



사망수준과 사망 원인관련 지표들 간의 관계에 대한 자료탐색 분석

이성용*

이 연구의 목적은 사망수준의 저하에 영향을 미치는 세 요소-사회경제적 발전, 공공 보건의 발달, 사회경제적 발달의 균등상태-의 상대적 중요성을 분석하는 것이다. 종속변수인 사망 수준의 지표로는 영아사망률과 출생시 기대수명 등 두 변수가 사용되었다. 국민총생산(GNP)은 사회경제적 발달지표로, 여성의 초등학교 취학률과 기니계수(GINI index)는 사회경제적 균등상태 지표로, 병원침대당 인구는 공공보건 지표로 간주되었다. 변수들에 대한 자료는 두 시점에 걸쳐 수집되었다. 하나는 1970년 이전 53개국에서, 다른 하나는 1970-80년대 55개국에서 수집되었다. 탐색적 자료 분석 방법이 통계 분석 방법으로 사용되었다. 이 기법은 종속변수와 독립변수와의 관계가 선형인지 아닌지, 그리고 우리 모형에서 어느 것이 유력 사례인지를 파악할 수 있는 장점이 있다.

분석결과에 따르면, 첫째로 영아 사망률과 세 요소의 관계가 선형이 아니라 비선형임이 밝혀졌다. 영아 사망률 저하에 국민총생산이 가장 많이, 여성의 초등학교 취학률이 두 번째, 기니계수가 그 다음으로 영향을 미치는 것으로 나타났다. 반면 병원침대당 인구는 통계적으로 유의미한 영향을 보여주지 않았다. 둘째, 출생시 기대수명은 여성의 취학률, 기니계수 등과 같은 변수와는 선형 관계를 가지는 반면 국민총생산 변수와는 비선형 관계를 가진다. 영아사망률 변수와는 달리 출생시 기대수명의 변이에는 여성의 초등학교 취학률이 국민총생산보다 더 커다란 영향을 미쳤다.

핵심단어: 자료탐색 분석, 사망 수준, 영아 사망률, 출생시 기대수명, 국민총생산, 기니계수, 여성의 초등학교 취학률, 병원침대당 인구수

I. 서론

1880년대 이래 일어난 급격한 사망수준 하강은 인류 역사상 가장 획기적인

* 강남대학교 교양교수부 전임강사

사건의 하나이다. 1880년 이후 오늘날까지 향상된 기대수명의 진보는 인류가 존재하여 1880년대까지 일어났던 기대수명의 진보를 훨씬 뛰어넘는다(Preston, 1976). 특히 1950년대 이래 일어난 개발도상국가의 사망수준 저하는 매우 놀라운 수준이다. 선진국가가 약 100년 내지 200년에 걸쳐 이룩했던 사망률 저하 수준을 일부 개발도상국가들은 약 30-40년 사이에 달성하였다.

사망수준 저하의 원인들을 밝히고 이해하는 것은 사망률 연구에서 매우 중요한 일이다. 사망률 저하의 원인으로는 (1) 사회경제적 발달(Mckweon, 1972, 1976; Preston, 1980), (2) 공공보건의 향상(Omran, 1973; Preston, 1980; Razzel, 1974), (3) 사회경제적 균등상태(Gwatkin, 1980; Palloni, 1981; Caldwell, 1986)가 주로 설명되어 왔다. 대부분의 학자들은 이 세 가지 요소들이 사망률 저하에 커다란 영향을 미쳤다는 사실에는 동의하지만, 어느 요소가 더 큰 영향을 미쳤는지에 대해서는 서로 다른 견해를 가지고 있다. 따라서 이 세 가지 요소들을 완벽하게 나타낸다고는 할 수 없지만 적합하다고 생각되는 경험적 지표들을 사용하여 어느 요소가 사망률 저하에 영향을 미쳤는지를 경험적으로 연구하는 것은 사망률 연구에서 바람직한 일이다.

현재까지의 사망률 연구 문헌들을 검토한 바에 의하면, 위의 세 요소 중 어느 것이 가장 사망률 저하에 영향을 미쳤는지를 경험적으로 분석한 연구는 없다. 세 요소들이 사망률 저하에 얼마나 상대적인 영향을 미쳤는지를 비교하기 위해, 본 논문에서는 그 요소들의 특성을 어느 정도 나타낸다고 할 수 있는 거시 지표들을 사용하여 탐색적 자료 분석을 하겠다. 본 연구의 거시 지표들은 두 시기에 걸쳐 수집되었다. 두 시기에 걸쳐 분석함으로써, 우리는 각 시기에 사망률 저하에 미치는 세 요소의 상대적 중요성이 변화했는지 아니면 그대로인지를 분석할 수 있다. 두 시기에 걸쳐 수많은 국가들에서 획득한 세 요소를 나타낼 수 있는 경험적 지표들로 사망률 저하에 미치는 상대적 중요성을 분석한 본 연구는 분명 사망 수준 저하에 미치는 세 요소들의 일반적 경향을 밝히는데 있어 도움을 줄 것이다.

많은 국가에서 수집된 경험적 지표들을 이용하여 사망수준 저하와 세 요소간의 관계를 회귀 분석을 할 때 두 가지 중요한 방법론적 문제가 고려되어야 한다. 하나는 독립변수와 종속변수간의 관계가 선형인지 비선형인지를 고려해야 한다. 이를테면 국제 자료들을 비교 분석한 프레스턴(Preston, 1975)은 국가간 개인 소득과 기대수명의 관계가 비선형임을 밝히고 있다. 따라서 본 연구는 세 요소들이 기대수명이나 영아사망률과 같은 사망수준 요인에 어떤 형태로 영향을 발견하기 위해 검정력 변형(power transformation) 기법을 사용한다. 둘째로 회귀 계

수는 비정상적인 값을 가진 한 두 개의 사례들로도 크게 영향을 받을 수 있기 때문에 이상값(outlier)이나 유력 사례(influential cases)를 회귀 계수의 추정에서 고려할 필요가 있다. 특히 표본의 수가 작을 경우에는 큰 경우보다 회귀 계수가 몇 개의 이상값이나 유력 사례에 민감하게 반응하기 때문에 이상값이나 유력 사례의 문제를 심각하게 고려해야 한다. 때때로 이상값이나 유력 사례들은 이론적이나 경험적인 함축을 내포할 수 있는데, 왜냐하면 그것들의 발견을 통해 모형이 제대로 특정화되었는지를 점검할 수 있기 때문이다(Bollen and Jackson, 1990). 때때로 우리는 이런 이상값이나 유력 사례의 발견으로 인해 변수간의 관계를 일반화하는데 고통을 받을 수 있다. 그러나 국제 지표들에 대한 통계 연구의 목표는 “변수들 간의 관계에 대한 일반화를 포괄적인 의미에서 서술하는 것이 아니라, 확률적인 의미로 밝히는(Jackson, 1985: 166)”데 있다. 결과적으로 우리는 이상값이나 유력 사례로 판정되는 예외적인 국가를 발견함으로써 변수들 간의 관계에 대한 버팀성(robustness)을 점검할 필요가 있다.

이런 관심 하에서 본 논문은 다음과 같은 연구 목표를 가진다. 첫째, 사망수준 저하에 미치는 요소들, 좀더 정확히 말해서 그 요소들과 관련된 네 개의 지표의 상대적인 중요성을 밝힌다. 둘째, 그 지표들과 영아사망률과 기대수명과 같은 사망률 요소 사이의 관계가 선형인지 비선형인지를 분석한다. 셋째, 이런 관계를 일반화하는데 있어 예외적인 국가가 되는 것은 어디인지를 밝힌다. 마지막으로 사망 원인 관련 지표들과 사망수준 요소들 사이에서 발견된 일반화 관계의 버팀성을 분석한다.

II. 사망수준 저하에 대한 이론적 설명

앞에서 말한 바와 같이, (1) 사회경제적 발달 (2) 공공 보건의 향상 (3) 사회경제적 균등 상태가 주로 사망률 저하의 주된 원인으로 설명되고 있다.

1. 사회경제적 발달

사망수준 저하에 대한 설명 가운데 가장 널리 알려져 있는 것은 맥쿠언(Mekcown, 1976)의 이론이다. 그는 유럽 사회를 연구한 것에 기초하여 인구변천 동안 사망률의 저하는 주로 전염병으로 인한 사망의 감소로 이루어졌다고 주

장한다. 그 전염병의 감소는 영양 섭취의 향상, 상수도 시설의 발전으로 인한 깨끗한 물의 공급, 멸균된 우유의 공급 등과 같은 생활수준의 향상, 즉 주위 환경의 개선을 통해 이루어진 것이다. 맥쿠언은 이런 주위 환경의 개선을 전체 사망률의 저하에 있어 가장 커다란 공헌을 한 요소라고 간주했다. 이런 주위 환경의 개선은 대체로 사회경제적 발전과 밀접하게 관련되어 있다. 즉 선진 산업화 사회에서 일어난 사망률의 저하는 주로 주위 환경 개선을 동반한 사회경제적 발달에 기인한 것이다. 게다가 이차세계대전이후 일어난 개발도상국가의 사망률 저하도 사회경제적 발달이 상당한 부분을 설명한다는 연구 결과도 있다 (Frederison, 1961).

2. 공공 보건의 향상

맥쿠언 이론과는 달리, 공공 보건과 위생에 대한 정부 기관의 적극적 개입이 유럽의 사망률 저하 원인의 대부분을 설명한다고 주장하는 연구도 있다(Szreter, 1988; Evans, 1987). 게다가 옴란(Omran, 1973)은 개발도상국가의 사망률 저하는 그 국가의 사회경제적 수준과는 상관없이 공공보건의 향상으로 인해 시작되었다고 주장한다. 말라리아 박멸 운동과 같은 공공보건 조치는 개발도상국가의 사망률 저하에 커다란 공헌을 하였다(Meegama, 1967; Preston, 1980). 요컨대 사회경제적 발달에 상관없이 공공보건의 향상은 독자적으로 사망률 저하에 커다란 공헌을 하였다.

3. 사회경제적 균등 상태

20세기 후반 들어, 몇몇의 개발도상국가들의 사망률 수준은 경제적 발달을 이룩하지 않은 상태에서 낮은 수준에 도달하였다. 카드웰(Caldwell, 1986)은 그렇게 될 수 있게 만든 주된 원인의 하나로 여성 교육의 향상(또는 충분한 여성의 자율성)을 제시한다. 특히 대중 교육을 통해 여성은 전통적인 확대 가족에 머무는 대신에 서구의 핵가족을 선택하기 쉽다. 전통적 가족에서 핵가족으로 전환하는 과정에서, 여성은 가족 내에서 보다 많은 자율성과 높은 지위를 차지할 수 있다. 여성은 가족 내에서 자율적인 행동을 할 수 있는 역량을 충분히 획득했기 때문에, 자신들이나 아픈 자녀들에게 자신의 행동을 취할 수 있다. 즉 사회의 경제적 발전과는 독립적으로, 개발도상국가 여성의 대중교육은 사망수준 저하에 큰 영향을 발휘하였다.

사회경제적 균등상태에 관련된 또 다른 중요한 요소는 그 사회의 소득 분포이다. 가트킨(Gwatkin, 1980)에 의하면, 사회경제적 발달을 통한 사망률 저하는 평등 상태에 의해 조건화된다. 사회가 경제적으로 발달됨에 따라 사망률은 저하되지만, 만약 경제적 발달이 균등 상태-즉 빈부 격차의 감소-를 동반하지 않는다면 어느 수준에서 멈춰 더 이상 저하되지 않는다. 일부 라틴 아메리카 국가들은 1960년대까지 사망률이 저하되었지만, 극심한 소득 격차로 인해 더 이상의 사망률 저하는 일어나지 않았다. 사망률 수준과 소득 분포 사이의 이런 관계는 개발도상국가에서뿐 아니라 영국(예: Pamuk, 1985)이나 미국(예: Duleep, 1989)과 같은 선진국가에서도 발견되었고, 또 국가간 비교 연구(Rodger, 1979)에서도 그런 관계가 입증되었다.

Ⅲ. 변수선정과 분석된 자료

<표1>과 <표2>에 이 연구에서 분석된 6개의 변수들이 각 국가별로 나열되어 있다. 6개의 변수 중 2개는 종속변수로 사용될 사망률 관련 지표들인 영아사망률과 기대수명이며, 나머지 4개는 독립변수로 사용될 소득불평등 지수인 기니지수, 국민총생산(GNP), 병원 침대 한 개당 인구 수, 여성의 초등학교 취학률이다. 영아사망률과 출생시 기대수명은 일반적으로 각 나라의 성·연령별 구조에 의해 영향을 받지 않기 때문에 국가간 사망률 수준을 비교하는데 적합하다. 국민총생산은 한 나라의 사회경제적 발달을 나타내는 지표로 자주 사용된다. 기니계수는 각 나라에서의 소득 불평등을 나타내는 계수이므로 사회경제적 불평등 지표로 간주될 수 있다. 여성의 초등학교 취학률은 흔히 사회경제적 발달지표로 간주될 수 있으나, 본 연구에서는 사회경제적 균등상태를 나타내는 변수로 간주된다. 왜냐하면 카드웰(Caldwell, 1986)이 언급한 바와 같이 영아사망수준 저하에 영향을 주는 것은 개별 여성의 교육수준이 아니라 사회적 차원에서 이루어진 여성의 대중교육이다. 대중적인 의무교육을 통해 여성들은 비서구 사회의 가족제도 특히 가부장적 제도의 문제점을 인식하게 되고, 또 여성과 아이들의 보건의 필요성도 깨닫게 된다. 물론 교육수준이 높은 여성일수록, 더욱더 서구식 핵가족 제도를 선호하여 자식을 더 독립적으로 돌볼 가능성이 높다. 그러나 사회적인 측면에서 영유아 사망수준 감소에 영향을 미치는 것은 대중교육의 보편화이므로 여성의 초등학교 취학률을 사회경제적 균등상태 변수로 간주한다. 마지막으로

병원 침대 한 개당 국민 수를 공공보건에 대한 지수로 사용하였다. 최근에는 다른 여러 변수들이 공공보건을 나타내는 지표로 간주될 수 있지만, 본 연구에 필요한 공공 보건 지표들은 본 연구에서 분석된 나라들에서 1970년 이전과 1970-80년 사이에 측정되어야 한다는 한계로 인해 병원침대당 인구수가 사용되었다.

회귀분석 결과에서 타당한 인과관계를 추론하기 위해서는 독립변수가 종속변수보다 선행되어야 한다. <표1>과 <표2>에서 세 원인 요소 지표들은 종속변수인 사망률 지표들보다 선행되어 측정되었다. <표1>에서 영아 사망률¹⁾과 기대수명에 대한 자료들은 1970년에 획득된 반면, 기니지수는 1948년과 1969년 사이에(참조 Paukert, 1973), 국민총생산은 1965년에, 그리고 병원 침대 한 개당 인구수와 여성의 초등학교 취학률²⁾은 1960년과 1965년 사이에 획득되었다. 동일한 절차가 <표2>에서도 사용되었다. 영아사망률은 1977년과 1980년 사이, 기대수명은 1980년, 기니계수는 1965-1975년(Simpson, 1990), 국민총생산은 1975년, 병원 침대 한 개당 인구수는 1970-76년, 그리고 여성의 초등학교 취학률은 1975-76년 사이에 획득되었다.

<표1>과 <표2>는 이 연구에서 분석될 국가들을 각기 다른 시점에서 보여주고 있다. <표1>에는 53개국의 사망률 지표들과 네 개의 가설적 원인 지표들을 보여주며, 그 지표들은 1948년과 1970년 사이에 획득된 것이다. 53개국은 11개의 유럽과 북미 국가, 16개의 중남미 국가, 5개의 중동국가, 12개의 아프리카 국가, 그리고 9개의 아시아 국가로 구성된다.

<표2>에는 55개국의 사망률 지표들과 네 개의 가설적 원인 지표들을 보여주는데, 그 지표들은 1965년과 1980년 사이에 획득된 것이다. 55개국은 15개의 유럽과 북미 국가, 16개의 중남미 국가, 4개의 중동국가, 9개의 아프리카 국가, 그리고 10개의 아시아 국가로 구성되어 있다. 동일한 국가들이 <표1>과 <표2>에 동시에 선정되지는 않았다. 하지만 그런 것이 국제 비교 연구에서 중대한 문제를 야기하지는 않는다. 왜냐하면 국제 비교 분석 결과가 단지 그 분석에 포함된 국가들로 국한되는 것은 아니기 때문이다(Bornschier and Chase-Dunn, 1985).

1) 본 글의 논평에서 한 논평자는 각 나라의 영아사망률 자료에 대한 타당성 검토를 요구하였으나, 그런 검토는 본 연구의 범위를 넘은 것이라 생각되어 생략되었다. 하지만 그런 검토는 매우 중요한 것이며 필요한 연구이다.

2) 표에서 여성의 초등학교 취학률은 많은 경우 100이 넘는다. 그것은 초등학교 취학 연령이 넘는 아이들이 초등학교에 등록하여 다닌 결과로 생긴 결과이다.

<표 1> 1970년 사망 수준과 사망관련 요인 변수 빈도표

	국가명	영아사망률	기대수명	기니계수	국민총생산	병원침대당 인구수	여성 초등학교 취학률
1	덴마크	14.2	73.3	37	2078	110	103
2	프랑스	18.2	72.1	50	1732	120	133
3	핀란드	13.2	70	46	1568	100	95
4	그리스	29.6	71.3	38	591	170	109
5	독일	23.6	70.5	45	1667	90	134
6	이탈리아	29.6	71.5	40	1011	110	110
7	네덜란드	12.7	73.7	42	1400	130	104
8	노르웨이	12.7	74.2	35	1717	100	100
9	스웨덴	11	74.3	39	2406	70	96
10	영국	18.4	71.8	38	1590	90	92
11	미국	20	70.9	34	3233	110	118
12	아르헨티나	53.9	67.9	42	782	160	102
13	바베이도스	40.1	68.3	45	368	170	107
14	볼리비아	154.4	45.9	53	132	410	50
15	브라질	98.6	58.9	54	207	370	93
16	칠레	82.3	62.4	44	486	240	122
17	콜롬비아	77	58.9	62	275	400	77
18	코스타리카	58.3	66.9	50	360	240	105
19	에콰도르	107.3	55.9	38	202	420	88
20	엘살바도르	106.5	57.5	54	249	460	79
21	자메이카	32.2	68.6	56	465	250	106
22	멕시코	73.6	61.3	53	441	500	90
23	파나마	48.9	66.1	48	490	290	99
24	페루	119.7	53.4	61	237	400	90
25	수리남	50.7	64.5	30	424	170	106
26	트리니다드 토바고	42.7	68.6	44	704	230	90
27	베네수엘라	58.7	63.2	42	904	320	94
28	이라크	103.5	51.4	60	285	520	45
29	이스라엘	23	71.2	30	1243	140	95
30	레바논	50	62.1	55	440	210	93
31	모로코	136.4	51.7	50	180	650	35
32	튀니지	131.3	54.2	53	187	390	65
33	차드	173	39.3	35	90	970	4
34	가봉	137.4	43.6	64	368	150	76
35	코트디부아르	148.6	41.5	43	213	540	41
36	마다가스카르	89.2	42.2	53	92	380	59
37	니제르	171.3	40.5	34	81	1950	7
38	나이지리아	158	43.7	51	74	2430	24
39	세네갈	164.4	40.3	56	192	810	29
40	시에라리온	227.6	42.1	56	142	1130	21
41	남아프리카공화국	114.5	57.5	58	521	190	85
42	수단	150.4	41.8	40	97	930	21
43	탄자니아	125.3	46.7	54	61	560	25
44	잠비아	125.2	44.6	48	207	390	46
45	미얀마	127.9	48.7	35	64	1210	65
46	피지	50	69.1	46	295	270	82
47	인도	139	48.1	33	95	1670	57
48	일본	13.1	72.4	39	838	110	100
49	말레이시아	45.6	59	36	278	290	84
50	파키스탄	143	46.2	37	101	1450	20
51	필리핀	75	59	48	240	710	93
52	스리랑카	58.5	63.6	44	140	320	90
53	한국	50.1	60.3	26	107	2560	99

출처: 영아사망률(1970), 출생 기대수명(1970) 병원침대당 인구수(1960-65), 여성 초등학교 취학률 (1960-65)은 World Bank(1983)에서, 그리고 기니계수와 국민총생산(GNP)(1965) 자료는 Paukert (1973)에서.

〈표 2〉 1977-1980 사망 수준과 사망관련 요인 변수 빈도표

국가명	영아사망률	기대수명	기니계수	국민총생산	병원침대당 인구수	여성 초등학교 취학률
1. 오스트리아	13.9	73	37.1	4,760	90	101
2. 벨기에	11.2	73	34	6,050	110	102
3. 캐나다	10.4	74.6	33.8	7,330	110	99
4. 덴마크	8.4	74.9	36.7	6,910	110	105
5. 프랑스	9.9	75.1	39.3	6,010	100	109
6. 헝가리	23.2	70.4	24.8	1,310	120	99
7. 아일랜드	11.2	73	30	2,610	90	106
8. 이탈리아	14.3	73.4	38	3,840	90	106
9. 네덜란드	8.6	75.7	30	6,070	100	100
10. 노르웨이	8.1	75.9	30	6,610	70	101
11. 포르투갈	26	71.5	40	1,480	180	112
12. 스페인	11.1	73.6	36	2,750	190	111
13. 스웨덴	6.7	76.5	33.9	8,27	70	102
14. 영국	12.1	73.3	31.4	3,910	110	106
15. 유고	32.8	70.4	34.7	1,380	170	101
16. 아르헨티나	45.2	70.4	41.4	1,810	180	109
17. 바베이도스	26.3	70.8	33.4	1,520	110	102
18. 볼리비아	131.3	50.2	49	360	460	77
19. 브라질	77.4	63.1	56.5	1,070	280	87
20. 칠레	43.2	67.1	46.5	870	270	118
21. 콜롬비아	56.4	62.9	52.1	560	530	120
22. 코스타리카	27.5	72.2	42.8	950	260	106
23. 엘살바도르	77.9	63	46	430	550	74
24. 과테말라	65.9	58.5	46	570	500	56
25. 온두라스	88.5	58.2	59.6	360	670	86
26. 멕시코	56	65.2	55.2	1,360	870	28
27. 니카라과	90.5	56.4	51	650	470	88
28. 파나마	21.7	70.4	54	1,030	280	120
29. 페루	87.7	57.5	54	1,000	440	105
30. 트리니다드 토바고	32.3	71.5	42	2,000	220	101
31. 우루과이	39.7	70.9	40	1,370	240	103
32. 베네수엘라	41.7	67.4	54.6	2,380	350	97
33. 이집트	110	56.6	40.5	310	470	57
34. 이란	107.8	58	52	3593.3	650	73
35. 튀니지	90	60.2	46	710	430	78
36. 터키	122.6	61.8	51.6	830	470	98
37. 가봉	116.8	48.5	55	2,660	150	98
38. 가나	102.8	54.1	39.1	280	610	61
39. 케냐	86.8	55.3	57.4	230	760	91
40. 말라위	171.9	43.9	40.5	120	540	44
41. 세네갈	146.9	43.4	47	340	880	31
42. 시에라리온	207.5	47	49.4	220	1110	28
43. 남아프리카	96.2	62	53	1,410	190	99
44. 수단	124.5	46.2	41	250	1000	34
45. 탄자니아	102.9	51.1	50	160	640	60
46. 호주	11	74	35.4	6,410	80	107
47. 인도	123.4	51.8	44	160	1310	60
48. 인도네시아	105	53.1	38.9	210	1560	76
49. 일본	7.4	76.4	33.2	4,490	100	99
50. 말레이시아	31.1	64.3	48.6	820	300	89
51. 뉴질랜드	12.6	73.7	34.5	4,520	100	106
52. 파키스탄	125.5	49.8	33	140	2090	28
53. 한국	34.1	65.4	31.9	580	640	107
54. 스리랑카	44.2	68.3	34.9	220	330	74
55. 아일랜드	54.8	62.6	44.2	360	810	81

출처: 영아사망률(1977-80), 출생 기대수명(1980), 국민총생산GNP(1975)는 World Bank(1989)에서, 병원침대당 인구수(1970-76), 여성 초등학교 취학률(1975-76-65)은 World Bank(1980)에서, 그리고 기니계수는 Simpson(1990)에서.

요컨대 본 연구에서 사용된 표본은 비록 전 세계를 대표하는 데는 문제가 있을지라도, 그것은 전 세계 국가들의 일부인 것이다. 따라서 약간 다른 국가들을 사용하여 두 시기에 세 가지 요소들이 사망률 저하에 미치는 상대적 영향력을 분석하는 데는 그다지 문제가 되지 않는다. 이 연구는 분석결과의 대표성보다는 분석결과의 특성에 더 관심을 가지기 때문에 표본의 대표성 문제는 그다지 중요하지 않다.

IV. 자료 분석

조심스런 자료 분석은 다음의 세 단계를 거쳐야 한다. 첫 번째 단계는 일차 준비 분석(preliminary analysis) 단계이고, 두 번째 단계는 모형을 검증하는 확인 분석(confirmatory analysis) 단계이며, 세 번째 단계는 모형의 내부 점검(internal checking) 단계이다. 불행히도 대부분의 자료 분석 연구들은 첫 번째와 세 번째 단계를 무시한 상태에서 독립변수와 종속변수 사이에서 어떤 인과관계를 가정하는 두 번째 단계에 초점이 맞추어져 있다(Franzosi, 1994). 일차 준비 분석은 변수들의 분포에 관한 정보-예를 들어, 대칭성이나 이상값에 관한-를 말하여 준다. 변수의 분포가 대칭적이지 않을 때는 그 변수의 변형이 요구된다. 결과적으로, 우리는 일차 준비 분석 단계를 통해 회귀분석 모형에서 어떤 형태의 함수가 적합한 형태인지를 발견할 수 있다. 우리는 이 연구에서 사망률 지표들과 독립변수 사이의 관계를 반드시 선형이라고 간주할 필요는 없다. 사망률 지표들과 독립변수 사이의 함수 관계가 어떤 형태일 것이라는 이론적 배경이 없기 때문에, 일차 준비 분석 단계 작업을 통하여 적합한 형태의 함수 모형을 찾는 것은 바람직한 일이다. 게다가 세 번째 단계에서는 우리 모형에 유력 사례들을 발견하는 모형의 내부 분석 작업이 시행되기 때문에 보다 일반화될 수 있는 관계들을 얻을 수 있다.

1. 일차 준비 단계

종속 변수의 편향된(skewed) 분포는 검정력(power)-즉 대립가설이 참일 때 귀무가설을 기각시킬 수 있는 확률-을 약화시킨다. 대칭성이 결핍되면 될수록, 검정력의 손실은 커진다. 종속변수의 일괄묶음(batch)이 대칭이 아니면 종속변수를 변형시킬 필요가 있다. 어떤 형태로 종속변수를 변형시킬 것인가는 다음의 검정

력 변형 공식에 의해 결정될 수 있다(Box and Cox, 1964).

$$T_{\lambda}(Y) = \begin{cases} \frac{(Y^{\lambda}-1)}{\lambda} & \text{for } \lambda \neq 0 \\ \log_e Y & \text{for } \lambda = 0 \end{cases} \quad [1]$$

Y는 종속변수이고 λ 는 검정력 사다리(ladder of power)이다. 우리는 최소 나머지 자승법(minimum residual sum of squares)에 근거해 λ 를 선택한다. 검정력 변형을 통하여 종속변수의 일괄묵음 모양은 비대칭형에서 대칭형으로 바뀐다.

정통(classical) 회귀 분석은 오차 항들이 정상적으로 분포되어 있다고 가정한다. 대칭성은 정상성(normality)의 중요한 부분이며, 통계 유의도 검정에 중대한 영향을 미친다. 특히 표본의 크기가 작을 때 종속 변수의 정상성은 매우 중요한 가정이 된다. 만약 일괄묵음의 모양을 대칭형으로 변형시키지 못한다면, 종속변수와 독립변수 사이의 관계가 참일 때 그 관계를 통계 분석에서 기각하기 쉽다. 일괄묵음이 대칭적일 때 평균과 표준 편차에 기초한 정통회귀분석 모형의 사용이 적합하게 되기 때문에, 일괄묵음의 대칭성은 바람직한 속성이 된다(Leinhardt and Wasserman, 1979). 결과적으로 검정력 변형을 통해 우리는 변수들 간의 참(true) 관계가 어떤 형태의 비선형 모형인지를 밝혀내는데 도움을 받을 수 있다.

종속변수가 대칭적으로 분포되어 있는지 아닌지를 보기 위해 자료탐색 분석방법(Exploratory Data Analysis, EDA)이 본 연구에 사용될 것이다. 자료탐색 분석방법에서, 변수의 일괄묵음은 사분값(Fourths)들과 중앙값 사이의 거리들이나 팔분값(Eights)들과 중앙값 사이의 거리가 동등할 때 대칭적이라고 간주된다(Leinhardt and Wasserman, 1979; Franzosi, 1994). 이 연구에서는 글자값 전시(letter-value displays), 잎줄기 전시(stem-and-leaf), 상자도면(boxplot), 그리고 정상확률 도면(normal probability plot) 등이 종속 변수의 일괄묵음이 대칭적인가 아닌가를 발견하기 위해 사용될 것이다.

2. 모형 내부 점검

유력 사례는 “다른 관찰값보다 그 자체나 다른 관찰값과 상호관련하여 회귀계수의 추정치 계산에 커다란 영향력을 발휘하는 관찰값이다(Belsley, Kuh and Welsch, 1980: 11).” 때때로 한 두 개의 사례만으로도 때로는 회귀 계수의 추정치 계산에서 영향을 미치곤 한다. 그 결과 한 두 개의 유력 사례를 빼버리고 다

시 분석한 회귀분석 추정치 계산에서 상당한 변화가 일어날 수 있다.

이런 유력 사례들이 존재하는지 안하는지를 분석하는 것은 매우 중요한데, 왜냐하면 그 특별 사례는 나머지의 다른 사례들과는 다른 이론적 함축을 포함할 수 있기 때문이다. 본 연구에서는 유력 사례를 발견하기 위해 회귀분석 진단 절차(regression diagnostic procedure)들을 사용할 것이다. 본 연구에서 사용될 회귀분석 진단 절차들에는 (1) 스튜던트 나머지(Studentized residuals), (2) 해트 행렬(Hat matrix), (3) 디휠(DFIT), 그리고 쿡스 디(Cook's D) 등이 있다.

유력 사례들은 지레거리(leverage)와 불일치(discrepancy)가 서로 결합한 영향력을 통해서 회귀분석 계수에 영향을 미친다(Bollen and Jackman, 1985; Fox, 1991; Franzosi, 1994). 해트 행렬의 대각선 요소(h_i)들은 일반적으로 회귀 계수에서의 지레거리에 대한 잣대로 사용되고, 스튜던트 나머지는 각 사례의 나머지가 자신의 분산에 비해 비교적 큰지 아닌지를 결정하는데 사용되는데 불일치가 큰 사례들은 커다란 나머지를 가진다. 베슬리(Belsely et al, 1980: 17)는 해트 행렬 값이 $2/n$ 을 초과하는 관찰값은 큰 지레거리를 가진다고 주장한다(k 는 독립 변수의 수에 1을 더한 수이고 n 은 표본의 수이다). 하지만 작은 수의 표본에 대해서는 $2k/n$ 보다 $3k/n$ 이 큰 지레거리를 밝히는데 보다 적합한 절단값(cutoff)이라는 주장이 있다(Velleman and Welsch, 1981: 234-5). 베슬리에 따르면, 스튜던트 나머지는 자유도($n-k$)인 t -분포에 적합하게 분포되어 있다고 한다. 하지만 볼렌과 잭슨(Bollen and Jackson, 1990)은 t -분포의 절단값을 낮은 절단값으로, α/n 에 상응하는 t -분포의 값을 스튜던트 나머지에 대한 높은 절단값으로 간주하였다.

그러나 이 두 잣대는 때때로 일치되지 않는다. 이를테면 큰 지레거리를 가진 사례가 보다 작은 불일치를 가질 수 있고, 또 작은 지레거리를 가진 사례가 커다란 불일치를 보일 수도 있다(Hoaglin and Welsch, 1978). 그러므로 우리는 큰 지레거리이나 불일치에 의해 영향을 받을 수 있는 다른 잣대들, 즉 디휠(DFIT)이나 쿡스 디(Cook's D)가 필요할 수 있다. 베슬리(Bersley et al, 1980)는 $2\sqrt{k/n}$ 을 디휠의 절단값으로 제시했다. 하지만 볼렌과 잭맨(Bollen and Jackman, 1990)은 $2\sqrt{k/n}$ 을 낮은 절단값 그리고 \sqrt{k} 를 높은 절단값으로 간주한다. 쿡스 디에 대해서는 볼렌과 잭맨이 $4/n$ 을 낮은 절단값으로, 1을 높은 절단값으로 제시한다. 그러나 폭스(Fox, 1991)는 1은 유력 자료를 놓칠 가능성이 있는 절단값이라고 경고한다.

V. 분석결과

사망률 지수들과 세 요소들 사이에서의 일반화된 패턴을 발견하기 위하여, 본 연구에서는 자료 분석에서 두 가지 중요한 측면이 고려되었다. 하나는 각 사망률 지수의 분포들이 대칭적인지 아닌지가 고려되었다. 다른 하나는 본 조사에서 제시한 세 가지 요소 이외의 다른 요소가 영향을 미친 국가가 있는지 없는지를 조사하기 위해, 회귀분석 진단 분석을 통해 유력 사례들을 발견하여 모형을 점검하였다.

1. 변형

먼저 글자값 전시(letter-value display)를 통해 변형의 필요성을 일차적으로 조사해 보았다. 본 연구에서는 7개의 요약값들이 글자값 전시를 위해 사용되었다. 글자값 전시에 사용된 7개의 요약값은 중앙값(median, M), 두 개의 사분값(hinges, H), 두 개의 팔분값(eighths, E) 그리고 두 개의 극단값(extreme)들이다. 일괄목록이 대칭적인가 아닌가를 보기 위해, 중앙요약값들의 추세(trend of mid-summaries)가 사용된다. 중앙요약값은 각 쌍 글자값의 평균이다. 예를 들어, 중앙사분값(mid H)는 상사분값과 하사분값의 평균값이고, 중앙팔분값(mid E)은 두 팔분값의 평균이다. 자료 분포가 대칭적이라면, 중앙요약값들이 어떤 추세도 보이지 않을 것이다(Leinhardt and Wasserman, 1979; Fox, 1990; Franzosi, 1994).

<표3>은 각 시기에서 영아사망률 변수가 변형된 것과 원래의 것에 대한 일괄 글자값이 전시되어 있다. 변형되지 않은 상태($\lambda=1$)에서 변수값의 추세는 위쪽을 향한다. 즉 정적으로 편향되어 있다. 원자료의 중앙요약값에 대한 추세가 위쪽으로 향하기 때문에, λ 가 1보다 작은 방향에서 사다리 검정력에 대한 중앙요약값들의 분석을 시도하는 것이 바람직하다. 1970년 영아사망률은 제공근을 씌워 변형시킨 결과 일괄목록이 대칭성을 보였다. 그러나 1977-80년대 영아사망률은 로그로 변형시켰을 때 일괄목록이 비교적 대칭적이 되었다. 우리의 분석 결과에 의하면, 1970년 영아사망률 변수에 적합한 변형 함수는 제공근인 반면 1977-80년의 것은 로그가 되었다. 영아사망률 변수의 분석 결과는 각 시기별로 각기 다른 형태의 함수를 적합한 변형함수로 제시하였다. 그러나 영아사망률의 변형을 분석한 또 다른 연구는 로그 변형을 가장 효율적인 변형함수로 지적하고 있다(Fox, 1990). 자료가 무작위로 추출되지 않았기 때문에 서로 다른 형태의 변형 함수 변수가 제시될 수 있다.³⁾ 따라서 본 분석은 1980년 영아사망률과 마찬가지로

지로 1970년대 영아사망률 변수도 로그 함수로 변형하여 분석한다.

<표 3> 영아사망률에 대한 글자값 전시

가. 1970년 영아사망률

	1/IMR ($\lambda=-1$)	\log_{10} (IMR) ($\lambda=0$)	Root(IMR) ($\lambda=1/2$)	IMR ($\lambda=1$)
하극단값(Min)	0.004	2.40	3.32	11
하팔분값(E)	0.007	2.805	4.3	18
하사분값(H)	0.008	3.47	5.7	32
중앙값(M)	0.014	4.30	8.57	73.6
상사분값(H)	0.031	4.83	11.3	128
상팔분값(E)	0.055	5.005	12.25	150
상극단값(Max)	0.091	5.43	15.1	228
중앙값	0.014	4.30	8.57	73.8
중앙 사분값	0.0195	4.15	8.5	80
중앙 팔분값	0.031	3.905	8.275	84
중앙 극단값	0.0475	3.915	9.21	119.5
추세	위로	약간 아래로	평평(flat)	위로

나. 1980년 영아사망률

	1/IMR ($\lambda=-1$)	\log_{10} (IMR) ($\lambda=0$)	Root(IMR) ($\lambda=1/2$)	IMR ($\lambda=1$)
하극단값(Min)	0.005	1.902	2.59	6.7
하팔분값(E)	0.008	2.400	3.30	11
하사분값(H)	0.017	2.645	3.50	14
중앙값(M)	0.023	3.789	6.65	44.2
상사분값(H)	0.076	4.545	9.95	99.5
상팔분값(E)	0.091	4.810	11.0	123
상극단값(Max)	0.149	5.5335	14.4	207.5
중앙값	0.023	3.789	6.65	44.2
중앙 사분값	0.0465	3.595	6.725	56.8
중앙 팔분값	0.0495	3.605	7.2	67
중앙 극단값	0.077	3.6185	8.495	107.1
추세	위로	평평	위로	위로

<표4>는 출생시 기대수명 자료들에 대한 추세가 약간 하향하고 있음을 보여 준다. 하지만 1980년의 경우 중앙값과 중앙 사분값은 거의 차이가 없다. 따라서

3) 변형된 함수변수들에 대한 일차준비 분석 단계에서 제공된 영아사망률 자료가 로그 영아사망률 자료보다 좀더 대칭적인 일괄류음의 형태를 띠지만, 정상확률도면에서는 로그 변형이 제공된 변형보다 정상분포에 더 근접하였다. 또 영아사망률을 제공근과 로그함수 형태로 변형하여 회귀분석을 시도하여 그 회귀분석 결과에 대해 모형 내부 점검을 해보았지만, 별다른 차이가 발견되지 않았다.

본 연구는 기대수명의 일괄 묶음은 거의 대칭적으로 분포되었다고 가정한다.

〈표 4〉 출생시 기대수명에 대한 글자값 전시

1970년 출생시 기대수명		1980년 출생시 기대수명	
	기대수명($\lambda=1$)		기대수명($\lambda=1$)
하극단값(Min)	38.3	하극단값	43.4
하팔분값	42.7	하팔분값	50.2
하사분값	48.1	하사분값	56.2
중앙값	60.3	중앙값	65.4
상사분값	68.6	상사분값	72.0
상팔분값	70.5	상팔분값	74.0
상극단값	74.3	상극단값	76.5
중앙값	60.3	중앙값	64.5
중앙 사분값	58.4	중앙 사분값	64.3
중앙 팔분값	56.6	중앙 팔분값	62.1
중앙 극단값	56.3	중앙 극단값	60.0
추세	아래로	추세	약간 아래로

변형된 종속변수의 일괄 묶음이 변형전 종속변수의 일괄 묶음보다 좀더 대칭적(또는 정상분포에 근접했는지)이 되었는지를 알아보기 위해 줄기잎 모형(stem-and-leaf display), 상자도면(boxplot) 그리고 정상확률도면(normal probability plot)을 분석하였다. 줄기잎 모형과 상자 도면은 보완적이 되는데, 왜냐하면 줄기잎 모형은 개별 자료들에 대한 정보를 제공하는 반면 상자 도면은 자료의 집합에 근거해서 중앙(center), 퍼짐(spread), 이상값(outlier)과 같은 통계값을 조명해 주기 때문이다(Fox, 1990; Leinhardt and Wasserman, 1979). 정상확률도면은 분석된 변수가 정상적으로 분포되었는지를 분석하는데 이용된다. 정상확률도면은 경험적 자료의 분포(즉 변형된 자료의 분포)가 이론적 분포에 얼마나 접근해 있는지를 보여준다. 만약 도면의 점들이 $y=x$ 선 가까이 분포되어 있다면, 우리의 자료들이 거의 정상분포 되었다고 볼 수 있다. 만약 도면의 점들이 $y=x$ 선에서 크게 떨어져 있다면, 우리가 그 변수를 변형하거나 이상값을 고려해야 한다는 사실을 암시한다(Chambers et al., 1983; Fox, 1990). <그림 1>과 <그림 2>는 1970년대와 1977-1980년대 영아사망률에 대한 원자료와 로그 변형된 자료에 대한 줄기잎 모형, 상자도면 그리고 정상확률도면을 보여준다. 특히 정상확률분포에서 변형된 변수는 변형 안된 변수보다 좀더 정상분포에 근접하여 분포되어 있음을 보여준다. <그림 1>에서 1977-80 영아사망률 일괄 묶음에 대해 로그로 변형된 일괄 묶음이 가장 대칭적이면서 정상분포에 근접함을 보여준다.

<그림 1> 1970년대와 1980년대 영아사망률 변수의 잎줄기, 상자, 정규 분포 플롯

1970 영아사망 Stem-and-Leaf Plot $\log_{10}(1970IMR)$ Stem-and-Leaf Plot $\text{Root}(1970IMR)$ Stem-and-Leaf Plot

Frequency	Stem & Leaf
18.00	0 . 11111111222234444
14.00	0 . 5555555777889
14.00	1 . 00011222333344
6.00	1 . 555677
1.00	2 . 2

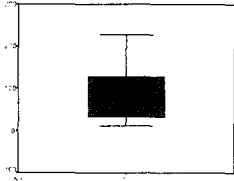
Stem width: 100
Each leaf: 1 case(s)

Frequency	Stem & Leaf
6.00	1 . 011111
5.00	1 . 22333
3.00	1 . 445
12.00	1 . 666666677777
6.00	1 . 888999
17.00	2 . 0000001111111111
4.00	2 . 2223

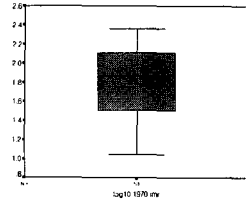
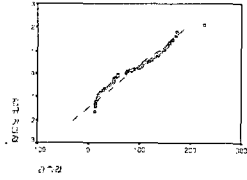
Stem width: 1.00
Each leaf: 1 case(s)

Frequency	Stem & Leaf
6.00	0 . 333333
8.00	0 . 44444555
12.00	0 . 6667777777
6.00	0 . 888999
13.00	1 . 000011111111
7.00	1 . 222233
1.00	1 . 5

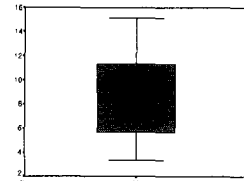
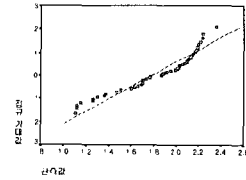
Stem width: 10.00
Each leaf: 1 case(s)



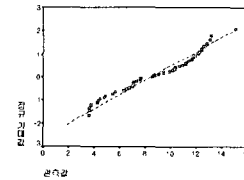
1970 영아사망의 정규 Q-Q 도표



\log_{10} 1970 imr의 정규 Q-Q 도표



RTIMR_{70} 의 정규 Q-Q 도표



1980 영아사망 Stem-and-Leaf Plot

Frequency	Stem & Leaf
29.00	0 .
00000011111111112222333334444	
12.00	0 . 555677888999
12.00	1 . 000011222234
1.00	1 . 7
1.00	2 . 0

Stem width: 100.0
Each leaf: 1 case(s)

$\log_{10}(1980IMR)$ Stem-and-Leaf Plot

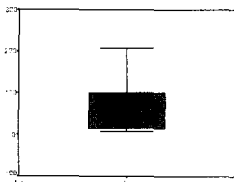
Frequency	Stem & Leaf
6.00	0 . 889999
15.00	1 . 000000111334444
20.00	1 . 5556666777888999999
14.00	2 . 00000000001123

Stem width: 1.00
Each leaf: 1 case(s)

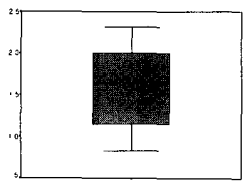
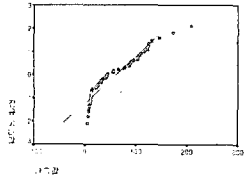
$\text{Root}(1980IMR)$ Stem-and-Leaf Plot

Frequency	Stem & Leaf
15.00	0 . 2222233333333333
9.00	0 . 44555555
8.00	0 . 6666777
9.00	0 . 888999999
11.00	1 . 0000011111
2.00	1 . 23
1.00	1 . 4

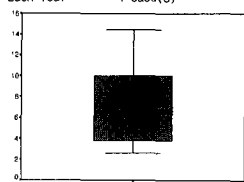
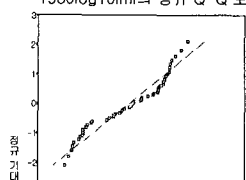
Stem width: 10.00
Each leaf: 1 case(s)



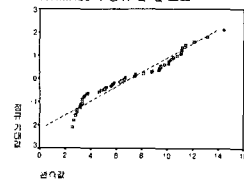
1980 imr의 정규 Q-Q 도표



\log_{10} 1980 imr의 정규 Q-Q 도표



RTIMR_{80} 의 정규 Q-Q 도표



<그림 2>에서, 1970년대 출생시 기대수명 자료들은 약간 음의 방향으로 편향되어 분포되었지만 1977-80년대 분포는 거의 대칭적으로 분포되어 있음을 보여준다. 반면 정상확률도면에서 두 시기의 기대수명 자료들 모두가 정상분포에 근접하여 분포되어 있음을 보여준다. 따라서 기대수명 자료들은 변형시킬 필요가 없다고 가정된다.

<그림 2> 1970년대와 1980년대 기대수명 변수의 잎줄기, 상자, 정규 분포 플롯

1970 기대수명 Stem-and-Leaf Plot

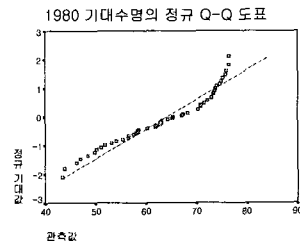
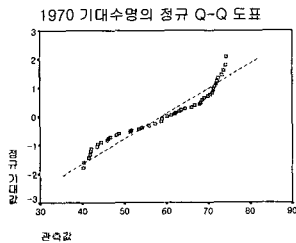
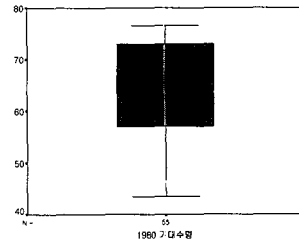
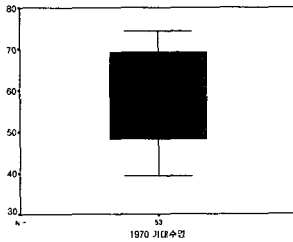
Frequency	Stem and Leaf
1.00	3 . 9
9.00	4 . 001122334
5.00	4 . 56688
4.00	5 . 1134
7.00	5 . 5778899
7.00	6 . 0122334
7.00	6 . 6678889
13.00	7 . 0001111223344

Stem width: 10
Each leaf: 1 case(s)

1980 기대수명 Stem-and-Leaf Plot

Frequency	Stem and Leaf
2.00	4 . 33
4.00	4 . 6789
5.00	5 . 01134
7.00	5 . 5667888
8.00	6 . 0122334
5.00	6 . 55778
19.00	7 . 0000001123333334444
5.00	7 . 55566

Stem width: 10
Each leaf: 1 case(s)



2. 회귀분석 진단과 발견들

앞에서 진술한 바와 같이, 영아사망률과 출생시 기대수명은 사망수준을 측정하는 지표로 간주된다. 본 연구가 국가간 비교를 하는 주된 목적의 하나가 사망수준 저하에 어떤 요인이 상대적으로 더 큰 공헌을 했는지를 분석하는데 있기 때문에, 표준화된 회귀 분석 계수가 이용된다. 또한 상대적 공헌도가 두 시기에 걸쳐 분석됨으로써, 시기에 따라 서로 다른 요소들이 가장 큰 공헌을 했는지 아니면 동일한 요소가 가장 큰 공헌을 했는지를 분석한다.

1) 영아사망률

1970년 로그로 변형된 영아사망률 변수에 여성의 초등학교 취학률(WEDU), 기니지수(GINI), 국민총생산(GNP) 그리고 병원 침대 한 개당 인구수(HBED)에 회귀하여 나타난 회귀 방정식 결과는 다음과 같다.

$$\log IMR = 1.98^{***} - 0.41WEDU^{***} + 0.18GINI^{**} - 0.51GNP^{***} + 0.08HBED \quad [2]$$

R^2 는 0.826이고, 여성의 초등학교 취학률과 국민총생산 회귀계수는 0.01수준에서, 기니계수 변수는 0.05 수준에서 유의미하였지만, 병원 침대 한 개당 인구수 변수는 0.1수준에서도 유의미하지 않았다.⁴⁾ 위의 분석 결과는 1970년 영아사망률 변수의 변이에 국민총생산이 가장 커다란 영향을, 여성 교육이 두 번째, 그리고 기니 계수가 세 번째로 큰 영향을 미쳤음을 나타낸다. 비록 국민총생산이 영아사망률에 가장 큰 영향을 끼치는 것으로 나타났지만, 여성 교육과 기니계수도 상당한 영향력을 미쳤다는 결과가 나왔다. 따라서 사회경제적 발달이 사회경제적 균등상태보다 영아사망률 저하에 더 큰 영향을 미쳤다고 주장하는 데는 좀 무리가 있다. 반면 영아사망률 저하에 있어 사회경제적 발달과 균등상태가 공공 보건보다는 더 큰 영향을 미쳤다고는 말할 수 있다. (반면 종속변수를 변형하지 않고 회귀분석을 하였을 때는 여성의 초등학교 취학률이 가장 중요한 요인으로 간주되었다. 하지만 본 모형에서 독립변수들과 영아사망률 간의 관계를 선형으로 가정하는 데는 많은 문제점들이 노출된다.)

4) ***, **, *는 각기 다른 유의미 수준을 나타낸다, ***는 0.01수준을, **는 0.05수준을, *는 0.1수준을 나타낸다.

로그로 변형된 모형에 대해 모형 내부 점검을 시도하였다. <표5>는 방정식 [2]의 회귀분석 진단 결과이다. <표5>의 첫 번째 열인 헤트 행렬의 대각선 요소들에서, 3개국-미국, 나이지리아, 한국-이 상한선을 넘는 큰 지레거리를 가지고 있음을 볼 수 있다. 두 번째 열은 스튜던트 나머지들을 보여준다. F 분포($n=53$, $k=5$)에서 하한 절단값은 2.02이고 상한 절단값은 3.46이다($\alpha=0.05$). 미국의 스튜던트 나머지값은 상한 절단값을 넘는 반면 일본의 값은 하한 절단값을 넘고 있다. 쿡스 디와 디휠은 <표 5>의 마지막 두 열에 제시되어 있다. 쿡스 디 검사에서 미국은 상한 절단값을 넘는 높은 값을 보여주었다. 이는 높은 이상 지레거리와 스튜던트 나머지값이 결합하여 나타난 값이라고 할 수 있다. 반면 지레거리나 스튜던트에서 높은 이상값을 보여준 나이지리아, 한국, 일본 가운데 나이지리아와 한국은 하한 절단값을 초과하는 쿡스 디 값을 보여주었지만 일본의 쿡스 디 값은 하한 절단값을 밑돈다. 한국과 나이지리아도 본 회귀모형에서 유력 사례로 간주될 수 있지만, 미국만큼은 못된다. 미국은 지레거리, 스튜던트 나머지 그리고 쿡스 디 모두에서 상한 절단값을 초과하는 값들을 보여준다. 사실 미국은 서구의 선진국가들에 비해 높은 영아사망률을 보인다. 그것은 다른 유럽국가들에 비해 높은 실업률, 높은 십대 미혼모의 출생률, 높은 흡연 및 음주율 그리고 마약 복용 등과 같은 요인으로 미숙아 출생률이 높기 때문이다(Pampel, Fred and Pillai, 1986; Eberstein, 1989). 이는 경제적 요인뿐만 아니라 높은 청소년 비행률과 같은 사회문제도 영아사망률에 영향을 미칠 수 있다는 사실을 말한다. 즉 미국의 영아사망률을 설명하는데 있어 본 모형의 세 가지 요소 이외의 요소가 중요할 수 있음을 의미한다. 이러한 사실은 본 연구의 영아사망률 국제 비교 분석에서 미국을 유력 사례로 간주하고 제외시킬 수 있는 이론적 근거를 제공한다.

앞의 분석 결과를 통해 미국은 본 회귀 모형에서 유력 사례로 간주되었다. 따라서 유력 사례가 1970년대 영아사망률 자료에서 배제되었을 때 회귀계수에 어떤 변화가 일어나는지를 분석해 볼 필요가 있다. 1970년 영아사망률 자료에서 미국을 빠뜨린 상태에서 다시 회귀분석을 시도한 결과는 다음과 같다.

$$\log IMR = 2.03^{***} - 0.37WEDU^{***} + 0.16GINI^{**} - 0.60GNP^{***} + 0.05HBED \quad [3]$$

R^2 는 0.868로 증가되었다. 회귀계수들의 유의도 수준은 변화하지 않은 반면, 영아사망률에 미치는 국민총생산의 상대적 영향력은 증가한 반면 여성의 초등학

<표 5> 1970 로그 영아사망률 회귀모형(방정식 2) 내부 점검 결과

국가명	지레거리	스튜던트 나머지	Cooks'D	DFIT
1. 덴마크	0.090080	-0.135130	0.000450	-0.002520
2. 프랑스	0.079170	-0.656700	0.009370	-0.010960
3. 핀란드	0.042320	-1.796370	0.042070	-0.018350
4. 그리스	0.043990	-0.780660	0.008180	-0.008200
5. 독일	0.049640	0.199270	0.000580	0.002290
6. 이탈리아	0.022030	-0.108570	0.000100	-0.000730
7. 네덜란드	0.026610	-1.852230	0.032690	-0.013940
8. 노르웨이	0.059590	-0.991950	0.016750	-0.013110
9. 스웨덴	0.142110	-0.494320	0.009380	-0.014050
10. 영국	0.045420	-0.569910	0.004460	-0.006120
11. 미국	0.320730	3.560960	1.304140	0.240640
12. 아르헨티나	0.019170	1.027820	0.008350	0.006450
13. 바베이도스	0.034350	-0.644540	0.004670	-0.005700
14. 볼리비아	0.034330	0.611690	0.004200	0.005410
15. 브라질	0.039710	0.858910	0.009180	0.008390
16. 칠레	0.025550	1.560020	0.022620	0.011460
17. 콜롬비아	0.068870	-0.602510	0.006980	-0.008950
18. 코스타리카	0.032020	0.124720	0.000170	0.001050
19. 에콰도르	0.050310	1.636420	0.039800	0.018970
20. 엘살바도르	0.022370	0.657160	0.003710	0.004480
21. 자메이카	0.050390	-1.616690	0.038900	-0.018770
22. 멕시코	0.029670	0.361670	0.001330	0.002910
23. 파나마	0.021770	-0.091610	0.000070	-0.000610
24. 페루	0.077940	1.046860	0.023490	0.017240
25. 수리남	0.126730	0.816610	0.022730	0.020800
26. 트리니다드 토바고	0.008270	-0.177130	0.000180	-0.000790
27. 베네수엘라	0.009170	1.220290	0.008590	0.005610
28. 이라크	0.066180	-0.750920	0.010480	-0.010800
29. 이스라엘	0.074900	-0.122920	0.000310	-0.001960
30. 레바논	0.033720	-0.597950	0.003970	-0.005220
31. 모로코	0.044710	-0.078030	0.000080	-0.000830
32. 튀니지	0.019200	0.744020	0.004380	0.004670
33. 차드	0.188440	-0.014390	0.000010	-0.000540
34. 가봉	0.080910	1.164510	0.030060	0.019800
35. 아이보리	0.052680	0.770490	0.009150	0.009250
36. 마다가스카르	0.026970	-0.670390	0.004320	-0.005090
37. 니제르	0.168040	-0.278890	0.003580	-0.009350
38. 나이지리아	0.310730	-1.046260	0.107640	-0.068110
39. 세네갈	0.069300	-0.061220	0.000070	-0.000910
40. 시에라리온	0.094100	0.393360	0.003940	0.007630
41. 남아프리카	0.041310	1.451010	0.026960	0.014570
42. 수단	0.090210	-0.053370	0.000070	-0.001000
43. 탄자니아	0.082110	-1.016250	0.023200	-0.017500
44. 잠비아	0.042250	0.269060	0.000940	0.002740
45. 미얀마	0.057140	0.996100	0.016320	0.012740
46. 피지	0.017680	-0.814960	0.005040	-0.004910
47. 인도	0.102270	0.975260	0.026220	0.020380
48. 일본	0.029770	-2.611200	0.069720	-0.021060
49. 말레이시아	0.066410	-0.609470	0.006930	-0.008790
50. 파키스탄	0.093440	-0.272190	0.001870	-0.005250
51. 필리핀	0.036820	0.310290	0.001140	0.002880
52. 스리랑카	0.036690	-0.331350	0.001290	-0.003060
53. 한국	0.501700	-0.707060	0.108560	-0.085960
절단값				
하한 절단값	0.189000	2.010000	0.075000	0.614000
상한 절단값	0.283000	3.460000	1.000000	2.237000

교 취학률과 소득불평등의 상대적 영향력은 감소되었다. 병원 침대 한 개당 인구수 변수는 여전히 0.1수준에서도 유의미하지 않았다. 방정식 [3]에 대한 모형 내부 점검 결과는 <표 6>에 나타나 있다. 지레거리 검사에서 세 국가-스웨덴, 나이지리아, 한국-가 상한 절단값을 넘고 있지만, 그 국가의 스튜던트 나머지 값들은 하한 절단값도 넘지 못하고 있다. 비록 나이지리아와 한국이 높은 지레거리 값으로 인해 콕스 디 검사에서 하한 절단값을 초과하였지만, 미국과 같이 이상적인 초과 현상을 보여주지는 않았다.

1977-80년 자료도 1970년 자료와 동일한 방식으로 모형 내부 점검을 시도하였다. 영아사망률 변수를 로그로 변형시킨 종속변수의 회귀분석 결과는 다음과 같다.

$$\log IMR = 1.61^{***} - 0.38WEDU^{***} + 0.35GINI^{***} - 0.44GNP^{***} + 0.06HBED \quad [4]$$

R^2 는 0.88이다. 1970년 자료의 방정식과 마찬가지로, 여성의 초등학교 취학률, 소득불평등과 국민총생산 회귀계수는 0.01수준에서 유의미한 반면, 병원 침대당 인구수의 회귀 계수는 여전히 0.1 수준에서도 유의미하지 않다. 방정식 [2]나 [3]의 결과와 비교해 보면, 우선적으로 회귀계수의 절편값이 감소되었음을 알 수 있다. 이는 약 10여 년 동안 감소된 영아사망률값을 반영한다. 1970년 자료 분석결과와 같이, 국민총생산이 영아사망률의 변이를 가장 많이 설명하고 있지만 그 영향력은 감소되었다. 1970년 자료에서 두 번째로 큰 영향력을 발휘했던 여성의 초등학교 취학률 변수는 여전히 두 번째의 위치를 차지하고 있는 반면, 세 번째의 영향력을 발휘했던 소득불평등 계수의 중요성이 급증하였다. 공공보건 변수는 여전히 우리의 회귀모형에서 여전히 유의미하지 않는데, 그것은 공공보건이 사회경제적 발달이나 균등상태와 상호관련하여 영아사망률에 영향을 미치기 때문일 것이다. 사실 공공보건은 사회경제적 발달과 균등상태가 활발하게 진행될 때 보다 효율적으로 시행된다는 연구 결과도 있다(Kuznet, 1975; Coale and Hoover, 1958).

<표 6> 1970년 로그 영아사망률 회귀모형(미국제외, 방정식 3) 내부 점검 분석 결과

국기명	지레거리	스튜던트 나머지	Cooks'D	DFITT
1. 덴마크	0.140580	0.850160	0.027500	0.020780
2. 프랑스	0.110900	0.017600	0.000010	0.000340
3. 핀란드	0.064760	-1.449710	0.038540	-0.017840
4. 그리스	0.044110	-0.994040	0.013360	-0.009120
5. 독일	0.075640	0.934180	0.018290	0.013060
6. 이탈리아	0.023980	0.076270	0.000050	0.000470
7. 네덜란드	0.039790	-1.659090	0.034530	-0.014150
8. 노르웨이	0.085580	-0.456220	0.004870	-0.007080
9. 스웨덴	0.222470	0.740450	0.034950	0.028810
10. 영국	0.067430	-0.022540	0.000010	-0.000290
12. 아르헨티나	0.018980	1.241860	0.012250	0.006780
13. 바베이도스	0.036840	-0.970380	0.011190	-0.007850
14. 볼리비아	0.035180	0.558660	0.003590	0.004380
15. 브라질	0.042700	0.747120	0.007370	0.006700
16. 칠레	0.026160	1.669340	0.026500	0.010870
17. 콜롬비아	0.069130	-0.802760	0.012490	-0.010410
18. 코스타리카	0.033880	-0.054890	0.000030	-0.000420
19. 에콰도르	0.054390	1.607630	0.041080	0.017240
20. 엘살바도르	0.023260	0.609910	0.003300	0.003710
21. 자메이카	0.050580	-1.966000	0.058020	-0.019950
22. 멕시코	0.029420	0.371080	0.001410	0.002590
23. 파나마	0.021980	-0.205330	0.000360	-0.001220
24. 페루	0.079400	1.023700	0.022940	0.014910
25. 수리남	0.129650	0.688670	0.016590	0.015580
26. 트리니다드 토바고	0.008080	-0.149200	0.000130	-0.000580
27. 베네수엘라	0.010920	1.601360	0.015950	0.006870
28. 이라크	0.065860	-0.839700	0.013120	-0.010470
29. 이스라엘	0.081540	0.220120	0.001090	0.003280
30. 레바논	0.033980	-0.794950	0.007100	-0.006090
31. 모로코	0.044390	-0.117690	0.000190	-0.001080
32. 튀니지	0.020330	0.697040	0.004000	0.003940
33. 차드	0.188090	0.003670	0.000000	0.000120
34. 가봉	0.080870	1.265730	0.035640	0.018720
35. 아이보리	0.052610	0.816470	0.010320	0.008530
36. 마다가스카르	0.029130	-0.986060	0.009880	-0.006850
37. 니제르	0.169600	-0.122300	0.000700	-0.003590
38. 나이지리아	0.314690	-0.880050	0.077650	-0.050480
39. 세네갈	0.069110	-0.015020	0.000000	-0.000190
40. 시에라리온	0.094390	0.564710	0.008180	0.009550
41. 남아프리카	0.040960	1.661340	0.035350	0.014460
42. 수단	0.089880	-0.085510	0.000180	-0.001390
43. 탄자니아	0.082110	-1.255330	0.035520	-0.018810
44. 잠비아	0.042480	0.207160	0.000560	0.001850
45. 미얀마	0.058250	0.985670	0.016320	0.011150
46. 피지	0.019420	-1.132920	0.010320	-0.006260
47. 인도	0.101960	1.091140	0.032840	0.019770
48. 일본	0.029730	-2.937220	0.088830	-0.020670
49. 말레이시아	0.069440	-0.955110	0.017750	-0.012440
50. 파키스탄	0.093360	-0.240840	0.001470	-0.004040
51. 필리핀	0.038250	0.179310	0.000390	0.001490
52. 스리랑카	0.042400	-0.713950	0.006700	-0.006370
53. 한국	0.501370	-0.849420	0.156700	-0.089530
절단값				
하한 절단값	0.192000	2.010000	0.070000	0.620000
상한 절단값	0.288000	3.460000	1.000000	2.230000

<표7>은 방정식 [4]에 대해 유력 사례의 존재 여부에 대한 모형 내부 점검을 시도한 결과이다. 헤트 행렬 검사에서 인도네시아는 하한 절단값을, 파키스탄은 상한 절단값을 초과하는 높은 지레거리를 가지고 있음을 볼 수 있다. 스튜던트 나머지 검사에서 세계의 국가-파나마, 가봉 말레이시아가 하한 절단값을 초과하지만 상한 절단값은 넘지 않는다. 쿡스 디 검사에서는 3개국-파나마, 가봉, 인도네시아이 하한 절단값을 초과했지만 상한 절단값은 넘지 않았다. 그리고 디웰 검사에서는 어느 사례도 하한 절단값도 넘지 않았다. 결과적으로 방정식 [4]에 대한 모형내부 검사는 1977-80년 자료에는 유력 사례가 없다는 사실을 말해준다.

2) 출생시 기대수명

앞의 준비분석 단계에서 출생시 기대수명자료는 거의 대칭적으로 분포되었음을 보았기 때문에 종속변수를 변형시키지 않았다. 하지만 기대수명과 국민총생산 사이의 관계는 선형이라기보다는 비선형-특히 로그형태-으로 간주된다 (Preston, 1975, 1980; Rodger, 1979). 따라서 국민 총생산 변수를 로그 형태로 변형시킨 뒤 회귀분석을 시도하였다.

$$e^0 = 23.6^{***} + 0.57WEDU^{***} - 0.13GINI^{**} + 0.42\log(GNP)^{***} + 0.08HBED \quad [5]$$

R^2 의 값은 0.88이며, 1970년 출생시 기대수명 변이의 설명에 있어 여성의 초등학교 취학을 변수가 가장 영향력 있고, 국민총생산이 두 번째, 소득불평등이 세 번째로 영향력 있는 것으로 나타났다. 그리고 병원 침대당 인구수는 유의미한 영향력을 발휘하지 않는 것으로 나타났다.

등식[5]에서 유력 사례가 있는지를 관찰하기 위해 모형 내부 점검을 시도하였다. 그 결과는 <표8>에 나타나 있다. 두 국가-나이지리아와 한국-가 상한 절단값을 초과하는 높은 지레거리를 가지는 것으로 나타났지만, 두 국가의 스튜던트 나머지값은 하한 절단값을 넘지 않았다. 4개국-가봉, 아이보리, 탄자니아, 피지의 스튜던트 나머지값은 하한 절단값은 초과했지만 상한 절단값은 넘지 않았다. 쿡스 디의 검사에서는 3개국-가봉, 나이지리아, 탄자니아-이 하한 절단값은 넘었지만 상한 절단값은 넘지 않았다. 또 디웰 검사에서는 5개국-미국, 가봉, 나이지리아, 탄자니아, 한국-이 하한 절단값을 넘었지만 상한 절단값을 넘지는 않았다. 어느 국가도 지레거리와 스튜던트 나머지에서 상한선을 넘고, 쿡스 디나 디웰에서 상한선을 넘는 유력 사례는 나타나지 않았다.

<표 7> 1980로그 영아사망률 회귀모형(방정식 4) 내부 점검 결과

국가명	지레거리	스튜던트 나머지	Cooks'D	DFITT
1. 오스트리아	0.026520	-0.238210	0.000530	-0.001660
2. 벨기에	0.056800	0.273040	0.001210	0.003240
3. 캐나다	0.107130	0.691570	0.013700	0.014110
4. 덴마크	0.089240	-0.277590	0.001850	-0.004810
5. 프랑스	0.060350	-0.427030	0.003110	-0.005320
6. 헝가리	0.148720	0.716140	0.020550	0.019950
7. 아일랜드	0.061580	-1.023620	0.018160	-0.012960
8. 이탈리아	0.016290	-0.549490	0.002160	-0.002940
9. 네덜란드	0.065980	-0.102320	0.000190	-0.001370
10. 노르웨이	0.080350	0.093280	0.000190	0.001470
11. 포르투갈	0.037690	-0.147870	0.000260	-0.001300
12. 스페인	0.028710	-1.486110	0.021730	-0.010870
13. 스웨덴	0.151430	0.059970	0.000150	0.001700
14. 영국	0.036210	-0.221770	0.000570	-0.001890
15. 유고	0.049820	0.611840	0.005460	0.006570
16. 아르헨티나	0.022950	1.344390	0.015510	0.008600
17. 바베이도스	0.060660	0.273650	0.001280	0.003420
18. 볼리비아	0.024370	1.120090	0.011150	0.007420
19. 브라질	0.066450	-0.377620	0.002640	-0.005090
20. 칠레	0.058850	0.455220	0.003460	0.005560
21. 콜롬비아	0.110230	0.368490	0.004000	0.007720
22. 코스타리카	0.029820	-0.933670	0.008790	-0.007000
23. 엘살바도르	0.015590	-0.182750	0.000230	-0.000960
24. 과테말라	0.066880	-1.431050	0.038080	-0.019380
25. 온두라스	0.076270	-1.004920	0.021070	-0.015190
26. 멕시코	0.073110	-1.537720	0.047510	-0.022430
27. 니카라과	0.020380	0.462580	0.001720	0.002770
28. 파나마	0.084270	-2.351400	0.126220	-0.038730
29. 페루	0.051730	1.024320	0.015770	0.011310
30. 트리니다드 토바고	0.009670	0.014280	0.000000	0.000060
31. 우루과이	0.023570	0.581040	0.002940	0.003780
32. 베네수엘라	0.051690	-0.787140	0.009310	-0.008690
33. 이집트	0.075930	0.682360	0.009670	0.010280
34. 이란	0.088070	1.801950	0.077200	0.030850
35. 튀니지	0.017030	0.633140	0.002930	0.003460
36. 터키	0.029570	1.844400	0.034120	0.013750
37. 가봉	0.068440	2.533760	0.121780	0.034980
38. 가나	0.047340	0.745080	0.007790	0.007690
39. 케냐	0.074030	-0.673520	0.009220	-0.009930
40. 말라위	0.122380	1.293030	0.054690	0.029860
41. 세네갈	0.126420	-0.679460	0.015610	-0.016180
42. 시에라리온	0.130690	-0.249730	0.002180	-0.006140
43. 남아프리카	0.043860	1.484760	0.029170	0.014490
44. 수단	0.102210	-0.392550	0.004220	-0.007680
45. 탄자니아	0.044960	-0.697750	0.006560	-0.006930
46. 호주	0.066940	0.508090	0.004800	0.006890
47. 인도	0.097090	0.253850	0.001680	0.004740
48. 인도네시아	0.261260	1.190070	0.110980	0.059680
49. 일본	0.030230	-1.864350	0.035370	-0.014090
50. 말레이시아	0.022750	-2.139290	0.039070	-0.013620
51. 뉴질랜드	0.026690	-0.123410	0.000140	-0.000860
52. 파키스탄	0.421170	-0.318080	0.015860	-0.028430
53. 한국	0.133100	0.698390	0.017390	0.017470
54. 스리랑카	0.081830	-0.542570	0.006540	-0.008710
55. 태일랜드	0.024680	-0.831240	0.006190	-0.005550
절단값				
하한 절단값	0.182000	2.020000	0.070000	0.060000
상한 절단값	0.273000	3.460000	1.000000	2.237000

〈표 8〉 1970 기대수명 회귀모형(방정식 5) 내부 점검 결과

국가명	지레거리	스튜던트 나머지	Cooks' D	DFITT
1. 덴마크	0.069140	-0.050320	0.000050	-0.019070
2. 프랑스	0.070360	0.393890	0.003040	0.151440
3. 핀란드	0.048310	0.080280	0.000090	0.022960
4. 그리스	0.031010	0.901460	0.008530	0.189690
5. 독일	0.048020	-0.174920	0.000440	-0.049810
6. 이탈리아	0.023830	0.431110	0.001660	0.077360
7. 네덜란드	0.034630	0.698300	0.005510	0.157890
8. 노르웨이	0.061200	0.313300	0.001710	0.107540
9. 스웨덴	0.083840	0.343420	0.002700	0.153110
10. 영국	0.054470	0.352580	0.001970	0.110450
11. 미국	0.123260	-1.333010	0.058880	-0.841090
12. 아르헨티나	0.015870	-0.096850	0.000070	-0.014080
13. 바베이도스	0.033460	0.971540	0.010420	0.214730
14. 볼리비아	0.045710	-0.469300	0.003040	-0.128840
15. 브라질	0.057270	0.005850	0.000000	0.001910
16. 칠레	0.018070	-0.851840	0.005570	-0.131840
17. 콜롬비아	0.068760	0.889200	0.015190	0.335430
18. 코스타리카	0.032630	0.847490	0.007800	0.184260
19. 에콰도르	0.065380	-1.121940	0.023160	-0.406130
20. 엘살바도르	0.023180	0.196520	0.000340	0.034720
21. 자메이카	0.049600	1.240160	0.022610	0.361760
22. 멕시코	0.031230	-0.129440	0.000180	-0.027360
23. 파나마	0.017380	0.269240	0.000550	0.040870
24. 페루	0.083020	-1.103590	0.027630	-0.487890
25. 수리남	0.107700	-0.780880	0.017670	-0.434860
26. 트리니다드 토바고	0.008390	0.801300	0.003600	0.091060
27. 베네수엘라	0.017950	-1.109700	0.009420	-0.171190
28. 이라크	0.079750	0.607970	0.008090	0.259670
29. 이스라엘	0.078280	-0.015020	0.000000	-0.006310
30. 레바논	0.032360	0.005060	0.000000	0.001090
31. 모로코	0.045690	1.316910	0.023940	0.361430
32. 튀니지	0.024240	0.407950	0.001500	0.073920
33. 차드	0.187260	0.033860	0.000060	0.032210
34. 가봉	0.081770	-3.214520	0.231260	-1.402720
35. 아이보리	0.047780	-2.048610	0.059940	-0.581120
36. 마다가스카르	0.075900	-1.501600	0.047210	-0.614980
37. 니제르	0.178210	0.224740	0.002480	0.203250
38. 나이지리아	0.316190	1.036680	0.108310	1.751520
39. 세네갈	0.081910	-1.078880	0.026090	-0.471450
40. 시에라리온	0.106230	0.161230	0.000740	0.088660
41. 남아프리카	0.044780	-0.785170	0.008380	-0.212350
42. 수단	0.091780	-0.141410	0.000500	-0.068230
43. 탄자니아	0.112630	2.068660	0.129590	1.200260
44. 잠비아	0.039470	-1.289120	0.020590	-0.318660
45. 미얀마	0.115430	-0.552160	0.009460	-0.327720
46. 피지	0.015000	2.401390	0.040430	0.340250
47. 인도	0.104980	-0.823610	0.019180	-0.448070
48. 일본	0.024700	0.824430	0.006190	0.151030
49. 말레이시아	0.060130	-0.556530	0.005310	-0.188380
50. 파키스탄	0.095140	0.855020	0.018810	0.425820
51. 필리핀	0.038980	-0.395660	0.001920	-0.096960
52. 스리랑카	0.095690	1.421910	0.052320	0.711830
53. 한국	0.502030	-0.645320	0.090550	-1.996910
절단값				
하한 절단값	0.189000	2.010000	0.075000	0.614000
상한 절단값	0.283000	3.460000	1.000000	2.237000

1977-80년도 출생시 기대수명에 대한 회귀분석은 다음과 같다.

$$e^0 = 43.6^{***} + 0.51WEDU^{***} - 0.31GINI^{***} + 0.291\log(GNP)^{***} + 0.01HBEDI [6]$$

R^2 의 값은 0.84이다. 여성의 초등학교 취학률은 1980년대에서 기대수명의 변이에 가장 큰 영향을 미치고 있다. 반면 1970년대에 두 번째로 영향을 미쳤던 국민총생산의 상대적 중요성은 크게 감소하여 세 번째가 된 반면, 소득불평등계수는 상대적 중요성이 증대하여 두 번째로 영향을 미치는 것으로 나타났다. 그러나 병원 침대 당 인구수는 등식 [5]와 마찬가지로 통계적으로 유의미하지 않았다. 이는 영아사망률에서와 마찬가지로, 공공 보건의 사회경제적 발달이나 사회경제적 균등 상태와 상호 관련하여 기대수명의 변이에 영향을 미치기 때문이라고 생각된다.

<표9>는 방정식 [6]에 대한 모형 내부 검사 결과이다. 헤트 행렬 검사에서 두 국가-인도네시아와 파키스탄-가 높은 지레거리를 가지는 것으로 나타났다. 특히 파키스탄은 상한 절단값을 초과하는 높은 지레거리를 보여준다. 스튜던트 나머지의 경우 하한 절단값은 2.22이고 상한 절단값은 3.46이다. 가봉은 상한 절단값을 초과하는 매우 높은 스튜던트값을 보여주었고, 스리랑카는 하한 절단값 값을 넘는 스튜던트 값을 보여주었다. 세 번째 열의 쿡스 디 분석에서 6개국-멕시코, 가봉, 말라위, 인도네시아, 파키스탄, 스리랑카-이 하한 절단값을 넘었고, 특히 가봉은 상한 절단값까지 초과하였다. 디휠 검사에서는 여덟 개의 사례가 하한 절단값을 넘었지만 상한 절단값을 넘는 사례는 없었다. 결과적으로 지레거리와 스튜던트 나머지에서 모두 절단값을 넘는 높은 값을 보여주는 사례는 없었다. 따라서 이 모형의 진단 분석은 유력 사례가 없음을 말해준다.

VI. 마치는 말

본 연구의 주목적은 1970년 및 1977-80년 시기 사망률 저하에 영향을 미치는 세 요소 중 어느 것이 가장 커다란 영향을 미쳤는지를 분석하는 것이다. 사회경제적 발달, 공공보건 수준, 사회경제적 균등상태는 분명 거시적 현상이므로 분석 단위 역시 거시 단위가 되어야 한다. 따라서 본 연구는 분석 단위를 국가로 규정

〈표 9〉 1980 기대수명 회귀모형(방정식 6) 내부 점검 결과

국가명	지레거리	스튜던트 나머지	Cooks'D	DFITT
1. 오스트리아	0.027770	-0.016360	0.000000	-0.003110
2. 벨기에	0.043210	-0.424920	0.002360	-0.108670
3. 캐나다	0.056080	-0.085430	0.000120	-0.026610
4. 덴마크	0.048390	0.224950	0.000720	0.062550
5. 프랑스	0.040350	0.583560	0.004230	0.142060
6. 헝가리	0.117370	-1.001160	0.031430	-0.589040
7. 아일랜드	0.045320	-0.300850	0.001230	-0.079670
8. 이탈리아	0.019990	0.286540	0.000650	0.045010
9. 네덜란드	0.055760	-0.095670	0.000150	-0.029670
10. 노르웨이	0.059030	-0.103010	0.000180	-0.033420
11. 포르투갈	0.015400	0.552170	0.002120	0.076110
12. 스페인	0.016560	0.380950	0.001040	0.054370
13. 스웨덴	0.061580	0.271310	0.001280	0.091040
14. 영국	0.036740	-0.325890	0.001230	-0.074310
15. 유고	0.035500	-0.154370	0.000270	-0.034380
16. 아르헨티나	0.009970	0.282110	0.000460	0.032510
17. 바베이도스	0.042140	-0.242020	0.000750	-0.060780
18. 볼리비아	0.033850	-1.805300	0.035770	-0.389310
19. 브라질	0.064170	0.882020	0.013960	0.306000
20. 칠레	0.029260	0.344680	0.001180	0.067620
21. 콜롬비아	0.064570	0.104880	0.000200	0.036570
22. 코스타리카	0.025270	1.256730	0.014350	0.225330
23. 엘살바도르	0.015140	1.276800	0.011240	0.174640
24. 과테말라	0.071420	1.056140	0.021960	0.400260
25. 온두라스	0.083450	0.716370	0.011610	0.310000
26. 멕시코	0.117570	1.698240	0.090600	1.000780
27. 니카라과	0.022320	-1.033810	0.009020	-0.172510
28. 파나마	0.049710	1.769860	0.045630	0.502290
29. 페루	0.052910	-1.478350	0.033450	-0.440070
30. 트리니다드 토바고	0.008550	0.565670	0.001760	0.061860
31. 우루과이	0.015400	0.462640	0.001490	0.063780
32. 베네수엘라	0.068920	0.775980	0.011490	0.285500
33. 아집트	0.082160	0.356070	0.002830	0.152010
34. 이란	0.170740	-0.632100	0.018610	-0.535110
35. 튀니지	0.009850	0.014920	0.000000	0.001710
36. 터키	0.039760	-0.347750	0.001490	-0.083780
37. 가봉	0.075030	-4.290270	0.378430	-1.694860
38. 가나	0.053430	-0.556130	0.004770	-0.166800
39. 케냐	0.111730	-0.258280	0.001990	-0.145170
40. 말라위	0.181930	-1.684110	0.141910	-1.520740
41. 세네갈	0.168150	-0.891210	0.036380	-0.742930
42. 시에라리온	0.160880	0.859570	0.032230	0.685550
43. 남아프리카	0.044010	-0.622830	0.005150	-0.161430
44. 수단	0.121140	-0.639600	0.013240	-0.387640
45. 탄자니아	0.072460	0.041990	0.000040	0.016110
46. 호주	0.043670	-0.087050	0.000100	-0.022430
47. 인도	0.100300	-0.116430	0.000360	-0.059300
48. 인도네시아	0.286600	-1.476190	0.191070	-2.177660
49. 일본	0.033410	0.599490	0.003910	0.128170
50. 말레이시아	0.018900	0.510420	0.002010	0.077850
51. 뉴질랜드	0.029720	-0.035550	0.000010	-0.007040
52. 파키스탄	0.437090	0.808500	0.109270	2.012690
53. 한국	0.130670	-1.093160	0.041800	-0.711790
54. 스리랑카	0.138840	2.095340	0.163560	1.446160
55. 타일랜드	0.035820	0.793370	0.007190	0.177780
절단값				
하한 절단값	0.182000	2.020000	0.070000	0.600000
상한 절단값	0.273000	3.460000	1.000000	2.237000

하였고, 1970년 사망률 분석에는 53개국, 1977-80년 분석에는 55개국이 포함되었다. 본 연구에서 분석된 변수는 영아사망률, 출생시 기대수명, 여성의 초등학교 취학률, 기니지수, 국민총생산, 병원침대당 인구수이다. 본 연구에서 영아사망률과 출생시 기대수명은 사망률 지표로, 여성의 초등학교 취학률과 기니지수는 사회경제적 균등상태 지표로, 국민총생산은 사회경제적 발달지수로, 병원 침대 당 인구수는 공공보건 지수로 간주하였다. 분석 방법으로는 탐색적 자료 분석 기법이 사용되었다. 발견된 사실은 다음과 같다.

첫째, 영아 사망률과 세 요소의 관계는 선형이 아니라 비선형이다. 가장 적합한 비선형 형태는 로그 함수 형태이다. 영아 사망률의 변이에 가장 영향을 미치는 요소로는 국민총생산으로 나타났고, 여성의 초등학교 취학률이 두 번째로, 그 다음은 기니계수이었다. 반면 병원침대당 인구수는 통계적으로 유의미한 영향을 보여주지 않았다. 이는 아마 공공보건 요소가 영아 사망률에 직접적인 영향을 미치기보다는 사회경제적 발달 요소나 사회경제적 균등 상태와 상호 관련하여 영향을 미치기 때문이라 생각된다. 이런 결과는 1970년대 자료와 1977-80년대 자료 모두에서 동일하였다.

둘째, 1970년 영아사망률 분석 자료에서 미국이 유력 사례임이 밝혀졌다. 그것은 미국이 경제적 발달이나 균등상태보다는 다른 사회문제로 인해 다른 서구의 선진국보다 높은 영아사망률을 보여주기 때문이다.

셋째, 출생시 기대수명은 여성의 취학률, 기니계수 등과 같은 변수와는 선형 관계를 가지는 반면 국민총생산(GNP) 변수와는 비선형 관계를 가진다. 영아사망률 변수와는 달리 출생시 기대수명의 변이에는 여성의 초등학교 취학률이 국민총생산보다 더 커다란 영향을 미쳤다. 1970년 자료와 1977-80년 자료에서 동일한 결과를 보여주었다. 반면 기니계수의 영향력은, 영아사망률과 마찬가지로, 1970년대보다 1977-80년대에 보다 커다란 영향을 미쳤다. 이는 사회경제적 요인이 사망률을 저하시키는 과정에서 어느 수준까지는 사회경제적 발달이 결정적으로 영향을 미치지만, 그 수준을 넘어 계속 사망률이 저하되기 위해서는 사회경제적 균등상태가 수반되어야 한다는 사실을 의미한다.

참고문헌

- Belsley, D.A., E. Kuh and R.E. Welsch, (1980), *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, New York: John Wiley.
- Bollen, K.A. and R.W. Jackman (1990), "Regression Diagnostics: An Expository Treatment of Outliers and Influential Cases," in Fox, J. and J.S. Long (eds.), *Modern Methods of Data Analysis*, Newbury Park: Sage Publication, 257-291.
- Bornschieer, V. and C. Chase-Dunn (1985), *Transnational Corporations and Underdevelopment*, New York: Praeger.
- Box, G.E.P. and D.R. Cox (1964), "An Analysis of Transformations," *Journal of the Royal Statistical Society, Series B* 26: 211-252.
- Caldwell, J.C. (1986), "Routes to Low Mortality in Poor Countries" *Population and Development Review* 12: 171-220.
- Chambers, J.M., W.S. Cleveland, B. Kleiner and P.A. Tukey (1983), *Graphical Methods for Data Analysis*, Belmont, CA: Wadsworth.
- Coale, A.J. and E. Hoover (1958), *Population Growth and Economic Development in Low-Income Countries*, Princeton: Princeton University Press.
- Duleep, H.O. (1989), "Measuring Socioeconomic Mortality Differentials Over Time," *Demography* 26: 345-351.
- Evans, R.J. (1987), *Death in Hamburg*, London: Penguin Books.
- Fox, J. (1990), "Describing Univariate Distribution," in Fox, J. and J.S. Long (eds.), *Modern Methods of Data Analysis*, Newbury Park: Sage Publication, 58-125.
- Fox, J. (1991), *Regression Diagnostics*, Newbury Park: Sage Publication.
- Franzosi, R. (1994), "Outside and Inside the Regression "Block Box" from Exploratory to Interior Data Analysis," *Quality and Quantity* 28: 21-53.
- Frederikson, H. (1961), "Determinants and Consequences of Mortality Trends in Ceylon," *Public Health Reports* 76: 659-63.

- Gwatkin, D.A. (1980), "Indications of Changing in Developing Country Mortality Trends: The End of an Era?" *Population and Development Review* 6: 615-644.
- Hoaglin, D.C. and R.E. Welsch (1978), "The Hat Matrix in Regression and ANOVA," *American Statistician* 32: 17-22.
- Jackman, R.M. (1985), "Cross-National Statistical Research and the Study of Comparative Politics," *American Journal of Political Science* 29: 161-182.
- Kuznets, S. (1975), "Population Trends and Modern Economic Growth: Notes towards an Historical Perspective," in *The Population Debate: Dimensions and Perspectives* papers of the World Population Conference pp. 425-433. New York: United Nations.
- Leinhardt, S. and S.S. Wasserman (1979), "Exploratory Data Analysis: An Introduction to Selected Method," in Schuessler (ed.) *Sociological Methodology*, D.C.: American Sociological Association, 311-365.
- McKeown, T. (1976), *The Modern Rise of Population*, New York: Academic Press.
- _____, R.G. Brown and R.G. Record (1972), "Interpretation of the Modern Rise of Population in Europe," *Population Studies* 26: 345-382.
- McKeown, T., R.G. Record and R.D. Turner (1976), "An Interpretation of the Decline of Mortality in England and Wales During the Twentieth Century," *Population Studies* 29.
- Meegama, S.A (1967), "Malaria Eradication and its Effect on Mortality." *Population Studies* 26: 207-238.
- Omran, A.R. (1973), "Epidemiologic Transition." in J. Ross (ed.), *International Encyclopedia of Population*, New York: Free Press, 172-183.
- Palloni, A. (1981), "Mortality in Latin America: Emerging Patterns," *Population and Development Review* 7: 623-648.
- Pampel, F. and V. Pillai (1986). "Patterns and Determinants of Infant Mortality in Developed Nations, 1950-1975," *Demography* 23: 525-542.
- Pamuk, E.R. (1985), "Social Class Inequality in Mortality from 1921 to 1972 in England and Wales," *Population Studies* 39: 17-31.

- Paukert, F. (1973), "The Distribution of Income at Different Levels of Developments," *International Labour Review* 108: 97-125.
- Potter, L.B. (1991), "Socioeconomic Determinants of White and Black Males' Life Expectancy Differential," *Demography* 28: 303-321.
- Preston, S. (1975), "The Changing Relation Between Mortality and Level of Economic Development," *Population Studies* 29: 231-248.
- _____ (1976), *Mortality Patterns In National Populations: With Special Reference to Recorded Causes of Death*, New York: Academic Press.
- _____ (1977), "Mortality Trend," *Annual Review of Sociology* 3: 163-178.
- _____ (1980), "Causes and Consequences of Mortality Declines in Less Developed Countries During the Twentieth Century," in R. Easterlin (ed.), *Population and Economic Change in the Developing Countries*, Chicago: University of Chicago Press, 289-360.
- _____ (1985), "Mortality and Development Revisited," *Population Bulletin of the United Nations* No.18: 34-40.
- _____ and M.R. Haines (1991), *Fatal Years: Child Mortality in the Late Nineteenth-century America*, Princeton: Princeton University Press.
- Razzel, P.E. (1974), "An Interpretation of the Modern Rise of Population in Europe-A Critique." *Population Studies* 28: 5-17.
- Rodgers, G.B. (1979), "Income and Inequality as Determinants of Mortality: An International Cross-Section Analysis." *Population Studies* 33: 343-352.
- Schellekens, J. (1989), "Mortality and Socio-Economic Status in Two Eighteenth-Century Dutch Villages," *Population Studies* 43: 391-404.
- Simpson, M. (1990), "Political Rights and Income Inequality: A Cross-National Test," *American Sociological Review* 51: 682-693.
- Szreter, S. (1988), "The Importance of Social Intervention in Britain's Mortality Decline, 1850-1914: A Re-Interpretation of the Role of Public Health," *Social History of Medicine* 1: 1-37.
- World Bank (1983), *World Tables: The Third Edition Volume II Social Data*, Baltimore: The John Hopkins University Press.
- _____ (1989), *World Tables: 1988-89 Edition*, Baltimore: The John Hopkins University Press.