Carte Method. Korea Journal of Population Studies. 34(3): 157-177.

Lee, J., Baek, J. Jeong M., Kim, S. and Oh, Y. 2013. Examination of Korean Mortality Forecasting Models. The 27 IUSSP International Population Conference. August 2013.

Li, N. and Lee R.D. 2005. Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method. Demography 2005 August; 42(3) : 575-594.

Swanson, D. A. and Tayman, J. 2012. Subnational Population Estimates, Dordrecht:Springer.

__________________________.2013. The accuracy of the Hamilton-Perry method for forecasting state population by age.

Swanson, D. A. and Tedrow, L. M. 2013. Exploring stable population concepts from the perspective of cohort change ratios, The Open Demography Journal, 6: 1-17.

Wilson, T. 2011. Modelling with NEWDSS: Producing state, regional and local area population projections for New South Wales, in Stillwell, J. and Clarke, M., eds., Population Dynamics and Projection Methods, Dordrecht: Springer, 61-97.

_________2014. New evaluations of simple models for small area population forecasts, Population, Space and Place, Published online: 24 March 2014.

Statistics Korea. 2011. *Population Projections for Korea, 2010~2060*.

____________ . 2012. *Provincial Population Projections for Korea, 2010~2040*.

____________ . 2014. *Provincial Population Projections for Korea, 2013~2040*.