디지털산업정보학회 논문지 제6권 제4호-2010년 12월

패널 데이터모형을 이용한 지역별 취업자 수 결정요인 추정에 관한 연구

이 현 주* · 김 희 철**

Estimating the Determinants for employment number by areas : A Panel Data Model Approach

Yi, Hyun Joo · Kim, Hee Cheul

〈Abstract〉

Employment number by areas is composed of various factors for groups and time series. In this paper, we use the panel data for finding various variables and using this , we analyzed the factors that is major influence to employment number by areas. For analysis we looked at employment number by areas , the region for analysis consist of seven groups, that is, the metropolitan city(such as Busan, Daegu, Incheon, Gwangiu, Daejeon, Ulsan.) and Seoul. Analyzing period be formed over a 63 time points(2005.01.- 2010.03). We examined the data in relation to the employment number by occupational job, unemployment rate, monthly household income, preceding business composite index, consumer price index, composite stock price index. In looking at the factors which determine employment number by areas job, evidence was produced supporting the hypothesis that there is a significant negative relationship between unemployment rate and monthly household income the consumer price index. The consumer price index and composite stock price index are significant positive relationship, preceding business composite index is positive relationship, it are not significant variables in terms of employment number by areas job.

Key Words : Fixed Effect Model, One-Way Error Component Regression Model, Lagrange Multiplier

Ⅰ. 서론

1.1 연구 배경 및 목적

통계청 발표에 의하면 젊은이들이 일자리를 못 찾고 있는 반면 55세 이상 고령층의 취업은 늘고 있는 것으로 나타났다. 고령층이 한 직장에서 평균 근속한 기간도 매년 꾸준히 줄어 2010년에는 평균 20년에도 못 미친 것으로 조사됐다.

통계청은 30일 청년층(15~29세)과 고령층(55~79세)을 대상으로 지난 5월 경제활동 부가조사를 한 결과 이렇게 나타났다고 밝혔다.

청년층 인구는 전년대비 6만9000명(0.7%) 줄었고 취업자 수는 1만5000명(0.4%) 감소했다. 취업자보다 인구가

* 숭실대학교 사회복지학과
** 남서울대학교 산업경영공학과(교신저자)

더 크게 줄면서 청년층 고용률은 41.4%로 지난해(41.3%)에 비해 0.1%p 개선됐다. 청년층 고용률은 2006년 이래 4년 연속 하락하다가 5년 만에 높아진 것이다. 청년층 중 재학. 휴학자 비중은 51.3%(498만7000명)으로 증가추세를 이어갔으며 이는 20대의 재학. 휴학 비중 증가에 기인한 것이다. 특히 이들 청년층은 힘들게 구한 첫 직장에서도 오래 근무하지 못하는 것으로 나타났다. 첫 직장의 평균근속 기간은 19개월로 전년 대비 1개월 줄었다. 첫 일자리 이직 경험자의 이직사유는 근로여건 불만족(42.5%)으로 다른 요인을 압도했다.

청년층 비경제활동인구(541만8000명) 중 취업관련 시험준비자는 54만 명이며 이들은 주로 '일반직 공무원'시험을 준비했고 고시 및 전문직, 기능분야 및 기타 분야도 늘어나는 추세였다. 반면 고령층의 경우 인구가 전년대비 37만명(4.1%) 늘면서 취업자 수도 32만3000명으로 7.3% 증가했다.

고령층 인구 10명 중 6명은 여전히 장래에 근로를 희망한다고 답했으며 일을 원하는 이유는 '돈이 필요해서'가 가장 많았다. 지난 5월을 기준으로 55~79세 고령층이 가장 오래 근무한 일자리에서의 평균근속기간은 19년8개월로 조사됐다. 이 기간은 2005년에는 20년10개월이었으나 2008년 20년8개월, 2009년 20년3개월 등으로 꾸준히 감소하면서 조사 이후 처음으로 20년 미만으로 내려섰다. 이들이 오랫동안 일 해왔던 직장을 떠나는 가장 큰 이유는 사업부진, 조업중단, 휴업, 폐업 등으로 조사됐다[1].

이러한 다양한 취업자 수 특성들을 이용하여 취업자 수 결정요인을 추정하는 연구는 일반적으로 등분산성, 독립성, 정규성 등의 가정과 다중 공산성 등을 해소하려고 노력하는 회귀모형에 의존하게 된다. 이러한 회귀문제에 대한 문제점을 해소하는 연구는 많이 이루어져 있지만 다양한 취업자 수 결정요인에 대한 모형 평가의 문제도 제기 할 수 있으리라 판단된다.

본 연구에서는 취업자 수 결정요인을 추정, 분석함에 있어서 패널 데이터모형(Panel data model)을 적용하고자 한다[2]. 패널 데이터모형은 시계열 자료와 횡단면 자료에 비해 자료 분석 결과 측면에서 볼 때 많은 장점을 가지고 있다. 즉, ① 패널 데이터모형은 정보측면에서 시계열 측면과 횡단면 측면을 동시에 고려함으로서 시계열 측면이나 횡단면 측면만 분석하는 것보다 더 많은 유용한 정보를 제공할 수 있는 모형이다. ② 패널 데이터모형은 분석과정에서 시차변수를 횡단면 자료와 함께 사용함으로서 시계열 자료가 흔히 발생하기 쉬운 다중 공선성을 감소시켜 주는 모형이다. ③ 패널 데이터모형에 의한 분석은 자유도를 증가시킴으로서 모수 추정치의 효율성을 향상 시킬 수 있게 해준다[3]

본 연구에서는 다양하고 높은 유용성을 가진 패널자료모형을 적용하여 7그룹(6개 광역시, 서울)을 분석대상으로 하였다[4]. 분석대상으로 실업률, 월평균가계소득, 선행(경기)종합지수 소비자물가지수, 종합주가지수 과의 관계를 파악함으로써 지역별 취업자 수 에 대하여 유효하게 영향력을 행사하는 제 결정 요인에 대하여 조사, 분석하고자 한다. 이러한 접근은 국내의 지역별 취업자 수 패턴에 영향을 미치는 제 변인에 대한 이해를 증진시켜 주는 계기를 마련해 줄 것이라 기대된다.

1.2 최근 연구 동향

이 분야에서 채창균과 김태기[5]은 "대졸 청년층의 취업 성과 결정 요인 분석이라는 논문" 에서 가구소득이나 출신대학, 전공과 같이 대학생 스스로의 노력으로 바꾸기 어려운 요인들의 영향이 매우 큰 것으로 나타난 반면, 학교의 교육적 지원 노력이나, 재학 중 일자리 경험, 해외어학연수, 자격증 취득 등과 같은 취업준비 노력은 그다지 긍정적인 영향을 미치지 못하는 것으로 확인되었다고 설명하고 있다. 그리고 남기곤[6]은 프라빗(Probit) 분석을 통하여 미취업 결정요인에 대해 분석한 결과, 학력과 연령, 기혼 여부와 같은 인적 속성은 미취업 혹은 실업에 유의한 영향을 미치는 경향이 강하지만, 직업훈련 경험이나 거주 지역의 실업률 그리고 성장단계의 가정환

경은 미취업이나 실업에 유의한 영향을 미치지 못함을 확인하였다고 하였다.

또한 엄동욱[7]은 "중 고령자의 취업 결정요인 "논문에서 가구소득이 높을수록 취업확률이 낮아지고 가구지출이 많아질수록 취업확률이 높아진다고 설명하고 이를 통해 가구소득뿐만 아니라 가구지출도 중 고령자 취업에 중요한 요인이라는 점을 확인할 수 있다고 설명하고 있다.

이현주, 나윤수와 김희철[8]은 가계소비지출 요인을 추정한 결과 월평균가계소득과 주택담보대출은 정(+)의 영향을 미치는 유의한 변수로 나타나 가계소비지출에 큰 영향을 주는 것으로 나타났다고 하였다.

한편, 본 연구에 사용하고자 하는 패널 모형은 조사, 방법의 새로운 툴(tool)로서 그 성능이 매우 우수하여 국내에서도 제 연구의 조사, 분석, 검증의 도구로 이용되고 있다. 패널모형을 이용한 연구로는 권남훈・고상원[9], 최충익[3]), 김태구・서용건[2] 등의 논문을 들 수 있다. 권남훈・고상원은 기업 R&D투자에 대한 정부 직접 보조금의 효과에 대하여 연구하였다. 이 연구에서 "패널자료 분석은 개별기업들이 고유한 특성으로 인해 발생 할 수 있는 변수들 간의 연관성을 제거 할 수 있다는 이점을 가지고 있고 대개 고정효과 모형이나 1차 차분(First difference) 모형이 사용 된다"고 하였다. 그리고 최충익은 패널모형에 의한 도시지역 수해결정요인 분석에 대한 연구를 하였는데 이 연구에서 패널모형은 패널 데이터가 가지고 있는 다양하고 풍부한 정보들을 가장 효과적으로 추출해내는 분석기법으로서 계량경제학에서 가장 이상적인 분석기법이라고 하였다. 또, 김태구・서용건은 패널 데이터 모형을 적용한 호텔 외국인 객실 수요 결정요인 추정에 관한 연구를 하였는데 이 연구에서 패널 데이터 모형(Panel data model)은 시계열 자료와 횡단면 자료에 비해 패널 자료가 분석 결과 측면에서 더 많은 유용한 정보를 제공할 수 있다고 하였다.

Ⅱ. 모형 및 자료

2.1 실증분석을 위한 모형의 탐색

본 분석에서 설정된 지역별 취업자 수 결정요인 추정을 시행하기 위하여 패널 데이터 모형을 구축하였다. 패널 데이터분석은 시계열과정에서 발생되는 추정오차와 지역별 단위에서 발생되는 추정오차를 통제 할 수 있다는 장점을 가지고 있다고 하였다[10]. 이러한 패널 자료 분석도 회귀방정식을 기본으로 하고 있다. 회귀방정식을 설정 할 때 종속변수에 영향을 미치는 모든 변수를 고려할 수는 없다. 만약 모든 변수들을 고려된다 하더라도 그것이 가장 좋은 모형 혹은 최적의 모형이라고 판단하기 어렵다. 그러나 종속변수에 매우 중요한 영향변수인데도 모형에 포함되지 않으면 심각한 오류를 범 할 수 있다. 패널분석은 이러한 누락변수(Omitted variable)에 대한 한계를 극복하는데 큰 의의를 가지고 있다[10]. 패널모형을 일반적 선형모형으로 표현하면 다음과 같이 표현 된다[9-10].

$$Y_{it} = a + X_{it}\beta + \varepsilon_{it} \tag{1}$$

단, α는 절편, β는 계수, 오차항, $\varepsilon_{it} = \mu_i + \lambda_t + \nu_{it}$, $i(\text{그룹}) = 1, 2, \cdots, N$, $t(\text{시점}) = 1, 2, \cdots, T$, μ_i =관찰되지 않은 그룹 특성 효과, λ_t =관찰되지 않은 시간 특성 효과, ν_{it} =확률적 교란항.

이러한 오차항의 형태(가정)에 따라 Random Effect Model(임의효과모형)과 Fixed Effect Model(고정효과모형)로 구분된다. 즉, 지역(그룹)별 차이가 지역의 고유한 고정된 특성으로 모두 설명된다고 가정하는 모형이 고정효과모형이고 지역의 차이가 지역별로 임의로 분포되어 있다고 가정 하는 모형이 임의효과모형이 된다. 오차항의 고려방식에 따라 OWECR모형(One-Way Error Component Regression Model) 과 TWECR 모형(Two-Way Error Component Regression Model)로 구분

된다. 즉 오차항을 그룹 특성효과와 교란항으로 분해하는 OWECR 모형과 그룹 특성 효과 및 시간 특성 효과와 교란항으로 분해하는 TWECR 모형으로 구분 한다[2, 10].

2.2 패널 모형의 구조

패널 데이터 분석모형은 다음과 같이 회귀모형의 형식을 기본으로 한다.

$$Y_{it} = \beta_0 + \gamma_i + \beta_1 X_{1it} + \beta_2 X_{2it} + \cdots + \beta_k X_{kit} + \epsilon_{it} \quad (2)$$

위 식에서 Y_{it}는 종속변수로서 i 번째 그룹의 t 시점 관측치를 의미하며 γ_i는 각 그룹의 특성효과, X_{kit}는 설명변수, β_0는 상수항, β_k는 변수의 파라미터 값, ϵ_{it}는 확률적 오차항을 의미한다. 이 모형에서 임의효과모형과 고정효과모형으로 구분 할 수 있으며 두 모형이 형태는 같지만 모형 설정 가정에서 차이가 난다. 그룹 특성효과를 의미하는 γ_i가 고정효과모형에서는 시간에 따라 고정불변하다고 가정하지만 임의 효과모형에서는 확률적으로 변한다는 의미로 해석 할 수 있다[2, 10].

고정 효과 모형의 추정원리는 관찰되지 않은 어떤 이질적인 요소(Unobserved heterogeneity)에 의해 발생되는 잠재적 편의(Bias)를 패널 데이터를 활용함으로서 제거하는 방식이다[14].

즉, N개의 개인에 대하여 시계열을 부여하여 계수값을 추정하게 된다. 이 추정방법을 수식으로 표현하면 다음과 같다[3].

$$Y_{it-1} = \beta_0 + \gamma_i + \beta_1 X_{1it-1} + \beta_2 X_{2it-1} + \cdots + \beta_k X_{kit-1} + \epsilon_{it-1} \quad (3)$$

$$Y_{it} = \beta_0 + \gamma_i + \beta_1 X_{1it} + \beta_2 X_{2it} + \cdots + \beta_k X_{kit} + \epsilon_{it} \quad (4)$$

위의 두 식은 시계열을 달리하는 동일한 식이며 (4)식에서 (3)식을 빼주면 다음과 같이 표현된다.

$$\Delta Y_i = \beta_1 \Delta X_{1i} + \beta_2 \Delta X_{2i} \cdots + \beta_k X_{ki} + \Delta \epsilon_i \quad (5)$$

단, $\Delta Y_i = Y_{it} - Y_{it-1}$이고 γ_i는 시간에 따라서 변하지 않는 일정한 값을 가지게 되므로 차분과정에서 없어지게 된다. 따라서 패널 데이터 분석에서 고정효과모형은 기본적으로 오차항에 대한 분해(Decomposition)를 기본 개념으로 하고 있다.

2.3 연구모형 및 변수

우리나라의 지역별 취업자 수 결정요인은 복잡하고 그룹(지역)별, 시간별로 다양한 원인에 의해서 취업자 수 결정이 이루어 질 수 있으므로 그룹 특성을 모두 변수화하여 분석하기란 무리가 따른다. 따라서 선정된 변수들로는 설명이 되지 않는 눈에 보이지 않은 어떤 변수가 존재함을 고려 할 필요가 있다. 본 연구에서는 이러한 제반 사항을 고려한 다음과 같은 분석모형의 필요성이 요구된다.

$$ENBAJ = f(UR, HMI, PBCI, CPI, CSPI, u_i) \quad (6)$$

이때 위 식 (6)에서 *ENBA*(Employment number by areas job)은 지역별 취업자 수(단위: 천명) 이고 *UR* (Unemployment ratios)은 실업률 (%)을 의미하며 *HMI* (Household monthly income) 는 월평균가계소득(전국, 2인 이상 가구: 단위: 원), *PBCI* (Preceding business composite index)는 선행(경기)종합지수(2005=100), *CPI* (Consumer price index)은 소비자물가지수(단위 : 2005=100), *CSPI*(Composite stock price index)는 종합주가지수(연월말, 1980. 1.4=100). u_i 는 모형내의 변수들이 설명하지 못하는 확률적 교란항을 의미 한다

본 연구에서 사용된 독립변수의 선택 근거는 최근 연구 동향에서 남기곤[6]은 미취업 혹은 실업에 유의한 영향을 미치는 경향이 강하다고 하여 실업률을 하나의 독립변수로 채택하였다. 그리고 엄동욱[7]은 가구소득이 높

을수록 취업확률이 낮아지고 가구지출이 많아질수록 취업확률이 높아진다고 설명하고 있기 때문에 가구당 월평균 가계소득을 또 하나의 변수로 채택하였다. 서론에서 서술 한바와 같이 경기가 회복세를 보이면 일자리가 늘고 실업률도 낮아지기 마련이기 때문에 선행(경기)종합지수와 물가로서 소비자 물가지수를 각각 독립변수에 포함하고 종합주가지수는 우리나라 경제를 총체적으로 보여주는 지표이기 때문에 종합주가지수 하나만 보면 그 나라의 경제상황을 알 수 있기 때문에 하나의 독립변수로 채택하였다.

독립변수들과 종속변수와의 관계를 선형관계로 정리하고 이에 대한 패널 분석을 시행하기 위하여 고전적 회귀분석에서 자주 발생하는 문제점중의 하나인 이분산성(Heteroscedasticity)을 해소하고자 다음과 같은 로그형태 모형을 선택하였다.

$$\ln(ENBAJ) = \beta_0 + \beta_1 \ln UR + \beta_2 \ln HMI + \beta_3 \ln PBCI + \beta_4 \ln CPI + \beta_5 \ln CSPI + \mu_i \quad (7)$$

단, $ENBAJ$:지역별 취업자 수, β_0 :절편,

$\beta_i\ (i=1,2,3,4,5)$: 회귀계수, UR: 실업률, HMI: : 월평균가계소득, $PBCI$:선행(경기)종합지수, CPI:소비자 물가지수, $CSPI$: 종합주가지수, u_i:확률적 교란항.

2.4 자료의 성격

본 연구에서는 지역별 취업자 수 결정요인을 추정하기 위하여 사용된 자료는 지역별 취업자 수(단위: 천명)을 종속변수로 독립 변수는 실업률 (%), 월평균가계소득(전국, 2인 이상 가구: 단위: 원), 선행(경기)종합지수(2005=100), 소비자물가지수(단위 : 2005=100), 종합주가지수(연.월.말, 1980.1.4=100))을 사용하였다. 이러한 결정요인을 추정하기 위하여 분석기간은 2005년 1월부터 2010년 3월(63시점, 월별) 까지 선정하였다. 한편, 본 연구에서는 통계청 국가통계포털 사이트를 통해서 분기별 자료는 월별 단위로 확대 해석하여 사용하였다[11]. 종속변수로 사용된 7 그룹(6개 광역시, 서울)의 취업자 수는 <표 1>에 요약 되었다.

<표 1> 지역별 취업자 수(단위: 천명)

년 도.월	서울 특별시	부산 광역시	대구 광역시	인천 광역시	광주 광역시	대전 광역시	울산 광역시
2005.01	4796	1605	1170	1175	606	630	484
2005.02	4807	1611	1154	1159	606	628	486
2005.03	4874	1600	1161	1178	610	637	491
2005.04	4892	1591	1180	1194	618	647	496
2005.05	4912	1625	1190	1215	622	650	497
2005.06	4936	1631	1188	1221	624	656	498
2005.07	4908	1641	1202	1213	629	659	496
2005.08	4858	1615	1178	1206	620	641	487
2005.09	4877	1644	1183	1208	624	648	490
2005.10	4919	1640	1184	1215	626	651	496
2005.11	4960	1651	1188	1219	627	661	500
2005.12	4938	1631	1192	1196	617	645	492
2006.01	4897	1625	1175	1186	608	641	483
2006.02	4862	1623	1161	1199	604	638	484
2006.03	4883	1623	1171	1209	616	651	488
2006.04	4918	1615	1180	1236	620	657	499
2006.05	4927	1637	1185	1244	621	670	499
2006.06	4937	1629	1179	1256	624	670	496
2006.07	4914	1637	1178	1246	630	665	498
2006.08	4862	1621	1156	1230	616	660	491
2006.09	4893	1640	1174	1231	623	669	497
2006.10	4919	1640	1178	1240	623	672	498
2006.11	4948	1630	1174	1242	622	674	498
2006.12	4907	1610	1162	1219	620	665	497
2007.01	4888	1602	1161	1201	623	655	500
2007.02	4846	1598	1142	1209	626	647	503
2007.03	4883	1609	1151	1236	635	663	508
2007.04	4939	1623	1165	1251	634	679	515
2007.05	4979	1639	1170	1259	637	692	522
2007.06	4972	1644	1171	1271	637	688	521
2007.07	4987	1628	1177	1283	638	689	518
2007.08	4946	1600	1138	1262	626	683	516
2007.09	4975	1626	1150	1276	633	687	515
2007.10	4980	1611	1155	1287	637	697	521
2007.11	4978	1615	1161	1285	642	697	522
2007.12	4913	1599	1147	1271	633	686	516
2008.01	4910	1584	1139	1244	638	689	514
2008.02	4861	1578	1147	1240	632	685	512
2008.03	4933	1609	1129	1253	638	687	524

2008.04	4949	1611	1149	1275	639	696	531
2008.05	4970	1625	1151	1276	645	700	526
2008.06	4961	1621	1146	1270	644	701	531
2008.07	4972	1607	1155	1278	642	693	534
2008.08	4929	1584	1132	1258	626	687	527
2008.09	4917	1592	1131	1265	627	702	533
2008.10	4908	1569	1139	1286	637	706	533
2008.11	4897	1585	1147	1289	645	709	528
2008.12	4853	1596	1134	1252	637	694	522
2009.01	4838	1572	1123	1234	631	681	522
2009.02	4743	1557	1102	1237	623	689	517
2009.03	4819	1565	1111	1252	623	690	515
2009.04	4833	1567	1136	1292	639	699	518
2009.05	4855	1577	1140	1294	649	707	518
2009.06	4889	1590	1156	1330	662	710	523
2009.07	4862	1572	1165	1335	654	710	517
2009.08	4788	1554	1163	1310	654	704	523
2009.09	4803	1564	1173	1310	661	701	523
2009.10	4840	1560	1185	1312	661	694	529
2009.11	4879	1571	1170	1316	662	696	525
2009.12	4877	1576	1150	1292	651	691	525
2010.01	4784	1569	1138	1273	645	681	518
2010.02	4803	1560	1137	1271	650	679	514
2010.03	4846	1574	1164	1291	666	687	522

Ⅲ. 실증분석 결과

본 연구에서는 실증분석을 위하여 LIMDEP 8.0 소프트웨어를 이용하였다[12-14].

<표 2>의 추정 결과표에서 Lagrange Multiplier(LM) 검정 통계량을 적용하였는데 1%유의수준에서 OWECR모형은 10148.53값으로 높게 나타났다. 이러한 결과는 그룹 더미(Dummy) 없는 POLS(pooled ordinary least square model) 모형보다 OWECR모형 이 더 적합한 모형임을 의미한다. 또한 전체적인 계수 값의 유의성과 고정효과 모형과 임의효과모형 중 적합한 모형을 채택하기 위한 Hausman검정[12]을 통해 판단해야 되는데 데이터의 특수성(예를 들면 분석기간 동안 각 그룹에 같은 종합주가지수 등을 반영) 때문에 OWECR 모형 중에서 임의효과 모형과 고정효과모형 중에서 어떤 모형이 데이터 적합도가 우수한 모형인가를 판단하기는 무리가 따른다(추정된 계수도 거의 동일). 그러나 그룹공통 특성을 파악하는 것이 본 논문의 주요 관심사이기 때문에 본 연구는 OWECR모형의 고정 확률모형 중심으로 해석을 전개하고자 한다. 따라서 도출된 지역별 취업자 수 결정요인 추정 식은 <표 2>에 요약된 것처럼 다음과 같다.

$$\begin{aligned}\ln(\widehat{ENBA.I}) = \\ -0.0039\ln UR - 0.4723\ln HMI + 0.0080\ln PBCI \\ +0.0628\ln CPI + 0.0254\ln CSPI\end{aligned} \tag{8}$$

지역별 취업자 수 결정요인을 추정하는데 있어서 5개의 독립변수 중에서 실업률(유의수준 0.01), 월평균가계소득(유의수준 0.01) 각각 유의적인 음(-)의 영향을 나타내는 유의적인 변수로 나타났다. 즉 실업률이 증가하면 취업이 어렵고 즉, 실업률이 1% 증가하면 취업자 수는 대략 0.0039% 감소됨을 나타내고 있고 다른 요인으로 가구소득이 높을수록 취업확률이 낮아진다고 평가할 수 있다. 이는 사전연구에서 엄동욱[7]의 결과와 유사하고 사회변수인 소비자물가지수(유의수준 0.01)와 종합주가지수(유의수준 0.05)는 각각 유의적인 정(+)의 영향을 미치는 것으로 나타나고 즉, 소비자물가지수가 높으면 그 만큼 지출이 늘어나 취업의도가 살아나고 종합주가지수가 증가하면 그 만큼 취업자의 수가 증가 할 확률이 높음을 시사한다. 선행(경기)종합지수는 정(+)의 영향을 미치지만 지역별 취업자 수에는 큰 영향을 주지는 않은 것으로 나타났다.

본 연구에서 적용된 패널 데이터 모형 중에서 각 모형들 간의 차이 검증은 Log-Likelihood Ratio검정을 통하여 시행되었다. Log-Likelihood Ratio 검정에서 상수항만을 고려한 모형(Model 1), 그룹 개별특성효과만을 고려한 모형(Model 2), 설명변수만을 고려한 모형(Model 3), 설명변수와 개별 그룹 특성효과를 동시에 고려한 모형(Model 4)의 각 모형 중에서 어떤 모형이 우수한지를 평가 할 필요가 있다[14].

<표 2> 패널 데이터모형을 적용한 지역별 취업자 수

변 수	OWECR Model (FE Model)
	Coefficient(t statistics)
ln UR	-0.0039*** (-3.405)
ln HMI	-0.4723*** (-5.157)
ln $PBCI$	0.0080 (0.957)
ln CPI	0.0628*** (5.681)
ln $CSPI$	0.0254** (2.414)
No. of obs.	441
No. of Grs.	7
df	429
R^2 ($adj\ R^2$)	0.999(0.998)
χ^2 Statistic	3056.16***
Log-likelihood	1930.509
Akaike Info. Criterion	-11.5386
LM	10148.53***

주) OWECR Model: one-way error component regression model, LM:group effects test of H_0(pooled) vs. H_1(unpooled), FE Model: fixed effect model, **$p<0.5$, ***$p<0.01$,

본 연구에서는 Log-Likelihood와 결정계수 R^2 (R-square)를 이용하여 평가하고자 한다. 평가한 결과는 <표 3>에 제시 하였다. 이 표에서 설명변수와 개별그룹 특성효과를 동시에 고려한 모형(Model 4)은 Log-Likelihood가 1930.509, 결정계수 R^2가 0.9986 으로 나타나 상대적으로 우수한 모형으로 평가 되었다. 즉, 설명변수 및 그룹특성효과를 동시에 고려한 모형인 패널자료모형이 우수하고 합리적임을 보여주고 있다.

<표 3> OWECR 모형의 모형별 우수성 평가결과

Model	Log-Likelihood	R-squared
(1) Constant term only	402.431	0.0000
(2) Group effects only	1847.967	0.9986
(3) X - variables only	441.193	0.1612
(4) X and group effects	1930.509	0.9990

주) OWECR Model: one-way error component regression model

3.2 OWECR 모형에 의한 그룹 특성효과 분석

<표 4> OWECR 모형에 의한 그룹 특성 효과 분석

Group(항목)	Coefficient	t-ratio
1(서울특별시)	3.04272***	14.32817
2(부산광역시)	2.90157***	13.66417
3(대구광역시)	2.85643***	13.45121
4(인천광역시)	2.86731***	13.50307
5(광주광역시)	2.76656***	13.02839
6(대전광역시)	2.77690***	13.07729
7(울산광역시)	2.73151***	12.86489

주) OWECR Model: one-way error component regression model, *** $p<0.01$.

본 연구에서 OWECR에 의한 그룹 특성효과를 <표 4>에 제시되었는데 추정된 계수값은 모든 계수가 유의적으로 나타났다. 추정된 계수값은 서울특별시(그룹의 계수: 3.04272)이 양(+)방향으로 가장 높다. 이러한 결과는 누락변수의 설명력이 서울특별시 그룹 계수가 3.04272가 된다는 의미이고 계수가 크다는 것은 본 연구에서 적용된 설명변수로서 지역별 취업자수 증가 양상을 설명하지 못하는 부분이 상대적으로 많음을 의미한다[14].

그러나 그룹특성 효과측면에서는 그룹특성효과 계수가 10을 초과하지 않으면 그룹특성효과가 있는 것으로 해석하기에는 무리가 따른다(Hasio, 2003 [15]) 결국 그룹 특성 효과는 거의 존재하지 않는다고 결론을 내려도 무방하다.

Ⅳ. 결론 및 한계

지역별 취업자 수 결정요인은 복잡하고 그룹(항목)별, 시간별로 다양한 원인에 의해서 취업자수결정이 이루어질수 있다. 본 연구는 지역별 취업자 수 추정을 위한 결정모형을 설정하고 패널 데이터 모형의 추정을 통해 기 취업자 수 결정요인을 파악하는데 연구의 목적을 두었다.

본 연구에서는 패널자료모형을 적용하여 7 그룹(6개 광역시, 서울)대상으로 하였고 분석기간은 2005년 1월부터 2010년 3월까지의 63시점 동안 의 월별자료 및 분기별 자료를 이용하였다.

5개의 독립변수 중에서 실업률(유의수준 0.01), 월평균 가계소득(유의수준 0.01) 각각 유의적인 음(-)의 영향을 나타내는 유의적인 변수로 나타났다. 즉 실업률이 증가하면 취업이 어렵고 다른 요인으로 가구소득이 높을수록 취업확률이 낮아진다고 평가할 수 있다. 그리고 사회변수인 소비자물가지수(유의수준 0.01)와 종합주가지수(유의수준 0.05)는 각각 유의적인 정(+)의 영향을 미치는 것으로 나타나고 즉, 소비자물가지수가 높으면 그 만큼 지출이 늘어나 취업의도가 살아나고 종합주가지수가 증가하면 취업자의 수가 높을 확률이 높아질 것이라고 판단된다. 선행(경기)종합지수는 정(+)의 영향을 미치지만 지역별 취업자 수에는 큰 영향을 주지는 않은 것으로 나타났다. 그룹특성 효과측면에서는 그룹특성 효과는 거의 존재하지 않는다고 결론을 내려도 무방하다.

본 연구에서는 복잡하고 다양한 취업자 수 결정요인에 대하여 관찰되지 않은 효과를 패널 데이터를 활용해 통제해줌으로서 누락변수에 따르는 통계적 위험을 일정부분 해소 할 수 있었고 복잡한 메카니즘을 가지고 있는 지역별 취업자 수 결정요인에 있어서 패널 데이터 분석이 얼마나 효율적인지를 그룹특성 분석과 모형별 우수성 평가를 통하여 단적으로 확인 할 수 있었다.

본 연구는 실증연구의 진행과정에서 다음의 한계성을 지니고 있다. 첫째, 사회 환경요인(예를 들면 기업 부패지수 등)도 지역별 취업자 수 결정요인에 직간접으로 영향을 줄 수 있기 때문에 이러한 사회 현상의 문제도 중요한 요소로 투입되어야 한다.

둘째, 모형 측면에서 고정효과모형과 임의효과모형 중 적합한 모형을 채택하기 위한 Hausman검정을 통해 판단해야 되는데 데이터의 특수성(예를 들면 분석기간 동안 각 그룹에 같은 종합주가지수를 반영) 때문에 OWECR모형 중에서 임의효과모형과 고정효과모형 중에서 적합한 모형 선택의 어려움을 해소하기는 무리가 따른다고 판단된다.

향후 관련연구의 방향은 위에서 기술한 한계들을 극복하면서 이루어 질 필요가 있으며 정책적인 활용도를 높이기 위해서는 보다 많은 실증연구가 이루어져야 하겠다.

참고문헌

[1] http://news.nate. com/view/20100729n13162?mid=n0300 (2010, 07, 29 방문).

[2] 김태구, 서용건, “패널 데이터 모형을 적용한 호텔 외국인 객실 수요 결정요인 추정,” 관광학 연구, 31권 59호, 2007, pp. 465-485.

[3] 최충익, “패널모형에 의한 도시지역 수해결정 요인 분석,” 국토계획, 39권 7호, 2004, pp. 49-67.

[4] http://www.kosis.kr/search/totalSearch2.jsp(통계청)

[5] 채창균, 김태기, “대졸 청년층의 취업 성과 결정 요인 분석,” 職業敎育硏究, Vol. 28, No. 2, 2009, pp. 89-107.

[6] 남기곤, “청년층 미취업자 특성에 관한 분석,” 산업노동연구, 제15권 2호, 2009, pp. 99-123.

[7] 엄동욱, “중 고령자의 취업결정요인,” 노동정책연구, 제8권 3호, 2008, pp. 17-38.

[8} 이현주, 나윤수, 김희철, “패널 데이터모형을 적용한 가계 소비 지출 결정요인 추정에 관한 연구,” 한국지식정보 기술학회 논문지, 제5권3호, 2010, pp. 123-181.

[9] 권남훈, “기업 R&D 투자에 대한 정부 직접 보조금의 효과,” 국제경제연구, 10권 2호, 2004, pp. 157-18.

[10] Green. W. H. LIMDEP : A user’s manual, Plainview. New York : Econometric Software. Inc, 1998.

[11] http://www.kosis.co.kr(통계청).
[12] Hausman, J. & Taylor, W., "Panel data and unobservable individual effects," Econometrica Vol. 49, No. 6, 1981, 1981, pp. 1377-1398.
[13] Ashenfelter, O., Zimmerman, D., & Levine, P. B, Statistics and econometrics: Methods and applications, New York : John Wiley & Sons, 2003.
[14] Green. W. H., Econometric analysis, New Jersey : Prentice-Hall, 2000.
[15] Hsiao. C., Analysis of Panel Data, New York : Cambridge University Press, 2003.

■ 저자소개 ■

이 현 주
Yi, Hyun Joo

2007년 숭실대학교 사회복지학과 박사 수료
2000년 강남대학교 사회복지학과 석사

관심분야 : 사회복지정책, 빈곤 정책, 사회보장, 발달장애청소년 교육정책
E-mail : yihyunjoo@paran.com

김 희 철
Kim, Heel Cheul

2005년 3월~현재
남서울대학교 산업경영공학과 교수
1998년 2월 동국대학교 통계학과(이학박사)
1992년 2월 동국대학교 통계학과(이학석사)

관심분야 : 소프트웨어 신뢰성공학, 웹 프로그래밍
E-mail : kim1458@nsu.ac.kr

논문접수일 : 2010년 11월 1일
수 정 일 : 2010년 11월 30일
게재확정일 : 2010년 12월 3일