

## 우리나라 가구의 자녀수 결정요인에 관한 Count 모형 분석 및 경제적 함의\*

김 현 숙\*\*

본 연구에서는 우리나라 가구의 자녀수 결정요인을 Count 모형을 이용하여 분석하였다. 분석 기법으로는 자녀수가 감마(Gamma) count 분포를 따른다는 가정 하에 기본적인 회귀분석과 40세 이하 기혼여성의 완결출산에 대한 대리변수로 예상출산자녀수에 대한 회귀분석, 허들모형 그리고 기혼여성의 노동시장 참여와 자녀출산간의 내생성을 고려한 모형을 각각 이용하여 다각적인 방법론으로 출산결정요인을 분석하였다.

분석결과, 가구의 소득수준이 높을수록 40세 이상 기혼여성의 자녀수는 많은 것으로 나타나 는 반면, 18~39세 기혼여성의 경우에는 소득이 자녀수에 큰 영향을 미치지 못하는 것으로 보인다. 기혼여성의 경제활동참여는 39세 이하의 비교적 젊은 여성들에게서 출산 예상자녀수에 부정적인 영향을 주는 것을 뚜렷이 확인할 수 있고, 출산과 경제활동참여 간에는 내생성이 존재하는 것으로 확인되었다. 허들모형을 이용한 결과, 기혼여성의 학력은 자녀출산 여부에는 긍정적으로 작용할 수 있으나 자녀수 결정에는 부정적으로 작용하여 기혼여성 학력이 높을수록 출산율은 감소하게 된다. 다양한 회귀분석 결과에 근거할 때 Becker의 Quantity-Quality 모형의 핵심내용은 우리나라의 자녀수 결정에 있어서도 대체적으로 적용됨을 확인할 수 있다.

**핵심단어 :** Becker의 Quantity-Quality 모형, Count 모형, 감마분포, 완결출산율, 예상출산 자녀수, 허들모형, 내생성

### I. 서론

출산을 결정모형에 대한 경제학적 연구는 오랫동안 진행되어 왔으나 그 연구의 근간은 Heckman & Willis(1975)의 생애효용모형과 Becker(1960)의 Quantity-Quality 모형으로 대표될 수 있다. Becker의 모형은 30여년이 지났지만 여전히

\* 본 연구는 한국조세연구원의 연구보고서인 <장기적 인적자본 형성을 위한 조세·재정정책: 출산율 결정요인의 경제학적 분석>의 저자 집필 내용일부를 수정한 것이다. 아울러 본 연구는 숭실대학교 교내연구비 지원으로 이루어졌음을 명기한다.

\*\* 숭실대학교 경제학과 교수, 이메일: annakim@ssu.ac.kr 전화: 02-828-7256

출산을 결정모형의 이론적 기초로 활용되고 있고, 각 나라들의 데이터를 이용하여 Becker의 이론을 검증하는 작업이 활발히 이루어져 왔다. Becker(1960)는 Quantity-Quality 모형의 기본적인 틀을 처음 제시한 논문으로 자녀의 질적 수준의 초기부존자원에 대한 고려가 없고 자녀수에 대한 소득탄력성이 자녀의 질적 수준에 대한 소득탄력성보다 크다는 가정을 이용하여 가구의 소득이 증가하면서 자녀수가 감소한다는 명제를 도출한다. Becker & Lewis(1973)는 소득탄력성과 가격탄력성 내용을 추가하였으며, 나아가 Becker & Nigel(1976)는 초기 부존자원에 대한 고려를 통해 자녀수와 자녀의 질적 수준의 소득탄력성의 크기에 대한 결론이 사전적인 가정 없이 도출되도록 수정하였다.

위와 같은 Becker 모형의 함의를 실증적으로 검증하는 것은 한 여성이 생애에 걸쳐 출산하는 자녀의 수를 나타내는 완결출산에 영향을 미치는 요소가 무엇인가라는 질문으로 귀결된다. 이와 같은 분석에 가장 널리 이용되는 것은 count 모형<sup>1)</sup>이다. 최근의 연구로는 Caudill & Mixon(1995), Gensler(1997), Nguyen-Dinh(1997), Wang & Famoye(1997), Melkersson & Rooth(2000), Silva & Covas(2000)을 들 수 있다. 먼저 Caudill & Mixon(1995)은 count 모형을 이용하여 Becker-Lewis의 가설<sup>2)</sup>이 실제 데이터(Consumer Expenditure Survey)와 일치하고 포아송 모형과 음이항(negative-binomial) 모형이 OLS 모형보다 통계적으로 우월함을 입증했다. Wang & Famoye(1997)은 기존의 count 모형이 분석적 단순함을 위해서 출산 행태에 대한 너무 강한 가정을 했다고 비판을 하면서, 그 대안으로 일반화된 포아송 회귀 모형(Generalized Poisson Regression Model, GPR)을 제시했다. Melkersson & Rooth(2000)의 연구는 자녀를 낳을 때 몇 명을 낳을 것인가와 자녀를 낳을 것인지 여부 결정은 질적으로 다른 성질의 결정으로 본다. 따라서 자녀의 수가 0인 경우에는 다른 결정 과정을 따르는 것으로 가정하고 inflated count data 모형을 발전시켰다. Silva & Covas(2000)는 좀 더 나아가 무자녀 또는 한 명의 자녀를 갖는 결정과 두 명 이상의 자녀를 갖는 결정 사이에는 큰 차이가 존재하고, 따라서 포아송 회귀모형으로는 설명할 수 없는 출산행태가 있다고 보았다. 이런 현상을

- 
- 1) count 모형이 아닌 다른 형태의 모형, 예를 들어 최소자승추정법을 이용하면 종속변수의 서포트(support)가 음이 아니어야 한다는 제약조건을 충족시킬 수 없다. 따라서 일치추정치(consistent estimator)를 얻기 위해서는 음(-)이 아닌 정수를 종속변수로 이용할 경우 count 모형을 이용하는 것이 바람직하다. 자세한 내용은 Cameron & Trivedi(1998) 참조.
  - 2) 통상적으로 Becker-Lewis 가설이라고 함은, 여성의 임금, 노동시장 참여 여부, 교육은 출산결정에, 기회비용의 상승으로 인해, 음(-)의 효과를 미치고, 가구의 소득은 자녀의 양-질의 trade-off로 인해 발생하는 소득의 대체효과와 소득효과의 크기에 따라 출산결정에 양(+) 혹은 음(-)의 영향을 미칠 것이라는 주장이다.

설명하기 위해서 허들(hurdle) 모형을 발전시키고 포르투갈의 출산 자료를 이용하여, 허들모형이 출산 행태를 좀 더 잘 설명해줄 수 있음을 보였다.

본 연구는 이와 같은 count 모형을 이용한 선행연구의 방법론을 우리나라 자료를 이용하여 검증하고 과연 우리나라 가구의 자녀수 결정에 영향을 미치는 요인은 무엇인지 분석하는 것을 목적으로 한다. 본 논문의 구성은 다음과 같다. II장에서는 기본적인 자녀수 결정의 count 모형을 설정하고 자녀수 분포가 감마분포를 따름을 확인하여 감마 count 회귀분석의 결과를 살펴본다. III장에서는 자녀수 결정모형을 확대하여 출산이 아직 완결되지 않은 18~39세 기혼여성의 생애에 걸친 출산자녀수에 대한 대리변수로 예상출산자녀수를 이용하여 젊은 세대의 자녀수 결정행태에 대해 분석한다. 아울러 허들모형을 도입하여 자녀출산여부와 출산자녀수를 결정하는 요인을 나누어 분석하고 기혼여성의 노동시장 참여와 출산기간의 내생성을 고려한 모형을 도입한다. 마지막으로 IV장에서는 본 논문의 연구결과를 요약하고 실증결과에 기초하여 Becker 이론의 성립여부를 검증해본다.

## II. 기본적인 자녀수 결정모형

### 1. 완결출산율

출산율을 측정하는 지표로 가장 많이 이용되는 것이 합계출산율이다. 그런데 합계출산율은 한 여성이 평생의 가임기간에 걸쳐 낳은 자녀수를 칭하는 완결출산율과는 달리 출산율을 측정하는 시점의 각 연령코호트별 자녀수의 평균으로 나타나기 때문에 아직 가임기간이 끝나지 않은 여성이 향후 추가적으로 출산할 자녀수나 결혼은 했으나 출산시기를 지연하거나 자녀 간 터울조절을 하고 있는 여성들이 조사 시점에 출산한 자녀수가 적을 경우 완결출산율과는 큰 차이를 보일 수 있다.

자녀수 결정모형의 큰 특징이면서 장점은 많은 연구들이 합계출산율에 영향을 주는 요인을 분석하는 것과 다르게 완결출산율을 결정하는 요인을 파악하는 것이다. 일반적으로 완결출산율에 대한 분석 시 사용되는 표본은 가임기간이 끝난 여성 집단이다. 여성의 가임기를 폭넓게 고려할 경우 15-44세(혹은 15~49세)로 보는 경우가 많다. 따라서 완결출산율을 분석할 때 45세 이상의 여성을

대상으로 하는 것이 가장 일반적일 수 있다. 단 45세 이상의 여성표본이 매우 적거나 연령을 다소 하향조정하여도 연구결과에 큰 영향이 없을 때는 이를 40세 전후까지 조정하기도 한다.

따라서 완결출산율 분석은 특정 연령계층 이상만을 대상으로 하므로 젊은 세대의 가치관이나 코호트별로 변화하는 출산행태를 포착하기 어려우며 출산시작부터 출산을 완결하는 시점까지의 여성의 출산행태의 연속적 흐름을 볼 수 없다는 단점이 있다. 비록 45세 이상의 여성들의 완결출산율이 높다고 해도 20대나 30대 여성들이 향후 출산할 자녀수가 적다면 장래의 완결출산율은 저하하게 된다. 따라서 20-30대 젊은 여성의 향후 완결출산율을 추정하여 이를 이용하여 전 연령층의 완결출산율을 추정하는 것이 필요하지만 아직 가임기간이 끝나지 않은 여성의 완결출산율에 대한 추정은 매우 어렵다.

본 연구는 우선 40세 이상의 여성에 대한 표본을 이용하여 우리나라 완결출산율에 영향을 주는 요인을 분석한다. 그러나 40세를 초과하는 여성만을 표본으로 할 경우, 젊은 세대의 출산율에 영향을 주는 요인들을 포착하기 어렵고 연령이 높은 세대의 출산율에만 분석이 집중되므로 향후 출산율 변화에 대한 예측이나 시사점을 제공하기 어렵다는 점을 고려하여 가임기간이 끝나지 않은 여성들이 생애동안 출산할 자녀수에 준하는 대체변수를 구하여 결혼한 부인의 전 연령에 대한 횡단면 분석도 시도한다.

또한 물리적 불임으로 인해 자녀를 출산할 수 없는 경우와 의도적으로 자녀를 출산하지 않는 경우가 혼재하고 있다는 사실에 착안하여 허들 모형을 고려하여 자녀의 존재유무와 1명 이상의 자녀의 수를 결정하는 요인을 분리하여 추정하는 방법도 시도하였다.

마지막으로 출산자녀수에 대한 설명변수로 이용한 기혼여성의 경제활동참가가 내생적인 변수일 수 있다는 가정 하에 count 모형에서 내생적 설명변수를 고려한 경우를 모형화하여 그 설명변수를 외생변수로 고려하고 추정한 모형과 비교하였다.

## 2. 자녀 수 분포결정 및 이용자료

자녀수 결정모형 시 가장 먼저 고려되는 문제는 계량경제학적 방법론에 따라 자녀수의 분포를 대표하는 적합한 분포를 찾는 것이다. 자녀수는 0,1,2,3,...의 정수 형태이므로 포아송 분포나 그와 유사한 분포를 고려해야 한다. 포아송 분포를 고려할 때 문제가 되는 것은 개인 간의 이질성을 포아송 분포가 제대로 포

착하지 못한다는 점이다. 이에 따라 음이항 분포가 고려되기도 하며 일반화된 포아송 분포나 허들모형도 고려된다.

본 논문에서는 선행연구들에 기초하여 자녀수 결정모형 구축 시 고려해야 하는 여러 가지 요건들을 고려하여 이에 적합한 다양한 분석을 시도함으로써 자녀수 결정요인에 대해 심도 있는 분석을 시도한다.

우선 출산을 결정요인에 대한 미시적 분석을 위해 이용한 자료는 <2003년 전국출산력 및 가족보건·복지실태조사>이다. 이는 전국의 13,000여 가구의 기혼 여성을 대상으로 가구의 인적사항, 경제관련 사항, 가족복지, 임신, 출산, 피임, 모자보건에 대한 사항을 조사한 것이다. 단년도의 횡단면 자료이므로 한 여성의 평생에 걸친 출산과 다른 의사결정의 궤적을 알 수 없다는 단점은 있으나 여성의 임신 및 출산, 취업여부, 가구의 경제적 상태, 가구의 인적 구성 등 출산을 결정하는데 영향을 미치리라고 추정되는 변수들을 비교적 소상하게 포함하고 있어 실증분석의 자료로 이용하였다. 표본자료에 포함된 여성은 18-65세 사이의 기혼여성(미망인 포함)이다. 이 중 자녀수 결정모형을 구축하기 위해 사용된 전국출산력 기초자료의 주요 변수들을 정리하면 다음과 같다(<표1 참조>).

우선 이용한 자료의 변수 중 자녀수가 어떤 분포를 따르는지 밝히는 것이 중요하다. 음(-)이 아닌 정수를 종속변수로 하는 count 모형에서는 흔히 종속변수가 포아송 분포를 따른다는 가정을 이용한다. 포아송 분포는 음이 아닌 정수가 각각 다음과 같은 확률분포를 가지는 경우를 의미하며 분산이 평균이 동일하여  $E(Y | x) = Var(Y | x) = \mu$ 인 특징을 지니고 있다. 확률밀도함수는 다음과 같다.

$$f(Y = y | x) = \frac{e^{-\mu} \mu^y}{y!}$$

그런데 평균과 분산이 일치하지 않아 분산이 평균보다 큰 경우나 평균보다 작은 경우에 포아송 분포를 사용하게 되면 추정치가 편의(bias)를 갖게 된다.

일반적으로 분산이 평균보다 커서 over-dispersion의 문제를 가질 경우에는 이항분포를 사용하며, 반대의 경우에는 감마(Gamma)분포를 이용하게 된다. 일반적으로 출산자녀 수의 분포는 평균보다 분산이 작은 under-dispersion 형태를 띤다는 연구결과가 많다. 예를 들어 Heckman and Walker(1990a, 1990b)는 스웨덴의 출산자료를 이용하여 출산사이의 간격이 일정하지 않고 간격사이에 양(+ )의 duration 의존성이 있다는 사실을 밝혔다. 이는 자녀수의 분포에 있어

&lt;표 1&gt; 전국출산력 자료 주요변수 기초통계

변수	관측치 수	평균	표준편차	최소값	최대값
출산자녀수(출생아수)	6,453	1.89	0.86	0	7
예정출산자녀수	6,453	2.08	0.72	0	8
주택 소유 여부(소유=1)	6,453	0.59	0.49	0	1
월 가구 총소득(원)	6,453	2,903,054	1,641,077	0	9,970,000
부인의 취업여부(취업=1)	6,453	0.52	0.50	0	1
부인의 교육수준(학력년수)	6,453	11.96	2.77	0	16
부인의 현재 연령	6,453	37.71	6.68	18	59
남편의 교육수준(학력년수)	6,453	12.61	2.95	0	16
남편의 현재 연령	6,453	41.01	7.28	20	68
인공유산 횟수	6,453	0.67	0.99	0	10
자연유산 횟수	6,453	0.27	0.63	0	8
부인의 혼인연령	6,453	24.15	3.64	14	47
남편의 혼인연령	6,453	27.45	3.81	13	62
부인 친정어머니 자녀양육 도움여부(도움=1)	6,453	0.013	0.11	0	1
부인 시어머니 자녀양육 도움여부(도움=1)	6,453	0.025	0.16	0	1
아들선호 여부(선호=1)	6,453	0.58	0.49	0	1
남편의 장남여부(장남=1)	6,453	0.29	0.45	0	1
부인의 형제자매 수	6,453	4.84	1.71	1	12
남편의 형제자매 수	6,453	4.74	1.78	0	12
대도시 거주여부	6,453	0.43	0.50	0	1
중소도시 거주여부	6,453	0.31	0.46	0	1

분산이 평균보다 작다는 것과 양립하는 것이므로 자녀 수 분포가 포아송 분포를 따르지 않는다는 추론과 연결된다.<sup>3)</sup> Winkelmann(1995)의 표현에 따라 각 부분포의 밀도함수를 나타내면 다음과 같다.

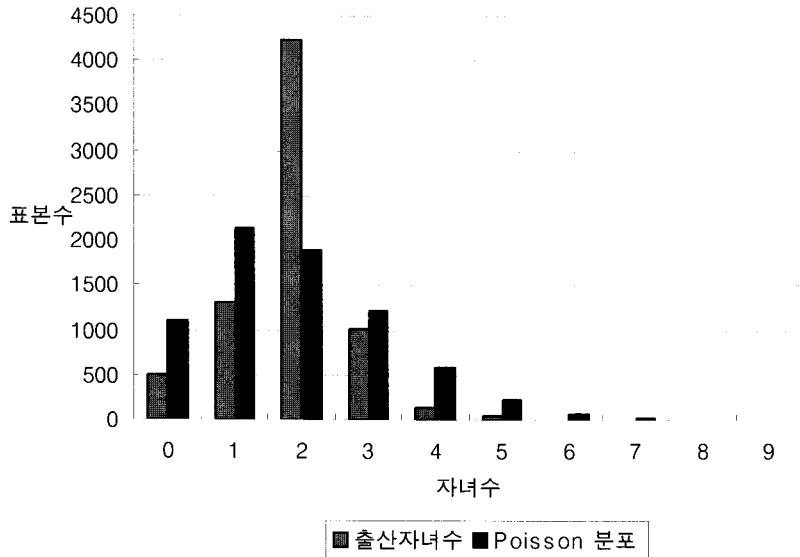
$$f(y | x) = \Gamma(\alpha y, \alpha \mu) - \Gamma(\alpha(y+1), \alpha \mu)$$

여기서  $\Gamma(\cdot)$ 는 감마 적분으로  $\Gamma(\cdot) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt$  이며,

$\int_0^{\infty} \Gamma(0, \alpha \mu) = 1$ 이 성립한다. 이러한 감마분포에서 평균은  $\mu$ 이고 분산은  $\frac{\mu}{\alpha}$ 이다. 따라서  $\alpha$ 가 1보다 작으면 분산이 평균보다 커서 넓게 퍼져있고, 반대로 1보다 크면 분산이 평균보다 작게 된다. 만약  $\alpha$ 가 1이면 자녀 사이의 출산간격

3) 자세한 내용은 Winkelmann(1995) 참조

<그림 1> 관측된 출생아수와 포아송 분포에 따른 출생아수 분포비교



이 지수함수형태로 유지되어 포아송 분포를 따르게 된다.

본 연구에서 사용하는 2003년 전국출산력자료의 출생자녀수에 대한 분포를 살펴보고 관측된 자녀수를 가장 잘 반영하는 분포의 형태를 찾는 것이 중요하다. 완결출산율을 파악할 때 가장 흔히 이용되는 변수는 출생 후 사망한 자녀의 수까지 합한 총 출생아 수이다. 그 외에 현존 자녀 수를 이용하는 경우도 있는데 이는 주로 출생아 수에 대한 정보가 부정확한 경우 대리변수로 이용되고 있다. 전국출산력 자료는 18~65세의 부인들을 대상으로 하여 조사된 서베이 자료로 이 중 40대 이상의 여성을 이용하여 완결출산율을 직접적으로 관측할 수 있고, 40대 미만의 기혼여성을 표본에 포함시킬 경우에는 불완결 출산율 (incompleted fertility rates)을 구할 수 있다.<sup>4)</sup> 본 연구는 40세 이상의 여성의 경우에는 총 출생아수를 변수로 이용하고, 18세 이상의 전체 기혼여성을 표본으로 이용할 경우에는 총 출생아수+ 현재 임신 중이며 출산예정인 자녀 수+ 향후 추가예정 자녀수를 예상 총 출생아 수라 보고 완결출산율에 대한 대리변수로 이용하였다.

4) Wang and Famoye(1997)은 미시간 패널자료인 PSID 자료 중 18~40세의 기혼여성 중 가구주가 아니고, 가구의 소득이 음이 아닌 경우를 표본으로 이용하였다.

위의 출생아수와 예상 출생아수의 평균을 유지하면서 포아송 분포를 따르는 자녀수 분포를 관측 치와 비교해 보기 위해 동일한 관측치 수를 가지면서 포아송 분포를 따르고 평균은 서베이 관측치와 동일한 값을 유지하는 무작위 수를 추출하였다. 포아송 분포를 따르는 무작위 추출의 경우에는 표준편차가 관측치의 2배 가까이에 이르고 있어 관측치의 경우 포아송 분포를 따른다고 보기 어렵다. 따라서 평균이 분산보다 큰 경우인 감마분포를 이용하여 모형을 추정하는 것이 적절한 것으로 보인다.

관측치에서 출산자녀수가 0인 경우는 전체 표본을 대상으로 할 경우 그 비중이 6.84%에 불과하여 포아송 분포를 따르는 무작위 추출경우의 15.12%와 대조되는 것으로 나타나 전국출산력 자료 서베이 관측치의 총 출산자녀수는 포아송 분포와 비교할 때 출산자녀가 없는 경우가 매우 과소하게 추출되었음을 확인할 수 있다. 출산 자녀수가 1명인 경우에도 관측치의 경우가 다소 과소하게 나타났다. 반대로 2명의 자녀를 출산한 경우는 관측치에서 58.4%를 차지하는 반면, 포아송 분포를 따른다고 가정할 경우에는 26.1%에 불과하여 두 배 이상 많은 것을 확인할 수 있다.

따라서 분산이 평균보다 크다는 추측을 할 수 있어 자녀 수 모형의 분산이 평균보다 큰지 여부를 통계학적으로 판별하기 위해 Cameron and Trivedi (1998)의 방법론을 이용하였다. 간단하게 소개하면 다음과 같이 정의되는 분산에 대해 월드(Wald) 검정을 통해 분산의 크기를 측정하는 것이다. 분산이 평균보다 크면  $\alpha$ 가 양수이고, 분산이 평균보다 작으면  $\alpha$ 가 음수로 나타난다.

$$V(Y_i | x_i) = \mu + \alpha g(\mu)$$

여기서  $\mu$ 는 평균이고,  $g(\mu)$ 는  $\mu$ , 혹은  $\mu^2$ 으로 정의된다. 먼저 포아송 모형을 이용하여 자녀수의 추정치 (fitted value)인  $\hat{\mu} = \exp(\hat{\beta}'x_i)$ 을 구하고 다음의 식을 OLS를 통해 추정한다. 여기서  $\alpha$ 의 값이 양수인지 음수인지 여부를 통해 분산의 정도를 확인할 수 있다.

$$\frac{(y_i - \hat{\mu})^2 - y_i}{\hat{\mu}} = \alpha \frac{g(\hat{\mu})}{\hat{\mu}} + \varepsilon_i$$

월드검정을 통해 자녀수 분포를 살펴보기 위해  $x_i$  즉 설명변수로는 기혼여성



의 연령, 결혼연령, 학력수준, 자녀수나 자녀성별에 대한 가치관에 영향을 줄 수 있는 변수들을 1차적으로 고려하였다. 가구의 소득이나 기혼여성의 직장유무 등도 매우 중요한 경제적 설명요인이지만 가구의 소득이나 기혼여성의 직장유무, 기혼여성의 근로소득 등은 서베이 자료를 이용할 경우 조사시점의 소득만 주어지며 항상소득을 알 수 없다는 단점을 지니고 있다. 따라서 생애주기에 따라 변화하는 현재소득변수를 이용할 경우에는 다소 그 결과를 해석하는데 신중할 필요가 있다. 기혼여성만이 아니라 남편의 학력, 연령, 결혼시 연령 등도 분석 시 이용해보았으나 그 유의성이 거의 없는 것으로 나타나 회귀분석 시 설명변수로 사용하지는 않았다.<sup>5)</sup>

본 연구에서는 기혼여성의 임금에 대한 관측치가 조사 시점에만 존재한다는 문제점을 보완하기 위해 기혼여성의 학력을 임금에 대한 대리변수로 사용한다.

<표 2> 출산자녀수 결정요인 분석 (Poisson 분포 가정)

변수	계수	z값
출산자녀 수 (종속변수): 0,1,2,...,7		
주택 소유 여부(소유=1)	0.052***	4.56
가구 총소득(원)	1.38e-08***	4.87
부인의 취업여부(취업=1)	-0.053***	-4.92
부인의 교육수준(학력년수)	-0.003	-1.16
부인의 현재 연령	0.022***	20.33
인공유산 횟수	0.027***	5.06
부인의 혼인연령	-0.031***	-16.02
부인 친정어머니 자녀양육 도움여부(도움=1)	0.099**	2.53
부인 시어머니 자녀양육 도움여부(도움=1)	0.136***	4.69
아들 선호 여부(선호=1)	0.076***	7.38
남편의 장남여부(장남=1)	0.031***	2.58
부인의 형제자매 수	0.015***	4.88
남편의 형제자매 수	0.009***	2.78
대도시 거주여부	-0.099***	-7.33
중소도시 거주여부	-0.081***	-5.54
상수항	0.397***	6.10
log likelihood	-11106.990	
표본수	6,453	

주: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

5) 가구소득변수를 제외하고 남편의 학력변수를 이용한 경우에는 남편의 학력이 출산자녀수에 유의한 양(+)의 영향을 미치는 것으로 나타났다. 따라서 가구소득변수를 사용한 후 남편학력변수의 유의성이 떨어지는 것은 일반적으로 남편의 학력이 높으면 가구소득이 높아지므로 두 변수간의 다중공선성(multi collinearity)의 문제인 것으로 보인다. 한편 남편의 연령은 출산자녀수에 큰 영향을 주지 않는 것으로 나타났다.

&lt;표 3&gt; 출산자녀수 분산분포 검정식

파라미터	추정치	t 값
$a_1(\mu)$	-0.590	-69.010
$a_2(\mu^2)$	-0.313	-69.010

일반적으로 학력수준이 높을수록 기혼여성의 항상소득은 증가할 것으로 보이므로 학력에 대한 추정치를 기혼여성의 임금에 대한 추정치의 대리변수로 해석한다. 여전히 가구소득도 한 시점에 관측된 소득이라는 문제점이 존재하나 적절한 대리변수가 없고 출산에 대한 소득효과를 포착하기 위해 서베이 자료의 소득을 이용하였다.<sup>6)</sup>

자녀수와 관련한 축약형 모형을 구성하여 포아송 분포를 따른다는 가정 하에 추정한 결과를 보면 <표 2>와 같다.

위의 출생아 수에 대한 추정치를 구해 위의 검정식에 대입하여  $a$ 값을 확인하면 <표 3>과 같다.

따라서 출산자녀 수는 분산이 평균보다 작은 분포를 보이고 있음을 확인할 수 있다. 이는 위의 출산자녀 수 중 두 자녀를 가진 가구가 많고 포아송 분포의 경우보다 자녀가 없거나 자녀가 들인 가구가 많다는 관측치의 분포로부터 도출했던 가설과도 일치한다.

이미 가입기를 지난 기혼여성들만을 대상으로 출산자녀수에 대한 회귀분석을 통해 추정치를 구한 후 위의 분산의 크기에 대한 검정식에 대입해도 동일한 결과를 얻는다. 참고로 40세 이상의 기혼여성의 출산자녀 수 분포에 대한 분산을 검정한 결과  $a_1$ 은 -0.726,  $a_2$ 는 -0.340의 유의한 음수 값을 보이고 있어 분산이 작음을 확인할 수 있고 이에 따라 포아송 분포나 음이항 분포보다는 감마 count 모형을 이용하는 것이 적절한 것으로 보인다. 감마 count 분포를 따르는 가정 하에 출산자녀 수의 결정요인의 계수들을 구하면 다음 <표 4>와 같다.

위의 추정계수를 비교해보면 두 모형은 유사한 결과를 보여주고 있다.<sup>7)</sup> 소득

6) 가구소득에 대한 대리변수로 거시통계에서 추출가능한 남성의 연령별·학력별 소득을 이해할 수는 있다. 그러나 이로부터 항상소득을 추출하는 문제는 여전히 존재한다. 여성과 달리 남성은 대부분 취업을 하고 따라서 실현된 소득은 자신의 기회비용과 큰 차이가 없다. 그러므로 40대 이후 남성의 소득에 기초한 가구소득은 가구의 경제적 능력을 나타내는 변수로서 기능할 수 있다. 반면 여성의 임금은 실제 취업을 할 경우에만 관측가능하고 경력단절 등이 있었을 때에는 자신의 학력이나 연령에 따른 기회비용과는 차이가 있으므로 기혼여성의 출산에 대한 기회비용으로 이용하기에는 다소 부적합하다.

7) 출산자녀수 분포가 감마 count 분포를 따르므로 포아송 분포를 이용하는 것은 부적절하다. 그러나 자녀수 결정요인의 설명변수들이 자녀수에 미치는 영향은 거의 유사하여 포아송 분포를 가정하고

<표 4> 출산자녀수 결정요인 분석 (감마 count 분포가정)

변수	계수	z값
출산자녀 수 (종속변수) : 0,1,2,...,7		
주택 소유 여부(소유=1)	0.056***	4.57
가구 총소득(원)	1.18e-08***	3.27
부인의 취업여부(취업=1)	-0.090***	-7.01
부인의 교육수준(학력년수)	0.001	0.43
부인의 현재 연령	0.029***	20.62
인공유산 횟수	0.038***	6.65
부인의 혼인연령	-0.035***	-15.68
부인 친정어머니 자녀양육 도움여부(도움=1)	0.148***	3.71
부인 시어머니 자녀양육 도움여부(도움=1)	0.171***	5.80
아들선호 여부(선호=1)	0.084***	7.19
남편의 장남여부(장남=1)	0.025*	1.91
부인의 형제자매 수	0.021***	5.89
남편의 형제자매 수	0.012***	3.42
대도시 거주여부	-0.109***	-7.33
중소도시 거주여부	-0.090***	-5.79
상수항	0.151**	2.01
log likelihood	-12773.680	
표본수	6,453	

주: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

이나 주택소유 여부 등 경제적 여건과 관련된 변수는 주택을 소유할수록, 가구의 총소득이 높을수록 자녀수가 많은 것으로 나타났다. 부인이 직업이 있는 경우에 자녀수가 적은 것도 유사하며 학력수준의 영향력은 포아송 분포를 가정할 경우에는 유의성이 낮지만 음(-)의 부호를 나타낸 반면, 감마 count 분포를 가정할 경우에는 출산자녀수에 거의 영향이 없는 것으로 나타나고 있다.

물리적인 가임력과 관련된 변수로 이용한 인공유산 횟수를 살펴보면 인공유산 횟수가 많을수록 자녀수가 많은 것으로 나타나 임신이 많이 할수록 출산도 많이 했음을 확인할 수 있다. 따라서 물리적인 가임력을 나타내는 대리변수로 는 인공유산횟수가 이용될 수 있을 것으로 보인다.

아들이 꼭 필요하다거나 아들이 있었으면 한다고 대답하여 남아선호 사상을 가진 가구로 분류된 경우에는 더 많은 자녀를 출산하는 것으로 나타났다. 이는 아들을 낳게 되면 추가적인 출산을 포기하는 반면, 딸을 낳은 경우에는 아들을 낳기 위해 추가자녀를 출산하기 때문인 것으로 보인다.

남편이나 부인의 가치관과 연결된 변수로 남편의 형제자매의 수, 부인의 형

회귀분석을 시도할 때 해석상의 오류를 크게 범하지는 않는다.

제자매의 수, 현재 거주하고 있는 지역이 대도시인지 중소도시인지, 시골인지를 더미 변수로 살펴보았다. 일반적으로 예측한 바와 같이 남편과 부인 모두 형제 자매가 많을수록 자신도 더 많은 자녀를 출산하는 것으로 나타났으며, 시골보다 대도시나 중소도시에 거주하는 경우에 자녀수가 적은 것으로 나타났다.

마지막으로 친정어머니나 시어머니의 육아도움이 있는 경우 출산자녀수가 많은 것을 확인할 수 있다. 그런데 이는 자녀수가 많아지면 부모님의 도움을 적극적으로 받으려는 동기가 발생하므로 자녀가 많은 가구에서 부모님의 적극적 도움을 요구하고 이로 인해 부모님의 육아도움이 발생할 수 있으므로 인과관계를 확인하기는 다소 어렵다.

위의 두 분포를 이용한 추정결과는 모든 연령대(18세 이상)의 기혼여성을 포괄하고 있어 출산자녀 수 자체가 완결출산율을 정확하게 나타내고 있다고 보기 어렵다. 따라서 흔히 가임기가 완료된 여성의 출산행태를 살펴보기 위해 40세 이상의 기혼여성 각각에 대해 출산자녀 수 결정에 대한 분석을 수행하였다(〈표 5〉참조).<sup>8)</sup>

40세 이상의 이미 출산을 거의 완료했다고 보이는 시점의 여성들은 주택소유 여부에는 큰 영향을 받지 않고 자녀를 둔 것으로 보이며 흔히 가치관과 관련된 변수들인 아들선호 여부나 남편이 장남인지 여부, 거주지역이 시골인 경우에는 두 분포 모두에서 출산자녀수에 강한 영향을 미친 것으로 보인다.

두 모형 모두에서 기혼여성의 취업여부는 자녀수에 유의하게 영향을 주지 못한 것으로 나타났다. 현재 직장을 다닌다는 것이 과연 출산시기에도 지속적으로 직장을 다니고 있었다는 것인지 아니면 자녀출산 및 육아 후 직장을 구한 것인지 확인할 수 없기 때문에 조사시점에 직업을 가진 것이 출산자녀수에 영향을 미치지 못한다고 해서 직장과 육아를 병행하는 것이 어렵지 않다는 것을 의미하지는 않는다.

인공유산횟수 변수는 감마 count 모형에서는 여성의 가임력을 나타내는 변수로서 양(+)의 유의한 값을 보이고 있으나 포아송 분포를 가정할 경우에는 유의성이 다소 떨어지는 것으로 나타났다.

부인의 혼인연령은 출산자녀수에 강한 음(-)의 관계를 보이고 있다. 이는 주로 결혼연령이 늦어질수록 가임기간이 짧아지기 때문인 것으로 보인다. 즉 늦은 결혼은 단지 자녀의 출산시기를 늦추어 합계출산율에 영향을 줄 뿐 아니라 가임기 전체에 걸쳐 출산하는 자녀수를 감소시킴으로써 완결출산율을 감소시킬

8) 45세 이상의 기혼여성만을 대상으로 회귀분석을 수행한 결과는 40세 이상의 기혼여성을 대상으로 한 경우와 매우 유사하게 나타나 결과보고를 생략하였다.

<표 5> 만 40세 이상 기혼여성의 출산자녀 수 결정모형

출산자녀 수 (종속변수)	만 40세 이상			
	포아송		감마	
	계수	z값	계수	z값
주택 소유 여부(소유=1)	0.022	1.39	0.023	1.46
가구 총소득(원)	2.25e-08***	5.69	2.33e-08***	5.78
부인의 취업여부(취업=1)	-0.007	-0.48	0.004	0.28
부인의 교육수준(학력년수)	-0.018***	-5.85	-0.017***	-5.66
부인의 현재 연령	0.0005	0.18	0.003	0.10
인공유산 횟수	0.011*	1.73	0.015**	2.21
부인의 혼인연령	-0.022***	-9.31	-0.022***	-8.27
부인 친정어머니 자녀양육 도움여부(도움=1)	-0.030	-0.49	-0.045	-0.63
부인 시어머니 자녀양육 도움여부(도움=1)	0.145*	1.86	0.1146*	1.71
아들 선호 여부(선호=1)	0.067***	4.85	0.070***	5.03
남편의 장남여부(장남=1)	0.056***	3.52	0.054***	3.47
부인의 형제자매 수	-0.001	-0.29	-0.00005	-0.01
남편의 형제자매 수	-0.002	-0.43	-0.001	-0.35
대도시 거주여부	-0.095***	-5.16	-0.095***	-5.01
중소도시 거주여부	-0.078***	-3.77	-0.080***	-3.78
상수항	1.389***	9.92	1.361***	8.87
Log Likelihood	-4997.180		-6099.350	
표본 수	2,892		2,892	

주: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

수 있음을 시사한다.

40세 이상 기혼여성의 출산자녀수는 완결출산율을 나타내는 것으로 위 모형으로부터 확인한 것처럼 가구소득 변수와는 강한 양(+의) 관계를 나타내고 있으나 주택의 소유여부는 큰 영향을 주지는 않는 것으로 보인다. 이는 가구소득의 크기와 주택소유여부간의 상관관계가 커서 나타나는 현상으로 볼 수 있다. 가구소득의 영향력이 큰 것은 가구의 경제력이 자녀수에 양(+의) 영향을 준다는 의미로 해석할 수 있다.

그 외 부인이나 남편의 형제 자매수는 큰 영향을 미치지 않는 것으로 나타나고 있다. 이는 전체표본의 경우 형제자매수가 일종의 가치관 대리변수로 출산자녀수에 양(+의) 영향을 주었던 점과 다르다.

### Ⅲ. 변형된 자녀수 결정모형

Ⅱ장에서는 출산자녀수가 감마 count 분포를 따른다는 가설검정에 기초하여 40세 이상 기혼여성의 자녀수 결정모형을 분석하였다. Ⅲ장에서는 자녀수 결정모형을 보다 심화하여 40세 미만 여성의 (예상)완결출산을 분석과 자녀가 0인 경우를 불임의 경우와 불임은 아니나 자녀를 출산하지 않은 경우로 나누어 자녀수만이 아니라 자녀 출산여부에 영향을 미치는 요인도 검토해본다. 나아가 출산과 여성의 경제활동간의 상관성<sup>9)</sup>을 고려한 모형도 분석하여 자녀수 결정모형을 다각적인 시각에서 검토한다.

#### 1. 불완결 출산율(incompleted fertility rates) 대 완결출산율

일반적으로 자녀수 결정모형은 가임기간이 끝난 여성만을 분석대상으로 한다. 그러나 젊은 세대의 출산력 패턴을 함께 분석하는 것은 향후 출산율이나 세대별 출산율의 차이와 이에 영향을 주는 변수들을 살펴본다는 의미에서 일정한 정보를 제공하므로 가임기간이 끝나지 않은 연령대의 기혼여성들도 포함한 분석을 시도해보는 것은 의의가 있다.

출산을 완결하지 않은 세대의 출산자녀수를 결정하는 요인을 분석할 때 사용하는 방식 중의 하나는 결혼 이후 혹은 가임기간 시작이후로 임신에 노출된 시기를 통제하는 방법이다. Winkleman and Zimmermann(1994)는 여성개인이 임신에 노출된 시간에 자연대수를 취한 값을 1로 정하여 각 개인들이 임신에 노출된 시간의 차이를 통제함으로써 센서링으로 인한 문제를 극복하려 시도하였다. 본 연구에서는 위의 방법론을 이용하면서 동시에 분석자료 상 기혼여성의 연령 외에 결혼 만연령이 제시되므로 연령과 결혼연령을 설명변수에 추가함으로써 임신에 노출된 기간의 차이를 통제하려고 시도하였다.

또한 기존연구와는 달리 완결출산율을 나타내는 가임기 전체를 통해 출산한 자녀수를 대신하는 변수로 예상 출산자녀수를 새롭게 정의하였다. 이미 언급하였듯이 예상 출산자녀 수는 기 출산자녀 수에 현재 임신하고 있으며 출산할 예정인 자녀를 더하고 향후 추가할 예정인 자녀수를 더하여 산출되었다.<sup>10)</sup> 40세

9) 전통적으로 여성의 출산율과 경제활동참여 간에는 음(-)의 상관관계가 있는 것으로 알려져 왔다. 그러나 최근 들어 북유럽 국가들에서는 출산율도 상당 수준을 유지하고 여성의 경제활동참여율도 높은 양(+)의 상관관계가 나타나고 있어 주목을 받고 있다. 본 논문에서도 두 변수간의 내생성을 고려한 모형을 추가하여 경제적 함의를 밝힌다.

이상의 여성의 경우에는 출산 자녀수와 예상 출산자녀수의 차이가 거의 없으나 전체 표본을 대상으로 할 경우, 각각의 평균은 1.885(표준편차 0.858), 2.074(표준편차 0.737)명으로 약 0.2명의 차이를 보이고 있다. 향후 추가할 예정인 자녀수가 1명 이상이라고 해도 물리적인 이유나 다른 경제적, 사회적 여건에 의해 자녀를 추가 출산하지 않을 수도 있으므로 본 연구에서 정의한 예상 출산자녀수는 전체 표본의 기혼여성을 대상으로 한 완결출산율의 최대치(upper bound)로 볼 수 있다. 한편 현 출산자녀 수는 완결출산율의 최소치(lower bound)로 해석해도 무방한 것으로 보인다.<sup>11)</sup>

위의 분산의 크기에 대한 검정식을 통해 예상출산자녀 수와 임신에 노출된 기간이 일정하다는 가정을 이용한 모형의 분산을 검사한 결과 분산이 작음을 확인하였다. 따라서 이 결과에 따라 전체표본을 이용하여 출산자녀수를 종속변수로 하고 임신에 노출된 기간의 로그값을 1로 통제하여 분석한 모형과 예상출생아 수를 결정요인을 각각 감마 count 분포를 따른다는 가정 하에 구해보면 다음 <표 6>과 같다.

예상출산자녀수는 만40세 이상의 기혼여성의 경우에 출산자녀수와 거의 동일하므로 만40세 이상의 기혼여성을 대상으로 완결출산율을 추정할 때 큰 의미가 없다. 대신 아직 출산이 완결되지 않은 젊은 세대인 18~39세의 기혼여성의 완결출산자녀수에 대한 대리변수로 예상출산자녀수는 큰 의미가 있으며 따라서 젊은 세대의 기혼여성을 대상으로 예상출산자녀수를 이용하여 회귀분석을 해 볼 필요가 있다.

아직 가임기간이 끝나지 않은 18-39세의 기혼여성을 대상으로 한 분석으로부터 40세 이상 기혼여성의 경우와 큰 차이점을 발견할 수 있다. 우선 가구의 총소득이 자녀수에 영향을 미치지 않는다.<sup>12)</sup> 주택소유여부는 양(+)의 영향을 주는 것으로 보이나 예상출산자녀수 모형에서는 유의성이 떨어진다. 부인의 취업여부는 두 모형 모두에서 강한 음(-)의 관계를 보이고 있으며 기혼여성의 교육

10) 추가예정 자녀수를 포함하여 완결출산율에 대한 적정 대리변수로 사용하는 것은 큰 문제가 되지 않는다. 이는 기혼여성이나 가구의 희망사항을 나타내기 보다는 현실적인 여건 등을 감안해 보고 된 것으로 실현될 자녀수에 가깝다고 볼 수 있다. 참고로 서베이 다른 문항인 이상적인 자녀수의 평균은 2.408 (표준편차 1.115)명으로 나타나 현재 출산 자녀수나 본 연구에서 이용한 예상 출산자녀 수의 크기보다 유의적으로 크다.

11) 전체 여성을 대상으로 한 완결출산율의 대리변수로 현 출산자녀 수 + 현재 임신 중이며 출산할 의도가 있는 자녀수를 사용할 수 있다 이를 이용하여 회귀분석을 하면 현 출산자녀 수를 이용한 경우와 거의 차이가 없는 것으로 나타났다.

12) 가구소득은 조사시점의 총소득이 아니므로 해석상 주의할 필요가 있다. 그러나 가구소득 대신 남편의 학력을 이용해본 결과, 그 영향력이 없는 것으로 나타나 가구소득이 자녀수에 양(+)의 영향을 미치는 40세 이상 표본에 비해 소득효과가 낮아짐을 확인할 수 있다.

<표 6> 가임기간 통제분석과 예상출산자녀 수 분석 비교(18-39세 대상)

종속변수	가임기간 통제		예상완결출산율	
	감마 count (18-39세)			
	출산자녀수		예상출산자녀수	
	계수	z값	계수	z값
주택 소유 여부(소유=1)	0.037**	1.98	0.019*	1.71
가구 총소득(원)	6.09e-09	0.10	4.50e-09	1.08
부인의 취업여부(취업=1)	-0.155***	-7.39	-0.059***	-5.05
부인의 교육수준(학력년수)	-0.018***	-3.22	0.004	1.18
부인의 현재 연령	-0.074***	-23.06	0.008***	5.27
인공유산 횟수	0.031***	3.18	0.002	0.29
부인의 혼인연령	0.079***	15.18	-0.022***	-10.00
부인 친정어머니 자녀양육 도움여부(도움=1)	0.242***	4.51	0.056*	1.72
부인 시어머니 자녀양육 도움여부(도움=1)	0.187***	5.10	0.049*	1.87
아들선호 여부(선호=1)	0.105***	5.85	0.089***	8.07
남편의 장남여부(장남=1)	0.026	1.19	0.015	1.07
부인의 형제자매 수	0.016***	2.58	0.012***	3.20
남편의 형제자매 수	0.009	1.44	0.002	0.63
대도시 거주여부	-0.110***	-4.87	-0.079***	-5.68
중소도시 거주여부	-0.084***	-3.75	-0.059***	-4.07
상수항	-0.832***	-6.26	0.855***	12.12
log(가임기간)=1	1로 고정		-	
Log Likelihood	-6445.75		-6552.95	

주: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

수준이 출산에 미치는 영향은 다소 다르게 나타난다. 임신노출기간 통제모형에서는 음(-)의 영향을 주는 것으로 나타난 반면, 예상자녀수 모형에서는 유의성이 없는 것으로 나타나고 있다.

40세 이상 기혼여성의 완결출산율과 비교하여 18~39세 기혼여성의 경우는 현재 기혼여성의 취업여부가 자녀수나 예상자녀수에 상당한 음(-)의 영향을 미치는 것으로 나타나며, 가구의 소득관련 변수의 영향력은 다소 감소하는 것으로 나타나고 있다.

## 2. 자녀출산여부와 자녀수 모형의 분리 : 허들(hurdle)모형

출산자녀수에 대한 결정요인을 고려할 때 간과해서는 안 될 문제는 자녀수가 0명인 가구 중 자녀를 출산하고자 해도 불임으로 인해 자녀를 출산할 수 없는 가구가 포함되어 있다는 점이다. 즉 자녀수가 0인 경우에는 의도적으로 자녀를



낳지 않은 가구와 자녀를 낳고 싶으나 낳지 못하는 가구가 혼재했다는 점을 주목해야 한다. 이 경우 자녀 수 0,1,2,3...을 결정하는 것이 유사한 요인에 의해 설명될 수 있다고 보기는 어렵다.

2003년도 전국출산력자료에는 불임여부를 직접 묻는 항목이 없다. 따라서 출산 자녀수가 0인 가구 중 어느 정도가 불임가구인지 정확히 알기는 어렵다. 다만 출생아 수가 없고, 현재 임신하지 않았으나 추가 자녀를 희망하는 부인의 수를 구하고 이 중 피임을 하지 않으며, 결혼 후 3년 이상이 경과한 부부를 불임부부라고 가정하고 가구 수를 구해보니 57가구가 불임인 것으로 나타났다. 이를 총 출생아 수와 비교해 보면 출산 자녀가 없는 전체 가구 496가구 중 11.5%를 차지하여 그 비중이 낮고, 예상 출산자녀 수와 비교해보면 전체 117개의 무자녀 가구표본 중 48.7%를 차지하여 그 비중은 다소 높은 편이다.

Becker의 Quantity-Quality 모형은 자녀수를 결정하는 데 있어 가구의 소득이나 여성의 경제활동 참여 등과 같은 경제적 변수들을 설명요인으로 이용하고 있는데 반해 Schoen et al (1997)이 주장하듯이 기혼자들이 자녀를 낳을지 여부를 결정하는 것은 경제적인 요인들보다는 사회적인 관계나 요소들에 의해 더 큰 영향을 받을 수 있다는 점을 고려해 보는 것이 필요하다. 따라서 출산자녀 수를 설명하는 모형을 구성할 때 허들 모형을 고려하여 자녀수가 0일 확률이 전체 자녀 수의 평균과 독립적이라는 가정을 이용해 볼 필요가 있다.

Cameron and Trivedi(1998)의 정의처럼 허들 모형은 0을 생산하는 과정과 양수를 생산하는 과정이 동일한 과정이 아닌 경우를 살펴보는 수정된 count 모형이다. 자녀가 없는 경우와 자녀를 한명 이상 출산하는 경우를 이항분포의 확률에 따라 고려하고 만약 자녀를 낳게 될 경우에 0이 절단된 count 모형을 구축하는 것이 허들 모형의 핵심이다.

허들 모형을 간략하게 소개하면 다음과 같다.  $F_1(\beta_1)$ 이 허들을 만나는 확률 즉 자녀를 낳게 될 확률이라고 하자.  $\frac{f_2(y, \beta_2)}{F_2(\beta_2)}$ 는 자녀수가 양수인 각각

의 경우의 조건부 분포를 나타내며  $f_2$ 는  $\sum_{y \in I^+} f_2(y, \beta_2) = 1$ 을 만족한다. 여

기서  $y$ 는 1이상의 양의 정수를 나타내며  $F_2$ 는 조건부 분포의 서포트(support)에 대한  $f_2$ 의 합이다. 따라서 허들모형의 우도함수는 다음과 같이 표현된다.

$$L = \prod_{i \in \Omega_0} \{1 - F_1(\beta_1)\} \prod_{i \in \Omega_1} \frac{\{f_2(y, \beta_2) F_1(\beta_1)\}}{F_2(\beta_2)}$$

여기서  $\Omega_0 = \{i \mid y_i = 0\}$  이고  $\Omega_1 = \{i \mid y_i \neq 0\}$  이다. 양변에 로그 값을 취하여 로그우도함수를 구하면

$$\ln(L) = \sum_{i \in \Omega_0} \ln \{1 - F_1(\beta_1)\} + \sum_{i \in \Omega_1} \ln \{F_1(\beta_1)\} + \sum_{i \in \Omega_1} [\ln \{f_2(y, \beta_2)\} - \ln \{F_2(\beta_2)\}]$$

이다. 로그우도함수가 분리 가능하므로 허들모형은 이항분포모형과 절단된 count 모형의 양자를 각각 극대화하는 값을 찾음으로써 로그우도 함수값을 극대화하는 파라미터를 구할 수 있게 된다.

출산자녀수나 예상출산 자녀수의 분포를 볼 때 자녀수가 0인 경우가 오히려 포아송 분포나 감마 count 분포의 경우보다 작으므로 0이 많은 분포 (zero inflated distribution)로 보기는 어렵다. Vuong Test 결과 역시 0이 적은 편이어서 이를 달리 취급하는 분포를 이용할 필요는 없는 것으로 나타났다. 따라서 위의 감마 count 분포를 이용한 결과와 허들 모형의 결과는 큰 차이가 없을 것으로 보인다. 다만 몇 가지 흥미로운 점을 발견하여 허들 모형을 이용한 결과를 <표 7>에 제공하였다.<sup>13)</sup>

위의 허들 모형의 결과로부터 몇 가지 흥미로운 사실을 발견할 수 있다. 첫째, 가구의 총소득과 기혼여성의 취업여부는 예상출산자녀수에 강한 영향을 미치는 것으로 나타났다.

둘째, 기혼여성의 교육수준이 높을수록 자녀를 낳는 경향이 있다. 그러나 하나 이상의 자녀를 보유한 가구를 대상으로 할 경우, 교육수준이 높을수록 출산하는 자녀수가 적은 것으로 나타나고 있다. 이는 여성의 학력수준이 높을수록 자녀출산 자체는 선호하는 경향이 있지만 기혼여성의 교육수준이 높을수록 적은 수의 자녀를 선호하는 것으로 해석할 수 있다.

셋째, 남편이 남자형제 중 장남인지 여부는 자녀의 출산여부에는 영향력이 떨어지지만 (예상)출산자녀수에는 유의한 양(+)의 영향을 미치는 것으로 나타나고 있다.

마지막으로 거주하는 지역은 자녀의 존재에는 영향을 주지 않지만 시골에 거

13) 본 논문에서는 Stata의 cloglog 명령어와 troipois0 명령어를 이용하여 추정하였다.

<표 7> 예상출산자녀수에 관한 허들모형

변수	1단계		2단계	
	무자녀 대 자녀		자녀 수 (1,2,3...)	
	계수	z값	계수	z값
주택 소유 여부(소유=1)	0.096	1.65	0.019**	1.99
가구 총소득(원)	4.34e-08***	3.17	1.18e-08***	3.93
부인의 취업여부(취업=1)	-0.151**	-2.48	-0.026***	-2.89
부인의 교육수준(학력년수)	0.041***	3.14	-0.013***	-5.71
부인의 현재 연령	0.013**	2.21	0.004***	5.17
인공유산 횟수	0.166***	3.62	0.008	0.17
부인의 혼인연령	-0.054***	-7.49	-0.017***	-10.48
부인 친정어머니 자녀양육 도움여부(도움=1)	1.670***	29.02	0.029	0.83
부인 시어머니 자녀양육 도움여부(도움=1)	0.205	1.00	0.041	1.55
남아선호 여부 (선호=1)	0.242***	4.21	0.074***	8.28
남편의 장남여부(장남=1)	0.112	1.62	0.025**	2.32
부인의 형제자매 수	0.027	1.51	0.004	1.35
남편의 형제자매 수	0.024	1.34	-0.001	-0.53
대도시 거주여부	0.060	0.89	-0.105***	-8.95
중소도시 거주여부	0.024	1.34	-0.084***	-6.54
상수항	1.256***	4.74	0.694***	13.27

주: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

주할수록 더 많은 수의 자녀를 두는 경향이 있음을 확인할 수 있다.

실증분석을 위해 이용한 자료에 자녀가 없는 경우가 많지 않고 불임이라고 해석할 수 있는 가구의 비중도 적어 허들 모형의 결과를 확대해석하기는 어렵지만 자녀를 낳을 것인가 말 것인가를 결정하는 요인과 자녀를 낳을 경우 몇 명을 낳을 것인가를 결정하는 요인 사이에 다소 차이가 있음을 확인할 수 있다.<sup>14)</sup>

### 3. 기혼여성 경제활동참여의 내생성 문제<sup>15)</sup>

기혼여성이 경제활동에 참여하여 직업을 가질 것인가와 몇 명의 자녀를 출산할 것인가에 대한 결정은 동시에 이루어지는 결정이라고 볼 수 있다. 위에서

14) 위와 같은 결과는 예상출산자녀수 대신 출산자녀수를 설명변수로 하여 분석할 때도 거의 유사하다. 다만 예상출산자녀수를 이용할 경우 자녀가 없는 가구수가 줄어들고 그중 불임으로 추측되는 가구가 많아 허들모형을 이용하는 것에 더 적합한 것으로 보여 예상출산자녀수 모형의 결과를 게재하였다.

15) 음(-)이 아닌 정수로 이루어진 종속변수에 대한 회귀분석 시 내생성 있는 설명변수를 다루는 데에는 Endogenous Switch Model이 일반적으로 사용된다. 단 내생성을 고려한 Gamma 분포모형을 추정할 것이 아니므로 해석시 Poisson 분포가정의 경우 도출한 결과와 비교해야 하는 문제점이 있다. Endogenous Switch Model에 대한 자세한 내용은 Miranda(2004) 참조.

다른 모형들은 출산 자녀수와 여성의 경제활동참여 여부에 대한 의사결정의 내생성을 고려하지 않고, 여성의 경제활동 참여에 대한 의사결정이 출산자녀수에 영향을 미치는 인과관계를 설정하였다. 즉 인과관계(causality)를 볼 때 기혼여성이 직장을 다니기로 결정하면 이것이 출산자녀수에 영향을 미쳐 그 결과로 자녀수가 결정된다고 가정한 것이다. 그러나 두 변수사이에 내생성이 존재한다면 위의 모형으로부터의 추정결과는 단지 많은 자녀를 출산하는 것과 직장생활을 지속하는 것을 양립하기는 어렵다는 해석수준 이상을 넘기 어렵다.

따라서 기혼여성의 경제활동을 출산자녀수 결정에 있어 내생적인 설명변수로 보고 기혼여성의 경제활동을 다른 독립적인 설명변수로 추정하여 내생성을 고려한 내생적 변환 포아송 (Endogenous Switch Poisson) 회귀분석을 시도하였다.

내생적 변환 포아송 모형은 포아송 분포를 따르는 피설명변수에 대한 회귀식에서 내생적인 더미 설명변수가 존재할 때 연립방정식의 편의를 교정하기 위해 사용되는 방법이다.  $x_i$  변수들과 더미 변수  $d_i$ , 교란항  $\epsilon_i$ 를 조건부로 하는 count 자료  $y_i$ 의 분포가 포아송 분포를 가진 경우를 고려해본다.

$$f(y_i | \epsilon_i) = \frac{\exp[-\exp(x_i'\beta + \epsilon_i)][\exp(x_i'\beta + \epsilon_i)]^{y_i}}{y_i!}$$

주어진 변수인  $z_i$  하에 더미변수  $d_i$ 는 다음과 같다.

$$d_i = \begin{cases} 1 & \text{if } z_i'\alpha + \nu_i > 0 \\ 0 & \text{otherwise} \end{cases}$$

여기서  $\epsilon_i$ 와  $\nu_i$ 는 평균이 0이고 공분산 매트릭스가  $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ 과 같은 결합정규분포를 따른다고 가정한다. 교란항  $\epsilon_i$ 에 조건부로 더미변수와 피설명변수는 독립이 되므로 결합밀도함수는 다음과 같다.

$$f(y_i, d_i) = \int_{-\infty}^{\infty} \{df(y_i | d_i = 1, \epsilon_i)Pr(d_i = 1 | \epsilon_i) + (1 - d)f(y_i | d_i = 0, \epsilon_i)Pr(d_i = 0 | \epsilon_i)\} f(\epsilon_i) d\epsilon_i$$

$w_i$ 를 모든 외생변수라 칭하고  $\eta_i = \frac{\epsilon_i}{\sigma\sqrt{2}}$ 의 변수치환법을 이용하면 피셀 명변수와 더미변수의 결합분포는 다음과 같다.

$$f(y_i, d_i | w_i) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \{f(y_i | d_i, w_i, \sigma\eta_i \sqrt{2}) \{d_i \Phi^*(\sigma\eta_i \sqrt{2}) + (1 - d_i) \Phi^*(-\sigma\eta_i \sqrt{2})\}\} \exp(-\eta^2) d\eta_i$$

위와 같은 결합분포 함수의 로그우도함수를 최우추정법을 통해 구하면 추정치를 구할 수 있다. 분석결과 만40세 이상의 기혼여성을 대상으로 할 경우 포아송

<표 8> 출산자녀수에 관한 내생적 Switch 모형(40세이상)

변수	Poisson 모형		Switch 모형	
	출산자녀수		부인의 취업	
	계수	z값	계수	z값
주택 소유 여부(소유=1)	0.027	0.91	-	-
가구 총소득(원)	1.81e-08**	2.22	-	-
남편의 근로소득	-	-	-3.80e-08***	-3.27
부인 가구주 여부(가구주=1)	-	-	0.540***	4.36
부인의 취업여부(취업=1)	0.008	0.23	-	-
부인의 교육수준(학력년수)	-0.017***	-3.18	-0.051***	-5.58
부인의 현재 연령	0.0006	0.14	-0.016*	-1.86
인공유산 횟수	0.009	0.86	-	-
부인의 혼인연령	-0.024***	-6.77	-	-
부인 친정어머니 자녀양육 도움여부(도움=1)	-0.083	-0.18	5.108***	10.97
부인 시어머니 자녀양육 도움여부(도움=1)	0.165	1.12	0.308	1.04
남아선호 여부(선호=1)	0.075***	2.80	-	-
남편의 장남여부(장남=1)	0.049	1.68	-	-
부인의 형제자매 수	0.001	0.09	-	-
남편의 형제자매 수	-0.001	-0.10	-	-
대도시 거주여부	-0.083**	-2.50	-0.589***	-8.47
중소도시 거주여부	-0.072**	-2.03	-0.468***	-5.85
상수항	1.395***	5.92	2.117***	4.97
$\sigma$		0.0020(0.13)		
$\rho$		0.931***(38.23)		

주: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

모형의 교란항과 내생적 변환 포아송 모형 교란항의 공분산 중  $\rho$ 값이 유의하여 두 모형의 교란항이 독립적이지 않음을 확인할 수 있어 기혼여성의 경제활동 참여가 내생성이 있는 설명변수임을 확인하였다. 40세 이상을 대상으로 출산자녀수에 대한 내생적 변환 포아송 회귀분석의 결과는 <표 8>과 같다.

$\rho$ 값의 결과에 근거할 때 기혼여성의 취업변수는 출산자녀수 결정에 있어 내생적인 변수라고 볼 수 있으나 추정결과는 기혼여성 취업변수의 내생성을 고려하지 않은 경우와 질적으로 큰 차이는 없다.

기혼여성의 취업여부를 결정하는 요인으로는 남편의 근로소득이 강한 음(-)의 영향을 주는 것으로 나타나고 있고, 교육수준이 높을수록 취업을 하지 않는다는 것을 확인할 수 있다. 이는 특히 40세 이상의 기혼여성 그룹에서 나타나는 현상으로 학력이 낮은 그룹에서 파트타임 직종 등에 취업하고 있음을 확인할 수 있다.

<표 9> 출산자녀수에 관한 내생적 Switch 모형 (18~39세)

변수	Poisson 모형		Switch 모형	
	예상 출산자녀수		부인의 취업	
	계수	z값	계수	z값
주택 소유 여부(소유=1)	0.017	0.72	-	-
가구 총소득(원)	8.92e-09	1.04	-	-
남편의 근로소득	-	-	-2.85e-08**	-2.52
부인 가구주 여부 (가구주=1)	-	-	1.234***	5.77
부인의 취업여부(취업=1)	0.0555**	-2.00	-	-
부인의 교육수준(학력년수)	0.0013	0.21	0.0090	0.87
부인의 현재 연령	0.0074**	2.29	0.0402**	7.90
인공유산 횟수	0.0028	0.20	-	-
부인의 혼인연령	-0.022***	-5.28	-	-
부인친정어머니 자녀양육 도움여부(도움=1)	0.052	0.63	0.641***	5.71
부인 시어머니 자녀양육 도움여부(도움=1)	0.039	0.65	0.656***	4.36
남아 선호 여부 (선호=1)	0.089***	3.81	-	-
남편의 장남여부(장남=1)	0.014	0.51	-	-
부인의 형제자매 수	0.0104	1.39	-	-
남편의 형제자매 수	0.0028	0.35	-	-
대도시 거주여부	-0.077**	-2.57	-0.374***	-6.93
중소도시 거주여부	-0.058*	-1.91	-0.224***	-4.02
상수항	0.888***	6.42	-1.322***	-6.01
$\sigma$	0.00069(0.07)			
$\rho$	0.482(0.10)			

주: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

한편 만39세 이하의 기혼여성의 예상출산자녀수를 내생적 변환 포아송 분포를 가정하여 회귀분석한 결과는 <표 9>와 같다.

만40세 이상의 경우와 달리 내생성은 크지 않은 것으로 보인다. 여성의 학력 수준이 취업에 거의 영향을 미치지 않는다는 결과와 친정어머니나 시어머니의 도움이 있는 경우 취업한 경우가 많다는 점 외에 만40세 이상과 비교하여 추정 결과가 큰 차이를 보이지 않는다. 다만 내생성이 크지 않다는 점을 고려할 때 18~39세 여성을 대상으로 할 경우 어머니의 취업여부를 외생적인 설명변수로 사용해도 예상출산자녀수를 과소추정하지는 않는다.

#### IV. 요약 및 결론

##### 1. 실증분석 결과요약 및 함의

다양한 방식에 의해 추정된 출산자녀수 결정모형의 분석결과를 요약하면 <표 10>과 같다.

1. 40세 혹은 45세 이상의 기혼여성만을 대상으로 완결출산율을 분석한 결과

<표 10> 출산자녀수에 관한 모형결과 비교

설명변수	만40세 이상기혼여성 출산자녀수	18~39세 기혼여성 예상출산자녀수	자녀출산여부
가구경제변수 :			
가구소득	양(+)	비유의적	양(+)
주택자산소유여부	비유의적	비유의적	비유의적
기혼여성 일반특징:			
연령	음(-) 혹은 비유의적	양(+)	양(+)
결혼연령	음(-)	음(-)	음(-)
친정어머니 탁아도움	비유의적	비유의적	양(+)
기혼여성 경제변수 :			
취업유무	비유의적	음(-)	음(-)
학력수준	음(-)	음(-) 혹은 비유의적	양(+)
가치관 변수:			
아들선호사상	양(+)	양(+)	양(+)
남편장남여부	양(+)	비유의적	비유의적
거주지역	시골지역 양(+)	시골지역 양(+)	비유의적
부인이나 남편의 형제자매 수	비유의적	일부 유의적	비유의적

가구의 소득은 출산자녀수에 양(+)<sup>1)</sup>의 영향을 미치며 기혼여성의 학력수준은 음(-)<sup>2)</sup>의 영향을 미친다. 그러나 기혼여성이 경제활동에 참여하는 지 여부는 출산자녀수와 큰 관계가 없는 것으로 나타나고 있다. 기혼여성의 경제활동과 예상출산자녀 수의 관계는 40세 미만의 여성을 대상으로 할 경우에는 유의한 음(-)<sup>3)</sup>의 상관관계를 가진 것으로 나타나고 있어 대조적이다.

40대 이상의 기혼여성의 경우에는 학력이 높을수록 오히려 경제활동참가율은 낮고 출산자녀 수는 적으므로 경제활동참가와 출산자녀 수 사이에 뚜렷한 관계를 발견하지 못했다. 그러나 30대 이하의 여성들의 경우에는 학력과 관계없이 경제활동참가가 이루어지고 있으며 경제활동참여를 하는 경우에는 출산하는 자녀수가 적은 것으로 나타났다. 이러한 경향은 가임기간이 완료되는 40세 이후에 완결된 출산자녀수를 추정할 때에도 유사한 경향을 나타낼 것으로 보인다.

2. 18~39세의 기혼여성을 대상으로 할 경우, 가구의 소득이 예상출산자녀수에 미치는 영향이 유의적이지 않은 것으로 나타나 이 또한 40세 이상의 기혼여성을 대상으로 한 분석결과와 대조적이다. 위와 같은 결과는 남편의 소득에 의존하여 경제활동 참여여부를 결정하던 40대 이상의 세대들보다는 남편 소득이 높다고 해도 경제활동에 참여하는 여성이 많아지면서 나타난 현상이다. 단 가구의 소득은 조사시점에 보고한 소득이므로 이를 생애소득으로 해석하기는 어렵고 아직 출산이 완결되지 않은 시점에서 나타나는 현상이므로 위와 같은 출산행태가 출산완결시점에도 동일하게 나타날지 여부는 조심스럽게 해석할 필요가 있다.

남편의 학력수준이 항상소득에 대한 대리변수라고 간주하고 가구소득대신 이용하여 예상출산자녀수에 대해 회귀분석을 시도한 결과 역시 예상출산자녀수에 큰 영향이 없어 자녀수에 대한 소득효과가 기혼여성 연령층이 낮아지면서 점점 사라져가는 것으로 보인다.

3. 허들모형을 이용할 경우 자녀를 출산할지 여부를 결정하는데 있어 기혼여성의 학력은 양(+)<sup>4)</sup>의 관계를 보이고 있어 기혼여성의 학력수준이 높을수록 자녀를 출산하는 경우가 많은 것으로 나타나고 있다. 한편 자녀를 낳은 경우에 자녀수에 대해서는 기혼여성의 학력이 높을수록 음(-)<sup>5)</sup>의 영향을 나타내는 점을 고려할 때 기혼여성의 학력은 자녀출산 여부에는 오히려 긍정적으로 작용할 수 있으나 자녀수 결정에는 부정적으로 작용하여 출산율을 감소시키게 된다.

여성의 학력이 시장임금에 대한 대리변수라고 본다면 여성의 시장임금의 증가가 적어도 두 번째 자녀이상의 출산에는 부정적으로 작용하여 완결출산율을 감소시킬 수 있다는 결론을 도출할 수 있다.



4. 기혼여성의 취업이 출산자녀수에 내생적인 더미변수라고 보고 이를 통제 한 결과, 40세 이상의 경우에는 내생성이 있고 18~39세 여성에게는 내생성이 크지 않은 것으로 나타났다. 전체표본을 고려할 경우 취업여부를 외생적인 더미변수로 보고 추정한 출산자녀수는 1.8919명으로 내생성을 고려한 경우의 출산자녀 수 추정치인 1.8922명보다 미세하나마 작아 내생성을 고려하지 않은 경우에는 출산자녀수를 과소추정하는 경향이 있음을 확인하였다. 이는 출산율이 기혼여성 노동공급에 내생적인 변수임을 간과한 채 출산율의 효과를 측정하면 출산율이 기혼여성 노동공급에 미치는 효과가 과대 추정된다는 사실과 일맥상 통한 것이다.

내생성이 있는 것으로 나타난 40세 이상 기혼여성의 경우에는 출산율이 노동공급에 미치는 영향이 과대추정될 가능성이 있으나 30대 이하의 경우에는 그 가능성이 적다고 볼 수 있다. 또한 30대 이하 기혼여성의 경우, 경제활동참여가 예상출산자녀수에 미치는 영향이 내생성을 고려하지 않아도 제대로 추정될 수 있음을 의미한다.

5. 40대 이상의 세대에서는 가구소득이 높을수록 출산자녀수가 커지는 자녀수에 대한 소득효과가 크게 나타나고 여성의 노동공급 참여로 인한 출산자녀수 감소효과는 작지만 30대 이하의 젊은 세대에서는 가구소득이 직접적으로 출산자녀수에 주는 영향은 작아져서 소득효과는 감소한다고 볼 수 있다. 30대 이하의 세대에서는 기혼여성 경제활동참여가 출산자녀수를 줄이는 역할을 하게 되므로 여성의 경제활동참여율이 높아질수록 출산율은 하락할 가능성이 크다. 이는 여성의 경제활동참여 욕구가 높아질수록 출산율이 더욱 감소할 수 있다는 가능성을 시사한다. 따라서 일하는 기혼여성의 출산부담을 경감할 수 있는 정부의 정책적 수단이 더욱 중요해질 수 있다.

본 논문에서 감마 count 모형 외에 다른 심화된 모형을 이용한 결과 발견하게 된 가장 중요한 점은 기혼여성이 자녀를 출산할 것인가 여부에 대한 결정은 자녀수 결정과 다소 다르다는 점이다. 특히 여성의 학력수준이 자녀수 자체에는 음(-)의 영향을 미치는 것과 달리 자녀출산여부에는 양(+)의 영향을 준다는 것이다.

이는 여성의 학력수준이 지속적으로 상승하는 가운데 출산율 제고를 위한 정부정책을 둘째아 이상의 출산장려에 집중하는 것이 필요하다는 점을 시사한다. 한 자녀 출산은 여전히 하나의 사회적 규범으로 자리 잡고 있으나 여성의 학력신장과 임금상승은 둘째아 이상의 출산을 저하시켜 결국은 출산율이 대체출산율 수준이하로 감소하는 주요한 요인으로 작용하고 있기 때문이다.

## 2. Becker 가설과 검증

본 연구결과가 Becker의 가설과 합치하는 지 여부를 가늠하기 위해 우선 Becker의 Quantity-Quality 모형의 핵심을 살펴보자. 가구의 효용함수  $U$ 는  $U = (n, w, y)$ 로 구성되고 여기서  $n$ 은 가구 내 자녀 수,  $w$ 는 자녀의 질적 수준(모든 자녀에 대해 동일하다고 가정),  $y$ 는 자녀를 제외한 기타 모든 재화를 나타낸다. 자녀의 질적 수준에 대한 생산함수  $w = e + q$ 로 표현할 수 있으며  $e$ 는 초기부존자원,  $q$ 는 자녀의 질적 수준에 대한 가구의 투자라고 정의한다.  $e$ 는 가구의 소득이 증가함에 따라 증가한다고 가정할 수 있다. 한편, 가구의 예산 제약식은  $p_y y + p_q nq = I$ 이며  $p_q$ 는 자녀의 질적 수준을 한 단위 높이는 데 필요한 평균비용,  $p_q nq$ 는 자녀에 대한 총 지출로 해석할 수 있다.

예산 제약하의 효용극대화 1계 조건을 통해  $\pi_w = np_q$ ,  $\pi_n = qp_q$ 가 도출되며  $\pi_w$ ,  $\pi_n$ 은 각각 자녀의 질적 수준에 대한 잠재가격, 자녀의 수에 대한 잠재가격을 의미한다. 위의 두 식을 통해 자녀의 수와 자녀의 질적 수준은 상호의존적임을 알 수 있다. 즉 자녀수가 증가하면 자녀의 질적 수준에 대한 잠재가격은 상승하며, 자녀의 질적 수준에 대한 가구의 투자가 증가하면 자녀의 수에 대한 잠재가격도 증가하게 된다. 소득이 증가하면 자녀 수에 대한 (관찰된) 소득탄력성은 낮은 소득에서는 감소하거나 음수이지만 소득이 증가하면서 증가한다. 반면 소득이 증가하면 자녀의 질적 수준에 대한 (관찰된) 소득 탄력성은 낮은 소득에서는 높으나 소득이 증가하면서 감소하게 된다.

자녀의 질에 대한 부존자원의 수준이 가구 소득에 크게 연동될수록 소득의 증가에 따라 자녀의 수는 증가하는 반면, 자녀의 질적 수준 증가 효과는 작다. 이는 자녀수와 사회계층의 이동성 간에 음(-)의 상관관계가 있음을 나타내는 것이다. 계층간 이동가능성이 크게 되면 소득이 증가함에 따라 자녀수의 감소 효과는 큰 반면, 자녀의 질적 수준을 유지하기 위한 가구의 지출은 증가하게 된다. 계층간 이동성이 작은 경우 가구소득이 높으면 이미 자녀의 질적 수준에 대한 부존자원이 높고 이것이 지속적으로 유지될 가능성이 크므로 자녀의 질적 수준보다는 자녀의 수를 유지하는데 가구의 예산제약을 사용할 수 있다. 따라서 소득수준이나 계층간 이동성 등에 따라 소득의 증가가 자녀수의 증가로 나타날 수도 있고, 감소로 나타날 수도 있다는 것이 Becker의 Quantity-Quality 모형의 핵심내용이다.

실증분석의 결과, Becker의 가설은 우리나라에서도 적용되는 것을 확인하였다. 만40세 이상의 기혼여성들에 있어서는 소득이 클수록 자녀수가 많아 소득 증가가 자녀수의 증가로 나타났다고 해석할 수 있고, 만39세 이하 기혼여성의 경우에는 소득증가가 자녀의 질적 수준에 대한 관심과 투자로 이어졌다고 해석할 수 있다. 즉 40세 이상의 기혼여성의 경우에는 소득효과가 대체효과보다 크고 39세 이하 기혼여성의 경우에는 소득효과의 상대적인 크기가 감소하고 있음을 의미한다.

여성의 경제활동 참여는 30대 이하 여성의 출산율에 강한 음(-)의 영향을 주는 것으로 나타났는데 이는 Becker의 가설과 일치하는 것이다. Becker 가설에 따르면 여성의 임금수준의 상승은 기타 재화의 잠재 가격을 감소시켜 다른 재화에 대한 지출을 증가시키고, 자녀 수와 자녀의 질적 수준에 대한 투자를 감소시킬 수 있는데 자녀수와 자녀의 질적 수준 중 자녀수에 대한 감소가 더 크다고 할 수 있다.

여성의 학력수준은 자녀의 질적 수준의 잠재 가격을 감소시켜 자녀의 질적 수준을 증가시킨다. 이는 결국 자녀수에 대한 잠재 가격을 상승시켜 자녀수를 감소시키게 된다는 것이 Becker의 가설이다. 실증분석의 결과, 감마 count 분포를 이용하여 기혼여성 전 세대를 대상으로 할 경우 여성의 학력수준이 자녀수에 음(-)의 관계가 있음을 확인하였다. 다만 아직 출산이 완결되지 않은 세대인 18~39세 여성을 대상으로 한 분석에서는 회귀분석 모형에 따라 여성의 학력수준이 자녀수에 미치는 영향의 유의성이 떨어지는 경우도 발생하였다.

### 3. 향후 연구과제

본 논문은 다양한 형태의 count 모형을 이용한 출산율에 대한 외국선행연구들을 체계적으로 종합하여 우리나라 자료에 응용한 국내 최초의 연구이다. 대부분의 선행연구가 포아송 분포나 감마분포, 허들 모형 등 하나의 모형에 기초하여 기혼여성 연령이나 학력 등 비교적 구하기 쉬운 변수가 자녀수 결정에 미치는 영향을 도출한 것과 달리, 본 논문에서는 여러 계량경제학적 기법을 통해 소득, 기혼여성의 학력, 경제활동참여 여부, 자녀에 대한 가치관 등 다양한 경제적·사회적 변수들이 여성의 생애전체에 걸쳐 출산하는 자녀수에 어떤 영향을 주는지를 다각적으로 검토하였다는 점에 의의가 있다. 특히 40세 미만의 아직 출산이 끝나지 않은 세대의 경우에도 예상출산자녀수 변수를 이용하여 출산 행태를 분석한 것은 최초의 시도이다. 나아가 다양한 count 모형 분석결과가

Becker가 제시한 가설과 크게 다르지 않음을 확인하여 출산을 결정모형의 기초 이론이 성립하는지 여부도 동시에 확인하였다.

그러나 여성의 출산행위를 보다 정확하게 포착하기 위해서는 완결출산자녀수만이 아니라 자녀터울 및 출산시기 결정에 영향을 주는 요소들을 파악할 필요가 있다. 특히 최근의 출산율 하락이 결혼 및 출산지연에 따른 것이라는 일부 연구결과를 고려할 때 정확모형을 넘어 동학모형으로 생애주기에 따른 기혼여성의 출산결정모형을 구축하는 연구를 통해 본 연구결과를 보완하는 작업이 추후 필요하다.

### 참고문헌

- Becker, Gary S. (1960) "An Economic Analysis of Fertility" in A. J. Coale, ed., *Demographic and Economic Changes in Developed Countries* Princeton University Press.
- Becker, Gary S. and H. G. Lewis (1973) "On the Interaction between the Quantity and Quality of Children" *Journal of Political Economy* 81(2).
- Becker, Gary S. & Tomes, Nigel (1976) "Child Endowments and the Quantity and Quality of Children" *Journal of Political Economy* 84(4): S143-62.
- Cameron C. & P. Trivedi (1998) *Regression Analysis of Count Data*, *Econometric Society Monographs* No. 30 Cambridge University Press.
- Caudill, S. B. and F. G. Mixon (1995) "Modeling Household Fertility Decisions: Estimation and Testing of Censored Regression Models for Count Data" *Empirical Economics* vol. 20: 183-196.
- Gensler, H. (1997) "Welfare and the Family Size Decision of Low-income, Two-parent Families" *Applied Economics Letters* vol. 4: 607-610.
- Heckman, J. and J. Walker (1991) "Economic Models of Fertility Dynamics: A Study of Swedish Fertility" in Schultz(ed), *Research in Population Economics* vol. 7.
- \_\_\_\_\_ (1990a) "Relationship between Wages and Income and the Timing and Spacing of Births: Evidence from Swedish Longitudinal Data" *Econometrica* vol. 58: 1411-1441.

- \_\_\_\_\_ (1990b) "The Third Birth in Sweden" *Journal of Population Economics* vol. 3: 234-275.
- Heckman, J. and R. Willis (1976) "Estimation of a Stochastic Model of Reproduction: An Econometric Approach" in Terleckyj(ed) *Household Production and Consumption* Columbia University Press, New York.
- Melkersson M. & D-Do (2000) "Rooth, Modeling Female Fertility using Inflated Count Data Models" *Journal of Population Economics* vol. 13.
- Miranda A. (2004) "FIML estimation of an endogenous switching model for count data" *Stata Journal* 4(1): 40-49.
- Nguyen-Dinh, H. (1997) "A Socioeconomic Analysis of the Determinants of Fertility: The Case of Vietnam" *Journal of Population Economics* vol. 10: 251-271.
- Silva, J. and F. Covas (2000) "A Modified Hurdle Model for Completed Fertility" *Journal of Population Economics* vol. 13: 173-188.
- Wang Famoye (1997) "Modeling Household Fertility Decisions with Generalized Poisson Regression" *Journal of Population Economics* vol. 10.
- Winkelmann R (1995) "Duration Dependence and Dispersion in Count Data Models" *Journal of Business and Economic Statistics* vol. 13.
- \_\_\_\_\_ (1996), "A Count Data Model for Gamma Waiting Times" *Statistical Papers* 37.
- \_\_\_\_\_ (1997) *Econometric Analysis of Count Data* 2nd revision, Springer Verlag, Berlin Heidelberg, New York.
- Winkelmann R & KF Zimmermann (1994) "Count Data Models for Demographic Data" *Mathematical Population Studies* 4: 205~221.

[2007. 9. 15 접수 | 2007. 11. 5 채택]

## The Economic and Social Implication of Count Regression Models for Married Women's Completed Fertility in Korea

*Hyun-Sook Kim*

This paper uses a static Gamma count model, a traditional hurdle model and an endogenous switching Poisson model, respectively for determining married women's completed fertility rates in Korea. This paper analyzes the impact of household income, women's wage and education, and women's job market participation on the number of children of married women above age 40 and on the expected number of children of women aged below 40.

The paper shows that a household income significantly increases the number of children for at least women aged above 40, however, this income effect is disappearing for younger generation. The empirical model suggests that women having a job tend to have fewer children for a group 39 years old and below and find that there is an endogeneity problem between child birth and labor force participation, too. The education level of married women gives a positive effect for giving a birth, itself, while it gives a negative impact on the number of children.

Based on the empirical results, it concludes that Becker's Quantity-Quality theory works for Korea, too.

**Key Words:** Becker's Quantity-Quality model, Count model, Gamma distribution, completed fertility rates, expected number of children, hurdle model, endogeneity