

## 저출생 문제해결을 위한 한자녀 기혼여성의 후속 출산의향 예측: 머신러닝 방법의 적용\*

Predicting the Subsequent Childbirth Intention of Married Women with One Child to Solve the Low Birth Rate Problem in Korea: Application of a Machine Learning Method

전효정<sup>1</sup>

Hyo Jeong Jeon<sup>1</sup>

### ABSTRACT

**Objective:** The purpose of this study is to develop a machine learning model to predict the subsequent childbirth intention of married women with one child, aiming to address the low birth rate problem in Korea. This will be achieved by utilizing data from the 2021 Family and Childbirth Survey conducted by the Korea Institute for Health and Social Affairs.

**Methods:** A prediction model was developed using the Random Forest algorithm to predict the subsequent childbirth intention of married women with one child. This algorithm was chosen for its advantages in prediction and generalization, and its performance was evaluated.

**Results:** The significance of variables influencing the Random Forest prediction model was confirmed. With the exception of the presence or absence of leave before and after childbirth, most variables contributed to predicting the intention to have subsequent childbirth. Notably, variables such as the mother's age, number of children planned at the time of marriage, average monthly household income, spouse's share of childcare burden, mother's weekday housework hours, and presence or absence of spouse's maternity leave emerged as relatively important predictors of subsequent childbirth intention.

\* 본 논문은 2024년 동아대학교 학술연구비 지원에 의하여 연구되었음.

<sup>1</sup> 제1저자(교신저자)

동아대학교 아동학과 교수  
(e-mail: hjeon@dau.ac.kr)

**key words** subsequent childbirth intention, married women with one child, low birth rate problem in Korea, machine learning method

## I. 서론

우리나라는 세계에서 출생률이 가장 급격하게 감소하는 국가로, 인구 감소를 넘어 국가 소멸을 걱정해야 할 정도로 전 세계에서 출생률 감소가 가장 심각한 수준이다. 이러한 저출생 현상은 더욱 악화될 것이라는 정부 추계가 나와 더욱 비관적이다. 통계청의 2022~2072년 장래 인구 추계 발표에 의하면, 2024년 합계출산율은 0.7명 선이 깨어져 0.68명을 기록하고, 2025년에 0.65명으로 바닥을 찍은 뒤에 2026년 0.68명, 2027년 0.71명으로 소폭 반등할 것으로 예측하였다(통계청,

2023). 저출생과 인구 감소에 대응하기 위하여 정부는 2005년 제1차 저출산고령사회 기본법을 시작으로 매 5년마다 구체적인 기본계획을 추진하면서 출산을 강력하게 촉진하는 인구정책을 마련하고, 이를 지원하기 위하여 280조의 막대한 재정을 투입해 왔으나, 합계출산율은 매년 최저치를 경신하고 있는 실정이다. 이러한 결과는 출생률을 높이기 위한 노력의 방향과 방법이 유효하지 않았기 때문일 것이다. 출생률을 높일 수 있는 타겟 대상을 정하고 보다 정교한 예측을 통해 유효한 정책을 제안할 필요가 있다.

아이를 갖더라도 한 명에 그치는 추세도 두드러져서 2022년 출생사망 통계에 의하면 출생아 24만 9000명 가운데 첫째 아이는 62.7%, 둘째 비율은 30.5%(7만 6000명), 셋째 이상은 6.8%(1만 6800명)에 그쳐, 둘째와 셋째는 각각 16.8%, 20.7% 감소했다(통계청, 2023). 한편 부모의 양육스트레스, 그리고 이와 관련된 아동학대 사건이 끊임없이 보도되고 있다. 따라서 이제는 출산 후 양육자를 위한 맞춤형 지원체계가 필요한 시대이며, 정확한 예측을 바탕으로 대상에게 유효한 지원 정책을 마련해야 할 것이다. 이제라도 출생률을 높일 수 있는 가장 효과적인 대상을 정하고 예측요인을 정확하게 파악하고, 이에 따라 유효한 방법을 모색하는 체계적인 노력이 절실한 시점이다.

저출산의 원인으로 자녀양육 부담과 비용 부담이 손꼽히면서 각종 자녀양육과 교육지원정책 및 아동수당, 출산지원금 등 출산에 대한 직접적 현금지원과 보육시설과 돌봄 지원 등 다양한 정책들이 쏟아져 나왔으나 지속적인 출산을 하락을 막기에는 역부족이었다. 또한 젊은 세대의 만혼화와 출산연기, 출산포기가 초저출생율의 원인으로 꼽히는 점(강은영 등, 2010; 차승은, 2008a)과 우리 사회에서 부모됨의 가치관의 변화도 고려해야 할 것이다(배혜원, 2017b). 2022년 혼인·이혼통계를 보면, 혼인건수는 19만2천건으로 1년 전보다 0.4% 감소했다. 이는 통계 작성이 시작된 1970년 이후 가장 적은 수치다. 가장 많은 혼인이 이루어졌던 1996년에는 43만5천건의 혼인신고가 이루어지기도 했는데, 1997년(38만9천건)에 30만건대로 내려온 뒤, 2016년(28만2천건)에 20만건대, 2021년(19만3천건)에 10만건대로 내려앉았다. 한창 혼인건수가 많았을 때와 견주면 반토막이 난 셈이다. 인구 1천명당 혼인건수인 조혼인율은 1년 전보다 0.1건 줄어든 3.7건이었다. 이 또한 통계 작성이 시작된 뒤 역대 최저치다.

혼인 건수의 감소는 25~49살 연령의 인구가 계속 줄어들고 있는 인구의 구조적인 원인이 있고, 더불어 혼인에 대한 가치관 변화도 영향을 미친 것으로 보인다. 한국 사회의 결혼 연령은 25-29세에서 30-34세로 높아졌으며, 2022년 초혼연령은 남녀 모두 역대 최고치를 기록했다. 평균 초혼연령은 남자 33.7살, 여자 31.3살로 1년 전보다 각각 0.4살, 0.2살 올라갔다. 남자의 경우 35살 이하에서 전년 대비 혼인건수가 줄고 35살 이상에서 증가하는 경향을 나타냈다. 특히 20대 후반(-8.4%)에서 가장 많이 줄고, 40대 초반(10%)에서 가장 크게 늘었다. 여성은 20대 후반(-7.2%)에서 가장 많이 줄었다(통계청, 2023). 이러한 만혼화로 출산 연령 역시 30-34세로 늦춰지면서 출산율도 함께 떨어지고 있다는 지적이다(배혜원, 2017a). 즉, 결혼 연령이 상승함에 따라 첫출산 시기가 늦어지고 이에 따라 다자녀출산의 가능성도 낮아지면서 저출산 현상이 가속화한다는 것이다(이완, 채재은, 2017).

한국의 저출생 문제를 해결하기 위한 정책 및 지원 방법에 대한 연구들은 다양한 측면에서 이루어져 왔다. 이러한 연구들은 정부, 학계, 비정부 기관 등에서 다루어지며, 다음과 같은 주요 내

용을 다루고 있다. 첫째 출산 장려 정책 평가와 관련하여 많은 연구가 한국 정부가 시행한 출산 장려 정책의 효과를 평가하였다(강경희, 진홍주, 2013; 송다영, 2013a; 이미옥, 명성준, 2015; 이철희, 2018; Lee, 2019). 이러한 연구들은 장려금, 육아휴직 지원, 양육비 지원 등의 제도가 출산율에 미치는 영향을 조사하고, 정책의 효율성과 효과적인 개선 방안을 모색하였다. 둘째 직장과 가족 생활의 균형과 관련하여 여성들이 직장에서의 경력과 가정생활을 균형 있게 유지할 수 있는 제도 및 지원 방안에 대한 연구도 많이 진행되었으며(최미란, 장경은, 2017; Choi & Kwon, 2018). 육아휴직, 유연한 근무 조건, 양립 지원 프로그램 등이 이에 해당된다. 셋째 교육과 인식 변화와 관련하여 저출생 문제를 해결하기 위해서는 출산과 관련된 인식과 가치관의 변화가 필요하다는 인식하에 교육 및 캠페인을 통한 인식 개선에 대한 연구도 이루어졌다(배혜원, 2017b; Jeong et al., 2022; Kwon & Chang, 2019). 넷째 경제적 인센티브의 측면에서 경제적 어려움이 출산 결정에 영향을 미칠 수 있기 때문에, 출산에 대한 경제적 인센티브를 제공하는 방안에 대한 연구도 진행되었다(Cho & Yoo, 2018; Lee, 2019). 출산 후 금전적 지원이나 세제 혜택 등이 여기에 해당된다. 다섯째 다문화 가정 및 저소득층 지원으로 다문화 가정이나 저소득층 가정에서의 출산지원에 대한 연구도 중요한 주제이다. 이들을 대상으로 한 정책 및 프로그램이 출산율 증가에 어떤 영향을 미치는지에 대한 연구가 진행되고 있다(Seo, 2019; Zhang et al., 2023). 그러나 선행연구들은 대부분 거시적인 영향이나 가치관과 같은 부분으로 정책의 주 대상 집단을 명확하게 규정하지 않아 구체적으로 유효한 정책으로 연결되기 어려워 실제 출산율을 증가시키는 정책적 해법을 제안하는데 한계가 있었다.

본 연구는 가장 효과적으로 출생률을 높일 수 있는 방법은 현재 한자녀를 둔 가정을 대상으로 후속 출산율을 높이는 것이 가장 즉각적이고 효율적으로 실질 출생률을 향상시킬 수 있는 길이라는 생각에서 시작되었다(백지희, 2020). 가족 안정성 이론(Family Stability Theory)에 의하면 가정 내 안정성이 출생률에 미치는 영향이 높다고 하겠다. 한자녀를 둔 가정은 이미 부모의 역할에 적응하고 있으며, 추가로 출생률을 높이는 데 더욱 열려있을 수 있으므로 이러한 가정을 대상으로 한 지원은 출생률을 향상시키는 데 효과적이라 하겠다. 또한 심리적 지원(Psychological Support) 관점에서 한자녀 가정은 추가 출산에 대한 심리적 부담이 비교적 적을 수 있다. 이미 부모가 되어 경험을 쌓았으며, 후속 출산에 대한 불안이나 걱정이 상대적으로 적을 수 있다. 따라서 이러한 가정을 지원함으로써 출산을 효율적으로 장려할 수 있다. 즉 한자녀 가정을 대상으로 후속 출산의향을 정확하게 예측하는 모델을 찾는다면 후속 출산율을 높일 수 있는 구체적인 방안을 제안할 수 있을 것이다.

출산계획과 관련된 개인적, 사회적 영향 요인들에 대한 선행연구들이 보고 되어왔다. 고용불안, 여성의 경제활동 참여 증가 등 경제구조의 변화, 급속한 고령화의 진행, 가족가치관의 변화 등 개인과 가족을 둘러싼 개인적, 사회 환경적 변화들이 젊은 세대가 가족을 형성하고 출산을 결심하기 부담스러운 '가족 과부화' 상황을 만들었다고 할 수 있다(송다영, 2013b). 개인적 특성으로서 연령, 교육수준, 경제적 수준, 경제활동 여부 등 개인의 사회인구학적, 그리고 경제적 요인들이 출산 결정에 중요하게 작용하는 것으로 알려졌다(이완, 채재은, 2017; 장한나, 이명석, 2013). 가족가치관 역시 출산결정에 중요하다. 선행연구들은 자녀가 부모에게 갖는 심리적, 사회적, 경

제적 가치, 즉 부모의 관점에서 자녀가 어떠한 의미를 갖는지가 출산의도와 매우 관련성이 높다고 보고했다(배광일, 김경신, 2011; 서정연, 김한곤, 2015b; 이완, 채재은, 2017; 차승은, 2008b). 이상 자녀수 역시 출산행동과 관련이 있는 것으로 알려졌다. 선행연구들(강유진, 2007; 김정석, 2007)은 이상 자녀수를 2명 이상으로 고려하는지가 추가 출산계획을 예측하는 변인으로 보고하였다. 이 외에도 가족 내 부부간 역할 관련 성평등 인식도 출산행동과 관련이 있었다. 서구의 연구들은(McDonald, 2000; Mills et al., 2008) 출산행동은 단지 여성의 행동에만 달린 것이 아니라, 배우자의 행동 역시 영향을 미친다고 주장하면서, 저출산 현상을 설명하는데 있어서 얼마나 가족 내 성평등적 역할분담이 필요하며, 성평등 인식이 중요한지를 주장한 바 있다. 특히 여성의 경제활동 참여가 많아지면서 일-가족 역할 갈등을 경험하는 기혼 직장여성들이 자녀출산을 꺼릴 수 있다. 그러므로 남편의 가사 분담, 집안일 참여 등의 공평성은 기혼여성의 출산의도를 높이는 데 기여한다(박수미, 2008; 송유진, 2014; 신인철, 2009; 장한나, 이명석, 2013; Anderson & Kohler, 2015; Duvander & Andersson, 2008; McDonald, 2000).

이와 같이 다양한 선행연구들이 있으나 관련된 변수들을 전체적으로 고려하여 출산의향을 구체적으로 예측한 연구는 없다. 그러므로 본 연구에서는 가구 특성, 첫 아이 양육경험, 가치관 및 인식 등 다양한 변수들을 전체적으로 고려하여 후속 출산의향을 예측하는 모델을 밝히고자 한다. 본 연구의 목적은 한국보건사회연구원의 2021년도 가족과 출산조사 자료(한국보건사회연구원, 2023)를 활용하여 한자녀 기혼 여성의 후속 출산의향을 예측하기 위한 기계학습 모델을 개발하는 것이다. 먼저, 1) 후속 출산의향 여부에 따라 후속 출산의향을 예측하는 변수에 차이가 있는지 알아볼 것이다. 2) 후속 출산의향을 예측하는 기계학습의 랜덤포레스트 모델을 전통적인 회귀 모델, 부스팅 모델과 비교할 것이다. 3) 후속 출산의향을 예측하는 중요 변수를 확인할 것이다.

## II. 연구방법

### 1. 연구 대상

본 연구의 분석자료는 온라인(<https://data.kihasa.re.kr/kihasa/kor/contents/ContentsList.html>)으로 한국보건사회연구원에서 공개한 2021년도 가족과 출산조사 자료이다(한국보건사회연구원, 2023). 이 자료는 2019년 인구주택 총조사 표본의 틀을 참고하여 전국 19~49세 성인과 그의 배우자를 대상으로 전국 17개 시도에서 층화표집 되었다(박중서 외, 2021). 배우자 상한 연령은 제한을 두지 않았으며, 배우자, 애인, 파트너도 함께 조사하였다. 전체 응답자는 가중치를 적용하여 14,538명(남성 7,498명, 여성 7,040명)이었다. 그중에서 본 연구의 대상은 현재 자녀(출생아 수 - 사망아 수)가 한자녀인 기혼 여성 1,614명의 자료를 분석하였다. 구 전국 출산력 및 가족보건복지 실태조사가 2021년도 가족과 출산조사로 개편되면서, 출산을 측정 그 자체가 아니라 출산 행동에 영향을 미치는 다양한 사회경제적 요인을 측정하고, 기혼과 미혼, 남성과 여성을 모두 조사하여 최종 출산 행동의 결과에만 집중하지 않고 행동을 결정해 가는 과정에도 초점을 두었다.

## 2. 투입변수

### 1) 결과변수

2021년도 가족과 출산조사 자료(한국보건사회연구원, 2023)에서 현재 임신인 경우를 제외하고 ‘앞으로 자녀를 (더) 낳을 계획이 있는지’에 대한 문항에 응답한 자료를 더미변수로 변환하여 출산의향 변수를 결과변수(0 = 추가 출산 의향 없음, 1 = 추가 출산의향 있음)로 예측모형에 투입하였다.

### 2) 예측변수

본 연구에서 한자녀 기혼여성의 후속 출산의향을 예측하기 위해 가구 특성(모 연령, 월평균 가구소득, 주택자가 여부), 첫 자녀 양육경험(첫 자녀 성별, 첫 자녀 대상 지출 총액, 모의 평일 가사노동 시간, 배우자 육아분담비율, 모의 출산전후 휴가 유무, 배우자 출산휴가 유무), 가치관 및 인식(결혼 당시 계획 자녀 수, 도구적 또는 본질적 자녀가치, 가족 및 동료에 대한 신뢰) 등의 정보를 활용하였다. 2021년도 가족과 결혼조사 자료(한국보건사회연구원, 2023) 중 선행연구를 근거로 후속 출산과 관련된 예측 변수를 가구 특성별, 양육경험, 가치관 유형별로 우선 선정하였다. 부분의존도표를 분석에 활용할 경우, 예측 변수 간 강한 상관을 가정하지 않기 때문에, 유형 내에서 대표적인 변수를 최종 예측 변수로 선정하였다. 선정된 예측 변수는 랜덤포레스트, 로지스틱, XGBoosting 모형에 투입하여 예측 성능을 비교하였다.

본 연구에서 예측 모형에 투입되는 결과 변수 및 예측 변수의 내용을 표 1에 제시하였다.

표 1. 투입 변수

구분	변수	내용(척도)
결과 변수	출산의향 <sup>†</sup>	(현재 임신인 경우는 제외) 앞으로 자녀를 (더) 낳을 계획(예 = 1, 아니오 = 2)
가구 특성	모 연령	2019년 인구주택 총조사의 표본 틀에서 19-49세 층화표집됨 (본 연구 대상 범위는 21-49세)
	월평균 가구소득	(지난 6개월간) 월평균 가구소득(세금 공제 전 근로소득, 사업·부업소득, 재산·금융·개인연금소득, 기저귀·분유 등 공적이전소득, 사적이전소득 등의 기타소득) 총액을 만원 단위로 개방적 응답
예측 변수	주택자가 여부 <sup>†</sup>	집의 점유 형태에 대한 문항에서 자가라고 응답한 여부
	첫 자녀 성별 <sup>†</sup>	지금까지 출산한 자녀 중 아들과 딸의 인원수를 개방형으로 응답
	첫 자녀 양육 경험	첫 자녀 대상 지출 총액
	모 평일 가사 시간	응답자의 평일의 1일 평균 가사 시간을 시간과 분 단위로 개방형 응답

표 1. 계속

구분	변수	내용(척도)	
첫 자녀 양육 경험	배우자 육아분담 비율	배우자와 육아분담 비율을 선택(1=응답자 0% 배우자100%, 3=응답자 20% 배우자 80%, 6=응답자 50% 배우자 50%, 9=응답자 80% 배우자 20%, 11= 응답자 100% 배우자 0%)	
	모 출산전후휴가유무 <sup>†</sup>	출산전후 휴가 이용(1=예, 2=아니오)	
	배우자 출산휴가유무 <sup>†</sup>	배우자의 출산휴가 이용(1=예, 2=아니오)	
예측 변수	결혼 당시 계획 자녀 수	결혼(초혼 또는 첫 사실혼 기준) 당시 몇 명의 자녀를 낳을 계획이었는데가에 인원수로 개방형 응답	
	가치관 및 인식	도구적 자녀가치	자녀에 대한 도구적 가치에 대한 문항(예: 자녀는 가족 간 유대와 애정을 강화한다)에 응답한 평균(4점 척도, 1=전혀 동의하지 않는다, 2=동의하지 않는다, 3=동의한다, 4=전적으로 동의한다)
		본질적 자녀가치	자녀에 대한 본질적 가치에 대한 문항(예: 자녀는 그 자체로 나에게 즐거움을 준다)에 응답한 평균
		가족에 대한 신뢰	가족/직장동료(비취업자의 경우 가상의 동료)에 대해 얼마나 신뢰하는지 응답(4점 척도, 1=전혀 신뢰하지 않는다, 2=신뢰하지 않는다, 3=신뢰한다, 4=매우 신뢰한다)
		동료에 대한 신뢰	

<sup>†</sup> 는 더미변수로 변환.

### 3. 분석방법

첫째, 분석에 앞서 결측이 있는 변수는 R의 *mice* 패키지를 활용하여 다중대체법으로 처리하였다. 투입 변수 중 결측 처리된 변수로는 결혼 당시 계획 자녀 수(결측율 21.7%), 배우자 육아분담 비율(20.0%), 첫 자녀 대상 지출 총액(3.8%), 모의 평일 가사 시간(3.2%)이었다. 또한, 출산의향, 첫 자녀 성별, 출산휴가유무는 더미변수로 변환하였다.

둘째, 결과 변수인 출산의향이 있는 집단과 없는 집단 사이에 예측 변수로 선정된 요인에 차이가 있는지 검정(*t*-test,  $\chi^2$  test)하였다. 변수들의 상관관계는 Spearman 방식으로 처리하여 히트맵으로 제시하였다.

셋째, 한자녀 기혼 여성의 후속 출산의향 예측을 위해, 예측과 일반화에 유리한(Fernández-Delgado et al., 2014) 랜덤포레스트 알고리즘으로 예측 모형을 개발하고 예측 성능을 평가하였다. 본 연구의 자료는 후속 출산의향이 없는 집단(73.7%)이 훨씬 많은 불균형 자료로서 후속 출산의향이 없는 집단에 맞추어 균형 자료(upsampling dataset)를 생성하고, 훈련용과 검증용으로 각각 70%, 30%로 구분하고 교차 검증을 수행하였다. 랜덤포레스트 모형의 예측 성능은 전통적 예측 모형인 로지스틱 모형과 높은 예측력으로 자주 활용되는 XGBoosting 모형과 비교하였다. 예측 성능 비교 지수는 정확도, 정밀도, 재현율, F1-score, AUC(Area Under the Receiver Operating Characteristic Curve)를 활용하였다. 특히, AUC 수치는 .90 이상이면 우수한(excellent) 성능으로 인정(Simundic, 2012)하고,

모형 간 AUC 수치에 차이가 있는지 검정(DeLong test)하였다. R의 *randomForest*, *xgboost* 패키지를 활용하여 모델을 개발하고, *caret* 패키지로 자료 균형화, grid search, 10-fold cross-validation으로 hyper parameter 확인하였다.

넷째, 예측 변수의 중요성과 영향은 중요도(importance)와 부분의존도표(partial dependence plot)를 활용하였다. 가장 중요도가 높은 예측 변수의 중요도를 100으로 설정하고, 다른 예측 변수의 중요도는 상대적인 값으로 제시하였다. 부분의존도표는 다른 예측 변수의 영향을 고정하고, 특정 예측 변수의 변화에 따라 결과 변수가 변화하는 관계를 파악하기 위해 분석되었다. 예측 변수의 단위별 변화에 따른 예측력을 확인하기 위해 예측 변수 자료를 표준화하지 않고 모형에 투입하였다. 본 연구의 분석 절차를 도식화하면 그림 1과 같다.



그림 1. 분석 절차 Flow Chart

### III. 결과 및 해석

#### 1. 후속 출산의향 여부에 따른 예측 변수의 차이

한자녀가 있는 기혼 여성 중 후속 출산의향이 있는 집단은 25.96%이다. 표 2는 후속 출산의향이 있는 집단과 없는 집단 간 후속 출산의향을 예측하는 변수의 평균 및 빈도에 차이가 있는지 검정한 결과이다.

표 2. 군집별 예측변수 평균 및 빈도 차이 (N = 1,614)

예측변수	후속 출산의향 없음 (n = 1,195)			후속 출산의향 있음 (n = 419)			전체			t/ $\chi^2$
	m	sd	%	m	sd	%	m	sd	%	
모 연령	39.33	6.02		34.15	4.99		37.99	6.20		17.31***
월평균 가구소득	590.80	288.10		573.00	328.90		586.20	299.20		1.05
주택자가 여부†			62.01			48.93			58.61	21.89***
첫 자녀 성별†			44.44			48.69			45.54	2.26
첫 자녀 대상 지출 총액	79.64	53.61		64.77	40.57		75.78	50.95		5.91**
모 평일 가사 시간	2.84	1.30		2.75	1.17		2.82	1.27		1.22
배우자 육아분담 비율	7.59	1.52		7.84	1.53		7.65	1.52		-2.94**
모 출산전후휴가 유무†			24.27			31.98			26.27	9.53**
배우자 출산휴가유무†			28.79			42.72			32.40	27.50***
결혼 당시 계획 자녀 수	1.61	0.70		2.07	0.68		1.73	0.72		-11.91***
도구적 자녀가치	2.17	0.66		2.17	0.65		2.17	0.66		0.01
본질적 자녀가치	3.43	0.49		3.50	0.45		3.44	0.48		-2.85**
가족에 대한 신뢰	3.58	0.53		3.64	0.52		3.60	0.52		-2.13*
동료에 대한 신뢰	2.63	0.59		2.70	0.56		2.65	0.58		-2.36**

주: † 이분변수, %는 해당 집단 중 (있음)이라고 응답한 비율

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$



표 2에서 월평균 가구소득, 첫 자녀 성별, 모 평일 가사노동 시간, 도구적 자녀가치는 두 집단 간 평균 및 빈도의 차이가 유의하지 않았다. 반면 후속 출산의향이 있는 집단의 가구 특성은 모 연령이 더 젊고( $t = 17.31, p < .001$ ), 주택자가 비율은 더 적었다( $\chi^2 = 21.89, p < .001$ ). 첫 자녀 양육과 관련하여, 첫 자녀 대상 지출 총액은 더 적고( $t = 5.911, p < .01$ ), 배우자의 육아분담 비율은 더 많고( $t = -2.94, p < .01$ ), 모의 출산전후 휴가 비율( $\chi^2 = 9.53, p < .01$ )과 배우자의 출산휴가 비율( $\chi^2 = 27.50, p < .001$ )도 더 많았다. 가치관 및 인식과 관련하여, 자녀에 대한 본질적 가치 수준은 더 높고( $t = -2.85, p < .01$ ), 가족( $t = -2.13, p < .05$ )과 동료( $t = -2.36, p < .01$ )에 대한 신뢰 수준도 더 높다.

상관분석(그림 4)에서, 후속 출산의향과 상관이 가장 높은 변수는 모 연령( $r_s = -.38, p < .001$ )과 결혼 당시 계획 자녀 수( $r_s = .31, p < .001$ )이다. 예측 변수들 사이에 상관이 높은 변수는 월평균 가구소득과 모의 출산전후 휴가유무( $r_s = .41, p < .001$ ), 월평균 가구소득과 첫 자녀 대상 지출총액( $r_s = .36, p < .001$ )이다. 그 외는 .30 미만의 약한 상관관계이다.

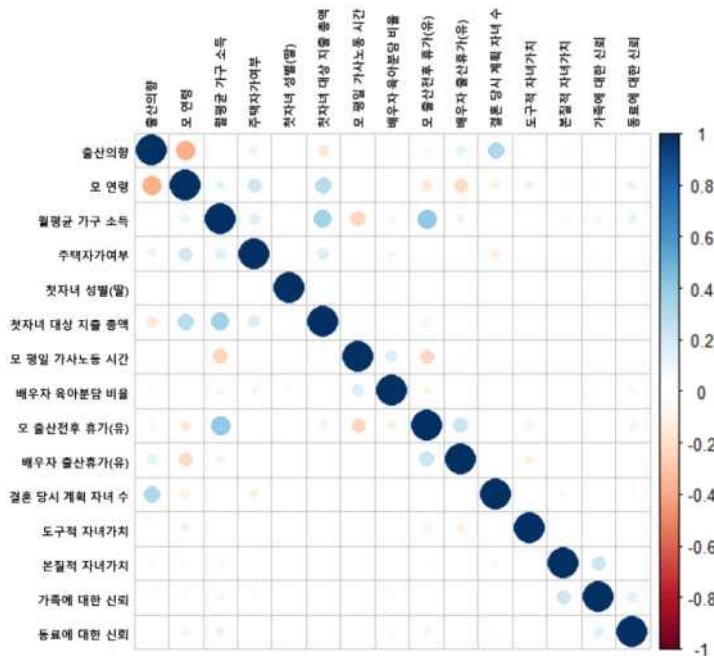


그림 2. 투입변수 간 Spearman 상관 Heatmap

## 2. 후속 출산의향 예측 모형의 성능

### 1) 성능평가 차이

목표 집단에 대한 지도학습 모형의 예측 성능을 표 3에 제시하였다.

표 3. 지도학습 모형의 성능평가

예측모형	정확도	정밀도	재현율	F1-score	AUC	DeLong test
Random Forest	.96	.94	.99	.96	.96	
Logistic	.75	.73	.78	.75	.82	11.72***
XGBoosting	.93	.89	.97	.93	.98	3.85***

주: DeLong test는 해당 모형을 참조로 Random Forest 모형과 AUC 값 차이 검정

\*\*\* $p < .001$

표 3에서, 모형에서 랜덤포레스트 모형의 예측 성능 지수가 가장 크고(정확도 .96, 정밀도 .94, 재현율 .99, F1-score .96), AUC도 .96으로 우수한(excellent) 수준이다. 또한, 랜덤포레스트 모형의 AUC는 로지스틱 모형과 XGBoosting 모형보다 유의하게 크다( $p < .001$ , 그림 3).

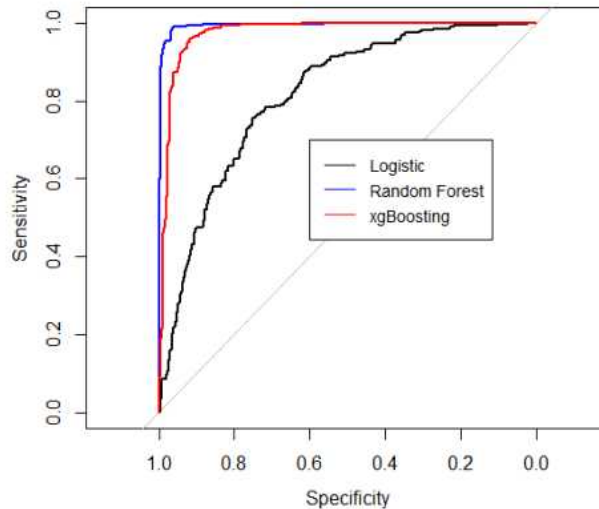


그림 3. 후속 출산의향 예측 모형의 AUC

## 2) 지도학습 모형의 예측 변수 중요도

### (1) 예측 변수의 중요도

랜덤포레스트 모형의 예측 변수 중요도(그림 4)와 부분의존도표(그림 5)를 제시하였다. 랜덤포레스트 모형에서 후속 출산의향을 예측하는 가장 중요한 변수는 모 연령(중요도 = 100)과 결혼 당시 계획 자녀 수(중요도 = 64.15)이다. 그 외에 상대적으로 중요한 변수는 첫 자녀 대상 지출총액(중요도 = 51.91), 월평균가구소득(중요도 = 39.26), 배우자 육아분담비율(중요도 = 18.02), 모의 평일 가사노동 시간(중요도 = 13.24), 배우자 출산휴가 유무(중요도 = 12.42)이다. 나머지 예측 변

수의 중요도는 10미만이다. 비교를 위해 함께 제시한 XGboosting 모형에서도 가장 중요한 변수는 모의 연령, 월평균 가구소득, 첫 자녀 대상 지출 총액, 결혼 당시 계획 자녀 수, 모 평일 가사노동 시간, 배우자 육아분담 비율이고, 그 이외 변수의 중요도는 10 미만이다.

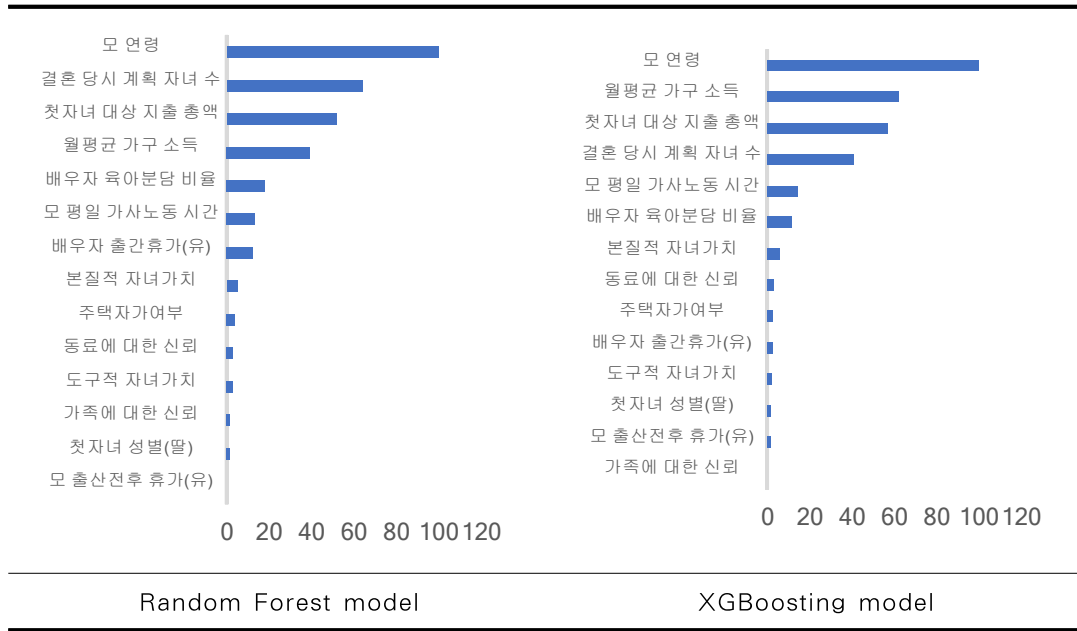


그림 4. 예측 변수의 중요도

(2) 예측 변수의 부분의존도표

랜덤포레스트 모형에서 예측 변수들의 부분의존도표를 변수 유형과 중요도 순서로 그림 5에 제시하였다. 가구 특성과 관련하여, 모의 연령 35세 이상이면 후속 출산의향이 낮아지고, 고소득이거나 주택이 자가인 경우에도 후속 출산의향이 낮아진다.

첫 자녀 양육경험과 관련하여, 후속 출산의향은 첫 자녀에 대한 지출총액이 월 60~70만원을 넘어서면 급격히 낮아지고, 배우자 육아분담 비율이 높을수록 높아진다. 모의 평일 가사노동 시간이 3시간 이상이면 급격히 낮아진다. 배우자의 출산휴가가 있고, 첫 자녀의 성별이 딸이며, 모의 출산전후 휴가가 있으면, 후속 출산의향은 높아진다.

가치관 및 인식과 관련하여, 결혼 당시 계획 자녀수가 2명 이상이면 후속 출산의향은 높아진다. 본질적 자녀가치가 높을수록, 동료에 대한 신뢰가 있으면, 자녀에 대한 도구적 가치를 동의하는 수준(3점)까지 높을수록, 가족에 대한 신뢰가 높을수록 후속 출산의향은 높아진다.

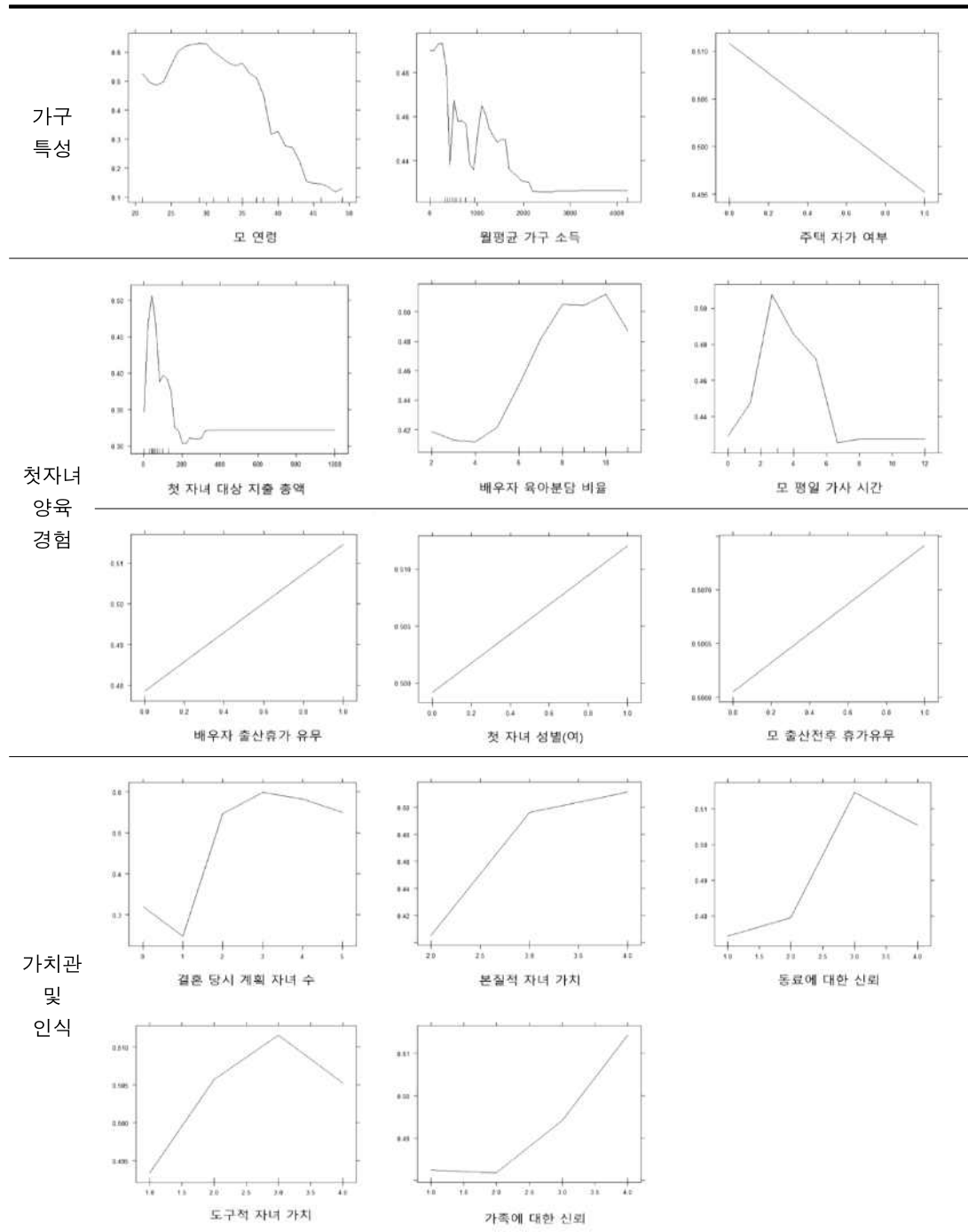


그림 5. 예측 변수의 부분의존도표

## IV. 논의 및 결론

본 연구에서는 한자녀를 둔 가정의 후속 출산 의향에 영향을 미치는 다양한 요인들을 분석하였다. 모의 연령, 결혼 당시 계획한 자녀 수, 월평균 가구소득, 주택자가여부, 첫 자녀에 대한 지출 총액, 모의 평일 가사노동 시간, 배우자 육아분담비율, 모와 배우자의 출산휴가 유무, 가족 및 동료에 대한 신뢰도 등이 중요한 변수로 확인되었다. 분석 결과를 요약하고 이를 토대로 한 논의는 다음과 같다.

첫째, 후속 출산의향이 있는 한자녀 기혼 여성 중 후속 출산의향이 있는 경우는 26%로 나타나 앞으로 둘째 또는 셋째 자녀의 비율은 더욱 낮아질 것으로 예상된다. 그러므로 후속 출산의향이 있는 집단의 특성을 파악하여 한자녀 가정을 대상으로 지원이 가능한 유효한 정책을 제안할 필요가 있다.

둘째, 가구 특성(모 연령, 월평균가구소득, 주택자가여부), 첫 자녀 양육경험(첫 자녀 성별, 첫 자녀 대상 지출 총액, 모의 평일 가사노동 시간, 배우자 육아분담비율, 모의 출산전후 휴가 유무, 배우자 출산휴가 유무), 가치관 및 인식(결혼 당시 계획 자녀 수, 도구적 또는 본질적 자녀가치, 가족 및 동료에 대한 신뢰) 등의 다양한 자료를 랜덤포레스트, 로지스틱, XGBoosting 알고리즘에 투입하여 출산의향 예측 성능을 비교한 결과, 랜덤포레스트 모형의 예측 성능이 두 가지 참조 모형에 비해 유의하게 우수하였다.

셋째, 랜덤포레스트의 예측 모형에서 중요한 역할을 하는 변수의 중요도를 확인한 결과, 모의 연령과 결혼 당시 계획한 자녀수가 후속 출산 의향 예측에 있어 가장 중요한 변수로 나타났다. 모의 연령이 낮을수록 후속 출산 가능성이 높기 때문에 가능한 빠른 시기에 영아를 둔 한자녀 가정을 대상으로 긍정적인 양육경험을 통해 후속 출산을 높이는 노력이 필요하다.

모의 출산전후 휴가유무 변수를 제외한 대부분의 변수가 후속 출산의향을 예측하는데 기여하였다. 예측 변수 중 후속 출산의향을 예측하는데 상대적으로 중요한 변수는 모의 연령, 결혼 당시 계획 자녀 수 외에 월평균 가구소득, 배우자의 육아분담 비율, 모의 평일 가사노동 시간, 배우자 출산휴가 유무였다. 모의 출산전후 휴가 유무보다 배우자의 출산휴가 유무가 더 중요한 것으로 나타나 배우자의 출산휴가와 같은 출산과 관련한 가족지원을 보다 강화할 필요가 있음을 시사한다. 또한, 중요 변수는 아니었지만 직장 혹은 가상의 동료에 대한 신뢰가 가족에 대한 신뢰보다 더 중요했다. 이와 같이 직장의 출산휴가 등 직장의 출산 친화적 환경이 중요한데 여전히 대기업과 중소기업의 차이, 그리고 정규직과 비정규직의 차별 문제가 있으므로, 이를 해결할 수 있는 구체적인 지원과 대책 마련을 통해 남성의 출산휴가와 직장 문화 등 직장의 출산 친화적인 환경을 증진하는 정책이 요구된다고 하겠다.

넷째, 부분의존도표를 통해 예측 변수의 변화에 따른 후속 출산의향을 예측한 결과에 의하면, 가구 특성과 관련하여 모 연령이 35세 이상부터 후속 출산의향이 낮아진다. 또한 고소득이면 후속 출산의향이 낮아진다. 자가주택인 경우 가구 소득과 약한 상관관계( $r_s = .15, p < .001$ )이나 후속 출산의향이 낮아진다. 즉 자가주택이고 고소득일수록 후속 출산의향이 오히려 낮은 것은 가구의

경제적 여건이 후속출산에 절대적인 영향은 아닌 것으로 해석할 수 있다.

다음으로, 첫 자녀 양육경험과 관련하여, 첫 자녀에 대한 지출총액이 월 60~70만원이 넘어서면 후속 출산의향은 급격히 낮아졌다. 배우자 육아분담 비율이 높을수록 후속 출산의향은 높아졌다. 모의 평일 가사노동 시간이 3시간 이상이면 후속 출산의향이 급격히 낮아진다. 배우자 및 모의 출산휴가는 후속 출산의향을 높인다. 첫 자녀가 딸인 경우에도 후속 출산의향이 높아진다. 첫 자녀 양육경험은 다양한 정책적 지원이 가능한 영역으로 자녀 양육 관련 경제적 지원과 직장의 출산 및 가족 친화적인 환경을 위한 정책에 대한 보다 적극적인 노력이 요구된다고 하겠다.

마지막으로, 가치관 및 인식과 관련하여, 결혼 당시 계획 자녀수가 2명 이상이면 후속 출산의향이 높아진다. 결혼과 가족조사 자료에서 결혼 당시 계획 자녀수는 이상적인 자녀 수( $r_s = .34$ ,  $p < .001$ )와 상관이 있었다. 본질적 및 도구적 자녀가치가 높아도 후속 출산의향이 높아진다. 동료와 가족에 대한 신뢰가 높아도 후속 출산의향이 높아진다. 이는 단기간의 정책으로 개선될 수는 없으나 결혼이나 가족 및 자녀가치에 대한 사회문화적 환경 개선을 위한 장기적인 노력도 필요할 것이다.

본 연구는 한자녀를 둔 가정의 후속 출산 의향에 영향을 미치는 가구 특성, 첫 자녀 양육경험, 가치관 및 인식 등 다양한 요인들을 분석하였으나, 이 외에 양육스트레스, 결혼만족도 등 심리적 요인들이 데이터는 부재하여 분석에 포함되지 않았다. 향후 다양한 가족의 심리적 요인들을 포함한 후속 연구가 요청된다. 그럼에도 불구하고 본 연구에서 밝힌 결과를 토대로 다음과 같은 결론을 내릴 수 있다. 저출생 문제에 대응하는 정책을 수립할 때, 단순히 경제적 인센티브를 제공하는 것을 넘어서 가족 구성원의 역할 분담, 일과 삶의 균형, 자녀에 대한 가치관 및 신뢰 구축 등 다차원적인 접근이 필요함을 시사한다. 특히, 모의 연령이 후속 출산 의향에 미치는 영향을 고려할 때, 젊은 부부의 출산 및 양육에 대한 긍정적인 인식을 증진시킬 수 있는 교육과 출산과 양육에 대해 긍정적인 결정을 내릴 수 있는 환경을 조성하는 것이 중요하다. 이들에게 긍정적인 양육경험을 할 수 있는 구체적인 양육지원시스템을 개발할 필요가 있다. 또한 일과 삶의 균형을 지원하기 위한 유연한 근무 환경 조성 및 양육 휴가제도 개선, 가족 내 역할 분담과 양성평등을 증진시키기 위한 사회적 캠페인 및 정책 마련 등 다각도의 접근을 통해, 한국 사회는 저출생 문제를 극복하고 지속 가능한 발전을 이룰 수 있을 것이다.

## 참고문헌

- 강경희, 전홍주 (2013). OECD 국가의 양육지원 정책과 출산을 분석: 현금지원 정책, 보육시설서비스 정책, 조세혜택 정책을 중심으로. **한국보육지원학회지**, 9(6), 197-221. <https://doi.org/10.14698/JKCCE.2013.9.6.197>
- 강유진 (2007). 둘째 자녀 출산계획에 영향을 미치는 자녀가치관 및 가족관계 요인. **한국지역사회생활과학회지**, 18(4), 639-653.
- 강은영, 진미정, 옥선화 (2010). 비혼 여성의 비혼 자발성 관련 요인 탐색: 결혼 및 가족 가치관을

- 중심으로. *Family and Environment Research*, 48(2), 135-144.
- 김정석 (2007). 기혼여성의 출산아수별 추가출산계획. *한국인구학*, 30(2), 97-116.
- 박수미 (2008). 둘째 출산계획의 결정요인과 가족 내 성 형평성. *한국인구학*, 31(1), 59-77.
- 배광일, 김경신 (2011). 생태학적 변인이 출산의도에 미치는 영향 연구. *한국가족관계학회지*, 16(2), 123-149.
- 배혜원 (2017a). 한국 저출산 문제와 가족 정책 방향. *가족과 문화*, 29(3), 211-234.
- 배혜원 (2017b). 기혼여성의 가족가치관 변화와 정책적 시사점. *보건복지포럼*, 251, 86-100.
- 백지희 (2020). 만 1세 첫째 자녀의 성별과 기질이 한 자녀 부모의 후속출산 계획 및 행동에 미치는 영향. *한국보육지원학회지*, 16(6), 19-34. <https://doi.org/10.14698/jkcce.2020.16.06.019>
- 서정연, 김한곤 (2015a). 한국의 결혼과 출산에 대한 생태계 동인성 분석. *한국가족관계학회지*, 19(3), 3-23.
- 서정연, 김한곤 (2015b). 출산자녀수별 기혼여성의 출산의지에 영향을 미치는 요인. *한국인구학*, 38(3), 1-24.
- 송다영 (2013a). 한국의 저출산 문제에 대한 진단과 정책적 대응. 한국보건사회연구원.
- 송다영 (2013b). 여성주의 관점에서 본 생애주기별 복지와 돌봄패러다임. *페미니즘연구*, 13(1), 93-129.
- 송유진 (2014). 기혼여성의 자녀출산계획에 영향을 미치는 요인. *한국지역사회생활과학회지*, 25(3), 339-347. <https://doi.org/10.7856/kjcls.2014.25.3.339>
- 신인철 (2009). 기혼여성의 자녀출산계획에 대한 공간효과 분석. *한국인구학*, 32(2), 59-85.
- 이미옥, 명성준 (2015). 출산장려정책의 효과성에 대한 연구. *한국자치행정학보*, 29(1), 331-350. <https://doi.org/10.18398/kjlgas.2015.29.1.331>
- 이완, 채재은 (2017). 기혼여성의 자녀가치관과 추가출산의향 간의 관계에서 출산장려정책의 조절효과 분석. *디지털융복합연구*, 15(9), 65-78. <https://dx.doi.org/10.14400/JDC.2017.15.9.65>
- 이철희 (2018). “한국의 출산장려정책은 실패했는가?” 2000년-2016년 출산율 변화요인 분해. *경제학연구*, 66(3), 5-42.
- 장한나, 이명석 (2013). 기혼 직장여성의 출산의도 결정요인에 관한 연구. *대한가정학회지*, 51(4), 425-438.
- 최미란, 장경은 (2017). 영아기 첫 자녀를 둔 취업모의 양육스트레스와 후속 출산의도의 관계: 직업만족도의 조절효과. *한국보육지원학회지*, 13(4), 53-73. <https://doi.org/10.14698/jkcce.2017.13.04.053>
- 차승은 (2008a). 한국의 저출산 문제: 출산행동에 대한 인과관계 모형 분석. *한국인구학*, 31(1), 1-28.
- 차승은 (2008b). 부모역할의 보상/비용과 둘째 자녀 출산계획: 사회경제적 특성에 따른 차이를 중심으로. *사회복지정책*, 33, 111-134.
- 통계청 (2023). 2023년 11월 인구동향 보도자료. [http://kostat.go.kr/assist/synap/preview/skin/doc.html?fn=synapview382662\\_1&rs=/assist/synap/preview](http://kostat.go.kr/assist/synap/preview/skin/doc.html?fn=synapview382662_1&rs=/assist/synap/preview) 2024년 1월 31일 인출함.

- 한국보건사회연구원 (2023). 2021년도 가족과 출산조사 자료. <https://data.kihasa.re.kr/kihasa/kor/contents/ContentsList.html> 2023년 10월 24일 다운로드함.
- Anderson, T., & Kohler, H. (2015). Fertility, Socioeconomic Development, and Gender Equity. *Population and Development Review*, 41(3), 381-407. <https://doi.org/10.1111/j.1728-4457.2015.00065.x>
- Choi, Y., & Kwon, S. (2018). Work-Family Balance Policies in South Korea. In *The Palgrave Handbook of Family Policy* (pp. 483-505). Palgrave Macmillan, London. <http://doi:library.oapen.org/handle/20.500.12657/46123>
- Cho, N., & Yoo, S. H. (2018). The Impact of the Financial Incentives for Childbearing on Fertility in South Korea. *Journal of Family and Economic Issues*, 39(1), 81-93. <http://doi.org/10.1186/s12889-022-14722-4>
- Duvander, A., & Andersson, G. (2008). Gender Equality and Fertility in Sweden  
A Study on the Impact of the Father's Uptake of Parental Leave on Continued Childbearing. *Marriage & Family Review*, 39(2), 121-142. [https://doi.org/10.1300/J002v39n01\\_07](https://doi.org/10.1300/J002v39n01_07)
- Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we Need Hundreds of Classifiers to Solve Real World Classification Problems? *Journal of Machine Learning Research*, 15, 3133-3181. <https://www.researchgate.net/publication/279556630>.
- Jeong, K., Yoon, J., Cho, H., Kim, S., & Jang, J. (2022). The relationship between changes in the korean fertility rate and policies to encourage fertility. *BMC Public Health*, 22, 2298. <https://doi.org/10.1186/s12889-022-14722-4>
- Kwon, E., & Chang, H. (2019). The Influence of Social Recognition on the Fertility Intentions of Young Korean Adults. *Population and Development Review*, 45(3), 597-620. <https://doi.org/10.1093/pubmed/fox1281258728K>
- Lee, S. (2019). The Effects of the Fertility Policy on Marriage and Fertility in South Korea. *Journal of Population Economics*, 32(4), 1261-1294.
- McDonald, P. (2000). Gender Equity in Theories of Fertility Transition, *Population and Development Review*, 26(3), 427-439. <https://doi.org/10.1111/j.1728-4457.2000.00427.x>
- Mills, M., Mencarini, L., Tanturri, M. L., & Begall, K. (2008). Gender Equity and Fertility Intentions in Italy and the Netherlands. *Demographic Research*, 18(1), 1-26. <https://www.jstor.org/stable/26347977>
- Seo, S. H. (2019). Low fertility trend in the Republic of Korea and the problems of its family and demographic policy implementation. *Population and Economics*, 3(2), 29-35. <https://doi.org/10.3897/popecon.3.e37938>
- Simundic AM. (2012). Practical recommendations for statistical analysis and data presentation in Biochemia Medica journal. *Biochemia Medica*. 22, 15-23. DOI: 10.11613/BM.2012.003
- Zhang, T., Cai, X., Shi, X., Zhu, W., & Shan, S. (2023). The Effect of Family Fertility Support Policies on Fertility, Their Contribution, and Policy Pathways to Fertility Improvement in OECD



Countries. *International Journal of environmental Research and Public Health*, 20, 4790.  
<https://doi.org/10.3390/ijerph20064790>

논문투고 : 24.03.13.  
수정원고접수 : 24.03.30.  
최종게재결정 : 24.04.04.