

생존모형을 활용한 한국의 출산력 변동 분석

우해봉¹⁾ · 장인수²⁾

요약

초저출산 현상이 장기화됨에 따라 출산에 대한 학술적 및 정책적 관심이 크게 증가하고 있다. 최근까지 국내에서 이루어진 출산 관련 경험적 연구들은 대체로 출산 계획 혹은 출산 행위에 유의한 영향을 미치는 요인들을 확인하는 작업에 초점을 맞추는 경향이 있었다. 선행 연구들과 비교하여 본 연구는 생존분석을 활용하여 합계출산율(TFR)과 같은 인구학적 지표를 산출하는 방식을 통해 우리나라의 출산력 변동을 분석하고 있다. 특히, 일반적인 연령별 출산율(ASFR) 대신 출산진도비(PPR) 지표에 기초한 분석을 진행함으로써 전체 가임 생애에 걸쳐 이루어지는 가족 규모의 분포와 관련된 정보를 제공하고자 하였다. 분석 결과는 출산율이 대체 수준 아래로 하락한 1980년대 이후의 출산력 변동이 그 이전 기간에 비해 더욱 가파른 모습을 보여 주고 있지만, 가족계획 사업이 본격적으로 추진되기 시작한 1960년대 중반부터 1970년대 기간에도 이미 셋째 이상 고순위 출산으로의 이행 과정에서 상당한 수준의 변화가 있었음을 보여 주고 있다. 교육 수준별 출산력 변동 분석 결과는 한국 사회가 보여 준 과거의 급격한 출산력 변천 과정에서 교육 수준이 높은 여성들이 선도적 역할을 하였음을 시사하고 있다.

주요용어 : 출산력 변동, 생존분석, 혼인 상태, 교육 수준

1. 서론

한국 사회는 2001년부터 현재까지 합계출산율이 1.3 아래의 초저출산(lowest-low fertility) 현상에 직면하고 있다. 저출산 현상이 심각한 사회적 문제로 부각됨에 따라 2006년부터 정부 주도의 제1차 저출산·고령사회 기본계획이 수립되었으며, 2016년부터는 제3차 저출산·고령사회 기본계획이 추진되고 있다. 제3차 기본계획의 경우 합계출산율(TFR) 지표 기준으로 2020년 1.5, 2030년 1.7, 그리고 2045년 2.1 도달을 최종적인 목표로 설정하고 있다(대한민국정부, 2015).³⁾ 정부 주도의 저출산 대응 정책이 추진된 지 이미 10년을 경과한 시점이지만, 현재까지도 출산율 회복의 징후는 관측되지 않고 있다. 더욱이 2017년의 경우 합계출산율이 역대 최저 수준인 1.05까지 떨어졌으며(잠정치), 출생아 또한 공식적인 통계 작성 이래 처음으로 40만 명 아래로 떨어졌다(통계청, 2018). 다른 한편으로 저출산·고령사회 기본계획이 저출산 문제 해결을 위한 뚜렷한 장기 방향을 제시하고 있는지도 불명확한 상황이다. 과거 제1~2차 기본계획은 물론 2015년에 수립된

1) 주저자 및 교신저자. 한국보건사회연구원 연구위원. E-mail: haebongwoo@kihasa.re.kr

2) 한국보건사회연구원 전문연구원. E-mail: sescis@kihasa.re.kr

3) 2015년에 수립된 제3차 기본계획은 기존 1~2차 기본계획이 미시적이고 현상적인 접근이었음을 지적하며 저출산·고령사회에 관한 종합적이고 구조적인 접근을 통해 저출산·고령사회 대응 패러다임 전환을 시도하고 있음을 표방하고 있다(대한민국정부, 2015, p.37). 그러나 제3차 저출산·고령사회 기본계획이 기존 1~2차 기본계획과 구분되는 패러다임 전환으로 평가될 수 있는지는 의문이다. 현 정부에서 제3차 기본계획을 재검토하고 있는 것은 이와 무관하지 않다고 볼 수 있다.

제3차 기본계획 또한 기본적으로 합계출산율 상승에 정책의 초점을 맞춘 반면, 현 정부는 출산율 상승 대신 「가족 행복」에 초점을 맞춘 정책으로의 방향 전환을 모색하고 있는 것으로 알려진다. 그러나 현 정부에서 추진 중인 저출산 대응 정책이 과연 기존 정책 방향과 어떠한 차별성을 지닐 것인지는 현재까지 명확하게 제시되지 않고 있다.

한국 사회에서 저출산 문제의 심각성으로 인해 다양한 정책적 논의가 이루어진 것과 비교하여 출산력 변천 과정과 최근 동향에 관한 경험적 연구는 상대적으로 제한적이다. 물론 이렇게 출산력과 관련된 경험적 연구가 부족한 것은 출산율이 대체 수준 아래로 떨어진 1980년대 중반 이후부터 2000년대 이전까지 저출산 현상이 갖는 사회적, 정책적 함의가 충분히 인식되지 못한 것과도 연관되지만, 다른 한편으로 활용 가능한 자료가 크게 부족하다는 점과도 관련이 있다. 우리나라의 경우 현 시점에서 활용 가능한 대부분의 사회조사 자료들은 혼인이나 출산과 관련된 정보를 「부수적」 정보 차원에서 수집하는 관계로 이에 기초하여 출산력 변동을 심층적으로 분석하기는 쉽지 않다. 이와 함께 출산 현상에 대한 분석이 이루어진 경우에도 분석 결과를 바탕으로 인구학적으로 의미 있는 지표들을 산출하는 대신 출산 계획(birth plan) 혹은 출산 행위(birth behavior)에 통상적인 계량분석 기법을 다소 기계적으로 적용하고 해석함으로써 경험적 분석이 갖는 효용성을 극대화하지 못한 측면이 있었다. 이러한 측면에서 본 연구는 출산력에 대한 경험적 분석을 진행하되 단순한 계량분석 기법의 적용을 넘어 합계출산율(Total Fertility Rate: TFR)이나 출산진도비(Parity Progression Ratio: PPR)와 같이 인구학적으로 보다 의미 있는 지표들을 산출함으로써 우리나라 출산력 변동에 관한 추가적인 정보를 제공하고자 하는 목적을 가지고 있다. 특히, 본 연구는 일반 서베이조사 혹은 패널조사 자료의 활용 가능성이 점차 높아지고 있는 상황에서 출산력 분석에서 생존분석(survival analysis)의 효용성을 보여 주는 데 초점을 맞추고자 한다.

논문의 구성과 관련하여 제2장에서는 국내에서 이루어진 출산 관련 경험적 연구들의 동향을 살펴보는 한편 계량분석 기법을 적용하여 인구학적 지표들을 산출하는 최근까지의 국외 선행 연구들을 간략히 살펴본다. 제3장에서는 본 연구에서 사용하는 자료에 대한 소개와 함께 출산력 변동을 살펴보기 위한 계량분석 방법 그리고 이러한 계량분석에 기초하여 인구학적 지표들이 산출되는 세부적인 절차에 대해 살펴보기로 한다. 제4장에서는 과거 50년에 걸친 전반적인 출산력 변동과 함께 개인들의 사회경제적 지위(Socio-Economic Status: SES)를 표상하는 핵심 지표 중의 하나인 교육 수준 변수를 중심으로 출산력 변동 과정에서 하위 집단별로 어떠한 차별적 패턴이 존재하는가를 살펴본다. 마지막으로 제5장에서는 본 연구의 분석 결과를 종합하는 한편 향후 저출산 대응 정책 수립과 관련된 시사점 및 후속 연구에서 검토될 필요가 있는 사항들을 간략히 정리하기로 한다.

2. 선행 연구

저출산 현상이 사회적 현안으로 등장함에 따라 출산 관련 인구학적 연구들도 최근 들어 크게 증가하는 모습을 보이고 있다. 1960~1980년대 출산 억제가 주요 정책 과제로 등장한 시기 이후 국내 인구학에서 출산 관련 연구는 상당히 제한적으로만 이루어졌다. 그러나 1990년대 중반 이후 정부가 공식적으로 출산 억제 정책을 폐기하는 한편 2000년대 중반 들어 저출산·고령사회 기본계획을 수립하면서 출산 관련 연구들이 본격적으로 증가하는 양상을 보여 주고 있다. 비록 국내에서 2000년대 이후 출산 관련 연구들이 활성화되는 양상을 보이고 있지만, 현재까지의 경험적 연구들은 주제 및 접근 방식에서 상당히 유사한 패턴을 보여 주고 있다.

출산과 관련하여 국내에서 이루어진 기존의 경험적 연구들은 기본적으로 출산 계획(birth plan)(김정석, 2007; 박수미, 2008; 우해봉, 장인수, 2017; 정은희, 최유석, 2013; 정혜은, 진미정, 2008)이나 출산 행위(birth behavior)(김현식, 2017; 류기철, 박영화, 2009; 민현주, 2007; 민현주, 김은지, 2011; 송헌재, 2012; 우해봉, 장인수, 2017; 은기수, 2001; 이성용, 2009; 정은희, 최유석, 2013)를 결과 변수(outcome variable)로 하여 출산 계획과 출산 행위에 유의한 영향을 미치는 변수들(determinants)을 확인하는 데 분석의 초점을 맞추고 있다. 출산 행위를 분석할 경우 전체 출생 순위를 모두 고려하는 대신 첫째 출산 혹은 둘째 출산 등 특정 순위의 출산에 초점을 맞추는 경우가 일반적이다. 저출산 현상이 장기간 지속됨으로써 고순위(high order) 출산 사례가 제한적이라는 점에서 출산 행위를 분석한 대부분의 연구들은 셋째 이하 출산을 분석 대상으로 하고 있다.

비록 최근까지 국내 연구에서 사용된 방식이 전통적으로 인구학적 논의에서 빈번하게 활용된 접근이기는 하지만, 이러한 접근을 통해 출산 생애과정(reproductive life-course)과 관련된 의미 있는 정보를 제공하는 것이 쉽지는 않다. 앞에서 언급했듯이, 대부분의 기존 출산 관련 경험적 연구들은 첫째 출산이나 둘째 출산과 같은 특정 순위의 출산 분석에 초점을 맞추고 있는데, 이를 통해서도 출산(율)과 관련된 전반적인 변화 양상을 판단하기가 쉽지 않다. 예컨대, 특정 순위에 한정된 출산 패턴에서의 변화만을 분석해서는 합계출산율(TFR)처럼 전반적인 출산율에서 어떠한 변화가 나타났는가를 판단하기가 쉽지 않다. 물론 합계출산율 지표의 경우 통계청이 지속적으로 작성 및 공표하고 있는 지표이다. 비록 개인 연구자 차원에서 인구주택총조사 표본과 인구동향조사를 활용한 사회경제적 특성별 차별 출산력에 관한 연구들이 이루어지고 있지만(예컨대, 최은영, 박영실, 2009), 통계청(KOSIS)이 공식적으로 공표하는 합계출산율은 1970년 이후 우리나라 전체의 출산율 정보만을 제공하는 관계로 사회 집단별로 출산율에서 어떠한 차이가 존재하는가를 보여주지는 않는다.

또한 통계청이 작성 및 공표하는 합계출산율은 연령별 출산율(Age Specific Fertility Rate: ASFR)에 기초하여 산출된 지표이다. 물론 출산력 연구에 있어서 연령의 중요성은 이미 잘 알려져 있다. 그러나 사망력에 비해 출산력 연구에서는 개인들의 의식적 선택 과정이 매우 중요한데 연령별 출산율(ASFR)에 기초한 접근의 경우 이와 관련된 정보는 극히 제한적이다(Hinde, 1998). 특히, 저출산 상황이 지속되고 있는 우리나라의 경우 생애에 걸쳐 이루어지는 출산과 관련된 일련의 선택 과정 및 이에 기초하여 완결되는 가족

규모의 분포(distribution of family sizes)를 이해하는 것은 매우 중요한 의미를 지닌다. 합계 출산율이 동일하더라도 자녀의 수적 분포는 매우 상이할 수 있다. 예컨대, 모든 여성들이 동일하게 일정 수준 이상의 출산을 하는 경우와 일부 여성들의 출산 횟수가 많은 반면 나머지 여성들은 생애 동안 전혀 혹은 상당히 적은 횟수의 출산을 하는 경우의 정책적 의미는 상이할 수 있는 것이다. 국내에서 이루어진 최근의 출산 관련 선행 연구들은 출산 계획 혹은 행위를 분석하는 통계적 방법 측면에서도 대체로 정형화된 모습을 보인다. 출산 계획의 경우 계획의 유무 측면에서 이항 변수(binary variable)를 분석하는 로짓분석(logit analysis)을 활용하는 것이 가장 일반적이다. 출산 행위의 경우 분석에 사용된 통계적 모형 측면에서는 출산 계획에 비해 상당히 다양하다. 출산 계획과 마찬가지로 출산 행위 분석에서도 통상적인 로짓 혹은 프로빗분석을 활용하는 연구도 있다(예컨대, 송헌재, 2012; 이성용, 2009; 정은희, 최유석, 2013). 반면 출산 행위 분석에서의 보다 일반적인 접근은 생존분석(survival analysis)을 사용하는 방식이라고 할 수 있다. 출산 행위 분석에서 생존분석의 활용도가 높은 것은 생존분석이 단순히 출산과 같은 사건(event)의 「발생 여부」뿐만 아니라 사건의 「발생 시점」에 관한 정보를 동시에 제공하는 것과 관련이 있다(Singer and Willett, 2003). 예컨대, 생존분석을 활용할 경우 가임 생애(예컨대, 15~49세)에 걸쳐 출산이 연령별로 어떠한 패턴을 보이는가를 파악할 수 있다. 물론 분석 대상 시간의 흐름(metric of time)이 반드시 연령일 필요는 없으며, 결혼 이후 출산까지의 기간 혹은 첫째 출산 이후 둘째 출산까지의 기간 등 상이한 방식으로 시간을 측정하는 것도 충분히 가능하다. 또한 생존분석은 관측 기간(observation window)의 절단/중결 관련 문제(truncation/censoring)를 통계적으로 적절히 처리함으로써 생애과정의 다양한 단계에 위치한 개인들을 효과적으로 분석 대상에 포함할 수 있는 장점이 있다(Von Hook and Altman, 2013). 생존분석의 구체적 모형화 과정에서는 대체로 콕스 모형(Cox hazard model)이나 이산형 모형(discrete-time hazard model)의 활용 빈도가 높은 모습을 보인다.

표본조사 자료에 생존분석 기법을 적용하여 합계출산율(TFR)과 같은 인구학적 지표들을 생산하기 위한 노력은 외국의 경우에도 상대적으로 최근에는 비로소 논의되기 시작한 영역이다(예컨대, De Santis, Drefahl, and Vignoli, 2014; Gray, Evans, Anderson, and Kippen, 2010; Hoem and Muresan, 2011a, 2011b; Retherford, Ogawa, Matsukura, and Eini-Zinab, 2010; Schoumaker, 2013; Van Hook and Altman, 2013). 그러나 출산력 정보를 제공하는 표본조사 자료의 활용 가능성 증대 및 통계적 방법론의 향상으로 인해 관련 연구들이 증가하고 있는 상황이다. 특히, 코호트 접근(cohort approach)에 기초하여 발전된 출산진도비 개념(Ryder, 1986)이 기간 접근(period approach)으로 확대(Feeney and Yu, 1987; Ni Bhrolchain, 1987)되면서 그 활용 가능성은 더욱 커진 상황이라고 할 수 있다. 활용 가능한 정보가 많을수록 보다 체계적인 분석이 가능하지만, 단순한 합계출산율(TFR) 지표 산출의 경우 특정 시점 기준으로 생애 동안의 출산 이력(날짜) 정보만 제공되면 충분하다는 점에서 자료에 대한 요구 수준 또한 높지 않다(Hinde, 1998). WFS(World Fertility Survey) 혹은 DHS(Demographic and Health Survey)와 같은 서베이 자료에 기초하여 합계출산율과 같은 인구학적 지표들을 산출하는 통계적 모형의 경우 포아송모형의 활용도가 높았지만(Pullum, 2004; Schoumaker, 2013), 최근에는 이산형 생존분석처럼 통상적인 생존분석 기법을 적용하는 사례 또한 증가하고 있다(Retherford et al. 2010; Schoumaker, 2013; Van Hook and Altman, 2013).

본 연구 또한 기존의 경험적 연구들처럼 표본조사 자료에 생존분석을 적용하여 출산 현상을

분석하지만, 이러한 분석 결과를 합계출산율(TFR)처럼 인구학적으로 보다 의미 있는 지표로 전환함으로써 국내에서 이루어진 기존 논의들을 확장한다. 다른 한편으로 본 연구에서는 합계출산율을 산출하는 과정에서, 전통적인 연령별 출산율(ASFR)에 기초한 접근 대신, 출산 생애과정에서의 선택적 과정을 보다 잘 보여 줄 수 있는 출산진도비(PPR)에 기초한 접근을 사용함으로써 우리나라 출산력 변천 과정에서 나타난 출생 순위별 분포에서의 특징을 살펴보고자 한다. 공통적으로 출산이라는 사건의 발생 여부를 분석하기는 하지만 기존 국내 연구들이 출산 간격에 초점을 맞춘 반면 본 연구는 상대적으로 출산의 양에 초점을 맞추는 차이가 있다. 기본적으로 출산 간격(birth interval) 분석이 출산의 속도(tempo of fertility)에 초점을 맞춘 반면 출산진도비(parity progression ratio)는 출산의 양(quantum of fertility)에 초점을 맞추는 차이가 있다(이흥탁, 1994). 출산진도비가 기간(period) 혹은 코호트(cohort) 접근에 기초하여 모두 산출 가능하지만, 가상 코호트(synthetic cohort)의 생애를 모형화하는 기간(period) 출산진도비의 경우에도 현 자녀 수(parity)를 조건으로 한 추가적인 출산 과정을 보여 줄 수 있다는 점에서, 연령별 출산율에 기초한 합계출산율과 달리, 코호트 출산율의 성격도 일정 부분 지니고 있다.

앞에서도 언급했듯이, 본 연구에서는 전반적인 출산력 변동과 함께 개인들의 사회경제적 지위를 표상하는 핵심 지표 중의 하나인 교육 수준별 출산력 변동 양상을 살펴본다. 외국과 마찬가지로 우리나라에서도 일반적으로 교육과 출산력은 부적(-) 관계를 갖는 것으로 논의된다. 그러나 교육과 출산 간의 관계를 매개하는 다양한 요인들이 존재하기에 그 관계가 항상 일정한 것은 아니다. 또한 부적(-) 관계가 시간의 경과에 따라 약화되거나 정적(+) 관계로 전환되는 경우도 발생함이 지적된다(이흥탁, 1994; 전광희, 2002). 과거 차별 출산력과 관련된 다양한 준거 변수들이 검토된 바 있지만(예컨대, 종교, 거주 지역, 직업, 가족제도 등),⁴⁾ 최근 들어 출산력에서의 집단 간 변이는 크게 줄어든 것으로 논의된다(최은영, 박영실, 2009). 본 연구와 마찬가지로 여성가족패널을 활용하여 출산 행위에 대한 다양한 요인들의 효과를 분석한 우해봉·장인수(2017)의 연구도 이러한 측면을 잘 보여 준다. 그럼에도 불구하고 교육 수준은 혼인 상태, 피임 실천과 같이 출산력을 결정하는 근접 요인들(proximate determinants)과 밀접히 연관되어 있음과 함께 한국 사회에서 출산과 관련된 사회적 규범 및 문화 측면에서도 그 파급 효과가 매우 크다는 점에서 여전히 중요한 함의를 가지고 있는 변수라고 할 수 있다. 또한 소득이나 자산 등 생애에 걸쳐 변동성이 강한 변수들과 달리 교육 수준은 청년기 이후 상대적으로 안정적인 성격을 지닌 장점이 있다. 마지막으로, 교육 수준은 시간적으로 직업이나 소득 활동에 선행하는 한편 이들 변수들에 유의한 영향을 미치는 요인이라는 점에서도 사회경제적 지위를 표상하는 변수로서 그 효용이 매우 크다.

4) 차별 출산력과 관련된 최근까지의 국내 논의들에 대해서는 이흥탁(1994), 전광희(2002)를 참고할 수 있다.

3. 자료 및 분석 방법

3.1 분석 자료

최근까지 한국 사회에서 진행된 출산력 변동을 분석하기 위해 본 연구에서 사용하는 분석 자료는 한국여성정책연구원이 2007년부터 구축해 오고 있는 「여성가족패널(KLoWF)」이다. 여성가족패널은 전국 대표성을 지닌 9천여 가구에 속한 19세 이상 64세 이하 여성들을 「개인조사」 대상으로 설정하고 있다. 현재까지 우리나라 출산력 분석과 관련하여 활용 가능한 자료는 제한적인데, 특히 생애에 걸친 여성의 출산 이력(birth history)과 관련된 정보를 제공할 수 있는 자료는 상당히 드물다. 이러한 측면에서 여성가족패널은 분석 대상자들의 최신 정보를 제공하는 동시에 과거 출산 이력에 관한 회고적(retrospective) 정보를 제공하고 있다는 점에서 그 효용이 상당히 높다고 할 수 있다. 여성가족패널을 활용한 출산력 분석은 2007년의 제1차 조사에서 수집된 개인조사 대상자들의 회고적 정보에 기초하여 진행할 수도 있지만, 본 연구에서는 회고적 자료는 물론 현 시점 기준 가장 최근 자료인 2014년의 제5차 개인조사까지의 모든 정보를 활용하기로 한다.

여성가족패널에 기초한 출산력 분석과 관련하여 언급할 사항으로 표집상의 특징이 있다. 비록 여성가족패널이 전국 대표성을 지니도록 설계되었지만, 우리나라 출산력 분석과 관련하여 매우 중요한 의미를 지니고 있는 「혼인 상태」 측면에서는 그 대표성이 다소 떨어지는 측면이 있다. 예컨대, 2005년 인구총조사 기준 19세 이상 64세 미만 중 미혼자의 비율이 24.0%인 반면 여성가족패널의 미혼자 비율은 19.1%(가중치 적용)에 그치고 있다. 또한 기혼자 중 출산 경험이 있는 개인들의 비율에서도 차이가 있는데, 2005년 인구총조사 10% 표본조사에서 20세 이상 64세 이하 여성 기혼자 중 출산 경험이 있는 여성의 비율이 93.2%인 반면 여성가족패널의 경우 기혼자 중 출산 경험이 있는 여성의 비율은 94.9%(가중치 적용)로 나타나고 있다(박수미 외, 2008). 비록 여성가족패널이 현 시점에서 우리나라 출산력 변동을 분석할 수 있는 중요한 자료원임에도 불구하고, 이러한 표집상의 특징으로 인해 여성가족패널에 기초하여 산출된 출산율은 통계청이 발표하는 출산율에 비해 높은 모습을 보인다(제4장 분석 결과 참고). 또 다른 관련된 특징으로 여성가족패널에서는 조사 과정상의 어려움으로 인해 기혼자에 한해 출산력 정보를 수집하고 있다. 물론 현재까지도 우리나라 전체 출생 건수에서 미혼자가 출산하는 출생아 수의 비율(1~2%)이 극히 낮다는 점에서 큰 문제를 초래하지는 않는다고 할 수 있다.

여성가족패널의 표집상의 특징을 고려하여 본 연구에서는 출산력 분석 모형 구축 과정에서 「혼인 상태」를 통제하는 방식을 취하기로 한다. 또한 통계청이 발표하는 수치에 상응하는 정확한 출산율 지표의 산출보다는 최근까지 진행된 출산율의 전반적인 변동 패턴(추이) 및 출산력 변동에서의 집단별(교육 수준별) 차이점 분석에 초점을 맞추기로 한다. 물론 이러한 접근은 본 연구만이 취하는 특징이 아니라 표본조사 자료에 통계적 모형을 적용하여 출산력을 분석하는 연구들이 일반적으로 취하는 방식이라고 할 수 있다(예컨대, Von Hook and Altman, 2013). 이는 표본조사에 기초하여 출산율 지표를 산출하는 접근의 한계일수도 있지만, 다른 한편으로 통계청과 같은 국가기관이 제공하는 출산율 지표와 달리, 출산과 유의하게 연관된 사회경제적 특성들을 통제한 상태에서 분석 대상 변

수(예컨대, 교육 수준)와 출산 행위 간의 연관성을 분석할 수 있다는 점에서 표본조사에 기초한 방식이 제공하는 장점 중의 하나라고 할 수도 있다.

최종 분석 표본의 구축과 관련하여 본 연구에서는 2007년 최초 조사 시점에서는 19세 미만이었지만, 「가구조사」의 가구원 추적 과정(제2~5차 조사)에서 19세 이상으로 「개인조사」 대상으로 새롭게 추가된 개인들을 모두 포함하고 있다(1941~1996년 출생자). 분석 방법 부분에서 보다 자세히 언급하겠지만, 본 연구에서 출산력 변동을 분석하기 위해 사용하는 통계적 모형은 이산형 생존분석(discrete-time survival analysis)이다. 본 연구에서 「이산형」 생존분석을 사용하는 것은 단순히 사건(출산) 발생 여부와 그 시점을 분석하는 것을 넘어 출산진도비(PPR)나 합계출산율(TFR)과 같은 인구학적 지표를 산출하는 것과 관련이 있다. 이산형 생존분석을 사용하는 본 연구에서 자료의 분석 단위는 개인-기간(person-period) 자료의 구조를 취한다. 결과적으로 개인-기간(연도) 자료의 형식에서 개인들은 복수의 레코드(에피소드)를 보유하게 되는데, 2014년 제5차 조사 시점까지 49세 이상 연령에 도달한 개인들의 경우 (출산력 분석 대상 연령 구간인) 15~49세까지 총 35개의 개인-기간(연도) 레코드가 관측될 수 있다. 물론 최종 관측 시점 기준으로 49세 미만의 경우 이보다 적은 수의 레코드가 존재하는데, 예컨대, 최종 관측 연령이 19세인 개인의 경우 15~19세까지 총 4개의 레코드만 존재하게 된다. 참고로, <표 3.1>은 분석 대상 표본 중 두 사례(교육 기간 16년(대졸) 및 12년(고졸))의 자료 구조를 예시적으로 보여 주고 있다.

제5차 기준으로 개인조사 연령 하한인 19세 이상 전체 개인들을 분석 대상으로 설정하고 있지만, 우리나라 출산력 변동의 시기별 패턴을 살펴보는 본 연구에서는 추가적으로 개인-기간 레코드를 특정 기간으로 한정하여 분석을 진행하기로 한다. 전체 분석 대상자들의 15~49세 연령대의 개인-기간 자료에 상응하는 역년(calendar year)을 살펴보면 최고 1956년에 해당하는 레코드까지 관측된다. 그러나 1950년대 후반과 1960년대 초반의 경우 분석 대상 레코드가 상대적으로 많지 않다는 점에서 본 연구에서는 2014년 기준으로 과거 50년에 해당하는 1965~2014년 기간에 걸친 개인-기간 레코드만을 분석하기로 한다. 추가적으로 1965~2014년의 전체 기간을 통합적으로 분석하는 대신 1965~2014년 기간을 하위 기간으로 구분하여 기간별로 독립적인 분석을 진행하기로 한다. 기간(시대) 구분에서 불가피하게 발생하는 작위성에도 불구하고 본 연구는 1965~2014년 기간을 1965~1983년, 1984~1997년, 1998~2014년의 세 기간으로 구분한다.

〈표 3.1〉 개인-기간(person-period) 자료의 기본 구조

개인 식별 번호	연도	연령	교육 수준(년)	혼인 상태	출산 여부	출생 순위 지표
302	1977	15	8	0	0	1
	1978	16	9	0	0	1
	1979	17	10	0	0	1
	1980	18	11	0	0	1
	1981	19	12	0	0	1
	1982	20	13	0	0	1
	1983	21	14	0	0	1
	1984	22	15	0	0	1
	1985	23	16	0	0	1
	1986	24	16	1	0	1
	1987	25	16	1	1	1
	1988	26	16	1	0	2
	1989	27	16	1	0	2
	1990	28	16	1	1	2
	1991	29	16	1	0	3
	1992	30	16	1	0	3
	1993	31	16	1	0	3
	1994	32	16	1	0	3
	1995	33	16	1	0	3
	1996	34	16	1	0	3
	1997	35	16	1	0	3
	1998	36	16	1	0	3
1999	37	16	1	0	3	
2000	38	16	1	0	3	
2001	39	16	1	0	3	
2002	40	16	1	0	3	
2003	41	16	1	0	3	
2004	42	16	1	0	3	
2005	43	16	1	0	3	
2006	44	16	1	0	3	
2007	45	16	1	0	3	
2008	46	16	1	0	3	
2009	47	16	1	0	3	
2010	48	16	1	0	3	
2011	49	16	1	0	3	
702	1988	15	8	0	0	1
	1989	16	9	0	0	1
	1990	17	10	0	0	1
	1991	18	11	0	0	1
	1992	19	12	0	0	1
	1993	20	12	0	0	1
	1994	21	12	0	0	1
	1995	22	12	0	0	1
	1996	23	12	1	0	1
	1997	24	12	1	0	1
	1998	25	12	1	0	1
	1999	26	12	1	0	1
	2000	27	12	1	1	1
	2001	28	12	1	0	2
	2002	29	12	1	0	2
	2003	30	12	1	0	2
2004	31	12	1	0	2	
2005	32	12	1	0	2	
2006	33	12	1	0	2	
2007	34	12	1	0	2	
2008	35	12	1	0	2	
2009	36	12	1	0	2	
2010	37	12	1	0	2	
2011	38	12	1	0	2	
2012	39	12	1	0	2	
2013	40	12	1	0	2	
2014	41	12	1	0	2	

우선, 1965~1983년 기간은 출산율이 지속적으로 감소하여 이른바 대체출산율(replacement-level fertility) 수준까지 감소하는 시기에 해당한다. 1984년 이후 합계출산율(TFR)은 2명 미만으로 떨어졌는데, 추가적으로 본 연구에서는 1997년 말의 IMF 경제 위기를 우리나라 출산력 변천에서 큰 변화가 이루어진 중요한 기점으로 보고 1984~1997년 기간과 1998~2014년 기간을 구분하고 있다. 참고로, IMF 경제 위기와 출산 간의 관계를 살펴본 선행 연구들은 경제 위기 이후 출산력 패턴에서 뚜렷한 변화가 존재함을 지적하고 있다(김두섭, 2007; 이성용, 2006). 앞서서도 이미 언급했듯이, IMF 경제 위기 이후 2001년부터 현재까지 합계출산율은 이른바 초저출산(lowest-low fertility)의 기준으로 지칭되는 1.3 미만 수준에서 고착된 모습을 보이고 있다는 점에서도 기존 시기와는 구분되는 측면이 있다고 볼 수 있다. 앞에서 언급한 조건들을 충족한 개인-기간 레코드는 총 284,065개에 해당하며, 이들 개인-기간 레코드를 개인 단위(person-level) 자료로 환원할 경우 분석 대상 개인들은 총 10,855명이다.

3.2 분석 방법

한국 사회의 출산력 변동과 관련하여 본 연구에서 사용하는 분석 방법은 생존모형(survival model)과 생명표모형(life table model)이다. 다른 사회과학 분야와 마찬가지로 생존분석은 인구학에서도 매우 중요한 위치를 점하고 있다. 기본적으로 생명표모형 또한 생존분석과 밀접히 연관되어 있다. 이미 수 세기 전부터 사용된 인구분석 도구인 생명표모형이 지난 20세기 후반부 이후 활용도가 크게 높아진 것은 바로 생존분석이라는 통계적 모형의 등장과 밀접한 관련이 있다고 할 수 있다. Cox(1972)에 의해 소개된 생존분석은 전통적인 생명표의 단순한 연령(기간)별 사건 발생 리스크에 대한 모형화를 넘어 다양한 준거 변수에 기초한 생명표(life-tables with covariates) 작성의 가능성을 열었다고 할 수 있다. 생명표 작성 과정에서 다양한 변수들을 고려하는 것은 출산 행위와 연관된 인구 구성에서의 차이(compositional differences)를 통제한 상태에서 집단별 혹은 시기별로 출산력을 비교할 수 있다는 점에서도 그 효용성이 상당히 크다고 할 수 있다.

앞에서 언급했듯이, 생명표모형을 활용하기 위해 본 연구에서는 다양한 생존분석 기법 중 이산형 생존분석(discrete-time survival analysis)을 사용한다. 이산형 생존분석에서의 핵심적인 고려 사항으로는 연계함수(link function)의 선택과 기준선 해저드 함수(baseline hazard function)의 모형화가 있다. 연계함수의 경우 일반적으로 로짓(logit) 혹은 보 로그-로그(complementary log-log) 연계함수 중 하나를 사용하는데 본 연구에서는 보 로그-로그 연계함수를 사용한다. 일반적으로 사건 발생 리스크가 상당히 높지 않을 경우 두 연계함수에 기초한 분석 결과에서는 큰 차이가 없는 것으로 알려진다(Singer & Willett, 2003). 모수 추정치를 지수화(exponentiation)한 값이 해저드 비(hazard ratio)라는 점에서 보 로그-로그 연계함수를 사용한 이산형 생존분석은 사회과학에서 광범위하게 활용되는 콕스 회귀모형(Cox regression model)의 이산형판이라고 볼 수 있다. 이러한 점에서 보 로그-로그 연계함수는 특히 생명표 작성과 관련하여 유용한 방식이라고 할 수 있다.

다음으로 기준선 해저드 함수의 선택 문제가 있다. 시간의 흐름을 측정하는 다양한 방식이 존재하지만, 통상적으로 출산력 분석에서 연령에 따른 변화가 관심의 초점이라는 점

에서 본 연구 또한 연령에 따른 출산 리스크에서의 변화를 살펴보기로 한다. 여성의 가임 생애 기간이라고 할 수 있는 15~49세에 걸친 출산 리스크의 형태를 통계적으로 모형화 하기 위해 우선적으로 고려할 수 있는 접근은 연령(범주) 더미 변수(dummy variable)를 활용하는 것이다(예컨대, Retherford et al. 2010; Von Hook and Altman, 2013). 각 연령(범주) 더미를 사용하여 기준선 해저드의 형태를 표시하는 방식은 상대적으로 기준선 해저드 형태에서의 유연성을 확보할 수 있는 장점이 있다. 본 연구 또한 연령(범주) 더미 접근의 사용 가능성을 검토하였지만, 셋째 출산 이상 고순위 출산(high-order birth)의 경우 분석 대상 시기에 따라 연령(범주)별 출생 건수가 제한적인 문제로 인해 그 적용이 쉽지 않은 것으로 판단하였다(특히, 1998~2014년 기간).

기준선 해저드를 표상하기 위해 더미 변수 접근에 대한 대안으로 본 연구에서는 함수형(functional) 형태의 기준선 해저드를 사용한다. 함수형 기준선 해저드는 더미 변수 접근에 비해 기준선 해저드의 형태와 관련된 제약이 큰 단점은 있지만, 연령별 출생 건수가 제한적인 경우에도 효과적으로 사용될 수 있다. 또한 비비례적 해저드 모형(non-proportional hazard model)을 추정하는 과정에서 수많은 상호작용항을 모형에 포함해야 하는 더미 변수 접근에 비해 보다 간명한 접근이 가능한 또 다른 장점이 있다. 본 연구의 관측 구간이 15~49세라는 점에서 연령 하한인 15세로 중심이동(centering)이 이루어진 기본 모형은 아래와 같이 나타낼 수 있다.

$$\log(-\log(1 - h(t_{ij}))) = \alpha_0 + \alpha_1(Age_{ij} - 15) + \alpha_2(Age_{ij} - 15)^2 + \sum_k^k \beta_k X_{ij}$$

위에서 논의한 이산형 생존분석에 기초하여 모수 추정치들이 산출되며, 역변환(inverse transformation) 과정을 거쳐 연령별 출산 확률의 도출이 가능하다. 여기에서 연령별 출산 확률은 해당 연령 도달 이전에 (특정 순위의) 출산을 하지 않았을 것을 조건으로 해당 연령에서 (특정 순위의) 출산을 할 조건부 확률(conditional probability; hazard)을 의미한다. 이 조건부 확률이 바로 생명표에 나타나는 q_x 지표에 해당한다. q_x 지표 산출과 관련하여 고려할 필요가 있는 사항으로 출생 순위의 문제가 있다. 사망과 같은 비반복적 사건(non-recurrent event)과 달리 출산은 반복 사건(recurrent event)으로 출생 순위(birth order)가 중요한 의미를 지니며, 출산진도비(*PPR*)와 같은 인구학적 지표 또한 이를 명시적으로 고려하고 있다. 반복 사건을 모형화하는 다양한 접근이 검토된 바 있지만, 본 연구에서는 인구분석 모형에서 가장 선호되는 「출생 순위별」 접근을 사용하기로 한다. 다만, 출생 순위별 출생 건수의 분포를 고려하여 1) 첫째 출산, 2) 둘째 출산, 3) 셋째 이상 출산의 세 범주로 구분하여 출생 순위별로 각각 독립적인 모형을 구축하기로 한다.

이산형 생존분석과 관련하여 본 연구에서는 예측 변수로 혼인 상태와 교육 수준 두 변수를 사용한다. 통계적 모형에 기초하여 인구학적 지표를 산출하는 본 연구의 특성상 이들 변수 외에도 다양한 변수들의 사용 또한 충분히 가능하지만, 1965~2014년 전체 기간 및 전체 출생 순위를 아울러 출산 행위를 설명하는 유의한 변수들을 찾기 쉽지 않다는 점에서(특히, 1998~2014년 기간) 본 연구에서는 상대적으로 간명한 모형을 구축하였다. 혼인 상태(결혼)는 현재까지도 한국 사회에서 출산을 위한 가장 중요한 전제 조건의

성격을 지니고 있으며, 개인들의 사회경제적 지위(SES)를 표상하는 대표적인 변수 중의 하나인 교육 수준은 전통적으로 출산력 변천을 설명하는 데 있어서 핵심적인 역할을 하는 변수로 논의된다. 참고로, 다른 사회과학 분야에서 사용되는 다변량 모형에 비해 상당히 단순한 모형으로 보이기는 하지만, 「생명표 분석」에서 연령 외에 혼인 상태나 교육 수준을 동시에 고려한 모형은 상당히 복잡한 모형임이 지적될 필요가 있다. 예컨대, 통계청이 공표하는 생명표의 경우 연령별 사건(사망) 발생 리스크의 패턴을 기술하고 있을 뿐(성별을 제외한) 추가적인 준거 변수는 사용되지 않는다.

교육 수준(years of education)과 혼인 상태(marital status) 변수 모두 시변(time-varying) 변수로 사용된다. 교육 수준은 연속형 변수, 혼인 상태는 더미 변수의 형태를 취한다(유배우 = 1)(앞의 <표 3.1> 참고). 교육 수준의 경우 여성가족패널에서는 최종 관측 기간까지의 교육 이수 기간(년) 정보가 도출되는 반면 교육 이수 시점(연도)에 관한 정보는 존재하지 않는다. 본 연구에서는 최종 교육 이수 기간(년) 정보에 기초하여만 7세부터 최종 관측 시점에서 보고된 (최종) 교육 이수 기간(년)까지 연속적으로 교육을 받은 것으로 가정하였다. 비록 연속적 교육 이수 가정에서 벗어나는 사례가 존재할 수는 있지만, 군복무로 인한 교육 단절의 개연성이 높은 남성에 비해 상대적으로 체계적 오류의 개연성은 크지 않을 것으로 추정해 볼 수 있다.

다음으로, 생애 전체의 출산 이력(birth history)을 분석하는 본 연구에서 현 시점 대신 생애 전체에 걸친 혼인 상태 변동을 고려하는 것이 바람직하다. 다만, 출산 이력과 달리 생애 전체에 걸친 혼인 상태 이력 정보가 제공되지 않는다는 점에서 여성가족패널에서 제공하는 정보를 최대한 활용하여 혼인 상태를 시변(time-varying) 변수로 구성하였다. 우선 초혼 시점과 현 배우자와의 혼인 시점 정보를 비교하여 생애에 걸쳐 혼인 상태에서 변동이 없는 개인들의 경우 초혼 시점 이후 현재까지 유배우 상태를 유지하는 것으로 변수가 구성된다. 여성가족패널의 경우 사별이나 이혼 등 혼인 상태에서 변동이 있는 경우 가장 최근 배우자와의 혼인 이력 정보만이 제공되는데, 본 연구에서는 해당 정보를 이용할 수 있는 경우 시변(time-varying) 형식의 혼인 상태 변수 구성 과정에서 이를 반영하였다. 결국 여성가족패널에서 제공되는 자료 제공상의 한계로 인해 생애에 걸친 혼인 상태 변동에 있어서 유배우 상태가 과대 평가될 개연성이 있다고 할 수 있다.

생존분석에 기초하여 후속적으로 이루어지는 출산력 변동 분석과 관련하여 본 연구에서는 통상적인 연령별 출산율(*ASFR*)에 기초한 분석 대신 출생 순위(birth order)에 기초한 분석을 진행한다. 통계청에서 제공하는 연령별 출산율 및 이에 기초하여 산출된 합계 출산율(*TFR_{asfr}*)은 출산력이 연령별로 변이를 보인다는 중요한 사실을 반영한다. 그러나 다른 한편으로 연령별 출산율에 기초한 출산력 분석은 상대적으로 출생 순위에 대한 정보를 충분히 제공하지 못하는 단점이 있다. 사망력과 구분되는 출산력의 가장 큰 특징 중의 하나는 출산 행위에는 개인들의 의식적 선택이 중요한 역할을 한다는 점이다(Hinde, 1998). 사실, 심각한 저출산 문제에 직면한 한국 사회에서 출산력 관련 더욱 큰 관심의 대상은, 단순한 출산율 수준이 아니라, 생애 동안 출산을 하지 않는 여성들의 비율, 첫째 자녀를 출산한 여성들 중 둘째 자녀를 출산하는 여성들의 비율 등과 같은 출생 순위(birth order) 관련 측정치라고 할 수 있다. 터울 조절(birth-spacing)과는 구분되는 이러한 출산 제한(birth-limiting) 행위의 패턴을 이해하는 데 있어서 매우 유용한 지표가 바

로 출산진도비(PPR)이다(Preston, Heuveline and Guillot, 2001). 이러한 측면에서 본 연구에서는 출생 순위에 기초한 출산진도비를 산출하고 이에 기초하여 합계출산율(TFR_{ppr})을 산출하는 접근을 취하기로 한다. 보다 구체적으로, t 번째 출산에서 $t+1$ 번째 출산으로 이어지는 출산진도비($P_{t \rightarrow t+1}$)는 다음과 같이 정의된다. 여기서 N_t 는 t 번째 자녀를 출산한 여성의 수를, 그리고 N_{t+1} 은 $t+1$ 번째 자녀를 출산한 여성의 수를 나타낸다.

$$PPR_{t \rightarrow t+1} = P_{t,t+1} = \frac{N_{t+1}}{N_t}$$

출산진도비와 합계출산율(TFR_{ppr}) 간의 관계를 이해하기 위해서는 t 번째 출생 건수는 t 번째 출산을 한 여성(母)의 수와 동일하다는 점을 인식할 필요가 있다. 설명의 편의를 위해 출생코호트 시각에서 합계출산율 산출 방식을 살펴보면, 합계출산율은 해당 코호트 여성이 출산한 전체 출생 건수를 해당 코호트 여성의 수로 나눈 값이다. t 번째 출생 건수가 t 번째 출산을 한 여성(母)의 수와 동일하다는 점을 고려하면 해당 코호트의 합계출산율은 아래와 같이 나타낼 수 있다. 여기서 N_0 은 해당 코호트에 속한 여성의 전체 수, N_t 는 해당 코호트 여성 중 t 번째 자녀를 출산한 여성의 수, w 는 자료에서 관측된 최고의 출생 순위를 각각 나타낸다. 참고로, 다태아 출산이 이루어질 경우 본 연구에서는 다태아 중 첫 번째 출산만이 분석 대상이 된다. 다태아 사례 수가 제한적인 동시에 출생 연도가 동일한 다태아의 경우 개인-기간(년) 자료 분석 측면에서 모형화가 가능하지 않다. 다른 한편으로 출산진도비가 기본적으로 개인들의 선택 과정에 초점을 맞추고 있다는 점에서 개인들의 의식적 선택이 작용하기 어려운 다태아 출산은 관심의 대상이 아니라고 할 수 있다(Hinde, 1998).⁵⁾

$$\begin{aligned} TFR &= \frac{\sum_{t=1}^w N_t}{N_0} = \frac{N_1 + N_2 + \dots + N_w}{N_0} = \frac{N_1}{N_0} + \frac{N_2}{N_0} + \dots + \frac{N_w}{N_0} \\ &= \frac{N_1}{N_0} + \frac{N_1}{N_0} \times \frac{N_2}{N_1} + \dots + \frac{N_1}{N_0} \times \frac{N_2}{N_1} \times \dots \times \frac{N_w}{N_{w-1}} \\ &= P_{0,1} + P_{0,1} \times P_{1,2} + \dots + P_{0,1} \times P_{1,2} \times P_{2,3} \times \dots \times P_{w-1,w} = \sum_{w=1}^w \left(\prod_{t=1}^w P_{t-1,t} \right) \end{aligned}$$

5) 본 연구처럼 생존모형 및 생명표에 기초하여 산출된 합계출산율(TFR_{ppr})과 연령별 출산율을 합산하는 통계청의 합계출산율(TFR_{asfr})은 동일한 자료를 사용하더라도 그 결과가 상이함이 지적될 필요가 있다. 통계청의 합계출산율 산출 방식처럼 연령별 출산율에 기초한 방식은 출생 순위와 관련 없이 해당 연령의 모든 여성들이 분모에 포함된다. 반면 생존모형 및 생명표에 기초한 방식에서는 분석 대상 사건(event)을 이미 경험한 여성들은 분모에 포함되지 않는다. 인구학에서는 전자를 제2종 율(rates of the second kind or incident rates) 그리고 후자를 제1종 율(rates of the first kind or hazard rates)로 구분한다. 출산, 사망, 혼인 등 인구학적 율(rate) 개념과 관련된 보다 자세한 내용에 대해서는 Bongaarts & Feeney(2006)를 참고할 수 있다.

4. 분석 결과

분석 결과는 출산율 지표의 산출 과정, 최근까지의 전반적인 출산력 변동, 그리고 출산력 변동에서 교육 수준별 패턴 분석의 순서로 살펴보기로 한다. 우선, <표 4.1>은 본 연구에서 고려한 일련의 생존분석 모형의 추정 결과를 보여 주고 있다. <표 4.1>의 상단은 본 연구에서 검토하는 기본 모형이다. 연령만을 고려하는 더욱 단순한 모형의 사용을 고려할 수도 있지만, 분석 방법 부분에서 언급했듯이, 우리나라 출산력 분석에서 혼인 상태의 중요성 그리고 기혼자가 과대 표집된 여성가족패널의 특성을 고려하여 본 연구에서는 「혼인 상태」 변수를 포함한 모형을 기본 모형으로 사용하였다. <표 4.1>의 하단은 검토 가능한 다양한 모형 중에서 선택된 최종 모형으로 모형의 분류상 혼인 상태와 교육 수준 간의 상호작용은 물론 이들 변수들의 효과가 시간(연령)의 흐름에 따라 상이하게 나타날 수 있음을 허용하는 비비례적 해저드 모형(non-proportional hazard model)이다. 앞에서도 언급했듯이, 본 연구에서는 출생 순위별 특징을 반영하는 모형을 구축하고자 각 모형별로 첫째, 둘째, 셋째 이상 출산으로 구분하여 별도의 모형을 추정하였다. 본 연구에서는 <표 4.1> 상단의 기본 모형에 기초하여 1960년대 중반 이후부터 최근까지의 전반적인 출산력 변동을 살펴보고, <표 4.1> 하단의 최종 모형에 기초하여 최근까지의 출산력 변천 과정에서 교육 수준별 변화 양상을 세부적으로 살펴보기로 한다.

앞의 분석 방법 부분에서 언급했듯이, 본 연구에서는 기준선 해저드를 모형화하는 과정에서 함수형 기준선 해저드를 사용하는데 <표 4.1>의 연령 및 그 제곱항의 계수는 통상적인 출산의 연령별 패턴을 반영하듯이 오목형의 포물선 형태를 취하고 있다. 혼인 상태(유배우) 변수의 계수에서 볼 수 있듯이 결혼과 출산은 시기 및 출생 순위를 가로질러 정적(+) 관계를 보여 주고 있는데, 모수 추정치를 지수화할 경우 해저드 비(hazard ratio)가 산출된다. <표 4.1> 상단의 기본 모형에 기초하여 출산진도비 및 이에 기초한 합계출산율을 산출하기 위해서는 연령별 (조건부) 출산 확률(해저드)을 산출할 필요가 있다. 물론 이 과정은 앞의 분석 방법 부분에서 제시된 모형식에서 역변환(inverse transformation) 과정을 통해 쉽게 도출할 수 있다. 다만, 기본 모형의 경우 연령 외에도 모형식에 혼인 상태 변수가 포함되어 있다는 점에서 혼인 상태 변수를 일정한 값으로 고정시킬 필요가 있다. 비록 통계적 통제 측면에서 평균값, 중앙값, 최솟값, 최댓값 등 어떠한 값을 사용할 수도 있지만, 본 연구에서는 표본에서 관측된 각 기간별-연령별 혼인 상태 평균값으로 고정하는 방식을 사용하였다.

이러한 과정을 거쳐 산출된 연령별 출산 확률(q_x)에 기초하여 합계출산율과 같은 출산력 지표를 산출하는 과정의 일부가 <표 4.2>에 예시적으로 제시되어 있다. 보다 구체적으로, <표 4.2>는 <표 4.3>에 제시된 1965~1983년 기간의 합계출산율(2.857) 산출 과정을 보여 주고 있다. <표 4.2>를 간략히 살펴보면 100,000명을 기준으로 하는 통계청 생명표(사망)와 달리 첫째 출산($t = 1$)의 기수(ladix; $l_{x=15, t=1}$)가 1명임을 살펴볼 수 있다. 이는 기본적으로 합계출산율이 가상 코호트(synthetic cohort)에 속한 여성 1명이 가임 생애 기간 동안 주어진 출산율 스케줄을 적용받게 될 경우 낳게 되는 자녀 수라는 점을 반영하고 있다. 본 연구에서 작성하는 생명표의 경우 출산 리스크가 최초 15세부터 시작되는 관계로 첫째 출산 리스크에 적용되는 기수(radix)가 1인 반면 둘째 출산부터 최종 순위

출산의 15세 기준 기수($l_{x=15,t=2} \sim l_{x=15,t=7}$)는 모두 0(zero)으로 고정되어 있다. <표 4.2>에서 $l_{x,1}$ 및 $d_{x,1}$ 지표는 통상적인 사망력 생명표 작성 방식과 동일하게 산출될 수 있다. 둘째 이상 출산의 16세 이상($t \geq 2; x \geq 16$) $l_{x,t}$ 는 직전 연령($x-1$)의 $l_{x-1,t}$ 에서 $d_{x-1,t}$ 를 차감한 후 직전 출생 순위($t-1$) 및 직전 연령($x-1$)의 $d_{x-1,t-1}$ 를 더하는 방식으로 산출된다. <표 4.2>에서 출생 순위별 합계출산율(Order Specific Total Fertility Rates; TFR_t)은 출생 순위별로 d_x 열(column)을 합산함으로써, 그리고 합계출산율은 출생 순위별 합계출산율을 합산함으로써 도출 가능하다($TFR = \sum TFR_t$).

<표 4.3>에서 살펴볼 수 있듯이, 본 연구의 분석 대상 표본에 기간별-연령별 혼인 상태(유배우) 평균값을 적용하여 산출한 1965~1983년, 1984~1997년, 1998~2014년 세 기간의 합계출산율(TFR)은 각각 2.857, 1.666, 1.394 수준으로 나타나고 있다. 비록 본 연구에서 사용한 분석 대상 표본(1941~1996년 출생자)이 각 기간의 전체 가입 여성(15~49세) 레코드를 모두 포괄하는 것은 아님에도 불구하고, 통계청이 발표하는 1998~2014년의 합계출산율 수준을 고려할 때 출산율이 다소 과대 추정되고 있음을 시사하고 있다.⁶⁾ 본 연구에서는 연령별 출산 확률(q_x) 산출 과정에서 혼인 상태를 기간별-연령별 평균값으로 통제하는 방식을 취하고 있는데, 여성가족패널이 기혼자 그리고 기혼자 중 출산 경험이 있는 개인들을 과대 표집함으로써 기간별-연령별 평균 혼인율이 과대 추정되는 측면이 있다. 앞의 분석 방법 부분에서 언급했듯이 본 연구에서 구성한 혼인 상태 변수 또한 자료 제공상의 한계로 인해 유배우 상태가 과대 평가될 개연성이 있다. 물론 본 연구에서 산출한 합계출산율이 모집단이 아닌 표본조사에 기초하고 있다는 점에서 표집상의 편이 문제, 표본 추출 과정에서 일반가구만이 추출 대상에 포함된다는 점(집단 거주 시설 등 제외), 그리고 연령별 출산율의 모형화 과정에서 함수형 형태의 제약을 부과하고 있다는 점 또한 일정 정도 영향을 미칠 수 있음을 생각해 볼 수 있다. 이러한 점들을 종합적으로 고려할 때 본 연구에서 산출하는 출산율은 개별 값의 구체적 수준보다는 전반적인 패턴 그리고 집단별 격차에 초점을 맞추어 해석될 필요가 있다.⁷⁾

6) 일반적으로 통계청의 연령별 출산율(ASFR)에 기초하여 산출된 합계출산율과 본 연구처럼 출산 진도비(PPR) 생명표에 기초하여 산출한 합계출산율은 동일하지 않다. 이는 분석 방법 부분에서 이미 언급했듯이, 출산율(fertility rate) 산출 과정에서 이들 두 접근의 분모 값이 상이한 것과 관련이 있다.

7) 일반 사회조사 자료나 패널조사 자료에 기초하여 합계출산율을 산출한 국내 연구를 찾기는 쉽지 않다. 예외적으로 노동패널 자료를 분석한 송헌재(2012)의 연구는 1997~2007년 기간의 출생아 수 및 연령별 여성(母) 수에 기초하여 연령별 출산율을 구한 후 11년간의 평균값을 계산하고 있는데 1997~2007년 평균 합계출산율은 1.36으로 나타나고 있다.

〈표 4.1〉 이산형 생존모형(clog-log model) 추정 결과

구분	기본 모형								
	1965~1983년			1984~1997년			1998~2014년		
	첫째	둘째	셋째+	첫째	둘째	셋째+	첫째	둘째	셋째+
상수항	-5.4523***	-2.8442***	-3.775***	-6.653***	-3.753***	-4.414***	-6.7747***	-4.9659***	-6.2581***
연령	.2740***	.3171***	.3412***	.2252***	.3242***	.2155***	.2184***	.3795***	.3182***
연령제곱	-.0149***	-.0153***	-.0173***	-.0113***	-.0151***	-.0122***	-.0099***	-.0165***	-.0140***
결혼(유배우)	3.5760***	.4953***	.6318***	4.8731***	.8982***	.4445	4.7877***	1.5355***	1.2912**
N	36095	7336	17844	47938	12054	45680	37243	18506	61369
AIC	12476.71	9074.33	12730.27	11830.71	10706.22	5433.20	10699.31	11483.64	5906.71
BIC	12510.68	9101.93	12761.43	11865.82	10735.81	5468.12	10733.41	11514.95	5942.81

구분	최종 모형								
	1965~1983년			1984~1997년			1998~2014년		
	첫째	둘째	셋째+	첫째	둘째	셋째+	첫째	둘째	셋째+
상수항	-4.7284***	-1.6267***	-1.940***	-4.397***	-2.433***	-4.028***	-5.9020***	-2.9127***	-1.5258
연령	.4751***	.1952***	.2323***	.2235***	.2037***	.2246***	.1242	.2649***	-1.509
연령제곱	-.0230***	-.0158***	-.0183***	-.0104***	-.0147***	-.0127***	-.0040	-.0172***	-.0021
교육	-.2111***	-.0762***	-.0574***	-.2786***	-.1098***	-.0448***	-.2451***	-.1687***	-.4108*
결혼(유배우)	3.4873***	-.2785	-1.027†	4.0138***	.8555***	.4860	7.0811***	1.5328***	1.3410**
연령×교육	.0061***	.0067***		.0115***	.0098***		.0243***	.0104***	.0411*
연령제곱×교육							-.0005*		-.0010*
연령×결혼	-.3101***	.0886*	.1366***	-.1555***			-.2088***		
연령제곱×결혼	.0109*								
교육×결혼	.1637***			.1688***					
N	36095	7336	17844	47938	12054	45680	37243	18506	61369
AIC	12317.19	9057.71	12633.54	11740.98	10680.72	5423.04	10585.50	11470.32	5905.97
BIC	12393.63	9106.01	12680.28	11811.20	10725.11	5466.69	10653.70	11517.27	5969.14

주: N=개인-기간(person-period) 자료; †p<.1 * p<.05 ** p<.01 *** p<.001.

〈표 4.2〉 이산형 생존모형에 기초한 출산을 산출 과정(1965~1983년 기본 모형 기준)

연령	출생 순위																				
	t = 1			t = 2			t = 3			t = 4			t = 5			t = 6			t = 7		
	$q_{x,1}$	$l_{x,1}$	$d_{x,1}$	$q_{x,2}$	$l_{x,2}$	$d_{x,2}$	$q_{x,3}$	$l_{x,3}$	$d_{x,3}$	$q_{x,4}$	$l_{x,4}$	$d_{x,4}$	$q_{x,5}$	$l_{x,5}$	$d_{x,5}$	$q_{x,6}$	$l_{x,6}$	$d_{x,6}$	$q_{x,7}$	$l_{x,7}$	$d_{x,7}$
15	.005	1.00	.005	.057	.000	.000	.023	.000	.000	.023	.000	.000	.023	.000	.000	.023	.000	.000	.023	.000	.000
16	.006	.995	.006	.077	.004	.001	.032	.000	.000	.032	.000	.000	.032	.000	.000	.032	.000	.000	.032	.000	.000
17	.009	.989	.009	.101	.010	.001	.043	.001	.000	.043	.000	.000	.043	.000	.000	.043	.000	.000	.043	.000	.000
18	.012	.981	.012	.129	.017	.003	.057	.002	.000	.057	.000	.000	.057	.000	.000	.057	.000	.000	.057	.000	.000
19	.019	.968	.018	.162	.026	.006	.073	.005	.001	.073	.000	.000	.073	.000	.000	.073	.000	.000	.073	.000	.000
20	.030	.950	.028	.198	.039	.010	.092	.010	.001	.092	.001	.000	.092	.000	.000	.092	.000	.000	.092	.000	.000
21	.051	.922	.047	.238	.056	.019	.114	.019	.003	.114	.002	.000	.114	.000	.000	.114	.000	.000	.114	.000	.000
22	.086	.875	.075	.278	.084	.034	.138	.035	.007	.138	.005	.001	.138	.001	.000	.138	.000	.000	.138	.000	.000
23	.139	.800	.111	.315	.126	.057	.160	.062	.014	.160	.011	.003	.160	.002	.000	.160	.000	.000	.160	.000	.000
24	.208	.689	.143	.346	.180	.087	.178	.104	.026	.178	.022	.006	.178	.004	.001	.178	.001	.000	.178	.000	.000
25	.262	.546	.143	.364	.236	.112	.188	.165	.041	.188	.042	.012	.188	.009	.003	.188	.002	.001	.188	.000	.000
26	.291	.403	.117	.369	.267	.120	.189	.235	.056	.189	.072	.019	.189	.018	.005	.189	.004	.001	.189	.001	.000
27	.301	.286	.086	.363	.264	.111	.182	.300	.065	.182	.109	.026	.182	.032	.008	.182	.008	.002	.182	.002	.001
28	.287	.200	.057	.345	.239	.092	.168	.346	.066	.168	.148	.030	.168	.049	.011	.168	.014	.003	.168	.003	.001
29	.261	.142	.037	.319	.204	.071	.150	.373	.061	.150	.184	.032	.150	.069	.013	.150	.021	.004	.150	.006	.001
30	.231	.105	.024	.288	.170	.052	.129	.382	.053	.129	.213	.031	.129	.088	.013	.129	.030	.005	.129	.009	.001
31	.197	.081	.016	.251	.142	.038	.108	.382	.043	.108	.235	.028	.108	.106	.013	.108	.039	.005	.108	.012	.002
32	.162	.065	.011	.213	.120	.027	.086	.376	.034	.086	.250	.023	.086	.121	.011	.086	.047	.005	.086	.015	.002
33	.126	.054	.007	.174	.104	.019	.067	.369	.025	.067	.261	.018	.067	.132	.009	.067	.054	.004	.067	.018	.001
34	.098	.048	.005	.138	.092	.013	.050	.363	.018	.050	.268	.014	.050	.141	.007	.050	.059	.003	.050	.021	.001
35	.073	.043	.003	.106	.084	.009	.036	.358	.013	.036	.272	.010	.036	.147	.005	.036	.063	.002	.036	.023	.001
36	.053	.040	.002	.079	.078	.006	.025	.354	.009	.025	.275	.007	.025	.152	.004	.025	.066	.002	.025	.024	.001
37	.037	.038	.001	.057	.074	.004	.017	.351	.006	.017	.277	.005	.017	.155	.003	.017	.069	.001	.017	.025	.000
38	.025	.036	.001	.039	.071	.003	.011	.349	.004	.011	.278	.003	.011	.157	.002	.011	.070	.001	.011	.026	.000
39	.016	.035	.001	.026	.069	.002	.007	.348	.002	.007	.279	.002	.007	.158	.001	.007	.071	.000	.007	.027	.000
40	.010	.035	.000	.017	.068	.001	.004	.348	.001	.004	.280	.001	.004	.159	.001	.004	.072	.000	.004	.027	.000
41	.008	.034	.000	.011	.067	.001	.002	.348	.001	.002	.280	.001	.002	.160	.000	.002	.072	.000	.002	.027	.000
42	.005	.034	.000	.007	.066	.000	.001	.348	.000	.001	.280	.000	.001	.160	.000	.001	.072	.000	.001	.027	.000
43	.003	.034	.000	.004	.066	.000	.001	.348	.000	.001	.280	.000	.001	.160	.000	.001	.072	.000	.001	.027	.000
44	.001	.034	.000	.002	.066	.000	.000	.348	.000	.000	.280	.000	.000	.160	.000	.000	.072	.000	.000	.027	.000
45	.001	.034	.000	.001	.066	.000	.000	.348	.000	.000	.280	.000	.000	.160	.000	.000	.072	.000	.000	.027	.000
46	.000	.034	.000	.001	.066	.000	.000	.348	.000	.000	.280	.000	.000	.160	.000	.000	.072	.000	.000	.027	.000
47	.000	.034	.000	.000	.066	.000	.000	.348	.000	.000	.280	.000	.000	.160	.000	.000	.072	.000	.000	.027	.000
48	.000	.034	.000	.000	.066	.000	.000	.348	.000	.000	.280	.000	.000	.160	.000	.000	.072	.000	.000	.027	.000
49	.000	.034	.000	.000	.066	.000	.000	.348	.000	.000	.280	.000	.000	.160	.000	.000	.072	.000	.000	.027	.000
TFR_t = $\sum d_x$.966			.900			.553			.272			.112			.040			.013

〈표 4.3〉 모형 추정 결과에 따른 기간별 및 교육 수준별 합계출산율

합계출산율	1965~1983년	1984~1997년	1998~2014년
기본 모형			
전체	2.857	1.666	1.394
최종 모형			
중졸(9년)	2.763	1.912	1.866
고졸(12년)	2.392	1.738	1.599
대졸(16년)	2.101	1.706	1.338
부가 모형			
중졸(9년)	2.749	1.666	1.419
고졸(12년)	2.583	1.725	1.421
대졸(16년)	2.434	1.827	1.440

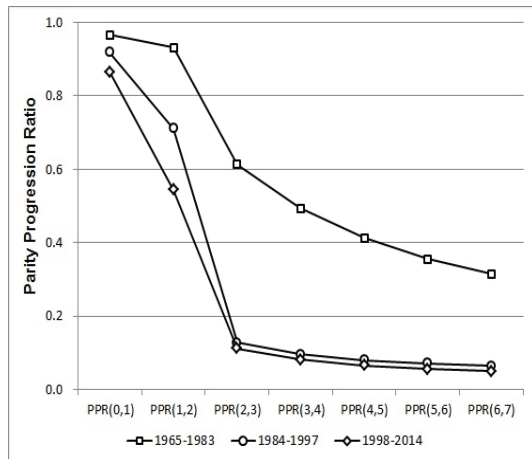
〈표 4.2〉와 같은 과정을 통해 산출된 〈표 4.3〉이 전반적인 출산력에 대한 요약 지표인 반면 〈그림 4.1〉은 추가적으로 출생 순위 정보를 포함하고 있는 기간별 출산진도비를 보여 주고 있다($PPR_{t,t+1} = TFR_{t+1}/TFR_t$). 출산진도비의 전반적인 패턴을 볼 때 1965~1983년 기간과 나머지 두 기간 간의 구분이 상대적으로 더욱 뚜렷함을 살펴볼 수 있다. 최근으로 올수록 무자녀 상태(childlessness)에서 첫째 출산으로 이어지는 비율($PPR_{0,1}$)이 감소하는 패턴(0.966 → 0.920 → 0.866)을 보임에도 불구하고 첫째 출산으로 이어질 개연성은 여전히 높은 모습을 살펴볼 수 있다(기간별-연령별 평균 유배우율 가정). 반면 첫째 출산에서 둘째 출산으로의 이행 그리고 둘째 출산에서 셋째 출산으로의 이행 단계에서는 상대적으로 큰 변화가 나타나고 있음을 살펴볼 수 있다. 1965~1983년 기간의 경우 첫째 출산을 한 여성 중 둘째 출산을 하는 여성의 비율이 93.2%였음에 비해 1984~1997년 기간에는 71.1%, 그리고 1998~2014년 기간에는 54.3%까지 떨어진 모습을 살펴볼 수 있다.

둘째 출산에서 셋째 출산으로의 이행 과정에서는 더욱 큰 변화가 관측되는데, 둘째 출산에서 셋째 출산으로의 이행 과정에서 관측되는 변화는 1965~1983년 기간과 나머지 두 기간 간의 차이가 상대적으로 더욱 뚜렷하다. 1965~1983년 기간, 즉 출산율이 대체출산율(RF) 이상으로 유지되었던 기간의 경우 둘째 출산에서 셋째 출산으로 이행하는 비율이 61.4%이었던 반면 그 이후 기간의 경우 10%대 초반 수준으로 크게 떨어진 모습을 살펴볼 수 있다. 1984~1997년과 1998~2014년 기간을 비교할 경우 셋째 이상 고순위 출산으로의 이행 비율에서는 큰 차이가 관측되지 않으며, 둘째 출산으로의 이행에서의 격차가 상대적으로 크게 나타남을 살펴볼 수 있다. 〈그림 4.2〉에서 볼 수 있듯이 1984~1997년 그리고 1998~2014년 기간의 경우 셋째 이상 고순위 출산율이 상당히 낮은 관계로 육안으로 식별하기가 쉽지 않다.

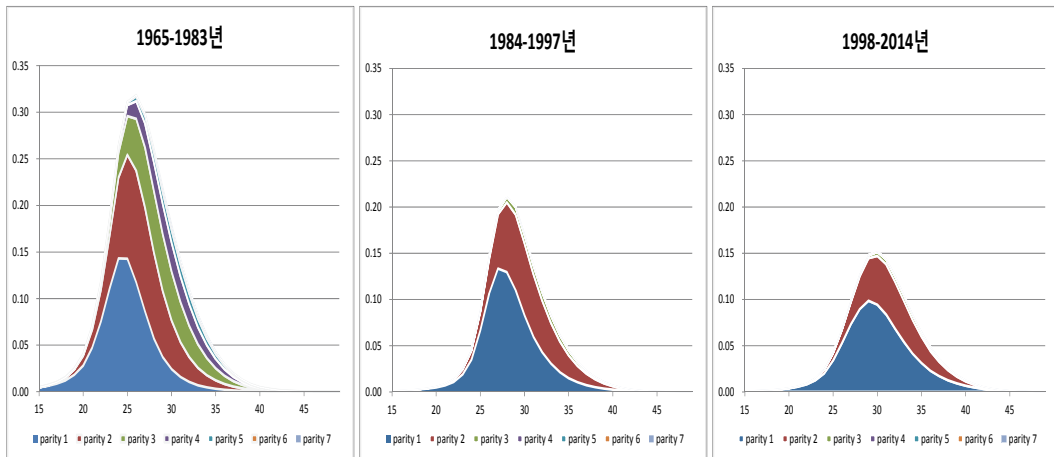
전반적으로 출산율이 대체출산율 아래로 하락한 1984년 이후 첫째 출산에서 둘째 출산으로의 이행, 그리고 둘째 출산에서 셋째 출산으로의 이행 비율이 더욱 가파르게 감소하는 모습을 살펴볼 수 있지만, 출산진도비 지표는 1965~1983년 기간 또한 우리나라 출산력에서 이미 상당히 큰 변화가 나타나고 있음을 시사하고 있다. 공식적인 통계가 존재

하지는 않지만 우리나라 합계출산율은 1950년대 중반부터 1960년대 중반까지 6명 이상으로 추정되고 있는데(권태환, 김두섭, 1990), <그림 4.1>은 1965~1983년 기간의 경우에도 이미 셋째 이상 출산으로의 이행 비율이 가파르게 감소하고 있음을 보여 주고 있다. 참고로, 1962년 경제개발5개년 계획의 일환으로 도입된 가족계획(family planning) 사업은 1960년대 후반부터 조직 및 재정 기반을 확대하여 1970년대 기간 동안 출산율을 두 명 수준까지 낮추는 데 초점을 맞춘 바 있다(인구정책50년사 편찬위원회, 2016).

출산진도비	1965~1983 년	1984~1997 년	1998~2014 년
PPR(0,1)	0.966	0.920	0.866
PPR(1,2)	0.932	0.711	0.543
PPR(2,3)	0.614	0.128	0.111
PPR(3,4)	0.493	0.095	0.081
PPR(4,5)	0.412	0.080	0.065
PPR(5,6)	0.356	0.071	0.056
PPR(6,7)	0.314	0.065	0.049



<그림 4.1> 기간별 출산진도비(PPR)의 변화 패턴(기본 모형 기준)



<그림 4.2> 기간별 출생 순위별 출산율의 분포(기본 모형 기준)

우리나라 출산력 변천에서의 전반적인 패턴을 살펴본 데 이어 집단별 분화 과정에 대해서 살펴보면, 본 연구에서는 개인들의 사회경제적 지위(SES)를 표상하는 핵심적인 지표 중의 하나인 교육 수준별 출산력 변천 과정에서의 차이를 살펴보기로 한다. 본 연구에서는 앞의 <표 4.1>의 최종 모형을 기준으로 교육 수준별로 출산력 변천 과정에서 어떠한 유사점과 차이점이 있는가를 살펴보기로 한다. <표 4.3>의 최종 모형은 <표 4.1>의

최종 모형에 기초하여 산출된 교육 수준별 합계출산율을 보여 주고 있다. 교육 수준 변수가 연속형이라는 점에서 본 연구에서는 특정 교육 이수 기간(년) 값을 <표 4.1>의 최종 모형에 투입하는 방식을 사용하였는데, 선정된 교육 이수 기간(년)은 9년(중졸), 12년(고졸), 16년(대졸)이다.⁸⁾ 비록 본 연구에서 중졸, 고졸, 대졸 집단이라는 용어를 사용하지만 이는 특정 교육 이수 기간(년)에 연계되어 있음에 유의할 필요가 있다. 최종 모형에 기초하여 교육 수준별 합계출산율을 산출하는 과정에서는 혼인 상태 변수를 기간별-교육 수준별(최종 교육 수준 기준)-연령별 평균값으로 고정하였다.

<표 4.3>의 최종 모형 기준 분석 결과를 보면 1965~1983년에서 1998~2014년 기간으로 진행될수록 합계출산율이 지속적으로 감소하지만, 모든 기간에 걸쳐 교육 수준이 낮을수록 출산율이 높은 패턴이 일관되게 관측됨을 살펴볼 수 있다. 여기에서 주목할 필요가 있는 사항은 <표 4.3>의 최종 모형에서 제시된 합계출산율을 도출하기 위해 기간별-교육 수준별-연령별 혼인 상태(유배우율) 평균값을 적용하고 있다는 점이다. <표 4.3>에는 최종 모형과 함께 부가 모형이 제시되고 있는데, 부가 모형에서는 혼인 상태 변수를 통제하는 과정에서 기간별-교육 수준별-연령별 혼인 상태(유배우율) 평균값 대신 기간별-연령별 혼인 상태(유배우율) 평균값으로 고정하는 방식을 사용하였다. 교육 수준별 혼인율 차이를 반영하지 않고 단순히 기간별-연령별 혼인율을 반영할 경우 교육 수준별 합계출산율에서 상이한 패턴이 출현함을 살펴볼 수 있다. 1965~1983년 기간의 경우 최종 모형과 마찬가지로 교육 수준이 높을수록 출산율이 낮은 패턴이 지속적으로 관측된다(교육 수준별 격차는 감소). 그러나 1984~1997년 그리고 1998~2014년 기간의 경우 교육 수준이 높을수록 오히려 출산율이 높은 상반된 패턴이 나타남을 살펴볼 수 있다. <표 4.3>의 최종 모형에서 1984~1997년 그리고 1998~2014년 기간에 걸쳐 교육 수준이 낮을수록 출산율이 높은 패턴은 교육 수준이 낮을수록 유배우율이 높은 상황을 반영하고 있는 것이다. 반면 부가 모형처럼 교육 수준을 가로질러 상이한 유배우율을 통제할 경우 1965~1983년 기간과 달리 합계출산율이 대체출산율 아래로 하락한 1984년 이후의 기간에서는 교육 수준과 출산율 간의 정적(+) 관계가 관측되고 있는 것이다.

<표 4.3>에서 살펴볼 수 있는 또 다른 점은 출산율 분석 대상 기간별 출산율 감소 폭이 교육 수준별로 차별적인 모습을 보인다는 것이다. 우선 1965~1983년 기간 대비 1984~1997년 기간의 출산율 감소 폭은 교육 수준이 낮을수록 더욱 큰 모습을 보인다. 1965~1983년 기간에 비해 1984~1997년 기간의 경우 고졸(12년)과 대졸(16년) 간 합계출산율 격차는 상당히 미미한 수준이다. 반면 1984~1997년 기간 대비 1998~2014년 기간의 경우 상반된 패턴이 관측되는데 교육 수준이 높은 집단의 출산율 감소 폭이 더욱 큰 모습을 살펴볼 수 있다. 교육 수준별 출산율 격차가 기간별로 상이하게 나타나고 있지만, 본 분석 결과에 기초하여 그 원인을 정확히 이해하기는 어렵다. 특히, 과거에 비해 IMF 경제 위기 이후 교육 수준별 혼인 및 출산에 있어서 어떠한 변화가 있었는가에 대해서는 추가적인 분석이 필요한 것으로 보인다.⁹⁾

8) 교육 수준을 연속 변수 대신 명목 변수로 구성하는 방법을 검토해 볼 수 있다. 그러나 우리나라의 경우 시기별 교육 수준 변동이 크다는 점에서 더미 변수 접근이 쉽지 않은 측면이 있다. 예컨대, 과거 대졸 집단의 비율이 상당히 낮은 반면, 최근 들어 중졸 이하 집단의 비율이 극히 낮은 모습을 보인다.

9) 교육 수준별 차별 출산력이 기간별로 어떻게 전개되어 왔는가를 정확히 살펴보기 위해서는 보다

마지막으로, <그림 4.3>은 최종 모형에 기초하여 분석 대상 기간별로 교육 수준별 출산 진도비가 어떠한 방식으로 전개되어 왔는가를 보여 주고 있다. 교육 수준을 가로 질러 전반적인 패턴은 앞에서 살펴본 <그림 4.1>과 유사하다. 교육 수준별 차이에 맞추어 분석 결과를 살펴보면 전반적으로 기간별 출산력 변천 과정에서 교육 수준이 높은 집단이 선도자로서의 역할을 하고 있음을 살펴볼 수 있다. 1965~1983년 기간의 경우 첫째 혹은 둘째 출산으로의 이행 비율이 여전히 높는데, 대졸 집단의 경우도 첫째 출산을 한 여성 중 88.9%가 둘째 출산으로 이행하는 모습을 볼 수 있다. 반면 둘째 출산에서 셋째 출산으로의 이행 과정에서는 교육 수준별로 커다란 격차가 존재함을 살펴볼 수 있다. 중졸 집단의 경우 둘째 출산을 한 여성의 56.2%가 셋째 출산으로 이행한 반면 대졸 집단의 경우 이 비율은 27.3%로 크게 떨어진 모습을 살펴볼 수 있다.

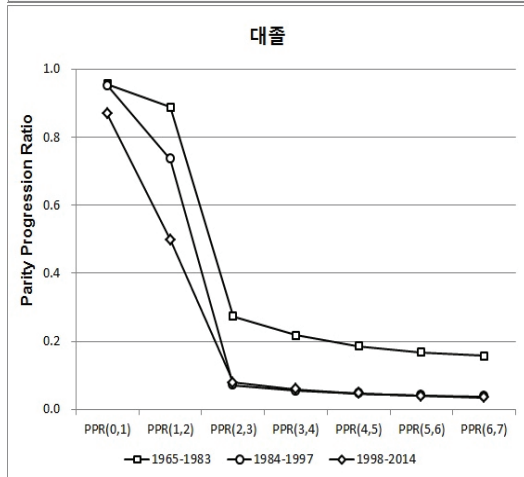
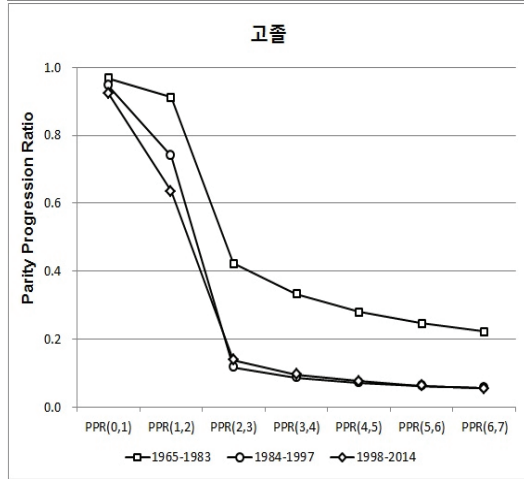
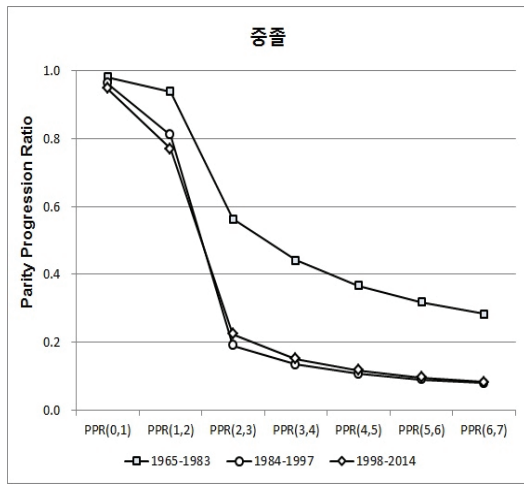
1984~1997년 기간의 경우 교육 수준을 가로질러 둘째 출산에서 셋째 출산으로 이행하는 비율이 크게 낮아진 모습을 볼 수 있다. 중졸 집단의 경우에도 둘째 출산을 한 개인들 중 20% 미만이 셋째 출산으로 이행하며, 대졸 집단의 경우 해당 비율은 10% 미만이다. 추가적으로, 1984~1997년 기간의 경우 첫째 출산에서 둘째 출산으로 이어지는 비율이 본격적으로 감소하기 시작한 시기라고 할 수 있는데, 중졸 집단의 81%, 대졸 집단의 73.5%가 첫째 출산에서 둘째 출산으로 이행하는 모습을 보이고 있다. 마지막으로 1998~2014년 기간의 경우 1984~1997년 기간의 추세가 더욱 심화되는 모습을 보이는데, 대졸 집단의 경우 첫째 출산 후 둘째 출산으로 이어지는 비율이 50% 아래로 떨어지는 동시에 무자녀 상태에서 첫째 출산으로 이행하는 비율 또한 87% 수준까지 감소한 모습을 살펴볼 수 있다. 교육 수준별 출산력 변천 분석 결과는 최근까지도 첫째 출산으로의 이행 비율이 상대적으로 높게 형성되어 있지만, 최근까지의 패턴이 향후에도 지속될 경우 첫째에서 둘째 출산으로 이어지는 비율이 크게 감소하는 동시에 생애 동안 자녀를 출산하지 않는 여성들의 비율 또한 작지 않을 수 있음을 시사하고 있다.

세부적인 기간의 구분이 필요한 것으로 보인다. 본 연구에서는 1965년에서 2014년까지의 기간을 살펴보고 있는 관계로 베이비붐 세대가 태어난 기간의 정보가 누락되어 있다. 자료가 활용 가능할 경우 베이비붐 세대가 태어난 시기에 비해 1965~1983년 기간의 출산율 감소가 교육 수준별로 어떻게 진행되었는가에 대해서도 살펴볼 필요가 있다. 또한 본 연구의 경우 1965~1983년 기간을 하나의 시기로 통합하는 방식을 사용하고 있는데, 출산력 변천에서 1975년 전후를 구분하는 기존 논의(전광희, 2002)를 고려할 때 보다 세부적인 구분 또한 검토 가능한 것으로 보인다.

중졸(9년)	1965~1983 년	1984~1997 년	1998~2014 년
PPR(0,1)	0.980	0.962	0.946
PPR(1,2)	0.938	0.810	0.770
PPR(2,3)	0.562	0.191	0.224
PPR(3,4)	0.442	0.135	0.153
PPR(4,5)	0.367	0.107	0.118
PPR(5,6)	0.318	0.091	0.097
PPR(6,7)	0.284	0.080	0.083

고졸(12년)	1965~1983 년	1984~1997 년	1998~2014 년
PPR(0,1)	0.968	0.947	0.924
PPR(1,2)	0.912	0.741	0.635
PPR(2,3)	0.422	0.117	0.138
PPR(3,4)	0.332	0.087	0.096
PPR(4,5)	0.279	0.072	0.075
PPR(5,6)	0.245	0.062	0.063
PPR(6,7)	0.222	0.056	0.054

대졸(16년)	1965~1983 년	1984~1997 년	1998~2014 년
PPR(0,1)	0.957	0.952	0.869
PPR(1,2)	0.889	0.735	0.498
PPR(2,3)	0.273	0.072	0.080
PPR(3,4)	0.218	0.055	0.059
PPR(4,5)	0.186	0.047	0.048
PPR(5,6)	0.168	0.042	0.040
PPR(6,7)	0.158	0.038	0.035



〈그림 4.3〉 기간별-교육 수준별 출산진도비(PPR)의 변화 패턴(최종 모형 기준)

5. 종합 및 결론

초저출산 현상이 장기화됨에 따라 출산에 대한 학술적 및 정책적 관심이 크게 증가하고 있다. 출산과 관련하여 최근까지 이루어진 기존 국내 연구들은 출산 계획 혹은 출산 행위와 유의하게 연관된 영향 요인들을 확인하는 작업에 초점을 맞추는 경향이 있었다. 선행 연구들과 달리 본 연구는 생존분석을 활용하여 합계출산율(TFR)과 같은 인구학적으로 보다 의미가 있는 지표를 산출하는 접근을 취하고 있다. 또한 연령별 출산율(ASFR) 대신 출산진도비(PPR)에 초점을 맞추므로써, 기존 연구들이 강조한 출산의 속도(tempo of fertility) 대신, 출산의 양(quantum of fertility)에 초점을 맞추었으며, 이에 기초하여 가임 생애에 걸친 가족(자녀) 규모의 분포와 관련된 정보를 제공하고자 하였다.

분석 결과는 상대적으로 최근까지도 첫째 출산으로 이어지는 비율이 상당히 높은 수준으로 유지되고 있지만, 1998~2014년 기간의 제로 출산진도비(zero parity progression ratio; $P_{0,1}$)가 0.866 수준으로 가임 생애 동안 자녀를 한 명도 출산하지 않는 여성의 비율이 작지 않은 수준까지 증가하고 있음을 시사하고 있다. 다른 한편으로, 출산진도비 분석 결과는 1965~1983년 기간에 비해 출산율이 대체출산율 아래로 떨어진 1984~1997년 기간 이후 출산력에서 더욱 큰 변동이 있었음을 보여 주는데, 특히 첫째 출산에서 둘째 출산 그리고 둘째 출산에서 셋째 출산으로의 이행 과정에서 상대적으로 큰 변화가 나타났음을 보여 주고 있다. 교육 수준별 출산력 분화와 관련하여 분석 결과는 관측 기간 전체에 걸쳐 교육 수준과 출산율 간에 부적(-) 연관성이 일관되게 나타남을 보여 주었다. 그러나 추가적인 분석 결과는 교육 수준과 출산율 간의 이러한 부적(-) 연관성은 교육 수준이 높을수록 혼인율(유배우율)이 낮은 점을 통해 매개되고 있음을 시사하고 있다.

마지막으로 교육 수준별 출산진도비 분석 결과는 과거 한국 사회의 출산율 감소 과정에서 교육 수준이 높은 집단이 전반적인 추세를 선도하는 반면 교육 수준이 낮은 집단이 일정한 시간적 간극을 갖고 뒤따르는 모습을 보여 주고 있다. 물론 한국의 출산력 변천 과정에서 나타난 선도자-추종자(leader-follower) 패턴이 최종적으로 교육 수준별 출산력 차이가 없어지게 될 것임을 의미하는 것은 아니다. 본 연구에서 대졸(교육 수준 16년) 여성의 경우 첫 출산으로 이어지는 비율이 90% 아래로 떨어지는 모습이 관측되었는데, 향후 제로 출산진도비($P_{0,1}$)에서의 교육 수준별 패턴이 어떠한 방향으로 전개될 것인가에 대해서는 추가적인 시간이 필요한 것으로 보인다.¹⁰⁾ 특히, 현재까지도 대졸 이상 고등교육 수료자의 비율이 절대적으로 높은 상황이라는 점에서 이들 집단의 제로 출산진도비($P_{0,1}$) 패턴에 대해 보다 세부적인 관심이 필요하다고 할 수 있다. 다른 한편 저출산 대응 정책의 영역에서는 현재까지도 결혼을 하면 출산을 한다는 측면에서 다자녀를 중심으로 지원 정책을 추진하는 경향이 강한 것으로 판단된다. 그러나 최근 대졸 집단이 보이는 제로 출산진도비($P_{0,1}$) 패턴은 향후 한국 사회에서 결혼이 반드시 출산으로 이어지지 않

10) 출산력 변천 과정에서 나타나는 교육 수준별 차별 출산력의 패턴과 관련하여 Bongaarts(2003)는 선도자-추종자(leader-follower) 가설과 항구적 격차(permanent difference) 가설을 제시한다. 선도자-추종자 가설의 경우 교육 수준별 격차가 최종적으로 해소됨을 상정함에 비해 항구적 격차 가설은 교육 수준별 격차가 유지될 것을 전망한다. 그의 경험적 자료(DHS) 분석 결과는 출산력 변천을 완료하는 시점에서 교육 수준별 격차가 감소하기는 하지만 교육 수준별 격차가 지속됨을 보여 주고 있다.

을 수 있다는 점을 시사함으로써 첫째 자녀 출산에 대한 정책적 배려 또한 매우 중요할 수 있음을 시사하고 있다.

전반적으로 본 연구는 현재 통계청이 생산하는 자료를 표본조사에 기초하여 산출하는 한편 출산진도비처럼 통계청이 공표하지 않는 자료 및 사회 집단별 출산력 차이를 비교하는 데 있어서 생존분석이 효과적으로 활용될 수 있음을 보여 주고 있다는 점에서 의의가 있다. 본 연구에서는 소수의 변수들만을 고려한 모형을 구축하였지만, 연구자의 관심에 따라 보다 복잡한 모형 설계 또한 충분히 검토 가능할 것이다. 다른 한편으로 본 연구의 분석 결과는 통계청 혹은 센서스에 기초한 분석 결과와 유사한 수준의 출산율 자료를 생산하기 위해서는 출산 관련 체계적인 데이터베이스 구축이 중요함을 시사하고 있다. 앞서도 언급했듯이, 본 연구에서 사용한 여성가족패널의 경우 생애에 걸쳐 출산 이력(birth history) 정보를 제공하는 중요한 장점에도 불구하고 기혼 여성, 특히 출산 경험이 많은 기혼 여성이 과대 표집되고 있다는 점에서 통계청 자료에 상응하는 수준의 값을 산출하기 어려운 측면이 있다. 물론 본 연구에서 검토한 기간(1965~2014년)의 연령별 혹은 교육 수준별-연령별 혼인율 자료가 외부적으로 활용 가능하다면 해당 값들을 사용하는 방식을 고려할 수 있을 것이다. 다른 한편으로 본 연구가 기본적으로 표본조사에 기초하고 있음을 고려한다면 생존분석에 기초한 출산력 분석은 전반적인 출산율 수준보다는 출산에 영향을 미치는 여타의 사회경제적 특성을 통제한 상태에서 분석 대상 변수와 출산 간의 관계 분석 측면에서 보다 유용한 정보를 제공할 수 있을 것이다. 다른 한편으로 본 연구가 기초하고 있는 여성가족패널은 출산력에 대해서만 생애 이력 정보를 제공할 뿐 출산과 밀접히 연관된 혼인력에 대한 이력 정보는 수집하지 않고 있다. 물론 이러한 문제는 여성가족패널에만 한정된 문제는 아니다. 출산에서 혼인 상태가 갖는 의미가 매우 큰 우리나라의 상황을 고려할 때 생애에 걸친 정확한 혼인 상태 변동 정보를 제공하는 작업 또한 매우 중요하다고 할 수 있다. 마지막으로 본 연구가 표본조사(서베이) 자료에 기초하고 있다는 점에서 출산력 관련 지표들의 표준오차를 제공하는 작업 또한 후속 연구에서 검토될 필요가 있는 사항이라고 할 수 있다.

(2018년 3월 16일 접수, 2018년 4월 17일 수정, 2018년 5월 11일 채택)

참고문헌

- 권태환, 김두섭 (1990). <인구의 이해>, 서울대학교출판부, 서울.
- 김두섭 (2007). IMF 외환위기와 사회경제적 차별출산력의 변화, <한국인구학>, 30(1), 67-95.
- 김정석 (2007). 기혼여성의 출산아수별 추가출산계획, <한국인구학>, 30(2), 97-116.
- 김현식 (2017). 자산과 소득에 따른 차별출산력 연구, <한국인구학>, 40(3), 51-78.
- 대한민국정부 (2015). <제3차 저출산·고령사회 기본계획(2016~2020)>.
- 류기철, 박영화 (2009). 한국여성의 출산율 변화와 출산간격 영향요인, <한국인구학>, 32(1), 1-23.
- 민현주 (2007). 엄마의 취업과 자녀터울에 관한 동태적 분석, <한국사회학>, 41(3), 106-12.
- 민현주, 김은지 (2011). 출산순위별 출산결정요인분석, <한국사회학>, 45(2), 198-222.
- 박수미 (2008). 둘째 출산 계획의 결정요인과 가족내 성 형평성, <한국사회학>, 31(1), 55-73.
- 박수미, 이택면, 김승연, 우원규, 강석훈 (2008). <2008년 여성가족패널조사: 사업보고서 및 제1차 기초분석보고서>, 한국여성정책연구원.
- 송헌재 (2012). 노동패널 자료를 이용한 우리나라 가구의 출산결정요인 분석, <응용경제>, 14(3), 51-78.
- 우해봉, 장인수 (2017). 출산에 대한 개인 특성과 부부 특성의 영향, <한국인구학>, 40(3), 79-105.
- 은기수 (2001). 결혼연령 및 결혼코호트와 첫 출산간격의 관계: 최근의 낮은 출산력 수준에 미치는 함의를 중심으로, <한국사회학>, 35(6), 105-139.
- 이성용 (2006). 경제위기와 저출산, <한국인구학>, 29(3), 111-137.
- 이성용 (2009). 출산순위별 출산증가 요인 분석, <한국인구학>, 32(1), 51-70.
- 이흥탁 (1994). <인구학: 이론과 실제>, 법문사, 서울.
- 인구정책50년사 편찬위원회 (2016). <한국 인구정책 50년: 출산억제에서 출산장려로>, 보건복지부·한국보건사회연구원, 세종.
- 전광희 (2002). 출산력, 김두섭, 박상태, 은기수 편, <한국의 인구 1>, 통계청, 대전.
- 정은희, 최유석 (2013). 기혼여성의 둘째자녀 출산계획 및 출산과 관련된 요인, <보건사회연구>, 33(1), 5-34.
- 정혜은, 진미정 (2008). 취업여부에 따른 기혼여성의 둘째자녀 출산의도, <한국인구학>, 31(1), 147-164.
- 최은영, 박영실 (2009). <사회경제적 특성과 지역별 차별 출산력 분석>, 통계개발원, 대전.
- 통계청 (2018). 2017년 인구동향조사 출생·사망통계 잠정 결과(보도자료).
- Bongaarts, J. (2003). Completing the fertility transition in the developing world: The role of educational differences and fertility preferences, *Population Studies*, 57(3), pp.321-336.

- Bongaarts, J. and Feeney, G. (2006). The quantum and tempo of life-cycle events, *Vienna Yearbook of Population Research*, 2006, pp.115-151
- Cox, D.R. (1972). Regression models and life-tables, *Journal of the Royal Statistical Society*, 34(2), 187-220.
- De Santis, G. Drefahl, S. and Vignoli, D. (2014). A period total fertility rate with covariates for short-panel data, *Population*, 69(3), 463-476.
- Feeney, G. and Yu, J. (1987). Period parity progression measures of fertility in China, *Population Studies*, 41, 77-102.
- Gray, E. Evans, A. Anderson, J. and Kippen, R. (2010). Using split-population models to examine predictors of the probability and timing of parity progression, *European Journal of Population*, 26(3), 275-295.
- Hinde, A. (1998). *Demographic Methods*, ARNOLD, London.
- Hoem, J.M. and Muresan, C. (2011a). The total marital fertility rate and its extensions, *European Journal of Population*, 27(4), 295-312.
- Hoem, J.M. and Muresan, C. (2011b). An extension of the conventional TFR, *European Journal of Population*, 27(4), 389-402.
- Ni Bhrolchain, M. (1987). Period parity progression ratios and birth intervals in England and Wales, 1941-1971: A synthetic life table analysis, *Population Studies*, 41(1), 103-125.
- Preston, S.H. Heuveline, P. and Guillot, M. (2001). *Demography: Measuring and Modeling Population Processes*, Blackwell Publishing, MA.
- Pullum, T.W. (2004). Natality: Measures based on censuses and surveys, in J.S. Siegel and D.A. Swanson (eds.), *The Methods and Materials of Demography*, Elsevier Academic Press, CA.
- Retherford, R. Ogawa, N. Matsukura, R. and Eini-Zinab H. (2010). Multivariate analysis of parity progression-based measures of the total fertility rate and its components, *Demography*, 47(1), 97-124.
- Ryder, N.B. (1986). Observations on the history of cohort fertility in the United States, *Population and Development Review*, 12(4), 617-643.
- Schoumaker, B. (2013). A Stata module for computing fertility rates and TFRs from birth histories: tfrs, *Demographic Research*, 28(38), 1093-1144.
- Singer, J.D. and Willett, J.B. (2003). *Applied Longitudinal Data Analysis: Modeling Changes and Event Occurrence*, Oxford University Press, New York.
- Van Hook, J. and Altman, C.E. (2013). Using discrete-time event history fertility models to simulate total fertility rates and other fertility measures, *Population and Policy Review*, 32(4), 585-610.

Fertility Transition in South Korea: Application of Survival Analysis

Haebong Woo¹⁾ · Insu Chang²⁾

Abstract

Due to persistent lowest-low fertility since 2000s, increasing attention has been paid to fertility-related academic and policy studies in South Korea. Past empirical studies addressing low fertility in South Korea have mainly focused on determinants of birth plans or birth behavior. Compared to prior research, this study uses survival analysis to produce demographically more meaningful information about lifetime childbearing such as total fertility rates (TFR). Particularly, instead of age-specific fertility rates (ASFRs), this study analyzes the fertility transition in terms of parity progression ratios (PPRs), which provide useful information on the distribution of family sizes over the reproductive life-course. The data show that the fertility transition became accelerated since the early 1980s when the total fertility rate dropped below the replacement-level. However, in case of third and higher-order birth, the transition already started from the middle of the 1960s when the government-sponsored family planning programs were implemented in a more organized way in terms of organization and finance. Finally, our analysis on educational differentials indicate that women with higher education have played a leading role in the rapid fertility transition over the past 50 years (1965–2014).

Key words : fertility transition, survival analysis, marital status, education

1) (Corresponding author) Research Fellow, Korea Institute for Health & Social Affairs, Building D, 370 Sicheond-daero, Sejong City. E-mail: haebongwoo@kihasa.re.kr

2) Senior Researcher, Korea Institute for Health & Social Affairs, Building D, 370 Sicheond-daero, Sejong City. E-mail: sescis@kihasa.re.kr