

기혼여성의 노동공급 결정요인: 패널데이터 분석

송 효 진*안 자 영**

여성의 노동 공급은 경제학, 여성학 등 다양한 분야에서 중요하게 다루어 진 이슈이다. 본 논문은 한국 기혼여성의 노동참여 결정과 노동공급곡선에 영향을 미치는 요소에 대하여 실증적으로 분석하고자 한다. 분석을 위해 2003년부터 2007년까지의 노동연구원의 한국노동패널자료를 이용하였다. 그간 기존의 여성 노동공급곡선 추정에 대한 연구에서 문제가 되어왔던 계량적인 문제를 해결하기 위해 표본 선택 편의(Sample Selection Bias), 시간당 임금 변수의 내생성(endogeneity), 경력 변수의 내생성의 문제를 모두 고려하였다. 기존의 연구가 횡단면분석 자료를 분석하거나, 여러 연도의 자료를 단순히 pooling하여 사용하였는데 반하여, 본 연구에서는 패널 데이터를 다루었다는 점이 다른 연구와 차별화 된 점이다. 또한 패널자료 방법론으로서 pooled probit, 도구변수 추정법 뿐만 아니라 다양한 panel data 방법론을 시도하였는데에도 의의가 있다.

I. 서론

1. 연구의 필요성

한국 여성의 경제활동 참가율은 상승추세를 보이고 있다. 남성과 비교하여 볼 때 남성의 경제활동 참가율 상승속도에 비하면 여성의 그것은 매우 빠른 속도로 진행되고 있음을 알 수 있다. 하지만 여전히 여성의 경제활동 참가율은 남성의 것보다 현저히 낮은 편이며, 선진국수준과 비교하였을 때도 우리나라 여성의 경제활동 참가율은 매우 낮은 편이라 할 수 있다.

아래 <표1>을 보면 2005년의 남성 경제활동 참가율은 74.4%로 1985년에 비하여 2.1%p 상승하는데 그친데 반하여, 여성의 경우 50.0%에서 41.9%로 8.1%p나 상승하였다. 또한 여성의 경제활동 참가율은 같은 유교문화권인 일본 등을 제외한 서구 선진국들과 비교하였을 때 2005년 기준으로 미국과는 9.3%p, 스웨덴과는 무려 10.6%p나 차이를 나타내고 있어 우리나라 여성의 경제활동 참가율이 저조하다는 것을 알 수 있다.

* 이화여자대학교 경제학과 석사과정

** 이화여자대학교 경제학과 석사과정

<표 1> 경제활동 참가율 국제비교

(단위:%)

	남성				여성			
	한국	일본	영국	미국	한국	일본	영국	미국
1985	72.3	77.9	76.1	76.3	41.9	47.6	50.3	54.5
1988	72.9	76.9	75.6	76.2	45.0	47.9	52.4	56.6
1989	73.4	76.8	75.8	76.4	46.6	48.5	53.4	57.4
1990	74.0	77.0	75.7	76.4*	47.0	49.1	53.7	57.5*
1991	75.0	77.4	75.1*	75.8	47.1	49.7	53.5	57.4
1992	75.7	77.8	73.7	75.8	47.1	49.9	53.4*	57.8
1993	75.7	77.8	72.7	75.4	47.1	49.6	53.4	57.9
1994	76.4	77.6	72.4	75.1*	47.8	49.5	53.5	58.8*
1995	76.4	77.5	72.1	75.0	48.4	49.3	53.5	58.9
1996	76.2	77.5	71.8	74.9	48.9	49.3	53.8	59.3
1997	76.1	77.6	71.6	75.0*	49.8	49.6	54.2	59.8*
1998	75.1	77.1	71.3	74.9*	47.1	49.4	54.4	59.8*
1999	74.4	76.7	71.5	74.7*	47.6	48.9	54.9	60.0*
2000	74.2	76.2	71.2	74.8*	48.6	48.6	55.2	59.9*
2001	74.2	75.5	70.9	74.4	49.2	48.5	55.1	59.8
2002	74.8	74.5	70.8	74.1	49.7	47.9	55.5	59.6
2003	74.6	73.9	70.8	73.5*	48.9	47.7	55.6	59.5
2004	74.8	73.2	70.5	73.3	49.8	47.6	55.9	59.2
2005	74.4	73.1	70.5	73.3	50.0	47.7	56.2	59.3

주: * 표시는 시계열이 절단됨을 의미함.

자료: 노동연구원, 해외노동통계 재인용

(BLS, Comparative Civilian Labor Force Statistics of Ten Countries(1960~2005), 2006.

또한 우리나라는 대표적인 저출산 고령화 국가로서 경제활동 참가인구의 장기적인 감소가 나타날 수 있다. 이러한 현상은 장기적으로 우리나라의 잠재성장력을 해칠 수 있으며, 성장 동력에 장애가 될 수 있다. 이런 문제에 대한 해결책으로서 육아나 보육으로 인해 노동시장에 참여하지 않는 여성 유희노동력을 최대한으로 사용할 수 있는 정책을 마련해야 한다.

여성의 노동공급에 있어서 가장 중요한 문제는 자녀보육으로 인한 휴직 또는 퇴직일 수 있다. 여성은 임신과 출산으로 인하여 30대를 전후하여 시장참여 경력이 단절되고, 자녀보육 및 보육관련 비용 등으로 인해 다시 노동시장에 참여하기가 어려운 실정이다. 국가경쟁력 강화와 유희노동력의 효과적인 활용을 위해서는 여성들이 일하기 좋은 환경을 구축함과 동시에 기혼여성들의 노동공급에 영향을 미치는 요인을 면밀히 분석하여 보육 및 노동정책에 활용토록 하여야 할 것이다. 이에 이 연구는 60세 미만의 기혼여성들이 노동시장에 참여를 결정하는 요인과 노동시간을 결정하는 요인을 계량적으로 분석하고자 한다.

2. 연구의 특징 및 의의

가. 특징

추정된 계수가 일치성(Consistency)을 갖도록 하는 것은 실증 분석의 목표 중 하나이다. 이를 위하여 본 연구에서 고려해야 할 사항은 크게 세 가지이다. 표본 선택편의(Sample Selection Bias), 시간당 임금 변수의 내생성(Endogeneity), 경력 변수의 내생성문제가 이에 해당한다.

먼저, 표본선택편의 문제를 고려하여야 한다. 우리가 관측할 수 있는 시장임금 자료는 이미 노동 시장에 참여하고 있는 사람들의 것이다. 이렇게 시장에서 노동공급을 하고 있는 여성만으로 추정된 노동공급곡선은 표본선택편의를 나타낼 수 있다. 다시 말해서 시장에서 일을 하는 사람에게서만 측정되는 시장임금을 고려하는 것은 일종의 Subsample만을 고려하게 된다는 것이다. 이때 일하는 여성만의 Subsample이 Random하게 뽑혀진 표본이냐는 것이 문제가 될 수 있다.

기존 연구들은 일하는 여성만으로 구성된 표본이 Random한 과정을 거쳐 선발된 것이 아닐 수 있다고 하였으며, 이를 무시할 경우 추정된 계수에 편의 및 불일치성이 발생할 수 있다는 것을 지적했다. Smith(1980)는 표본 선택 편의를 고려하지 않을 경우 임금탄력성이 약 5.3%~11%의 상향편의(Upward Bias)를 갖게 됨을 보였다.¹⁾ 이러한 표본 선택 편의를 고려하기 위하여 다양한 방법이 연구되고 있으며 모수 추정법(Heckman, Tobit)과 준모수 추정법이 그 예이다.

또한, 시간당 임금 변수의 내생성을 고려해야 한다. 시간당 임금 변수가 내생적이라는 사실은 많은 논문을 통하여 지적되어 왔다. 설명변수가 내생성을 갖게 되면 예측된 결과가 편의(Bias) 또는 불일치성(Inconsistency)이 발생하며 유의미하지 않은 결과를 만들 수 있다. 이를 고려하기 위하여 도구 변수 추정법이 쓰일 수 있다. 이때 도구변수로는 교육연수나 나이, 경력등이 이용될 수 있는데, 이때 경력변수에 대해서는 더 발전된 논의가 필요하다.

최근 연구에서는 시간당 임금 변수의 내생성의 교정을 위해 사용되는 경력변수 또한 내생적일 수 있다는 가능성을 가지고 있다고 지적한다. 기존의 연구에서는 임금의 내생성을 해결하기 위한 도구변수로 경력변수를 사용하고 있는데, 이 경우 많은 주의를 기울여야 한다는 것이다. 즉, 도구변수 추정법을 사용하기 위해서는 도구변수가 외생적이어야 한다는 가정이 필요한데 실제로 관찰되지 않는 변수와 경력 변수간의 상관관계가 존재할 수 있으므로 경력변수의 외생성이 성립하지 않을 수 있다. 본 연구에서는 이러한 점을 고려하기 위하여 임금변수의 내생성 교정을 위하여 경력 변수를 포함한 모형과 포함하지 않은 모형을 함께 추정하였다. 또한 이 둘이 유의한(Significant) 차이를 나타내는지를 테스트함에 따라 경력변수의 내생성을 검증해볼 수 있다.

또한 노동수요곡선과 노동공급곡선간의 연립방정식(Simultaneity)문제를 해결하기 위한 방안으로 고용선택에 있어서 회사 측은 고려하지만, 노동자는 고려하지 않는 변수로 기업의 성과지표인 ROE(Return on Equity)를 상정하여 산업별 ROE를 도구변수로 추가한 모형도 고려해 보았다..

1) 황윤재·최강식(1999) 재인용.

나. 의의

본 논문의 의의는 여성의 노동공급요인 결정분석에 있어서 본격적으로 Panel 형식의 Data를 활용하였고, 이와 관련된 다양한 Panel Data 분석기법을 사용하였다는 것이 특징이다. 대부분의 국내 논의는 횡단면 자료(Cross Section Data)를 사용하였거나 Pooled Cross Section Data를 Pooled Probit(Logit) 또는 Pooled OLS분석에 한정되어 있다. 이에 반하여 우리의 분석은 2003년부터 2008년까지 5년간의 한국 노동 패널 데이터를 패널 데이터(Panel Data)²⁾로써 이용하였다. 패널 데이터를 이용할 경우 일반적인 횡단면 자료나 시계열 자료를 사용하는 것보다는 표본의 수가 증가하게 되고, 독립변수간의 다중공선성이 감소하여 추정계수의 효율성이 개선된다는 장점이 있다.³⁾ 그러나 패널자료를 사용하여 분석할 경우 오차항이 횡단면상으로는 이분산현상, 시계열상으로는 자기상관현상이 나타날 가능성이 높기 때문에 이를 통제하는 것이 요구된다.

또한, 패널 데이터는 Pooled Probit이나 Pooled OLS와 다르게 동일한 개인에 대하여 분석하는 것이 가능하다. Pooled OLS는 동일한 개인을 고려하지 않기 때문에 각 년도의 표본이 모두 Random하게 추출되었다는 것을 가정한다. 단, 전기의 자료와 당기의 자료와 상당한 상관관계를 갖기 때문에 자기상관현상이 나타난다.⁴⁾ 이에 반해 Panel Data는 각 개인의 특성을 고려하는 고정효과분석(Fixed Effect Estimation)이나 확률효과분석(Random Effect Estimation)과 같은 다양한 방법을 해결책으로 사용하는 것이 가능하다. 또한, Cluster Sample 분석을 함께 하였는데 Cluster Sample 분석이란, 분석 대상을 각 개인으로 하지않고 특정 그룹을 설정하는 방법을 의미한다. Geronimus and Korenman(1992) Ashenfelter and Krugger(1994)등이 이 방법을 사용하였으며 또한 이는 고정효과분석과 확률효과분석에 응용될 수 있다. 본 논문에서는 Cluster Sample을 고정효과모형에 적용하였다. 또한 개개인의 관측치와 설명변수가 관계가 있지 않다면 더 효율적일 수 있는 확률효과모형을 추가로 분석하였다.

본 연구에서는 경제학에서 가정하고 있는 ‘대표적 개인’을 가정하며 우리는 각 개인의 성격을 각종 변수를 통해 통제한다. 그러므로 각 개인이 갖고 있는 특성보다 거주하고 있는 시·도별로 갖고 있는 특성이 더 클 수 있다는 생각 때문에 시·도별 지역적 차이를 통제하는 Cluster Sample 분석의 시도가 의미를 갖을 수 있다고 생각한다. 이처럼 다양한 패널자료 분석을 통해 관측계수의 강건성(Robustness)을 측정하고자 하였다.

노동공급곡선 추정의 도구변수로서 산업별 ROE를 사용하였다는 것 또한 본 연구의 의의라고 할 수 있다. 노동공급곡선 추정시, 문제점은 우리가 관측할 수 있는 자료는 노동 공급 함수와 수요 함수가 만나는 균형에서의 값이라는 것이다. 이러한 소위 연립 방정식(Simultaneous Equation)문제를 해결하기 위해서는 노동수요측면(회사의 입장)에서는 중요하지만 노동공급측면(노동자의 입장)에서는 중요하지 않은 변수가 필요하다. 이에 우리는 한국노동패널 데이터에서 제공하는 ‘개인이 종사

3) Hsiao, analysis of panel data(1986)

4) 우리의 분석에서는 Pooled OLS와 Pooled Probit분석시의 자기상관현상과 이분산성을 고려해주었음.

하는 산업별 코드'를 이용, 산업 별 ROE(Return on Equity)자료를 구성하였다. 이 변수를 이용한 것은 산업의 자기자본 순이익률이 각 회사가 노동 수요 함수를 결정하는데 영향을 줄 수 있을 것이라는 아이디어를 반영한 결과이다. 또한 자료를 사용할 때 자기자본 순이익률자료의 전년도 값을 이용하였다. 그 이유는 자기자본 순이익률 수치는 재무제표를 통해 계산되어지는 수치로 회계 연도인 12월 31일 이후 계산되어지는 수치이기 때문이다. 기업입장에서 자기자신의 기업성과를 알 수 있는 것이 아니고, ROE 자체가 그 해가 지난 후 회계결산이 끝난 이후 계산된다는 점을 고려할 때 ROE변수는 그 다음해의 노동수요량 결정에 영향을 줄 수 있으리라는 생각에 ROE 변수는 전년도 값을 이용하였다. 물론, ROE 데이터가 가진 한계나 한국노동패널 데이터상에서의 한계가 있으며 이에 대해 후에 자세하게 언급하도록 하겠다.

또한, 표본 선택 편의를 교정하기 위한 변수로서 시·도 별 여성의 평균 임금과 여성의 평균 실업률을 사용했다는 것은 본 논문의 새로운 시도라고 볼 수 있다. 황운재·최강식(1999)에서 표본 선택 편위의 교정을 위하여 시·도별 평균 실업률과 서울 더미를 사용하였다. 본 논문에서는 시·도별 여성의 평균 실업률과 시·도별 여성의 평균 임금을 사용하여 예측의 정확도를 높이기 위해 노력했다.

II. 기존문헌 고찰

1. 외국의 연구동향

기혼여성의 노동공급에 대한 외국의 경험적 연구는 최근까지 활발하게 진행되어 오고 있다. 기혼여성의 경제활동 참가에 관한 관심은 Mincer(1962)의 연구에서 시작된다. Mincer의 문제의식은 다음과 같은 실증결과로부터 출발한다. 남성의 노동시장 참가율과 실질임금의 관계는 "Backward-bending Supply Curve"를 잘 설명하고 있으며, 여성의 횡단면자료 분석 또한 이를 잘 설명한다. 그러나 기혼여성의 시계열 자료의 경우 이러한 이론적 분석과 배치되는 결과가 발견된다. 즉, 실질임금이 증가함에도 불구하고, 기혼여성의 노동시장 참가율은 상승하는 현상이 발견된 것이다.

Mincer는 이 문제를 분석하기 위하여 대체효과와 소득효과 개념을 이용하였고, 여성의 경우 실질임금 상승으로 인한 대체효과가 훨씬 크기 때문에 여성의 경제활동 참가율이 상승하였다고 결론 내렸다. Mincer는 분석에 있어서 기혼여성의 노동공급 분석 시에는 단순히 여가와 노동 간의 선택에 관한 분석만으로는 부족하다고 지적하였다. 이에 기혼여성의 여가는 가정 내에서의 노동을 포함하고 있으며, 노동공급곡선 도출 시에는 여가에 대한 수요뿐 아니라 가정에서의 노동에 대한 수요도 포함되어야 한다고 하였다.

이러한 이론적 논의는 Becker(1965)에 의해 발전된다. Becker는 가계를 소비자나 아닌 공급자로 가정한다. 가계는 일반재화와 시간을 투입하여 생산하는데, 전통적인 기업의 비용극소화 문제의 해로서 상품(Commodity)을 생산한다. 상품의 양은 가계의 가격과 가계의 임금으로 측정되는 자원을 제약으로 하는 효용극대화 문제를 통해 결정된다. 또한 노동공급을 결정하는 주체는 개인이 아닌 가계로 본다. 가계 내에서는 모든 구성원들의 일하지 않는 시간의 효율성을 극대화 하는 시간배분이 이루어진다. 가사노동 중에는 소득 그리고 가사노동중의 일부는 소득증가에 의해 쉽게 대체될 수 있는 것이 있지만, 자녀양육의 경우 그 대체성이 매우 낮기 때문에 일종의 사치재로 볼 수 있다. 결국 자녀양육 같은 가사노동은 여성의 경제활동 참가에 상당한 영향을 미칠 수 있다고 하였다.

이러한 논의 이후 기혼여성의 노동공급곡선을 추정하는 다양한 연구들이 진행되었다. 그런데 이러한 실증분석에 있어서 노동시장에 참여하지 않는 여성이 많이 존재한다는 사실은 이를 고려하지 않고 분석할 경우 편의를 일으킬 수 있다는 점이 지적되었다. 이에 Heckman(1980)은 여성의 취업 결정과 노동시간결정을 동시에 고려하고 표본선택에 따른 편의를 수정하였으며 임금의 내생성을 고려한 축약식을 사용하여 노동공급 구조방정식을 추정하였다.

그에 의하면 노동공급함수 추정에 있어서 표본선택에 따른 편의는 심각한 것으로 나타났으며 노동공급의 탄성치가 큰 값을 갖는 것으로 추정되었다. 이러한 Heckman의 접근방법과는 달리 Barton and Zabalza, Layard는 Tobit 분석을 통해 미취업여성의 경우 노동시간 자료가 0의 노동시장에서 절단되어 있다는 전제하에 임금이 상승하는 경우 비경제활동 상태에 있던 일부여성이 노동

공급을 시작하게 되고, 이미 취업한 여성은 노동시간을 늘리게 되므로 임금탄력성이 매우 탄력적이라는 결론을 내렸다.

또한 자녀양육비용이 여성의 노동공급에 미치는 영향을 고려하여 분석한 Hotz and Miller(1988)의 연구가 있다. 이들은 자녀의 수와 자녀의 나이가 기혼여성 노동공급행위에 유의한 영향을 주는 것을 고려하여 자녀양육 비용함수를 분석에 추가하였다. Cogan(1981)의 경우에는 자녀양육비용을 일종의 노동공급에 있어서의 고정비용으로 간주하고 사람들이 일종의 유보근로시간이 있는 것처럼 행동할 것이라는 가정 하에 분석을 시도하였다. 이러한 고정비용을 무시한 노동공급에 대한 추정 은 실제 노동공급 계수보다 과대 추정할 가능성이 있다는 점을 지적하였다.

그 외에 Mroz(1987)는 1976 PSID 자료를 이용하여 여성의 노동공급곡선을 분석한다. 그는 Tobit 가정을 통해 노동공급의 Self-selection을 통제하였고, 부인의 임금이 외생적이라고 가정한다. 또한 부인의 임금의 내생성을 통제하기 위하여 시장참가경력을 도구변수로서 사용한다. Tobit 모형은 이러한 소득과 임금을 과대추정할 수 있다고 지적하였다. 이러한 편의는 시장경력을 외생적으로 보기 때문이라고 밝혔다.

2. 국내의 여성노동공급에 관한 연구

국내의 여성노동공급에 대한 연구에는 황운재·최강식(1999)가 있다. 이들은 대우경제연구소의 한국가구패널조사자료를 이용하여 60세 이하의 미혼을 제외한 기혼, 이혼, 별거 및 사별한 여성을 대상으로 실증분석을 하고 있다. 이 연구는 표본의 선택성 및 임금변수의 내생성을 명시적으로 고려하고 있으며, 임금변수의 도구변수로서 경력변수의 포함여부를 살펴보았다. 또한 연구방법의 특징으로 준모수 추정법을 사용하였다는 것을 들 수 있다. 이 연구에서는 기혼여성의 노동시장 참여와 근로시간에 가장 중요한 영향을 미치는 요소로서 8세 미만의 자녀수를 들고 있다.

김현숙·원중학(2004)의 연구에서는 노동패널 1차~5차 자료를 Pooling하여 사별 및 이혼을 포함한 기혼여성을 대상으로 자녀보육비용이 여성 노동공급에 미치는 영향을 분석하고 있다. 이때 자녀 시간당 보육비 지출의 경우 가구의 소득, 가구의 총지출등과 상관관계가 크므로 내생성이 있을 가능성이 있으므로 이를 고려하여 분석하고 있다.

김현숙·성명재(2007)의 연구는 기혼여성 중 18세 이하의 자녀가 1명이라도 있는 가구의 기혼여성만을 대상으로 한 실증분석을 하였다. 이 결과를 기초로 하여 기혼여성의 보육비 지출에 대한 세액공제시 어떠한 영향을 받을 지를 정책시뮬레이션을 통해 분석하였다. 그 결과 근로소득장려세제와 유사한 내용으로 면세점 이하의 기혼여성 근로자만을 대상으로 세액공제를 허용해주는 방안의 경우 기혼여성의 노동공급시간 확대는 9.5%, 신규노동시장 참여비율은 19.6% 증가하는 것으로 분석하였다.

최형재(2008)는 설문조사 자료를 이용하여 자녀보육비용의 연장선상에서 자녀교육과 노동공급간의 관계를 분석하였다. 분석결과 자녀의 존재가 기혼여성의 노동시장 참가에 상당한 부의 영향을 미치고 있으며, 자녀의 사교육 지출에 막대한 금전적 투자를 하고 있는데, 이러한 사교육 지출이

기혼여성의 학력과 가구소득에 비례하며 계층 간 격차가 현저히 드러나고 있으며 자녀의 사교육이 일부 기혼여성들의 소득활동 참가유인을 저하시키는 것으로 분석하고 있다.

III. 자료

1. 『한국노동패널』

실증분석을 위한 자료로는 한국 노동연구원의 『한국노동패널』 자료로서 2003년 실시한 6차부터 10차까지의 패널자료를 이용한다. 분석기간을 2003년 이후로 결정한 것은 시간당임금변수를 생성하기 위한 연간소득자료가 2003년 이전자료에는 존재하지 않기 때문이다. 연간소득자료 대신 월소득자료를 12배하여 사용하는 방법을 택할 수도 있지만, 이런 방식에는 심각한 측정오차가 발생할 것이라 생각되었다. 대부분의 사람들이 자신의 월평균임금을 정확하게 알 수 있을 것이라 생각되지 않는데 대표적인 이유로는 상여금과 각종 보너스 혜택을 들 수 있다.

분석대상이 되는 표본은 2003년부터 2007년까지 총 18,341명의 기혼자이면서 배우자가 있는 여성이다. 기존의 국내문헌에서는 주로 분석대상이 되는 기혼여성을 미혼을 제외한 별거, 이혼, 사별의 경우를 포함한 경우를 분석대상으로 삼고 있다. 그러나 우리의 연구에서는 자녀의 유무와 관계없이 배우자가 있는 기혼여성의 경우만을 분석대상으로 삼고자 한다.

기존 연구들에 의하면 타가구원의 소득과 자녀의 수와 나이가 기혼여성의 노동공급에 유의한 영향을 미친다. 그런데 별거, 이혼, 사별의 경우 타가구원중 주요 소득원으로 예상되는 배우자의 부존재는 타가구원 소득에 유의한 차이를 갖게 할 것이고 이 결과 배우자가 있는 경우의 노동공급 패턴과는 다를 것으로 예상된다. 또한 자녀가 있는 기혼여성으로 한정할 경우 자녀의 수와 나이가 미치는 영향을 분석하는데 있어서 편의가 나타날 것으로 예상되므로 분석대상은 자녀의 유무와 관계없이 배우자가 있는 총 18,341명의 기혼여성이다. 이 중 만 60세 이하의 여성들을 경제활동 인구로 생각할 수 있고, 아웃라이어라고 생각되는 표본을 제외하고 최종적으로 15,701명을 분석대상으로 한다.

<표 2 > 한국노동패널 제 6-10차 표본 중 여성의 혼인상태

	6차연도	7차연도	8차연도	9차연도	10차연도	합계
미혼	1,421	1,391	1,349	1,355	1,338	6,854
기혼	3,606	3,664	3,657	3,699	3,715	18,341
별거	35	34	37	40	43	189
이혼	121	130	137	154	160	702
사별	798	808	831	844	846	4,127

자료: 한국노동연구원 한국노동패널 제 6차-10차

2. 『사업체 임금근로 시간조사』, 『경제활동 인구조사』

노동부의 『사업체 임금근로 시간조사』 통계에서 ‘시도별 여성 실업률’과 통계청의 『경제활동 인구조사』에서 ‘시도별 여성의 평균임금자료’를 사용하였다. 기혼여성의 노동공급분석을 위해서는 표본선택편의를 고려하여야 하는데, 이때 노동참가결정식과 노동공급식을 추정한다. 두 식을 추정하기 위해서는 노동공급시간 결정에는 영향을 미치지 않지만, 노동시장 참여여부를 결정할 때에는 영향을 미치는 변수를 추가해주어야 한다. 이러한 변수로서 황윤재·최강식(1999)은 기혼여성의 경제활동참가율은 전반적인 실업수준에 영향을 받을 수 있으므로 경기상태를 대표하는 변수인 실업률을 노동참가결정식의 설명변수로 포함시킨다.

이와 유사한 이유로 본 연구에서는 시·도별 여성의 실업률과 시·도별 여성의 평균임금자료를 사용하는데, 이때 각 개인이 거주하는 시·도별 자료를 이용하여 노동참가결정식의 설명변수로 포함시킨다. <표 3>과 <표 4>에서는 2003년부터 2007년까지 각 시·도의 여성의 실업률과 여성의 평균임금자료이다. 이 표를 보면 각 시·도별로 실업률과 평균임금에 있어서 유의한 차이를 나타내고 있음을 알 수 있다.

본 연구에서는 또한 시·도별 자료와 함께 시·도별 더미 변수를 사용하였다. 황윤재·최강식(1999)은 노동참가결정식에 서울거주여부를 더미변수로 추가하였었는데, 우리나라의 상황에서 서울이라는 지리적 변수가 갖는 특징이 있기는 하지만, 부산 등 다른 대도시와 비교하여 볼 때 기혼여성의 고용과 관련하여 차별화되는 이점이 어느 정도가 될지 알 수 없으므로 서울더미가 아닌 시·도별 더미를 이용하였다. 또한 시·도별 더미 변수를 사용하는 이유는 각 시·도별 여성의 고용상의 특징이 있을 수 있다는 것을 가정하는 것으로서 이는 고정효과 모형, Cluster Effect를 이용하여 분석한다.

〈표 3〉 시도별 여성실업률

단위: %

시도별	2003	2004	2005	2006	2007
전국평균	3.3	3.4	3.4	2.9	2.6
서울	4.3	4.3	4.6	4.1	3.3
부산	3.8	3.6	3.2	3.1	3.3
대구	3.9	4.2	3.7	2.9	3.2
인천	3.8	4.1	4	3.5	3.8
광주	4	3.5	3.6	3.1	2.8
대전	4.3	4.3	4.3	4.2	3.4
울산	3.4	4.1	3.8	3.2	2.8
경기도	3.4	3.4	3.9	3.2	2.6
강원도	2	2.5	1.7	0.9	1.4
충청북도	2.7	2.4	2.2	1.7	1.9
충청남도	2.6	2.1	2.8	2.3	2.2
전라북도	2.4	2.5	1.4	1.5	1.7
전라남도	2	2.1	1.8	1.3	1.3
경상북도	1.7	2.1	1.7	1.9	1.8
경상남도	2.2	2.2	2.5	2.1	1.5
제주도	2	2.3	1.8	1.9	1.7

자료: 통계청, 경제활동 인구조사

〈표 4〉 시도별 여성 월급여총액

단위: 원

시도별	2003	2004	2005	2006	2007
전국평균	1,176,575	1,256,585	1,362,262	1,461,042	1,561,520
서울	1,424,288	1,514,932	1,605,654	1,706,978	1,806,423
부산	1,020,462	1,113,277	1,179,039	1,256,005	1,376,752
대구	1,070,833	1,096,629	1,164,429	1,261,102	1,355,605
인천	1,073,097	1,161,778	1,247,978	1,329,807	1,458,695
광주	1,082,984	1,167,088	1,151,498	1,320,583	1,416,203
대전	1,222,500	1,316,452	1,382,955	1,538,232	1,657,196
울산	1,159,086	1,210,676	1,282,385	1,366,396	1,417,791
경기도	1,141,488	1,241,739	1,324,392	1,416,944	1,541,312
강원도	1,111,075	1,140,849	1,302,467	1,350,947	1,441,017
충청북도	1,043,296	1,132,043	1,203,343	1,312,589	1,391,851
충청남도	1,097,039	1,161,832	1,260,805	1,362,998	1,438,932
전라북도	1,030,612	1,095,410	1,151,289	1,291,014	1,410,140
전라남도	1,093,714	1,141,648	1,193,042	1,337,662	1,432,705
경상북도	1,027,055	1,135,725	1,209,553	1,292,813	1,347,430
경상남도	1,042,558	1,130,194	1,225,716	1,319,640	1,408,379
제주도	1,083,292	1,158,451	1,179,760	1,268,849	1,354,700

자료: 노동부, 사업체임금근로 시간조사 통계

3. 『기업경영분석』

기혼여성의 노동공급곡선을 추정하기 위해 한국은행 『기업경영분석』의 산업별 ROE(return on equity) 변수를 사용하였다. 관측되는 기혼여성의 노동공급에 대한 자료는 실제로 노동공급과 수요가 일치하는 지점에서 결정된 것이다. 그러나 실제 관측 가능한 자료는 각각의 노동공급곡선과 수요곡선이 아닌 최종 균형결과이다. 공급곡선을 도출하기 위해서는 기업의 노동수요에는 영향을 미치지만, 기혼여성의 노동공급에는 영향을 미치지 않는 변수가 분석에 사용되어야 한다. 이러한 변수로서 고려된 것이 산업별 ROE자료이다. 이때 산업은 한국표준산업분류상 중분류(2digit)로 분석하였다.

이 자료를 사용함에 있어서의 문제점은 각 개별기업의 응답 중 자신이 종사하고 있는 사업체가 영위하는 산업을 알지 못하거나 다른 이유로 보고하지 않은 경우가 존재한다는 거이다. 또한 기업경영분석 자체는 농업, 임업, 수도 사업, 음식점업, 금융·보험업, 공공행정·국방 및 사회보장 행정, 교육 서비스업, 보건 및 사회복지사업, 개인 및 가사서비스업 등을 영위하는 산업에 대해서는 조사를 하지 않아 이 산업들에 대해서는 자료가 미비하다.

금융·보험업의 경우 금융감독원 금융통계시스템을 통해 각 금융업별 ROE를 구할 수는 있으나, 금융·보험업의 경우에는 각 금융업별 ROE를 계산하는 방식이 다르고, 각 특성이 달라 이를 평균내는 것이 무의미할 수 있다. 예를 들어 은행업의 경우는 예금자로부터 수신한 예금의 경우 모두 은행의 부채로 계상되기 때문에 자기자본비율은 다른 산업에 비해 낮은 편이라 할 수 있으므로, 다른 금융업과 ROE를 평균하는 것은 의미가 없으므로 금융업의 경우 분석대상에서 제외하였다.

기타 위에서 언급한 산업들 또한 ROE 자료가 없으므로 분석대상에서 제외되었다. 이러한 자료의 결측은 표본선택상의 또다른 편의를 야기할 수 있다는 점을 우선 지적한다. 이 부분에 대해서는 다음장에 걸쳐 더 자세히 논의하도록 하겠다.

IV. 실증분석

1. 모형 및 변수

가. 노동참여결정식

$$\text{식(1)} \quad \text{Employed}_{ti} = 1[\text{Employed}_{ti}^* > 0]$$

$$\text{식(2)} \quad \text{Employed}_{ti}^* = \beta_0 + \beta_1 \text{edu}_{ti} + \beta_2 \text{age}_{ti} + \beta_3 c3\text{vm}_{ti} + \beta_4 c410\text{vm}_{ti} + \beta_5 c1118\text{vm}_{ti} \\ + \beta_6 \text{unemp}_{ti} + \beta_7 \text{provinc}_{ti} + \sigma(\text{year}) + \omega(\text{area}) + u_{ti}$$

$Employed_{it}$ 는 여성의 노동공급여부를 나타내는 변수로서 t기에 i번째 여성이 취업을 하기로 결정한 경우 $Employed_{it} = 1$ 이고, 그렇지 않은 경우 $Employed_{it} = 0$ 이 된다. u_{it} 는 교란항이다.

노동시장 참여 결정식에 포함되는 대표적인 변수는 각 연령대별 아이의 수, 타가구원의 소득, 학력, 연령이다. 이 변수들에 대한 기본 통계값은 다음과 같다.

〈표 5〉 노동시장 참여 결정식 변수의 기초통계량

변수	내용	평균	최소값	최대값	표준편차
Employed	취업=1, 미취업=0	0.478	0.000	1.000	0.500
Edu	교육을 받은 년수 (예: 석사=18)	11.640	0.000	23.000	3.143
Age	나이	41.820	19.000	60.000	9.337
othinc	여성의 근로소득이 제외된 가정 내 소득 (단위: 연, 백만원)	30.76	-119	1048.26	29.29
C3num	만 0-3세 아이의 수	0.103	0.000	3.000	0.321
C410num	만 4-10세의 아이의 수	0.223	0.000	3.000	0.484
C1118num	만 11-18세의 아이의 수	0.258	0.000	3.000	0.530
unem_p	시·도별 여성의 평균 실업률	3.246	0.900	4.600	0.895
prov_inc	시·도별 여성의 평균 임금 (단위: 월, 만원)	134	102	181	20
σ (year)	연도 더미	-	0	1	-
ω (area)	시·도별 더미	-	0	1	-

이하에서는 기존 문헌들을 바탕으로 위의 변수들이 실증분석 결과 어떤 추정치를 갖게될 것인지 생각해보도록 하겠다.

1) 학력

여성의 학력은 황윤재·최강식(1999)의 경우 부의 상관관계를 갖는 것으로 나타났다. 선진국의 경우 교육수준과 기혼여성의 경제활동 참가율 사이에 강한 양의 상관관계를 갖는 것으로 나타나는 것과는 상반되는 결과이다. 이에 대하여 황윤재·최강식(1999)은 이러한 단순한 상관관계 분석은 결과를 오도할 가능성이 있다고 지적한다. 학력이 경제활동에 긍정적인 영향을 미치는 이유는 학력이 높을수록 시장에서의 임금이 높아지기 때문인데, 실제로 학력은 기대비용으로서 의중임금수준을 상승시키는 요인으로 작용하기도 하므로 학력의 효과는 단순하게 추측할 수 없다고 설명하고 있다.

노동경제학에서 교육은 생산성을 직접 향상시켜 주는 방법 또는 노동 수요자에게 자신의 생산성을 나타내는 일종의 신호역할을 하는 것으로 알려져 있다. 이 중 두 가지 모두는 임금을 상승시키는 방향으로 작용하게 되는데, 이는 또한 노동자의 의중임금을 상승시키게 되므로, 실제로 학력수준이 노동시장의 참가여부에는 어떠한 영향을 미치는지는 알 수 없다.

2) 연령

우리는 표본을 만 60세 이하의 기혼 여성으로 설정하였다. 이는 경제활동참가인구 통계를 고려하여 결정하였다. 남성의 경우 대부분 취업자의 경우 특별한 사정이 없는 한 자신의 정년때까지는 일을 그만두지 않는 경향을 보이거나, 여성의 경우는 자신의 생애주기 관점에서 노동공급을 결정하는 경향을 보인다. 예를 들어 30대를 전후한 가임기의 여성의 경우 출산, 육아등으로 인하여 일을 그만두는 경향이 많으며, 아이가 학교에 갈 나이가 되어 일을 할 수 있는 경우 다시 노동시장에 참여하게 된다. 단, 이 경우 30대를 전후한 시기에 경력의 단절은 여성의 노동시장 재참여를 어렵게 만드는 요인으로 작용하기도 한다. 또한 최근은 사교육비의 증가로 인하여 중년층 여성의 노동공급이 증가하는 경향을 보이기도 하는데, 여성의 노동공급에 있어서 연령은 단순한 선형 관계를 갖는 것이 아니라, 그들의 생애주기에 따라 다른 패턴을 보이게 될 것이라고 예상이 된다. 그래서 본 분석에서 연령변수는 단순히 선형관계만을 나타내지는 않을 것이라는 가정하에 예를들어 연령² 변수를 추가함으로써 비선형의 관계를 나타낼 경우 이를 추정할 수 있도록 하였다.

3) 타가구원의 소득

타가구원의 소득은 비노동소득이 증가할 경우 정상재인 여가의 소비가 늘어나는 소득효과를 잘 반영하는 결과를 나타낸다. 즉, 타가구원의 소득은 여성의 노동공급에 대체로 부정적인 영향을 미치는 것으로 나타났다. 황수경(2004)의 연구에서는 기타소득의 경우 소득효과를 가져와 기혼여성의 취업에 부의 효과를 가져오는 것으로 나타났다. 이에 반하여 배우자 소득의 경우에는 소득효과를 통해 기혼여성의 경제 참가에 부의 효과를 미칠 것이라는 이론적 예측과 상반되는 결과를 나타냈다. 이러한 결과에 대해서 황수경(2004)는 배우자소득이 교육 이외의 관측되지 않은 여성의 능력이나 자원의 측면을 대표하고 있을 수 있다고 분석하고 있다.

타가구원의 소득중 가장 중요한 역할을 하는 것은 배우자의 소득일 것이라고 생각된다. 이에 본 연구에서는 기혼여성중에서도 별거, 사별, 이혼한 경우를 제외하고 배우자가 있는 표본들만을 대상으로 분석하였다.

4) 자녀 수

우리의 연구에서는 만 0-3세, 4-10세, 11-18세의 세 가지의 그룹으로 자녀를 나누었다. 그룹을

나누게 된 기준은 자녀보육기관 등의 이용 가능 나이이다. 만 0-3세의 자녀는 실제로 탁아소나 유치원 등을 이용하기 어려운 나이이며 가능하더라도 비용이 매우 많이 든다. 만 4-10세는 유치원 및 초등학교 저학년으로 보육기관을 이용하는 것이 가능하므로 여성의 노동 참가가 늘어날 수도 있다. 하지만 여전히 자녀들에게는 부모의 손길이 필요하다는 점에서 여성의 노동 참가에 부정적인 영향을 미치는 것도 가능하다. 만 11-18세는 초등학교 고학년, 중학교, 고등학교를 의미하며 상대적으로 부모가 해주어야 하는 일이 줄어든다. 또한 상당수의 가정에 큰 부담이 될 수 있는 자녀의 사교육비는 기혼 여성이 노동 결정의 큰 유인이 될 수 있다.

기존의 연구에 있어서 미취학 연령의 자녀는 기혼여성의 경제활동 참가에 부정적인 영향을 주는 것으로 나타났다. 황윤재·최강식(1999)에 따르면 8세 미만의 자녀의 존재는 기혼여성의 유보임금을 높이는데 반해 8세 이상 19세 이하의 자녀들은 여성의 노동시장 진출에 큰 영향을 주지 않는 것으로 나타났다.

5) 시·도별 여성의 실업률

실업률은 기혼여성에게 여러 가지 측면으로 영향을 줄 수 있다. 먼저, 실업률이 너무 높을 때에는 진입 전, 스스로 노동 시장의 진입을 포기할 수 있으며 동시에 직업을 갖길 원해도 찾기 힘들 수 있다는 것이다. 본 논문은 실업률이 여성의 노동 참여 결정에 어떠한 영향을 미치는지에 대하여 검증하고자 하며 통계청의 시·도별 여성의 실업률 데이터를 사용하였다.

6) 시·도별 여성의 평균 임금

각 개인은 자신의 유보임금이 시장임금보다 클 때, 노동 시장의 참가를 결정한다. 노동 참가 결정에서 임금이 매우 중요한 역할을 함에도 불구하고 우리가 갖고 있는 데이터는 이미 노동 시장에 참가한 사람들만의 시장 임금뿐이다. 실제 사람들이 받고 있는 임금은 유보임금보다는 크거나 같을 것이기 때문에 실제 받는 임금이 유보임금을 대체할 수는 없다. 이에 여성의 노동공급에 있어서 주거지역별로 특징이 있을 것이라는 가정하에 여성의 유보임금의 대용치로서 시·도별 여성의 평균 임금 자료를 사용하였다. 또한 시·도별 여성의 평균임금 대신, 예측된 유보임금을 사용하는 것을 고려해볼 수 있다.

7) 연도 더미

분석 기간인 2003년부터 2007년까지 5년간의 데이터가 사용되었으며 2003년을 베이스 년도로 삼고 연도별 특성을 통제하기 위해서 4개의 연도 더미(y_7-y_{10})를 생성하였다. (예: $y_7=1$ if $year=2004$, $y_7=0$ otherwise)

8) 지역 더미

우리가 사용하는 자료는 패널 형식의 자료이며, 표본은 사별·이혼 등을 미포함시킨 기혼 여성이 분석 대상이다. 5개년간의 총 관측치는 15,701개이다. 이러한 패널자료를 사용할 때는 개인의 특성을 어떠한 방식으로 통제해야 할지를 결정하여야 한다. 관찰되지 않는 개인의 특성이 설명변수와 상관관계를 갖게 된다면 편의를 만들 수 있다. 이러한 경우 일계차분(First Differencing)이나 고정효과(Fixed Effect) 분석이 해결책으로 쓰일 수 있다.

우리의 데이터에서는 개인의 수는 많되 관찰년도는 짧은 특성을 가졌기 때문에 일계차분방법은 적절하지 않다고 볼 수 있다. 각 개인의 특성을 고려해주는 고정효과 분석은 개인별로 3000여개의 더미를 만드는 것과 같은 효과를 줄 수 있다. 모든 개인의 특성을 각각 고려하는 것은 너무 많은 자유도(Degree of Freedom)를 잃을 가능성이 있다. 이러한 이유로 이론적 근거로서 미시경제학에서 분석시 사용하는 기본적 전제는 “대표적 개인”이라는 전제하에 거주지역을 시도별로 구분하여 거주지의 특징을 고려하여 주는 Cluster Effect를 고려하여 주었다.

이러한 이유로 본 논문에서는 각 지역을 기준으로 하는 고정 효과 분석을 하였다. 이 경우 통제되는 특성변수로서 시·도 별 15개의 지역 더미(area_1-area_15)가 이용될 수 있으며, 대표적 개인이라는 가정과 시·도별 여성 노동 참가 결정에 미치는 환경이 다를 수 있다는 것을 암묵적으로 가정하고 있다고도 볼 수 있다.

<표 6> 시·도 별 코드

코드	시·도
1	서울
2	부산
3	대구
4	대전
5	인천
6	광주
7	울산
8	경기
9	강원
10	충청북도
11	충청남도
12	전라북도
13	전라남도
14	경상북도
15	경상남도
16	제주도

나. 노동공급곡선

이론적으로 노동공급곡선과 노동수요곡선이 만나는 점에서 균형이 결정되고, 그 점에서 균형 노동 시간과 임금이 결정된다. 본 논문을 통해 예측하고자 하는 것이 노동공급 곡선이지만 실제로 데이터를 통해 우리가 관찰할 수 있는 것은 균형에서의 값이다. 이를 연립방정식 문제 (Simultaneous Equation)라고 하며 노동공급곡선을 식별(identify)하기 위해서는 *the Order Condition*을 만족시켜야 한다. 이는 노동의 공급식에 존재하지 않지만, 노동의 수요식에서는 존재하는 변수가 최소한 한 개가 있어야 한다는 것을 의미한다. 이 때 노동공급곡선은 노동자의 입장이라고 할 수 있으며, 노동수요곡선은 회사 측 입장이라고 볼 수 있다.

즉, 회사는 심각하게 고려하지만 각 노동자는 고려하지 않거나 덜 신경 쓰는 요소가 있으며 그 변수를 임금의 도구 변수로 이용한다는 것이다. 연립방정식 문제는 임금의 내생성 문제와 유사한 문제로 볼 수 있다. 기존 연구에서는 여성의 교육과 나이 변수의 제곱과 삼승, 남편의 교육과 나이 변수 및 그의 제곱, 경력 변수를 이용해왔다. 최근에는 임금의 도구변수로 이용되는 경력변수가 외생적이지 않을 수 있다는 가능성이 제기되고 있다. 이에 우리의 연구에서도 임금의 도구변수로서 경력변수가 외생적인지 여부를 판단하기에 앞서 경력변수를 사용하는 모형과 그렇지 않은 모형을 비교하여 분석하고자 한다.

또한 시·도 별 실업률을 도구 변수로 사용한 몇몇의 기존 연구가 있었으나 본 논문에서는 이 변수를 노동참여결정식에 이용하는 것을 채택하였다.⁵⁾ 그 이유는 시·도별 실업률이 여성의 노동공급

량을 결정하기 보다는 노동시장에 참여할지 말지를 결정하는 요소로 더 큰 작용을 할 것으로 생각되었기 때문이다. 이미 노동시장에 참여한 사람의 입장에서는 거시경제변수인 실업률 자체가 본인의 노동공급시간을 결정하는 요인으로 작용하지 않으리라는 판단하였다. 단, 노동시장의 경직성으로 인하여 기업입장에서 경기 호황기에 신규직원을 추가 채용하여(실업률의 감소) 대응하지 않고, 기존 노동자의 추가근무를 통해 생산량을 증가시키고자 할 경우 실업률이 변하지 않고 근무시간이 변하게 될 가능성이 존재한다. 그러나 추가근무는 일반근무에 비해 높은 시간당임금을 지불하므로 평균시간당임금이 상승하게 되므로 높아진 시간당임금에 의한 노동공급증가 효과로 포착할 수 있을 것으로 생각된다.

본 연구는 개인이 종사하고 있는 산업의 ROE(Return on Equity)변수⁶⁾를 도구 변수로 추가하는 시도를 하였으며 도구변수로서 적절한지 테스트하였다. 단, 앞에서 언급한 것과 같이 금융업 등의 산업의 ROE가 기업경영분석에 보고되지 않아 이용할 수 없다는 점과 자신이 종사하고 있는 산업의 종류를 보고하고 있지 않은 개인들이 다수 있다는 점이 큰 한계로 작용할 수 있다.

임금의 내생성 문제를 위해 사용되어진 도구 변수법은 먼저 1단계로 도구변수 집합을 이용하여 임금변수를 예측한다. 이 때 예측되어진 임금변수는 2단계에서 설명변수로서 사용된다.

1) 1단계

도구변수 추정법(2SLS)의 1단계는 도구변수를 이용하여 시간당 임금변수를 추정하는 것이다. 본 연구에서는 보다 안정적인 예측을 위하여 2단계에서 사용할 변수를 1단계의 도구변수 추정을 위한 집합에 사용하였으며 이는 2단계에서의 변수가 모두 외생적이라는 도구변수법의 가정과 같은 내용을 담고 있다.⁷⁾ 아래의 <표 7>은 도구 변수 집합을 나타낸다.

$$\begin{aligned} \text{식(3) } lwage_{ti} = & \delta_0 + \delta_1 edusq_{ti} + \delta_2 agesq_{ti} + \delta_3 eduage_{ti} + \delta_4 educb_{ti} + \delta_5 agecb_{ti} + \delta_6 edusqage_{ti} \\ & + \delta_7 eduagesq_{ti} + \delta_8 hage_{ti} + \delta_9 hedu_{ti} + \delta_{10} hagesq_{ti} + \delta_{11} hedusq_{ti} + \delta_{12} hagehedu_{ti} \\ & + \delta_{13} roe_{ti} + \delta_{14} exp_{ti} + \delta_{15} exogenous_{ti} + w_{ti} \end{aligned}$$

5) 표본선택편의의 교정을 위한 Heckman의 2단계 추정법의 사용을 위해서는 2단계에서 노동공급곡선의 추정을 위해 사용되지 않는 변수가 최소한 하나 이상 1단계, 노동참가결정식에 존재해야함.

6) 한국은행, 기업경영분석

7) 통계 프로그램인 stata는 이를 자동으로 실행함.

〈표 7〉 도구변수 목록

	설명	관측치수	평균	표준편차	최소값	최대값
edusq	교육연수 ²	15685	145.37	69.97	0	529
agesq	나이 ²	15701	1836.05	796.25	361	3600
eduage	교육연수*나이	15685	470.62	126.83	0	1081
educb	교육 ³	15685	1906.79	1311.72	0	12167
agecb	나이 ³	15701	84173	53103	6859	216000
edusqage	교육 ² *나이	15685	5737.06	2632.24	0	24863
eduagesq	교육*나이 ²	15685	19975	8121	0	64800
hage	남편나이	12387	45.55	10.52	21	88
hedu	남편교육연수	12549	12.33	3.62	0	23
hagehedu	남편나이*교육	12378	554.36	164.30	0	1403
exp	경력(단위:월)	12558	37.51	73.94	0	546
roe	Return on Equity	3902	8.31	10.26	-42.82	137.85

우리는 경력변수의 내생성을 테스트하기 위하여 경력변수를 포함한 도구변수 집합과 경력변수를 포함하지 않은 집합을 이용하여 분석한다. 만약 두 결과가 상당한 정도로 다르다면 경력이 외생적이라고 볼 수 없다. 상당히 다른 정도의 기준을 위해서는 Hausman Test 등을 이용할 수 있으나 Hausman Test는 검정력이 낮다는 문제를 갖고 있다. 그러므로 Hausman Test가 기각 되지 않더라도 경력변수가 외생적이라는 결론을 내리는 것은 성급한 판단일 수 있다.

노동공급곡선을 추정하기 위하여 추가로 임금의 도구변수로 사용하는 산업별 ROE 변수의 경우 실제 자료의 사용에 있어서 결측이 많이 발생할 수 있다. 첫 번째 이유는 자신이 종사하는 산업코드를 보고하지 않은 경우가 많이 있다는 것이다. 또한 산업별 ROE 변수를 구할 수 있는 기업경영분석이 제공하는 산업통계가 제한적이기 때문이다. 기업경영분석은 농업, 임업, 수도사업, 음식점업, 금융·보험업, 공공행정·국방 및 사회복지 행정, 교육서비스업, 보건 및 사회복지사업, 개인 및 가사서비스업 등을 영위하는 산업에 대해서는 표본조사를 시행하지 않는다. 이러한 이유로 자신이 종사하는 산업이 한국표준산업분류 중분류(2digit)상 기업경영분석에서 조사하지 않는 분류에 해당하는 표본은 결측이 생겨날 수 밖에 없다.

<표 8> 산업별 ROE 변수 사용가능한 표본의 기초통계량

	최종샘플 (1=2-3)	전체 (2)	결측 (3)	금융업 (4)
lwage (ln(시간당임금))	-0.88	-0.64	-0.24	-0.03
whour (연간근로시간)	2627.38	2510.98	2313.85	2318.54
edu (교육연수)	11.02	12.08	13.88	12.88
age (연령)	41.42	40.27	38.32	36.14
othinc (타가구원소득)	21.53	24.16	28.65	31.35
c3num (3세미만자녀수)	0.03	0.05	0.09	0.14
c410num (4-11세자녀수)	0.16	0.17	0.20	0.13
c1118num (11-18세자녀수)	0.31	0.26	0.18	0.15
관측치수	1751	2785	1034	172

주: (4)의 금융업의 경우 기업경영상 자료가 없으므로 결측치에 포함되는 수치임.

이 문제에 대해서 결측치가 원래 본 표본의 37%가까이 되는 관계로 처음에는 산업전체 평균 ROE를 사용하고자 하였으나, ROE라는 것 자체가 비율변수이므로 평균을 할 경우 그 의미가 퇴색될 우려가 있고, 산업별 가중치를 고려한 가중평균을 이용할 수도 있으나, 현실적으로 계산이 어려워 사용할 수 없었다.

이에 표본의 탈락과 편의가 예상됨에도 불구하고, 산업별 ROE 변수를 사용할 수 없는 변수는 타락시켜 이용해보았다. 위의 표는 이러한 경우 어떠한 표본이 최종적으로 포함되게 되는지 기초통계량을 살펴본 것이다. 표에서 알 수 있듯이 산업코드 미보고 및 기업경영분석 자료 미비로 인하여 결측되는 표본은 최종적으로 분석되는 표본에 비하여 시간당 임금은 높고, 연간 근로시간은 적으며 교육연수는 높고, 연령은 낮다. 또한 타가구원 소득은 높으며, 3-11세 자녀의 수는 많지만, 11-18세 자녀의 수는 더 적다.

(4)에 보고한 금융업의 경우 산업코드를 보고하였지만, 기업경영상 자료가 없는 표본 중 가장 많은 비율을 차지하고 있어 그들의 기초통계량을 따로 보고해 보았다. 이들은 다른 이유로 결측된 표본에 비하여 시간당 임금은 높고, 연간 근로시간은 유사하나 비교적 높고, 교육연수와 연령이 낮다. 타가구원 소득은 매우 높은 편이며, 타가구원 소득은 최종표본과 비교하여 연간 1000만원 가까이 차이가 난다. 이러한 점들을 고려해 볼 때 산업ROE자료를 이용하기 어려운 표본을 탈락시키고 진행하는 본 분석은 그 방향에 대해서는 확실하게 말할 수는 없지만, 편의가 나타날 것으로 예상된다.

2) 2단계

도구변수법의 2단계에서는 1단계에서 추정된 임금 변수를 실제 임금데이터 대신 설명변수로서 사용한다. 단, 우리는 표본선택편의의 수정을 위하여 Heckman의 2단계 추정법을 사용한다.. 즉, Heckman 추정법의 1단계인 노동참여 결정식에서 Mill's Inverse Ratio를 추정한 후, 이를 도구변수법의 2단계에서 설명변수로 추가 한다. 우리가 추정 할 식은 다음과 같다.

$$\text{식(4) } \textit{whour} = \gamma_0 + \gamma_1 \textit{lwagehat} + \gamma_2 \textit{edu} + \gamma_3 \textit{age} + \gamma_4 \textit{othinc} + \gamma_5 \textit{c3num} + \gamma_6 \textit{c410num} + \gamma_7 \textit{c1118num} \\ + \gamma_7 \lambda + \sigma(\textit{year}) + \omega(\textit{area}) + v$$

<표 9> 공급곡선 추정 2단계에 사용되는 변수의 정의

변수	설명	관측치	평균	표준편차	최고값	최대값
whour	연간 노동 시간	3971	2500.01	698.95	52	14248
lwagehat	1단계 예측된 시간당 임금 변수 ⁸⁾	3033	-	-	-	-
edu	교육을 받은 년수	3967	12.18	3.22	0	23
age	나이	3971	40.21	8.48	21	60
Othinc	타가구원 소득	3931	2511.5	2082.6	-5080	39040
c3num	만 0-3세의 아이의 수	3971	0.08	0.28	0	2
c410num	만 4-10세의 아이의 수	3971	0.22	0.47	0	3
c1118num	만 11-18세의 아이의 수	3971	0.31	0.59	0	3
$\sigma(\textit{year})$	년도 더미	-	-	-	0	1
$\omega(\textit{area})$	지역 더미	-	-	-	0	1
λ	Mill's Inverse Ratio (노동참여결정식에서 계산된 결과)	-	-	-	-	-

2. 분석방법

가. 노동참가결정식

본 연구에서는 여성의 노동참가결정식을 추정하기 위하여 Logit모형이 아닌 Probit모형을 사용하였다. 이는 기존 연구에서 Probit은 계산의 복잡함 때문에 보다 정확한 예측을 할 수 있음에도 시행하지 못하였으나, 최근에는 통계프로그램으로 이러한 단점을 극복할 수 있다(Wooldridge(2008)). 실증 분석을 위한 통계 프로그램으로는 Stata가 사용되었으며, Probit의 가장 큰 단점이었던 복잡한 계산을 쉽게 해주는 역할을 하였다.

노동참가결정식의 추정을 위해 사용된 방법은 세 가지로 다음과 같다. Pooled Probit과 Population Average Estimation, Random Effect Estimation이다. Pooled Probit은 횡단면 분석에 사용된 Probit 방법을 Panel자료에 적용한 것으로서 연도별 특성의 고려를 위해 연도 더미를 추가로 사용한다. 단, 여기에는 전기의 종속변수 및 설명변수가 결정에 영향을 미치지 않는다는 가정이 필요하다. 이러한 가정은 예측된 잔차항의 전기값을 설명변수로 사용하여 이 값의 유의한지 테스트해 볼 수 있다.⁹⁾

Probit 모형을 해석할 때 주의해야 할 점은 통계 프로그램에서 보고되어지는 계수를 상수로 해

8) 1단계에서 예측된 log(시간당 임금)은 도구 변수 집합에 따라 차이를 보임.

9) 이를 확인하는 방법은 Wooldridge(2002)를 참고할 수 있음.

석해서는 안 된다는 것이다. 이는 Probit 혹은 Logit 모형의 장점이기도 한 것으로, 각 점에서 Partial Effect의 크기가 다르다는 특성 때문이다. 이를 해석하기 위해서는 예측되어진 계수에 Scaling Factor를 곱해야 하는데 Scaling Factor가 각 점에서 다른 값을 갖는다.¹⁰⁾ 예를 들어 만 0-3세의 아이가 한 명이 있을 때와 2명, 혹은 3명이 있을 때 모든 Partial Effect는 다르다는 것이다. 이러한 Probit(Logit)의 장점은 해석시 어려움을 겪을 수 있다. 그렇기 때문에 각 설명변수의 중간값 및 평균에서 어떤 계수를 갖는지를 관찰해볼 수 있다. 그러나, 본 논문에서 중심으로 연구하고자 하는 부문은 기혼 여성 중 특정 그룹의 특징을 보는 것으로 한정하는 것이 아니기 때문에 이런 방식의 해석이 가장 적절하다고 할 수만은 없다.

또한 자주 사용되어지는 방법 중 하나로서 Population Average Estimation은 해당되는 샘플들의 Partial Effect를 모두 평균 낸 값을 의미하며 대략적인 분포 및 성향을 확인하기에 적당하다. Random Effect Probit Estimation이란 관측되지 않은 개인의 특성에 대하여 가정이 필요하며 이 가정이 성립하지 않을 경우, 불일치성을 갖고 올 수 있다.

본 연구에서는 Pooled Probit, Population Average Estimation, Random Effect Estimation을 통해 각각 노동참여결정식을 추정한다. 이를 통해 예측되어진 Mill's Inverse Ratio를 노동공급곡선의 표본선택편의의 교정을 위하여 사용하며, 각 방법론이 이끄는 결과의 차이를 보는 것 또한 의미가 있을 것이다. 다양한 계량방법론을 통해 도출된 결론이 유사하다면 본 연구의 추정이 강건성(Robustness)를 갖는다고 해석할 수 있다.

나. 노동공급식

노동공급곡선의 추정시, 임금의 내생성을 고려하기 위하여 도구변수추정법이 사용되었으며 경력, ROE, 교육 및 나이의 제곱 및 삼승 등을 도구변수집합으로 설정하였다.¹¹⁾ 본 연구에서 다루고 있는 자료는 패널 데이터이기 때문에 패널 데이터를 다룰 수 있는 다양한 방법이 사용되었다. 우리는 개인별로 관측되지 않은 특성과 설명변수와의 상관관계를 제어하기 위한 방법으로 고정효과 분석(Fixed Effect Estimation)을 이용하였다. 또한 미시경제학의 “대표적 개인”이라는 가정 하에 개인별 특성을 제어해주는 대신 거주하고 있는 시·도의 Cluster Effect를 제어해주는 Cluster Sample¹²⁾의 고정효과분석 방법을 사용하였다. 또한, 미관측 변수의 효과(Unobserved Effect)와 설명변수의 상관관계가 없다는 가정 하에, 확률효과분석을 사용하였다.

앞 장에서 언급한 바와 같이 표본선택편의를 교정하기 위해 노동참가결정식에서 예측된 Mill's Inverse Ratio가 설명변수로 추가되었다. 만약 Mill's Inverse Ratio의 계수가 0이라는 귀무가설을 기각한다면 우리의 자료가 표본선택편의를 갖고 있다고 볼 수 있으며, 기각하지 못한다면 표본선택편의가 없으며 이를 고려할 필요가 없다는 것이다.

10) Probit의 Scaling Factor의 경우 $x=0$ 에서 약 0.4이며, Logit의 경우, 약 0.25의 값을 갖음.

11) 도구 변수 집합에 대한 설명은 앞 장, 분석 모형 및 변수에서 자세하게 다루었음.

12) Wooldridge(2008), p50

3. 결과

가. 노동참가결정식

아래의 <표 10>은 세 가지 방법, Pooled Probit Estimation과 Population Average Estimation 그리고 Random Effects Estimation을 통해 노동참가결정식을 추정한 것이다. 단, Probit 모형은 Partial Effects는 예측된 계수가 나타내는 것이 아니라 각 점에서의 Scaling Factor와 예측된 계수의 곱이 이를 나타낸다. 이를 분석하기 위해서 Stata에서 mfx compute라는 명령어를 사용하여 Partial Effects를 예측하였다.

<표 10> 표본선택 고려하는 경우 probit 분석: 노동시장 참여식

	(1) Pooled Probit estimates		(2) GEE population-averaged		(3) Random-effects probit regression	
edu	0.0055	(1.11)	0.0019	(0.59)	0.0000	(0.00)
age	0.0033	(2.02)**	0.0014	(1.32)	0.0072	(3.69)***
othinc	-0.0059	(-5.25)***	-0.0024	(-3.68)***	-0.0021	(-6.32)***
c3num	-0.5257	(-13.92)***	-0.2084	(-10.82)***	-0.2433	(-8.57)***
c410num	-0.0710	(-3.11)***	-0.0277	(-2.01)**	0.0410	(1.95)*
c1118num	0.2625	(12.85)***	0.1045	(8.41)***	0.1340	(7.08)***
unem_p	-0.0408	(-0.95)	-0.0277	(-2.62)***	-0.0481	(-2.80)***
prov_inc	0.0000	(1.20)	0.0000	(0.41)	0.0000	(0.73)
관측치수	15479		15479		15479	
연도더미	포함		포함		포함	
지역더미	포함		미포함		미포함	

주: () 안은 z-value 이며, **,***은 각각 10%, 5%, 1% 수준에서 유의함을 뜻함.

<표 10>의 결과를 보면 세 가지 다른 방법을 통해 추정된 값은 만 11-18세 자녀수에 대한 확률 효과 분석을 제외하고 모두 같은 부호의 계수값을 갖는다. 단, 통계적으로 유의한 정도는 세 조금씩의 차이를 보이고 있다.

먼저, 교육 및 나이는 기혼 여성이 노동 시장에 참여하는 것에 정의 효과를 갖는다. 여기에서, 유의한 값을 갖는 것은 Pooled Probit과 Random effects probit에서의 나이 변수(age)로 각각 5%와 1%에서 유의한 결과를 갖는다.

타가구원 소득은 세모형의 예측 결과가 모두 1%에서 통계적으로 유의하며 타가구원의 소득이 높을수록 기혼여성은 노동시장에 덜 참여하려고 한다고 해석할 수 있다. 이는 우리가 예측하던 결과 및 기존 연구의 결과와 일치한다.

자녀의 수를 보면 만 0-3세의 자녀가 있을 때, 약 20-52%로 여성이 노동시장에 참가하는 것에 부정적인 영향을 미치며 세 결과 모두 1% 유의수준에서에서 통계적으로 유의하다. 만 4-10세 자녀의 수의 경우 확률효과 Probit분석에서 양의 부호를 갖으며 그 외의 두 모형은 음의 부호를 갖는

다. 이러한 상반되는 결과는 각 분석상의 기본 가정 때문에 발생한 것으로 해석할 수 있다. 만약 확률효과분석에서의 가정이 성립된다면 만 4-10세의 자녀는 유치원, 학교 등을 이용할 수 있는 나이로 기혼 여성이 노동 참가를 촉진시킬 수도 있다는 것으로 해석할 수 있다. 그러나 확률효과분석의 가정이 성립하지 않을 경우, Cluster Effect를 고려한 Pooled Probit이나 Population Average Estimation이 보다 일치성 있는 계수를 추정한다고 볼 수 있을 것이다. 이 경우에는 만 4-10세의 자녀는 비록 유치원 등의 보육기관이 있을 지라도 부모의 역할이 필요할 수 있다고 해석할 수 있다.

만 11-18세의 자녀의 경우, 세 모형으로부터 예측된 결과는 모두 양의 부호를 갖으며 1%유의수준에서 통계적으로 유의하다. 이는 자녀가 초등학교 고학년부터 고등학생일 경우 기혼 여성이 노동 시장에 참여하려는 비율이 높아진다는 것을 의미한다. 또한 이 결과는 자녀가 자람에 따라, 학교, 학원 등 가정 외에서 보내는 시간이 많아지게 되며 기혼여성이 갖는 여가시간이 늘어나 노동참가에 긍정적인 영향을 미친다고 해석할 수 있다. 또한 학년기 자녀의 경우 학원, 과외비등을 포함한 사교육비용의 증가는 기혼여성이 일을 하도록 하는 것에 긍정적인 영향을 미치는 큰 요인으로 볼 수 있다.

시·도별 여성의 평균 실업률의 경우 세 결과 모두 음의 부호를 갖으며, Pooled Probit에서는 유의하지 않으나 그 외 두 모형에서는 1% 유의수준에서 통계적으로 유의하다. Pooled Probit분석 결과가 유의하지 않은 것은 거주지역을 기준으로 시·도별 Cluster Effect를 고려한 Pooled Probit(고정효과분석)을 사용하여 시·도별 여성의 평균 실업률이 이미 지역더미를 통해 통제되었기 때문이라고 해석해볼 수 있다. 시·도별 여성의 평균실업률이 기혼 여성이 노동 시장에 참가하는 것에 부정적인 영향을 미치는 것은 우리가 미리 예상했던 결과와 일치한다. 이는 실업률이 높을 때, 노동 시장에 참여하는 것을 포기하거나 혹은 원하더라도 할 수 없는 상황을 의미한다.

시·도별 여성의 평균 임금은 세 모형의 결과에서 모두 유의하지 않은 결과를 보였다.

나. 노동공급식

1) 표본선택편의 고려하지 않은 경우

Heckman의 2단계 추정에 앞서, 표본선택을 고려하지 않은 Pooled OLS(IV)를 사용한 결과를 보는 것은 의미가 있다. 이는 표본선택을 고려할 경우와 비교하는 것이 가능하며, 실제로 우리의 자료가 얼마만큼의 차이를 보이는지를 알 수 있기 때문이다.

<표 11> 표본선택편의를 고려하지 않은 경우 도구변수 추정법

	(1) 경력변수 미포함		(2) 경력변수 포함	
lwage	-211.80	(-2.25)**	-115.61	(-2.85)***
othinc	-1.00	(-1.34)	-1.24	(-1.64)
edu	-12.09	(-1)	-23.29	(-3.02)***
age	5.95	(2.89)***	5.43	(2.47)**
c3num	-47.53	(-1.15)	-74.42	(-1.22)
c410num	-65.96	(-2.15)**	-50.73	(-1.23)
c1118num	15.06	(0.54)	30.97	(0.95)
관측치수	3545		2758	
연도더미	포함		포함	
지역더미	포함		포함	

주: () 안은 t-value 이며, *,**,***은 각각 10%, 5%, 1% 수준에서 유의함을 뜻함.

<표 11>은 표본선택을 고려하지 않은 것으로 노동참가결정식에 추정된 Mill's Inverse Ratio를 이용하지 않았다. 단, 경력변수의 내생성 외부를 검증하기 위해서 경력변수를 포함한 도구변수 집합과 포함하지 않은 집합을 이용하여 추정해보았다. 이론에 따르면, 도구변수로 쓰인 경력변수가 내생적이라면 두 예측치가 상이하게 된다.

만약, 표본선택편의가 있다면 위의 <표 11>의 결과는 편의 또는 불일치성을 갖게 되므로 여기에서 이에 대하여 자세하게 분석하는 것은 의미가 크지 않을 수 있다. 그러므로 자세한 설명은 표본의 편의를 교정한 결과의 비교와 함께 하도록 하겠다.

2) 표본선택편의 고려: Heckman의 2단계 추정법

(1) Pooled Probit Estimates 이용한 경우¹³⁾

표본선택편의를 고려하기 위하여 우리는 Heckman의 2단계 추정법을 사용하고자 한다. 이를 위

13) 각 방법은 Heckman의 2단계 추정법 중 1단계에 해당하는 노동참가결정식의 예측에 쓰이는 방법을 의미함.

해 1단계로 노동참가결정식을 예측하며 2단계에서는 노동공급식을 예측한다. 1단계에서 예측된 노동참가결정식은 <표 10>에 나타난 결과이다. 이때 1단계에서 Mill's Inverse Ratio를 추정하고, 이는 2단계의 노동공급식에서 설명변수로 쓰이며 표본선택편의를 교정해주는 역할을 한다.

<표 12> 표본선택 고려하는 경우 - (1단계: Pooled Probit estimates, 2단계: 도구변수 추정법)

	(1) 경력,ROE 미포함		(2) 경력변수 포함		(3) ROE 포함		(4) 경력, ROE 포함	
lwage	-225.89	(-2.85)***	-130.55	(-3.29)***	-296.97	(-2.58)***	-263.70	(-3.82)***
othinc	-11.74	(-2.53)**	-13.08	(-2.82)***	-13.74	(-2.20)**	-14.12	(-2.32)**
edu	-0.13	(-0.01)	-10.27	(-1.20)	2.21	(0.19)	0.22	(0.02)
age	11.95	(3.65)***	11.82	(3.51)***	11.76	(2.44)**	12.06	(2.60)***
c3num	-1055.50	(-2.58)***	-1143.89	(-2.76)***	-1244.63	(-2.33)**	-1267.87	(-2.41)**
c410num	-195.78	(-3.12)***	-189.48	(-2.99)***	-225.93	(-2.58)***	-218.74	(-2.56)**
c1118num	441.86	(2.45)**	488.71	(2.69)***	455.14	(1.86)*	469.15	(1.97)**
λ	2721.38	(2.46)**	2925.57	(2.60)***	2814.11	(1.89)*	2870.01	(1.96)*
관측치수	2721		2721		1724		1724	
연도더미	포함		포함		포함		포함	
지역더미	포함		포함		포함		포함	

주: () 안은 t-value 이며, *,**,***은 각각 10%, 5%, 1% 수준에서 유의함을 뜻함.

<표 12>를 4가지로 분류한 것은 임금의 내생성 교정을 위한 도구 변수로 사용된 경력변수의 내생성 검증과 ROE의 도구변수로의 적절성에 관한 논의를 위해서이다. 먼저, (1)의 경우, 경력, ROE를 도구변수집합에 포함시키지 않은 것이며, (2)는 경력변수만을 포함, (3)은 ROE만을 포함, (4)는 경력변수와 ROE변수를 모두 포함시킨 것이다. 연도별 특성과 지역별 특성을 고려하기 위하여 연도 더미와 지역 더미는 4가지 경우에 모두 포함시켰다.

먼저, 기존 연구에서 지적되고 있는 경력변수의 내생성에 관하여 우리의 자료를 테스트하기 위해서 (1)과 (2)를 비교하는 것이 필요하다. 두 결과는 예측된 계수의 부호와 통계적 유의성이 크게 다르지 않다. 이를 보다 정확하게 테스트하기 위해서 Hausman Test가 가능한데, 심지어 20%의 신뢰구간에서도 두 모형의 예측된 계수가 같다는 귀무가설을 기각하는 것에 실패했다. 즉, 1단계에서 Pooled Probit을 사용하고 2단계에서는 도구변수법을 사용할 경우, 경력변수가 내생성을 갖고 있다고 보기 어려웠다. 결과적으로 경력변수를 도구변수 집합에 넣는 것이 가능할 것으로 보인다. ROE변수의 적절성을 보기 위해 시험한 (3)과 (4)의 결과를 보면 이 결과는 앞의 (1), (2)와 크게 다르지 않다. 단, 교육 변수가 (1), (2)에서는 음의 부호이나 ROE변수를 도구변수집합에 포함시키면 양의 부호를 갖는다는 것이 차이점이다. 그러나 4가지 모두 통계적으로 유의하지 않기 때문에 이를 비교하는 것은 큰 의미를 갖지는 않는다고 할 수 있다. ROE가 일하는 기혼여성 표본에 모두 이용 가능하지 않아 표본선택편의를 일으킬 수는 있지만, 최소한 외생적인 변수로 생각된다. 단, 여기에서 Over-identification Test가 행해질 수 있으며, 본 논문에서는 이에 대한 Test는 생략하였다.

표본선택 편이가 있는지를 확인해보기 위해서는 Mill's Inverse Ratio가 유의한지를 보아야 한다.

<표 12>의 4가지 경우 모두 귀무가설, $H_0: \hat{\lambda}=0$ 을 대립가설을 기각할 수 있다. 이 때, 대립가설, $H_1: \hat{\lambda} \neq 0$ 에 대하여 양측검증으로 (1)은 5%, (2)(3)은 1%, (4)는 10%유의수준 하에서 기각할 수 있다. 이때 귀무가설을 기각한다는 것은 <표 11>의 표본선택편의를 교정하기 전 예측된 노동공급 곡선이 편의를 갖고 있다고 해석할 수 있다.

앞서 표본선택 편의를 교정하기 전인 <표 11>과 비교를 해보면 먼저 가장 큰 특징은 자녀의 수가 노동 공급에 미치는 영향이 상당히 커졌다는 것이다. <표 11>에서 만 0-3세의 자녀를 갖고 있는 기혼 여성은 약 40-70시간의 노동 시간을 줄였으나 표본선택편의를 고려한 후에는 연간 약 1000시간을 덜 일하는 것으로 분석되었다. 또한 통계적으로 유의하지 않던 결과가 유의수준 1%에서 유의해졌다. 타가구원의 소득효과 또한 표본선택편의 후 커졌으며 통계적으로 유의한 결과로 나타났다.

<표 12>의 결과를 해석해 보면, 시간당 임금이 1% 증가할 때, 여성들은 연간 약 130-293시간의 노동 공급을 줄인다고 볼 수 있으며 이는 1%의 유의수준에서 통계적으로 유의하다. 이때 여성의 임금탄력성이 음이라는 것은 재미있는 결과로 볼 수 있다. 왜냐하면 과거에는 여성의 임금탄력성은 항상 양일 것이라고 예상되었기 때문이다. 그러나 Heckman(1976)의 연구를 시작으로 여성의 노동 공급은 남성의 것과 같이 음이 나오는 것이 가능하다는 것을 보였다.

타가구원의 소득이 100만원 증가시, 여성은 연간 약 11-14시간의 노동을 덜하며 이는 우리가 예측했던 결과와 같다. 교육은 ROE를 포함시키느냐 시키지 않느냐에 따라 다른 결과를 보이거나 4가지 경우 모두 통계적으로 유의하지 않았다. 교육변수의 계수가 음의 부호를 갖는다는 것은 여성이 교육을 더 받을수록 연간 일하는 시간이 줄 것이라는 것이며 이러한 결과는 황윤재·최강식(1999)에서 교육수준의 상승이 의증임금 또한 상승시킬 수 있다고 하였으므로 이로 해석해볼 수 있겠다. 또한 사회적으로 교육을 더 받은 여성이 더 좋은 직업을 갖게 됨에 따라 실질임금이 상승하고, 이의 소득효과 일하는 시간을 줄일 수 있다는 것이다. 단 <표 11>의 결과로는 어느 쪽의 영향이 큰지를 확인할 수는 없다. 기혼 여성의 연령이 증가함에 따라 연간 약 11시간의 노동 시간을 증가시킨다고 볼 수 있다.

앞서 예상했던 것과 같이 기혼여성의 노동 공급에 큰 영향을 미치는 것은 자녀의 수로 볼 수 있는데 특히 만 0-3세의 자녀가 있는 경우, 연간 약 1000-1200시간의 노동공급을 줄인다. 이것은 만 4-10세 자녀의 수나 만 11-18 세의 자녀의 수의 영향에 비해 매우 크다고 볼 수 있다. 이 결과는 만 3세 미만의 자녀의 양육을 위한 적당한 보육 기관 등이 없다는 현실과도 일치한다. 만 4-10세의 자녀의 수는 역시 연간 노동 시간을 줄인다고 할 수 있다. 만 4세 이후 이용 가능한 보육 시설이 보다 손쉽게 이용 가능하기 때문에 만 0-3세의 아이의 수로 인한 효과보다 노동 시간의 감소를 적게 야기한다고 예상할 수 있다. 만 11-18세의 자녀를 갖고 있는 기혼 여성은 연간 약 400 시간의 노동 시간을 증가시키며, 이 결과는 자녀가 부모의 도움을 덜 필요로 한다는 것과 사교육비 등의 비용이 상당히 든다는 것으로 해석해볼 수 있다.

<표 13> 표본선택 고려하는 경우 - 1단계: Pooled Probit estimates, 2단계: 고정효과모형

	(1) 경력,ROE 미포함		(2) 경력변수 포함		(3) ROE 포함		(4) 경력, ROE 포함	
lwage	-824.30	(-2.92)***	-379.49	(-2.70)***	-844.82	(-2.37)**	-333.93	(-1.81)*
othinc	4.53	(1.26)	6.57	(2.04)**	-0.90	(-0.14)	3.90	(0.75)
edu	11.80	(0.30)	-6.21	(-0.17)	26.05	(0.41)	-15.36	(-0.28)
age	51.62	(1.78)*	9.51	(0.56)	58.04	(1.34)	0.59	(0.02)
c3num	412.02	(1.22)	663.91	(2.28)**	-250.13	(-0.40)	319.54	(0.67)
c410num	51.37	(0.62)	80.84	(1.05)	3.95	(0.03)	39.87	(0.36)
c1118num	-187.11	(-1.19)	-309.16	(-2.30)**	2.21	(0.01)	-237.02	(-1.12)
λ	-1285.98	(-1.45)	-1899.91	(-2.46)**	-145.09	(-0.10)	-1368.46	(-1.11)
관측치수	2721		2721		1724		1724	
연도더미	포함		포함		포함		포함	
지역더미	미포함		미포함		미포함		미포함	

주: () 안은 t-value 이며, *,**,***은 각각 10%, 5%, 1% 수준에서 유의함을 뜻함.

<표 14> 표본선택 고려하는 경우 - 1단계: Pooled Probit estimates 2단계: 확률효과모형

	(1) 경력,ROE 미포함		(2) 경력변수 포함		(3) ROE 포함		(4) 경력, ROE 포함	
lwage	-286.55	(-3.22)***	-135.09	(-2.96)***	-303.37	(-2.67)***	-254.00	(-3.60)***
othinc	1.46	(1.29)	1.47	(1.24)	0.18	(0.11)	0.41	(0.26)
edu	-8.74	(-0.71)	-27.27	(-3.03)***	-11.57	(-0.96)	-15.35	(-1.48)
age	5.44	(2.26)**	3.95	(1.63)	5.55	(1.79)*	5.67	(1.75)*
c3num	173.45	(1.69)*	223.39	(2.11)**	33.57	(0.21)	66.52	(0.43)
c410num	-20.85	(-0.52)	14.50	(0.38)	-33.85	(-0.56)	-13.49	(-0.25)
c1118num	-97.23	(-2.07)**	-101.77	(-2.06)**	-91.94	(-1.48)	-99.17	(-1.54)
λ	-590.18	(-2.38)**	-727.09	(-2.85)***	-581.95	(-1.61)	-656.65	(-1.87)*
관측치수	2721		2721		1724		1724	
연도더미	포함		포함		포함		포함	
지역더미	미포함		미포함		미포함		미포함	

주: () 안은 t-value 이며, *,**,***은 각각 10%, 5%, 1% 수준에서 유의함을 뜻함.

<표 13>과 <표 14>는 <표 12>와 마찬가지로 1단계에서는 Pooled Probit을 사용하였으나 2단계에서 사용한 방법론이 다르다. <표 13>의 경우, 각 개인의 특성을 고려한 고정 효과모형을 사용하였으며 <표 14>는 확률효과모형을 사용한 결과이다. <표 13>과 <표 14>의 결과를 보면, 만 0-3세의 자녀의 수가 오히려 노동 공급 시간을 증가시킬 수 있으며 심지어 몇몇 경우에는 통계적으로 유의했다. 이는 기존의 연구의 결과나 우리가 예측한 결과와 매우 상반되는 것으로 2단계에서 사용한 고정효과분석과 확률효과분석의 한계를 의심해볼 수 있다. 고정효과분석의 경우, 각 개인에게 하나씩의 더미를 주므로 약 2000개-3000개의 더미를 형성하게 된다. 확률효과분석의 경우에는 미관측 요소와 설명변수가 상관관계를 갖지 않는다는 가정 하에 GLS를 이용한 것이다. 이러한 두 모형은 미관측 요소(Unobserved Effect)를 심각하게 고려할 때, 의미 있는 모형이다. 만약, 미관측 요소가 없거나 다른 변수에 의해 통제될 수 있다면 Pooled OLS(우리의 모델에서는 도구변수법)이 더욱 Robust한 결과를 예측할 수 있다.

(2) GEE population-averaged model를 이용한 경우

두 번째로, Heckman의 2단계 추정법 중, 1단계에서 Population Average Model을 사용한 예측 결과에 대하여 논의하고자 한다. Population Average Model은 우리의 Sample을 대상으로 Partial Effect를 평균을 내어 계산하는 방법을 의미한다. (1)의 Pooled Probit을 사용할 때의 방식과 같게, 이때 계산된 Mill's Inverse Ratio가 2단계에서 설명변수로 쓰인다. 이것이 표본선택편의를 교정하는 역할을 한다. 아래의 <표 15>, <표 16>, 그리고 <표 17>은 각각 2단계에서 도구변수 추정법, 고정효과분석, 확률효과분석 모형을 사용한 것에 따른 결과이다.

<표 15> 표본선택 고려하는 경우 - 1단계:GEE population-averaged model 2단계:도구변수 추정법

	(1) 경력,ROE 미포함		(2) 경력변수 포함		(3) ROE 포함		(4) 경력, ROE 포함	
lwage	-229.66	(-2.90)***	-132.70	(-3.34)***	-299.67	(-2.61)***	-265.50	(-3.85)***
othinc	-8.16	(-1.95)*	-9.57	(-2.34)**	-11.36	(-2.06)**	-11.68	(-2.18)**
edu	-3.94	(-0.38)	-14.33	(-1.76)*	-1.07	(-0.10)	-3.20	(-0.32)
age	10.36	(3.31)***	10.29	(3.21)***	10.76	(2.4)**	11.03	(2.57)***
c3num	-726.59	(-2.02)**	-820.28	(-2.27)**	-1027.65	(-2.24)**	-1045.57	(-2.31)**
c410num	-153.60	(-2.72)***	-147.49	(-2.58)***	-196.52	(-2.41)**	-188.32	(-2.38)**
c1118num	298.87	(1.84)*	348.09	(2.17)**	358.50	(1.67)*	370.27	(1.77)*
λ	1822.34	(1.88)*	2040.44	(2.08)**	2208.06	(1.72)*	2248.90	(1.77)*
관측치수	2721		2721		1724		1724	
연도더미	포함		포함		포함		포함	
지역더미	포함		포함		포함		포함	

주: () 안은 t-value 이며, *,**,***은 각각 10%, 5%, 1% 수준에서 유의함을 뜻함.

<표 16> 표본선택 고려하는 경우 - 1단계:GEE population-averaged model 2단계:고정효과모형

	(1) 경력,ROE 미포함		(2) 경력변수 포함		(3) ROE 포함		(4) 경력, ROE 포함	
lwage	-795.43	(-2.84)***	-367.00	(-2.55)**	-813.51	(-2.27)**	-307.11	(-1.59)
othinc	-0.41	(-0.11)	1.20	(0.33)	-5.11	(-0.89)	-2.36	(-0.46)
edu	14.80	(0.38)	-1.98	(-0.05)	26.89	(0.43)	-11.07	(-0.20)
age	58.77	(2.06)**	18.94	(1.09)	62.67	(1.54)	9.78	(0.39)
c3num	-24.12	(-0.07)	185.22	(0.57)	-638.11	(-1.15)	-240.66	(-0.51)
c410num	-4.78	(-0.06)	24.35	(0.31)	-39.08	(-0.32)	-19.86	(-0.18)
c1118num	-12.72	(-0.08)	-109.26	(-0.74)	141.37	(0.59)	-9.08	(-0.04)
λ	-102.54	(-0.11)	-598.30	(-0.70)	824.91	(0.61)	111.38	(0.09)
관측치수	2721		2721		1724		1724	
연도더미	포함		포함		포함		포함	
지역더미	포함		포함		포함		포함	

주: () 안은 t-value 이며, *,**,***은 각각 10%, 5%, 1% 수준에서 유의함을 뜻함.

<표 17> 표본선택 고려하는 경우 - 1단계:GEE population-averaged model 2단계:확률효과모형

	(1) 경력,ROE 미포함		(2) 경력변수 포함		(3) ROE 포함		(4) 경력, ROE 포함	
lwage	-278.84	(-3.09)***	-135.09	(-2.96)***	-309.45	(-2.62)***	-250.80	(-3.47)***
othinc	-1.85	(-0.57)	1.47	(1.24)	-6.28	(-1.41)	-5.92	(-1.35)
edu	-6.12	(-0.50)	-27.27	(-3.03)***	-5.70	(-0.47)	-10.12	(-0.94)
age	7.64	(2.60)***	3.95	(1.63)	9.33	(2.39)**	9.34	(2.35)**
c3num	-127.92	(-0.44)	223.39	(2.11)**	-551.42	(-1.40)	-502.92	(-1.29)
c410num	-57.21	(-1.12)	14.50	(0.38)	-102.36	(-1.41)	-78.42	(-1.14)
c1118num	26.84	(0.21)	-101.77	(-2.06)**	157.22	(0.93)	144.63	(0.86)
λ	206.69	(0.26)	102.97	(0.13)	993.13	(0.94)	876.11	(0.84)
관측치수	2721		2721		1724		1724	
연도더미	포함		포함		포함		포함	
지역더미	포함		포함		포함		포함	

주: () 안은 t-value 이며, *,**,***은 각각 10%, 5%, 1% 수준에서 유의함을 뜻함.

1단계에서는 Population Average Estimation을, 2단계에서 도구변수법을 사용한 결과인 <표 15>는 앞의 <표 12>와 매우 비슷한 양상을 띄고 있다. 또한 Mill's Inverse Ratio가 통계적으로 유의함에 따라 표본선택편의가 있다고 할 수 있다. 이와 비교해서 <표 16>, <표 17>의 경우에는 표본선택편의가 없다는 결과가 나오며 임금 변수 이외에 신뢰 구간을 10%로 해도 유의하지 않다는 결과를 볼 수 있었다.

(3) Random Effects Probit Model를 이용한 경우

마지막으로 1단계에서 고려할 수 있는 방법은 Random Effects Probit Model이다. Random Effects Probit Model에서는

<표 18> 표본선택 고려하는 경우 - (3-1) Random-effects probit regression 이용 도구변수 추정법

	(1) 경력,ROE 미포함		(2) 경력변수 포함		(3) ROE 포함		(4) 경력, ROE 포함	
lwage	-234.70	(-2.98)***	-135.77	(-3.43)***	-303.16	(-2.64)***	-266.80	(-3.84)***
othinc	-3.54	(-1.48)	-4.20	(-1.78)*	-5.68	(-1.82)*	-5.79	(-1.87)*
edu	-8.96	(-0.86)	-20.26	(-2.65)***	-7.57	(-0.70)	-9.96	(-1.06)
age	15.66	(2.03)**	15.49	(1.97)**	16.77	(1.66)*	16.80	(1.66)*
c3num	-416.49	(-1.48)	-446.55	(-1.56)	-641.47	(-1.73)*	-638.54	(-1.72)*
c410num	-17.54	(-0.28)	1.12	(0.02)	-34.02	(-0.37)	-24.29	(-0.28)
c1118num	175.01	(1.29)	197.43	(1.44)	204.18	(1.15)	207.03	(1.17)
λ	742.53	(1.30)	775.28	(1.32)	873.95	(1.16)	860.79	(1.14)
관측치수	2721		2721		1724		1724	
연도더미	포함		포함		포함		포함	
지역더미	포함		포함		포함		포함	

주: () 안은 t-value 이며, *,**,***은 각각 10%, 5%, 1% 수준에서 유의함을 뜻함.

<표 19> 표본선택 고려하는 경우 - (3-2) Random-effects probit regression 이용 고정효과모형

	(1) 경력,ROE 미포함		(2) 경력변수 포함		(3) ROE 포함		(4) 경력, ROE 포함	
lwage	-800.52	(-2.86)***	-367.98	(-2.56)***	-819.58	(-2.28)**	-305.87	(-1.59)
othinc	0.15	(0.06)	1.02	(0.49)	-1.31	(-0.40)	0.03	(0.01)
edu	15.21	(0.39)	-0.38	(-0.01)	26.02	(0.42)	-10.35	(-0.19)
age	56.71	(1.95)*	15.80	(0.89)	56.47	(1.39)	3.01	(0.12)
c3num	59.60	(0.20)	253.59	(0.98)	-299.97	(-0.66)	35.32	(0.09)
c410num	-27.34	(-0.33)	-43.93	(-0.56)	-10.57	(-0.08)	-50.53	(-0.44)
c1118num	-51.32	(-0.36)	-144.83	(-1.15)	-6.79	(-0.03)	-134.80	(-0.75)
λ	-252.92	(-0.44)	-596.72	(-1.16)	-85.11	(-0.10)	-494.99	(-0.68)
관측치수	2721		2721		1724		1724	
연도더미	포함		포함		포함		포함	
지역더미	포함		포함		포함		포함	

주: () 안은 t-value 이며, *,**,***은 각각 10%, 5%, 1% 수준에서 유의함을 뜻함.

<표 20> 표본선택 고려하는 경우 - (3-3) Random-effects probit regression 이용 확률효과모형

	(1) 경력,ROE 미포함		(2) 경력변수 포함		(3) ROE 포함		(4) 경력, ROE 포함	
lwage	-277.37	(-3.07)***	-132.92	(-2.90)***	-310.31	(-2.62)***	-250.68	(-3.46)***
othinc	-0.44	(-0.23)	-0.42	(-0.22)	-2.64	(-1.02)	-2.30	(-0.90)
edu	-7.02	(-0.58)	-23.93	(-2.94)***	-8.74	(-0.75)	-12.88	(-1.26)
age	5.28	(0.81)	2.46	(0.38)	8.55	(0.98)	7.38	(0.85)
c3num	17.98	(0.08)	82.66	(0.35)	-247.32	(-0.76)	-185.92	(-0.59)
c410num	-58.73	(-1.17)	-41.87	(-0.83)	-50.13	(-0.71)	-39.61	(-0.58)
c1118num	-37.82	(-0.35)	-51.62	(-0.48)	27.63	(0.19)	7.91	(0.05)
λ	-147.43	(-0.31)	-282.37	(-0.59)	116.64	(0.18)	0.35	(0.00)
관측치수	2721		2721		1724		1724	
연도더미	포함		포함		포함		포함	
지역더미	포함		포함		포함		포함	

주: () 안은 t-value 이며, *,**,***은 각각 10%, 5%, 1% 수준에서 유의함을 뜻함.

<표 18>, <표 19>, 그리고 <표 20>은 1단계에서 Random Effects Probit 분석을 한 결과이다. 세 가지의 결과 모두 임금에 대한 효과 이외의 대부분의 계수는 통계적으로 유의하지 않은 결과가 나왔으며 Mill's Inverse Ratio도 유의하지 않은 것으로 나왔다. 그러나, 표본선택편의를 고려하지 않아도 된다는 판단에 앞서서 Random Effects Probit Estimation을 위한 몇 가지의 가정이 성립하지 않을 가능성을 의심해보아야 한다.

단, 2단계에서 도구변수법을 사용한 <표 18>의 경우, 유의하지는 않으나 대부분의 계수가 우리가 앞에서 예측한 결과 <표 12>나 <표 15>와 일치하였다. 앞에서 언급한 바와 같이 2단계에 사용한 고정효과분석과 확률효과분석이 갖고 있는 한계가 크며, 1단계에 사용하는 방법론과 관계없이 유사한 결과를 볼 수 있었다.

V. 결론

본 연구는 한국 기혼 여성의 노동 참여 결정과 노동공급곡선에 대한 연구로써 2003년부터 2007년까지 5년간의 한국노동패널데이터를 이용하였다. 표본선택편의, 임금 변수의 내생성 등의 계량적 문제를 교정하기 위한 다양한 방법을 사용하였다. 표본선택편의를 위해 Heckman의 2단계 추정법이 이용되었으며 임급의 내생성의 교정을 위해 도구변수법이 사용되었다. 본 논문은 횡단면 분석이나 Pooled Cross Section Data가 아닌 Panel Data로 분석을 하였다는데 의미가 있다. 또한 패널 데이터를 다루기 위한 다양한 방법론을 사용하였다.

기존 문헌에서 지적되고 있는 도구 변수 집합 중, 경력변수의 외생성 여부를 검증하였으며 ROE, 시·도별 여성의 평균 임금 등 새로운 변수의 사용· 테스트했다. 특히 경제학에서 가정하는 대표적 개인을 고려하여 시·도별 Cluster Sample을 다루었다는 점은 새로운 시도라고 볼 수 있다.

예측된 결과를 보면, 여성의 노동참가결정에서는 타가구원의 소득과 만 0-3세의 자녀의 수, 만 4-10세의 자녀의 수, 실업률이 부정적인 영향을, 만 11-18세의 자녀의 수는 긍정적인 영향을 갖고 있었다. 여성의 노동공급식의 분석에서는 표본선택편의를 고려한 Heckman의 2단계 추정법의 결과로 1단계에서 Pooled Probit, Population Average Estimation 과 Random Probit Estimation이 사용되었으며 2단계에서 시·도 별 Cluster Effect를 고려한 도구변수법, 고정효과분석, 그리고 확률효과 분석을 사용하였다. 각각의 경우에 대하여 모두 고려하여 추정한 결과, 1단계에서 사용한 방법의 종류와 무관하게 2단계에서 사용된 고정효과모형과 확률효과분석의 결과는 기존 연구 및 Cluster Effect를 고려한 도구변수법의 결과와 매우 다르다는 것을 알 수 있었다. 즉, 개인별 미관측 변수(Unobserved Effect)를 고려하기 위한 방법을 사용하기 위해 개인별 더미를 사용하거나 미관측 변수와 설명변수가 상관관계를 갖지 않는다는 가정하의 GLS를 사용하는 것이 가능하지만 이 역시 한계가 있을 수 있다는 것이다. 오히려, 각 개인의 특성보다는 시·도별의 여성을 중심으로 하여 고정효과분석을 한 결과가 기존 연구의 결과와 같은 방향을 갖는 것으로 볼 수 있다.

만약 2단계에서 시·도별 특성과 연도별 특성을 고려한 도구변수법이 적절한 방법이라고 한다면 기혼 여성은 자녀의 수, 임금, 타가구원의 소득이 여성의 노동 공급 곡선에 중요한 변수이다. 경력 변수는 우리의 분석에서는 내생적이라고 할 수 없었다. 그리고 새로운 도구 변수로 사용한 ROE는 이론상으로 보다 좋은 도구 변수를 찾고자하는 시도와는 일치하지만 자료의 한계로 인한 문제를 갖고 있었다고 할 수 있다.

본 연구의 결론으로는 기혼여성의 노동공급에 있어서 중요한 요소는 자녀의 수, 특히 어린자녀의 수라는 결론을 내릴 수 있었다. 그 외에도 타가구원의 소득 또한 노동공급에 중요한 영향을 미칠 수 있는 요소이기는 하다. 하지만, 노동정책적인 차원에서 볼 때 타가구원의 소득은 이미 결정된 측면이고, 정책적 고려를 할 수 있는 부분은 어린 자녀에 대한 보육비용이라고 생각할 수 있다. 국가차원에서 30대 전후반의 고학력의 여성은 가장 손쉽게 이용할 수 있는 유희노동력인데 반하여 그 이용이 이루어지지 못하고 있는 실정이다. 이런 상황을 고려할 때 이들의 보육비용 감소와 출산 및 육아로 인한 경력의 단절을 이들이 다시 노동시장에 진입할 때 걸림요소가 되지 않도록 배

려하는 정책이 요구된다고 할 수 있다.

참고문헌

- 김정연, 「한국에서의 기혼여성 노동공급에 관한 연구」, 1998.
- 김현숙·원종학, 「여성인력공급과 조세 재정정책: 자녀보육비용을 중심으로」, 한국조세연구원, 2004년.
- 김현숙·성명재, 「자녀세액공제제도 도입이 기혼여성 노동공급에 미치는 영향」, 한국조세연구원, “공공경제” 제 12권 제 1호, 2007년 4월.
- 최형재, 「자녀교육과 기혼여성의 노동공급」, 한국노동연구원, 2008년.
- 황윤재·최강식, 「기혼여성의 노동공급 결정요인: 비모수적 접근」, 계량경제학보, 1999년 6월.
- 황수경, 「기혼여성의 노동공급과 단시간근로」, 한국노동연구원, 2004년.
- Ashenfelter, O and A. B. Krueger, "Estimates of the Economic Return to Schooling from a New Sample of Twins," *American Economic Review*, Dec. 1994, pp. 1157-1173.
- Barry, K. G. and T. H. Matthew, "Parametric and Semi-parametric Modeling of the Off-Farm Labor Supply of Agrarian Households in Transition Bulgaria," *American Journal of Agricultural Economics*, Vol.84, No.1, Feb. 2002.
- Becker, G.S., "A Theory of the Allocation of Time," *The Economic Journal*, Vol. 75, No. 299, Sep. 1965.
- Blau, F.D. and L.M.Kahn, "Changes in the Labor Supply Behavior of married women: 1980-2000," IZA DP No.2180, Jun. 2006.
- Bruce, D. M. and D.T. Rosenbaun, "Welfare, The Earned Income Tax Credit, and The Labor Supply of Single Mothers," *The Quarterly Journal of Economics*, Aug. 2001.
- Cogan, J.F., "Fixed Costs and Labor Supply," *Econometrica*, Vol.49, No. 4, Jul. 1981.
- Eissa, N and J. B. Liebman, "Labor Supply Response to the Earned Income Tax Credit," *The Quarterly Journal of Economics*, Vol.111, May. 1996.
- Geronimus A. T., and S. Korenman, "The Socioeconomic Consequences of Teen Childbearing Reconsidered," *The Quarterly Journal of Economics*, Nov. 1992.
- Heckman, J. J, "The Common Structure of Statistical Models of Truncation, Sample Selection Bias and Limited Dependent Variables and a Simple Estimator for Such Models," *Annual of Economic and Social Measurement*, 1976a.
- _____, "What Has Been Learned About Labor Supply in the Past Twenty Years?," *The American Economics Review*, Vol. 83, No. 2, May. 1993b.
- Hotz, V. J. and R. A. .Miller, "An Empirical Analysis of Life Cycle Fertility and Female Labor Supply," *Econometrica*, Vol. 56, No. 1, Jan. 1988.
- Hsiao, C., "Analysis of Panel Data," Cambridge university press, 2003.
- Lee, M., "Semi-parametric Estimation of Simultaneous Equations with Limited Dependent

- Variables: A Case Study of Female Labour Supply," *Journal of Applied Econometrics*, Vol. 10, No. 2, 1995.
- Mincer, J., "Labor Force Participation of Married Women: A Study of Labor Supply," Columbia university and national bureau of economic research, 1962.
- Mroz, T. A., "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumption," *Econometrica*, Vol. 55, No. 4, Jul. 1987.
- Ribar, D. C., "A Structural Model of Child Care and the Labor Supply of Married Women," *Journal of Labor Economics*, Vol. 13, No. 3, Jul. 1995.
- Schultz, T. P., "Women's Changing Participation in the Labor Force—A World Perspective," Working Papers. The World Bank, Dec. 1989.
- Wooldridge, J.M., *Econometric Analysis of Cross Section and Panel Data*, MIT Press, 2002a.
- _____, *Introductory Econometrics: A Modern Approach 4th Edition*, Thomson South-western, 2008b.