

제3장

지역통계 작성을 위한 소지역 추정 방법 연구

권순필 · 김서영 · 이정희

제1절 서론

1. 연구 목적

통계청은 지난 수십 년간 국가통계 전반에 걸쳐 큰 틀에서 필요 통계를 지속적으로 작성하고 발전시켜오고 있다. 사회적인 인식도 정책 수립을 위해 직관과 경험에 의존하던 과거보다는 과학적인 데이터에 기반한 정책의 개발·개선·시행을 요구하고 있으며, 공적인 영역 바깥에서도 우리 사회 전반의 사회·경제적 변화들은 이에 대한 해석과 대응을 위한 다양한 통계를 요구하고 있다. 그러나 현실적으로는 이처럼 다양하면서 급속도로 늘어나고 있는 통계 수요에 일일이 대응하도록 통계를 공급할 수는 없다. 공식통계의 대부분을 조사통계가 차지하는 현실에서 예산과 자원이 한정되어 있다는 것을 무시할 수 없기 때문이다.

지방자치제 실시 이후 지역통계에 대한 수요가 지속적으로 증가하고 있다. 지방 정책 수립의 주체가 지자체가 된 이후 지역정책을 수립하는데 있어서 지역통계의 필요성이 상당히 높기 때문이다. 그러나 지역 발전 정책 수립에 필요한 분야별 지역통계는 여전히 부족한 실정이다. 2010년을 기준으로 생산되고 있는 지역통계의 대부분은 인구현황, 기후, 주택종류, 지방세 세입세출, 사업체수 등 지역의 기본사항만을 파악하는 기본통계이다(통계청, 2010). 지자체의 입장에서는 물론 이들 통계를 이용한 큰 틀에서의 정책 수립이 중요하다. 그러나 보다 직접적이고 다양한 이해관계를 가지는 주민들이 체감



할 수 있는 정책을 수립하고 시행하는 것 역시 중요하기 때문에 환경, 보건복지, 지역개발 등 다양한 분야에서 지역통계의 부재를 크게 느끼고 있다.

통계청은 지역통계의 개선과 확충을 위하여 다양한 측면에서 지원 노력을 수행하고 있다. 일례로, 지방 정부의 수요가 상당한 고용 분야 지역통계 생산을 위하여 지역별 고용조사를 2010년 3분기부터 연간 조사에서 분기 조사로 확대하였고, 광역시도 별로 자체 시군구 GRDP(Gross regional domestic product; 지역내총생산)를 작성할 수 있도록 기술 지원을 하고 있다. 또한, 전국 수준에서 제공되고 있는 서비스업생산지수는 2011년 시험 조사를 거쳐 시도까지 확대 제공하는 계획이 확정되었다. 그리고 포항시광업제조업동향조사와 포항시사회조사 2종을 신규 개발하는 등 지자체의 지역통계 직접 생산 역량 강화에도 힘을 기울이고 있다(통계청, 2010).

그러나 이런 적극적인 노력에도 불구하고 통계 생산에 있어 가장 필수적인 예산과 자원이 한정되어 있다는 점은 극복할 수 없는 문제이다. 또한 지자체들의 통계에 대한 인식 역시 매우 낮은 수준이다. 최근 들어 보다 체계적으로 수집된 지역 자료를 요구하는 지자체장이 늘어나고는 있지만 그 동안 통계 업무는 중요하지 않은 업무라는 인식이 강했고, 눈에 띄는 정책 업무들에 비해 후순위로 밀리면서 중앙정부의 위임사무로만 생각되는 경향이 강하다.

이런 현실적인 어려움을 극복하기 위하여 본 연구에서는 소지역 추정 기법 등을 이용하여 지역통계 수요에 대응할 수 있는 방법을 모색하게 되었다. 즉, 조사를 통한 지역통계의 직접 생산보다는 현재 제공되고 있는 다양한 통계 자료들을 이용하여 비용 효율적이면서 적시에 제공될 수 있는 추정 통계를 생산하는 등의 검토가 필요하기 때문이다. 지자체에도 조사 통계 방법에 대한 지원뿐만 아니라 통계의 가공·추정 등에도 적용될 수 있는 기법적 측면의 지원이 늘어난다면 지자체의 자체 통계 역량 강화 효과도 얻을 수 있을 것이다.

통계개발원에서는 지난 몇 년간 고용 분야에 대해 깊이 있는 소지역 추정 연구를 수행해왔다. 본 연구에서는 국가통계의 수요자들이 가장 필요하다고 느끼는 분야(통계청, 2010) 중에서도 가계소득소비 분야에 대하여 시도별 통계 생산 가능성 검토를 위한 소지역 추정을 연구하고자 한다.

2. 연구 내용과 범위

“2011 지역통계발전계획(안)(통계청, 2010)”의 증장기 수요 목록 17종을 확인하고, 대응하는 지역 통계를 생산하기 위한 소지역 추정 검토 방법을 제안한다. 소지역 추정 가능성 검토는 해당 통계의 개요, 즉, 표본 크기, 조사 항목, 추정을 시도해 볼 수 있는 소지역(소영역) 분야 등을 확인함으로써 수행할 수 있을 것이다. 또한, 소지역 추정은 조사통계처럼 다양한 항목의 통계를 생산할 수는 없기 때문에 특별히 활용성이 높은 분야에 기초하여 검토되어야 한다.

그 중 가장 수요 순위가 높은 가계소득소비 분야의 가계동향조사에 대한 소지역 추정 가능성을 검토하고, 더불어 이 분야에 대한 국·내외 국가통계기관 작성 사례를 확인해본다. 마지막으로 검토 결과를 활용하여 가계의 소득소비에 대한 소지역 추정을 직접 시도한다. 이 분야에 대한 소지역 추정 연구는 이미 일정 정도의 발전이 있어 왔기 때문에 이 추정을 보다 개선시킬 수 있는 보조정보 확보에 노력을 기울이며, 자주 비교의 대상이 되곤 하는 지역내총생산 결과(GRDP)와의 구체적인 비교를 통해 추정 결과에 대한 평가를 한다.

3. 연구 기대 효과와 제한

이 연구는 조사 통계에 대응하는 지역 통계를 생산하기 위한 소지역 추정 검토 방법의 안을 제안하고, 가계소득소비 분야에 대한 추정 통계를 직접 시범 생산한다는 데 그 의의가 있다. 특히, 추정 결과에 대한 설명력을 높이기 위하여 이전보다 다양하고 많은 보조정보를 확인한다는 점이 그렇다. 이를 통해 기존 통계의 표본크기로 인한 제한, 원시자료 사용의 제한 등으로 인해 지역통계 생산에 어려움을 겪어왔던 기관들이 본 연구에서 제공된 가계소득소비 분야 소지역 추정 통계의 실례를 통하여 스스로의 통계 작성 의지와 능력을 고양시킬 수 있는 계기가 될 것으로 기대한다.

그러나 본 연구에서는 국민여행실태조사와 보육실태조사에 대한 원시 자료(micro data) 대신 공표된 보고서(macro data) 내용을 검토의 근거로 삼았기 때문에 향후 원시 자료에 대한 구체적인 검토가 이뤄진다면 소지역 추정 방법 검토 결과는 달라질 수도 있다.

본 연구는 제2절에서 지역통계의 구체적인 수요 내용 확인과 가계소득소비 분야에 대한 소지역 추정 적용 검토를 하고 제3절에서는 가계소득소비 분야에 대한 소지역 추정을 시도하고, 마지막으로 결론과 시사점을 기술하면서 마무리한다.



제2절 지역통계 수요와 가계소득소비 분야 소지역 추정 검토

1. 지역통계 수요

늘어나는 지역통계 수요에 비해 현재 지역발전정책 수립에 필요한 분야별 지역통계는 상당히 부족한 상황이다(통계청, 2010). <표 3-1>의 지역통계 부문별 작성 현황을 살펴보면, 2010년 12월 현재 생산되고 있는 지역통계는 모두 514종인데, 그 중에서 대부분은 인구현황, 기후, 주택종류, 학교총개황, 지방세 세입세출, 사업체수 등 지역의 기본사항만을 파악하는 기본통계(47.2%)이다. 뒤이어 큰 차이로 보건·사회·복지 분야 통계(13.6%)가 생산되고 있다. 그러므로 각 지자체들은 환경, 보건복지, 지역계정 등 다양한 분야에서 지역통계의 부재를 늘 크게 느끼고 있다.

<표 3-1> 지역통계 부문별 작성 현황

(2010년 12월 1일 현재)

부 문	종류별(종)		부 문	종류별(종)	
	통계수	구성비(%)		통계수	구성비(%)
계	514	100.0			
인구	19	3.7	교통·정보통신	18	3.5
고용·임금	8	1.6	도소매·서비스	8	1.5
물가·가계·소비	6	1.2	경기·기업경영	31	6.0
보건·사회·복지	70	13.6	국민계정·지역계정	17	3.3
환경	9	1.8	재정·금융	2	0.4
농림·수산	37	7.2	교육·문화·과학	12	2.3
광공업·에너지	12	2.3	기본통계	244	47.5
건설·주택·토지	19	3.7	기타	2	0.4

출처 : 통계청, 2011 지역통계발전계획(안), 2010.12

지역통계에 대한 구체적인 수요 내용은 “2011 지역통계발전계획(안)”의 중장기 수요 목록에 잘 나타나있다(<부표> 참고). 우선 정리된 중장기 수요 목록은 17종으로, 필요 분야는 가계소득소비, 문화, 복지, 주택, 사회, 인구, 교육, 교통 등 다양하다. 각 통계들의 기본 공표 범위는 인터넷이용실태조사와 배출시설단속조치현황이 광역시도인 것을 제외하고는 모두 전국이다. 이런 넓은 공표 범위는 자신들이 속한 지역의 상황에 맞는 정책을 세우고자 하는 지자체나 소속 연구기관에게는 큰 문제가 될 수 있다.

수요의 우선순위가 가장 높은 통계는 가계소득소비 분야의 가계동향조사이다. 이 조사를 요청한 기관들은 주로 서울·울산·강원·경남 등 지방자치단체, 한국은행·국토연구원·한국조세연구원 등 국가기관, 서울시정개발연구원·대구경북연구원·전북발전연구원 등 지방자치단체의 연구기관들이다. 이들은 가계동향조사의 지역자료(권역별) 공개를 가장 많이 요구하였고, 특히 경남, 울산, 전북발전연구원은 시군단위 GRDP 지출 및 분배계정 추계 기초자료로 활용하고, 지역주민의 사회복지정책수립 및 예산편성 기초자료로 활용하고자 해당 자료를 요청하였다.

그 다음으로는 문화체육관광부의 국민여행실태조사와 보건복지부의 보육실태조사에 대한 수요가 많았다. 국민여행실태조사의 경우 지역 경제에서 관광업의 비중이 큰 강원 이 세부 비용 단위로 시도별 숙박, 당일여행으로 구분하여 자료를 제공받기를 원하는 등 모든 지자체가 관광 통계에 대해 높은 수요를 가지고 있었다. 더불어 지자체가 직접 작성하는 통계에는 허수가 많다는 인식을 가지고 있어 중앙기관이 지역자료도 정확히 작성해서 제공하기를 기대하고 있다. 이를 통하여 지자체들은 지역별 관광실태와 시민들의 여행실태를 파악하여 지역 관광정책 수립에 활용하고자 한다. 보건복지부 보육실태조사의 지역통계를 필요로 하는 기관은 서울, 울산, 충남, 전북, 경남 등 지자체들로 보육실태조사 자료를 이용하여 지역의 아동보육실태를 파악하고 저출산 대책 수립에 활용하려고 한다. 보육시설 및 보육 수요량 조사 결과는 자치단체별 보육사업계획수립 시 유용하게 활용될 수 있다.

그 외에도 국토해양부의 주거실태조사, 농촌진흥청의 농촌생활지표조사, 보건복지부의 전국 결혼 및 출산 동향조사 등과 같이 주민의 생활 전반에 대한 판단 지표로서의 자료들을 요구하고 있다.

2. 소지역 추정 검토 방법

수요 통계에 대한 소지역 추정 검토 방법은 다음과 같은 항목을 기준으로 대략적으로 우선 검토할 수 있다. 이해를 돕기 위하여 수요 통계 중 ① 가계동향조사, ② 국민여행실태조사, ③ 보육실태조사를 이용하여 가능한 실례를 들어본다. 소지역 추정 가능성 검토 결과 소지역 추정이 가능하다고 판단되는 조사와 항목은 해당 통계를 필요로 하는 기관이나 지자체에서 실제로 추정 통계를 생산할 수 있을 것이다. 물론 구체적인 추정 방법의 결정은 보다 구체적인 자료 검토를 통해서 이루어져야 한다. 그러나 기존 통계의 표본크기로 인한 제한, 원시자료 사용의 제한 등으로 인해 지역통계 생산에 어려움을 겪어왔던 기관들이 본 연구에서 제공된 가계소득소비 분야 소지역 추정 통계의 실례를 통하여 자체적으로 지역 통계를 생산할 수 있는 계기가 될 수도 있다.



• 표본크기

소지역 추정에서 소지역은 일반적으로 표본 규모가 작은 지역을 말한다. 그렇더라도 안정적인 추정 결과를 얻기 위해서는 전국 표본의 크기가 어느 정도 규모 이상은 되어야 하며 관심값에 대해서도 포착이 가능해야 한다. 그러므로 표본의 지역적 편중 현상도 가급적 발생하지 않아야 한다. 관심값을 포착할 수 없는 경우에는 신뢰할 수 있는 추정 결과를 얻기가 어렵기 때문이다. 지역별로 배정된 표본의 크기가 너무 작다면 지역의 수요에 맞추어 표본의 크기를 조금 더 증가시켜 줄 필요가 있다.

① 가계동향조사

- 표본크기 : 전국 8,700여 가구로 16개 광역시도에 평균적으로 544가구
- 표본오차 : 소득, 소비 1순위의 대부분 항목이 3.0% 이하(2010년 전국 연간 CV)
 - ※ 출처 : 통계청, KOSIS
- 광역시도에 대한 소득 소비 1수준 항목까지는 추정 가능

② 국민여행실태조사

- 표본크기 : 전국 2,812 가구의 만 15세 이상 동거 가구원 약 6천여 명
- 표본오차 : 항목별로 대체로 25% 미만(2010년)
 - ※ 출처 : 문화체육관광부, 2010년 국민여행실태조사 보고서(2011.05)
- 표본크기는 작지 않으나, 광역시도별×월별 통계처럼 보다 작은 소영역에 대한 통계 생산을 위해서는 표본을 조금 더 확보할 필요가 있다고 판단됨

③ 보육실태조사

- 표본크기 : 약 15,000가구 (전국 257개 조사구, 조사구당 평균 60가구)
 - 가구 : 총 4,901가구에서 7,455 명 아동조사
 - 보육시설 : 전국 대상 3201개 어린이집
 - ※ 출처 : 보건복지가족부, 2009년 전국보육실태조사 - 가구조사 보고서(2009.11)
- 표본크기는 작지 않으나 조사대상인 아동의 지역 편중 우려

- 조사 항목

소지역 추정은 “양(量)” 추정에 검토 초점을 맞춘다. 의식 조사(혹은 만족도 조사) 항목 추정은 어려움이 있다. 왜냐하면 보조정보가 없으면 소지역 추정 방법 적용이 불가능하고 대지역에 대한 비추정도 여의치 않기 때문이다. 또한 만족도 조사에 대한 추정 결과는 해석에 있어 “양”을 추정하는 것에 비해 이론의 여지가 많다.

① 가계동향조사

- "가구당 월평균 소득 389만 8천원"과 같은 양 조사

② 국민여행실태조사

- “2006년 1년 동안 우리나라 만 15세 이상 국민의 1인 평균 국내여행 참가 횟수는 7.20회“와 같은 양 조사와

- “국내여행에 대한 전반적 만족도는 5점 만점 기준 3.79점”과 같은 만족도 조사 동시 수행

③ 보육실태조사

- “미취학 아동의 경우 33.7%는 아무 곳도 다니지 않지만 57.7%의 아동은 1종의 기관 서비스를 이용하고~” 등의 양 조사와

- “보육시설 비용에 대해서 매우 만족 13.4%, 만족 32.8%, 보통 33.9% ~” 등의 만족도 조사 동시 수행

- 소지역(소영역) 추정 범위

세 통계에 대하여 소지역 추정 분야를 보도자료, 보고서와 KOSIS 제공 범위 등을 기준으로 검토하여 선택할 수 있다. 이들 자료에서는 다양한 항목들 중에서도 사용 빈도가 높거나 정책 판단에 중요한 역할을 하는 부분이 주로 인용되기 때문이다. 물론 가장 중요한 판단 기준은 수요자의 요구 수준일 것이다.



① 가계동향조사

- (현재) 전국, 도시 → (소지역) 광역시도

② 국민여행실태조사

- (현재) 전국 → (소지역) 광역시도 혹은 조금 더 큰 관광권역

③ 보육실태조사

- (현재) 전국 → (소지역) 광역시도

- 보조정보

최근의 소지역 추정은 모형을 주로 이용한다. 모형추정의 경우에는 좋은 보조정보의 존재가 중요하다. 왜냐하면 소지역의 적은 표본으로 인해 크게 발생하는 변동을 좋은 보조정보의 지지 혹은 “borrow strength”로 줄일 수 있기 때문이다. 이렇게 보조정보(혹은 covariate)로 사용될 수 있는 자료는 일반적으로 센서스 자료나 국세청 자료처럼 표본오차가 없는 자료이면서 다음의 조건도 최대한 만족해야 한다.

- 1) 가급적 보조정보의 포함 범위(커버리지)가 주요정보만큼은 넓어야 한다.
- 2) 지속적인 추정을 위해 소지역별 자료 입수가 수월하면서 계속 입수가 가능해야 한다.
- 3) 이용하는 보조정보의 수가 너무 많거나 적지 않도록 적정 수의 보조정보를 확보한다.
- 4) 표본조사 자료라도 표본의 규모가 크거나 비표본오차가 적을 거라는 기대를 가질 수 있는 자료라면 복합 추정량 형태로 그 이용을 고려해 볼 수도 있다.

① 가계동향조사

- 가계금융조사 등 조사통계, 상속세 등 국세자료, 지역내총생산(GRDP) 자료, 복지통계 자료 등

② 국민여행실태조사

- 외래관광객실태조사 등 조사통계, 관광지방문객보고통계 등 행정자료

- 국내외 적용 사례

추정이 가능하다고 판단되면 비슷하거나 혹은 동일한 분야에 대한 국·내외 적용 사례를 검토해 본다. 이를 우리 사례에 맞춰 적용할 수 있을 것이라 기대하기 때문이다.

① 가계동향조사

- 한국 : “가계조사 지역별 추정기법 개발(통계청 용역 연구 보고서, 2004)”, “가계조사 지역별 추정기법의 타당성 검증(김혜원과 정동명, 2005)”
- 미국 : Small Area Income and Poverty Estimates (SAIPE)
- 영국 : Small Area Estimation Project (SAEP)

다음에는 이와 같은 소지역 추정 검토 방법을 이용하여 실제로 가계소득소비 분야에 대한 소지역 추정 방안을 검토해본다.

3. 소지역 추정 검토 방안

가. 소지역 추정 검토

가계소득소비 분야 추정을 위해서 사용될 주요 자료는 통계청에서 조사·작성하고 있는 가계동향조사 자료이다. 가계동향조사는 공표 범위가 도시와 전국이기 때문에 이보다 작은 지역에 대한 통계 수요가 지속적으로 있어왔다. 구체적인 수요 내용을 보면 다음과 같다

- 지역별 맞춤형 정책수립 (지역위원회)
- 지역관련 동향지표 부족 (산업연구원)
- 현행 공표 범위로는 지역의 결과 추론에만 이용 (지방행정연구원)
- 16개 시도별로 구분하여 공표 (한국은행, 국토연구원, 한국조세연구원, 강원 등)
- GRDP 지출 및 분배 계정 추계 기초 자료 활용 (경남, 울산 등)
- 시군구 단위 지역계정의 지출부문 추계에 필요 (대구 등)



가계동향조사의 표본은 전국 8,700여 가구로, 16개 광역시도에 평균적으로 544가구의 표본이 있다고 볼 수 있다. <표 3-2>는 가계동향조사의 가계수지 항목에 대한 2010년 연간 CV 자료다. 이를 통해 현행 표본 규모로 제공되고 있는 주요 항목에 대한 정도 (precision)를 확인할 수 있다. 이 조사는 전국이 공표 범위가기 때문에 CV가 대체로 작다. 그럼에도 소득부문에서 재산소득이 11.0%, 비경상소득이 8.0%로 크고, 지출부문에서는 교육이 2.9%로 상대적으로 큰 편이다. 전국 수준에서의 CV 정도에 비추어 봤을 때 동 자료를 이용한 시군구 통계나 소비지출 2수준 항목처럼 너무 작은 영역에 대한 추정은 상당히 큰 변이를 가질 것으로 예상할 수 있다.

<표 3-2> 가계동향조사의 가계수지 항목과 2010년 CV

가계수지 항목	CV	가계수지 항목	CV
소득(원)	1.3	가계지출(원)	1.3
경상소득(원)	1.3	소비지출(원)	1.2
근로소득(원)	1.9	01. 식료품·비주류음료(원)	0.9
사업소득(원)	2.9	02. 주류·담배(원)	2.1
재산소득(원)	11.0	03. 의류·신발(원)	1.7
이전소득(원)	3.1	04. 주거·수도·광열(원)	1.4
비경상소득(원)	8.0	05. 가정용품·가사서비스(원)	2.8
		06. 보건(원)	1.7
		07. 교통(원)	2.3
		08. 통신(원)	0.8
		09. 오락·문화(원)	2.1
		10. 교육(원)	2.9
		11. 음식·숙박(원)	1.3
		12. 기타상품·서비스(원)	1.6
		비소비지출(원)	2.1

출처 : 통계청 KOSIS.

주 : 전체가구(근로자가구 + 근로자외가구) 대상

가계동향조사 자료를 이용한 소지역 추정에 이용할 수 있는 보조정보를 검토한 결과는 <표 3-3>과 같다. 보조정보는 일반적으로 표본오차가 없는 자료를 이용하지만 다양한 방법의 적용을 위해서 이용 가능한 조사통계와 행정자료를 모두 검토해 보았다.¹⁾ 조사통계 자료는 통계청에서 조사하고 있는 가계금융조사, 농가경제조사, 여가경제조사, 서비스업통계조사, 도소매업통계조사 등의 활용성이 클 것으로 판단된다. 행정자료는 국세청의 법인납세조합관련 자료, 재산세세, 근로소득관련세금 등을 활용한다. 행정자료의 경우 2004년 가계동향조사 자료를 이용한 “가계조사 지역별 추정기법 개발(2004)”에서 그 유용성이 확인된 자료이다.

추가적으로 본 연구에서는 지역의 소비지출 부분을 설명할 수 있는 자료로 GRDP와 복지통계 자료를 활용해보고자 한다. 이는 가계동향조사 자료로 도출된 지역통계 결과가 GRDP의 비목 순위와 크게 달라지는 결과를 보완하고 가구의 이전소득에 대한 설명력을 높여보고자 하는 시도로 소지역 추정 결과에 대한 이론의 여지를 조금이라도 줄여보고자 하는 노력이다. 2004년의 연구에서도 인구1인당지역내총생산 자료를 이미 활용한 바 있지만, 이는 총량 지표로써 각 지역별로 서로 다른 특성의 지출 비목 설명력을 결과에 적절히 반영할 수 없는 문제점이 있었다고 판단된다. 그러므로 이를 구성하는 요소들을 지역과 부문별로 쪼개어 추정에 적용한다면 지출 비목별 추정 결과의 설명과 해석이 보다 용이할 수 있을 것이다.

가계동향조사 자료를 이용한 소지역 추정(안)을 대략적으로 검토해 본 결과는 <표 3-4>와 같다. 가구의 소득·소비 분야에 대하여 가계수지 1수준 항목까지 연평균 값을 광역시도 범위까지 추정해 볼 수 있다. 시군구 수준까지 지역통계를 요구하는 경우가 있었지만, 이는 앞서 살펴본 바와 같이 추정 결과에 대한 정도를 허용 가능한 수준까지도 기대할 수 없기 때문에 고려하지 않기로 한다.

1) 본 연구의 “제3절 가계소득소비 분야 소지역 추정”은 조사통계를 보조정보로 활용하기 어려운 모형을 사용했기 때문에 본 연구에서는 행정자료만을 보조정보로 이용했다. 검토된 조사통계는 향후 조사통계 결과를 추정에 활용하는 복합추정량이 개발된다면 그 유용성을 확인할 수 있을 것이다.

〈표 3-3〉 가계동향조사 소지역 추정을 위한 보조정보 검토

통계 구분		이용 자료
조사통계		가계금융조사, 농가경제조사, 어가경제조사, 서비스업통계조사, 도소매업통계조사, 사교육비실태조사
행정 자료	인구	총인구, 60세 이상 인구
	GRDP	인구1인당 지역내총생산 가계동향조사, 연말정산자료, 현금영수증 및 신용카드사용액 산업연관표, 상품군별 소매판매액, 담배판매액 자가소유주택서비스, 전기요금, 연료비, 가스 공급량, 수도 사용액 의약품, 의료급여통계, 지역별 진료실적, 운수업 산출액 자동차 등록대수, 서비스 수출입, 유선전화, 별정통신 매출액 복권, 경정 경륜, 경마, 카지노, FISIM, 보험료
	국세	조합원당징수세액, 법인수당법인세, 상속세납부인원 상속세납부자당과세표준, 인구1인당상속세 과세표준 인구1인당상속세 총결정세액, 양도소득세대상인원 양도소득금액, 양도세납부자당양도소득금액 양도세납부자당총결정세액, 인구1인당 양도소득금액 재산세세납부자당과세표준, 재산세세납부자당총결정세액 인구당지급액, 인구당원천징수세액
	복지	국민기초생활급여, 국민기초생활보장수급자가구 장애수당수급자수

출처 : 행정자료는 가계조사 지역별 추정기법 개발 (2004.12.)

추정 방법에서 또 하나 고민해 볼 수 있는 부분은 소득, 경상소득, 가계지출, 소비지출 등 하위 항목에도 추정 대상이 있는 경우에 각 하위 항목을 추정한 후 상위 그룹으로 취합하여 상위 항목을 추정할 지 여부이다. 예를 들면, 근로소득, 사업소득, 재산소득, 이 전소득을 추정하여 그 합으로 경상소득을 추정하는 것이다. 이 경우에는 상위 항목에 대한 CV 계산이 어렵다는 문제가 있다.



가계동향조사 자료를 이용한 소지역 추정은 다음 절에서 구체적으로 시도한다.

〈표 3-4〉 가계동향조사 소지역 추정 검토(안)

구분	공표	소지역 추정 검토	비고
분야	가구 소득, 소비	가구 소득, 소비	가계수지 1수준 항목
주기	분기, 연	연	매월조사. 분기 소지역 추정은 추후 검토
단위	도시, 전국	광역시, 도	시군구 추정에 대해서는 추후 검토
추정 방법	가구별 가중치를 이용한 추정	직접추정 or 모형기반추정	실제 자료 검토가 필요하지만 기존 연구 결과 소득 부문은 광역시, 도에 대해서 직접추정도 상당 정도 타당하다는 결과가 있음

나. 국내외 가계소득소비 추정 사례

가계소득소비 분야에 대한 국내외 추정 사례는 많지 않다. 국내에서는 2004년과 2005년에 진행된 바 있다(통계청, 2004; 김혜원과 정동명, 2005). 국외에서는 미국과 영국에서 소득 부문에 대한 소지역 추정 결과가 활용되고 있으나 소비 부문에 대한 적용이나 연구는 현재까지 뚜렷한 사례를 확인하지 못하였다.

1) 국내

국내에서 가계소득소비 분야에 대하여 통계청 주도로 이루어진 연구는 “가계조사 지역별 추정기법 개발(통계청 용역 연구 보고서, 2004)”이다. 이 연구에 대한 활용성 검토를 위하여 “가계조사 지역별 추정기법의 타당성 검증(김혜원과 정동명, 2005)” 이 후속 연구된 바 있다.

주요 연구인 2004년 연구의 개요는 다음과 같다.



- 사용통계 : 가계동향조사
- 추정 시점 : 2003년 연평균 자료 이용
- 소지역 정의 : 광역 시도
- 추정 내용 : 1수준 가계수지 항목 20개, 검토 CV : 10% or 20%
- 추정 방법 : 직접추정법, 이중추출추정법, 지역단위모형기반추정법(EBLUP1·2, HB), 조사구단위모형기반추정법 등
 - EBLUP1 : 회귀계수를 단순최소제곱법(OLS)으로 계산
 - EBLUP2 : 회귀 계수를 일반화최소제곱법(GLS)으로 계산
 - 지역단위모형 : 관심 변수에 따라 다른 보조정보 이용

연구는 다음과 같은 결과를 얻었다(<표 3-5>).

- 25개 지역(광역시도의 동부/읍면부) 단위보다는 16개 지역(광역 시도) 단위로 추정 발표하는 것이 바람직
- 총소득과 경상소득, 근로소득, 비근로소득(=사업소득 + 재산소득 + 이전소득), 가계 지출은 직접 추정량으로도 발표 가능
- 지출 10대 비목 중 식료품, 광열수도, 피복, 보건의료, 교양오락, 교통통신 등은 직접 추정량으로도 발표 가능하나 그 외 비목은 지역단위모형기반 추정량을 사용하는 것이 바람직
- 중앙값은 항목별로 발표하는 것은 의미가 없어 보이고 필요하다면 총소득과 총지출 같은 합계에서만 발표 가능.
- 문제점 : 직접추정법을 이용한 광역시·도 추정값은 가계소득소비지출의 비목 순위와 GRDP의 항목별 순위의 차이가 나는 경우가 발생

〈표 3-5〉 2004년 연구의 직접추정량과 지역단위모형기반추정량의 CV값 결과

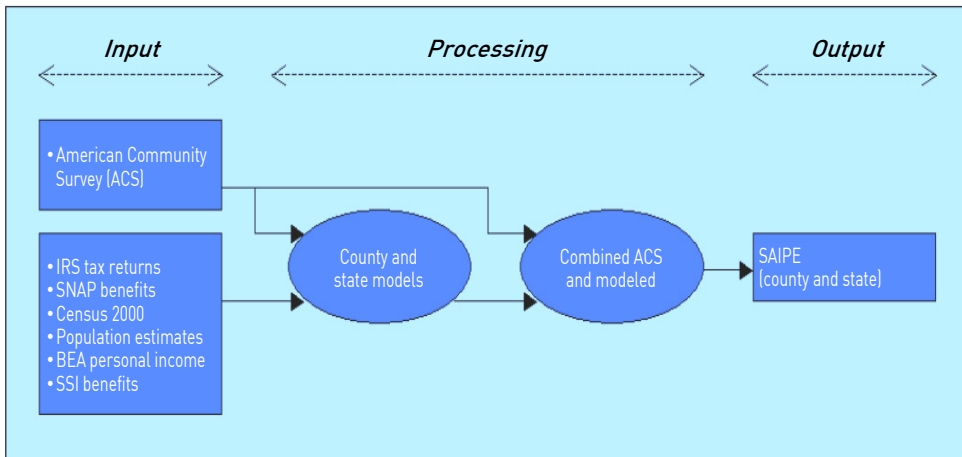
지역	CV	직접 통계량	지역 단위 모형 기반 추정량		
			EBLUP1	EBLUP2	HB
25개 시도	모두 10% 이내	식료품	사업소득, 식료품, 가구가사용품, 교통통신, 기타소비지출	사업소득, 식료품, 교통, 기타소비지출	식료품, 광열수도, 교통통신, 기타소비지출
	모두 20% 이내	근로소득, 광열수도, 가구가사용품, 보건의료, 교양오락, 교통통신, 기타소비지출	근로소득, 이전소득, 비경상소득, 광열수도, 피복, 교양오락, 비소비지출	근로소득, 비경상소득, 광열수도, 피복, 가구가사용품, 교양오락, 비소비지출	근로소득, 사업소득, 가구가사용품, 피복, 보건의료, 교육, 교양오락, 비소비지출
	모두 30% 이내	사업소득, 비경상소득, 피복, 교육, 비소비지출	주거, 보건의료, 교육	재산소득, 이전소득, 주거, 보건의료, 교육	이전소득, 비경상소득
	30% 이상	계산소득, 이전소득, 주거	재산소득		재산소득, 주거
16개 시도	모두 10% 이내	식료품, 광열수도, 보건의료, 기타소비지출	근로소득, 식료품, 광열수도, 가구가사용품, 보건의료, 교양오락, 교통통신, 기타소비지출, 비소비지출	근로소득, 식료품, 광열수도, 보건의료, 교양오락, 교통통신, 기타소비지출, 비소비지출	근로소득, 식료품, 광열수도, 보건의료, 교육, 교양오락, 교통통신, 기타소비지출, 비소비지출
	모두 20% 이내	근로소득, 사업소득, 주거, 가구가사용품, 교육, 교양오락, 교통통신, 비소비지출	사업소득, 이전소득, 비경상소득, 주거, 교육	사업소득, 이전소득, 비경상소득, 주거, 가구가사용품, 교육	사업소득, 이전소득, 주거, 가구가사용품
	모두 30% 이내	이전소득	재산소득	재산소득	비경상소득
	30% 이상	재산소득, 비경상소득			재산소득



2) 국외

미국은 센서스국에서 총인구에 대한 빈곤 추정값, 모든 군(county)과 주(state)에 대한 연간 가구 소득 중앙값을 생산하기 위하여 Small Area Income and Poverty Estimates (SAIPE) 프로젝트를 시행한다. 주요 목적은 연방 정부에 적시에 믿을만한 소득과 빈곤 통계를 제공하고 지역정부에 대한 연방 기금을 할당하며 지역 정부 활용 자료를 제공하기 위함이다(U.S. Census Bureau, 2010).

American Community Survey(ACS)로부터 측정되어진 빈곤을 모형화하고 학군(school district) 수준에서의 빈곤 추정은 Synthetic-share 방법을 이용한다. 이를 통해 서로 다른 지리적 수준의 추정값은 상위 수준으로 집계될 때 상위 수준의 값과 같아지도록 제어(benchmarking)된다. [그림 3-1]은 미국 SAIPE 방법을 간략하게 그림으로 나타낸 것이다.



출처 : 미국 센서스국, 2010

[그림 3-1] 미국 SAIPE 방법의 개요도

미국 SAIPE를 위한 기본 활용 자료는 다음과 같다.

-
- ACS의 소득과 빈곤 직접 추정값
 - 2000년 센서스 표본 자료의 소득과 빈곤 직접 추정값
 - 개인 식별 정보를 제거한 연방 개인 소득세 집계자료
 - 음식·영양 서비스(Food and Nutrition Service; FNS)로부터 보충영양보조프로그램 (SNAP) 수급자 수
 - 경제분석국(BEA)에서 집계된 개인 소득 추정값
 - 사회보장부(SSA)로부터 보조적 소득 보장(SSI) 수급자
 - 연령·그룹별 인구추계값
-

SAIPE은 FH 모형을 이용한다. 주(state) 모형은 선형 결합을 이용하고, 카운티 모형은 변수를 로그 변환하여 사용한다. 분산은 연속차분반복법(Fay and Train, 1995)을 이용한다. 주 추정에는 베이지안 접근법을 이용하고 카운티 모형은 반복가중최소제곱 알고리즘으로 수행한다. 관측값이 0인 경우에는 모형에서 제외한다.

영국은 2001년 센서스에 소득 관련 질문을 포함하지 않기로 한 정부의 결정에 따라 Small Area Estimation Project(SAEP)를 통해 Wards를 위한 모형 기반 추정값을 생산하게 되었다. 추정하고자 하는 값은 총 주간 가구 소득(불균등), 순 주간 가구 소득(불균등), 주거비 사용 전 순 주간 가구 소득(균등), 주거비 사용 후 순 주간 가구 소득(균등) 등이며 이들 소지역 추정 값과 95% 신뢰구간을 동시에 제공한다.

영국의 SAEP 기본 활용 자료는 다음과 같다.

-
- Family Resources Survey; FRS
 - 2001 센서스
 - 노동연금부 수급자수
 - HM 토지 등록 주거 가격 자료
 - 소득에 대한 국세/관세청(Her Majesty's Revenue and Customs service;HMRC) 자료
 - 2004 의회 세금 자료
 - 지역/국가 지표 - GOR 지표 + 웨일즈의 country 지표
-



<표 3-6>에 미국과 영국의 소득 부문에 대한 소지역 추정 방법을 정리해서 비교해 놓았다. 두 국가의 가장 큰 공통점은 모형 추정을 이용하였다는 것과 상위 수준에 대해서 일치하는 결과를 갖도록 하위 수준 추정 결과에 대해 벤치마킹을 시도하였다는 것이다. 또한 두 통계 모두 공식 통계의 지위를 확보하였다. 그러나 본 연구의 진행 과정 동안 가계소비나 지출과 관련된 해외 공식 통계 적용 사례는 아직 찾아볼 수 없었다.

<표 3-6> 미국 SAIPE와 영국 SAEP 비교

	미국 SAIPE	영국 SAEP
목적	<ul style="list-style-type: none"> 연방 정부에 적시에 믿을만한 소득과 빈곤 통계를 제공 지역 정부에 대한 연방 기금의 할당 지역 정부 활용 자료 제공 	<ul style="list-style-type: none"> 2001년 센서스에 소득 문항을 포함하지 않으면서 대안 필요
생산 통계	<ul style="list-style-type: none"> 총인구에 대한 빈곤 추정값 모든 군과 주에 대한 연간 가구 소득 중앙값 연간 통계 	<ul style="list-style-type: none"> 총 주간 가구 소득(불균등) 순 주간 가구 소득(불균등) 주거비 사용 전 순 주간 가구 소득(균등) 주거비 사용 후 순 주간 가구 소득(균등) <p>※ 95% 신뢰구간 동시 제공</p>
이용 통계	American Community Survey; ACS	Family Resources Survey; FRS
보조 정보	<ul style="list-style-type: none"> ACS의 소득과 빈곤 직접 추정값 2000년 센서스 표본 자료의 소득과 빈곤 직접 추정값 개인 식별 정보를 제거한 연방 개인 소득세 집계자료 음식·영양 서비스(Food and Nutrition Service; FNS)로부터 보충영양보조프로그램 (SNAP) 수급자 수 경제분석국(BEA)에서 집계된 개인 소득 추정값 사회보장부(SSA)로부터 보족적 소득 보장(SSI) 수급자 연령·그룹별 인구추계값 <p>※ 2009년 활용 자료임</p>	<ul style="list-style-type: none"> 2001 센서스 노동연금부 수급자수 HM 토지 등록 주거 가격 자료 소득에 대한 국세/관세청(Her Majesty's Revenue and Customs service; HMRC) 자료 2004 의회 세금 자료 지역/국가 지표 - GOR 지표 + 웨일즈의 country 지표
추정법	county : FH 모형 기반 추정 school district : Shynthetic-share 방법	모형 기반 추정
소지역	school district, county, state	Middle Layer Super Output Area; MSOA ※ 통계 보고용 지역

벤치 마킹	서로 다른 지리적 수준의 추정값은 더 높은 수준으로 함께 제어됨	좌동
비고	<ul style="list-style-type: none"> ○ 비슷한 단년도 조사에 비해 SAIPE가 정도가 높음 ○ 비슷한 소지역 추정 통계로는 SAHIE (Small Area Health Insurance Estimates)도 있음 	

출처 : 1. 미국 센서스국 Small Area Income and Poverty Estimates (SAIPE) : 2009 Highlights
 2. 영국 Model-Based Estimates of Income for Middle Layer Super Output Areas in England and Wales, 2004/05 User Guide

제3절 가계소득소비 분야 소지역 추정

이 절에서는 가계소득소비 분야 소지역 추정을 직접 시도한다. 주요 자료는 통계청의 가계동향조사이고, 보조정보는 인구, GRDP, 국세, 복지자료이다. 본 연구에서는 가계동향조사 소지역 추정을 함에 있어 다음과 같은 목적을 가지고 있다.

- 모형을 이용한 소지역 추정 결과의 정도(CV) 개선
- 다양한 보조정보 탐색으로 시도 단위 추정 결과의 정확성과 신뢰성 향상 기대
- 추정 결과의 활용
 - 더불어 가계동향조사와 GRDP 차이에 대한 인지
 - 가계동향조사 지출 부문과 GRDP 가계최종소비지출 항목간의 항목별 순위 차이 최소화를 통한 GRDP 지출 및 분배 계정 추계 기초 자료로 활용 기대

추정은 가구의 소득·소비 분야에 대하여 가계수지 1수준 항목까지 연평균 값을 광역시도 범위까지 한다. 소득, 경상소득, 가계지출, 소비지출 등 하위 항목에 추정 대상이 있는 경우에는 각 하위 항목만 추정한다. 상위 항목은 예를 들면, 근로소득, 사업소득, 재산소득, 이전소득을 추정하여 그 합으로 경상소득을 추정하는 것처럼 추정한다. 이 경우에는 상위 항목에 대한 CV는 계산되지 않는다.

1. 소지역 추정 모형

가. 모형

소지역 추정에 적용한 모형은 1979년에 Fay와 Harriot에 의해 제안된 Fay-Herriot 모형 (FH 모형)이다(Fay와 Herriot, 1979). 지난 2004년 연구에서 다양한 방법을 적용한 연구를 수행한 결과, FH 모형의 사용이 상당히 타당하다는 결론이 내려진 바 있으며(통계청 용역 연구 보고서, 2004), 미국의 SAIPE도 county 단위 추정시 이 모형을 사용하고 있다.

FH 모형은 지역 단위 직접추정량 $\hat{\mu}_h$ 을 지역 단위 보조변수 x_h 를 이용하여, 혼합모형을 통해 평활시켜 구현하는데 구체적인 모형의 형태는 다음과 같다.

$$\begin{aligned}\hat{\mu}_h &= \theta_h + e_h \\ \theta_h &= x_h' \beta + u_h\end{aligned}$$

여기서, 표본오차 $e_h \sim iid N(0, \hat{V}_h)$ 이고, 모형오차 $u_h \sim iid N(0, \tau^2)$ 임을 가정한다. 그러면 두 식을 결합한 최종 모형은 다음과 같은 혼합 모형의 형태가 된다.

$$\hat{\mu}_h = x_h' \beta + u_h + e_h$$

이에 대한 EBLUP 추정량은 다음과 같이 주어진다.

$$\hat{\theta}_h = (1 - \hat{B}_h) \hat{\mu}_h + \hat{B}_h x_h' \hat{\beta}$$

여기서 $\hat{B}_h = \frac{\hat{V}_h}{\tau^2 + \hat{V}_h}$ 이다.

따라서 우리가 기대하는 $\hat{\theta}_h$ 를 구하기 위해서는 β 와 τ^2 을 추정해야 하며, 이들을 추



정하는 방법은 다음의 두 방법이 주로 사용된다.

- 방법 1 (OLS). 다음의 두 식을 순차적으로 계산한다.

$$\textcircled{1} \hat{\beta} = \left[\sum_{h=1}^H x_h x_h' \right]^{-1} \left[\sum_{h=1}^H x_h \hat{\mu}_h \right]^{-1}$$

$$\textcircled{2} \hat{\tau}^2 = \max \left(0, \frac{\sum_{h=1}^H (\hat{\mu}_h - x_h' \hat{\beta})^2 - \sum_{h=1}^H \hat{V}_h (1 - r_h)}{H - p} \right)$$

여기서 $r_h = x_h' (X'X)^{-1} x_h$

- 방법 2 (GLS). 다음의 두 식을 반복적으로 계산한다.

$$\textcircled{1} \tilde{\beta}(\tau^2) = \left[\sum_{h=1}^H x_h x_h' / (\hat{V}_h + \tau^2) \right]^{-1} \left[\sum_{h=1}^H x_h \hat{\mu}_h / (\hat{V}_h + \tau^2) \right]^{-1}$$

$$\textcircled{2} h(\tau^2) = H - p$$

여기서 $h(\tau^2) = \sum_{h=1}^H (\hat{\mu}_h - x_h' \tilde{\beta})^2 / (\tau^2 + \hat{V}_h)$ 이다.

방법 1은 회귀계수를 단순최소제곱법(OLS)으로 계산하여 계산이 간단하고 방법 2는 회귀 계수를 일반화최소제곱법(GLS)으로 계산하는 것이다. 같은 모형을 이용한 베이지안 추정도 가능하다. 본 연구에서는 추정 결과의 활용에 더 초점을 맞추기 때문에 이에 대한 자세한 언급은 하지 않는다. 보다 자세한 내용을 확인하기 위해서는 통계청의 2004년 용역보고서를 참고할 수 있다.

나. 보조정보

가계동향조사 자료를 이용한 소지역 추정에 이용할 수 있는 보조정보를 구체적으로 검토한 결과는 <표 3-7>~<표 3-9>와 같다. 보조정보는 일반적으로 센서스 자료나 국세청 자료처럼 표본오차가 없는 자료를 이용한다. 앞서 <표 3-3>에서는 조사통계와 행정자료를 모두 검토했는데, 본 연구에서 사용한 FH모형에서는 조사통계가 가지는 오차를 계산하기가 어려워 행정자료만을 보조정보로 이용했다.

본 연구에서는 추정 결과에 대한 설명력을 높이기 위하여 보다 많은 보조정보를 찾는데 노력을 기울였다. 크게 4가지 분야에서 보조정보 탐색을 수행했는데, 인구통계자료, 지역내총생산(GRDP) 자료, 국세자료, 복지자료이다. 인구통계자료는 통계청에서 제공하는 성별 연령별 추계 인구 자료를 사용한다. GRDP 자료는 통계청의 지역소득통계팀에서 “가계최종소비지출” 작성 시 활용하는 자료 중 일부이다. 국세 자료는 국세청에서 연간 발행하는 국세통계연보에서 각 가구의 소득과 지출에 연관이 있을 만한 자료를 검토하였다. 마지막으로 복지 자료는 보건복지부에서 제공하는 급여종류별 국민기초생활급여 자료를 활용한다.

국세 자료의 경우 2004년 가계동향조사 자료를 이용한 “가계조사 지역별 추정기법 개발(2004)”에서 그 유용성이 확인된 자료이다. 추가적으로 본 연구에서는 지역의 소비지출 부분을 설명할 수 있는 자료로 GRDP 자료를 활용해보고자 한다. 이는 가계동향조사 자료로 도출된 지역통계 결과가 GRDP 지출 비목 순위와 크게 달라지는 결과를 보완하여 가계동향조사 소지역 추정 결과에 대한 이론의 여지를 조금이라도 줄여보고자 하는 노력이다.

물론, 두 조사의 포괄범위, 조사 대상 등이 달라서 생기는 근본적인 차이는 여전히 크게 존재한다. 2004년의 연구에서도 인구1인당지역내총생산 자료를 이미 활용한 바 있지만, 이는 총량 지표로써 각 지역별로 서로 다른 특성의 지출 비목 설명력을 결과에 적절히 반영할 수 없는 문제점이 있었다고 판단된다. 그러므로 이를 구성하는 요소들을 지역과 부문별로 쪼개어 추정에 적용한다면 지출 비목별 추정 결과의 설명과 해석이 보다 용이할 수 있을 것이다.



〈표 3-7〉 가계동향조사 소지역 추정을 위해 고려된 보조정보 : 인구, GRDP 분야

부문	변수명	보조정보
인구 (4)	x1_1	총인구
	x1_2	남자인구
	x1_3	여자인구
	x1_4	60세 이상 인구
GRDP (25)	x2_1	지역내총생산
	x2_2	인구1인당지역내총생산, $x2_1/x1_1$
	x2_3	담배판매액
	x2_4	자동차등록대수 중 민간소비율
	x2_5	복권판매현황
	x2_6	경정 매출
	x2_7	경정 환급금
	x2_8	경륜 매출
	x2_9	경륜 환급금
	x2_10	경마 매출
	x2_11	경마 환급금
	x2_12	GRDP 보험료
	x2_13	GRDP 의료비
	x2_14	GRDP 초중고 교육비
	x2_15	GRDP 대학(원) 교육비
	x2_16	시도별 진료비
	x2_17	시도별 급여비
	x2_18	판매전력량
	x2_19	도시가스공급량
	x2_20	수도부과금액
	x2_21	초중고교육비
	x2_22	공동주택관리비
	x2_23	초중고유학생수
	x2_24	지역별 가정용 석유류제품 소비 비중
	x2_25	보육시설 보육아동

출처 : 1. 인구 : 통계청 추계인구 자료, KOSIS
 2. GRDP : 지역소득통계팀 제공 자료

〈표 3-8〉 가계동향조사 소지역 추정을 위해 고려된 보조정보 : 국세 분야

부문	변수명	보조정보
국세 (31)	$x3_1$	법인수
	$x3_2$	인구1인당 법인수, $x3_1/x1_1$
	$x3_3$	피상속인수
	$x3_4$	상속세과세표준
	$x3_5$	상속세총결정세액
	$x3_6$	상속세납부자당과세표준, $x3_4/x3_3$
	$x3_7$	상속세납부자당총결정세액, $x3_5/x3_3$
	$x3_8$	인구1인당상속세과세표준, $x3_4/x1_1 \times 1,000,000$
	$x3_9$	인구1인당상속세총결정세액, $x3_5/x1_1 \times 1,000,000$
	$x3_10$	양도소득세결정건수
	$x3_11$	양도소득금액
	$x3_12$	양도세총결정세액
	$x3_13$	양도세납부자당양도소득금액, $x3_11/x3_10 \times 1,000,000$
	$x3_14$	양도세납부자당총결정세액, $x3_12/x3_10 \times 1,000,000$
	$x3_15$	인구1인당양도소득금액, $x3_11/x1_1 \times 1,000,000$
	$x3_16$	인구1인당양도세총결정세액, $x3_12/x1_1 \times 1,000,000$
	$x3_17$	증여세결정인원
	$x3_18$	증여세과세표준
	$x3_19$	증여세총결정세액
	$x3_20$	증여세납부자당과세표준, $x3_18/x3_17 \times 1,000,000$
	$x3_21$	증여세납부자당총결정세액, $x3_19/x3_17 \times 1,000,000$
	$x3_22$	인구1인당증여세과세표준, $x3_18/x1_1 \times 1,000,000$
	$x3_23$	인구1인당증여세총결정세액, $x3_19/x1_1 \times 1,000,000$
	$x3_24$	소득총지금액
	$x3_25$	원천징수부과세액중소득세등
	$x3_26$	인구당지금액, $x3_24/x1_1 \times 1,000,000$
	$x3_27$	인구당원천징수세액, $x3_25/x1_1 \times 1,000,000$
	$x3_28$	종합소득세 소득결정금액
	$x3_29$	종합소득세 결정세액
	$x3_30$	인구1인당종합소득세 소득금액, $x3_28/x1_1 \times 1,000,000$
	$x3_31$	인구1인당종합소득세 고지세액, $x3_29/x1_1 \times 1,000,000$

출처 : 국세청 국세통계연보



〈표 3-9〉 가계동향조사 소지역 추정을 위해 고려된 보조정보 : 복지 분야

부문	변수명	보조정보
복지 (14)	x4_1	국민기초생활급여합계
	x4_2	1인당국민기초생활급여합계, $x4_1/x1_1 \times 1,000,000$
	x4_3	국민기초생활급여생계급여
	x4_4	1인당국민기초생활급여생계급여, $x4_3/x1_1 \times 1,000,000$
	x4_5	국민기초생활급여주거급여
	x4_6	1인당국민기초생활급여주거급여, $x4_5/x1_1 \times 1,000,000$
	x4_7	국민기초생활급여교육급여
	x4_8	1인당국민기초생활급여교육급여, $x4_7/x1_1 \times 1,000,000$
	x4_9	국민기초생활급여재산장제급여
	x4_10	1인당국민기초생활급여재산장제급여, $x4_9/x1_1 \times 1,000,000$
	x4_11	국민기초생활급여저소득층에너지보조금
	x4_12	1인당국민기초생활급여저소득층에너지보조금, $x4_11/x1_1 \times 1,000,000$
	x4_13	국민기초생활보장수급자가구
	x4_14	장애수당수급자수

출처 : 보건복지부 급여종류별 국민기초생활급여, KOSIS

다. 모형 선정 결과

총 18개의 관심 변수에 대한 보조정보 탐색 결과는 <표 3-10>과 같다. 보조정보는 일차적으로 단계적 선택 방법으로 선택되었다. 그러나 재산소득 같은 관심 변수는 단계적 선택 방법으로 모형에 적합한 보조정보를 탐색할 수 없었다. 이런 경우에는 이차적으로 재산소득에 연관이 많을 것으로 판단되는 상속세, 양도세 등 정보를 연구자의 판단으로 모형에 보조정보로 활용하였다. 또한, 이런 방식으로도 적절한 보조정보를 찾지 못한 경우에는 불가피하게 같이 조사된 다른 관심 변수를 보조정보로 이용하였다. 이런 경우는 사업소득, 재산소득, 이전소득, 비경상소득, 주류·담배 지출, 기타 상품·서비스 지출 항목에서 발생하였다.

〈표 3-10〉 16개 지역 단계적 선택 방법에 의한 지역단위모형 변수 선택 결과

종속변수	설명변수	결정 계수		
		R^2	$Adj R^2$	
소득	근로소득 (y11_1)	인구1인당국민기초생활급여합계 (x4_2) 인구1인당지역내총생산 (x2_2) 인구1인당증여세과세표준 (x3_22)	0.79	0.74
	사업소득 (y11_2)	인구1인당국민기초생활급여생계급여 (x4_4) 가계동향조사교육지출 (y21_10)	0.60	0.53
	재산소득 (y11_3)	인구1인당상속세총결정세액 (x3_9) 가계동향조사사업소득 (y11_2) 양도세납부자당총결정세액 (x3_14) 양도세납부자당양도소득금액 (x3_13)	0.34	0.10
	이전소득 (y11_4)	인구1인당국민기초생활급여교육급여(x4_8) 가계동향조사재산소득 (y11_3)	0.36	0.26
	비경상소득 (y12)	인구1인당지역내총생산 (x2_2) 가계동향조사교통지출 (y21_7)	0.54	0.47
가계 지출	식료품·비주류음료 (y21_1)	양도세납부자당총결정세액 (x3_14) 양도세납부자당양도소득금액 (x3_13) 자동차등록대수증민간소비율 (x2_4)	0.67	0.59
	주류·담배 (y21_2)	인구1인당지역내총생산 (x2_2) 가계동향조사기타상품·서비스지출 (y21_12)	0.77	0.74
	의류·신발 (y21_3)	양도세납부자당총결정세액 (x3_14) 양도세납부자당양도소득금액 (x3_13)	0.53	0.46
	주거·수도·광열 (y21_4)	인구1인당국민기초생활급여주거급여 (x4_6)	0.62	0.60
	가정용품·가사서비스 (y21_5)	수도부과금액 (x2_20), 지역내총생산 (x2_1) 남자인구 (x1_2)	0.48	0.35
	보건 (y21_6)	양도세납부자당총결정세액 (x3_14) 상속세과세표준 (x3_4) 인구1인당종합소득세소득금액 (x3_30)	0.74	0.68
	교통 (y21_7)	시도별급여비 (x2_17), 여자인구 (x1_3) 인구1인당지역내총생산 (x2_2) 인구1인당종합소득세고지세액 (x3_31)	0.63	0.49
	통신 (y21_8)	양도세납부자당총결정세액 (x3_14) 상속세총결정세액(x3_5) 인구1인당종합소득세소득금액 (x3_30) 양도세총결정세액 (x3_12), 양도소득세결정건수 (x3_10)	0.85	0.78
	오락·문화 (y21_9)	인구1인당종합소득세소득금액 (x3_30) 인구1인당법인수 (x3_2), 인구1인당지역내총생산 (x2_2)	0.68	0.60
	교육 (y21_10)	양도세납부자당총결정세액 (x3_14) 인구1인당법인수 (x3_2), GRDP의료비 (x2_13)	0.74	0.67
	음식·숙박 (y21_11)	양도세납부자당총결정세액 (x3_14) 상속세총결정세액 (x3_5)	0.53	0.45
	기타 상품·서비스 (y21_12)	인구1인당지역내총생산 (x2_2) 가계동향조사교통지출 (y21_7)	0.68	0.64
	비소비지출 (y22)	인구1인당지급액 (x3_26), 피상속인수(x3_3) 소득총지급액 (x3_24)	0.77	0.71

결정계수는 종속변수 값을 로그 변환하여 구한 결과임

2. 추정 결과 분석

직접추정량과 지역단위모형기반추정량 EBLUP1, EBLUP2의 추정 결과를 추정값과 CV 측면에서 비교해보면 <표 3-11>, [그림 3-2] ~ [그림 3-19]와 같다. 18개 관심 변수 중 16개 시도 모두 CV가 10% 이내인 경우는 직접추정량이 7개(38.9%), EBLUP1이 12개(66.7%), EBLUP2가 11개(61.1%)로 직접추정량에 비해 모형추정에서 뚜렷한 CV 개선 효과가 있었음을 알 수 있다. 범위를 넓혀 CV를 20% 이내 수준으로 제한해 살펴보면, 모형추정량은 재산소득 추정값만 제외하면 모두 해당 수준 안으로 들어온다는 것을 확인할 수 있다. 재산소득의 경우에는 직접추정량도 CV가 크고 EBLUP2가 CV 개선 효과는 있으나 시도별로는 대전 지역이 30%에 가까운 등 안정성을 확보하는 데는 어려움이 있었다.

EBLUP1과 EBLUP2는 거의 같은 추정 성능을 보여준다. 따라서 보다 계산이 쉬운 EBLUP1을 이용하는 것이 나을 것이라 판단되며, 2004년의 연구에서도 같은 결론을 제시한 바 있다. 따라서, 이후 추정 결과 분석은 직접추정량과 EBLUP1을 비교하기로 한다.

추정 항목별로 간단히 요약하면, 추정값의 변화와 이에 따른 순위 변동은 소득 부문에서 컸고 지출 부문에서는 주류·담배, 주거·수도·광열, 비소비지출 부문이 다른 항목에 비해 컸다. CV 개선 효과는 소득 부문은 ① 사업소득, ② 비경상소득, ③ 근로소득, ④ 재산소득, ⑤ 이전소득 순으로 컸고, 지출 부문은 ① 주거·수도·광열, ② 주류·담배, ③ 기타 상품·서비스, ④ 오락·문화, ⑤ 보건, ⑥ 음식·숙박, ⑦ 비소비지출, ⑧ 교통, ⑨ 통신, ⑩ 교육, ⑪ 의류·신발, ⑫ 식료품·비주류음료, ⑬ 가정용품·가사서비스 순으로 컸다. ① 주거·수도·광열과 ② 주류·담배는 50% 이상의 개선 효과가 있었고, ⑥ 음식·숙박 이하 항목은 CV 개선 효과가 10% 이하였다.

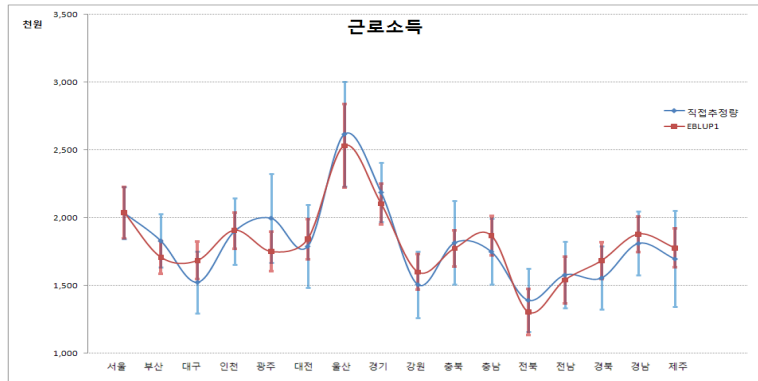
〈표 3-11〉 직접추정량과 지역단위 모형기반 추정량의 CV 값 결과

	직접추정량	EBLUP1	EBLUP2
16개 시도 모두 10% 이내		근로소득	근로소득
		사업소득	
			이전소득
	식료품·비주류음료	식료품·비주류음료	식료품·비주류음료
		주류·담배	
	의류·신발	의류·신발	의류·신발
	주거·수도·광열	주거·수도·광열	주거·수도·광열
	보건	보건	보건
	통신	통신	통신
		오락·문화	오락
	음식·숙박	음식·숙박	음식·숙박
	기타상품·서비스	기타상품·서비스	기타상품·서비스
		비소비지출	비소비지출
모두 20% 이내	근로소득		
	사업소득		
	이전소득	이전소득	
		비경상소득	비경상소득
	주류·담배		
	가정용품·가사서비스	가정용품·가사서비스	가정용품·가사서비스
	교통	교통	교통
	오락		
	교육	교육	교육
	비소비지출		
모두 30% 이내	비경상소득		
			재산소득
30% 이상	재산소득	재산소득	

EBLUP2 추정 예러 : 사업소득, 주류·담배

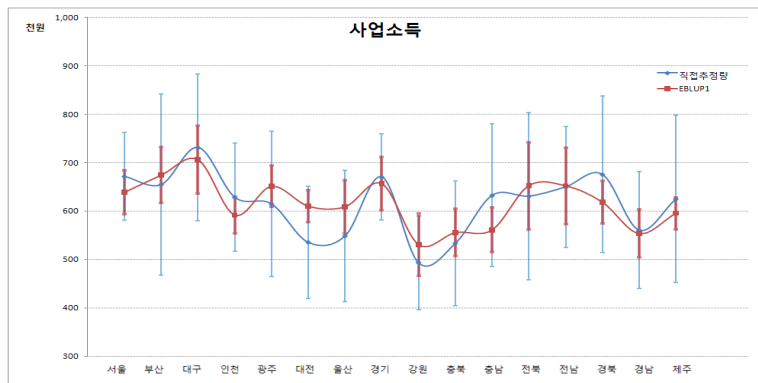


가. 소득 부문



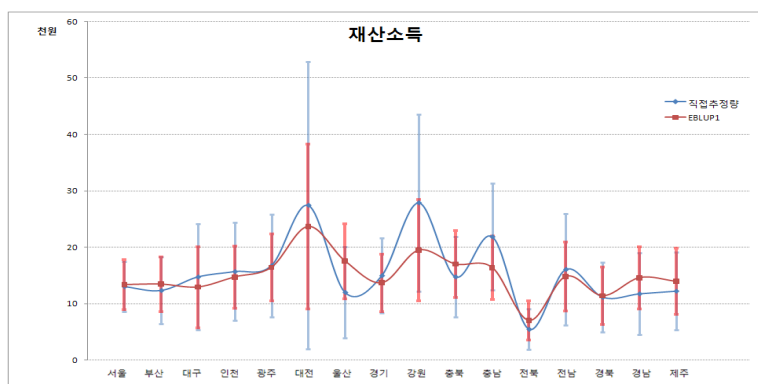
[그림 3-2] 근로소득 추정 결과

근로소득은 직접추정을 이용한 것에 비해 모형추정으로 대구, 광주, 강원, 제주 등 지역에서 CV가 개선되었다. 직접추정량의 CV가 4.81~10.66%인데 비해 모형추정량은 3.57~6.71%까지 감소하였다. 또한, 지출 항목에 비해 소득 항목은 추정에 따른 추정값의 변화도 크게 나타났는데, 이에 따른 지역간 소득 순위 변동이 보다 크게 나타났다. 부산과 광주는 모형추정 전에는 각각 6위와 4위였으나, 추정 후에는 11위와 4위로 떨어지고, 대신 제주, 충남 등의 상대적 순위가 올라갔다.



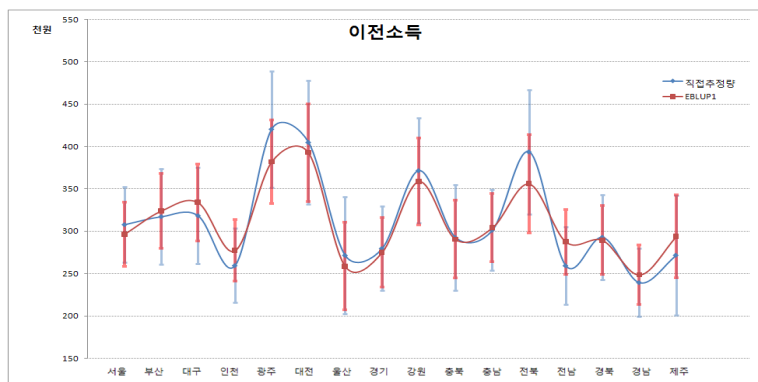
[그림 3-3] 사업소득 추정 결과

사업소득은 근로소득에 비해서도 모형추정에 따른 CV 개선 효과가 크다. 직접추정량의 CV가 6.77~14.56%까지 분포되어 있는 반면, 모형추정량 EBLUP1은 2.75~7.03%까지 줄어든다. 순위 변동이 큰 지역은 광주, 대전이 보조정보를 이용하여 추정한 결과 직접추정값일 때보다 순위가 상승하였고, 충남, 경북 등은 순위가 하강하였다.



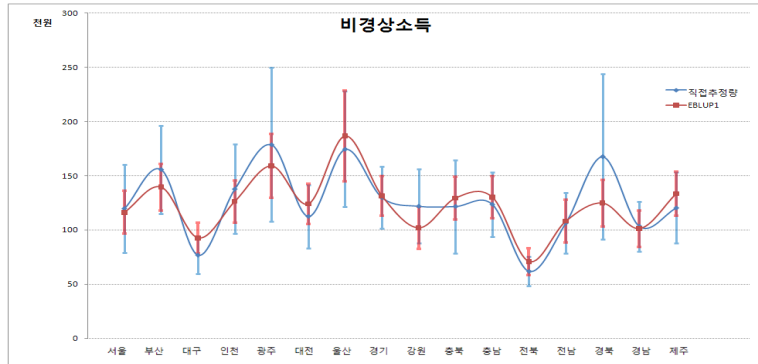
[그림 3-4] 재산소득 추정 결과

재산소득의 CV 분포를 보면 시도별 직접추정량이 17.22~47.33%에서 모형추정량이 16.95~31.43%까지 개선되었다. 그러나 재산소득은 앞서서도 보았듯이 전국 CV도 11.4%로 상당히 크다. 순위 변동이 큰 지역은 울산, 경남이 보조정보를 이용하여 추정한 결과 직접추정값일 때보다 순위가 큰 폭으로 상승하였고, 대구 등은 순위가 하강하였다.



[그림 3-5] 이전소득 추정 결과

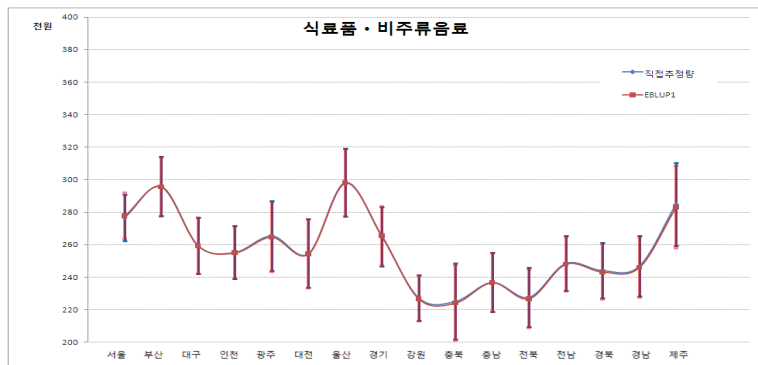
이전소득의 CV 분포를 보면 직접추정량이 7.44~13.28%에서 모형추정량이 6.50~10.24%로 감소하였다. 광주, 대전, 강원 전북 등이 타 지역에 비해 이전소득이 높았지만, 추정에 따른 순위 변동은 크지 않았다.



[그림 3-6] 비경상소득 추정 결과

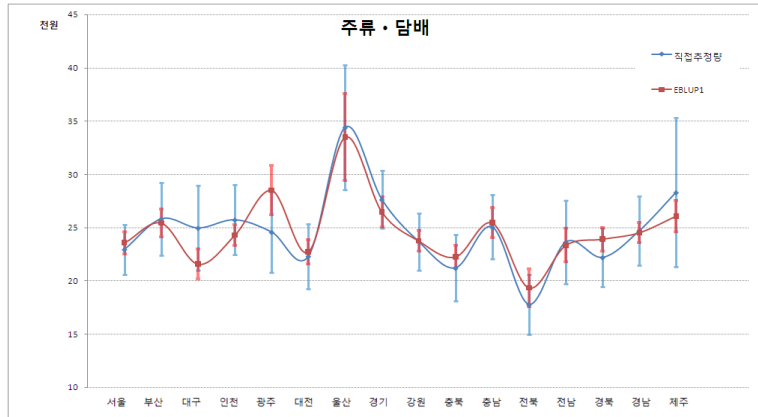
비경상소득의 CV 분포를 보면 직접추정량이 11.20~23.27%로 분포하는데 추정 후에는 7.10~11.43%로 CV 개선 효과가 있었다. 특히 광주와 경북은 CV가 50% 이상 개선되는 효과를 보여준다. 이는 재산소득에 비해서도 큰 개선 효과를 보인 것인데, 비경상소득의 전국 CV는 8.8%이었다. 순위 변동이 큰 지역은 인천, 강원, 경북이 모형추정 결과 직접추정값일 때보다 순위가 하강하였고, 제주 지역은 순위가 상승하였다.

나. 소비 부문



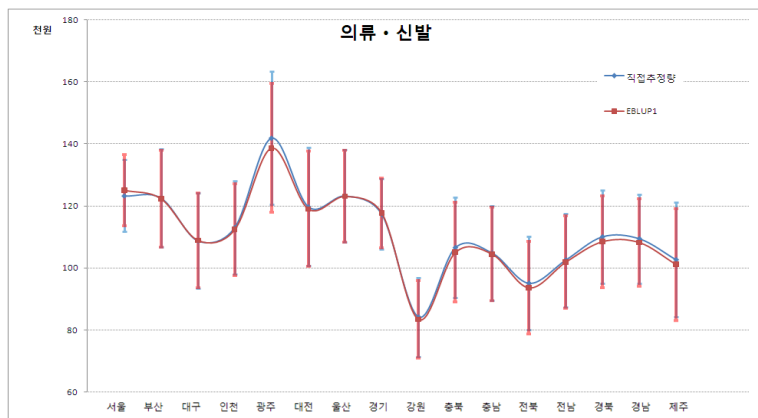
[그림 3-7] 식료품비주류 음료 지출 추정 결과

식료품·비주류 음료 지출 항목은 직접추정과 모형추정에 큰 차이가 없었다. 직접추정량의 CV가 2.62~5.27%까지 분포하는데, 모형추정도 2.60~5.27%로 거의 근사하고 추정값도 거의 비슷하기 때문에 순위 변동도 크지 않았다.



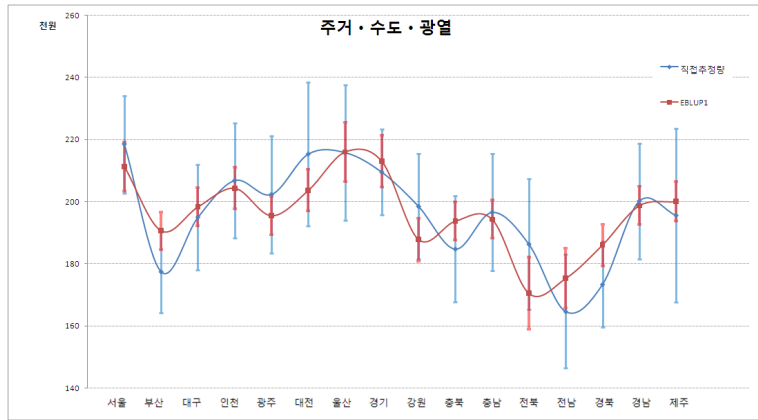
[그림 3-8] 주류담배 지출 추정 결과

주류·담배 지출 항목은 직접추정보다 모형추정으로 CV 개선 효과가 컸다. 직접추정량의 CV가 5.02~12.63%까지 분포하는데, 모형추정은 1.94~6.26%로 줄어들었다. 또한 지역간 지출 순위 변화는 대구가 추정 후 순위가 크게 떨어진 반면, 광주와 경북은 상승하였다.



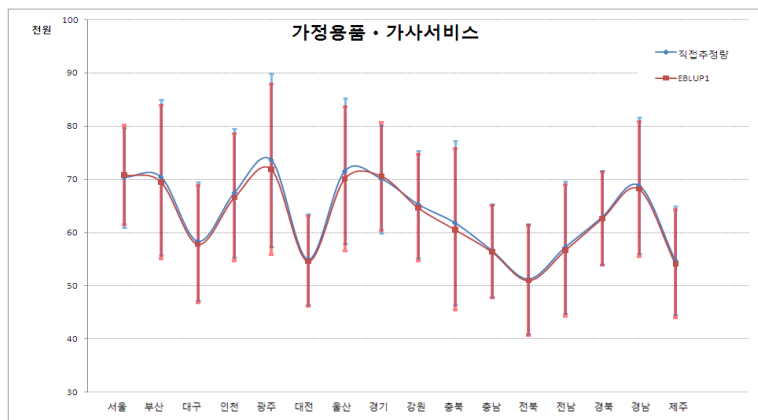
[그림 3-9] 의류신발 지출 추정 결과

의류·신발 지출 항목은 직접추정과 모형추정에 큰 차이가 없었다. 직접추정량의 CV가 4.80~9.14%까지 분포하는데, 모형추정도 4.70~9.08%로 거의 근사하고 추정값도 비슷하기 때문에 순위 변동도 거의 없다.



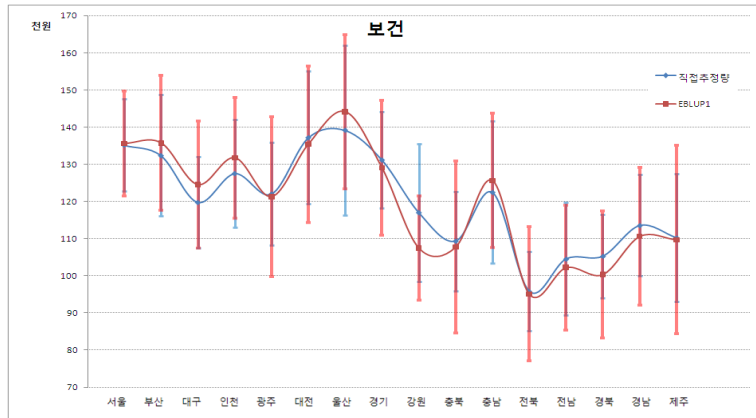
[그림 3-10] 주거·수도·광열 지출 추정 결과

주거·수도·광열 지출 항목은 직접추정보다 모형추정으로 CV 개선 효과가 컸다. 직접추정량의 CV가 3.35~5.75%까지 분포하는데, 모형추정은 1.57~3.46%로 줄어들었다. 또한 지역간 지출 순위 변화는 강원, 전북이 추정 후 순위가 떨어진 반면, 대구, 광주, 제주는 상승하였다.



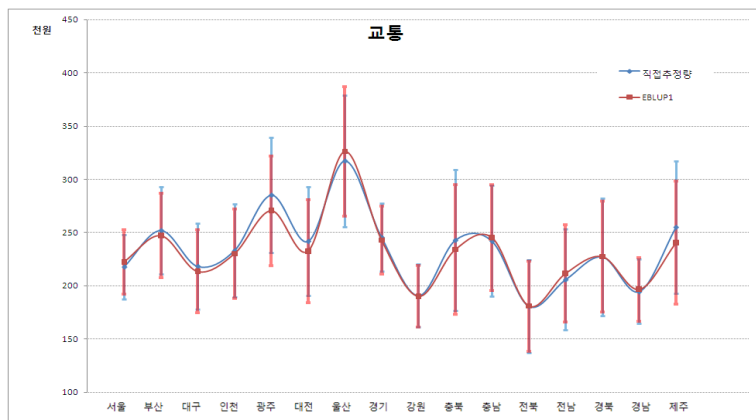
[그림 3-11] 가정용품·가사서비스 지출 추정 결과

가정용품·가사서비스 지출 항목은 직접추정과 모형추정에 큰 차이가 없었다. 직접 추정량의 CV가 6.78~12.10%까지 분포하는데, 모형추정도 6.72~12.78%로 거의 근사하고 추정값도 비슷하기 때문에 순위 변동도 거의 없다.

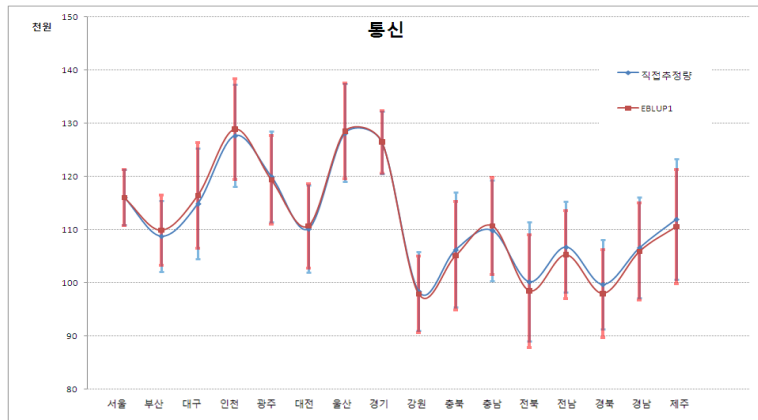


[그림 3-12] 보건 지출 추정 결과

보건 지출 항목은 CV 개선 효과가 어느 정도 있으면서 순위 변동은 크지 않은 경우에 속한다. 직접추정량의 CV가 4.68~8.40%까지 분포하는데, 모형추정도 4.53~7.06%로 감소하면서 추정값은 비슷하기 때문에 순위 변동은 거의 없다.

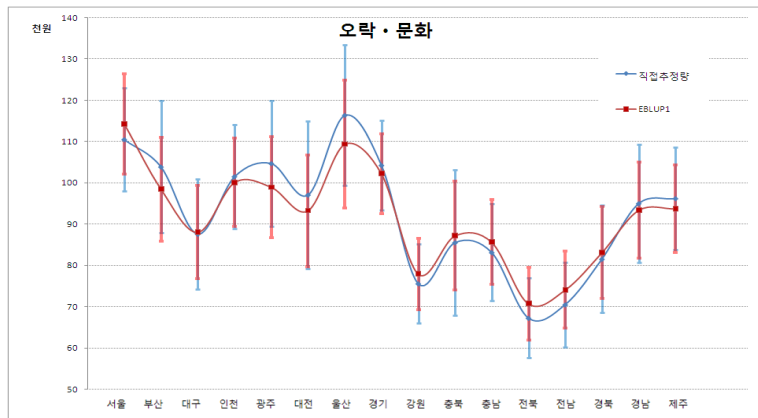


[그림 3-13] 교통 지출 추정 결과



[그림 3-14] 통신 지출 추정 결과

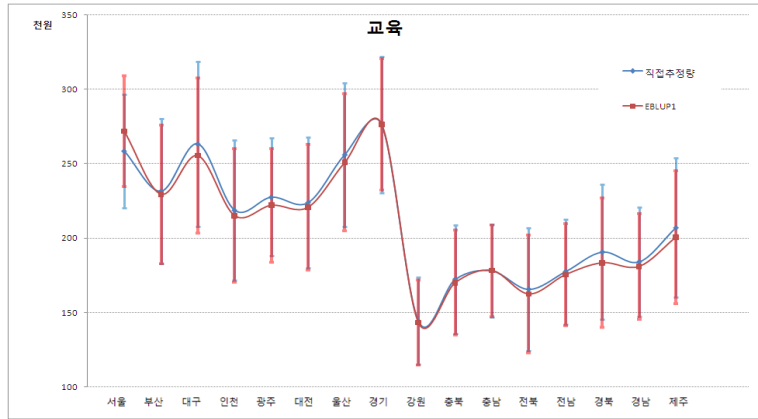
교통과 통신 지출 항목은 직접추정과 모형추정에 큰 차이가 없었다. 교통 부문은 직접추정량의 CV가 6.65~13.90%까지 분포하는데, 모형추정도 6.70~13.27%, 통신 부문 역시 직접추정량의 CV가 2.31~5.72%인데, 모형추정도 2.31~5.49%로 거의 근사하고 추정값도 비슷하기 때문에 순위 변동도 거의 없다.



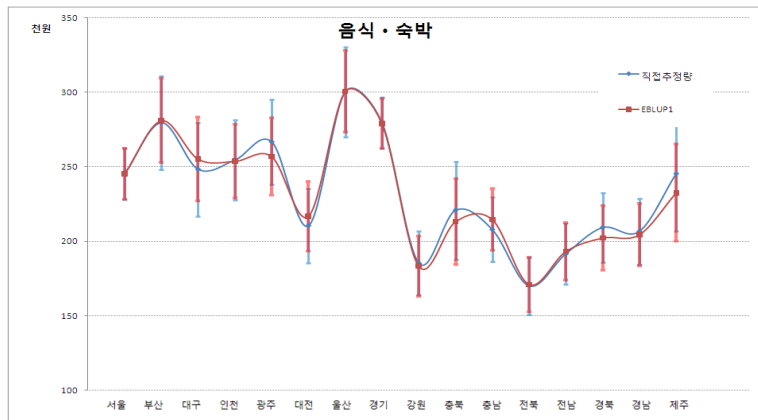
[그림 3-15] 오락문화 지출 추정 결과

오락·문화 지출 항목은 CV 개선 효과가 어느 정도 있으면서 순위 변동은 크지 않은

경우에 속한다. 직접추정량의 CV가 5.33~10.47%까지 분포하는데, 모형추정도 4.81~7.74%로 감소하면서 추정값은 비슷하기 때문에 순위 변동은 거의 없다.



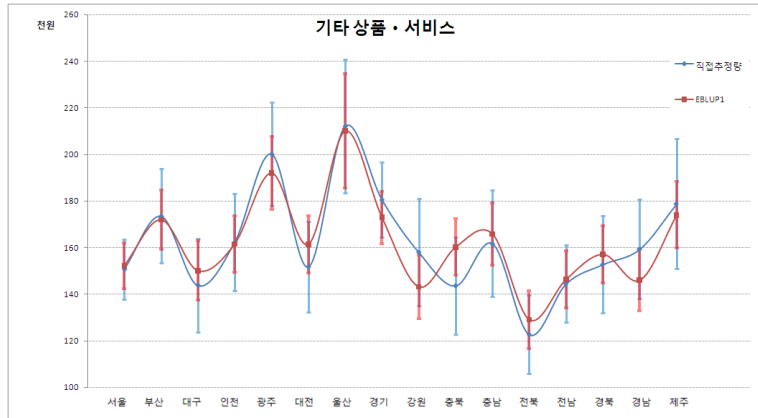
[그림 3-16] 교육 지출 추정 결과



[그림 3-17] 음식·숙박 지출 추정 결과

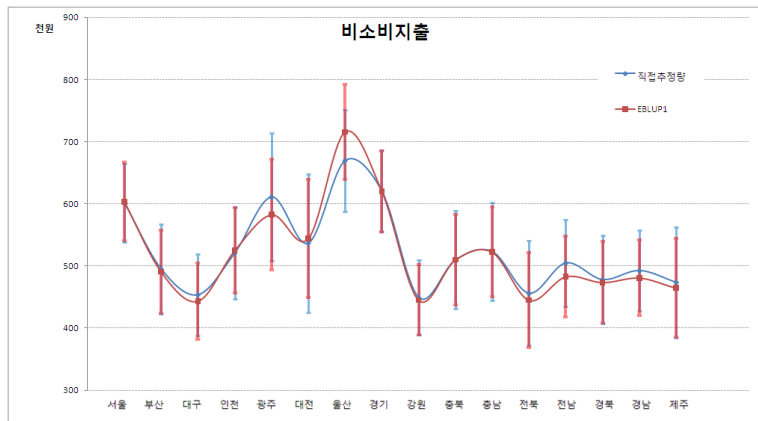
교육과 음식·숙박 지출 항목은 직접추정과 모형추정에 큰 차이가 없었다. 교육 부문은 직접추정량의 CV가 7.52~12.71%까지 분포하는데, 모형추정도 6.99~12.48%, 음식·숙박 부문 역시 직접추정량의 CV가 3.14~7.99%인데, 모형추정도 3.04~7.14%로 거의 근

사하고 추정값도 비슷하기 때문에 순위 변동도 거의 없다.



[그림 3-18] 기타 상품서비스 지출 추정 결과

기타 상품·서비스 지출 항목은 직접추정보다 모형추정으로 CV 개선 효과가 컸다. 직접추정량의 CV가 4.35~7.97%까지 분포하는데, 모형추정은 3.29~5.97%로 줄어들었다. 또한 지역간 지출 순위 변화는 강원, 경남이 추정 후 순위가 떨어진 반면, 충북은 상승하였다.



[그림 3-19] 비소비지출 추정 결과



비소비지출 항목은 직접추정과 모형추정에 큰 차이가 없었다. 직접추정량의 CV가 5.35~10.58%까지 분포하는데, 모형추정은 5.33~8.89%로 거의 근사하고 추정값도 비슷하기 때문에 순위 변동도 거의 없다.

지금까지 소득과 소비 부문에서 추정 결과를 확인했는데, 이를 간략히 요약하면 다음 <표 3-12>와 같다. 소득 부문은 CV 개선효과와 함께 대체로 순위가 변하는 경향이 있고, 지출 부문은 식료품·비주류음료, 의류·신발통신, 음식·숙박 지출 부문은 직접추정값의 CV도 10% 이하이고 소지역 추정 결과도 CV 개선효과가 크지 않으면서 순위 변동도 없는 지역으로 이런 지역은 소지역 추정을 굳이 수행하지 않고 직접추정값을 그냥 이용하는 방안도 좋은 대안이 될 수 있을 것이라 본다.

<표 3-12> 소지역 추정 결과 요약

		CV 개선 효과		
		10% 이하	10~30%	30% 이상
순위 변동	있음		근로소득 사업소득 재산소득 비경상소득	주류·담배 주거·수도·광열 기타 상품·서비스
	없음	식료품·비주류음료 의류·신발 가정용품·가사서비스 교통 통신 교육 음식·숙박 비소비지출	이전소득 보건 오락·문화	

1. 순위 변동 '있음'은 한 지역이라도 추정으로 인해 5단계 이상 순위 변화가 있을 때
2. 굵은 글씨는 직접추정값의 CV가 10% 이하인 경우
3. 이탤릭 글씨는 EBLUPI의 CV가 10% 이하인 경우

3. 제한점과 시사점

가계동향조사에 대한 모형기반 소지역 추정 결과 CV는 상당 정도 개선효과가 있다는 것을 확인할 수 있었다. 그럼에도 이 자료의 사용에 대해서 우려를 가지는 이유는 동

일 항목에 대한 GRDP 결과와의 차이 때문이다. 두 값을 단순 비교하기에는 두 값을 얻기 위한 전제와 과정의 차이가 크다. 이 절에서는 이런 점에 대한 간단한 고찰과 이를 통해 얻을 수 있는 몇 가지 시사점을 확인한다.

가계동향조사 추정과 GRDP 추계 사이에는 크게 다음과 같은 세 가지 차이점이 있다.

① 포괄범위

이론적으로 가계동향조사는 농림어가를 제외한 도시가구만을 대상으로 하고 GRDP의 가계최종소비지출은 전체 가구를 대표한다. 즉, GRDP는 우리나라 전체 천 7백여만 가구를 대상으로 하고 있으나 가계동향조사는 천 5백여만 가구를 대상으로 한다.

<표 3-13> 가계동향조사의 포괄범위

구분	표본 규모	대표성	표본 변동
가계동향조사	약 8,700 가구 (응답 : 7,100 가구)	1천 6백만 가구 (비농가)	연동 표본 (매년 표본의 1/3 교체)
농가경제조사	약 2,800 가구 (응답 : 2,800 가구)	1백 26만 가구 (농가)	동일 표본 (2008년 이후 동일)

가계동향조사 이용자 가이드 (2011. 8.)

② 추계 방식

가계동향조사는 FISIM, 자가소유부문을 반영하지 않고 가구 가중치를 이용하여 표본 추정하지만, GRDP는 이 부분을 반영하며 상품흐름법을 이용하여 가구가 아닌 1인당 추계 값 × 인구수 방법으로 추계한다(통계청, 가계최종소비지출 추계 개요).

③ 이용 자료

본 연구에서 사용된 보조정보와 GRDP 추계에 이용하는 정보가 다르다. 가계동향조사와 GRDP 간 차이를 극복하기 위하여 보조정보를 가능한 많이 탐색하였으나 한계가 있다.

이런 차이가 두 통계의 비교를 어렵게 하기 때문에 본 연구에서는 GRDP 결과를 최대한 가계동향조사의 포괄 범위와 맞추어 보고자 다음과 같은 GRDP 비교 자료를 새로 생성해 보았다.²⁾ 이를 통해 구체적으로 지출 항목별 추정 결과 순위와 GRDP 순위 비교

2) GRDP1과 GRDP2는 가계동향조사와 GRDP 사이의 포괄 범위 차이를 줄여보고자 연구자가 관련 자료를



를 통해 추정 결과에 대한 평가를 시도한다. 비교하고자 새로 생성한 GRDP 자료는 다음과 같은 두 가지 결과를 이용하였다.

- GRDP1

$$\frac{\text{지출비목에 따른 시도별 지출}}{\text{추계가구수} \times 12}$$

- GRDP2

$$\frac{\text{지출비목에 따른 시도별 지출} \times \text{비농림어업 비중}}{\text{추계가구수} \times \text{비농가 비중} \times 12}$$

여기서, 각 항목은 다음과 같이 계산하였다.

- 지출비목에 따른 시도별 지출 : 지역내 총 생산에 대한 지출 중 가계최종소비지출
- 비농림어업 비중 : 경제활동별 지역내총생산의 총부가가치에서 비농림어업이 차지하는 비중
 - 그러나, 생산 계정은 지역에 기반한 가계, 사업체 등의 모든 생산 활동이 포함되기 때문에 가계최종소비지출에 그대로 비중을 반영하는 것은 무리가 있고, 비농림어업 비중을 지출 비목별로 구분하여 적용하지 않았다는 점을 유의해야 한다.
 - 도 지역으로 갈수록 농림어업 비중은 커진다.
- 비농가 비중 : 2010년 인구주택총조사의 비농가 비중
 - 도 지역으로 갈수록 농가 비중이 커진다. 다만 임가와 어가 비중은 그렇게 크지 않기 때문에 비교에 직접 이용하지 않기로 한다.

※ 참고 : 새로 생성한 GRDP와 가계동향조사 지출 총액 비교

$$\frac{\text{GRDP1 가구 지출 총액}}{\text{가계동향 지출 총액}} = \frac{559\text{조}}{452\text{조}} = 1.24$$

$$\frac{\text{GRDP2 가구 지출 총액}}{\text{가계동향 지출 총액}} = \frac{535\text{조}}{452\text{조}} = 1.18$$

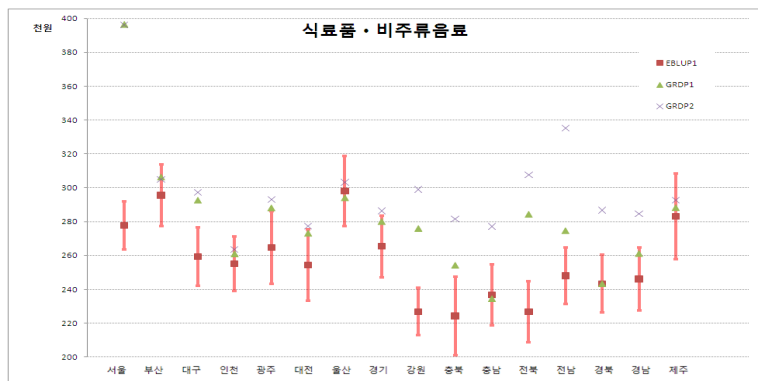
가계동향조사와 GRDP는 포괄 범위, 추계 방식, 이용 자료가 다르기 때문에 단순 비교는 불가능하다. 다만, GRDP의 지역간 상대적인 순위를 계산하기 위해서만 총량값과 비율을 이용하고(<표 3-14>), 구체적인 비교는 지역간 상대적인 순위를 이용한다.

<표 3-14> 2009년 지역내총생산 및 지역내총지출의 관계와 각 항목 비중

(단위 : %)

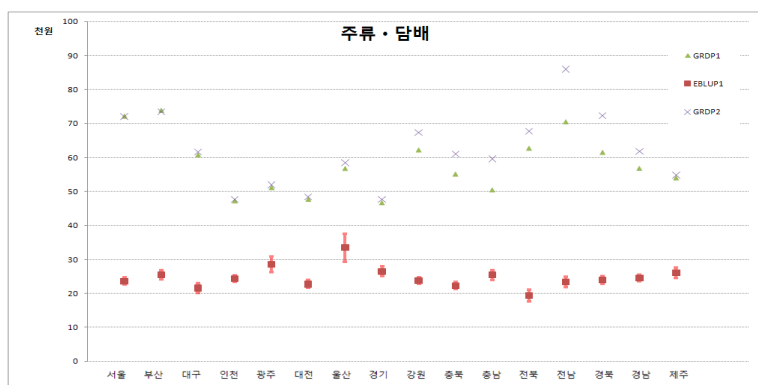
생산 GRDP			지출 GRDP								
순 생산 물세	총 부가 가치		최종소비지출			총자본형성				재화와 서비스 의 순이출	통계적 불일치
			민간최종소비 지출		정부 최종 소비 지출	총고정자본형성			재고 증감		
	비 농림 어업	농림 어업	가계 소비	NPI SHS 지출		건설 투자	설비 투자	무형 투자			
10.03	87.53	2.44	52.46	1.53	15.93	18.17	9.09	1.80	-2.88	3.70	0.21

간단히 결과를 요약하면, 현재의 추정 결과와 GRDP결과는 비교가 될 수 없다. 추정 과정의 상이함 때문에도 직접 비교는 어렵지만 실제 자료를 이용하여 비교한 결과에서도 좋은 결론을 얻지 못했다. 이를 위해 12개 지출 항목 중 식료품·비주류음료, 주류·담배, 교통 3개 항목의 결과를 그림으로 확인한다.



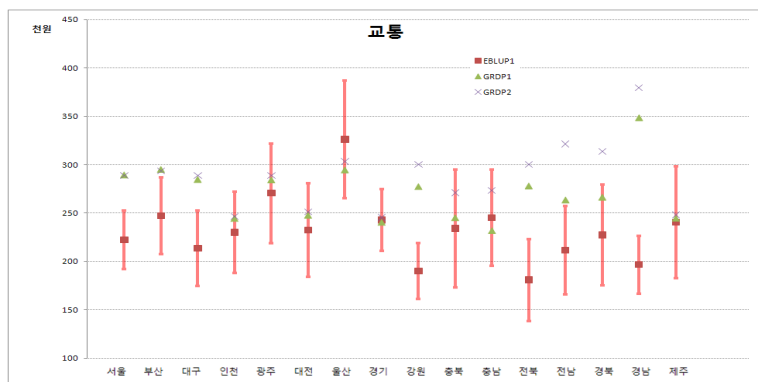
[그림 3-20] 식료품비주류음료의 가계동향조사와 GRDP 분포

식료품·비주류음료 지출 부문에서 가계동향조사 추정 결과와 GRDP의 총량과 순위는 틀리다. 서울 지역을 보면, 모형추정 결과 16개 시도에서 4위의 지출 순위를 보였으나, GRDP1과 GRDP2에서는 모두 1위였다. 전북은 모형추정 결과 15위였던 지출 순위가 GRDP1에서는 7위 GRDP2에서는 3위까지 올라갔다. 이런 현상은 전남도 마찬가지로 나타났다.



[그림 3-21] 주류·담배의 가계동향조사와 GRDP 분포

주류·담배 지출 부문에서 가계동향조사 추정 결과와 GRDP의 총량과 순위도 역시 틀리다. 역시 서울 지역을 보면, 모형추정 결과 16개 시도에서 11위의 지출 순위를 보였으나, GRDP1과 GRDP2에서는 각각 2위와 4위였다. 모형추정에서 1위였던 울산은 GRDP1과 GRDP2에서 9위와 11위까지 떨어졌다.



[그림 3-22] 교통의 가계동향조사와 GRDP 분포

교통은 앞의 두 항목과 다르게 총량 규모에서 모형추정과 GRDP 자료간 차이가 크지 않다. 모형 추정에서 상위권인 울산(2위), 부산(3위)이 GRDP1에서는 상위권을 차지하고 있다. 그러나 농림어업 비중을 제외하고 재계산된 GRDP2에서는 이전 항목들과 마찬가지로 순위가 떨어진다. 이는 농림어업 비중을 제외하고 GRDP를 계산하면 도 지역의 도시 가계 지출 총량이 커지기 때문이다. 따라서, 경남은 모형 추정 결과가 14위인데 GRDP1, GRDP2에서는 모두 1위가 된다.

본 연구에서는 최대한 많은 보조정보를 탐색하였다. 그러나 기대와는 달리 보조정보가 모두 모형에서 효율적으로 이용된 것은 아니다. 예를 들면, 사업소득, 이전소득, 비경상소득 등의 경우, 표본오차가 없는 보조정보만으로는 모형 설명력을 높일 수 없었기 때문에 불가피하게 같이 조사된 가계동향조사 자료를 이용할 수밖에 없었다. 심지어 재산소득은 상관 정도가 높은 가계동향조사의 사업소득 결과를 보조정보로 이용했음에도 모형 적합도를 크게 높일 수 없었다. 이는 좋은 보조정보를 찾아내지 못했을 때 모형추정이 가지는 큰 단점이 될 수 있다. 그러므로 향후에는 다양한 방법의 적용을 위해서 이용 가능한 조사통계와 행정자료를 모두 검토할 필요가 있다. 대표적으로 조사통계 자료는 통계청에서 조사하고 있는 가계금융조사, 농가경제조사, 어가경제조사, 서비스업통계조사, 도소매업통계조사 등의 활용성이 클 것으로 판단된다. 이 경우에는 물론 표본오차를 제어할 수 있는 새로운 방법이 제시되어야 할 것이다.

가계동향조사 자료를 이용한 소지역 추정 결과를 효율성 면에서 본다면 다음과 같은 결론을 내릴 수 있다.

-
- 모형추정은 직접추정의 정도를 분명하게 개선한다
 - 직접추정도 공표 가능한 수준에 있는 항목이 있다.
 - GRDP와의 직접 비교는 어렵다. 즉, 추가적인 보완 연구가 수행되지 않고 현재의 방법을 이용한다면 직접 비교는 하지 않아야 한다.
-



제4절 결 론

본 연구의 목표는 지역통계 수요에 대한 통계청의 대응 방안을 고민해 보는 것이다. 이를 위해 소지역 추정 검토 방안을 제시하였고 통계청이 확보한 17개 지역 통계 수요 중 가장 높은 우선순위를 갖는 가계동향조사에 대해 시험적으로 시도 소득소비 분야에 대한 추정 통계를 작성하였다.

가계동향조사 추정 결과를 요약하면 다음과 같다.

- 가계동향조사의 소지역 통계(직접추정 혹은 모형추정) 생산이 방법적으로 많이 어렵지는 않다. 그러므로 도시가구를 대상으로 하는 광역시도 통계 생산은 현재도 가능하다. 그러나 이 경우에는 조사통계 결과를 추정에 활용하는 것과 같이 보다 다양한 자료를 추정에 이용할 수 있는 방법이 검토되어야 한다.
- 그러나, 추정 결과를 생산 과정이 전혀 다른 GRDP 결과와 비교하는 것은 여전히 문제가 있다. 총액, 순위 등이 모두 다르기 때문이다. 두 통계를 비교하기 위해서는 현재의 도시가구를 대상으로 한 가계동향조사가 전체가구를 대표할 수 있도록 가계동향조사와 농림어업가구조사 결과를 결합하는 것과 같은 통계를 만들어내야 한다.
- 그러므로 이를 위한 다양한 추정 방법이 더 연구되어야 할 필요가 있다. 즉,
 - 가계동향조사와 농림어업가구조사 결과를 결합하거나
 - 조사통계 결과를 추정에 활용하는 복합추정량의 개발
 - 혹은 두 결과 간 혼란 방지를 위한 ranking calibration 방법 개발 등을 말한다.

늘어나는 지역통계 수요에 비해 예산과 인력은 감소하고 있지만, 그럼에도 시의성 있는 통계는 계속 요구되고 있고 제공되어야 한다. 그런 의미에서 소지역 추정은 좋은 대안 중 하나이다. 소지역 추정이 물론 완벽한 대안은 될 수 없다. 소지역별로 너무 작은 표본 규모가 문제이며, 표본은 가급적 소지역(영역)에 골고루 분포해야 하기 때문이다. 또한 조사 통계에 비해 소지역 추정에 대한 약한 신뢰감도 문제가 된다. 그러나 이는 예산과 인력의 절감, 시의성 확보 등 소지역 추정이 가지는 장점을 고려한다면 충분히 극복 가능하다.

이와 같은 장점을 발현하기 위해서는 분명히 일정 기간 이상의 소지역 추정 연구 기간이 필요하다는 것에는 이의가 있을 수 없다. 또한 소지역 추정 연구가 늘어나는 통계 수요에 대비하기 위한 미래지향적 작업임을 분명히 인식해야 할 필요가 있다.

참고문헌

- 김서영, 권순필 (2008). 고용통계 소지역 추정 연구 I, 통계개발원.
- 김서영, 권순필 (2009). 시군 실업에 대한 조사통계와 추정통계의 비교, 통계개발원.
- 김서영, 권순필 (2010). 고용통계 소지역 추정 연구 II, 통계개발원.
- 김혜원, 정동명 (2005). 가계조사 지역별 추정기법의 타당성 검증. 통계개발팀.
- 문화체육관광부 (2011). 2010년 국민여행실태조사.
- 보건복지가족부 (2009). 2009년 전국보육실태조사 - 가구조사 보고서.
- 통계청 (2004). 가계조사 지역별 추정기법 개발, 용역보고서.
- 통계청 (2010). 2008년 지역소득통계.
- 통계청 (2010). 2011 지역통계발전계획(안), 조사기획과.
- 통계청 (2011). 가계동향조사 이용자 가이드.
- 통계청. 가계최종소비지출 추계 개요.
- 통계청 KOSIS, <http://kosis.kr>.
- Fay, R.E. and George F. Train (1995). Aspects of Survey and Model-Based Postcensal Estimation of Income and Poverty Characteristics for States and Counties, 1995 Proceedings from the American Statistical Association.
- Heady, P., Clarke, P. et al (2003). Model-Based Small Area Estimation Series No.2: Small Area Estimation Project Report, Office for National Statistics (UK) publication.
- U.S. Census Bureau (2010). Small Area Income and Poverty Estimates (SAIPE): 2009 Highlights.
- ONS (Office for National Statistics), Model-Based Estimates of Income for Middle Layer Super Output Areas in England and Wales, 2004/05 User Guide

<부록>

<부표> 지역통계 증장기 수요 목록(17종)

우선순위	분야	통계명	작성기관	작성방법	주기	공표범위	수요(지자체)	조사목적	조사항목
1	가계소비	가계동향조사	통계청	조사	월	전국	16(6)	가구의 생활수준실태와 그 변동사항을 파악하기 위하여 가계의 수입과 지출을 조사하여 국민소비수준변화의 측정 및 분석 각종 경제, 사회정책에 필요한 자료 제공	가구실태(가구주위의 관계, 성별, 연령, 교육정도, 취업여부, 산업, 직업, 주택, 소유여부 등), 가계수지(가구주소득, 가구원 소득, 주거비/방송/교통/통신, 의료/보험/신용카드 등)
	문화	국민여흥실태조사	문화체육관광부	조사	반기	전국	7(6)	우리나라 국민의 소득증대, 생활태도 및 가치관 등의 변화에 따른 여흥실태를 파악하고 국민관광 등 총량 및 관련 지표 조사 분석하여 국민관광에 필요한 기초자료 제공	국내 숙박여행(14개 문항) - 국내 당일여행(13개 문항) - 방문지 및 정보원(8개 문항) - 국내여행 만족도(18개 문항) - 해외여행(19개 문항) - 라이프스타일(1개 문항) - 일반사항(8개 문항)
2	복지	보육실태조사	보건복지부	조사	5년	전국	6(6)	초등학교 이하 아동의 보육교육서비스 이용실태, 이용에 따른 만족도, 요구도 파악 및 보육시설 운영의 전반적 현황 파악을 통한 보육교육 정책 수립을 위한 수요자와 공급자의 보육실태 기초자료 생산	(가구조사) 가구특성 : 가구원 특성, 가구의 주택 및 경제적 특성, 보육에 관한 의견, 아동 부모에 관한 질문, 다문화 가정 질문, 초등학교 이하 자녀의 보육교육이용, 보육아동조사 : 보육 이용 기관 장려, 미취학 아동 하루 중 시간 단위 이용 기관 및 돌보는 사람, 미취학 아동 현재 보육 또는 교육 기관 이용, 전체 아동 특기교육학원 및 보습학원 이용, 전체 아동 개별/그룹 교육지도 및 학습지 활용, 교육 이용, 전체 아동 개인 양육지원 서비스
	주택	주거실태조사	국토해양부	조사	2년	시도	6(5)	민중사는 주택법 제5조에 의해 실시되는 조사로 지역별, 계층별 주거실태 파악, 주택 수요량 및 진후도 파악을 통해 공급계획 지원, 주거복지 수요 파악, 정부정책 수립에 필요한 기초자료 제공	주거 및 주거환경 - 주거이동 및 주택구입계획 - 공공임대주택 관련 - 가구특성 생활실태 및 구매행
3	사회	농촌생필지표조사	농촌진흥청	조사	1년	전국	3(3)	농촌주민의 생활실태를 파악하여 농촌생활에 대한 장기적 변화를 예측하고 농촌주민의 삶의 질 향상을 위한 각종 정책 및 사업에 필요한 기초자료 제공	일반가구현황, 인구 및 사회, 가족생활, 여가생활, 사회복지, 지역개발 및 정보화 부문 등
	인구	전국결혼및출산동향조사	보건복지부	조사	4년	전국	6(3)	저출산의 미시적 원인을 파악하고, 저출산을 극복하기 위한 대책을 마련하는 기초자료를 산출	- 가구조사표, 동거가구인사상 7항목, 비동거가족사상 8항목, 가구일반사항 4항목 - 만20~44세 미혼 자혼사표, 경제활동이력, 현재의 경제상황, 결혼에 대한 가치관, 자식의 결혼, 자녀에 대한 가치관, 향후 출산, 출산정책 등과 관련된 사항 - 만20~44세 기혼여성조사표, 본인과 남편의 현재 경제활동, 남편과의 역할분담, 결혼당시 상황, 부모님과 의 등, 관계, 부부의 형제자매 등

<부표> 지역통계 증장기 수요 목록(17종) (계속)

우선순위	분야	통계명	작성기관	작성방법	주기	공표범위	수요(지자체)	조사목적	조사항목
1	교육	평생학습개인실태조사	한국교육개발원	조사	연	전국	3 (2)	○ 학습사의 학습속도, 학습의 활동패턴 이해 ○ 평생교육과 노동시장과의 이행과정에 대한 심층적 이해를 통해 모든 연령대에 걸친 효과적인 인적자원개발 계획 수립 ○ 우리나라 현황 진단 및 국가간 비교를 위해 OECD 등 국제기구 표준과 호환 가능한 조사 실시	평생교육 참여실태, 비평생교육 참여실태, 무형식 학습 참여실태, 장애요인 및 접근성, 학습유형, 학습습관량, 학습의 사회적 성과, 학습에 대한 태도 등 8개 영역별 지표에 관련된 사항과 응답자 통계적 배경 질문으로 구성
2	교통	대중교통통원황조사	국토해양부	조사	연	5대 광역권	3 (3)	대중교통통원황 조사 결과를 체계적이고 효율적으로 작성 및 관리하여 지방자치단체 및 관련기관의 대중교통기본계획 수립 등을 위한 기초자료로 제공	- 대중교통관련 사회 경제지표 - 대중교통수요자의 경영여건 - 대중교통수단 및 대중교통시설의 현황 - 대중교통의 이용실태 - 대중교통 운행현황 - 대중교통의 개선에 위하여 필요한 사항
3	농림	농산물소득조사	농촌진흥청	조사	연	전국, 도별	5 (5)	지역별 작목별 소득을 조사분석하여 농작경영진단 및 설계 등 경영개선 지도와 농업경영 연구를 위한 기초자료 제공	- 조수입 : 담배연도 생산량, 부산물생산량, 농기수취가격 - 생산비 : 종묘비, 비료비, 농약비, 광열동력비, 수리(水利)비, 제제료비, 소농구비, 수선(修繕)비, 조성비, 임차료, 기타요금, 고용노력비, 자가노력비, 토지자본용역비와 대농기구 및 영농시설에 대한 감가상각비 및 자본용역비 등
4	보건	농업인 직업병 예방활동조사	농촌진흥청	조사	2년	전국	3 (3)	농업인에게 발생된 사고, 질병, 손상 등에 대한 실태를 파악하고 분석하여, 농업활동으로 인한 재해를 예방하고 관리하기 위한 정책 수립 등에 필요한 기초자료 제공	기구의 농업종류와 규모(4항목), 기구원 사항(5항목), 기구 구성원별 농업활동 특성(8항목), 사고/중독/질병 경험(1항목), 농작업과 관련된 사고 및 질병 부문(9항목), 음만성 질환 및 농작업과 관련된 사고/손상 부문(5항목)
5	보건	장애통신서비스 실태조사	보건복지부	조사	3년	전국	4 (4)	출산력 변화에 따른 가족관계 및 가족부양 변화양상 파악 및 관련대책 수립을 위한 자료 생산	- 기구조사 : 기구원 사항, 출생에 관한 사항, 기구에 관한 사항 등 - 부인조사 : 임신 및 출산, 피임, 모자보건, 가족복지, 가족 가치관 등 - 미혼자조사 : 미혼남녀의 가족형성에 관한 의식조사(미혼자조사표)
6	경제	장애인 실태조사	보건복지부	조사	3년	전국	3 (3)	장애유형별 장애인 및 장애특성을 파악하고, 장애인의 복지욕구 복지서비스 수혜 실태 등을 파악함으로써 장애인복지정책 수립의 기초자료로 활용	장애인의 일반특성, 장애특성, 보건의료, 일상생활지원, 장애인복지기구, 보육교육, 취업 및 직업생활, 사회 및 직업활동, 결혼생활, 여장장애인, 생활만족도 및 복력차별 경험, 주거, 복지서비스, 경제상태
7	정보통신	인터넷 이용실태조사	방송통신위원회 (한국인터넷진흥원)	조사	1년	시도	도시 통계	국내 기구의 인터넷 이용환경, 국민의 인터넷 이용 현황과 이용형태 등을 파악하여 관련 정책 수립 및 연구, 국제기구 ICT 관련 국제지수 산출 등의 기초자료 제공	가, 컴퓨터 이용여부, 인터넷 이용여부 및 이용행태, 주요 인터넷 서비스(이메일, 메신저, 블로그, 인터넷쇼핑, banking 등) 이용현황 등 82개 항목

〈부표〉 지역통계 증장기 수요 목록(17종) (계속: 행정자료)

우선 순위	분야	통계명	작성 기관	작성 방법	주기	공표 범위	수요 (지자체)	조사목적	조사항목
3	고용	외국인 근로자(고용허가제)고용동향	고용노동부	보고	월	전국	4 (3)	고용허가제에서의 외국인근로자 취업 추이를 파악하여 외국인근로자의 체계적 도입·관리 및 인력수급의 기초자료로 활용	외국인근로자의 제도별 업종별, 월별 취업자 수 등(외국인고용관리시스템(EPS)을 이용]
		배출시설단속 조치현황	환경부	보고	분기	시도	연계성 연구	폐수 및 대기 배출시설 지도·점검실적 및 행정처분실적을 파악하여 수질·대기 보전대책수립을 위한 기초자료로 활용	지도점검 업소 수, 위반업소 수, 위반내역 및 처분 내역 등 (작성대상:폐수 및 대기 배출시설 설치·허가 신고를 한 모든 사업장)
	에너지	에너지 수급 통계	에너지 경제 연구원	보고	월	전국	도시 통계 연구 /연계성 연구	매월 모든 정유사 및 석유제품 수입사를 대상으로 원유도입에서부터 원유처리, 제품생산, 제품 수출입, 제품 수요, 제고 등에 이르기까지 전반적인 수요·공급 상황을 조사함으로써 정부의 장단기 석유수급계획 및 미장시 석유수급대책 수립 등 석유산업관련 정책의 입안 및 집행, 학계 및 업계 석유수급 관련 자료 제공	<ul style="list-style-type: none"> 정유수급상황(제품별 제고량, 수입량, 임하량, 생산량, 정제량, 출하량, 수출량, 거래회사별 출하량) 수출입상황(제품, 국가별 수출입량, FOB, 운송비, 보험료) 기본자료 기록부(거래정유회사, 자본금, 주유기 수, 저장시설, 수송장비) 거래상황기록부 : 프로판용, 대리점용, 부탄용, 주유소용
4	인적	재난발생 현황	행정안전부	보고	1년	전국	도시 통계	재난관리정책수립과 연구기관, 학계, 각계각층의 재난관련 연구·개발자료 등으로 활용하여 급격한 재난환경 변화에 대응적으로 대처하고, 국가재난관리 역량제고에 기여하고자함	화재, 도로교통, 산불, 폭발, 환경오염, 해난, 기타 재난 발생 현황 집계 통계