

제3장 장래가구추계방법 개선 연구



김수영 · 박영실

제1절 서론

1. 연구배경 및 목적

출산력 및 사망력 감소, 혼인의 지연 등에 의한 인구 및 사회변화는 한국의 가구 현황을 급격하게 변화시키고 있다. 그 변화 양상은 총 가구 수 뿐만 아니라 가구의 형태 및 구성원수의 측면에서 나타나고 있는데, 최근의 두드러진 변화는 1-2인가구가 급속하게 증가하여 주된 가구형태로 자리 잡고 있다는 점이다(통계청, 2011).

통계청에서는 이러한 가구의 변화를 예측하기 위해 2002년 최초로 장래가구추계를 작성한 이래로 2007년 제2차 장래가구추계를 작성하였다. 2007년 장래가구추계는 시도별 및 세분화된 가구유형별로 추계자료를 제공하면서 다양한 수요에 부응하고 있다는 평가를 받고 있다. 그러나 2007년에 예측된 총가구수 및 가구구성별 가구수가 2010년 총조사 자료와 일부 차이를 보여 통계의 신뢰성에 대한 문제가 제기된 바 있다. 또한 1-2인 가구의 급격한 증가 등의 변화를 반영하는데 한계가 있었음이 나타나면서, 가구규모 및 가구유형 등에 따른 세분화된 통계에 대한 예측력 향상이 필요하다는 지적이 있었다(통계청, 2011; 최은영 외, 2011). 그러나 가구추계에 영향을 미치는 변수들은 출산, 사망, 이동, 혼인 등 구체적인 인구학적 변수 이외에도, 개인의 가치관이나 경제활동패턴의 변화, 복지정책 등 사회문화적 변수 등 복잡하고 다양하다. 따라서 이런 다양한 변수들을 모두 고려하여 장래의 가구규모 및 구조를 예측하는 일은 쉽지 않은 일이다(김형석, 2002).

이에 본 연구에서는 급격하게 변동하고 있는 한국사회에서의 장래가구추계의 타당성을 높이기 위해 해외사례를 포함한 다양한 가구추계방법에 대한 검토 및 2007년 장래가구추계의 단계별 방법을 검토하고, 이를 토대로 2012년 장래가구추계를 위한 개선안을 제시하고자 한다.

2. 연구범위

본 연구의 구성내용은 다음과 같다. 우선 2절에서는 대표적인 가구추계방법 및 주요국의 가구추계작성방법을 살펴보고 시사점을 찾고자 한다. 3절에서는 2007년 장래가구추계 방법을 작성단계별로 검토할 것이다. 그리고 대안적인 방법들을 적용함에 따라서 2010년 총조사 자료와의 적합도를 검증한 후 이를 토대로 각 단계별 개선안을 제시하고자 한다. 마지막으로 제 4절에서는 결과 요약 및 제언을 하고자 한다.

제2절 가구 추계방법의 이해

1. 가구 추계방법의 종류

대부분의 국가에서는 취사, 취침 등 생계를 같이 하는 생활단위인 가구를 추계대상으로 하고 있다. 이는 혈연 및 일정한 친족관계를 가진 가족의 경우 직업, 학업 등의 이유로 함께 살지 않는 경우가 많기 때문에 가족관계의 변화를 파악하는 것이 기술적으로 매우 어렵기 때문이다.

가구를 추계하는 방법은 작성기법 및 이용 자료의 종류 등에 따라서 달라진다. 작성기법에 따른 가구추계방법은 정태모형과 동태모형으로 나뉜다. 정태적 모형은 독립적으로 추계되는 인구의 구성비나 가구의 비율에 따라 가구를 할당하여 추정하는 방식이며, 동태적 모형은 일정기간동안 개인이나 코호트의 행태를 추적하여 가구상태 간 변화나 전환을 추정하는 방식이다. 이용 자료의 종류에 따른 구분으로는 집합 자료를 사용하는 거시모형과 개별 자료를 사용하는 미시모형으로 구분할 수 있다(Bell et al, 1995). 따라서 어떠한 종류의 자료를 이용하여 어떠한 작성기법을 사용하느냐에 따라서 가구추계방법이 구분되는데, 여기에서는 주로 활용되고 있는 세 가지 가구추계 방법을 소개하고자 한다(Suzuki, 1999; Yi et al. 2003).



가. 거시-정태모형 : 가구주율법에 의한 접근방법

거시-정태모형 중 대표적인 것으로 가구주율법(headship rate method)이 있는데, 이는 가구주 및 비가구주의 변동 등은 고려하지 않고 성 및 연령 또는 혼인상태별로 가구주의 수를 예측하는 것이다. 이 방법은 1930년 미국의 국가자원기획위원회(National Resources Planning Committee)가 센서스 자료를 이용하여 처음 사용하였다. 가구주율은 센서스 자료에서 쉽게 입수가 가능한데, 이러한 자료 접근의 용이성은 가구주율법이 대표적인 추계방법으로 자리 잡는데 중요한 기여를 하고 있다.

가구주에 대한 개념이 바뀌지 않는다면 시계열 자료를 이용하여 가구주율의 미래의 외삽경향을 추정할 수 있다. 추정된 가구주율은 미래의 성 및 연령별 혼인상태별로 인구에 각각 적용되어 장래의 가구수를 추계한다. 가구주율의 외삽법 추정을 위해서는 여러 가지 수학적 방법이 적용되고 있는데, 수정지수식, 선형로그식, 선형회귀식, ARIMA, Random Walk(이하 RW) 등의 시계열모형이 그 예이다. 한편, 미국의 센서스국(Census Bureau)에서는 가장 최근인 1996년 가구추계에 Arima모형을 이용하여 가구주율을 추정하고 있다.

나. 거시-동태모형

거시-동태모형은 총조사 또는 표본조사의 개별 자료를 매칭하여 인구학적 상태 간 변동을 추적하여 가구를 추계하는 방법으로 가구의 형성 및 해체를 반영한다는 특징을 갖고 있다. 이 모형은 인구변화요인이 가구에 미치는 영향을 포착하는 데 적절한 모형으로 평가되고 있다. 예컨대, 1인가구의 증가를 결혼 전 분가 및 이혼 등 혼인의 해체 또는 배우자의 사망 등 인구학적 요인으로 분해할 수 있다. 개인이 가구에서 점하고 있는 지위를 유동적으로 볼 수 있다는 장점이 있으나, 전이확률을 정확하게 산출하기 위해서는 광대한 양 및 종류의 자료를 필요로 한다는 단점이 있다. 두 센서스 자료 또는 동태등록자료를 매칭시킨다거나 특별히 설계된 표본조사를 수행하는 것이 요구된다. 이 방법은 주로 유럽에서 1980년 이후 시도되었다. LIPRO(Lifestyle PROjection) 모델 및 다차원 모델(multi-dimensional model) 등이 여기에 해당한다.

1) LIPRO 가구추계모형

LIPRO 모델은 동거 및 혼외자녀의 증가 등 최근 네덜란드의 생활패턴 변화를 반영하기 위해 개발된 것으로 개인의 가구 내 지위사이에서의 변동을 추적한다. 추이행렬 작성을 위해 주택수요에 관한 별도의 표본조사를 실시하였는데, 이 조사를 통해 네덜란드

46,730가구에 대해 조사시점 및 1년 전 가구내에서의 개인의 지위자료를 산출하였다 (Van Imhoff and Keilman, 1991, p.61, p73-74).

1. CMAR : 결혼한 가구내의 자녀
2. CUNM : 동거가구내의 자녀
3. CIPA : 한부모가정의 자녀
4. SING : 1인가구
5. MAR0 : 부부가구
6. MAR+ : 부부 및 자녀가구
7. UNM0 : 동거부부(자녀는 없음)
8. UNM+ : 동거부부 및 자녀가구
9. HIPA : 한부모가구의 가구주
10. NFRA : 가족이 아닌 성인
11. OTHR : 기타가구

이 모델의 추이행렬은 $11 \times 11 = 110$ 의 원소를 갖고 있으며, 가구추이율을 얻는 방법에 따라 선형모델 및 지수모델 등 두 가지 종류가 있다. 선형모델은 모든 사건이 추계기간에 균등하게 분포한다는 가정에 따른 것이며, 지수모형은 마르코프 추정이론으로부터 산출될 수 있다. LIPRO 모델은 부부가구의 남자 및 여자가 동일할 수 있어야 한다거나, 부모 및 어린이에 대한 조건 등을 부여하여 일관성을 만족시키고 있다.

2) 다차원 모델(Multi-Dimensional model)

Zeng 외(1997)는 앞의 LIPRO 모델과 달리 특별한 표본조사자료를 요구하지 않고, 일반적인 인구자료로 적용이 가능한 동태모형을 작성한 바 있다. 인구를 $N_{k,m,p,c}(x,t,s)$ 라 했을 때 (x =연령, t =연도, s =성별) 각 연도의 성 및 연령별 인구는 부모와의 거주여부 k , 혼인상태 m , 자녀수 p , 자녀와의 동거여부 c 등 네 가지 상태에 의해 분류된다. 이 모델에서의 가구 내 지위 변동은 부모-자녀, 또는 파트너와의 관계의 시작 및 종결로부터 작성된다. 예를 들면, 결혼 시 부모의 집을 떠나는 여자의 결혼은 단지 “아이”에서 “부인”으로의 지위 변화뿐만 아니라, 혼인상태 및 부모와의 동거에 대한 변화를 동시에 보여준



다. 이러한 특성은 이 모델이 가구 내 지위의 변동을 인구학적 요소로 분해해 준다는 장점이 있다.

그러나 다차원 모델의 이러한 전략은 엄청나게 큰 규모의 전이확률을 요구한다. 거주상태(도시 및 농촌), 혼인상태(미혼, 유배우, 사별, 이혼, 동거), 부모와의 동거상태(두 부모와 동거, 한부모와 동거, 부모와 동거하지 않는 경우), 여섯 경우의 출산순위, 여섯 경우의 자녀와의 동거여부에 따라 개별 성 및 연령별로 630개의 복합적인 결과가 생성되며, 그 결과 396,900 (=630*630)개의 원소를 가진 전이확률행렬을 필요로 한다.

이런 광대한 데이터 요구를 피하기 위해 이 모델은 강한 가정을 내포하고 있다. 예를 들면 이 모델에서는 성, 연령별 분가 비율을 사용하고 있는데, 분가와 혼인은 상관성이 높은 사건임에도 불구하고 두 사건을 독립적이라고 가정한다. 또한, 어린이는 부와 모 사이에 균등하게 분포하며, 이혼 후 자녀는 모와 함께 거주하는 것으로 가정한다. 뿐만 아니라, 이 모델에서는 작성된 현재의 전이확률행렬이 변하지 않는다는 가정 하에 장래의 전이확률을 예측하고 있다. 그런데, 장래확률의 변화를 예측하는 것이 추계의 중요한 부분이라는 점을 감안한다면, 현재의 확률을 장래추계에 그대로 이용한다는 점은 이 모델의 단점으로 지적되고 있다.

결론적으로 말하면, 다차원모델은 일반적인 인구자료를 이용한 상세한 전이확률을 필요로 하며, 이를 위한 거대한 전이확률행렬 요구는 이 모형의 적용을 어렵게 하고 있다. 이 모델 적용을 위해 Profamy라는 시스템을 유료로 제공하고 있으나, 구체적인 작성 절차는 제공하고 있지 않다.

3) 가구전이율법(Household Transition Method)

이 모델은 1998년 일본의 전국 가구추계에 사용된 방법으로 LIPRO모델과 유사하다. 이 모델에서도 전이확률 산출을 위해서 특별히 디자인된 표본조사 자료를 이용하고 있다. 국립사회보장·인구문제연구소(National Institute of Population and Social Security Research)에서는 5년마다 가구동태조사를 실시하고 있으며, 이 자료에서 5년 전 및 현재 조사시점의 가구 내 지위 자료를 제공한다.¹⁾ LIPRO 모델은 네덜란드의 동거 및 혼외자의 증가를 반영하도록 설계되어진 반면, 일본에서는 혼인을 주거의 중요한 지표로 설정하고 있다. 한편, 가구전이율법에서는 전이확률산출시 선형모형을 사용하고 있다.

1) 일본의 가구전이율법은 다음절의 해외사례에서 보다 자세하게 소개하고자 한다.

다. 미시-동태모형 : 가구시물레이션통합모델

미시-동태모형은 개인의 상태가 동태적 사건 또는 다른 사건에 의해 변동하는 확률을 반영하는 방법이다. 이 때 특정전이를 결정하기 위해 Monte Carlo 표본을 이용하며, 출생, 사망 등 특정사건에 대해 조건부 확률을 추계하여 적용하고 있다. 이 모형은 일관성 및 유동성 측면에서는 가구 및 가족의 동적 변동을 관찰하는데 적합한 것으로 평가받으면서, 유럽이나 오스트레일리아 등에서 사회정책을 계획하고 평가하는데 많이 활용되어 왔다(Fukawa, 1995).

이 모형에는 가구시물레이션통합모델(INAHSIM; Integrated Analytical Model for Household Simulation)이 있는데, 개인의 출생·사망·혼인·이혼·별거·가족의 합성·부모가구로의 편입 등을 시물레이트시키는 것이다. Fukawa(2007)가 일본의 가구추계에 적용한 적이 있다. 미시-동태모형은 가구구조에 있어 인구의 효과를 검증하기 위해 여러 시도가 있어야 하므로 거시-동태모형에 비해 고비용의 문제가 있으며, 특히, INAHSIM 모델에서는 자료의 질 문제가 심각하다.



[그림 3-1] 가구시물레이션통합모델 2006/2007 업무흐름도



2. 해외 통계작성기관의 가구추계 작성 사례

다음은 주요 국가별 장래가구추계 현황을 살펴본 것이다.

〈표 3-1〉 국가별 장래가구추계 현황

원계	한국	일본	미국	영국		호주 ¹⁾
				잉글랜드 웨일즈	스코틀랜드	
작성기관	통계청	국립사회 보장·인구 문제 연구소	센서스국	부총리실	행정부	통계청
작성주기	5년	5년	부정기	4년 또는 6년	2년	부정기
최근작성 년도 및 추계기간	2007 (2005-2030)	2008 (2005-2030)	1996 (1995-2010)	2010 (2008-2033)	2010 (2008-2033)	2011 (2006-2031)
작성방법	가구주율법 - 혼인상태 별 인구 및 장래의 가구 주율을 추정 하여 추계	가구전이율법 - 개인의 배 우관계 및 가구내 지위 추이 확률 행 렬에 의해 추계 ²⁾	가구주율법 - 가구변화의 과거 및 미 래의 경향을 타임 시리즈 모델에 기초 하여 추계	가구주율법 - 성, 연령, 혼인 및 동 거상태별 가 구주율 추계	가구주율법 - 연령 및 가 구구성 그룹 별 가구주율 추계	Propensity method - 동거형태 별 경향의 미래 변화율 을 가정하여 추계

1) 뉴질랜드도 호주와 동일하게 동거형태율법(Propensity Method) 가구추계를 작성하고 있음

2) 최근의 도도부현별 추계, 1993년까지의 전국추계에 가구주율법 사용

가. 미국

가장 최근의 가구추계는 1990년 센서스를 이용하여 1996년에 작성된 것으로 가구구성비를 시계열 모델을 이용하여 추정하는 가구주율법을 적용하고 있다.²⁾ 그 작성방식을

2) 이후 가구추계를 작성하지 않았으며, 향후 1-2년 내에는 작성할 계획이 없다는 답변을 미국 센서스국으로부터 받았다. 추가적인 가구추계 미작성에 대해서는 가구추계에 대한 수요 감소, 인력부족을 원인으로 들고 있다.

보면, 혼인상태인구 산출→결혼한 가구를 산출→전체 가구에서 결혼한 가구를 제외하는 방식으로 다른 가족유형 및 비가족가구 산출 순이다.

1) 장래가구구성비 산출

장래의 가구의 변화에 대한 서로 다른 패턴을 보여주기 위해 3가지 방법을 제시하고 있다. 첫 번째는 시계열 모델을 이용하여 가구구성비 변화를 예측하는 것이며, 다른 두 가지는 가구구성비 변화는 없다고 가정한 후 성 및 연령별 인구구조의 변화만 반영하는 방법과 성 및 연령별 구조, 인종 및 출신 구성비 변화를 고려하는 방법 등이다.

1959-1993년 센서스 데이터에 기초한 혼인상태 및 Current Population Survey(CPS) 데이터에 기초한 연령 및 성별 가구구성비 자료에 기초하여 100세트의 연간자료 시계열을 산출한 후 시계열 모델에 기초하여 추정하였다. 이 때 100세트의 비율은 정상분포에 근접시키기 위해 다음과 같이 로지스틱변환을 이용하여 변형하였다.

$$y_t = \log(x_t / (1 - x_t))$$

시계열은 각각의 y_t 시리즈에 적합시킨 후 x_t 를 예측하기 위해 다시 변환되었다.

2) 혼인상태인구 추계

혼인상태인구의 성·연령·인종·히스패닉별 초기비율을 산출한 후, 가구구성비와 동일하게 시계열 모델을 이용하여 추계하였다. 추계인구에 이 비율을 곱하여 각 특성별 혼인상태별 인구를 산출한다.

3) 결혼한 가구 추계

결혼한 가구의 현재비율은 결혼건수에 기초한다. 결혼한 가구의 추계된 비율의 변화는 순차적 연도에 대한 현재의 비율에 곱해져서 매년 추계 비율을 산출한다. 다음 결혼한 가구의 각 연도의 연령, 성, 인종, 히스패닉계별 비율은 미래의 결혼한 인구에 곱해져서 결혼한 가구의 연도별 수를 산출한다.

4) 다른 가족유형 및 비가족가구의 추계

여자 가구주 가구, 남자 가구주 가구 및 비가족가구는 전체가구인구에서 결혼한 가

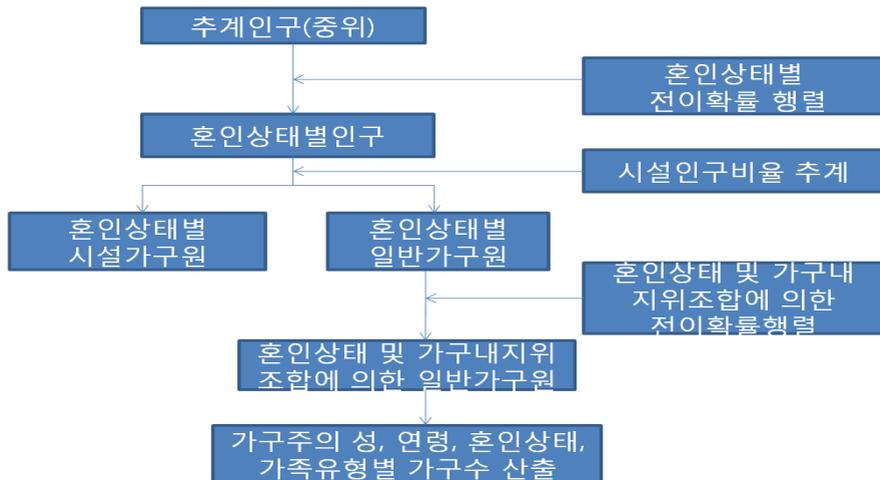


구의 인구수를 빼서 산출한다. 1990년 센서스로부터의 시작비율에서 각 기준년도의 연령, 성별, 인종에 따라 산출된다. 다음 추계된 비율은 전년도비율에 유형별 가구형성의 전년도 변화율을 곱함으로써 구해진다. 유형별 가구수는 각 년도별 인구에 연도별 비율을 곱해서 구해진다.

나. 일본

일본은 전국 및 지역별 가구추계를 위해 가구전이율법 및 가구주율법을 사용하고 있다. 전국가구추계는 1993년까지는 가구주율법, 1998년 이후 최근(2008년)까지는 가구전이율법을 사용하여 추계하고 있다. 지역별 가구추계는 최근(2010년)까지 가구주율법을 적용하고 있다.

일본의 전국가구추계에 적용한 가구전이율법에 대해 개략적으로 설명하면 다음과 같다. 가구전이율법은 생존자를 복수의 상태로 분할해, 상태간의 추이 확률 행렬을 통해 장래의 상태별 인구를 추계하는 방법이다. 추계해야 할 내용은 배우자 관계와 가구내 지위 구성의 조합이다. 센서스 및 5년마다 실시하는 가구동태조사의 가구주에 대해, 규칙에 의해 선정된 마커(대상가구의 참조인)의 지위와 성·배우자 관계의 구성을 남자 12종류, 여자 11종류로 정해서 전이확률행렬에 의해 성·배우자 관계 및 가구내 지위 구성을 추계하였다. 가구전이율법에 의한 전체적인 절차는 [그림 3-2]와 같다.



출처 : Household Projections for Japan:2005-2030(Nishioka et al., 2011)

[그림 3-2] 가구전이율법에 의한 가구추계 업무흐름도

기준 인구는, 국세조사의 남녀별, 5세 계급별, 가족 유형별 가구주수 및 비가구주수를 남녀별, 5세 계급별, 가족유형별 마커수 및 비마커수로 변환해 얻었다. 시설가구원의 전이확률산출은 어려우므로 시설인구의 구성비를 성·연령·혼인상태별로 외삽법을 이용하여 추계하였다.

전이확률에 의해 산출한 마커 및 비 마커의 인구를 기준 인구시 사용한 방법을 역으로 산출하여 가구주와 비가구주 자료를 산출하였다. 이것과 추이 확률 행렬에 의해 구한 잠정적인 일반 가구 인원을 구성하여, 남녀별, 5세 계급별, 배우자 관계와 가구 내 지위(시설을 포함한)의 구성별 인구를 매5년마다 추계했다. 여기로부터 남녀별, 5세 계급별로, 배우자 관계와 가구내 지위(시설을 포함한)의 구성 분포를 구해 그것을 전국 인구의 장래 추계의 중위 추계에 있어서의 남녀·5세 계급별 인구를 곱해서, 남녀·5세 계급·배우 관계·가구내 지위(마커·비마커)별 인구를 얻었다.

연간자료는 5년간격의 추계 결과에 근거해, 선형보간에 의해 산출하였다. 또한 기준 인구 작성시의 가구주·비가구주로부터 마커·비마커에의 변환을 반대로 적용해, 남녀별, 5세 계급별, 배우자 관계별, 가구내 지위(가구주·비가구주) 별 인구를 얻었다. 추계 작업은 [그림 3-2]에 나타난 순서에 따라 이루어졌다.

1) 장래의 배우 관계간 추이 확률의 설정

배우 관계간 추이 확률 추계는, 2005년 센서스 및 동태자료에 기초한 전이확률을 산출한 후 이 행렬을 2000년 인구에 적용하여 2005년 성, 연령 및 혼인상태별 인구가 산출될 수 있도록 확률을 조정하였다. 이 행렬을 기초로 장래의 초혼율, 재혼율 및 이혼율의 변화 등을 고려하여 미래의 전이확률행렬을 추계하였다

2) 시설가구원의 추계

가구동태조사에는 일반가구원의 변동은 포함되어 있으나, 시설가구원의 변동은 포함되어 있지 않아 장래의 시설가구원의 비율은 2000년 및 2005년 성, 연령 및 혼인상태별 시설가구원의 비율을 기초로 이 비율이 선형적으로 감소한다는 가정을 바탕으로 추정했다.



3) 배우관계 및 가구내 지위조합 전이확률행렬의 작성

별거부부를 제외한 일반 가구의 가구내 지위간 추이확률은, 배우 관계간의 각종 추이확률과 가구동태조사에서 관찰된 추이패턴으로부터 산출하였다. 예측된 초혼, 재혼, 이혼, 사별 확률을 가구동태조사에서 관찰된 추이패턴에 적용했다. 남녀·5세 계급별 가구내 지위간 추이확률행렬을 작성했다. 2000년 센서스 자료에서 산출된 전이확률행렬을 적용하여 2005년 혼인상태 및 가구내 지위별 자료를 산출하여 2005년 센서스 자료와 비교하여 전이확률행렬을 조정하였다.

4) 가구주의 성, 연령, 배우관계 및 가족유형별 가구수 산출

배우관계 및 가구내 지위조합 전이확률행렬을 배우관계별 인구에 곱하여 배우관계 및 가구내 지위조합 일반가구원을 산출한 후 마커 및 비마커를 가구주 및 비가구주로 전환한 후 가구주의 성, 연령, 배우관계 및 가족유형별 가구수를 산출한다.

다. 영국

영국(잉글랜드, 스코틀랜드 포함)은 수정지수식을 이용해 추계한 가구주율법을 이용해서 가구추계를 작성하고 있다. 이 때 가구주율은 노동력조사(Labor Force Survey)를 이용하여 보완하고 있다. 잉글랜드의 가구추계 작성단계는 일차적으로 전국 및 지역별로 혼인상태별 추계인구에서 시설가구의 추계인구를 제외하여 일반가구의 추계인구를 확정한다. 혼인상태별 인구는 조성법(Component Method)에 의해 작성하고 있는데, LIPRO 모형을 이용하고 있다. 다음, 각 연령·성·혼인상태 또는 동거상태별 가구주율을 추계하여 일반가구추계인구에 곱하여 일반가구를 추계한다. 지역별 추계는 전국과 독립적으로 작성되지만 전국추계와 일치하도록 조정하였다. 스코틀랜드는 혼인상태별 인구는 추계하지 않고 지역 및 가구유형별 가구수를 추계하여 제공하고 있다.

라. 호주 및 뉴질랜드

호주 및 뉴질랜드는 동거형태의 인구비율을 추계하여 추계인구에 적용한 후 가족 및 동거형태별 가구를 추계하는 propensity method에 의해 장래가구를 추계하고 있다.

1) 호주

호주통계청은 가족 및 가구동거형태에 대한 가구추계자료를 주 및 준 주별로 제공하고 있다. 가족 및 가구동거형태율은 세 가지 가정을 토대로 산출하고 있는데, 2011년 가구추계에 적용된 가정은 다음과 같다.

- 가정 1. 동거형태율은 2006년과 동일하다.
- 가정 2. 동거형태율 변화 감소, 2011년까지는 1991-2006년까지와 동일한 변화율, 2016년까지는 변화율의 1/2, 2021년까지는 변화율의 1/4, 이후 2031년까지는 동일하게 유지된다.
- 가정 3. 1991-2006년간의 동거형태변화율이 지속된다는 가정, 즉 2031년까지 1991-2006년의 경향이 선형적으로 지속된다.

중위가정에 의한 인구추계자료에 동거형태별 비율추계를 적용하였다.

호주의 가구추계단계는 다음과 같다.

- 단계 1 : 4회의 센서스자료를 이용하여 주거단위별 인구분포를 산출한다.
- 단계 2 : 단계1에서 다른 주거단위에 있는 사람들의 동거형태경향이 5세 단위별로 결정된다. 각 가족 및 가구 추계는 동거형태의 미래변화율에 대한 가정에 기초한다.
- 단계 3 : 연령별 동거형태율에 대한 연간변화율을 계산하여 2006년에 관찰된 경향에 적용한다.
- 단계 4 : 추계된 인구에 추계된 동거형태율을 적용하여 5세 연령 계급별 동거형태별 인구를 산출한다.
- 단계 5 : 속주 및 주별 합을 호주 전체 합에 일치하도록 조정한다.
- 단계 6 : 가족 및 가구 수를 계산한다.
- 단계 6.1 : 각 추계연도에 대하여, 가족 및 가구 수는 추계된 동거형태에서부터 계산된다. 부부가족 수는 부부가족의 파트너수의 절반이고, 한부모 가족 수는 남자한부모수와 여자한부모수를 합친 수이다. “다른 가족”수는 “다른 가족”에 속하는 개인을 2006년 센서스의 이 가족유형의 평균크기를 나누어서 산출한다.

단계 6.2: 가족을 가족가구로 전환한다. 가족가구는 하나이상의 가족을 포함한다. 가구 수를 산출하기 위해서는 2006년 센서스로부터 계산된 비율을 이용하여 가족은 가구로 전환한다.

단계 6.3: 비가족가구의 숫자를 산출한다. 비가족가구수는 주거단위로부터 산출 된다. 단독가구의 숫자는 혼자 사는 남자 및 혼자 사는 여자의 합에 의해 계산된다. 2006년 센서스의 평균가구수는 그룹가구원수를 그룹가구수로 나누어 계산한다. 평균가구원수는 추계에서 그룹가구의 수를 계산하는데 사용된다.

단계 6.4: 총가구수를 산출한다.

가족가구 및 비가족가구가 총가구수에 더해진다.

2) 뉴질랜드

뉴질랜드는 호주와 동일하게 동거형태율을 추계하는 propensity method를 적용하여 작성하고 있다. 다만 호주의 경우는 가구형태율에 대한 세 가지 가정으로 세 가지 가구 추계결과를 제시하고 있지만, 뉴질랜드의 경우는 인구에 대한 세 가지 가정 및 가구형태에 대한 두 가지 가정을 포함하여 여섯 종류의 추계결과를 제공하고 있다. 최근의 가구 추계는 2006년 센서스를 기초로 한 2006-2031년 가구추계이다. 인구에 대한 가정 및 동거형태율에 대한 가정은 다음과 같다.

인구가정 1 : 저위출산력, 고위사망력, 저위이동력

인구가정 5 : 중위 출산력, 중위사망력, 중위이동력

인구가정 9 : 고위출산력, 저위사망력, 고위이동력

동거형태율 가정 A : 동거형태율이 2006년 수준에서 일정하게 유지된다.

동거형태율 가정 B : 성, 각세별 1986-2006에 관찰된 율을 평가하여 2006 - 2031년 기간 동안 선형적으로 변한다.

가정 A 및 B에 모두 평균가구원수는 동일하다는 가정을 적용하였다. 추계는 여섯 가지의 가정을 적용하여 산출하였으나, 실제로는 5B가정(중위 출산력, 중위사망력, 중위이동력 및 동거형태율이 선형적으로 변한다는 가정)이 가장 적합하다고 평가하고 있으며, 5B가정에 의한 결과를 주로 제시하고 있다.



동거형태율을 추계한 후 가구를 작성하는 과정은 다음과 같다.

1. 어린이가 없는 부부가구= (어린이가 없는 가구의 남자배우자수 + 어린이가 없는 가구의 여자배우자수)/2
2. 두부모 가구 = (두부모가구의 남자배우자+두부모가구의 여자배우자)/2
3. 한부모 가구 = 한부모 가구의 남자부모 +한부모 가구의 여자부모
4. 가족 가구 = 가족의 구성원수/ 가구당 평균 가구원수
5. 1인 가구= 1인가구의 숫자
6. 다른 다인가구 : 다른 다인가구 / 다른 다인가구의 평균 가구원수

마. 캐나다

Population Extrapolation for Organizational with Less Error (P.E.O.P.L.E) 모델에 의해 산출하며, 소지역별 평균가구원수를 추계한 후 추계인구에 평균가구원수를 곱하여 가구수를 산출한다. 지역별 가구원수는 주택가격, 개인별 소득, 도시화율, 여성의 경제활동 참가율, 혼인율, 이혼율의 영향을 받을 수 있다고 보고, 여러 변수를 고려한 선형회귀식으로 소지역별 가구원수를 추계하였다. 1998년 소지역 가구원수 추계에 적용한 선형회귀식은 다음과 같다.

$$\ln PPH = 0.65 + 0.38 \ln CPOP - 0.14 \ln MPOP - 0.22 \ln SDPOP$$

여기서 PPH = 평균가구원수

CPOP = 19세미만 어린이수

MPOP = 결혼한 인구

SDPOP = 별거 및 이혼 인구

3. 소결

지금까지 살펴본 거시적 동태모형과 미시적 동태모형을 비교해 보면 다음과 같다. 거시-동태 모형이 인구의 분포가 구조적이며 동태사건은 전 인구에 걸쳐서 발생한다고 가정하는 것에 비해, 미시-동태 모형은 해당 사건의 발생 및 시기가 개별적으로 발생한다는 점에 차이가 있다. 또한, 마이크로 데이터에 포함된 개인의 출생·사망·교육·혼



인·이혼 등의 다양한 사건의 경험은 해당 국가에서 발생하는 사건의 확률과 동일하다고 본다(Harding, 1996; Fukawa, 2007).

그러나 각 방법은 한계점을 갖고 있다. 미시-동태모형의 경우 적정한 초기인구를 얻거나 전이확률을 추계 혹은 랜덤수를 이용하는 것에 따른 표본오차 등 여러 가지 요인으로 공식 추계에 사용되는 경우는 드물다. 거시-동태모형 중 다차원 모델의 경우 특별한 표본조사를 이용하지 않는 것으로 제안하였으나, 개인의 특성별 전이확률 생성을 위해 396,900개의 원소를 가진 거대한 전이확률행렬을 필요로 하고, 이를 극복하기 위한 상태 간 독립성가정이 비현실적이라는 단점이 크다. 또한 장래확률의 변화를 예측하는 것이 추계의 중요한 부분이라는 점을 감안한다면, 현재의 확률을 장래에 그대로 이용한다는 점은 이 모델의 단점으로 지적될 수 있을 것이다. 제안자들이 제시하고 있는 유료 소프트웨어인 Profamy를 이용하지 않고는 작성이 불가능하며, 이 소프트웨어의 내부의 프로시저가 검증되지 않았으므로 공식추계에 사용하기 어렵다고 볼 수 있다. 가구전이율법은 가구내 지위변화 작성을 위한 추계행렬을 필수로 한다. 일본의 경우는 이를 위하여 매 5년마다 5년전 가구내 지위 및 현재의 가구내 지위를 조사하는 가구동태조사를 실시하고 있으나, 우리나라의 경우 센서스 자료만으로는 적용의 어려움이 있다.

따라서 본 연구에서는 2012년 장래가구추계 시 미국, 영국, 일본의 지역추계 시 사용 중인 거시-정태 모형의 가구주율법을 적용하고자 한다. 가구주율법은 계산과정이나 자료요구가 다른 방법에 비해 단순하고, 가구구성비 변화 등에 대한 다양한 결과제공이 가능하여 국가차원의 가구추계에 가장 많이 활용되고 있는 방법이다. 물론 이 방법은 2002년 및 2007년 장래가구추계 시 활용된 바 있으나, 당시 적용하였던 혼인상태별 인구추계, 가구주율 추정방법, 가구유형 및 가구원수별 가구구성비 작성 등 각 단계별 작성방법을 검토 개선방향을 찾아 2012년 장래가구추계의 적합도를 높이고자 한다.

제3절 2007년 추계방법 검토 및 2012년 개선안 제안

2007년 장래가구추계 시 가구주율법을 적용하였는데, 이 방법은 특정집단의 인구에 성, 연령, 혼인상태별 집단의 가구주율을 곱하여 장래가구를 추계한 것으로, 구체적인 식은 다음과 같다.

$$H^s(x, t) = \sum_m P^s(m, x, t) \times h^s(m, x, t)$$

여기서, $H_s(x,t)$ 는 연도 t 의 가구주의 성 s , 연령 x 의 가구수,
 $P_s(m,x,t)$ 는 연도 t 의 성 s , 혼인상태 m , 연령 x 의 인구,
 $h_s(m,x,t)$ 는 연도 t 의 성 s , 혼인상태 m , 연령 x 의 가구주율

위 식에 의하면, 혼인상태별 가구를 산출한 후 가구유형별로 가구구성비 및 가구원 수별 가구구성비를 추계하여 세부특성별로 가구자료를 산출한다. 따라서 가구주율법의 가구추계의 적합도는 혼인상태별 인구추계단계, 혼인상태별 가구수 추계단계, 각 가구 구성별 가구구성비 추계단계에 의해 영향을 받는다. 2007년 장래가구추계 작성절차는 [그림 3-3]과 같다.



[그림 3-3] 장래가구추계 작성 업무 흐름도

1. 2007년 장래가구추계 작성현황

가. 혼인상태별 인구추계

혼인상태별 인구는 혼인상태별 전이확률을 이용한 조성법에 의해 추계하였다. 추계된 혼인상태별 인구구성비를 장래인구추계 결과에 적용하여 장래인구추계결과와 일관성을 갖는 혼인상태별 인구를 추계한 것이다. 이 방법에 의한 혼인상태인구추계는 혼인

의 생애주기 변화 파악이 가능하고, 혼인상태의 변동을 반영한다는 점에서 혼인상태구성비를 이용한 외삽법만으로 추계하는 정태적 혼인상태인구추계에 비해 강점을 갖고 있다.

① 1단계(연령별 변동율)

혼인상태변동 발생율을 총조사 및 인구동태자료로부터 작성($nM_x(ij)$)하고, 변동율($n\pi_x(ij)$)과 관찰치 $nM_x(ij)$ 가 같다는 가정에서 변동율을 근사적으로 산출한다.

- 미혼인구의 초혼, 사망률
- 유배우인구의 이혼율, 사별율, 사망률
- 사별인구의 재혼율, 사망률
- 이혼인구의 재혼율, 사망률

② 2단계(상태간 전이확률)

- 배우관계상태 i 에서 j 로의 변동확률: x 에서 $x+n$ 세까지의 사건 발생이 선형이라고 가정하여 아래의 식으로 변환하고

$${}_n\pi_x^{ij} = \frac{n \times {}_n m_x^{ij}}{1 + \frac{n}{2} \times \sum_{i \neq j}^{k+1} {}_n m_x^{ij}}$$

${}_n m_x^{ij}$: n 연령구간 중 혼인상태변화율

${}_n \pi_x^{ij}$: n 연령구간 중 혼인상태전이확률

- 배우관계상태 i 에서 i 로 남아있을 확률

$${}_n\pi_x(ii) = \frac{1 - \frac{n}{2} \times \sum_{i \neq ij}^{k+1} {}_n m_x(ij)}{1 + \frac{n}{2} \times \sum_{i \neq j}^{k+1} {}_n m_x(ij)}$$

혼인상태별 전이확률을 각 세 기준으로 산출한 후, 각 세별 확률자료의 불안정을 보완하기 위해 여러 가지 보정방법을 사용하였다. 각 세별 자료를 3세 이동 평균을 이용하



여 보정하였으며, 저 연령 및 고 연령에서는 자료에 따라 지수모형, 로그모형 등을 이용하여 개별적 보정을 실시하였다.

③ 3단계(장래의 혼인상태별 전이확률 추계)

장래의 혼인상태별 인구산출을 위해서는 장래의 혼인상태별 전이확률 추계가 필요한데, 이를 위해서는 기준년도의 혼인상태별 전이확률변화율을 고려해야 한다. 대상기간 동안(1990년~2005년)의 혼인상태전이확률 변화율을 고찰하여 대상기간의 평균변화율 또는 변화율이 지속적으로 감소한다는 가정 등을 적용하여 추정하였다 즉 1990년 이후의 연령구간별 초혼, 재혼, 이혼확률 변화율을 검토하여 매 5년마다 변화율이 1/2로 감소한다는 가정 및 2020년 이후는 변화율이 동일하다는 가정으로 추계하였다. 단, 연도별 증감 변화폭이 큰 사별 후 재혼확률 추정을 위해서는 연도별 평균 변화율을 적용하였다.

혼인상태별 장래 사망확률 산출은 1990-2005년 혼인상태별 사망확률에 일관적인 차이가 있다는 점을 고려하여 전체사망확률에 대한 혼인상태별 사망확률비를 적용하였다. 뿐만 아니라, 각세 사망확률비 산출시에도 각세 연령별 불안정을 보완하기 위해 3세보정, 일부연령 선형보정, 지수함수 보정 등 각 자료에 대한 개별적 보정을 실시하였다. 즉 전체사망확률에 대한 혼인상태별 사망확률비를 산출하여 장래추계인구의 혼인상태별 사망확률에 적용하여 각세별 혼인상태별 사망확률을 추계하였다.

④ 4단계(혼인상태별 인구추계)

혼인상태별 전이확률을 기준 혼인상태별 인구에 곱하여 혼인상태간 변동계수를 산출하고, 혼인상태간 변동계수를 가감하여 다음연도의 혼인상태별 인구를 산출하였다. 혼인상태별 인구산출과정은 다음과 같다.

- t 년 $x+1$ 세 미혼인구 : $P_{x+1}^s = P_x^s \cdot (1 - \pi_x^{sm} - \pi_x^{sd})$
- t 년 $x+1$ 세 유배우인구 : $P_{x+1}^m = P_x^m \cdot (1 - \pi_x^{mw} - \pi_x^{mv} - \pi_x^{md}) + P_x^s \cdot \pi_x^{sm} + P_x^w \cdot \pi_x^{wm} + P_x^v \cdot \pi_x^{vm}$
- t 년 $x+1$ 세 사별인구 : $P_{x+1}^w = P_x^w \cdot (1 - \pi_x^{wm} - \pi_x^{wd}) + P_x^m \cdot \pi_x^{mw}$
- t 년 $x+1$ 세 이혼인구 : $P_{x+n}^v = P_x^v \cdot (1 - \pi_x^{vm} - \pi_x^{vd}) + P_x^m \cdot \pi_x^{mv}$

P_x^s : 미혼인구	P_x^m : 유배우인구	P_x^w : 사별인구	P_x^v : 이혼인구
π_x^{sm} :미혼인구초혼확률	π_x^{mw} :유배우인구사별확률	π_x^{wm} :사별인구재혼확률	π_x^{vm} :이혼인구재혼확률
π_x^{sd} :미혼인구사망확률	π_x^{mv} :유배우인구이혼확률	π_x^{md} :사별인구사망확률	π_x^{md} :이혼인구사망확률
	π_x^{md} :유배우인구사망확률		

이후 작성된 혼인상태별 인구로부터 혼인상태별 인구구성비를 산출하고, 장래인구추계에 적용하여 혼인상태별 인구를 산출하였다.

나. 혼인상태별 장래가구주를 추계

혼인상태별 가구주율 추정에는 가구주율의 변화 형태를 수학적 모형을 적용, 외삽법으로 연장하는 방식이 주로 이용되어 왔다. 2002년 추계 시에는 이 때, 연령별로 다른 방식을 적용하였는데 35세 이상에서는 동일한 연령 및 성, 코호트의 가구주 간 생존율로 작성하는 순천이율법을 적용하였고, 34세미만은 선형로그식을 사용하였다. 그러나, 2007년 장래가구추계에서는 순천이율 적용시 시도 및 각세별 가구주율이 비정상적인 시계열을 보임에 따라 전 연령에 선형로그식을 사용하였다(통계청, 2007).

다. 가구원수별 가구구성비, 가구유형별 가구구성비 추계

1인~ 7인 이상 가구, 부부 및 한부모 가구 등 17개 가구유형별 가구구성비를 장래가구주율과 동일하게 선형로그식을 이용하여 추계하였다.

참고로 혼인상태별 인구구성비 추정, 가구주율 추정, 세대구성별 구성비 추정에 이용되고 있는 수학적모형은 다음과 같다.

- 1) 수정지수식
 - i) 비율이 증가할 때,

$$h^s(m, x, t+k) = 1 - [1 - h^s(m, x, t-n)] \times \left[\frac{1 - h^s(m, x, t)}{1 - h^s(m, x, t-n)} \right]^{\frac{-(t+k)-(t-n)}{n}}$$



ii) 비율이 감소할 때,

$$h^s(m, x, t+k) = h^s(m, x, t-n) \times \left[\frac{h^s(m, x, t)}{h^s(m, x, t-n)} \right]^{\frac{-(t+k)-(t-n)}{n}}$$

여기서, $h^s(m, x, t)$ 는 연도 t 의 성 s , 혼인상태 m , 연령 x 의 가구주율

2) 선형로그식

$$h^s(m, x, t) = a \times \text{Ln}(n) + b$$

3) 순천이율

$$C_i^s(x, t) = \frac{H_i^s(x+5, t+5)}{H_i^s(x, t)} - \frac{P^s(x+5, t+5)}{P^s(x, t)}$$

여기서, $C_i^s(x, t)$ 는 연도 t 의 가구형태 i , 성 s , 연령 x 의 순천이율, $H_i^s(x, t)$ 는 연도 t 의 가구형태 i , 성 s , 연령 x 의 가구수, $P_i^s(x, t)$ 는 연도 t 의 가구형태 i , 성 s , 연령 x 의 인구수

4) 로지스틱모형

$$\ln(h^s(m, x, t)/(1 - h^s(m, x, t))) = a \times n + b$$

5) 시계열 모형

Arima 모형 및 Random Walk 모형이 주로 사용되고 있다. 미국의 1996년 가구원 구성비 추계 시에 Arima(0,2,1)모형을 각 연령군에 개별적으로 적용하였다. 그러나 혼인상태의 개별연령군에 각각 시계열 모형을 적용할 경우 일부 연령에는 적합한 모형을 찾을 수 없는 경우가 있고, 추계작업이 지나치게 복잡해지는 경향이 있다. 이런 한계를 피하기 위해 연령별 평균 및 연령별 평균변화율을 산출한 후 연도별 변화만을 시계열 모형으로 추정하는 모델인 단일변수에 의한 시계열 모형을 적용할 수도 있으나, 연령별 증감의 방향이 다를 경우 연령별 오차의 상쇄로 현재자료에서 변동이 없는 미래자료를 산출하는 결과를 줄 수 있다. 단일변수 시계열모형은 아래의 식으로 표시할 수 있으며, k_i 를 시계열 모형에 의해 산출한다.

$$\log(h^s(m, x, t)) = a_x + b_x * k_t$$

a_x : 연령별 평균

b_x : 연령별 평균변화율

k_t : 연도별 변화율



2. 혼인상태별 인구 추계방법 검토 및 개선안

가. 혼인상태별 인구구성비 추이(1990-2010)

혼인상태별 인구추계 작성방법을 검토하기에 앞서 1990년 이후의 혼인상태별 인구구성비의 추이를 살펴보고자 한다. 혼인상태별 인구추이가 성 및 연령대별로 어떻게 변화하고 있는지를 <표 3-2>와 [그림 3-4]를 통해서 제시하였다. 총인구에서 미혼인구가 차지하는 비율은 1990년 이래로 지속적으로 감소하고 있었으며, 그 감소폭이 2005~2010년간 큰 것으로 나타났다. 최근 미혼율이 증가하고 있는 것으로 알려져 있는 점을 고려해 볼 때 출산율 감소에 따른 혼인연령인구의 감소에 의한 영향임을 알 수 있다. 마찬가지로 유배우 인구 비중 또한 줄어들고 있다. 반면, 사별 및 이혼인구의 비중은 2005~2010년간 증가하였다. 특히, 사별인구 구성비는 2010년에 큰 폭으로 증가하였는데, 이는 기대수명의 증가로 사별한 이후 생존연수가 증가함에 따라 전체인구에서 차지하는 비중의 증가로 나타난 것이다.

<표 3-2> 혼인상태별 인구구성비 추이(1990-2010)

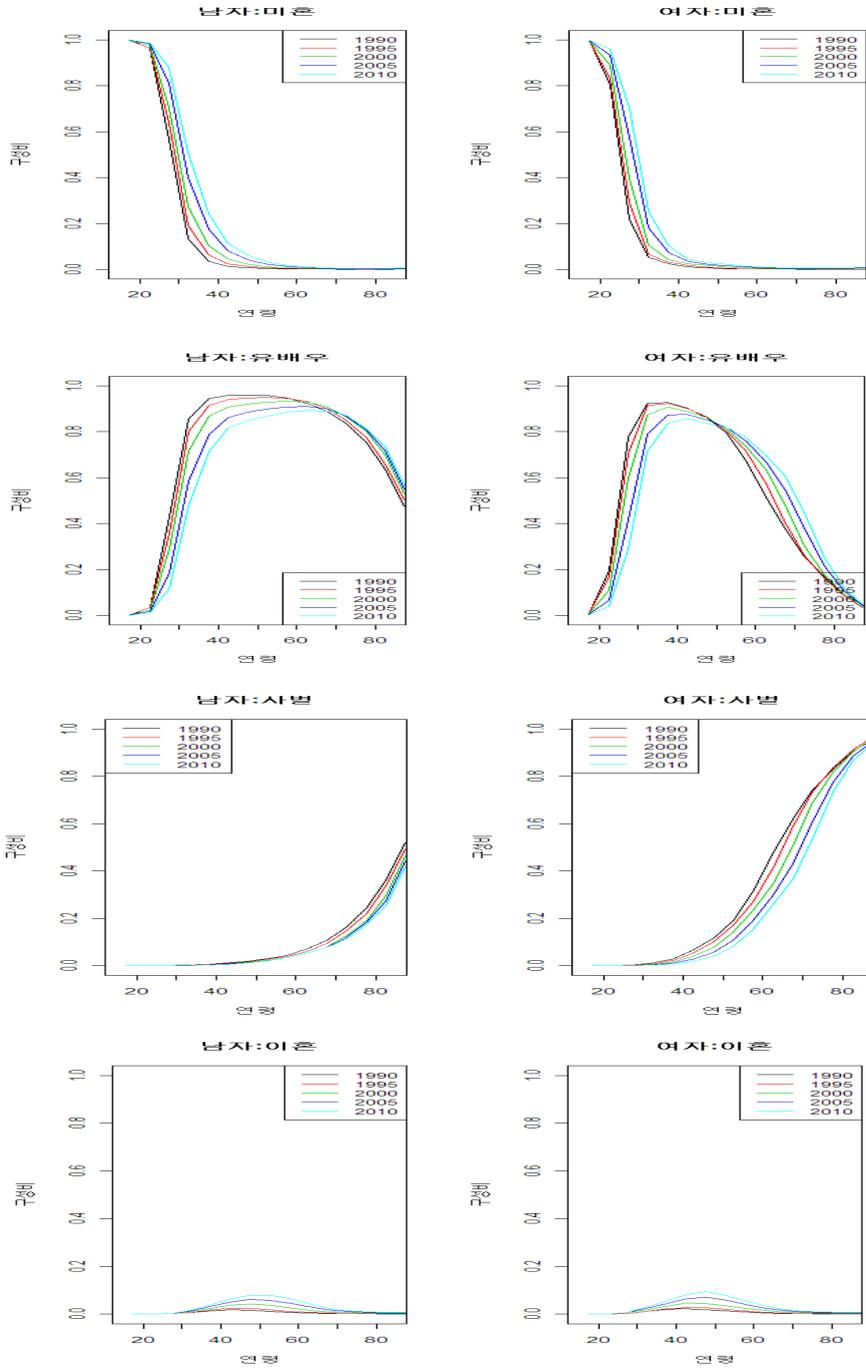
(단위 : %)				
연도	미혼	유배우	사별	이혼
1990	33.08	58.95	7.21	0.76
1995	30.79	60.68	7.49	1.04
2000	29.99	60.62	7.49	1.90
2005	29.95	59.49	7.63	2.93
2010	25.83	56.89	12.92	4.36

다음은 1990년 이후 혼인상태별 인구구성비가 연령대별로 어떻게 다른지 살펴본 것이다. [그림 3-4]에 의하면 미혼인구구성비는 1990~2010년 사이 남녀 모두 큰 폭으로 증가하고 있음을 알 수 있다. 특히, 남자는 25세 초반에서 55세까지 여자는 45세까지 미혼인구 구성비가 지속적으로 증가하고 있으며, 특히 2000-2005년에 그 비중이 가장 큰 폭으로 증가하였음을 보여주고 있다. 이는 출산율 저하의 직접적인 원인으로 꼽히기도 한다.

반면 연령대별 유배우인구의 구성비는 역 U자형과 유사한 패턴인데, 1990~2010년간의 시계열 변화를 보면, 남자는 70세 이전에, 여자는 50세 이전에 지속적인 감소를 보이는 것으로 나타났다. 이러한 감소 경향은 미혼율 및 사별 또는 이혼율의 증가와 관계가 있는 것으로 보인다. 한편, 남자의 경우 70세 이후에는 유배우 인구구성비가 약간 증가하다가 2005~2010년은 거의 변화가 없는 것으로 나타나고 있으며, 여자의 경우에는 50세~80세까지의 유배우 인구구성비가 지속적으로 증가하고 있다. 이는 기대수명의 변화와 관계되는데, 남자 기대수명의 지속적인 증가로 인해 여자의 유배우구성비는 증가하는 반면, 여자 기대수명의 증가 둔화는 남자 유배우구성비의 소폭증가 및 정체로 나타나고 있다.

연령대별 사별인구 구성비를 보면 고연령층으로 갈수록 구성비가 급속하게 증가하는 패턴을 보여주고 있는데, 그 추세를 보면 남자와 여자 모두 구성비가 1990~2010년간 급속하게 감소하고 있는 것으로 나타났다. 그러나 감소폭에 있어서는 남자가 여자에 비해 적으며, 특히 2000년 이후에는 거의 변화가 없는 것으로 나타난다. 이는 앞서 유배우 인구 구성비에서 언급한 기대수명의 변화와 관련된다. 남자의 기대수명증가가 여자 사별인구의 감소를 상쇄하고 있으며, 반면에 여자의 기대수명 증가폭이 남자에 비해 둔화되고 있는 점은 남자 사별인구의 적은 감소폭에 대한 원인으로 설명할 수 있을 것이다.

혼인상태별 인구구성비 추이를 살펴본 결과, 1990~2010년간 미혼 및 유배우 인구는 감소, 사별 및 이혼인구의 비중은 증가하는 양상을 보여주었다. 그리고 이러한 혼인상태별 구성비는 연령 및 성에 따라서도 차별적으로 나타났다. 따라서 혼인상태인구추계의 타당성을 높이기 위해서는 이러한 변화율에 대한 합리적인 추계가 선행되어야 함을 알 수 있다.



[그림 3-4] 혼인상태별 인구구성비 추이(1990~2010)

나. 2007년 혼인상태별 인구추계의 타당성 검토

다음으로 2007년 장래가구추계 시 작업한 혼인상태별 인구추계 결과의 타당성을 검토하고자 한다. <표 3-3>은 2010년 총조사 혼인상태별 인구구성비와 2007년 추계한 2010년 혼인상태별 인구구성비의 오차율을 나타내고 있다. 혼인상태별 인구추계 오차율은 평균 절대오차율로 측정했을 때 남자 4.7%, 여자는 3.4%로 남녀 모두 5%미만으로 나타났다.

다만 남녀모두 미혼율은 과소추계 유배우율은 과대추계 되었는데, 이는 미혼율 증가 폭이 완화되리라는 가정이 적합하지 않았음을 나타낸다. 또한, 사별율은 과소추계 된 것으로 나타났는데 이는 인구추계의 사망률이 높게 추계되었음을 일부 설명하고 있으며, 동시에 인구추계 시 사망률 추계의 개선이 사별인구구성비 개선으로 반영될 수 있음을 시사한다. 남녀의 이혼인구구성비의 다른 오차방향은 이혼율 추계에 의한 영향보다는 다른 혼인상태인구구성비 추계 차이에 의한 영향으로 보인다.

<표 3-3> 혼인상태별 인구추계 오차율

(단위 : %)

	평균절대오차율 ¹⁾	미혼	유배우	사별	이혼
남자	4.7	-4.1	2.3	-5.9	6.6
여자	3.4	-4.6	2.7	-0.8	-5.3

1) $\frac{1}{n} \sum |((m_i - M_i)/M_i) * 100|$ 여기서 M_i 는 총조사 혼인상태 구성비, m_i 는 추계 혼인상태구성비

〈표 3-4〉 연령별 혼인상태별 인구구성비 비교(2010)

〈미혼〉

(단위 : %)

	남자				여자			
	총조사 (A)	추계 (B)	차이 (B-A)	오차 ((B-A)/A) *100	총조사 (A)	추계 (B)	차이 (B-A)	오차 ((B-A)/A) *100
계	35.6	34.1	-1.5	-4.1	25.8	24.6	-1.2	-4.6
15~19세	99.7	99.9	0.2	0.2	99.6	99.7	0.0	0.0
20~24세	98.8	98.7	-0.1	-0.1	96.0	94.3	-1.7	-1.7
25~29세	85.4	85.5	0.1	0.1	69.3	63.8	-5.5	-7.9
30~34세	50.2	46.9	-3.2	-6.5	29.1	24.7	-4.4	-15.1
35~39세	26.7	22.8	-4.0	-14.8	12.6	11.2	-1.4	-10.9
40~44세	14.4	12.7	-1.7	-11.6	6.2	5.9	-0.3	-4.8
45~49세	7.5	6.4	-1.1	-14.9	3.3	3.0	-0.3	-8.8
50~54세	4.1	3.4	-0.8	-18.3	2.3	2.0	-0.3	-12.4
55~59세	2.3	1.8	-0.6	-24.2	1.7	1.4	-0.3	-18.1
60~64세	1.3	1.0	-0.3	-20.3	1.2	0.9	-0.2	-19.5
65~69세	0.8	0.6	-0.2	-22.5	0.8	0.6	-0.2	-25.0
70~74세	0.5	0.4	-0.1	-26.6	0.6	0.4	-0.2	-28.9
75~79세	0.4	0.3	-0.1	-29.0	0.6	0.3	-0.2	-37.9
80~84세	0.3	0.2	-0.1	-32.1	0.6	0.3	-0.3	-44.5
85세이상	0.5	0.3	-0.2	-37.2	0.8	0.4	-0.3	-41.3

〈유배우〉

(단위 : %)

	남자				여자			
	총조사 (A)	추계 (B)	차이 (B-A)	오차 ((B-A)/A) *100	총조사 (A)	추계 (B)	차이 (B-A)	오차 ((B-A)/A) *100
계	58.6	60.0	1.3	2.3	56.9	58.4	1.5	2.7
15~19세	0.2	0.1	-0.2	-77.4	0.3	0.3	0.0	0.6
20~24세	1.1	1.2	0.1	10.6	3.9	5.4	1.5	38.5
25~29세	14.3	14.0	-0.2	-1.7	30.0	35.1	5.1	16.8
30~34세	48.6	51.1	2.4	5.0	68.5	72.2	3.7	5.4
35~39세	69.9	72.9	2.9	4.2	81.6	83.3	1.7	2.1
40~44세	79.4	80.6	1.2	1.5	84.0	85.7	1.7	2.0
45~49세	84.0	85.1	1.1	1.3	83.3	85.1	1.7	2.1
50~54세	86.1	87.3	1.2	1.4	80.7	82.7	2.0	2.4
55~59세	87.7	88.9	1.2	1.3	77.2	78.5	1.3	1.7
60~64세	88.7	89.8	1.1	1.2	70.2	70.4	0.1	0.2
65~69세	88.5	89.5	1.0	1.1	59.5	59.3	-0.2	-0.4
70~74세	86.9	87.6	0.7	0.8	46.3	46.4	0.1	0.3
75~79세	83.0	83.0	-0.1	-0.1	29.6	28.2	-1.4	-4.8
80~84세	74.9	75.3	0.3	0.4	14.5	11.3	-3.2	-22.4
85세이상	59.9	54.6	-5.2	-8.7	6.1	1.3	-4.7	-78.2

〈사별〉

(단위 : %)

	남자				여자			
	총조사 (A)	추계 (B)	차이 (B-A)	오차 ((B-A)/A) *100	총조사 (A)	추계 (B)	차이 (B-A)	오차 ((B-A)/A) *100
계	2.2	2.1	-0.1	-5.9	12.9	12.8	-0.1	-0.8
15~19세	0.0	0.0	0.0	-100.0	0.0	0.0	0.0	-100.0
20~24세	0.0	0.0	0.0	-94.4	0.0	0.0	0.0	-82.3
25~29세	0.0	0.0	0.0	-2.2	0.1	0.0	0.0	-11.0
30~34세	0.1	0.1	0.0	-2.8	0.2	0.2	0.0	-7.6
35~39세	0.2	0.2	0.0	-12.4	0.7	0.6	-0.1	-16.9
40~44세	0.5	0.4	-0.1	-14.4	1.8	1.6	-0.2	-13.0
45~49세	1.0	0.8	-0.2	-16.0	4.0	3.6	-0.4	-9.9
50~54세	1.8	1.6	-0.2	-13.2	7.8	7.2	-0.7	-8.7
55~59세	2.9	2.6	-0.3	-11.0	14.1	13.5	-0.6	-3.9
60~64세	4.7	4.0	-0.7	-14.8	24.0	23.9	0.0	-0.1
65~69세	7.0	6.5	-0.6	-8.0	36.8	37.2	0.4	1.1
70~74세	10.5	10.0	-0.5	-4.8	51.6	51.5	0.0	0.0
75~79세	15.4	15.6	0.2	1.1	68.9	70.5	1.6	2.3
80~84세	23.9	23.6	-0.3	-1.4	84.3	87.8	3.5	4.1
85세이상	39.1	44.4	5.4	13.7	92.7	97.8	5.0	5.4

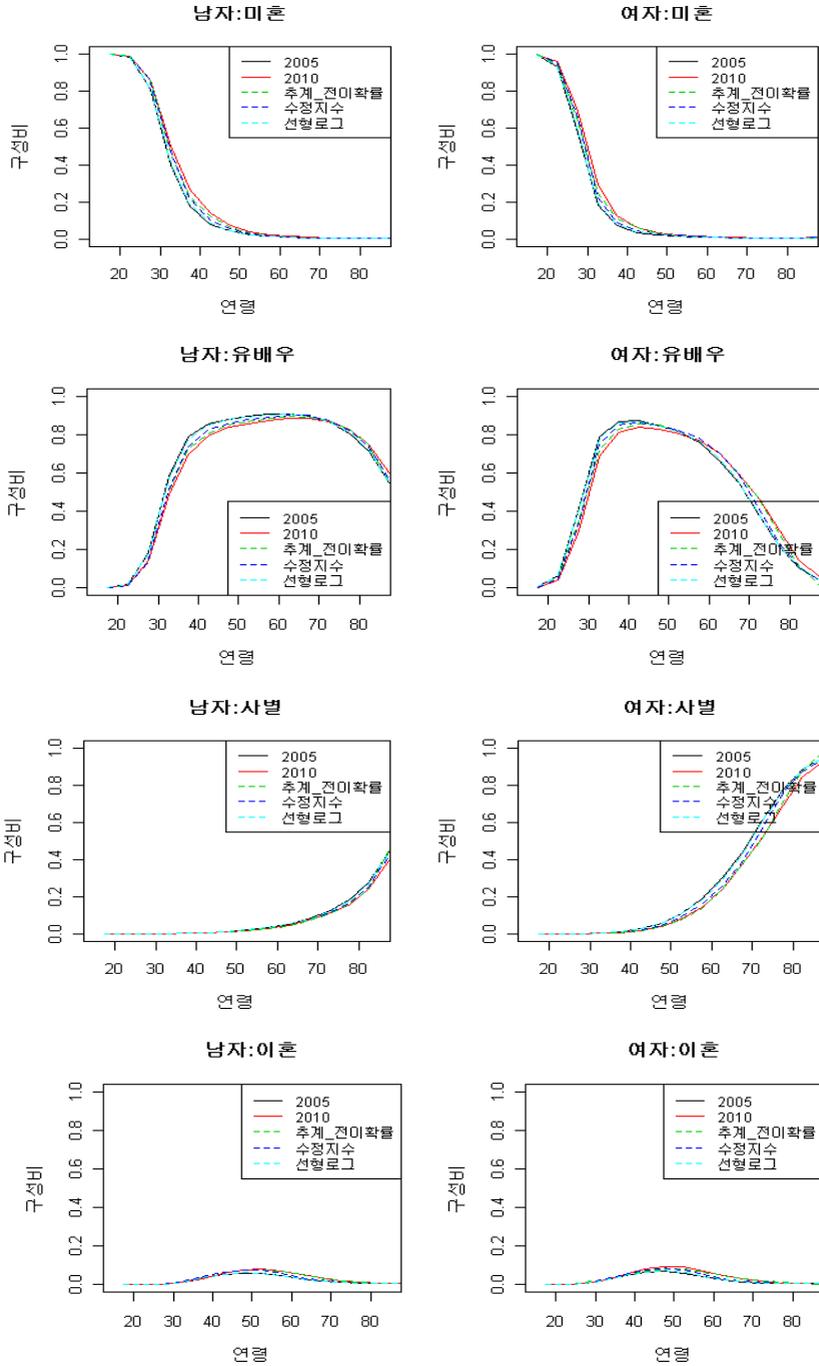
〈이혼〉

(단위 : %)

	남자				여자			
	총조사 (A)	추계 (B)	차이 (B-A)	오차 ((B-A)/A) *100	총조사 (A)	추계 (B)	차이 (B-A)	오차 ((B-A)/A) *100
계	3.6	3.9	0.2	6.6	4.4	4.1	-0.2	-5.3
15~19세	0.0	0.0	0.0	-98.4	0.0	0.0	0.0	-49.3
20~24세	0.0	0.1	0.0	51.9	0.1	0.3	0.2	200.8
25~29세	0.3	0.4	0.2	54.1	0.6	1.1	0.4	68.1
30~34세	1.1	1.9	0.8	71.8	2.2	2.9	0.7	34.2
35~39세	3.2	4.2	1.1	33.3	5.1	4.9	-0.2	-4.2
40~44세	5.8	6.3	0.5	8.7	8.0	6.9	-1.1	-14.2
45~49세	7.5	7.7	0.2	2.1	9.4	8.3	-1.1	-11.3
50~54세	8.0	7.8	-0.2	-2.9	9.1	8.1	-1.0	-10.9
55~59세	7.0	6.8	-0.3	-3.9	7.0	6.6	-0.4	-5.9
60~64세	5.3	5.2	-0.1	-2.2	4.7	4.8	0.1	2.3
65~69세	3.7	3.4	-0.2	-6.8	2.9	2.9	0.0	0.2
70~74세	2.1	2.0	-0.1	-3.5	1.6	1.6	0.1	4.1
75~79세	1.2	1.2	0.0	-0.2	0.9	1.0	0.0	2.3
80~84세	0.8	0.9	0.1	16.7	0.6	0.6	0.0	5.1
85세이상	0.6	0.6	0.1	10.3	0.4	0.4	0.0	0.7

다. 혼인상태별 인구추계 방법별 적합도 비교

혼인상태별 인구추계의 방법별 적합도를 비교하기 위하여, 2007년 추계 시 적용하였던 전이확률법에 의한 추계와 수학적모형(수정지수 및 선형로그)에 의한 결과를 제시하였다. 모형별 일관적 비교를 위해 1990-2005년 자료를 투입하였다 [그림 3-5].



[그림 3-5] 혼인상태별 인구 추계방법별 구성비

[그림 3-5]의 모형별 혼인상태 인구 구성비를 비교해 보면, 미혼, 유배우, 사별, 이혼 모두 전이확률법에 의한 추계가 2010년에 관찰된 혼인상태별 인구구성비에 가장 근접한 것으로 보이고 있으며, 다음으로 수정지수에 의해 추계된 구성비가 2010년 관측치와의 차이가 적었다. 2010년 총조사 자료와 가장 차이가 큰 모형은 선형로그에 의한 추계 결과로 나타났다. 세부적으로 살펴보면, 선형로그방법은 총조사 자료에 비해 미혼인구 구성비는 전 연령에서 낮게, 사별인구 구성비는 전 연령에서 높게, 이혼인구구성비는 40대 이후 연령에서 낮게 추정하고 있다. 이는 연도에 로그값을 취하면서 구성비의 연도별 변화율을 급격하게 감소시키는 선형로그식의 특성이 반영된 것으로, 혼인상태별 변화추이가 지속되고 있는 최근의 가구관련 지표의 특성을 고려할 때 선형로그식은 적합하지 않을 수 있음을 시사한다.

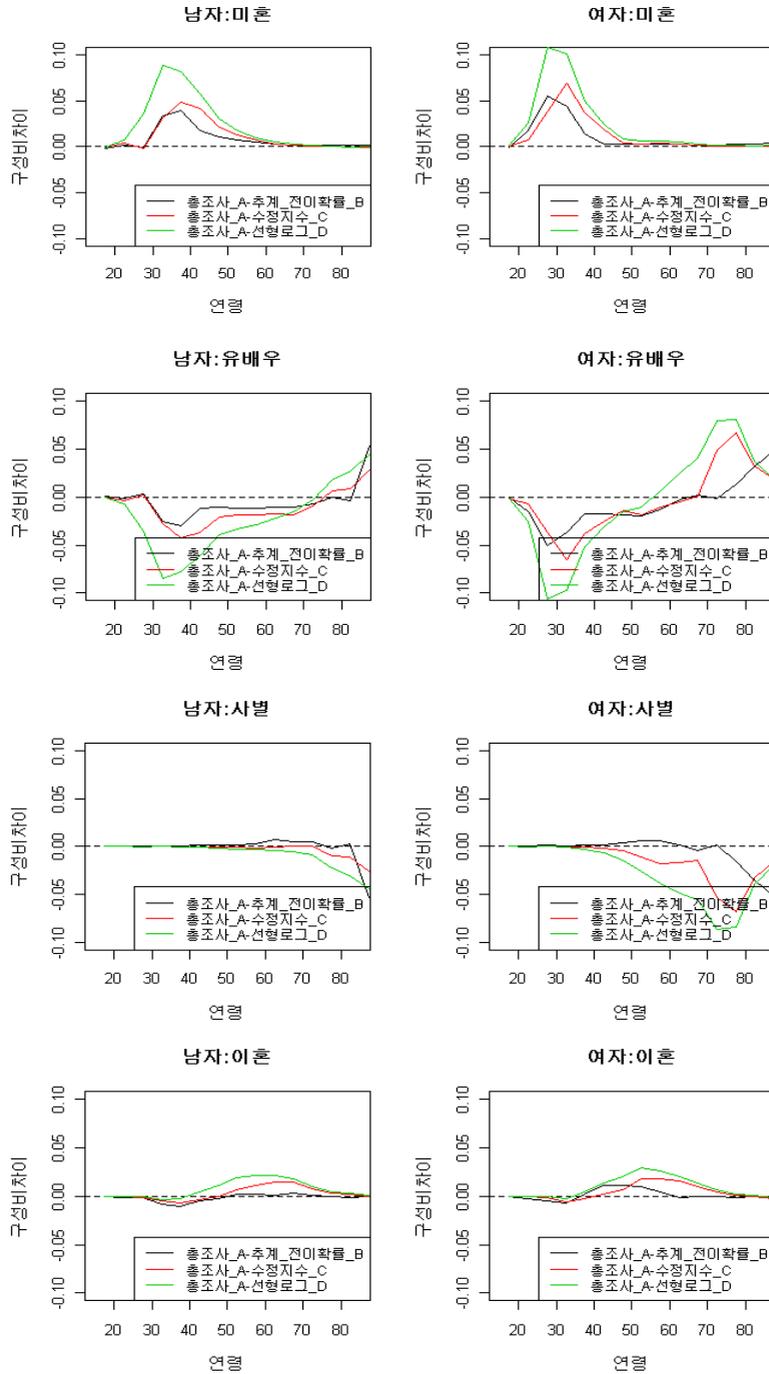
[그림 3-6]에서는 구체적으로 2010년 혼인상태별 인구구성비에 대한 2010년 총조사 결과와 2007년 추계 결과의 차이를 제시하였다. 모든 혼인상태에서 방법별 오차의 방향은 유사하나, 오차의 크기는 차이가 있었다. 모든 혼인상태에서 전이확률모형에 의한 오차가 가장 작고, 다음으로 수정지수모형이며 선형로그모형에 의한 오차가 가장 크게 나타났다. 그러나, 이러한 오차의 크기는 성 및 연령에 따라서도 차이가 있었다.

미혼인구구성비의 경우에는 특히 20-50대 연령에서 모든 방법에서 낮게 추계되었는데, 이는 이 연령층의 미혼율 증가가 추계에 덜 반영되어진 결과이며, 향후 미혼전이확률 추정에서 참고하여 경향을 반영해야 할 것으로 판단된다.

유배우인구구성비는 세 방법 모두 남자는 70세 미만에서, 여자는 60세 미만에서 총조사인구에 비해 높게 추정되었다. 그러나 여자의 경우 60세 이상에서는 낮게 추정되고 있다. 저 연령층의 유배우인구 과대추정과 관련하여, 선형로그모형 등이 현재까지 변화가 급격하게 진행되고 있는 혼인상태인구구성비 추계에는 적절하지 않음을 시사한다. 이는 선형로그모형의 경우 미혼율 증가가 과거 변화를 보다 크게 둔화될 것을 가정하고 있기 때문이다. 전이확률모형 또한 미래의 확률 증가율 또는 감소율이 크게 둔화되리라고 가정할 경우 단기간의 추계 불일치를 가져올 수 있음을 보여준다.

사별 및 이혼구성비에서도 선형로그식에 의한 구성비가 총조사 자료에 비해 가장 차이가 큰 것으로 나타났고, 전이확률법에 의한 구성비가 남자 85세 이상, 여자 75세 이상 사별인구에서의 오차를 제외하고는 전 연령에서 적합도가 매우 높은 것으로 나타나고 있다. 고연령층에서 사별인구가 과대추정된 것은 이 연령층 남자의 사망확률 감소에 따른 기대여명증가가 추계에 적게 반영된 것으로 보인다.

전이확률법에 의한 혼인상태 인구구성비는 전 연령에서 오차가 가장 적으나, 85세 이상에서는 수정지수모형 및 선형로그모형에 비해 오차가 두드러지게 크게 나타나고 있다. 이는 최종 연령의 전이확률이 과소 또는 과대 추계되었음을 의미한다. 이의 개선을 위해 혼인상태별 전이확률 작성 최종 연령을 현재 85세 이상에서 100세 이상으로 연장하여 고령인구의 혼인상태 전이확률을 세밀하게 추정할 필요가 있는 것으로 판단된다.



[그림 3-6] 혼인상태별 인구추계방법별 구성비 차이

<표 3-5>는 각 작성방법별 2010년 총조사 혼인상태 인구구성비와의 적합도 검토를 위해 제공근오차(RMSD; Root Mean Square Deviation) 수치를 제시한 것이다. 그런데, 자료의 투입기간에 따라 모형 적합성이 달라질 수 있으므로, 비교적 장기간의 자료를 활용해야 하는 전이확률모형은 1990-2005년 자료를 활용하고, 수학적 모형(수정지수 및 선형로그)은 1990~2005년 자료 뿐 아니라 최근 5년간의 자료에 대한 추가비교를 실시하였다. [그림 3-5] 및 [그림 3-6]에서도 나타난 것과 같이 1990-2005년 자료에 대한 모형적합도를 비교해 보면 전이확률법에 의한 오차가 남녀 모두 가장 작았다. 선형로그모형에 의한 오차가 가장 커, 적합도가 가장 낮은 것으로 나타났다.

<표 3-5> 혼인상태별 인구구성비 추계방법별 제공근오차(RMSD¹⁾)

구분	전이확률 ²⁾	1990-2005		2000-2005	
		수정지수	선형로그	수정지수	선형로그
남자	0.1060	0.1157	0.2223	0.0828	0.1717
여자	0.1376	0.1864	0.3039	0.1070	0.2263

1) $\sqrt{\sum_{x,i} E(m_{xi} - M_{xi})^2}$ M_{xi} 는 총조사 연령별 혼인상태별 인구구성비, m_{xi} 는 추계방법별 연령별 혼인상태별 인구구성비

2) 1990-2005년 자료 이용

반면, 최근 5년 자료를 활용한 분석결과에서는 수정지수식의 적합도가 최근 15년 자료를 활용한 전이확률모형에 비해 적합도가 약간 높은 것으로 나타났다. 그러나 다양한 동태자료를 활용하고, 다른 혼인상태간의 변동을 장래의 혼인상태 구성비 추계에 반영하는 동태적 전이확률모형이 장기추계에 보다 합리적 모형으로 판단됨에 따라 전이확률모형을 주 모형으로 하고자 한다. 다만, 최근 5년 자료 수정지수 모형을 단기추계 타당도 비교를 위해 활용할 것이다. 또한, 전이확률모형을 이용하더라도 각세자료의 불안전성 보정 방법 개선, 고 연령 및 저 연령 전이확률 작성방법 개선, 장래전이확률 추정을 위한 모형 적용 등의 방법 개선이 혼인상태인구추계의 정확성을 향상시키기 위해 필요한 것으로 보인다. 다음에서는 이 부분에 대해 추가적으로 검토해 보고자 한다.



라. 혼인상태별 전이확률 보정 및 추계방법 개선사항

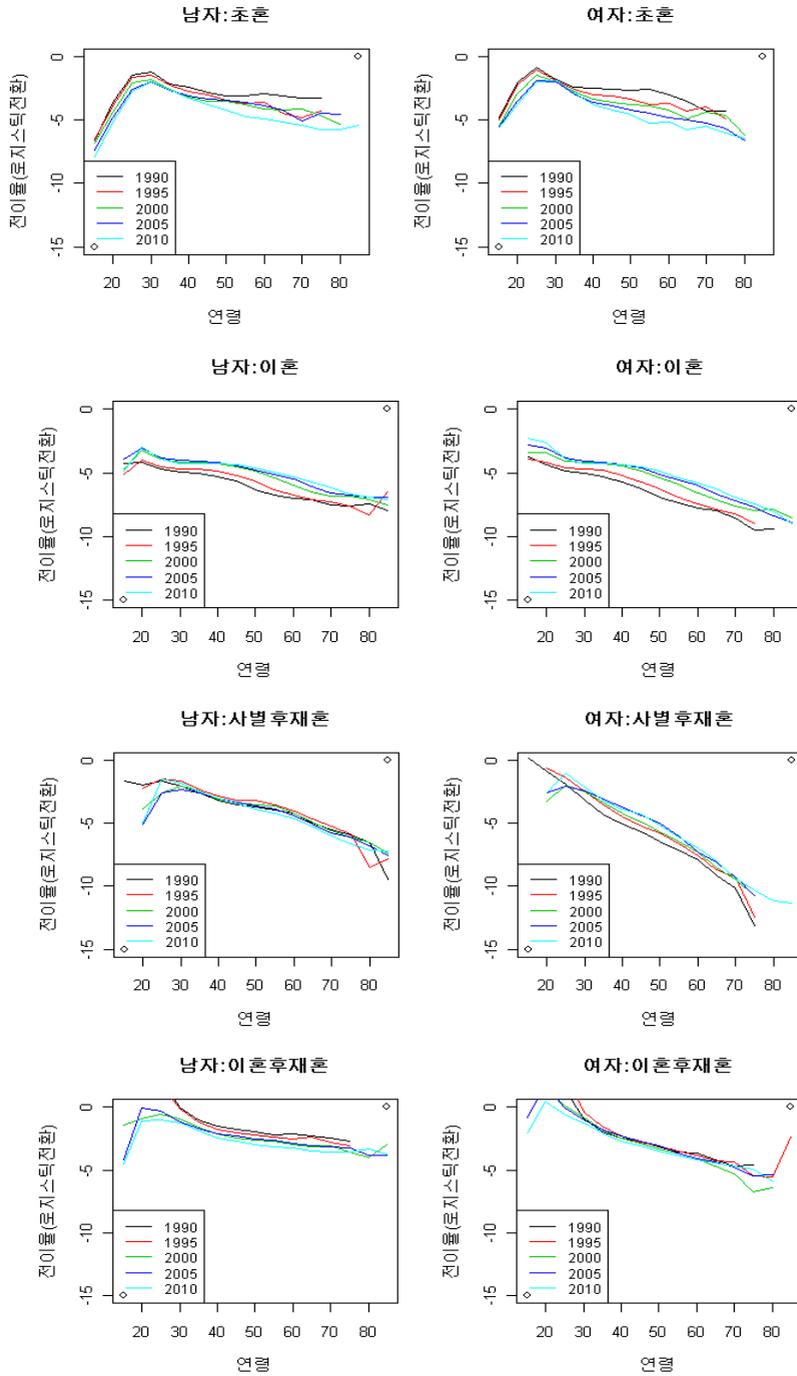
앞서 언급했듯이, 여기에서는 혼인상태별 인구추계를 위한 적합모형으로 판단된 전이확률 모형의 부분적 개선을 위해 각세자료의 불안정성 보정 방법, 고 연령 전이확률 작성방법 개선, 장래의 전이확률 추계의 모형 적용 등 개선 가능성에 대해서 살펴보고자 한다.

초혼, 이혼, 재혼(이혼 및 사별 후) 전이확률은 총조사의 혼인상태별 인구 및 인구동태자료의 연령별 혼인 및 이혼 건수를 기초로 산출한 혼인상태별 전이율로부터 전환한다. 각세별 전이율 산출시 혼인상태별 전이건수가 적어 전이확률이 매우 불안정한 형태를 보이는데, 이런 불안정을 피하기 위해 혼인상태별 전이율을 5세 간격으로 산출한 후 5세 간격 연령별 자료 또는 일부 불규칙한 자료를 단일연령으로 전환하는데 보편적으로 사용되는 spline 함수를 이용하는 방법을 제안한다(Smith 외 2004).

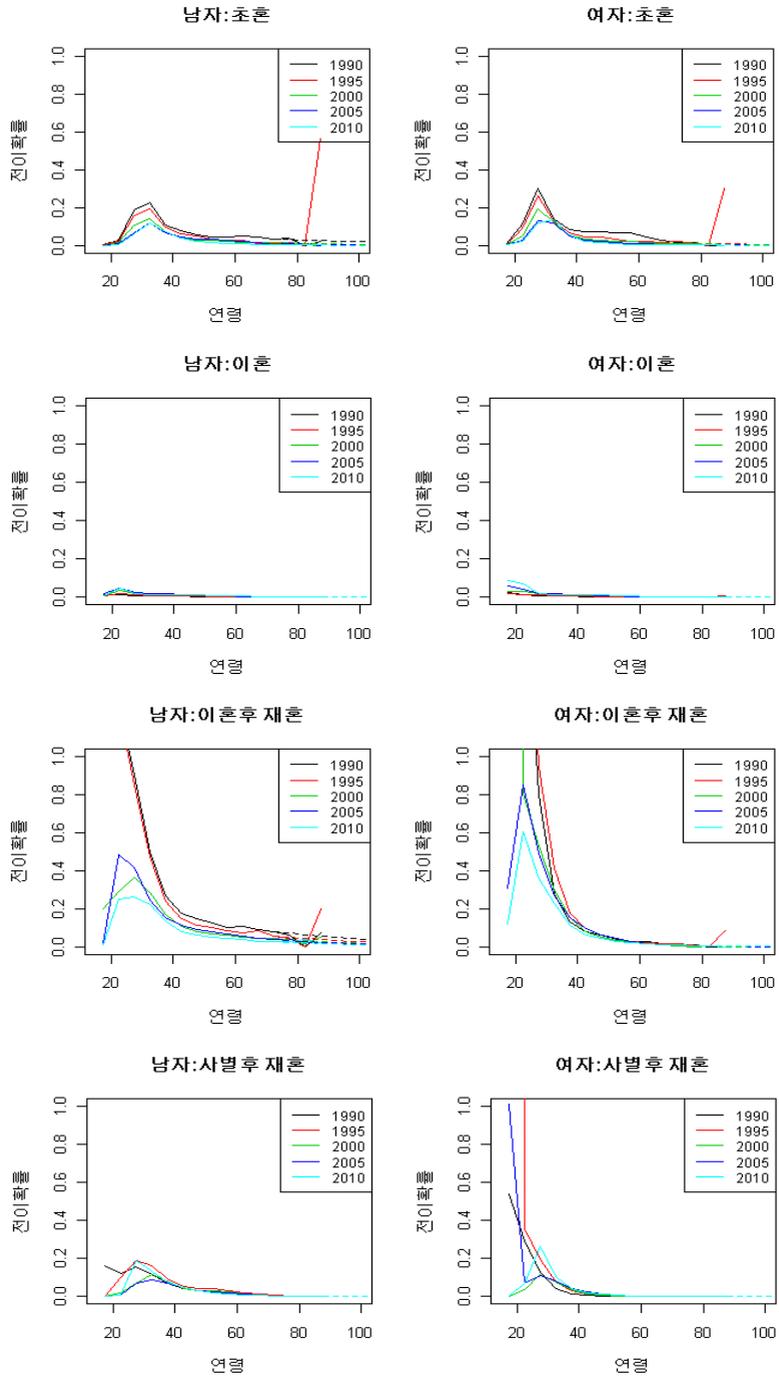
5세 간격 자료의 경우도 고 연령 및 저 연령에서는 전이건수가 적어 매우 불안정한데, 이의 보정을 위해 비교적 안정적인 연령(40-75세까지)의 전이율을 기초로 추정된 로지스틱 모형 등을 이용하여 이후 연령의 전이율을 보정하는 방법이 적합한 것으로 보인다. 2007년 혼인상태별 추계 시 중간 연령층에서는 적합도가 비교적 좋고, 고령층에서 오차가 많이 발생한 점을 감안할 때 100세 이상으로 연장하는 방법도 필요한 것으로 보이는데, 이를 위해서 앞에서 언급한 고령 전이확률 보정 모형을 이용할 수 있을 것이다.

[그림 3-7]은 40-75세의 전이율을 기초로 고연령 전이율을 보정하기 위해 로지스틱함수로 전환한 자료이며, [그림 3-8]은 원자료 및 추정된 모형으로 75-100세까지의 전이율을 보정한 결과이다. 이밖에, 전 연령의 혼인상태별 전이율을 살펴보면, 30세 미만 저 연령층에서는 혼인상태 전이건수가 적어 매우 불안정하게 나타나고 있어 추가 검토 사항으로 보인다.

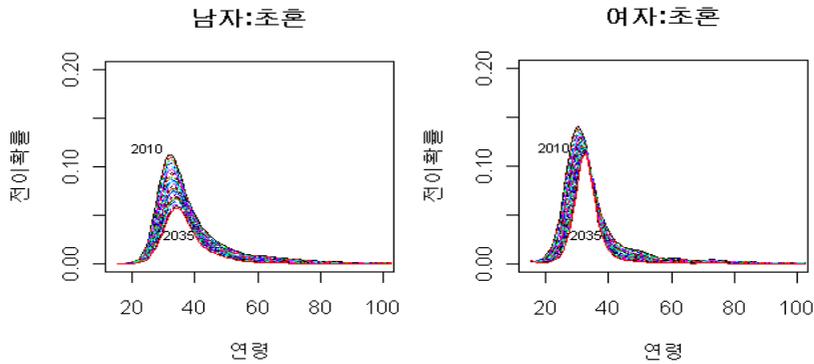
장래전이확률 추계를 위해서는 단일변수 시계열모형 등을 적용할 수 있을 것이다. 다만, 2007년 추계 시 대상기간동안의 변화율이 5년마다 1/2씩 감소한다는 가정 등을 적용한 추계도 비교적 적합도가 좋은 것으로 나타났으므로, 상호 비교 후 적용하는 것이 바람직한 것으로 보인다. [그림 3-9]는 초혼 전이확률을 단일변수 시계열모형에 의해 추계된 결과를 보여주고 있다.



[그림 3-7] 혼인상태별 전이율(1990-2010)



[그림 3-8] 고령 혼인상태별 전이율 보정(1995-2010)

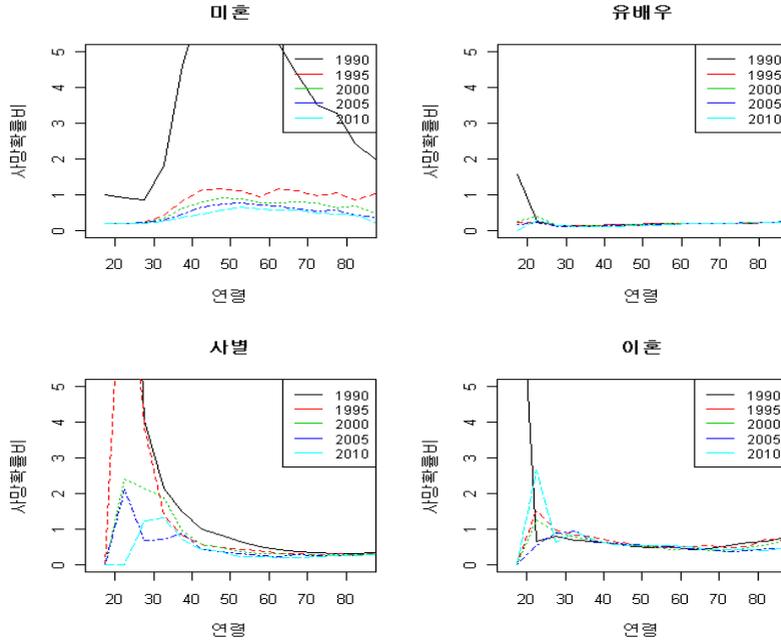


[그림 3-9] 초혼확률(2010-2035)

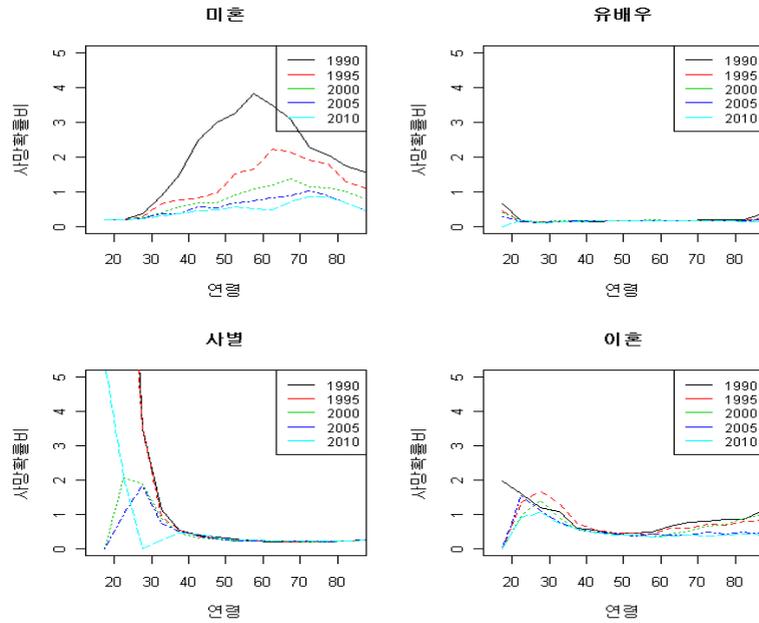
혼인상태별 전이확률 산출을 위해서는 혼인과 이혼의 전이확률 이외에, 혼인상태별 사망확률 및 유배우 사별확률 산출이 필요하다. 2007년 추계에서는 혼인상태별 사망확률의 상대비를 기초로 장래의 혼인상태별 사망확률비를 추계하였는데, [그림 3-10]에서 나타난 것과 같이 혼인상태별 사망확률비로 상대적 비율을 유지하고 있는 것으로 볼 때 합리적인 접근으로 보인다. 1990년 이후 유배우, 사별, 이혼의 사망확률비는 일정한 수준을 유지하고 있는데 반해, 미혼 사망확률비의 경우, 남자는 1990년 이후 급격하게 낮아진 이후 점차 감소하고 있고, 여자는 1990년 이후 지속적으로 낮아지고 있다. 이런 추세를 감안할 때 유배우, 사별, 이혼의 사망확률비는 2005년 및 2010년 사망확률비의 평균비를 이용하고, 미혼 사망확률비는 2005-2010년 감소율을 고려하는 추계 등이 적합해 보인다.

[그림 3-11]은 위에서 제시한 것과 같이 유배우, 사별, 이혼의 사망확률비는 2005년 및 2010년 사망확률비의 평균을 이용하고, 미혼사망확률비는 2005-2010년 감소율이 1/2로 감소한 이후 지속한다고 가정하여 추계한 2030년 혼인상태별 사망확률비를 2010년 혼인상태별 사망확률비와 비교하고 있다.

<남자>

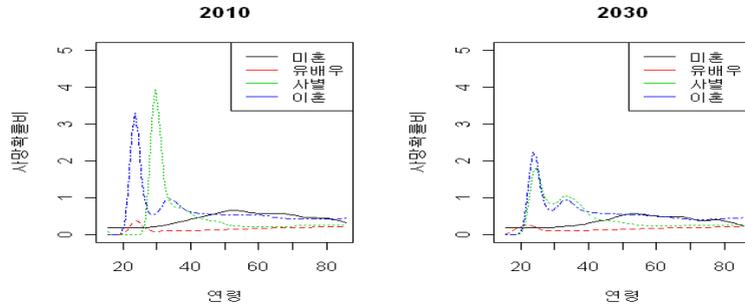


<여자>

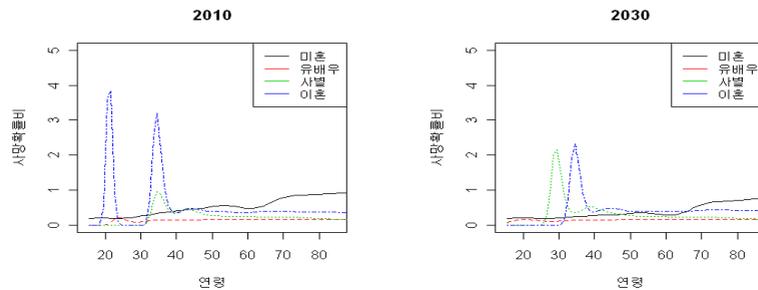


[그림 3-10] 혼인상태별 사망률비(1990-2010)

<남자>



<여자>

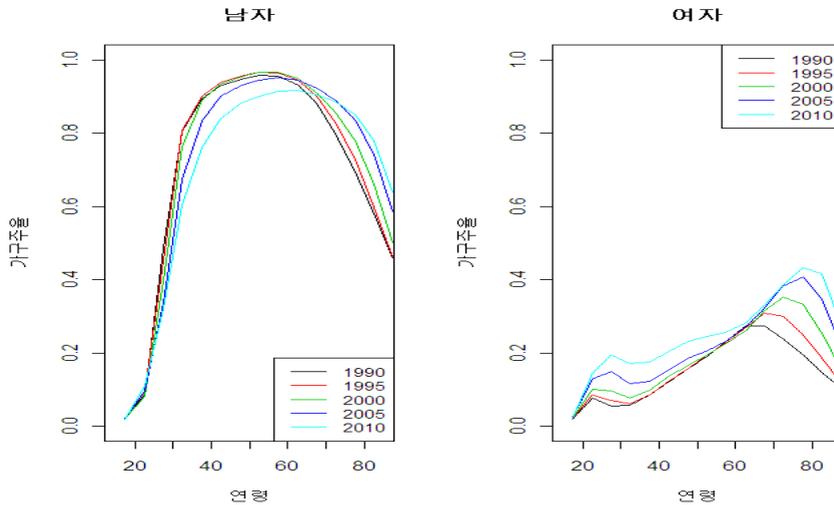


[그림 3-11] 혼인상태별 사망확률비(2010, 2030)

3. 혼인상태별 가구주율 추계방법 검토 및 개선안

가. 혼인상태별 가구주율 추이(1990-2010)

[그림 3-12]에서는 1990-2010년의 성별에 따른 가구주율 추이를 나타내고 있는데 그 패턴이 뚜렷한 차이를 보인다. 남자 가구주율은 65세를 기점으로 그 미만에서는 감소하고, 이상에서는 증가하고 있다. 이는 각각 다음과 같은 인구사회학적 특성 변화와 관계되는 것으로 보인다. 65세 미만에서의 가구주율 감소는 전반적으로 혼인율이 감소한 것과 함께 가구 내 여성의 지위, 즉 여자 가구주의 증가와 관련되는 한편, 65세 이상에서의 가구주율 증가는 기대수명의 증가 및 자녀가구와 독립적으로 거주하는 노인가구의 증가의 영향을 받은 것이다. 여자 가구주율 또한 65세를 기점으로 변화가 달라지는데, 그 변화폭이 시계열별로 차이를 보이고 있다. 65세 미만에서는 2000년까지는 증가폭이 완만하였으나, 그 이후에는 큰 폭으로 증가하고 있는 한편, 65세 이상에서는 지속적으로 큰 폭으로 증가하고 있다.



[그림 3-12] 가구주율 추이(1990-2010)

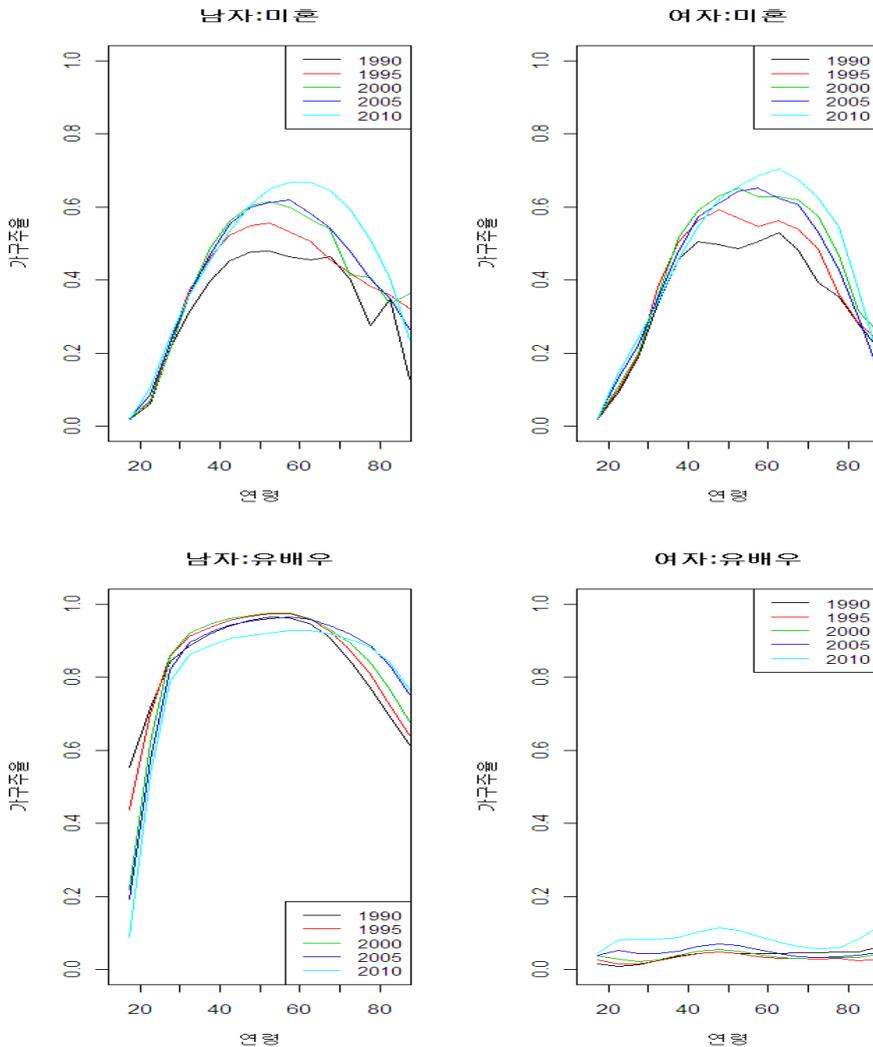
다음은 연령별 가구주율 추이를 혼인상태별로 살펴본 것이다 [그림 3-13]. 먼저, 미혼 가구주율 추이 변화를 보면, 남자와 여자 모두, 청년층에서는 연도별 차이가 거의 없으나, 중장년층으로 갈수록 그 비율이 급격하게 증가하고 있음을 알 수 있었다. 혼인의 지연 및 결혼포기의 증가 등에 따라서 생성되는 미혼인구가 독립가구를 형성하는 비율이 증가하고 있는 것으로 예상해 볼 수 있다.

유배우 가구주율은 남자와 여자에서 차이를 보였다. 남자의 경우 중년층에서는 큰 변화가 없는 것과 달리, 청년층에서는 감소, 노년층에서는 증가하는 패턴을 보이고 있다. 이러한 경향은 시점에 따라서 다른데, 특히 2005년과 2010년 사이에 변화가 있었다. 즉 20-65세의 유배우 가구주율은 과거의 증가추이와는 달리 2005년에 비해 크게 감소하였고, 65세 이상의 유배우 가구주율은 2005년까지 큰 폭으로 증가하다가 2010년은 2005년과 유사한 수준을 보인다. 이는 기대수명 증가 및 노인부부가구의 증가와 관련이 있음을 예상해 볼 수 있다. 한편, 여자 유배우 가구주율은 연령과 무관하게 거의 일관적인 것으로 나타났으며, 그 추세 또한 2005년까지 큰 변동이 없었다. 그러나, 2010년에 그 비율이 거의 모든 연령대에서 크게 증가하였음을 주목해 볼 만하다.

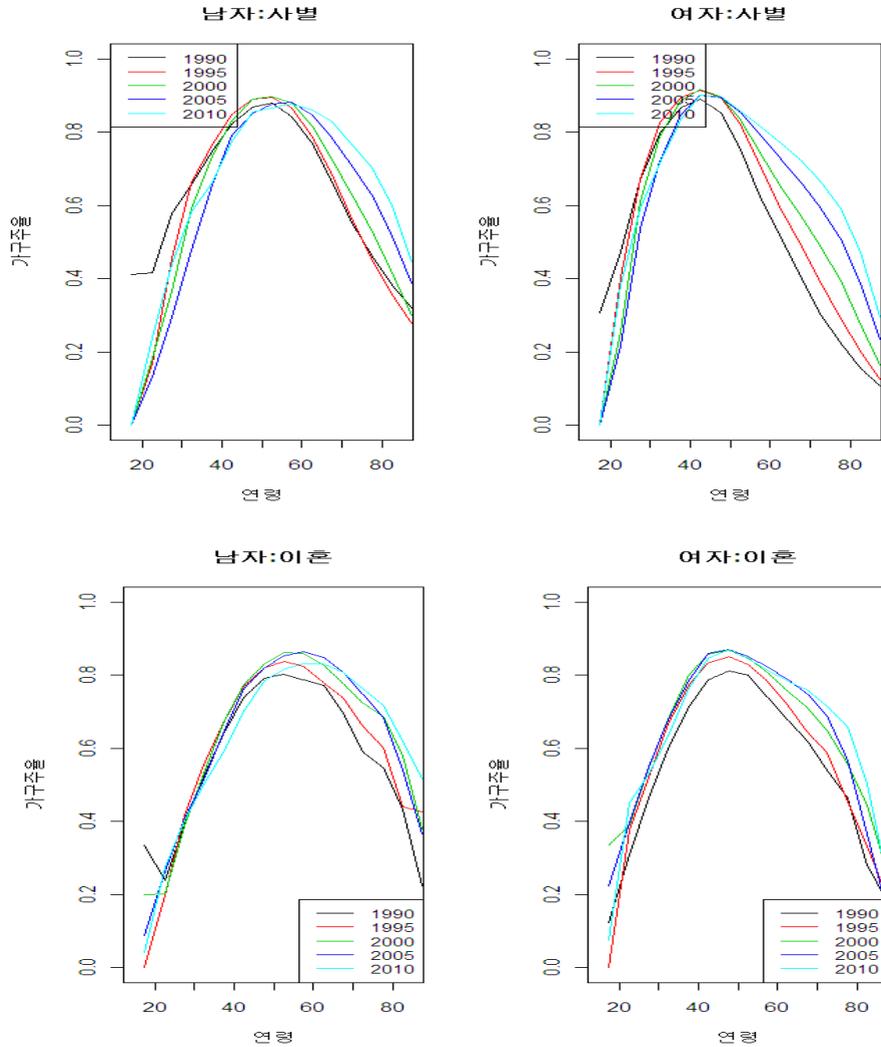
사별 및 이혼 가구주율의 비율은 남녀 모두 중장년층이상에서 증가하고 있었다. 그런데 사별 가구주율이 최근 시점으로 올수록 증가하는 패턴은 꽤 명확한데 반해, 이혼 가구주율의 경우는 시점 및 연령에 따라서 차이를 보여주었다. 2000년까지는 50세까지

의 이혼가구주율이 증가하였으나, 2000년 이후에는 남자의 경우는 감소, 여자는 정체를 보이고 있고, 50세 이상에서는 약간 증가를 보이고 있다. 이혼 및 사별 등으로 노년에 단독가구로 형성되는 비율이 높아지고 있음을 알 수 있다.

혼인상태에 따른 가구주율의 변화를 살펴본 결과, 노년층의 변화가 매우 두드러진 것으로 나타났다. 기대수명의 증가에 따라서 이 집단의 유배우가구주율이 증가한 동시에 미혼 및 이혼, 사별가구주율의 증가도 함께 보이는 등의 복잡한 양상을 띠고 있다. 향후 가구 추계 시 이 집단에 대한 심도 있는 분석이 요구됨을 보여주는 대목이다.



[그림 3-13] 혼인상태별 가구주율 추이(1990-2010)



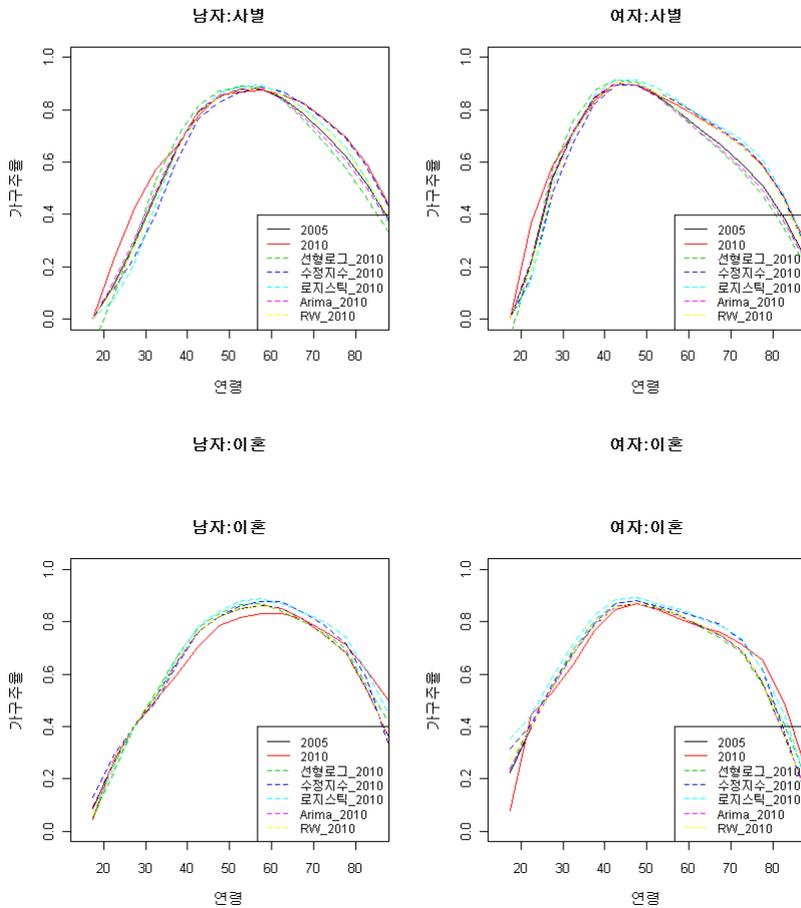
[그림 3-13] 혼인상태별 가구주율 추이(1990-2010)(계속)

나. 혼인상태별 가구주율 추계방법별 적합도 비교

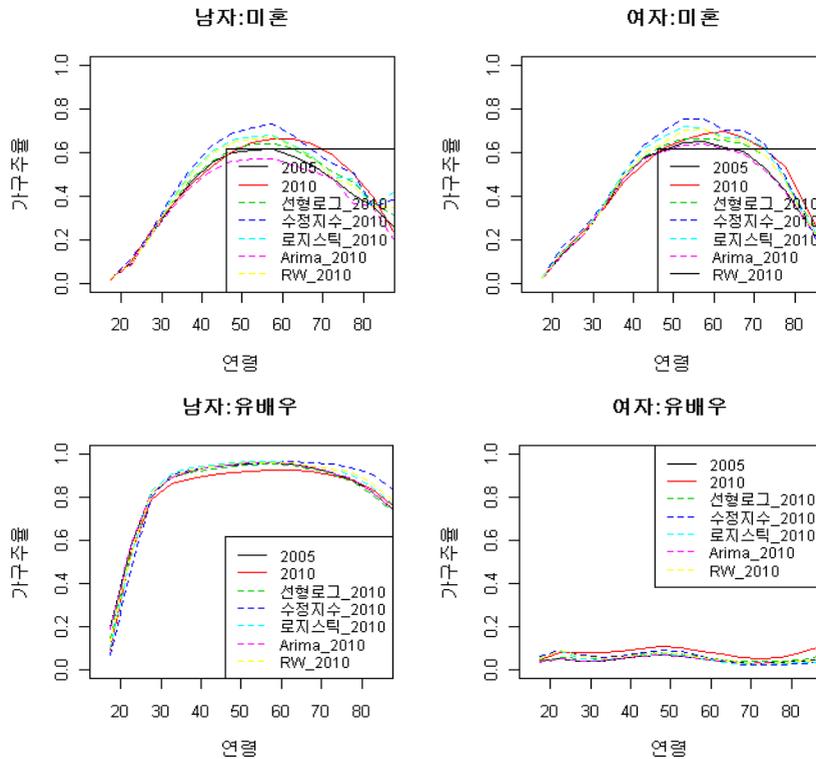
이 절에서는 가구주율 추계에 주로 이용되고 있는 수학적 모형으로 2007년 장래가구 추계에서 적용한 선형로그식과 함께, 수정지수식, 로지스틱모형, 시계열모형 등 다양한 방법으로 가구주율을 추계하여, 2010년 총조사의 혼인상태별 가구주율과의 적합도를 비교하고자 한다. 1990-2005년 혼인상태별 가구주율 자료를 투입하여 적합도가 좋은 모형

(1-2순위)을 선택한 다음, 투입기간을 다음과 같이 변경해가면서 1990-2005(15년 자료), 1995-2005(10년 자료), 2000-2005(5년 자료) 최종 적합도를 비교하고자 한다.

[그림 3-14]는 1990-2005년(15년)의 자료를 이용한 혼인상태별 가구주율 추계결과를 작성방법별로 나타낸 것인데, 혼인상태별 모형별 적합도가 다르며, 모든 혼인상태에서 일률적으로 잘 적합되고 있는 모형은 없는 것으로 보인다. 특히 2010년의 남자 유배우율의 감소 및 여자 유배우율의 증가는 모든 모형에서 적합도가 떨어졌는데, 이는 과거 추이와 달리 2010년 급격한 변화를 보인 구성비의 변화가 모형에 반영되기 어려웠음을 반영하는 것이다.



[그림 3-14] 모형별 혼인상태별 가구주율(2005, 2010)



[그림 3-14] 모형별 혼인상태별 가구주율(2005, 2010)(계속)

<표 3-6>과 <표 3-7>은 1990-2005년 자료를 이용한 혼인상태별 추계방법별 가구수 및 추계방법별 가구수 차이와 오차율을 나타내고 있다. 가구주 성별에 따른 오차율을 보면, 남자 가구주 가구는 1~3%, 여자 가구주 가구는 -6~-18%를 보이긴 하나, 남녀전체는 2.5% 미만 이었다³⁾. 구체적으로 오차율을 모형별로 비교해 보면, 2007년 추계 시 적용한 선형로그모형은 Arima 모형 다음으로 과소추계오차가 크고, 수정지수모형 및 Random Walk 시계열모형은 실제 가구수보다 크게 추계하고 있지만, 절대오차의 크기는 Random Walk 시계열 모형, 수정지수모형순으로 적다.

다음으로, 투입자료기간에 따른 차이를 비교하기 위하여 적합도가 높은 순인 RW시계열모형 및 수정지수모형을 1990-2005년(5년), 1995-2005년(10년), 1990년-2005년(15년) 자료를 이용하여 추가적으로 모형 및 투입기간별 가구수 차이 및 오차를 비교하였다

3) 현 자료는 총조사 자료에 사후조사 누락률 및 기준일 변경(총조사 기준 11. 1. -> 추계기준 7. 1.) 을 실시하지 않은 자료이므로 추계가구와의 정확한 비교는 어렵다

<표 3-8>. 투입기간별 6개의 비교자료 중 최근 5년 자료를 이용한 수정지수모형이 오차가 가장 적은 것으로 나타났다. 2007년 추계 시 적용한 최근 15년 자료를 이용한 선형로그식에 비해 절대가구수 오차가 50% 수준으로 감소하였다.

결론적으로 말하면, 2007년 혼인상태별 가구주율 추계(선형로그모형)는 여자 가구주율의 과소추계 및 남자 가구주율의 과대 추계로 남녀 가구주별 가구수 차이는 비교적 큰 편이나, 남녀전체 가구주율 추계단계의 오차는 2.3% 수준이다. 다만, 2007년에 적용한 이 모형은 15년 자료를 이용한 적용방법 중 과소오차가 가장 크므로 모형을 변경하여 적용할 필요가 있는 것으로 판단된다. 이에 투입기간을 변경하여 모형 적합도를 비교해 본 결과, 최근 5년 자료를 이용한 수정지수모형이 적합도가 가장 높은 것으로 나타나고 있다.

<표 3-6> 혼인상태별 가구주율 추계방법별 가구수 차이 (2010년)¹⁾

(단위 : 천 가구)

구분		선형로그	수정지수	로지스틱	Arima	RW
전 체	미 혼	-90	18	28	-159	-22
	유배우	-80	-8	-75	-55	33
	사 별	-242	-26	16	-214	-37
	이 혼	11	24	36	-2	3
	차 이	-401	8	6	-430	-28
	절대오차합	1332	848	888	1252	847
남 자	미 혼	-45	13	19	-109	-13
	유배우	448	386	354	400	400
	사 별	-29	-9	-12	-24	-14
	이 혼	18	20	26	11	10
	차 이	391	409	387	278	383
	절대오차합	540	428	411	544	437
여 자	미 혼	-45	5	9	-49	-8
	유배우	-528	-394	-429	-455	-367
	사 별	-213	-17	28	-191	-22
	이 혼	-6	4	11	-13	-13
	차 이	-792	-402	-381	-708	-411
	절대오차합	792	420	477	708	410

1) $\sum_{x,m,i} (P_{xm} * h_{xm}^i - H_{xm})$, 여기서 P_{xm} : 총조사인구 h_{xm}^i : 가구주율 H_{xm} : 총조사 가구, x 는 연령, m 은 혼인상태, i 는 추계방법

2) 남자 및 여자의 절대오차합의 합

3) 1990-2005년 자료 이용

〈표 3-7〉 혼인상태별 가구주율 추계방법별 가구수 오차율¹⁾

(단위 : %)

구분		선형로그	수정지수	로지스틱	Arima ²⁾	RW
전체	미 혼	-3.6	-0.5	1.1	-6.3	-0.9
	유배우	-0.7	-0.6	-0.7	-0.5	0.3
	사 별	-12.0	1.0	0.8	-10.6	-1.8
	이 혼	0.9	-0.2	2.9	-0.2	-0.2
	전 체	-2.3	-0.3	0.0	-2.5	-0.2
남자	미 혼	-3.0	-1.0	1.3	-7.3	-0.9
	유배우	4.3	2.2	3.4	3.8	3.8
	사 별	-9.0	-0.7	-3.6	-7.2	-4.4
	이 혼	3.2	1.3	4.7	1.9	1.9
	전 체	3.0	1.7	3.0	2.2	3.0
여자	미 혼	-4.5	0.3	0.9	-4.9	-0.8
	유배우	-48.9	-27.7	-39.7	-42.1	-34.1
	사 별	-12.5	1.3	1.6	-11.2	-1.3
	이 혼	-0.9	-1.3	1.5	-1.8	-1.8
	전 체	-17.6	-6.3	-8.5	-15.7	-9.1

1) $\sum_{x,m,i} (P_{xm} * h_{xm}^i - H_{xm}) / H_{xm}$, 여기서 P_{xm} : 총조사인구 h_{xm}^i : 추계 가구주율 H_{xm} : 총조사 가구, x 는 연령, m 은 혼인상태, i 는 추계방법

2) Arima(1,1,0)모형

3) 1990-2005년 자료 이용

<표 3-8> 자료투입기간별 혼인상태별 가구주율 추계방법별 가구수 차이¹⁾

(단위 : 천 가구)

구분		수정지수_5년	수정지수_10년	수정지수_15년	RW_5년	RW_10년	RW_15년
전 체	미혼	-12	-49	18	14	-37	-22
	유배우	-67	-23	-8	18	55	33
	사별	21	-7	-26	-175	119	-37
	이혼	-3	13	24	-3	9	-3
	전체	-61	-67	8	-146	146	-28
	절대오차합 ²⁾	590	739	848	703	805	847
남 자	미혼	-14	-40	13	1	-34	-13
	유배우	232	322	386	257	342	400
	사별	-2	-5	-9	-20	4	-14
	이혼	7	14	20	7	14	10
	전체	223	291	409	245	327	383
	절대오차합	255	381	428	285	394	437
여 자	미혼	3	-9	5	14	-3	-8
	유배우	-299	-345	-394	-239	-288	-367
	사별	23	-2	-17	-155	115	-22
	이혼	-10	-2	4	-10	-5	-13
	전체	-283	-358	-402	-391	-180	-411
	절대오차합	335	358	420	418	411	410

1) $\sum_{x,m,i} (P_{xm} * h_{xm}^i - H_{xm})$, 여기서 P_{xm} : 총조사인구 h_{xm}^i : 가구주율 H_{xm} : 총조사 가구, x 는 연령, m 은 혼인상태, i 는 추계방법

2) 남자 및 여자의 절대오차합의 합

3) 1990-2005년 자료 이용



4. 가구원수별 가구구성비 추계방법 검토 및 개선안

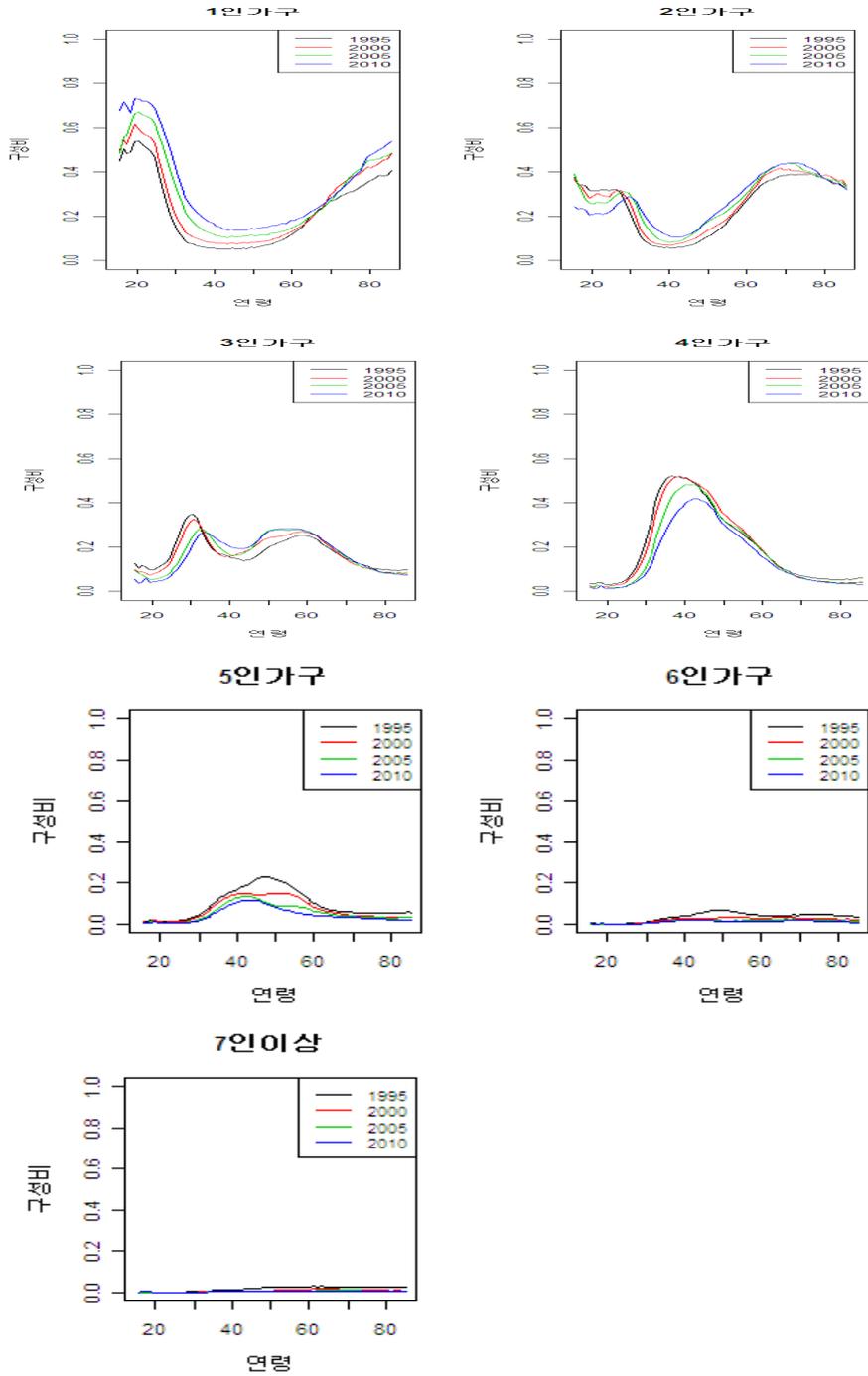
가. 가구원수별 가구구성비 추이

가구구성과 관련하여, 2010 총조사에서 나타난 가장 뚜렷한 변화는 1-2인가구의 급격한 증가와 4인가구의 감소로 요약될 수 있다(통계청, 2011). 1인가구는 1995년 이후 65세 전후를 제외하고는 모든 연령에서 큰 폭으로 증가하고 있다. 마찬가지로 2인 가구 또한 30세 미만을 제외하고는 전 연령층에서 증가하고 있었는데, 이와 같이 1-2인 가구가 증가하는 것은 혼인지연 등에 따른 가구형성의 감소와 이들의 독립가구 형성, 그리고 혼인 이후의 저출산에 따른 부부가구의 증가, 기대수명의 증가에 따른 노인가구의 증가 등에서 그 원인을 찾을 수 있겠다. 이러한 경향은 동시에 3인 및 4인가구의 구성비 감소로 이어지고 있다.

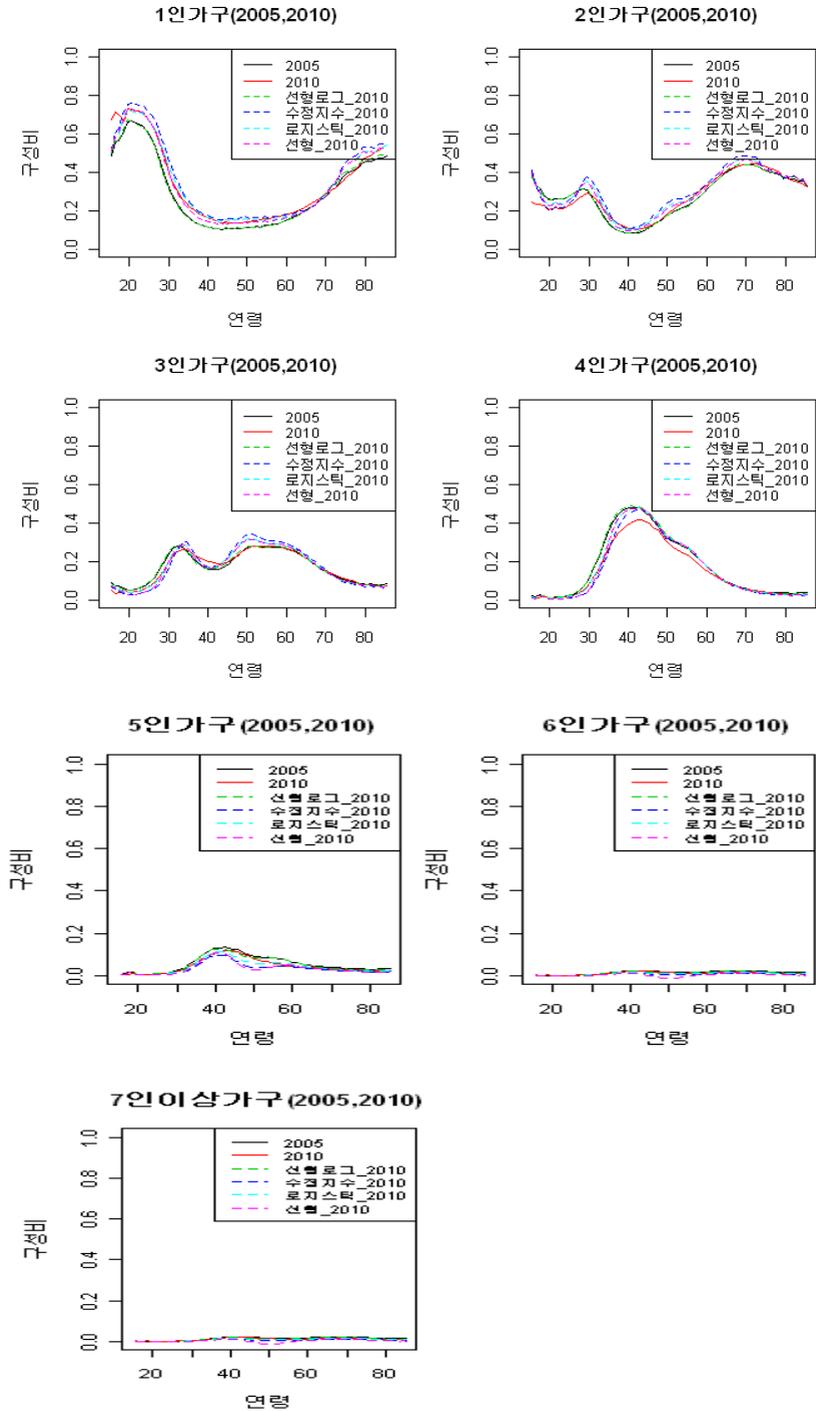
나. 가구원수별 가구구성비추계의 적합도 검토

[그림 3-15]는 추계방법별 가구구성비 적합도를 나타내고 있는데, 1인가구 구성비는 선형로그식을 제외하고는 모든 모형에서 잘 적합하고 있는 것으로 보이고, 4인가구 구성비는 모든 모형에서 적합도가 낮은 것으로 나타난다. 2인 및 3인가구는 변화의 폭이 적어 모형 간 적합도 구분이 명확하게 드러나지 않지만 큰 오차는 없는 것으로 보인다. 4인 가구 구성비의 적합도가 낮은 이유는 4인가구 구성비 감소가 2005-2010년에 급격하게 나타난 변화이므로 과거추이를 이용하는 모형에서 반영이 잘 되고 있지 않은 것이다.

다음으로 가구원수별 구성비추계의 적합도 검증을 실시하였다. 먼저, 2010년 총조사의 가구원수별 구성비 및 추계방법별로 산출한 가구원수별 구성비를 비교하였다 [그림 3-16]. 다음으로, 2010년 총조사의 연령별 가구수에 추계방법별 연령별 가구원수별 구성비를 곱하여 추계방법별 가구원수 구성비를 산출하고, 총조사와 추계 결과의 차이를 가구원수별 가구 및 추계방법에 따른 오차를 산출하였다 <표 3-9>. 각 방법별 구성비 차이의 절대오차합을 비교해 보면, 로지스틱 모형의 적합도가 가장 높은 것으로 보이고, 수정지수식에 의한 추계가 두 번째로 적합도가 높은 것으로 나타났다.



[그림 3-15] 가구원수별 가구구성비 추이(1990-2010)



[그림 3-16] 추계방법별 가구구성비

〈표 3-9〉 추계방법별 가구원수 구성비 및 오차(2010)

구분	총조사(A)	선형로그(B)	수정지수(C)	로지스틱(D)	선형회귀(E)
1인가구	0.2388	0.2017	0.2258	0.2344	0.2326
2인가구	0.2424	0.2304	0.2446	0.2443	0.2502
3인가구	0.2131	0.2115	0.2137	0.2117	0.2184
4인가구	0.2247	0.2665	0.2451	0.2420	0.2509
5인가구	0.0622	0.0719	0.0567	0.0542	0.0455
6인가구	0.0140	0.0138	0.0107	0.0102	0.0024
7인이상가구	0.0048	0.0041	0.0033	0.0033	0.0000
		B-A	C-A	D-A	E-A
1인가구		-0.0371	-0.0130	-0.0044	-0.0062
2인가구		-0.0120	0.0022	0.0018	0.0078
3인가구		-0.0015	0.0006	-0.0014	0.0054
4인가구		0.0418	0.0204	0.0173	0.0262
5인가구		0.0097	-0.0055	-0.0079	-0.0167
6인가구		-0.0002	-0.0033	-0.0038	-0.0117
7인이상가구		-0.0007	-0.0015	-0.0015	-0.0048
절대오차합 ¹⁾		0.1031	0.0465	0.0383	0.0787

1) $\sum_{n,i} |P_{ni} - p_{ni}|$, 여기서 P_{ni} : 총조사가가구구성비, p_{ni} : 추계가가구구성비, n 은 가구원수, i 는 추계방법

*1995-2005년 자료이용

분석 시 투입된 자료의 기간이 모형 적합성에 영향을 줄 수 있으므로, 비교모형 중 적합도 1,2순위인 수정지수식 및 로지스틱 모형에 대해 자료 투입기간을 변경시켜 보았다 <표 3-10>. 투입기간별 적합도를 살펴보면, 수정지수식 및 로지스틱 모형 모두 최근 10년(1995-2005년) 자료를 이용한 모형보다는 최근 5년(2000-2005년) 자료를 이용한 모형이 잘 적합하였다. 모형으로는 2000-2005년 자료를 이용한 로지스틱모형의 절대오차합이 가장 적어 적합도가 가장 좋은 것으로 나타났으며, 특히 2010년 총조사의 가구원구성비 규모순인 2인, 1인, 4인, 3인과 정확히 일치하는 구성비를 추계하고 있다. 이러한 분석결과는 우리나라의 급격한 가구원구성비 변화를 반영하기 위해서는 단기변화를 적용하여 추계할 필요성을 시사하고 있다.

부부가구, 한부모가구, 자녀가구 등 가구유형별 구성비추계는 가구원수별 구성비 추계와 유사한 방법으로 작성되고, 2010년 총조사 자료부터 세대유형 구분이 2005년 기준

총조사의 19가지 유형에서 일부 통합하여 집계될 예정이므로 본 연구에서는 검토를 생략하고자 한다.

〈표 3-10〉 추계방법별 가구원수별 구성비 추계(2010)

	총조사(A)	수정5년(B)	수정10년(C)	로지스틱5년(D)	로지스틱10년(E)
1인가구	0.2388	0.2311	0.2258	0.2396	0.2344
2인가구	0.2424	0.2484	0.2446	0.2494	0.2443
3인가구	0.2131	0.2126	0.2137	0.2099	0.2117
4인가구	0.2247	0.2350	0.2451	0.2292	0.2420
5인가구	0.0622	0.0578	0.0567	0.0563	0.0542
6인가구	0.0140	0.0115	0.0107	0.0113	0.0102
7인이상	0.0048	0.0036	0.0033	0.0043	0.0033
		B-A	C-A	D-A	E-A
1인가구	-	-0.0077	-0.0130	0.0008	-0.0044
2인가구	-	0.0060	0.0022	0.0070	0.0019
3인가구	-	-0.0005	0.0006	-0.0032	-0.0014
4인가구	-	0.0103	0.0204	0.0045	0.0173
5인가구	-	-0.0044	-0.0055	-0.0059	-0.0080
6인가구	-	-0.0025	-0.0033	-0.0027	-0.0038
7인이상	-	-0.0012	-0.0015	-0.0005	-0.0015
절대오차합	-	0.0327	0.0465	0.0245	0.0383

* 5년(2000-2005), 10년(1995-2000)

5. 소결

이 절에서는 혼인상태별 인구추계 및 가구주율법에 의한 가구추계의 각 단계별 적합도를 검토해 보았다. 1990-2005년 자료를 이용한 혼인상태별 인구추계 적합도는 전이확률법에 의한 추계에서 오차가 가장 적었다. 이 방법은 고연령층의 일부를 제외하고는 전 연령에서 비교적 잘 적합하였다. 그러나 전이확률 모형 적용 시에도 여러 개선이 필요한 것으로 나타났다. 저 연령 및 고 연령 전이확률 보완 및 장래 전이확률 추계를 위한 합리적 모형 적용 등의 개선이 요구되는데, 로지스틱 함수 보정 및 5세 간격 전이율 작성 후 스플라인 함수 보정 등의 방법이 검토될 필요가 있다. 또한, 단일변수 시계열모형에



의한 장래전이확률 추계방법 등도 고려될 수 있다.

혼인상태별 가구주율 및 가구원수별 구성비추계에서는 2007년 추계 시 적용한 선형 로그모형은 그 특성상 최근의 급격한 가구변화에 대한 적합도가 가장 낮은 것으로 나타났다. 가구주율 추계 적합도를 보면 1990-2005년 자료를 이용한 모형에서는 RW 시계열 모형, 수정지수모형 순으로 오차가 적은 것으로 나타났으나, 적합도가 높은 1,2 순위 모형에 대해 자료투입기간을 변경시켜 비교해 본 결과, 최근 5년 자료를 이용한 수정지수식이 최근 15년 자료를 이용한 선형로그모형에 비해 오차가 50% 미만 수준으로 감소하면서 적합도가 가장 좋은 모형으로 검증되었다.

가구원수별 구성비 추계를 검증한 적합도 검증에서는 1995-2005년 자료를 이용한 모형에서는 로지스틱모형 및 수정지수 모형의 오차가 적은 것으로 나타났고, 적합도가 높은 1,2 순위 모형에 대해 자료투입기간을 변경시켜 살펴본 결과, 최근 5년(2000-2005년) 자료를 이용한 로지스틱모형의 절대오차합이 가장 적어 적합도가 가장 좋은 것으로 나타났다. 특히, 이 모형은 2010년 총조사의 가구원구성비 규모순인 2인, 1인, 4인, 3인과 정확히 일치하는 구성비를 추계하고 있다.

제4절 결론 및 제언

본 연구는 장래가구추계의 예측성 제고를 목적으로 실시되었으며, 다음과 같은 내용을 포함하고 있다. 첫째, 가구추계방법의 이론적 배경 및 유형, 해외통계작성기관의 가구추계사례를 고찰하였다. 둘째, 2007년 가구추계의 작성단계별 방법 및 오차를 검토하였고, 셋째 각 단계별로 다양한 모형을 적용, 2010년 총조사 자료와의 적합도 검증을 실시하였다. 넷째, 이 결과를 기초로 2012년 장래가구추계의 개선방안을 제시하였다.

가구추계방법은 작성기법 및 이용 자료의 종류에 따라서 정태모형과 동태모형, 거시모형과 미시모형으로 구분할 수 있는데, 본 연구에서는 거시-정태 모형에 속하는 가구주율법을 기본으로 그 개선방안을 제시하였다. 가구주율법은 계산과정이나 자료요구가 다른 방법에 비해 단순하고, 가구구성비 변화 등에 대한 다양한 결과 제공이 가능하여 국가차원의 가구추계에서 가장 많이 활용되고 있는 방법이다. 미시-동태 모형은 가구나 가족의 동적변동을 관찰하는데 적합하다는 장점이 있으나, 랜덤수를 이용함에 따른 표본오차 증가 등의 장애요인으로 인해 공식추계에 사용되는 경우는 드물다. 또한 거시동태모형은 개인의 가구변화 전이확률 생성을 위해 특별한 표본조사를 요구한다는 점, 특별한 표본조사를 이용하지 않는 것으로 제안된 다차원 모델의 경우에도, 개인의 특성별 전이확률 생성을 위해 거대한 전이확률행렬을 필요로 한다는 점에서 제약이 있는 것으로 평가되고



있다. 이외에도 다차원 모델의 경우 기존 장래인구추계와의 일관성 확보가 어려운 점, 장래전이확률이 일정하게 유지된다는 가정을 적용하는 점 등의 단점이 제기된다.

2007년 장래가구추계의 단계별 타당성 검토 결과, 혼인상태별 인구추계, 혼인상태별 가구주를 추계 단계에서는 오차가 비교적 적게, 가구원수별 구성비 추계 단계에서 오차가 크게 발생한 것으로 검증되었다. 혼인상태별 인구추계 오차율은 이혼율의 과소추계 및 유배우율의 과대추계 등의 오차가 있었고, 평균 절대오차율로 측정했을 때 남자 4.7%, 여자는 3.4%로 남녀 모두 5%미만으로 나타났다. 2007년 적용하였던 전이확률법이 가장 오차가 적은 것으로 나타났으므로, 2012년 추계에서도 동일방법을 적용하되, 오차가 비교적 크게 발생하였던 고연령층 부분의 전이확률 오차를 줄이기 위해서 전이확률 작성 최종연령을 기존 85세 이상에서 100세 이상으로 연장하는 방법이 바람직한 것으로 판단된다. 이와 함께 고령 전이확률 불안정성 보완, 각세별 전이확률 산출의 불안정성 보완, 장래전이확률 추계를 위해 각각 로지스틱 함수 보정, 5세 간격 전이율 작성 후 spline 함수 등으로 보정하는 방법 및 단일변수 시계열모형에 의한 장래전이확률 추계방법 등을 제안한다.

2007년 혼인상태별 가구주를 추계에서는 남녀별 가구주를 추계 시 일부 오차가 발생하였으나, 전체 가구 수 추정에서는 약 2%내외의 과소추정이 있었던 것으로 검증되었다. 2012년 혼인상태별 가구주를 추계에는 2010년 총조사 자료와의 적합도 검증에서 가장 적은 오차를 갖는 것으로 검증된 최근 5년 자료를 이용한 수정지수모형 이용을 제안하였다.

2007년 가구원수별 가구구성비 추계는 2010년 총조사의 가구구성비와 큰 차이가 있는데, 특히 1인 가구구성비오차가 15%이상 크게 발생하였다. 여러 가지 모형으로 2010년 총조사 자료와 적합도를 비교한 결과, 최근 5년 자료를 이용한 로지스틱 모형의 오차가 가장 작았고, 2010년 총조사에서 나타난 2인→1인→3인→4인의 가구구성비순과 정확히 일치하는 추계결과를 제공하고 있어 최근 5년 자료 로지스틱 모형이용을 제안하였다.

마지막으로 향후 과제와 관련하여 몇 가지 사항을 언급하고 마무리하고자 한다. 첫째, 혼인상태에 의한 가구변화외의 가구의 형성 및 해체에 대한 연구가 필요한 것으로 보인다. 이는 결혼연령이 늦어지면서 자녀가 부모로부터 독립하여 가구를 형성하는 등, 혼인상태의 변화 없이도 가구가 형성되는 경우의 비율이 증가할 것이라는 예측에 따른 것이다. 둘째, 외국인 가구추계에 대한 수요가 증가하고 있는 현실을 감안할 때 외국인가구추계방법에 대한 검토가 요청된다. 이를 위해서는 타당성 높은 외국인 장래인구추계자료의 구축이 선행되어야 하며, 외국인 가구형성비율을 비롯한 심층적 연구가 필요하다. 셋째, 2015년 이후의 등록센서스 자료를 이용한 장래가구추계방법에 대한 대비가 있어야 한다. 등록센서스 기준가구의 정확한 설정이 선결요건인데, 이를 위해서는 총조사의 가구 및 주민등록자료의 세대간의 차이에 대한 검토를 비롯하여, 주민등록자료 및 가족관계등록부 자료 등 다양한 행정자료를 이용한 가구추계방법에 대한 연구가 필요하다.

참고문헌

- 김형석. 2002, "가구주율법에 의한 장래가구추계", 한국조사연구학회
- 최은영, 강창덕, 권순필, 김경환, 김동희, 김용창, 김준형, 민인식, 이상일, 이희연, 조만. 2011, 《주
택수요변화에 대한 연구》, pp. 283-307, 국토해양부
- 통계청, 2007, 《2005-2030 장래가구추계》
- 통계청, 2011, 《2010 인구주택총조사 가구주택부문》
- Bell, M. Cooper, J., and Les, M. 1995. *Household and Family Forecasting Models: A Review*.
Commonwealth of Australia: Canberra.
- Van Imhoff, E., Kuijsten, A., Hooimeijer, P., Van Wissen, L., 1991. "National household forecasts
for the Netherlands", *Household Demography and Household Modeling*, Newyork,
Plenum Press, pp. 251-272
- Fukawa, T., 1995, "Future trends of Japanese households through micro-simulation model- An
Application of INAHSIM", *The Journal of Population Studies*, No.18, pp13-28
- 2007, " Household projection 2006/7 in Japan using a micro-simulation model", National
Institute of Population and Social Security Research
- Harding, A. 1996, Introduction and Overview. in Harding(ed.), *Microsimulation and Public
Policy*, North-Holland.
- Nishioka, H. Suzuki, T., Koyama, Y., Yamato C., Kojima K. 2011, "Household Projections for
Japan, 2005-2030 : Outline of Results and Methods", *The Japanese Journal of Population*.
Vol 9. No. 1.
- Smith, L., Hyndman, R.J., Wood. S. N, 2004 "Spline Interpolation for Demographic Variables:
the Monotonicity Problem", *Journal of Population Research*, Vol. 21, No.1.
- Suzuki, T., 1999. *Projection of Households in Japan with a Dynamic Macro-Simulation Model* (Ph.D
Dissertation, University of California at Berkeley)
- Yamauchi, M., Suga, K. 2011. "Household Projections by Prefecture in Japan : 2005-2030 Outline
of Results and Methods"
- Zeng, Y., Vaupel, J.W. Wang, Z. 1997. "A Multidimensional Model for Projecting Family
Households -- with an illustrative numerical application." *Mathematical Population
Studies* 6(3)
- Zeng, Y., Land, K.C. Wang, Z. Danan. G, 2003, "US. Family Household Dynamics and
Momentum-Extensions of Profamy Method and Application"
- National Institute of Population and Social Security Research. 2008 "Household Projections for
Japan :2005-2030"
- National Institute of Population and Social Security Research. 2010 "Household Projections by
Prefecture in Japan :2005-2030"
- US Bureau of Census. 1996. *Projections of the number of Households and Families 1995-2010*.

Washington

- The Office of the Deputy Prime Minister. 2010, *Updating the Department for Communities and Local Government's household projections to a 2008 base : methodology*, DCLG Publications
- General Register Office for Scotland. 2010, *Household Projections for Scotland 2008-based*
- Office for National Statistics. 2010, "*Marital Status Population Projections, 2008-based*"
- Australian Bureau of Statistics. 2010, "Household and Family Projections, Australia, 2006-2031"
- Statistics New Zealand. 2010, *New Zealand Family and Household Projection 2006(base)-2031 update*
- Len, S. Hyndman, R. Wood, S. 2004 "Spline Interpolation for Demographic Variables: the Monotonicity Problem", *Journal of Population Research*. Vol 21, No.1.