

연구보고서 | 2009-06

시군 실업에 대한 조사통계와 추정통계의 비교

권순필 · 김서영

시군 실업에 대한 조사통계와 추정통계의 비교(요약)

□ 목적

- 대규모 표본조사인 「지역별 고용조사」와 소지역 추정 기법에 의한 추정통계의 실증 비교를 통하여 추정 통계의 활용성을 확인해보고자 함

□ 내용

- 소지역 추정량 : 직접, 합성, 복합, EBLUP, HB 등 5개 추정량
- 추정 모형 선택
 - 2변, 4변 모형 선택 : AIC, BIC가 작을수록 좋은 모형임
 - 실제 4개 모형추정량을 모두 계산해 본 결과 EBLUP이, 그 중에서도 EBLUP4가 가장 좋은 결과를 보여줌.

	모형	추정량 표기	AIC	BIC
1	$\hat{y}_i = x_{1i}^T \hat{\beta}_1$, x_{1i} : 실업급여 등록자수	EBLUP1, HB1	-112.215	-110.159
2	$\hat{y}_i = x_{2i}^T \hat{\beta}_2 + x_{3i}^T \hat{\beta}_3 + x_{4i}^T \hat{\beta}_4 + x_{5i}^T \hat{\beta}_5$, $x_{2i}, x_{3i}, x_{4i}, x_{5i}$: 성별 연령별 실업급여 등록자수	EBLUP2, HB2	-119.443	-117.083
3	$\hat{y}_i = x_{6i}^T \hat{\beta}_6$, x_{6i} : 대지역 실업자수 정보	EBLUP3, HB3	-112.526	-110.470
4	$\hat{y}_i = x_{1i}^T \hat{\beta}_1 + x_{6i}^T \hat{\beta}_6$, x_{1i} : 실업급여 등록자수, x_{6i} : 대지역 실업자수 정보	EBLUP4, HB4	-118.401	-116.274

○ 소지역 정의 방법

- ① 특광역시를 제외한 143개 시군 지역
- ② 직접추정량(실업자수)이 0인 지역을 제외한 98개 시군 지역
- ③ 각 도의 군지역을 하나의 소지역으로 집락한 83개 소지역

□ 결과

○ 표본조사구별 143개 지역에 대한 소지역 추정량의 평균ARB와 CV

표본 조사 구수	평균 ARB						평균 CV						
	직접	합성	EBL UP2	EBL UP4	HB2	HB4	고용	직접	합성	EBL UP2	EBL UP4	HB2	HB4
2	1.13	0.62	0.59	0.56	1.01	0.83	39.68	87.14	33.43	20.26	16.24	42.42	37.01
3	1.04	0.50	0.51	0.41	0.77	0.58	31.90	106.65	18.32	16.62	12.85	35.82	28.18
4	0.67	0.42	0.30	0.29	0.45	0.38	26.13	100.46	8.88	11.44	8.07	22.41	15.01
5~6	0.64	0.33	0.38	0.37	0.44	0.55	26.65	80.40	4.84	9.71	6.93	20.71	11.32
7~13	0.39	0.24	0.22	0.23	0.29	0.53	20.48	55.99	7.14	7.40	5.10	16.77	5.67
14+	0.28	0.21	0.18	0.20	0.24	0.48	18.94	36.97	8.46	7.51	4.86	16.02	4.42
평균	0.75	0.42	0.39	0.37	0.59	0.59	28.88	75.48	11.35	13.15	9.90	27.75	19.45

※ 직접, 합성추정량의 지역수는 98개, 나머지 추정량은 143개임

- 표본조사구수가 늘어남에 따라 소지역 추정량들의 CV가 안정적이 됨
- 표본조사구수가 4개 이상이면 보다 안정적인 추정 결과를 얻을 수 있을 것으로 보임

○ 소지역의 정의별 추정량의 평균ARB와 CV, CV≤25% 이하 지역 수

소지역	평균 ARB						평균 CV						
	직접	합성	EBLUP2	EBLUP4	HB2	HB4	고용	직접	합성	EBLUP2	EBLUP4	HB2	HB4
①	0.75	0.42	0.39	0.37	0.59	0.59	28.88	75.48	11.35	13.15	9.90	27.75	19.45
②	0.65	0.28	0.31	0.27	0.84	0.65	24.26	75.48	11.35	11.96	8.94	16.61	13.61
③	0.51	0.43	0.41	0.45	1.54	0.98	22.69	61.96	5.44	12.50	9.77	17.11	24.17

소지역	CV ≤ 25% 인 지역 수						
	고용통계	직접	합성	EBLUP2	EBLUP4	HB2	HB4
①	63	3	85	126	136	86	101

3개 소지역 정의에 5가지 추정량을 적용해 본 결과,

- 추정량 : ARB, CV 측면 모두에서 고용보험 실업급여 등록자수와 대지역의 실업자수를 소지역에 인구 비례로 배분한 정보를 결합하여 모형에 적용한 EBLUP4가 가장 우수
- 소지역 : ②직접추정량(실업자수)이 0인 지역을 제외한 98개 시군 지역에 대한 추정시 추정 결과가 가장 우수
- 인구 규모가 작거나, 표본이 적은 지역 등 관심값(실업자수)의 관측이 쉽지 않은 지역의 추정에 대한 구체적인 고민 필요
- 최소 표본조사구수 확보, 혹은 비슷한 특성의 지역 집락을 고려할 수 있음

□ 논의

- 많은 표본으로 직접 조사를 수행한 결과와 소지역 추정 결과를 실증적으로 비교해 보았다는 데 큰 의의가 있음
 - 그 동안의 연구들이 종합적인 진단과 검토보다는 개괄적인 방법 적용에 치중
 - 소지역 추정치를 비교할만한 적절한 대상의 조사 통계가 없었음.
 - 대부분의 다른 나라들도 모형 또는 방법의 타당성은 센서스 자료를 기준으로 평가하는 경향이 있음(개발원도 동일 연구 기수행 : 김서영과 권순필(2009))
 - 그런 면에서 대규모의 표본을 이용한 「지역별 고용조사」 결과를 대상으로 소지역 추정 결과를 비교해 보게 된 것은 연구, 연구자, 사용자 모두의 입장에서 흔치 않은 좋은 기회임
- 조사의 목적과 활용성 측면에서 조사통계와 추정통계를 단순 비교하고 효율을 논하는 것은 적절하지 않지만, 세부적으로 실업자 수 추정과 같은 단순한 목적으로 두 결과를 비교해본다면, 추정통계가 가지는 장점이 조사통계의 그것보다 훨씬 많음
- 또한, 추정 통계의 개발과 사용은 세계적인 추세이며, 조사 환경 등 여러 측면에서 전통적인 조사 방법보다 비용효율적인 추정 방법으로의 변화를 요구하고 있음
- 그러나 내적 평가와 같은 수치적 연구만으로 추정 통계 사용을 요구할 수는 없음. 추정량을 이용하고자 하는 이용자 입장에서도 충분히 납득할 수 있는 결과를 얻어야 하기 때문임
- 그러므로 학계 전문가, 실무 전문가, 실제 사용자 등과 함께 추정 결과에 대한 기술적, 활용성 등 다양한 측면에서 평가를 진행해야 함.
- 더불어, 가장 중요한 것은 추정 통계를 이용하겠다는 정책적 판단임

□ 기타사항

- 중간발표 때는 모형1과 2를 중심으로 연구 결과를 정리하였으나, 재검토 결과 모형2와 4가 가장 좋은 결과를 보여서 내용 및 결과 중 일부가 수정되었음

목 차

제1절 서론	1
1. 연구배경	1
2. 소지역 추정	3
3. 연구목표와 범위	3
4. 기대효과와 주의	5
제2절 소지역 추정방법	5
1. 직접추정법과 간접추정법	5
2. 모형추정법	7
3. 평가방법	9
제3절 소지역 추정 및 결과	12
1. 자료설명	12
2. 추정결과 진단	20
제4절 지역별 고용조사와 소지역 추정 결과 비교분석	25
1. 지역별 고용조사 자료	26
2. 평가 기준	26
3. 소지역 종류	27
4. 결과 비교	29
제5절 결 론	48
1. 요약	48
2. 향후 연구 및 논의	49

표 목 차

<표 1> 시도별 경제활동인구조사 표본조사구수 분포	31
<표 2> 각 모형의 AIC, BIC	16
<표 3> 직접추정량에 대한 소지역 추정량의 회귀절편과 기울기	20
<표 4> 지역별 조사구수 분포	29
<표 5> 표본조사구수별 143개 지역에 대한 소지역 추정량의 \overline{ARB}	31
<표 6> 표본조사구수별 143개 지역에 대한 소지역 추정량의 \overline{CV}	31
<표 7> 143개 지역 중 $CV \leq 25\%$ 인 지역 수	33
<표 8> 소지역 추정량들의 신뢰구간 비교	34
<표 9> 143개 지역에 대한 소지역 추정량의 \overline{ARB}	43
<표 10> 143개 지역에 대한 소지역 추정량의 \overline{CV}	43
<표 11> 직접추정량이 0인 지역을 제외한 98개 지역에 대한 소지역 추정량 의 \overline{ARB}	45
<표 12> 직접추정량이 0인 지역을 제외한 98개 지역에 대한 소지역 추정량 의 \overline{CV}	45
<표 13> 군지역 집락 후 소지역 추정량의 \overline{ARB}	47
<표 14> 군지역 집락 후 소지역 추정량의 \overline{CV}	47
<부표> 소지역 추정 결과	53

그림 목차

<그림 1> 각 보조정보와 직접추정량과의 관계	51
<그림 2> 모형 추정을 위한 자료 변환	9
<그림 3> 소지역 추정치와 직접추정치와의 관계도	2
<그림 4> EBLUP4와 HB4의 잔차플랏	23
<그림 5> EBLUP4와 HB4의 정규확률플랏	24
<그림 6> 143개 지역에 대한 소지역 실업지수 추정량 : 표본조사구수 2개	5
<그림 7> 143개 지역에 대한 소지역 실업지수 추정량 : 표본조사구수 3개	6
<그림 8> 143개 지역에 대한 소지역 실업지수 추정량 : 표본조사구수 4개	7
<그림 9> 143개 지역에 대한 소지역 실업지수 추정량 : 표본조사구수 5~6개	8
<그림 10> 143개 지역에 대한 소지역 실업지수 추정량 : 표본조사구수 7~13개	9
<그림 11> 143개 지역에 대한 소지역 실업지수 추정량 : 표본조사구수 14개 이상	10
<그림 12> 143개 지역에 대한 소지역 추정량의 ARB	14
<그림 13> 143개 지역에 대한 소지역 추정량의 CV	2

제0장

시군 실업에 대한
조사통계와 추정통계의 비교

권순필 · 김서영

제1절 서론

1. 연구배경

최근 학계에서는 지역경제의 불균형을 분석하는데 관심이 모아지고 있다. 이는 국가 경제 발전과 사회적 결속에 대한 도전이라 할 수 있다. 이러한 관심은 우리나라는 물론이고 북미, 유럽 등 선진국가를 중심으로 높아지고 있으며, 지역 불균형은 경제 전망의 핵심 요인이 되었다. 이와 같은 지역 경제에 대한 관심은 지역통계 수요를 증가하게 하였고 전국이 아닌 국가내 작은 지역 수준에서 그 지역의 고용현황, 소득분포, 빈곤정도, 사회적 배제 등에 대한 연구를 자극하였다고 볼 수 있다.

우리나라는 2005년을 전후로 지역사회의 여러 문제 중 특히 소지역

(시군)의 고용문제에 대한 관심이 높아져 왔다. 통계청은 통계작성에 대한 오랜 전통과 많은 예산을 가지고 있고, 대부분 전국 수준의 조사기반 통계를 생산하는 책임을 지고 있다. 그러나 전국 수준의 조사통계가 때로는 지역수준에서는 만족스럽지 못한 경우가 발생한다(Costa et al., 2006). 그 이유는 다음의 3가지를 들 수 있다: 1) 전국 수준이 아닌 지역에 관련된 문제가 조사를 통해서 잘 조사되지 않는다; 2) 전국 수준에서 수집된 자료는 지역수준에서는 신뢰하기 어렵다; 3) 전국 수준에서 집계된 통계는 소지역에 대한 믿을 만한 정보를 제공하기 어렵다.

2005년 이후 경기도, 경상도, 전라도 일부지역의 지자체에서는 지역의 고용통계(실업률, 취업률 등)를 생산하기 위해 직접 조사를 실시하였다. 그러나 이는 엄청난 예산 소비와 응답자 부담이 가중된다는 측면에서 부정적인 시선도 적지 않다. 어떤 가구나 개인은 사실상 유사한 조사 질문지를 받게 되어 국가 통계청과 지자체 통계 당국의 행동이 일치하지 않는다는 인상을 받기 쉽다. 게다가 응답자들로 하여금 통계청에 대한 부정적인 태도를 야기하고 응답자들은 응답 비용을 중복하여 지불한다는 인상을 주기 쉽다.

그런 가운데 정부는 지자체의 고용현황 파악과 지원 등을 목적으로 시군구 단위의 고용통계를 필요로 하였다. 이에 통계청은 2008년 10월 2주간에 걸쳐 전국 특광역시의 구(Gu)를 제외한 161개 시군(Si-Gun)을 대상으로, 약 17만 표본에 대해 표본조사를 실시하였다. 조사통계의 대안이 준비되어 있지 않은 상태에서 조사 이외의 방법으로 정부의 요구를 수용하기는 현실적으로 어려웠을 것으로 보인다.

이런 점에 비추어 조사통계의 부담을 해결하기 위해 소지역 통계 생산을 위한 대안 마련은 시급한 실정이다. 이에 대해 통계개발원은 고용통계의 소지역 통계 생산에 대한 조사통계의 대안으로 소지역 추정방법을 연구하고 있다. 그렇지만 지금 당장 소지역 추정에 의한 고용통계를 공식통계로서 공표하기에는 몇 가지 검토할 사항들이 남아 있다. 이를 두고 혹자들은 소지역 추정의 편리함과 장점에도 불구하고 추정방법의 선택이라든가 보조정보의 선택문제에 대해 강한 의심을 제기하는 경우가 있다. 이에 다양한 경로에서 소지역 추정에 대한 평가와 검토가 필요하다 하겠다.

2. 소지역 추정

소지역 추정 전문가들의 생각을 빌리자면 우리가 운이 좋게 충분한 예산이 있어서 많은 표본으로 조사할 수 있다면 또는 소지역이 상당히 넓은 지역이라면 표본조사에 의한 직접추정치의 사용도 괜찮다고 말한다(Daniel Elazar, Denise Silver 등과의 e-mail 자문 내용 중). 그러나 대부분의 상황에서 소지역 추정에 의한 상대 MSE(relative Mean Squared Error)가 직접조사 추정치에 의한 랜덤 표본오차보다 항상 좋다(즉, 항상 작다). 물론 모형이 잘못 선택되었다면 이는 추정치와 상대 MSE가 편향될 수 있다는 것은 사실이다. 그러나 모형이 잘 선택되었다면 추정치와 상대 MSE의 편향은 크게 문제되지 않는다. 따라서 소지역 추정에 있어서 모형의 잘못된 선택에 의한 영향이 어느 정도인지를 파악하는 것은 결과의 활용을 위한 중요한 절차이다.

소지역 추정의 또 다른 이점은 행정자료(administrative data)와 조사자료(survey data)를 동시에 사용하는 편리한 방법이라는 것이다. 이렇게 두 자료를 사용하는 것은 조사자료에서 발생하는 비표본오차를 조정하는데 도움이 된다. 충분한 예산이 항상 좋은 자료 생산과 관련되는 것은 아니다. 때때로 소지역에서의 조사는 꽤 복잡해지고, 자료를 정리하고 에디팅(editing)하는 시간이 길어져서 조사 결과의 발표가 지연되기도 한다. 이런 측면에서 보면 자료 품질(data quality)은 다차원의 문제이고, 풍부한 재원이 항상 자료품질을 결정한다고 장담할 수는 없다.

3. 연구목표와 범위

본 연구의 최종 목표는 소지역 추정결과에 대해 평가하는 것이다. 소지역 추정결과에 대한 평가는 다양한 방법을 통해 할 수 있고, 반드시 필요하다. 평가는 크게 내적평가(internal validation)와 외적평가(external validation)로 구분해서 할 수 있다. 내적평가는 다양한 모형 진단(model diagnostic)을 통해 모형의 적정성 평가를 하는 것이고, 외적평가는 사용

자 또는 외부 전문가들을 통해 추정방법과 결과를 평가하는 것이다. 본 연구에서는 다른 조사 자료를 이용한 외적 평가에 초점을 맞추되, 모형 선택에 대해서 몇가지 내적평가도 동시에 수행한다. 평가방법은 2008년 10월 통계청에서 실시한 「지역별 고용조사」와 동시점의 경제활동인 구조사를 이용한 소지역 추정 결과를 비교하는 것이다.

내적 평가에 대해서는 통계개발원은 2008년도 연구를 통해서 모형과 방법에 대한 진단통계량 등을 통해 평가를 수행하였다. 또한 센서스 자료를 이용한 모의실험 연구에서 센서스 값과의 비교를 수행한 바가 있다. 본 연구는 외적 평가의 또 다른 과정으로 실제 조사값과의 비교를 수행하고자 한다. 이는 표본조사 결과에 대해 소표본에 의한 추정결과를 비교해본다는 점에서 그 의미가 있다.

연구방법으로 대표본 조사에 의한 「지역별 고용조사」를 기준값으로 하여 소지역 추정결과를 비교하였다. 비교방법으로는 상대표준오차(CV)와 편향(Bias) 정도를 사용하여 추정결과를 평가하였고, 다양한 시각적 그림(visual plot)들을 이용하여 이용자의 이해를 쉽게 하고자 하였다. 추정방법은 직접추정량과 전통적인 간접추정량, 모형추정량을 고려하였다. 직접추정량은 통계청 경찰조사의 비추정량(ratio estimator)을 사용하였고, 전통적인 간접추정량은 합성추정과(synthetic estimator)과 복합추정량(composite estimator)을 사용하였다. 모형추정량은 최량선형불편추정량(Empirical Best Linear Unbiased Predictor: EBLUP)과 계층적 베이지스 추정량(Hierarchical Bayes estimator: HB)을 사용하였다. 모형은 일반선형모형의 일종인 Fay-Herriot(FH)모형을 사용하였다. FH 모형은 지역효과를 랜덤효과로 사용하여 지역에 따른 효과를 반영하고자 한 모형이다(Rao, 2003). 모형에서 설명변수는 고용보험 실업급여 등록자 수, 대지역(시도)의 실업자 수를 사용하였다.

고용보험 실업급여 등록 자료는 선행연구(김서영과 권순필, 2009)에서 이미 소지역 추정의 보조정보로서 검토한 바 있다. 이 자료는 임금근로자만을 대상으로 한다는 점에서 모든 실업자들을 포함할 수 없다는 한계는 가지고 있다. 그러나 모형추정에서 설명변수가 주된 정보인 종속변수와 완벽하게 설명되는 자료일 수는 없고 그럴 필요도 없다. 즉, 설명변수는 종속변수를 부가적으로 설명해주는 자료로서 상관관계가

높은 자료일수록 좋다. 이런 측면에서 실업급여 등록 자료는 소지역 모형 추정의 보조정보로서 충분한 가치가 있다는 점에 초점을 맞춰야 한다. 추정방법, 모형, 보조정보에 관한 자세한 내용은 선행연구인 김서영과 권순필(2009)을 참고할 수 있다.

4. 기대효과와 주의

본 연구는 소지역 추정 결과를 대표본조사와 비교함으로써 추정에 대한 신뢰성을 높일 수 있을 것으로 기대된다. 또한 조사통계와의 비교 관점에서 소지역 추정 과정에서 발생하는 현실적인 문제를 인지하여 이를 개선할 수 있는 계기가 될 수 있다.

한편 조사통계와의 비교는 조사 추정치가 모수는 아니라는 측면에서 절대적인 비교는 되지 못한다는 제약이 있다. 따라서 조사 추정치와 모형추정치가 적정 CV내에서 또는 신뢰구간 범위 내에서 유사하다는 정도에 의미를 두는 것이 바람직하다 하겠다.

제2절 소지역 추정방법

본 절에서는 연구자들이 고려한 소지역 추정방법과 평가방법을 설명한다. 소지역 추정 방법은 직접, 합성, 복합, EBLUP, HB 등 5가지 방법을 사용하였다. 이들 추정방법에 관한 내용은 선행연구에 자세하게 설명되어 있다(김서영과 권순필, 2009). 여기서는 이들 방법들의 핵심적인 특징만 간단하게 언급하기로 하고, 구체적인 수식이나 이론적 측면은 생략하기로 한다.

1. 직접추정법과 간접추정법

가. 비추정량(Ratio estimator)

경찰조사는 매월 실시되는 연속조사로 비추정량을 사용한다. 소지역 i 에 대해 실업자 수를 y_i 라 하면, 비추정량은 다음과 같다.

$$\hat{y}_i = \sum_{sa=1}^{22} {}_{sa}\hat{y}_i = \sum_{sa=1}^{22} \sum_{h=1}^{n_i} {}_{sa}\hat{y}_{ih} = \sum_{sa=1}^{22} \sum_{h=1}^{n_i} {}_{sa}m_i \cdot {}_{sa}y_{ih}$$

여기서, sa 는 성별(남녀)·연령별(15세 이상 인구를 5세 간격으로 11개 그룹 형성) 특성을 나타내며 총 22개의 범주를 갖는다. n_i 는 i 소지역의 표본조사구수를 나타내고, ${}_{sa}m_i = {}_{sa}\hat{X}_i / {}_{sa}X_i$. 여기서 ${}_{sa}\hat{X}_i$ 은 소지역 i 에 거주하는 성별·연령별 범주의 추계인구값이고, ${}_{sa}X_i$ 은 소지역 i 의 경제활동인구조사에서 조사된 성별·연령별 범주의 표본조사 인구값을 나타낸다. 소지역 i 의 \hat{y}_i 에 대한 분산 추정엔 잭나이프 방법을 사용하였다. 직접추정량은 불편추정량이기 때문에 \hat{y}_i 의 MSE는 \hat{y}_i 의 분산과 같다.

나. 합성추정량(Synthetic estimator)

소지역 i 에 대한 합성추정량은 다음과 같다. 이때 소지역 i 가 포함되어 있는 대지역 안에는 I_k 개의 소지역이 포함되어 있다고 가정한다. 즉, $i = 1, 2, \dots, I_k$ 이고, k 는 대지역을 나타내는 첨자이다.

$$\hat{y}_i^{syn} = \sum_{g=1}^G \frac{{}_g\hat{X}_i}{{}_g\hat{X}} \cdot {}_g\hat{y}_{dir}$$

여기서 ${}_g\hat{X}$ 은 성별-연령별 그룹을 G 개로 나누었을 때, g 번째 그룹에 해당되는 추계인구값을 말한다. 이때 성은 남자와 여자 2개 그룹, 연령은 15~29세와 30세 이상 2개 그룹을 사용하였다. 이런 범주를 사용한 이유는 한국의 경찰조사에 대해 실업률이 성별과 30세 전후 연령별로 나누었을 때 명백한 차이를 보였기 때문이다. ${}_g\hat{y}_{dir}$ 은 경찰조사에서 소지역 i 를 포함하는 대지역의 g 번째 성별-연령별 경찰조사의 비추정값을

나타낸다. 합성추정량 \hat{y}_i^{sym} 의 MSE 추정은 잭나이프 방법을 사용하였다.

합성추정량의 장점은 분산이 작다는 것이지만, 만약 대영역내에서 여기에 포함된 소지역이 대지역의 특성과 유사하다는 가정이 위배될 경우에는 심각한 편향(bias)이 발생할 수 있다. 그런 면에서 합성추정치는 대체로 넓은 지역에 대해 추정할 경우 유리하게 작용할 수 있다.

다. 복합추정량(composite estimator)

복합추정량은 합성추정량에서 편향이 크게 발생하는 것을 보완하기 위해 고안된 방법이다. 그러나 편향이 줄어드는 대신에 MSE가 다소 커지는 문제를 지니고 있다. 복합추정량 \hat{y}_i^{com} 은 다음과 같이 합성추정량과 직접추정량의 가중합의 형태를 사용한다.

$$\hat{y}_i^{com} = w_i \hat{y}_i + (1 - w_i) \hat{y}_i^{sym}$$

w_i 는 직접추정량과 합성추정량의 공분산 $Cov(\hat{y}_i, \hat{y}_i^{sym})$ 이 0이라는 가정 하에 $MSE(\hat{y}_i^{com})$ 를 최소화시키는 값을 사용한다. 연구에서는 근사식으로

로 $\hat{w}_i = \frac{MSE(\hat{Y}_i^{sym})}{MSE(\hat{Y}_i^{sym}) + MSE(\hat{Y}_i^{dir})}$ 을 사용하였다. 이 추정량은 편향

과 표본분산 측면에서 극단적인 값이 발생할 가능성이 적다는 특성을 갖는다. 복합추정량의 MSE 추정은 합성추정량과 마찬가지로 잭나이프 방법을 사용하였다.

2. 모형추정법

본 연구에서는 FH(Fay와 Herriot, 1979)라는 혼합선형모형을 사용하였다. FH 모형은 일반적으로 많이 사용되는 선형모형 중 하나로, 소지역 추정에서 지역효과를 랜덤효과로 가정하여 모형에 반영하였다. 연구에서는 지역기반모형을(area based model)사용하였고, 구체적인 형식은 다음과 같다.

$$y_i = x_i^T \beta + \nu_i + e_i, \quad \nu_i \sim N(0, \sigma_\nu^2), \quad e_i \sim N(0, \psi)$$

여기서 y_i 와 x_i 는 소지역의 종속변수와 설명변수를 나타낸다.

우리 연구에 적용하여 설명하면 종속변수는 경제활동인구조사의 실업자수에 대한 직접추정값, 설명변수로는 실업급여 등록자수, 대지역의 실업자수 등이다. 특히, 실업급여 등록자 수는 고용정보원에서 관리하는 자료로서 임금근로자에 한정되고 있지만, 행정자료로서 표본오차가 없고, 경찰조사의 실업자 수와 밀접한 상관관계가 있다는 점에서 좋은 설명변수 역할을 할 수 있다(김서영과 권순필, 2009).

가. EBLUP 추정법

EBLUP 추정량은 조사값과 FH 모형에 의한 추정값과의 가중합 형태이다. EBLUP 추정량의 형태는 다음과 같다.

$$\hat{y}_i^{eb lup} = \gamma_i y_i + (1 - \gamma_i) x_i^T \hat{\beta}$$

여기서 $\hat{\beta}$ 과 $\hat{\sigma}_\nu^2$ 은 FH에 의한 최우추정량(maximum likelihood estimator)이고, $\gamma_i = \hat{\sigma}_\nu^2 / (\hat{\sigma}_\nu^2 + \psi_i)$ 로 γ_i 는 추정된 분산 성분에 의존한다. 즉, 소지역 i 의 표집오차 ψ_i 가 크면, 그 지역의 EBLUP 추정치는 추정치에 더 많은 영향을 받게 되고, 반대로 지역분산인 $\hat{\sigma}_\nu^2$ 이 커지면 EBLUP추정치는 직접조사값의 영향을 더 받게 된다. EBLUP의 MSE는 $\hat{\sigma}_\nu^2$ 을 최우추정법으로 계산했다는 가정 하에 다음과 같이 구할 수 있다(Rao, 2003).

$$MSE(y_i^{eb lup}) = g_{1i}(\hat{\sigma}_\nu^2) + g_{2i}(\hat{\sigma}_\nu^2)$$

$$\left(\begin{array}{l} g_{1i}(\hat{\sigma}_\nu^2) = \frac{\hat{\sigma}_\nu^2 \psi_i}{\psi_i + \hat{\sigma}_\nu^2} (= \gamma_i \psi_i) \\ g_{2i}(\hat{\sigma}_\nu^2) = (1 - \gamma_i)^2 \mathbf{x}_i^T \left[\sum_{i=1}^m \frac{\mathbf{x}_i^T \mathbf{x}_i}{\psi_i + \hat{\sigma}_\nu^2} \right]^{-1} \mathbf{x}_i \\ \bar{V}(\hat{\sigma}_\nu^2) = [I(\hat{\sigma}_\nu^2)]^{-1} = \left[\sum_{i=1}^m \frac{1}{(\psi_i + \hat{\sigma}_\nu^2)^2} \right]^{-1} \end{array} \right)$$

나. HB 추정법

혼합모형에서 HB는 직접적으로 유도가 가능하다. 특히 FH 모형은 다음과 같이 계층적 모형으로 표현된다.

$$y_i | x_i^T \beta + \nu_i, \psi_i \sim N(x_i^T \beta + \nu_i, \psi_i)$$

$$x_i^T \beta + \nu_i | \beta, \sigma_\nu^2 \sim N(x_i^T \beta, \sigma_\nu^2), \sigma_\nu^2 \sim IG(a, b)$$

여기서 우리는 소지역 i 에 대한 $x_i^T \beta$ 의 사후추정값을 얻기 위해 깃스프본자(gibbs sampler)를 사용하였다. 이때 실제 추정에서 a, b 는 모두 0.01을 사용하였다. 그리고 추정치에 대한 사후평균과 분산은 MCMC 방법에 의해 라오-블랙웰 추정량(Rao-Blackwellized estimator)으로 계산되었다(Rao, 2003). HB방법에서 사용된 이론, 수식 및 프로그램 작성에 관한 자세한 설명은 Rao(2003), 김달호(2005), 김서영과 권순필(2009)을 참고하기 바란다.

3. 평가방법

평가(validation)는 소지역 수준에서 추정치 작성 방법에 대한 평가와 그 추정치 결과 자체의 수용을 보장하기 위해 반드시 필요하다. 평가에는 크게 두 가지 방법이 사용될 수 있다. 즉, 추정방법 및 결과에 대한 내적평가와 외적평가가 그것이다. 내적평가는 사용된 방법에 대한 검증이 핵심이 된다. 즉, 모형추정법의 경우, 편향(bias), 모형의 적합성(goodness of fit), 커버리지(coverage), MSE, 조정(calibration) 진단 등을 통해 내적 평가를 할 수 있다. 본 연구에서도 몇 가지 내적 평가 방법에 의해 모형 추정법을 평가하였다. 본 연구는 사용자들을 위해 내적 평가와 외적 평가방법을 간략하게 소개하기로 한다.

가. 내적평가(Internal validation)

서로 다른 모형과 추정법을 비교하고 그들의 수행능력을 평가하기

위해 다양한 진단방법들이 사용되고 있다. 이러한 방법들은 Brown 등 (2001)이 제안하였고, 간단하게 요약하면 다음과 같다.

1) 편향 진단

모형기반 추정치는 직접추정치(direct estimate)에 대해 불편추정량이어야 한다. 모형기반 추정치에서 예측 편향 검사는 적절하게 크기 조정(scale transformation)된 모형추정치와 직접추정치를 그림으로 표현하는 것이다. 즉, 자연로그 변환인 경우 Y축은 $\ln(\text{모형추정치})$ 로 하고, X축은 $\ln(\text{직접추정치})$ 로 하여 그래프를 그린다. 만약 편향이 거의 발생하지 않는다면, 모형추정치와 직접추정치의 회귀선은 거의 $Y=X$ 라는 회귀선과 일치하게 될 것이다. 이를 통해 편향을 시각적으로 설명할 수 있고, $Y=X$ 라는 회귀선과 비교함으로써 모형추정치의 편향에 대해 유의성 검정을 할 수 있다.

2) 적합성 진단

적합성(goodness of fit) 검정은 모형의 적절성 또는 추정치의 적합성을 검정하기 위해 사용되는 방법이다. 우리는 실제로 우리가 사용한 모형이 만약 직접추정치가 믿을만하다고 했을 때 모형추정치가 직접추정치에 근접해지기를 원한다. 본 연구에서는 모형추정치의 편향에 대한 검사로서 적정 표본규모가 유지되도록 성-연령별, 인구규모별, RMSE 크기별로 그룹으로 나누어 전체 시군에 대해 χ^2 통계량을 사용하여 직접추정치의 기댓값과 모형추정치의 차에 대한 유의성을 검증한다. 그리고, 모형추정에 대한 잔차검정과 잔차들의 Q-Q plot(Quantile-Quantile plot) 등을 통해서 모형 가정을 검증한다.

3) Coverage 진단

이 진단법은 모형에 의한 소지역 추정 절차에 의해 생성된 신뢰구간의 유효성을 평가한다. 이 통계량은 직접추정치와 모형추정치 간에 95% 신뢰구간의 겹치는(overlap) 정도를 측정한다. 이때 직접추정치의 신뢰구간은 모형추정치의 신뢰구간을 포함해야 한다. 검정은 직접추정치의 신

퇴구간 커버지리가 옳다는 전제하에, 이항분포 가정에서의 비모수적 검정방법을 사용한다.

4) Calibration 진단

이 진단은 보다 큰 지역 또는 영역으로 소지역 값들이 합산되었을 때 모형추정값이 직접추정값과 얼마나 다른가를 보여주는 방법이다. 이러한 방법은 모형 추정의 공간적 편향 또는 자기상관성 정도에 대한 증거가 될 수 있다. 이 진단을 통해 어떤 지역 또는 영역으로 소지역 추정값을 조정할 것인지에 대한 기준을 정할 수도 있다. 그러나 본 연구에서는 이 진단 방법을 적용하지 않는다. 그러나 이러한 조정기법은 동일한 내용에 대해 두 통계가 있을 경우 반드시 어떤 수준에서든지 조정이 되어야 하며, 이는 통계작성기관의 신뢰성 문제와 관련이 있다.

나. 외적 평가

외적평가는 특정한 이론적 근거보다는 국내외 학계 전문가, 실무 전문가, 또는 실제 사용자를 통해 할 수 있다. 뿐만 아니라, 비교가능한 자료가 있을 경우, 추정결과와 믿을 만한 다른 통계와의 비교도 하나의 평가 방법이 될 수 있다.

본 연구는 외적 평가의 일환으로 대표본에 의한 표본조사와의 비교를 목표로 한다. 물론 위에서 언급한 내적 평가도 동시에 수행함으로써 추정방법에 대한 내적 타당성 또는 추정결과에 대한 신뢰성을 확인하였다.

관련 전문가 평가에 대해서는 사후적으로 진행할 계획이다.

제3절 소지역 추정 및 결과

본 절에서는 우리가 고려한 소지역 추정을 위해 사용된 기본 자료들과 추정을 위한 이들 자료의 변환 및 처리 방법에 대해서 설명하고, 다양한 진단방법을 통해 추정결과를 평가한다. 고려한 소지역 추정량은 앞에서 설명한 바와 같이 직접, 합성, EBLUP, HB추정량이다. 사용된 보조정보는 고용보험의 실업급여 등록자수 자료와 대지역의 실업자수 정보를 인구크기에 비례하여 소지역으로 할당한 자료이다.

1. 자료설명

가. 경제활동인구조사 자료

소지역 추정을 위한 핵심정보(종속변수)는 경제활동인구의 월별 조사 자료이다. 경제활동인구조사는 국민의 경제활동(취업, 실업, 노동력 등) 특성을 조사함으로써 거시경제 분석과 인력자원의 개발정책 수립에 필요한 기초 자료를 제공하고자 하는 조사이다. 이를 통해 노동공급, 고용구조, 가용 노동시간 및 인력자원의 활용정도를 제공하고 고용정책의 정부입안 및 평가에 필요한 기초자료 제공을 목적으로 한다. 이 조사는 1962년 8월부터 시작된 우리나라의 고용 동향을 보여주는 가장 대표적인 고용통계 표본조사라 할 수 있다.

본 연구의 목적은 전국의 시군구가 아닌 도 단위 이하 지역의 시군별 실업자수에 대한 소지역 통계를 추정하는 것이다. 이를 위해 경제활동인구조사에 대한 표본조사구 현황을 살펴보면 <표 1>과 같다. <표 1>에서 보는 바와 같이 2008년 10월 기준으로 전체 표본 조사구수는 1,629개이고, 특광역시를 제외하면 도단위 지역의 표본조사구수는 891개이다. 본 연구는 구를 제외한 시군별 실업자수 추정을 위해 891개 표본조사구만을 사용한다.

〈표 1〉 시도별 경제활동인구조사 표본조사구수 분포

시도	조사구수	시도	조사구수
서울	207	경기	198
부산	108	강원	90
대구	90	충북	90
인천	99	충남	90
광주	81	전북	81
대전	81	전남	90
울산	72	경북	108
소계	738	경남	99
		제주	45
		소계	891
합계			1,629

나. 보조정보

보조정보는 크게 고용보험의 실업급여 등록자수 자료와 소지역(시·군)이 포함된 대지역(도)의 경제활동인구조사 실업자수 결과를 고려하였다. 이 자료들을 성별·연령별 그룹으로 나누거나, 서로 결합하여 모형에서 설명변수를 2개 이상 사용함으로써 모형의 설명력을 높이고자 하였다.

1) 고용보험 자료

고용보험 자료는 통계개발원의 선행 연구(김서영과 권순필, 2009)에서 사용했던 실업급여 수급자수 자료 대신 실업급여 신규신청자수와 실업급여 순수지급자수(중복자 제외) 자료를 이용하였다. 이는 권혜자와 노현국(2007)이 다음과 같이 계산한 고용보험 실업률 식에 따른 것으로 이 자료가 현재 실업에 대해 보다 설명력있는 자료가 될 것으로 판단된다.

$$\text{고용보험 실업률} = (\text{월 실업급여 등록자수} / \text{월 피보험자 수}) \times 100$$

여기서, 실업급여 등록자수가 실업급여 신규신청자수와 실업급여 순수 지급자(중복자 제외)수의 합을 말한다. 또한, 보조정보의 설명력을 높이기 위하여 고용보험 자료를 성별(남, 여), 연령별(30세 미만과 이상) 4개 범주로 나누어 모형 추정에 이용하였다.

2) 대지역 정보

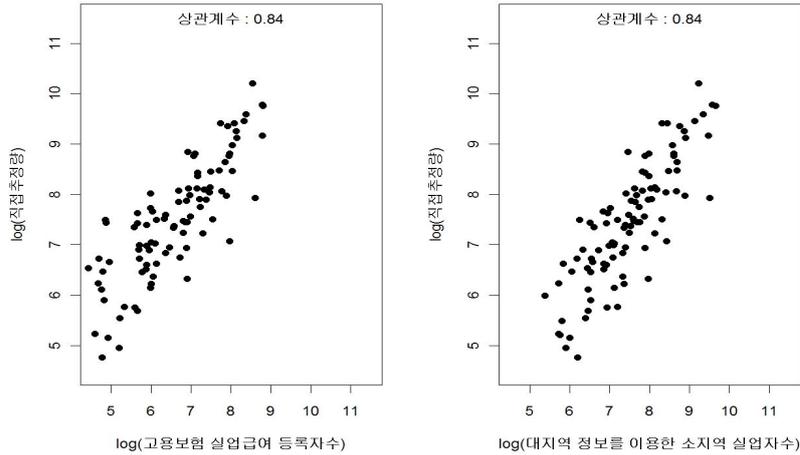
대지역 정보는 정도 높게 추정된 대지역의 실업자수를 소지역의 추계인구에 비례하게 소지역에 배분한 자료이다. 이 경우에 사용가능한 보조정보는 다음과 같이 계산될 수 있다.

$$\text{보조정보} = \frac{\text{소지역의 추계인구}}{\text{대지역의 추계인구}} \times \text{대지역의 추정 실업자수}$$

이는 대지역의 고용동향과 소지역의 고용동향이 강한 상관관계를 가질 것이라는 가정하에 이용한 정보이다. 이 경우는 소지역이 포함된 대지역의 정보를 빌려다 쓴다는 점과 대지역의 추정 실업자수 정보 역시 표본조사라는 측면에서 분산이 과소하게 계산될 수 있다는 한계가 있다.

<그림 1>은 설명변수(보조정보)와 종속변수(직접추정값) 간의 선형 관계를 나타낸 것이다. 그림에서 x 축과 y 축은 각각 자연로그변환된 보조정보와 직접추정값을 의미한다. 왼쪽 그림은 실업급여 등록자수와 직접추정값과의 관계이고, 오른쪽 그림은 대지역 정보를 이용한 소지역 실업자수와 직접추정량과의 관계를 보여준다. 각 그림에서 보조정보와 직접추정값의 상관관계도 각각 0.84로 설명변수로 의미가 있어보인다.

한편, 자료 분석에 따르면 고용상황은 성별과 연령에 따라 차이가 있다. 따라서 이를 모형에 반영하기 위해 고용보험 자료를 성별과 연령별(30세 미만과 이상) 4개 범주로 나누거나, 고용보험 자료와 대지역 정보를 모형에 동시에 적용하는 등 모두 4개의 모형을 추정해 보았다.



<그림 1> 각 보조정보와 직접추정량과의 관계

3) 보조정보 선택

다양한 변수에 대해 최적 모형을 선택하기 위해 AIC(Akaike Information Criteria)와 BIC(Sawa's Bayesian Information Criterion)를 사용하였다. AIC와 BIC는 모두 작은 값을 가질수록 좋은 모형으로 평가한다. 여기에서 AIC와 BIC는 SAS에서 제공되는 값을 사용하였다.

각 모형들은 간단하게 $\hat{y}_i = x_i^T \hat{\beta}_1$ 와 같은 선형회귀식으로 표현할 수 있는데 보조정보의 종류에 따라 다음과 같은 모형을 설정할 수 있다.

모형 1: $\hat{y}_i = x_{1i}^T \hat{\beta}_1$, 여기서, x_{1i} : 실업급여 등록자수,

모형 2: $\hat{y}_i = x_{2i}^T \hat{\beta}_2 + x_{3i}^T \hat{\beta}_3 + x_{4i}^T \hat{\beta}_4 + x_{5i}^T \hat{\beta}_5$,

여기서, $x_{2i}, x_{3i}, x_{4i}, x_{5i}$: 성별 연령별 실업급여 등록자수,

모형 3: $\hat{y}_i = x_{6i}^T \hat{\beta}_6$, 여기서, x_{6i} : 대지역 실업자수 정보,

모형 4: $\hat{y}_i = x_{1i}^T \hat{\beta}_1 + x_{6i}^T \hat{\beta}_6$,

여기서, x_{1i} : 실업급여 등록자수, x_{6i} : 대지역 실업자수 정보

1) 본 연구에서는 종속변수와 설명변수 모두 자연로그변환한 값을 이용하였다. 앞으로 특별한 언급이 없는 한 모형에 이용된 추정값은 자연로그변환한 값을 밝혀둔다.

각 모형에 대한 AIC와 BIC값은 <표 2>와 같이 나타났다.

<표 2> 각 모형의 AIC, BIC

모형	사용한 보조 정보	추정량 표기	AIC	BIC
1	x_1 : 고용보험 실업급여 등록자수	EBLUP1, HB1	-112.215	-110.159
2	x_2, x_3, x_4, x_5 : 성별·연령별(30세 미만과 이상) 4개 범주 실업급여 등록자수	EBLUP2, HB2	-119.443	-117.083
3	x_6 : 대지역 실업자수 정보	EBLUP3, HB3	-112.526	-110.470
4	x_1 : 고용보험 실업급여 등록자수 x_6 : 대지역 실업자수 정보	EBLUP4, HB4	-118.401	-116.274

<표 2>의 AIC, BIC 기준에서 보면 고용보험 실업급여 등록자수를 4개 범주로 나누어서 적용한 모형 2가 가장 좋다. 모형 2와 모형 4는 AIC, BIC값 모두 큰 차이가 없기 때문에 이 기준에서만 본다면 어떤 모형을 사용하더라도 큰 차이는 없다. 그러나 가능한 변수의 개수가 적고 다양한 보조정보를 사용하였기 때문에 추후 모형 설명에 더 유리할 것으로 보이는 모형 4가 본 연구에는 더 적합할 것으로 보인다.

이러한 선택 결과에도 불구하고 본 연구에서는 <표 2>의 4개 모형을 모두 추정하였고, ARB와 CV도 비교해보았다. 그러나 모든 내용을 담기에는 지면이 충분하지 않기 때문에 본 보고서에서는 직접추정량(이하 dir), 합성추정량(이하 syn) 그리고 이들 모형추정 결과(이하 EBLUP2, HB2, EBLUP4, HB4)에 대해서만 소지역 추정량을 평가하도록 하겠다. 모든 모형 추정 결과는 <부표>에 수록해 놓았으니 이를 이용할 수 있다.

복합추정량은 합성추정량과 직접추정량의 가중합의 형태인데, 추정 결과에 따르면, 추정값들의 성향(ARB, CV)이 거의 합성추정량에 유사하였다. 이러한 이유에서 최종 결과 설명시에 복합추정량의 설명은 생략하였다.

다. 자료의 로그변환과 대체

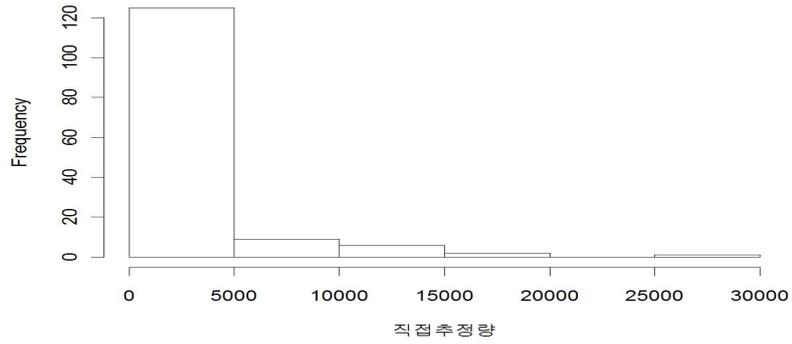
우리가 고려한 FH 모형은 종속변수 y_i 에 대해 정규분포를 가정한다. 이때 사용되는 종속변수 \hat{y}_i 는 경제활동인구조사 시군의 실업자수에 대한 직접추정값으로 연속형 자료가 아니다. 또한 상당히 한쪽으로 치우친 경향을 보인다. 그러므로 대칭성을 가정한 FH 모형에 맞지 않는다. 이런 경우에 많은 연구에서는 자료의 변환을 통해 대칭성을 확보하게 된다. 두 번째로 소지역은 일반적으로 적은 표본수를 갖는다. 게다가 실업자수와 같은 특성치는 쉽게 관측되어지지 않는다. 그러므로 소지역의 실업자수가 전혀 관측되지 않는 경우가 적지 않다. 이 경우 문제는 해당 소지역의 직접추정값에 대한 분산추정값을 구할 수 없고 편향의 가능성이 상존한다는 것이다. 분산추정값이 없을 경우에는 모형에 적용이 어렵기 때문에 해당 소지역에 대한 모형추정도 여의치 않게 되기 때문이다.

그래서 본 연구에서는 소지역의 직접추정값 \hat{y}_i 와 분산추정값 \hat{e}_i 를 모형에 적용하기 전에 이와 같은 문제를 해결할 두 가지 사전 처리를 했다. 첫 번째로는, 소지역의 직접추정량 \hat{y}_i 를 자연로그 변환하였고, 두 번째로는 직접추정량 \hat{y}_i 가 0인 지역에 대하여 실업자수와 분산 추정값을 대체하였다.

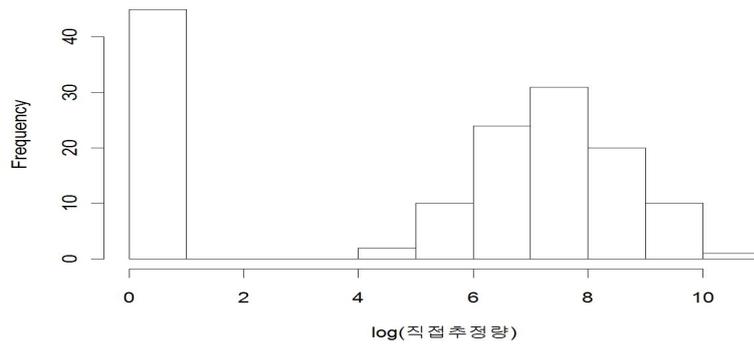
직접추정값을 변환한 이유는 <그림 2>에서 찾을 수 있다. <그림 2>는 각 지역별 직접추정량의 실업자수 \hat{y}_i 들의 분포를 나타낸다. 첫 번째 그림은 직접추정량의 로그 변환 전 값의 분포로 왼쪽으로 치우친 경향을 나타낸다. 이는 자료의 정규성을 가정해야 하는 모형에는 적합하지 않은 형태이다. 따라서, 자연로그 변환을 이용하여 보다 정규성을 가질 수 있도록 변환해 줄 필요가 있다. 로그 변환을 하면 자료의 분포는 <그림 2>의 두 번째 그림과 같다. 이처럼 추정 과정에서 로그 변환하여 얻어진 최종 결과 값에 다시 지수(exponential) 변환을 해주면 실업자수에 대한 추정 결과가 얻어진다.

그러나 직접추정량이 0인 지역 때문에 두 번째 그림처럼 로그 변환된 직접추정값이 두 그룹으로 극단으로 나뉘는 것을 알 수 있다. 이처럼 직접추정량이 0인 경우, 해당 지역의 MSE 계산이 어렵기 때문에 1시점의 추정량인 FH 모형의 EBLUP이나 HB에 적용이 어렵다. 실제적으로도 실업자수가 0이라는 것은 해당 소지역의 실제 값과 명백하게 같지 않다. 즉, 편향이 발생했다는 뜻이기도 하다. 이런 경우에는 해당 소지역의 합성추정값을, 분산값이 0인 경우에는 같은 표본조사구수를 갖는 시군들 중에서 랜덤하게 MSE를 선택하여 대체하는 방법을 사용하였다. 이와 같은 3단계 변환을 통해 모형 추정에 적용되는 값은 <그림 2>의 세 번째 그림처럼 보다 대칭적인 형태로 변화하게 된다.

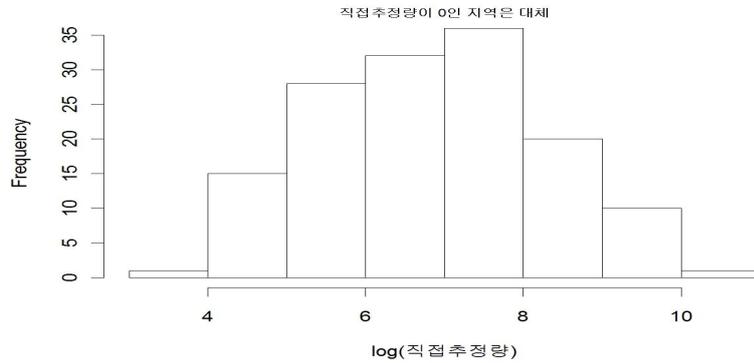
(a) Histogram of 직접추정량



(b) Histogram of log(직접추정량)



(c) Histogram of log(직접추정량)



<그림 2> 모형 추정을 위한 자료 변환

2. 추정결과 진단

가. 편향 진단

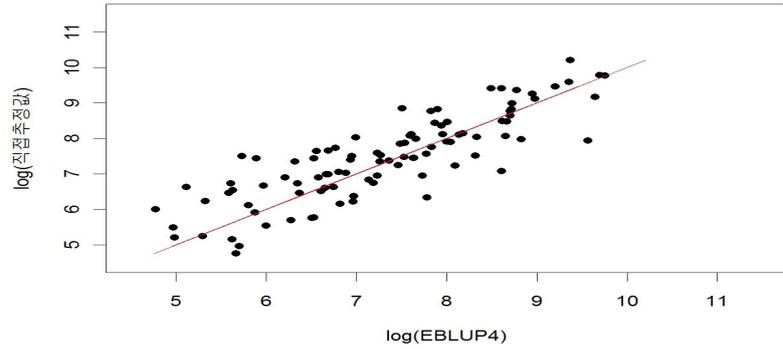
이 방법은 경찰조사 추정치를 종속변수로 하여 소지역 추정을 하였기 때문에 각 추정방법에 의한 추정치들이 경찰조사추정치와 어느 정도 일치적인지를 진단한다. 즉, 경찰조사 추정치(dir)를 X축으로 하고, 각 소지역 추정치를 Y축으로 하여 회귀분석을 수행하여 회귀선 $Y=X$ 와 얼마와 일치하는가를 판단하게 된다. <표 3>은 회귀분석결과로 회귀절편과 기울기 값을 나타낸다. 이때 괄호는 각 계수 값에 대한 표준오차를 의미한다. <표 3>에서 경찰조사 추정치에 대해 가장 일치성을 보이는 것은 HB4로 회귀식은 $Y=-8.034+1.848X$ 이다.

<그림 3>의 (a)~(c)는 EBLUP, HB, 고용조사 추정값과 경찰조사 추정값과의 산점도를 나타낸다. 그림에서 직선은 회귀선 $Y=X$ 이고, 두 추정치가 일치한다면 각각 추정들의 회귀선은 $Y=X$ 에 완전히 겹치게 된다. <그림 3>의 (a)~(c)에 보듯이 각 소지역 추정치들이 경찰조사 추정치에 대해 다소 퍼져있다는 것을 알 수 있다.

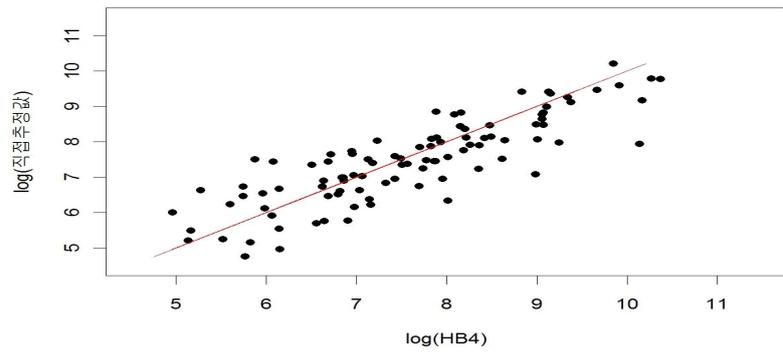
<표 3> 직접추정량에 대한 소지역 추정량의 회귀절편과 기울기

추정량	절편	기울기
EBLUP2	-8.792(1.180)	2.028(0.169)
EBLUP4	-8.512(1.199)	1.987(0.172)
HB2	-11.260(1.373)	2.292(0.190)
HB4	-8.034(1.174)	1.848(0.162)

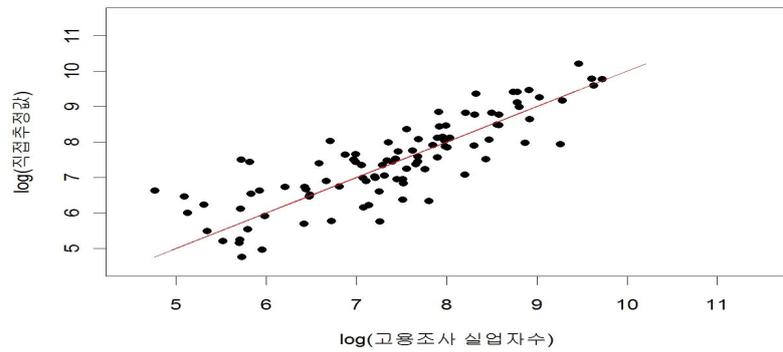
(a) EBLUP4와 직접추정값



(b) HB4와 직접추정값



(c) 고용조사 실업자수와 직접추정값



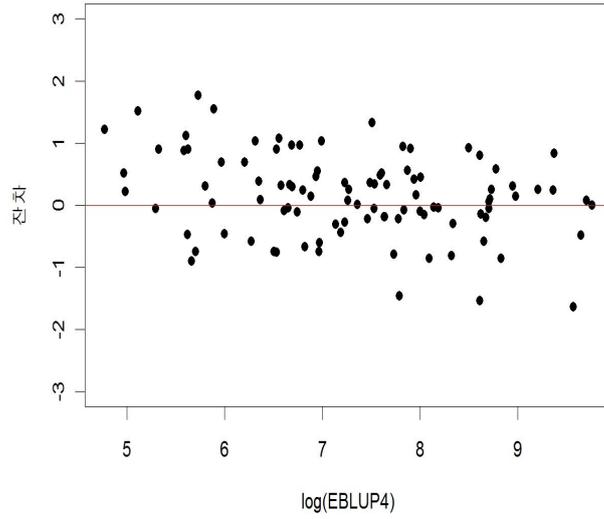
<그림 3> 소지역 추정치와 직접추정치와의 관계도

<그림 3>의 (c)는 지역별고용조사와 경찰조사와의 산점도를 나타낸다. 이는 대표본조사와 소표본조사의 경향을 이해하는데 도움이 된다. X축과 Y축이 작은 값일 경우 다소 산포가 크기도 하지만, 두 조사값은 다른 소지역 추정값과 경찰조사값과의 경향과 크게 다르지 않다.

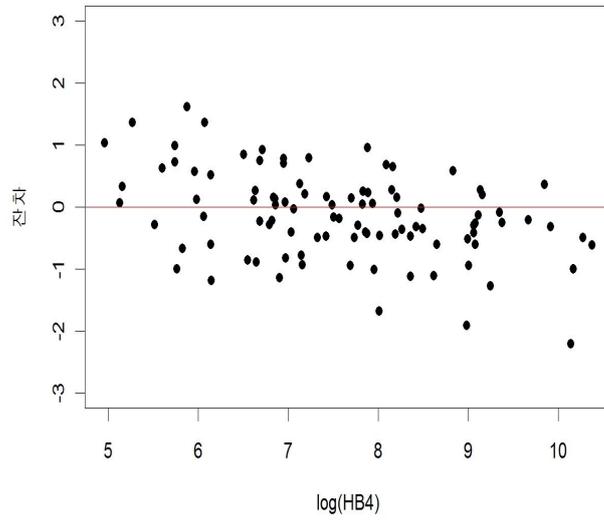
나. 적합성 진단

모형추정치에 대한 적합성 검토로 추정치들의 잔차그림과 정규확률플롯을 이용하였다. <그림 4>는 EBLUP4와 HB4의 잔차플롯, <그림 5>는 정규확률플롯을 나타낸다. 잔차플롯을 보면 '0'을 중심으로 위아래로 ± 2 내에서 잔차가 고르게 분포하고 있어 등분산성에는 큰 문제가 없어 보인다. 또한 각 정규확률플롯은 잔차가 직선 근처에 분포하고 있어 정규분포가정을 대체로 만족하고 있음을 알 수 있다. 그러나 양쪽 꼬리부분에서는 약간의 극값이 존재하는 것으로 보여지는데, 이는 특별한 지역적 특성을 갖는 지역들에서의 추정치들도 전체 평균으로 조정되는 경향이 있기 때문으로 보인다.

(a) EBLUP4의 잔차 플랏

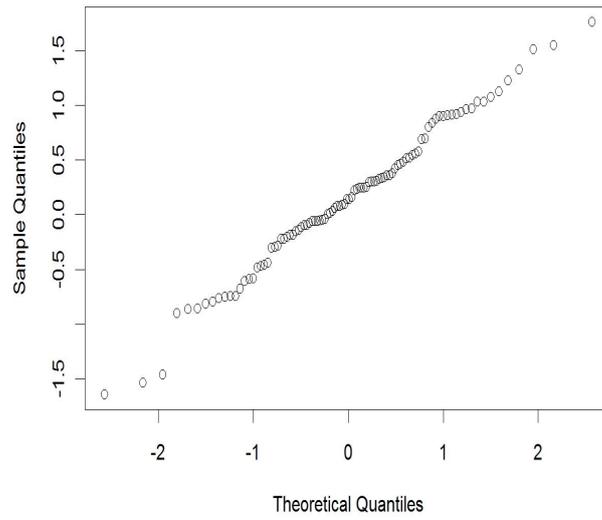


(b) HB4의 잔차 플랏

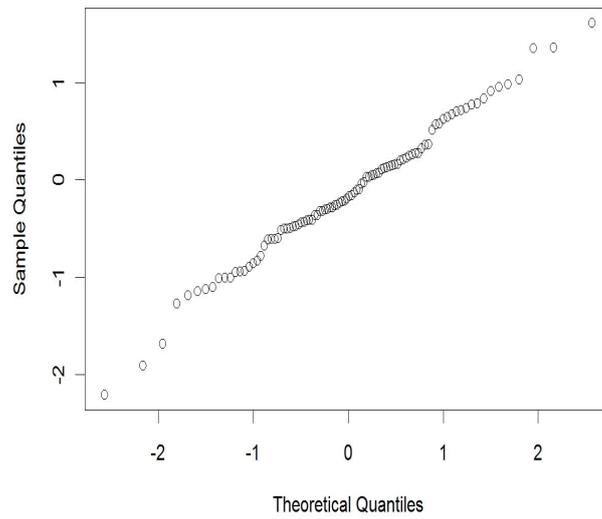


<그림 4> EBLUP4와 HB4의 잔차플랏

(a) EBLUP4의 Q-Q Plot



(b) HB4의 Q-Q Plot



<그림 5> EBLUP4와 HB4의 정규확률플랏

다. 커버리지 진단

커버리지 진단은 신뢰할만한 조사통계에 대한 95% 신뢰구간을 구했을 때, 각 소지역 추정방법에 의한 추정치의 95% 신뢰구간이 조사통계의 신뢰구간과 어느 정도 겹치는지를 판단한다. 그런데 경찰조사의 비추정치 CV가 크고, 신뢰구간의 폭이 매우 크기 때문에 경찰조사의 신뢰구간을 이용한 비교는 어렵다. 대안으로, 지역별 고용조사결과를 이용하고자 한다. 물론, 소지역 추정이 지역별고용조사를 이용하여 추정한 것이 아니기 때문에 이 또한 비교에 완전한 자료는 아니다. 그럼에도 불구하고, 지역별고용조사를 비교대상으로 사용한 것은 이 조사는 경찰조사와 마찬가지로 동일한 모집단에 대한 모수를 추정한다는 점에서 유사하면서 표본크기와 CV가 믿을만하다는 측면에서 간접적인 비교의 의미를 가질 수 있을 것으로 판단되었기 때문이다. 이때 소지역 추정의 CV가 25%를 넘지 않은 지역만을 비교대상으로 하였다.

방법은 소지역 추정에 의한 신뢰구간이 지역별고용조사의 신뢰구간에 포함된 지역의 수에 대해 이항분포 검정을 수행하였다. EBLUP4의 경우 $CV \leq 25\%$ 를 만족하는 136개 지역 중 두 신뢰구간이 포함관계에 있지 않은 지역은 46개 지역으로 약 33.8%를 차지하였다. 이러한 포함관계가 95%의 신뢰도를 유지할 것이라는 통계적 귀무가설은 기각되어 두 조사의 포함관계는 95%의 신뢰도 유지에 대해 통계적으로 유의하지 않은 것으로 나타났다($p=0.001$). 통계적으로 95% 신뢰구간내에서 통계적으로 유사하다고 볼 수는 없지만 $CV \leq 25\%$ 를 만족하는 136개 지역 중 약 66.2%가 두 통계간에 유사한 결과값을 제공한다고 볼 수 있다. 다른 추정방법에 대해서도 유사한 결과를 보였다. 물론 이 결과는 소지역 추정에 사용된 경찰조사와의 비교가 아니라는 점에서 크게 신경쓸 필요는 없고, 다만 추정결과의 진단을 위해 이러한 방법이 사용될 수 있다는 점을 이해하는데 의미를 둘 수 있다.

제4절 지역별 고용조사와 소지역 추정 결과 비교분석

본 절에서는 조사통계와 추정통계의 결과를 실증적으로 비교해보기

로 한다. 비교대상 자료는 「지역별 고용조사」이다. 모형 추정을 위해 경제활동인구조사 자료를 핵심정보(종속변수)로 하고, 고용보험 자료를 보조정보(설명변수)로 사용한 4개의 소지역 추정량을 이용하였다. 조사 결과의 비교 시점은 2008년 10월이다.

결과 비교시 유의할 점은 대규모 표본조사라는 점에서 지역별 고용 조사와 추정통계의 비교는 유의미한 작업이지만, 지역별 고용조사 역시 전수조사가 아니기 때문에 기본적인 오차를 포함하고 있다는 것이다. 그러므로 ARB만 가지고 어느 추정량이 단정적으로 우수하다고 말하는 것은 무리가 있다. 추정량들의 경향, ARB, CV, 신뢰구간 비교 등 다양한 측면에서 모두 비교 검토해보는 것이 필요하다.

1. 지역별 고용조사 자료

지역별 고용조사는 지역 고용정책 수립에 필요한 시군 단위의 세분화된 고용현황을 파악할 수 있는 기본 통계를 생산·제공하기 위해 2008년 10월을 기준으로 조사된 고용통계이다. 이 조사의 경우 광역 시·도에 대해서만 신뢰할 수 있는 추정량을 제공할 수 있었던 기존의 경제활동인구조사(약 3만 가구)에 비해 6배 정도인 약 17만 가구를 표본으로 조사를 수행했다. 조사는 경제활동인구조사와 마찬가지로 조사원이 각 표본 가구를 직접 방문하여 면접조사를 통하여 가구원의 인적사항, 취업자, 실업자, 비경제활동인구와 관련된 31개 항목을 조사한다. 이 조사는 실제적으로 조사 당시의 고용 상황을 알 수 있는 가장 대규모의 표본조사이다. 본 연구에서는 이 지역별 고용조사 결과 중 시군별 실업자수 조사 결과를 이용하여 소지역 추정 통계와의 결과를 비교해본다.

2. 평가 기준

소지역 추정량은 절대상대편향(Absolute Relative Bias: ARB)과 오차

율(Coefficient of Variation: CV), 이를 이용한 신뢰구간의 비교를 통해서 평가하였다. ARB는 통계청에서 실시한 2008년 10월 지역별 고용조사 결과 θ_i 를 기준으로 각 추정량들이 다음과 같은 계산으로 구해진다.

$$ARB_i = \frac{|\hat{\theta}_i - \theta_i|}{\theta_i}$$

물론, 지역별 고용조사 결과 역시 소지역들의 실업자수에 대한 참값은 아니다. 그러므로 ARB에 대한 해석에는 주의해야 할 필요가 있다. 즉, ARB는 두 통계를 용이하게 비교하기 위한 방편으로만 사용하기로 한다. 전체 m 개의 소지역 추정량에 대한 절대상대편향의 평균은 다음과 같이 구할 수 있다. 즉,

$$\overline{ARB} = \frac{1}{m} \sum_{i=1}^m ARB_i, \quad m \text{은 소지역수.}$$

ARB는 0에 가까울수록 좋고 이는 상대적으로 두 추정량의 차이(향후 편향이라고 통칭하겠음)가 작음을 의미한다.

오차율 CV는 추정량 계산시 다음과 같이 구해진다.

$$CV_i = \frac{\sqrt{MSE_i}}{\theta_i} \times 100$$

CV도 역시 0에 가까울수록 좋다. 전체 m 개의 소지역 추정량에 대한 평균 오차율은 다음과 같이 구할 수 있다.

$$\overline{CV} = \frac{1}{m} \sum_{i=1}^m CV_i, \quad m \text{은 소지역수.}$$

그리고 각 추정량별로 $CV \leq 25\%$ 이하인 지역 수를 비교해보았다. 통상적으로 $CV \leq 25\%$ 이면, 해당 지역에 대한 추정 결과는 발표할 만하다고 볼 수 있고, 이 조건에 해당하는 지역이 많을수록 보다 나은 추정량으로 간주하였다.

3. 소지역 종류

추정을 위한 소지역은 9개 도의 하부 단위인 158개 시군으로 단순 정의하였다. <표 4>는 각 지역의 조사구수 및 직접추정량이 0인 지역

현황을 나타낸다. 이 중 조사구수가 1개 이하인 15개 지역은 MSE를 계산할 수가 없다. 또한 MSE를 대체하기도 쉽지 않기 때문에 모형 추정에서는 이들 지역을 제외하고, 나머지 143개 지역에 대해서만 추정하였다.

143개 소지역 중, 45개 지역에서는 경제활동인구조사에서 실업자수가 전혀 관측되지 않아 직접추정량은 0으로, MSE는 계산되지 않는 현상이 발생하였다. 조사구수가 1개인 지역과 마찬가지로 MSE가 계산되지 않으면 이들 자료는 모형 추정에서 사용될 수 없다. 그러나 조사구수가 2개 이상인 지역들은 비슷한 규모의 다른 지역의 MSE를 이용해 랜덤 대체할 수 있었다.

그래서 본 연구에서는 이런 여러 가지 상황을 모두 가정하고, 다양한 해법을 찾아보기 위해 다음과 같이 3개의 방법으로 소지역을 다시 정의한 후 각각 추정하였다.

- ① 자료 대체를 통한 143개 지역에 대한 추정
- ② 직접추정량이 0인 지역을 제외한 추정 결과(98개 지역)
- ③ 군지역 집락 추정(83개 지역으로 소지역 추정 후 143개 지역에 대한 추정 결과 배분)

첫 번째 지역 정의는, 직접추정값이 0으로 나타난 지역에 대하여 해당 실업자수는 해당 지역의 합성추정값으로, MSE는 같은 표본조사구수를 가지는 타 지역의 MSE 결과를 랜덤 대체하여 모형 추정에 적용한 것이다. 그러므로 143개 지역 모두에 대해서 추정이 가능하였다. 두 번째 지역 정의는 첫 번째 지역 정의에서 사용한 대체가 원 자료를 훼손할 수 있고, 이와 같이 직접추정량이 0인 경우가 없을 때, 소지역 추정 결과는 어떻게 나타날 수 있는지 확인해보고자 직접추정량이 0으로 나온 지역은 제외하고 98개 지역을 대상으로만 소지역 추정을 시도해보았다. 그러나 이 방법은 연구 목적으로서만 가능할 뿐 실제적으로 매 추정 때마다 직접추정량이 0으로 나오지 않는 지역에 대해서만 추정한다면 소지역 추정 통계로서 제공되기에는 어려움이 있다. 또한 직접추정량이 0으로 나온 지역의 특성이 반영되지 않는 문제가 있다.

마지막으로, 조사구수가 1개 이하이거나 직접추정량이 0으로 나오는

경우는 대체로 군 단위 지역인 경우가 많다(9개 시, 36개 군). 그래서 군 지역을 하나의 소지역으로 묶어서 143개 소지역을 83개로 줄여서 추정을 시도해보았다. 보다 작은 군 지역의 추정량은 집락화된 지역의 추정 결과를 각 군의 추계인구에 비례하여 재할당하였다. 이 경우에는 조사구수가 1개 이하인 경우에도 추정량을 제공할 수 있다는 장점이 있었으며, 표본조사구 수가 증가되지 않는 경우에 시도해 볼 수 있는 좋은 대안이 될 수 있을 것으로 보인다. 그러나 보다 작은 군 지역의 추정 결과에 대해서는 오차율을 제공할 수 없다는 단점도 있다.

〈표 4〉 지역별 조사구수 분포

도	계	시	군	조사구수	조사구수 1 인 시군	조사구수 0 인 시군
				≥ 2 인 시군		
경기	31	27	4	30	4	1
강원	18	7	11	17	7	1
충북	12	3	9	11	2	1
충남	16	7	9	14	4	1
전북	14	6	8	14	4	
전남	22	5	17	15	5	6
경북	23	10	13	20	10	2
경남	20	10	10	20	9	
제주	2	2		2		
계	158	77	81	143	45	11

4. 결과 비교

모의실험을 수행해 본 결과를 요약하면 다음과 같다. 가장 중요한 것은 우리나라 소지역 고용통계 작성시에 소지역 추정량(특히, 모형추정량)의 활용성이 충분하다는 것이다. 특히 EBLUP 추정량의 경우, ARB, CV 측면에서 고용조사, 직접추정량, 합성추정량에 비해 뚜렷한 개선 효과가 나타났다. HB 추정량은 지역의 수가 감소함에 따라 ARB가 다소 증가하는 경향이 나타났는데, 이는 추정의 문제라기보다는 로그 변환된 자료를 다시 지수 변환하는 과정에서 편향이 발생한 것으로 판단된다.

이 부분에 대해서는 향후 추가 연구가 필요한 것으로 보인다.

또한 이러한 소지역 추정법들이 표본의 인위적인 증가가 없어도 정확성과 정도 측면에서 신뢰할만한 추정 결과를 제공한다는 것을 알 수 있다. 그러나 실업자수와 같이 관측이 쉽지 않은 특성값을 추정하기 위해서는 추정 기법과는 별개로 특성값을 관측할 수 있는 조치가 필요한 것으로 보인다. 그러한 방법으로는 최소 표본조사구수를 확보하거나 혹은 비슷한 특성의 지역을 집락하여 특성값을 관측할 수 있는 기대를 높이는 것이 필요하다.

가. 자료 대체를 통한 143개 지역에 대한 추정

우선 경제활동인구 표본조사구수가 2개 이상인 143개의 모든 시군에 대하여 소지역 추정을 시도해보았다. 자세한 지역별 추정 결과는 <부표>를 참고할 수 있다²⁾. <표 5>와 <표 6>는 표본조사구수별 143개 지역에 대한 소지역 추정량의 \overline{ARB} 와 \overline{CV} 결과를 요약해 보았고, <그림 6> ~ <그림 11>에는 표본조사구수별 143개 지역의 고용조사 결과와 직접, 합성, EBLUP4, HB4의 실업자수 추정값을 그림으로 표현해 보았다. <표 5>와 <표 6>에 나타난 바와 같이, 적어도 표본조사구수가 4개 이상이면 3개 이하일 때보다 \overline{ARB} 와 \overline{CV} 가 크게 감소하는 것을 알 수 있다.

<표 7>은 143개 지역 중 $CV \leq 25\%$ 인 지역수를 도별로 요약해 보았다. <표 7>에서 $CV \leq 25\%$ 를 공표수준으로 볼 때 직접추정량을 제외한 모든 소지역 추정값의 $CV \leq 25\%$ 인 지역수는 고용조사 63개보다 훨씬 많다. 가장 CV 결과가 좋은 추정량은 EBLUP4(143개 지역 중 136개)와 EBLUP2(143개 지역 중 126개) 순으로 나타났다.

2) 지면의 한계상 ② 직접추정량이 0인 지역을 제외한 추정 결과(98개 지역)와 ③ 군지역 집락 추정(83개 지역으로 소지역 추정 후 143개 지역에 대한 추정 결과 배분) 결과는 본 보고서에서는 수록하지 않았다. 자세한 결과를 알고 싶은 경우 연구자에게 개별 연락해 주시기 바란다.

<표 5> 표본조사구수별 143개 지역에 대한 소지역 추정량의 \overline{ARB}

표본조사구수	\overline{ARB}					
	직접	합성	EBLUP2	EBLUP4	HB2	HB4
2	1.13	0.62	0.59	0.56	1.01	0.83
3	1.04	0.50	0.51	0.41	0.77	0.58
4	0.67	0.42	0.30	0.29	0.45	0.38
5~6	0.64	0.33	0.38	0.37	0.44	0.55
7~13	0.39	0.24	0.22	0.23	0.29	0.53
14+	0.28	0.21	0.18	0.20	0.24	0.48
평균	0.75	0.42	0.39	0.37	0.59	0.59

<표 6> 표본조사구수별 143개 지역에 대한 소지역 추정량의 \overline{CV}

표본조사구수	\overline{CV}						
	고용통계	직접	합성	EBLUP2	EBLUP4	HB2	HB4
2	39.68	87.14	33.43	20.26	16.24	42.42	37.01
3	31.90	106.65	18.32	16.62	12.85	35.82	28.18
4	26.13	100.46	8.88	11.44	8.07	22.41	15.01
5~6	26.65	80.40	4.84	9.71	6.93	20.71	11.32
7~13	20.48	55.99	7.14	7.40	5.10	16.77	5.67
14+	18.94	36.97	8.46	7.51	4.86	16.02	4.42
평균	28.88	75.48	11.35	13.15	9.90	27.75	19.45

* 직접, 합성추정량의 지역수는 98개, 나머지 추정량은 143개임

<표 7> 143개 지역 중 $CV \leq 25\%$ 인 지역 수

도	CV ≤ 25% 인 지역 수						
	고용통계	직접	합성	EBLUP2	EBLUP4	HB2	HB4
경기	27		24	30	30	26	29
강원	1		9	14	15	7	7
충북	6	1	7	11	11	6	10
충남	5	1	8	13	14	9	13
전북	4		8	8	12	7	7
전남	4		10	12	14	6	6
경북	7		8	16	18	11	13
경남	7	1	9	20	20	12	14
제주	2		2	2	2	2	2
계	63	3	85	126	136	86	101

<표 8>은 지역별 고용조사와 각 추정량들의 신뢰구간을 비교해본 것이다. 여기서 신뢰구간이 겹친다는 것은 소지역 추정량들의 신뢰구간 중 일부(100%, 80%, 50% 등)가 지역별 고용조사의 신뢰구간과 겹친다는 것을 의미한다. <표 8>에서 알 수 있듯이 고용조사의 신뢰구간과 소지역 추정 결과가 완전히 겹치거나 80% 이상 겹치는 지역은 합성추정량이 72(73.47%), EBLUP2가 98(68.53%), EBLUP4가 102(71.33), HB2가 65(45.45%), HB4가 73(51.05%)개로 상당히 유사한 경향을 보인다는 것을 알 수 있다. 그러나 두 통계 중 어느 한 쪽의 CV가 너무 클 경우의 포함 관계를 단순 비교할 경우에는 그 신뢰성이 다소 떨어질 수도 있다. 왜냐하면 표본이 작은 지역은 CV가 크기 때문에 이런 지역도 신뢰구간 비교에서 좋은 결과를 보이는 것으로 판단될 수 있기 때문이다. 그래서 추정량의 CV가 25%이하인 지역에서 80%이상 겹치는 경향을 살펴보면 합성추정량은 77.65%, EBLUP2는 70.63%, EBLUP4는 72.06%, HB2는 45.35%, HB4는 47.52%로 합성추정량과 EBLUP 추정량이 상당히 고용조사의 신뢰구간과 유사한 것으로 나타났다. 이 중에서 합성추정량은 143개 지역 중 표본 규모가 정도 이상인 98개 지역에 대해서만 추정되었기 때문에 다른 모형추정량들의 신뢰구간 유사 정도와 단순 비교하는 것은 무리가 있다.

전체적으로는 직접추정량을 제외하면 지역별 고용조사 결과와 많이 유사한 것으로 나타났다. 평균 ARB를 살펴보면 그 중에서도 EBLUP2와 EBLUP4가 고용조사와 가장 비슷한 경향을 갖는 것으로 나타났다(참고 <표 9>). 그리고, EBLUP과 HB 같은 모형추정량의 경우, <그림 12>에 보이는 바와 같이 강원도 삼척에서 ARB 값이 높게 나와서 평균이 전체적으로 높게 나타났으나, 이 지역에 대한 추정 결과를 제외하면 평균 ARB 값이 크게 떨어지면서 고용조사 결과와 견주었을 때 그 정확성을 비교할 만 하다.

CV 측면은 <표 10>에서 보면 EBLUP4, EBLUP2, 합성추정량이 각각 9.90, 13.15, 11.35로 좋은 결과를 보여준다. 그러나 합성추정량의 경우 실업자수가 전혀 관측되지 않은 45개 지역에 대한 결과를 포함하지 않은 값으로 다른 추정량의 경우, 이 값들을 제외한 결과를 보면 모형추정량들의 평균 CV도 EBLUP2가 10.73, EBLUP4가 7.76, HB2가 21.46, HB4

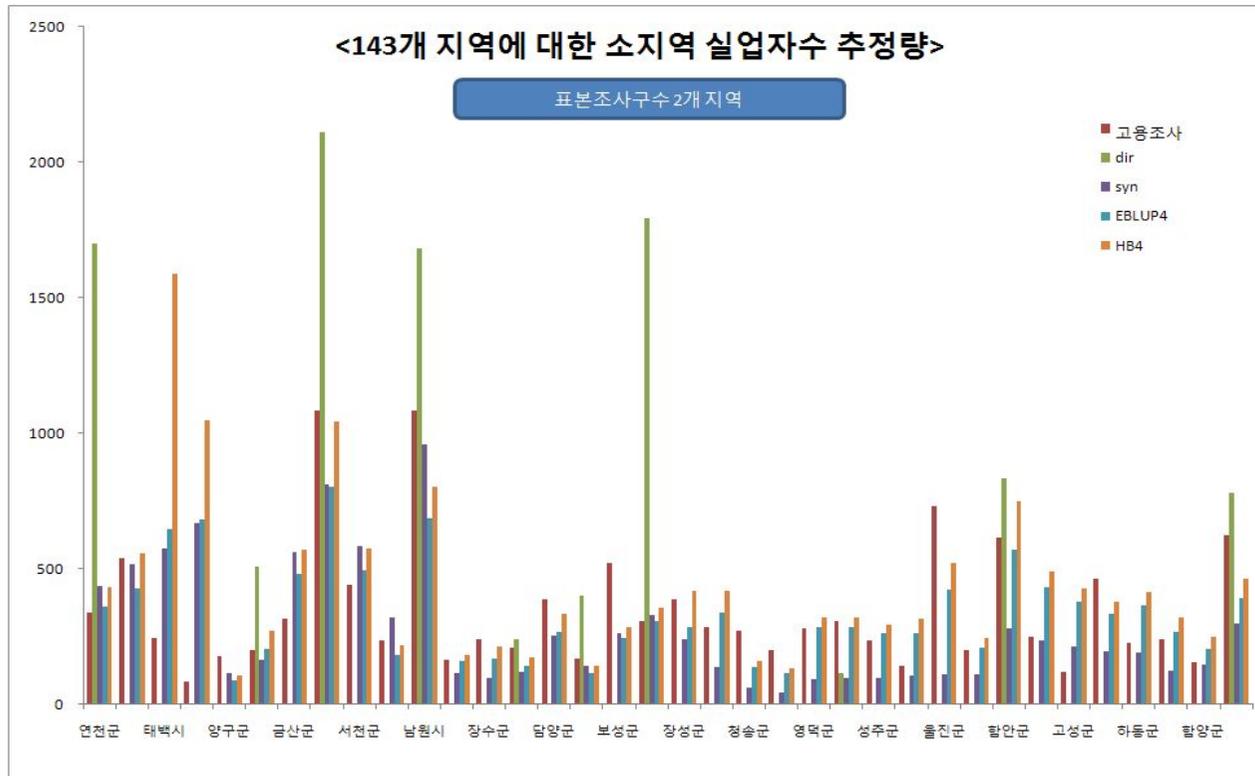
가 31.72로 좋은 결과를 보여준다. 특히 순창군 같은 지역의 경우에는 고용조사를 포함하여 모든 소지역 추정량의 CV가 25%를 넘었다. 고용조사에서 순창군은 33개의 표본으로 조사를 수행하였고, 경제활동인구조사에서는 2개의 표본이 사용되었다. 모든 추정량의 CV가 25%를 넘지는 않았지만 이런 현상은 연천군, 양양군, 예산군 등에서도 확인이 가능하다. 즉, 이는 추정 뿐만 아니라 조사통계 역시 특성값의 관측이 어려운 지역에서 표본을 무조건 확보하는 것만으로 통계값의 정도를 높일 수 없다는 것을 의미한다.

결과적으로 <표 8>을 보면 143개 지역 중 CV가 25%이하인 지역수는 고용통계조사 결과가 63개인데 비해 소지역 추정량들은 합성추정량이 85개, EBLUP2가 126개, EBLUP4가 136개 등으로 대체로 지역별 고용조사보다는 많은 지역이 CV 측면에서 공표 가능하다고 보인다. 특히 강원도는 CV가 25%이하인 지역이 지역별 고용통계에서 1개 뿐이었지만, 소지역 추정법을 이용하면 7개 이상으로 늘어난다.

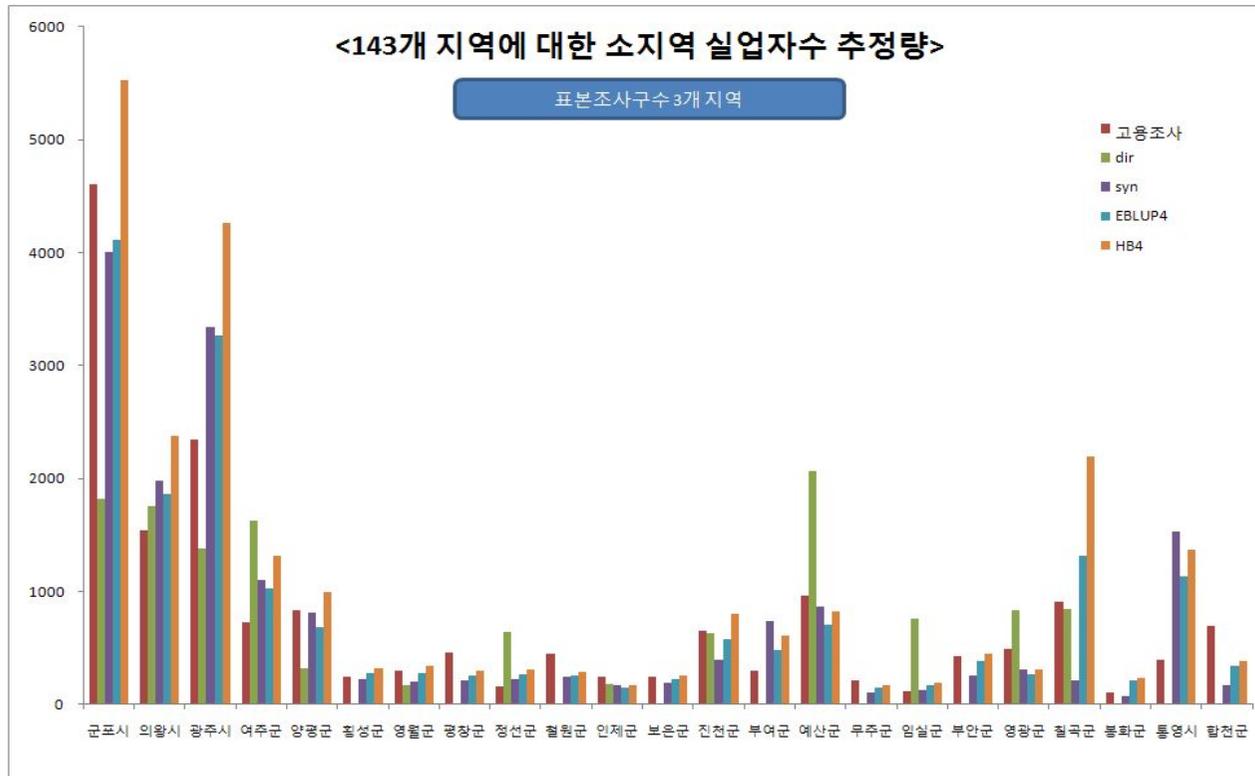
또한 <그림 13>를 보면 표본조사구수가 늘어남에 따라 소지역 추정량들의 CV가 안정적으로 변하는 것을 알 수 있다. 특히, 표본조사구수가 4개 이상이면 지역별 고용조사 보다도 안정적인 결과를 얻을 수 있다. 이는 역으로 표본이 너무 작은 지역에 대해서는 소규모의 표본 증가를 통해서 보다 안정적인 소지역 추정량을 얻을 수 있다는 것을 의미한다. 그러므로 표본을 무조건 확보하는 것으로 추정량의 정도를 높일 수는 없지만, 소지역 추정을 위해 기본적으로 일정 규모 이상의 표본이 필요하다는 것은 확인할 수 있다.

〈표 8〉 소지역 추정량들의 신뢰구간 비교

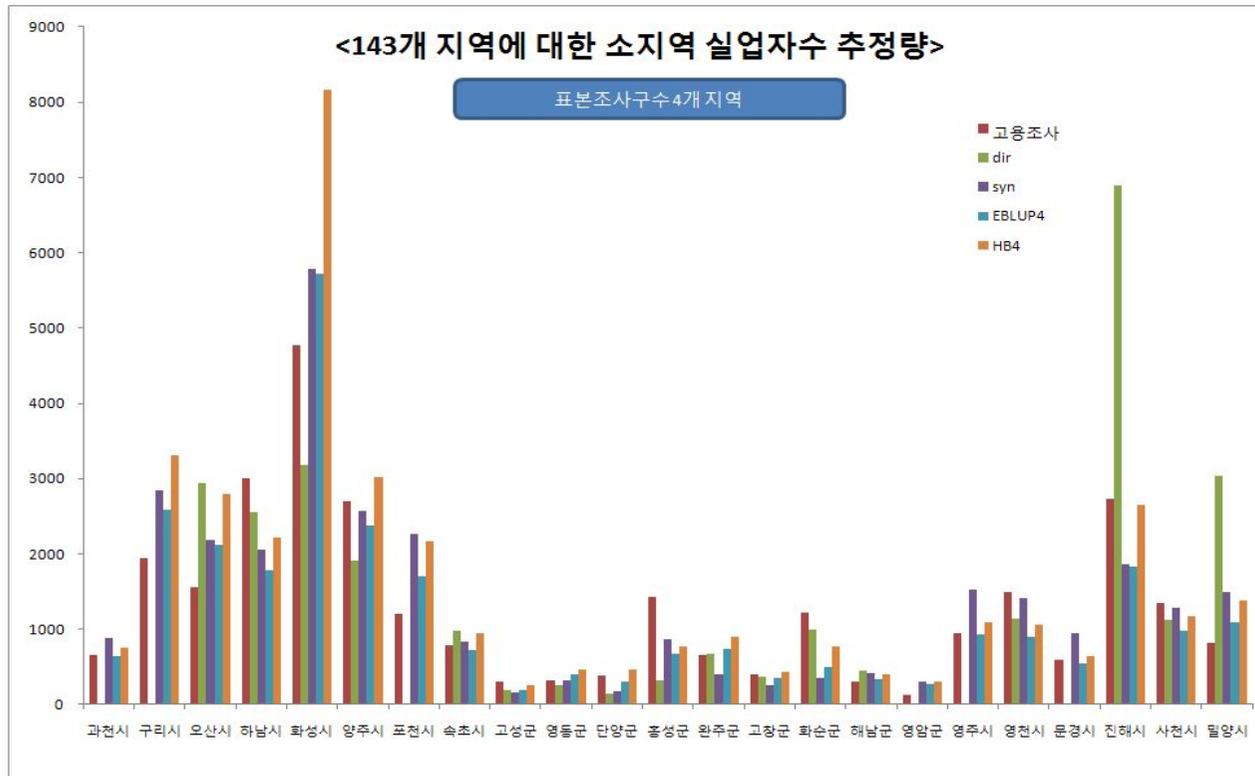
구분	syn		EBLUP1		EBLUP2		EBLUP3		EBLUP4		HB1		HB2		HB3		HB4	
	CI 갯수	CV 25% 이하																
거의 유사 (완전겹침)	62 (63.27)	57 (67.06)	75 (52.45)	41 (47.67)	79 (55.24)	74 (58.73)	67 (46.85)	52 (48.60)	91 (63.64)	90 (66.18)	76 (53.15)	36 (43.37)	37 (25.87)	15 (17.44)	60 (41.96)	44 (45.83)	50 (34.97)	37 (36.63)
유사 (80%이상겹침)	10 (10.20)	9 (10.59)	17 (11.89)	13 (15.12)	19 (13.29)	15 (11.90)	21 (14.69)	15 (14.02)	11 (7.69)	8 (5.88)	14 (9.79)	9 (10.84)	28 (19.58)	24 (27.91)	20 (13.99)	15 (15.63)	23 (16.08)	11 (10.89)
약간유사 (50%이상겹침)	11 (11.22)	8 (9.41)	32 (22.38)	14 (16.28)	22 (15.38)	15 (11.90)	29 (20.28)	18 (16.82)	19 (13.29)	16 (11.76)	29 (20.28)	19 (22.89)	39 (27.27)	23 (26.74)	30 (20.98)	14 (14.58)	20 (13.99)	9 (8.91)
유사하지않음	15 (15.31)	11 (12.94)	19 (13.29)	18 (20.93)	23 (16.08)	22 (17.46)	26 (18.18)	22 (20.56)	22 (15.38)	22 (16.18)	24 (16.78)	19 (22.89)	39 (27.27)	24 (27.91)	33 (23.08)	23 (23.96)	50 (34.97)	44 (43.56)
총계	98	85	143	86	143	126	143	107	143	136	143	83	143	86	143	96	143	101



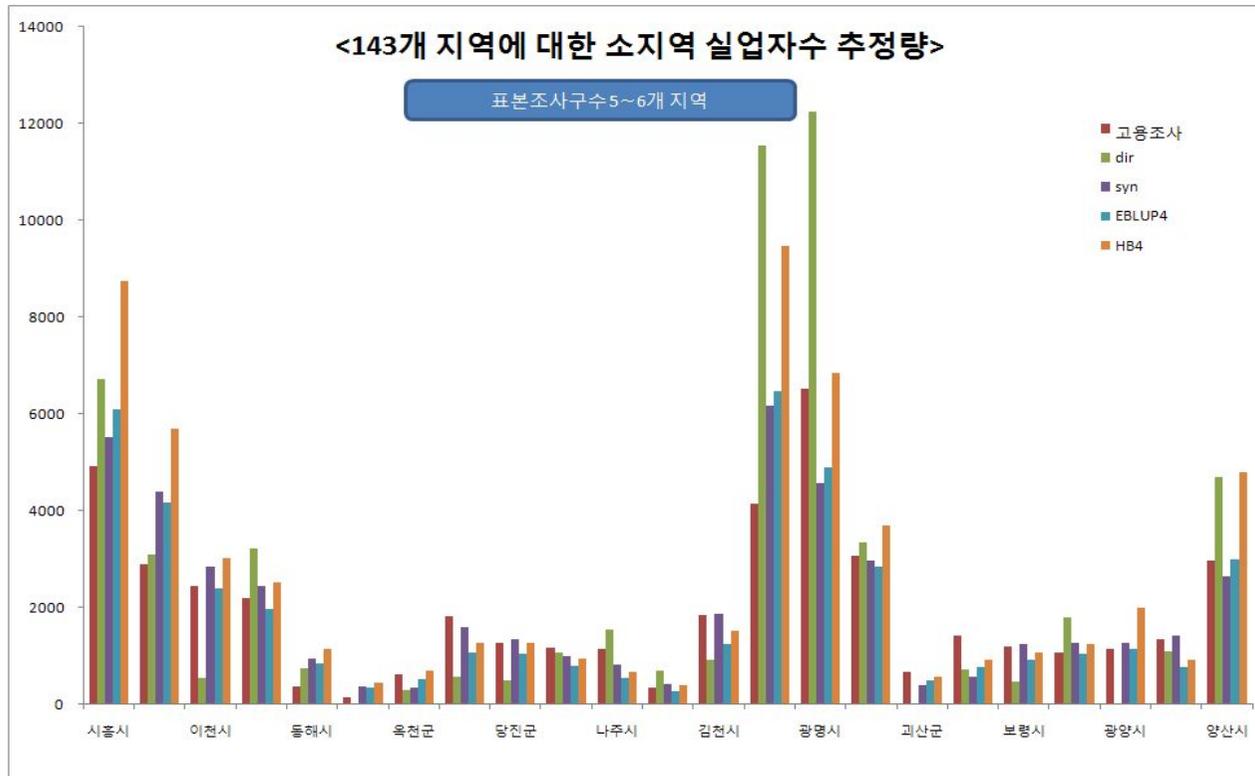
<그림 6> 143개 지역에 대한 소지역 실업자수 추정량 : 표본조사구수 2개



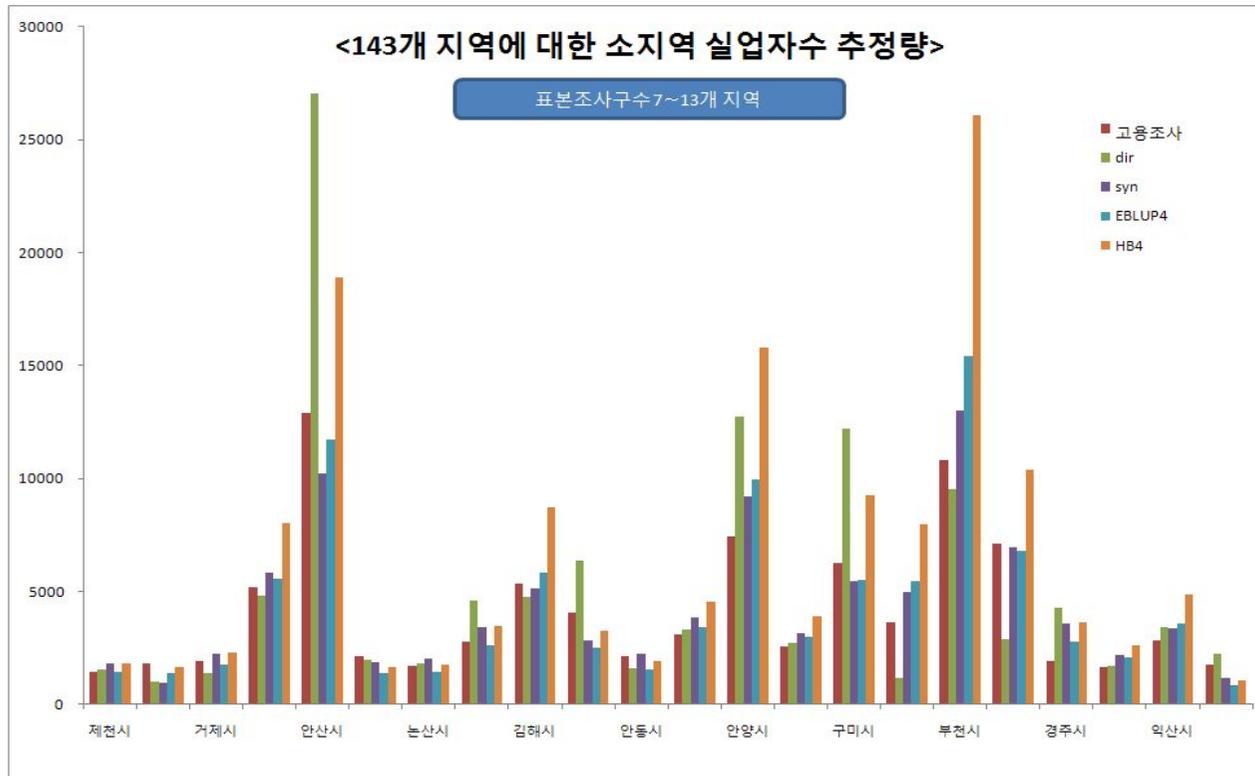
<그림 7> 143개 지역에 대한 소지역 실업자수 추정량 : 표본조사구수 3개



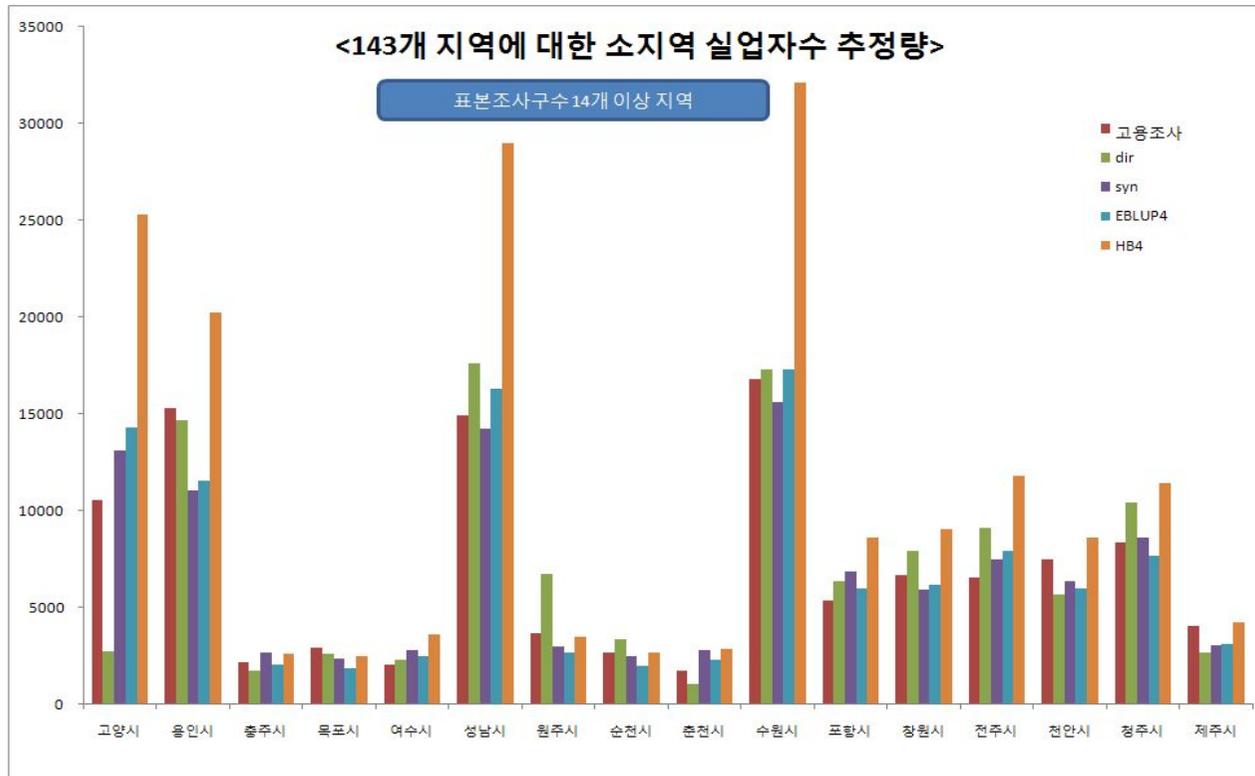
<그림 8> 143개 지역에 대한 소지역 실업자수 추정량 : 표본조사구수 4개



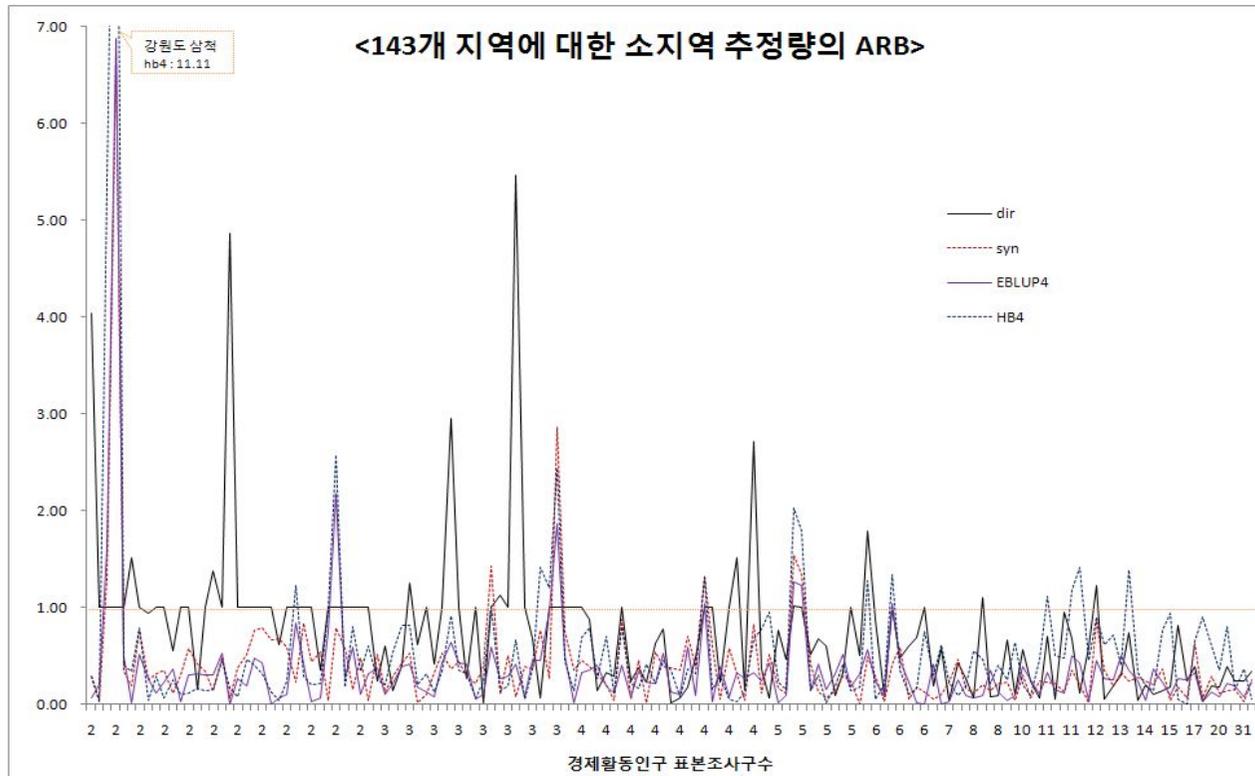
<그림 9> 143개 지역에 대한 소지역 실업자수 추정량 : 표본조사구수 5~6개



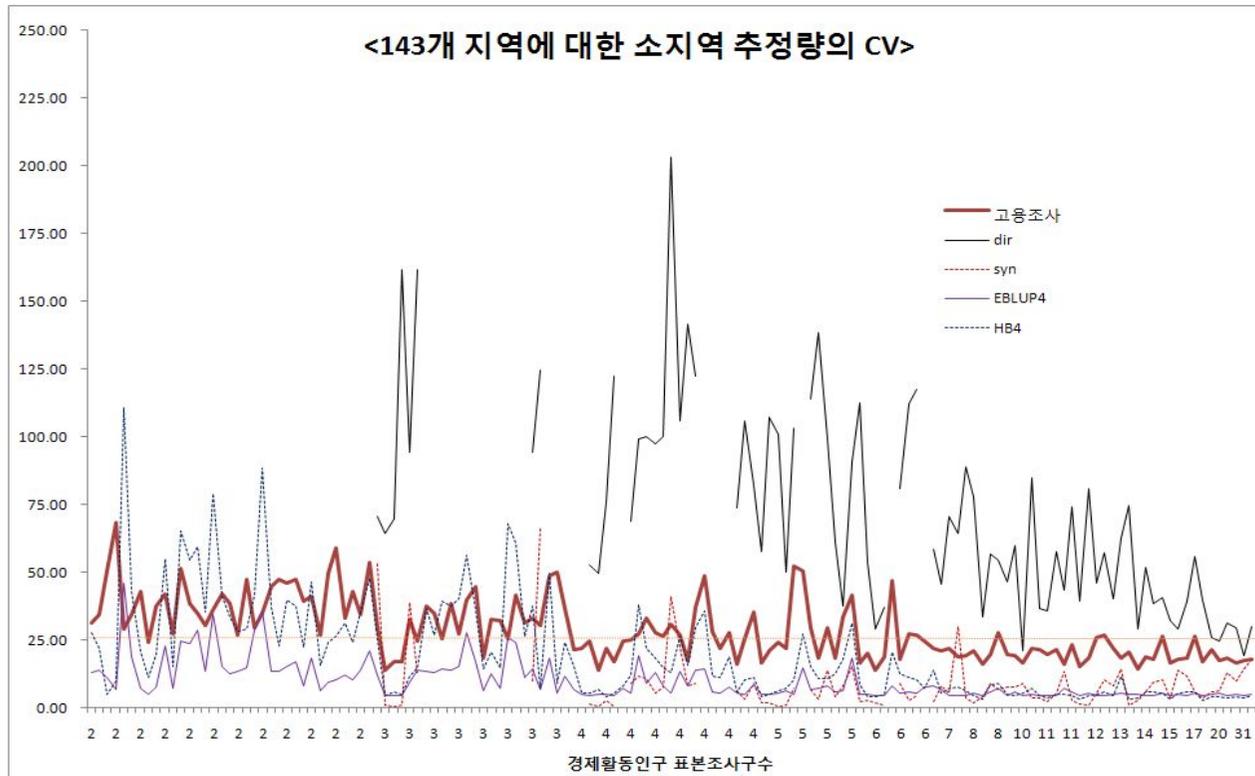
<그림 10> 143개 지역에 대한 소지역 실업자수 추정량 : 표본조사구수 7~13개



<그림 11> 143개 지역에 대한 소지역 실업자수 추정량 : 표본조사구수 14개 이상



<그림 12> 143개 지역에 대한 소지역 추정량의 ARB



<그림 13> 143개 지역에 대한 소지역 추정량의 CV

〈표 9〉 143개 지역에 대한 소지역 추정량의 \overline{ARB}

도	\overline{ARB}					
	직접	합성	EBLUP2	EBLUP4	HB2	HB4
경기	0.69	0.25	0.25	0.23	0.31	0.59
강원	0.89	0.89	0.95	0.88	1.46	1.48
충북	0.44	0.33	0.23	0.24	0.35	0.33
충남	0.69	0.34	0.33	0.31	0.32	0.30
전북	0.84	0.29	0.23	0.22	0.45	0.31
전남	0.96	0.37	0.40	0.34	0.63	0.37
경북	0.74	0.45	0.36	0.31	0.57	0.45
경남	0.79	0.47	0.44	0.44	0.74	0.67
제주	0.32	0.29	0.25	0.37	0.25	0.23
평균	0.75	0.42	0.39	0.37	0.59	0.59

〈표 10〉 143개 지역에 대한 소지역 추정량의 \overline{CV}

도	\overline{CV}						
	고용통계	직접	합성	EBLUP2	EBLUP4	HB2	HB4
경기	20.39	69.31	4.96	7.92	6.04	18.69	6.97
강원	35.97	84.79	16.10	16.36	14.05	32.85	29.99
충북	26.83	76.74	13.04	11.84	7.61	24.05	14.87
충남	28.04	84.35	12.24	11.47	8.10	27.15	15.44
전북	29.65	88.48	14.77	18.40	13.49	35.92	30.35
전남	31.55	74.81	11.22	17.33	12.45	30.73	27.26
경북	31.93	72.50	14.75	15.06	11.52	32.65	23.27
경남	32.77	69.32	13.01	12.09	9.09	27.55	18.36
제주	18.10	46.35	16.17	6.91	5.12	18.23	8.15
평균	28.88	75.48	11.35	13.15	9.90	27.75	19.45

※ 직접, 합성추정량의 지역수는 98개, 나머지 추정량은 143개임

나. 직접추정량이 0인 지역을 제외한 추정 결과(98개 지역)

<그림 12>를 보면 ARB가 1인 지역이 다수 보인다. 이는 표본조사로는 실업자수를 관측하지 못한 경우이다. 이처럼 소지역에서 한 명의 실업자도 관측되지 않은 경우에는 실업자수의 직접추정량이 0이 된다. 또한 이 값들은 표본조사구수가 작은 지역일수록 자주 나타난다. 이렇게 되면 표본조사 자료를 추정에 이용하는 모형추정량 계산시 해당 지역에 대한 관측값을 이용할 수 없기 때문에 추정에 어려움이 있다. 그래서 앞서 143개 지역에 대한 추정시에는 직접추정량이 0인 지역에 대해서는 합성추정값을, MSE는 같은 표본조사구수를 갖는 시군들 중에서 MSE를 선택하여 대체하였다. 그러나 이는 대체가 원 자료를 훼손할 수 있는 가능성도 있기 때문에 이번에는 이와 같은 지역은 제외하고 98개 지역을 대상으로만 소지역 추정을 시도해보았다.

<표 11>을 보면 ARB 측면에서 보면 전체적으로는 합성추정량, EBLUP2, EBLUP4가 가장 나은 것으로 나타났다. 이는 앞서 강원도 삼척에서 특히 높은 ARB가 나와서 전체적으로 모형추정량의 ARB가 높게 나타났던 부분이 해소된 결과이다. 삼척의 경우 직접추정량이 0으로 이번 추정에서는 제외되었다. CV 측면에서도 <표 12>를 보면 합성추정량, EBLUP2, EBLUP4가 좋은 성능을 보여준다.

현재의 경제활동인구조사의 표본시스템으로는 이렇게 직접추정량이 0인 지역이 계속적으로 발생할 것이다. 이럴 때마다 해당 지역을 제외하고 추정한다는 것은 소지역 추정에 대한 기본적인 요구에 부합되지 않고, 추정을 위한 자료 규모에도 영향을 미치게 된다. 그러나, 143개 지역 모두를 추정한 경우에 비해서는 평균 ARB, 평균 CV 모두 뚜렷하게 개선된 것으로 나타났다. 이는 당연한 결과이다. 즉, 표본이 너무 작은 지역, 특성치 관측이 어려운 지역들은 소지역 추정으로도 정확하고 정도높은 추정에는 어느 정도의 어려움이 존재한다는 것이다. 그러므로 소지역 추정을 위해서도 일정 규모 이상의 표본을 확보하거나 특성치를 관측할 수 있는 환경이 필요하다. 따라서 이같은 경우에는 해당 지역의 표본수를 일정 규모 이상으로 늘려주거나 비슷한 지역의 집락화를 통한

규모의 확대로 추정을 용이하게 할 수 있는 등의 추정 방법에 대한 연구가 좀 더 필요할 것으로 보인다.

<표 11> 직접추정량이 0인 지역을 제외한 98개 지역에 대한 소지역 추정량의 \overline{ARB}

도	\overline{ARB}					
	직접	합성	EBLUP2	EBLUP4	HB2	HB4
경기	0.64	0.22	0.25	0.22	0.72	0.64
강원	0.81	0.44	0.73	0.50	1.48	1.21
충북	0.31	0.33	0.21	0.19	0.60	0.60
충남	0.61	0.22	0.26	0.22	0.18	0.25
전북	0.77	0.22	0.32	0.30	1.22	0.85
전남	0.94	0.29	0.32	0.32	1.25	0.67
경북	0.49	0.32	0.30	0.26	0.78	0.54
경남	0.63	0.30	0.19	0.18	0.78	0.56
제주	0.32	0.29	0.22	0.33	0.33	0.14
평균	0.65	0.28	0.31	0.27	0.84	0.65

<표 12> 직접추정량이 0인 지역을 제외한 98개 지역에 대한 소지역 추정량의 \overline{CV}

도	\overline{CV}						
	고용통계	직접	합성	EBLUP2	EBLUP4	HB2	HB4
경기	19.60	69.31	4.96	8.59	7.31	11.43	8.63
강원	32.33	84.79	16.10	14.78	11.97	21.05	20.00
충북	22.62	76.74	13.04	13.35	7.84	16.37	12.61
충남	23.77	84.35	12.24	10.83	7.63	17.33	12.60
전북	26.82	88.48	14.77	15.64	10.06	23.51	18.56
전남	30.02	74.81	11.22	17.00	12.71	25.05	22.77
경북	24.40	72.50	14.75	10.65	8.24	12.83	10.71
경남	23.17	69.32	13.01	11.24	8.81	14.06	11.56
제주	18.10	46.35	16.17	7.78	6.34	15.23	11.13
평균	24.26	75.48	11.35	11.96	8.94	16.61	13.61

다. 군지역 집락 추정

대체를 통한 소지역 추정, 실업자수가 관측되지 않은 지역의 제외를 통한 소지역 추정을 모두 시도해보았다. 각 방법에는 장단점이 존재하

는데, 이번에는 대체와 집락을 통해 앞의 방법들에 대한 단점을 보완해 보고자 한다. 즉, 조사구수가 작고, 직접추정량이 0으로 자주 관측되는 ‘군’ 단위 지역들을 하나의 지역처럼 집락화한 것이다. 그리고 “시” 단위 지역들(과천시, 구리시, 포천시, 태백시 등)에도 실업자수가 관측되지 않은 경우가 발생하는데, 이를 위해서는 앞 절에서 사용한 대체 방법을 이용하여 대체 후 소지역 추정을 시도하였다. 이를 기반으로 소지역 추정량을 계산한 후 다음과 같이 각 ‘군’ 지역들의 추계 인구에 비례하여 추정량을 배분하는 방법을 이용하였다.

하부 ‘군’ 지역의 실업자수

$$= \frac{\text{하부 ‘군’ 지역의 추계인구}}{\text{‘군’ 집락 소지역의 추계인구}} \times \text{‘군’ 집락 소지역의 추정 실업자수}$$

이 방법의 경우, 표본이 없거나 조사구수가 1개인 지역도 추정할 수 있다는 장점이 있다. ‘군’ 지역을 하나의 소지역으로 집락하면 표본 크기는 최소 경기도 10개에서 최대 충청북도 38개까지 확보된다. 이처럼 ‘군’ 지역을 집락한 후의 83개의 소지역에 대해 추정을 완료하고, 그 결과를 하부 ‘군’ 지역에 배분하여 143개 지역에 대한 결과를 도출해보았다. 그러나 이 경우에는 집락 후 추정된 결과를 이용한 하부 ‘군’ 지역들에 대해서는 오차를 계산 결과를 내놓을 수 없다는 단점이 있다.

<표 13>을 보면 ARB 측면에서 보면 전체적으로는 합성추정량, EBLUP2, EBLUP4가 가장 나은 것으로 나타났다. CV 측면에서도 <표 14>을 보면 합성추정량, EBLUP4가 좋은 성능을 보여준다. 이는 ‘군’ 지역 집락화를 통해 소지역의 추정 안정성을 높여준 때문이다. 그러므로 특별히 추정에 영향을 받는 지역들이 없어지고 대체로 안정적인 추정 결과를 보여준다.

이와 같은 결과는 143개 지역을 모두 추정한 결과에 비해선 CV 측면에서 상당한 개선이 있었다고 평가된다. 그리고 직접추정량이 0인 지역을 제외한 98개 지역에 대한 추정 결과가 제시하지 못한 군 단위 지역에 대한 추정 결과도 제시할 수 있었다. 물론 군 단위 지역에 대한 추정 결과의 오차를 제공하지 못한다는 단점은 존재한다. 그러나 너무 작은 소지역의 추정을 위해선 고려할 수 있는 수준의 대안이라고 보인다. 실제로 미국에서도 county와 같은 아주 작은 지역에 대해서는 모형추정

법을 적용하는 대신에, 대안으로 적절한 부차 정보를 활용한 할당 또는 분해 기법을 사용하는 것으로 보인다(김서영과 권순필, 2009).

<표 13> 군지역 집락 후 소지역 추정량의 \overline{ARB}

도	시군 수	\overline{ARB}					
		직접	합성	EBLUP 2	EBLUP 4	HB2	HB4
경기	30	0.48	0.24	0.24	0.22	0.92	0.59
강원	17	0.64	0.90	1.07	1.29	6.57	3.13
충북	11	0.34	0.35	0.36	0.36	0.78	0.50
충남	14	0.41	0.35	0.30	0.31	0.49	0.52
전북	14	0.35	0.30	0.18	0.20	0.62	0.61
전남	15	0.48	0.39	0.38	0.38	0.78	0.75
경북	20	0.68	0.48	0.31	0.39	0.91	0.72
경남	20	0.59	0.48	0.48	0.50	1.29	1.07
제주	2	0.32	0.29	0.17	0.28	0.41	0.11
평균	143	0.51	0.43	0.41	0.45	1.54	0.98

<표 14> 군지역 집락 후 소지역 추정량의 \overline{CV}

도	\overline{CV}						
	고용통계	직접	합성	EBLUP2	EBLUP4	HB2	HB4
경기	18.81	64.14	1.73	9.76	7.58	13.63	19.21
강원	38.15	62.82	9.21	18.22	12.43	18.95	25.97
충북	19.33	47.22	8.55	10.47	6.84	14.72	18.95
충남	20.96	68.73	6.76	12.64	10.11	18.84	26.03
전북	22.37	63.55	7.00	12.86	11.05	19.79	29.12
전남	24.16	38.97	9.86	16.34	12.15	17.71	23.13
경북	22.22	65.51	5.35	14.21	11.60	23.46	31.97
경남	24.83	65.76	5.64	11.81	9.88	15.42	24.02
제주	18.10	46.35	16.17	11.72	10.92	18.62	31.45
평균	22.69	61.96	5.44	12.50	9.77	17.11	24.17

* 고용통계, 직접, 합성추정량의 지역수는 75개, 나머지 추정량은 83개임

제5절 결론

본 절에서는 지금까지의 연구결과를 종합적으로 요약하고, 연구를 통해 드러난 몇 가지 문제점들에 대한 해결방안을 모색하고자 한다. 그리고 소지역 추정 통계를 소지역 통계로 이해하고 인정하기 위한 방법에 대하여 논의한다.

1. 요약

본 연구에서 비교를 위해 사용된 소지역 추정량은 직접, 합성, 복합 추정량과, 모형추정량 EBLUP, HB 등 5개이다. 모형 적용을 위한 소지역은 다음과 같이 3가지 방법으로 정의하였다.

- ① 조사구수 2개 이상의 전국 143개 시군(실업자수 0지역 대체)
- ② 실업자수 0지역을 제외한 98개 시군
- ③ 각 도내 군단위 지역을 하나로 통합하고, 시단위 지역은 그대로 사용한 83개 지역.

모형 추정을 위해 사용한 보조정보는 고용보험 자료, 대지역 실업자 수를 소지역의 인구에 비례하여 배분한 정보이다.

분석결과 고용조사와의 비교에서 전체적으로 EBLUP이 편향이나 상대오차(CV)측면에서 가장 좋았다. 소지역 정의 측면에서는 ② 실업자수가 0인 지역을 제외하고 분석했을 때가 소지역 추정치 차원에서 편향이나 CV 측면에서 가장 좋았다. ③의 군지역을 통합하여 추정할 경우는 표본이 1개이거나 직접추정량이 0으로 나온 지역에 대해서도 추정이 가능하다는 장점이 있었고, 편향이나 CV 측면에서 143개로 추정하는 경우에 비해 결과가 나쁘지 않았다. 그러나 이 경우에는 통합된 군보다 작은 군단위로 추정값을 배분하는 방법에 대한 연구가 필요하고 작은 군단위에 대해서는 CV를 제공할 수 없다는 단점이 있었다.

HB 추정량은 ARB 측면에서 큰 개선 효과가 나타나지 않았는데, 이는 추정의 문제라기보다는 로그 변환된 자료를 다시 지수 변환하는 과정에서 편향이 발생한 것으로 판단된다. 이 부분에 대해서는 향후 추가

연구가 필요하다.

연구 결과를 이해함에 있어 지역별 고용조사 결과나 소지역 추정 결과 모두 추정값으로서 참값이 아님을 주지할 필요가 있다. 그리고 절대적으로 소지역 추정값이나 고용조사 추정값이 동일한 수준의 결과를 준다고 볼 수는 없다. 왜냐하면, 지역별 고용조사도 군 지역에서는 CV가 매우 크고, 소지역 추정값들도 군 지역에서 매우 큰 CV를 갖는다는 단점이 있기 때문이다.

소지역 추정만 놓고 보면, 경제활동인구의 직접추정값이 '0'으로 관측되는 지역들에 대해서 해결방안을 찾는다면 소지역 추정은 훨씬 정도 높은 결과를 줄 수 있을 것으로 보인다. 그에 대한 대안으로 소지역 정의 ①처럼 '0'값에 대한 좋은 대체방안을 강구하거나, ③처럼 군 지역을 통합하여 추정하는 것도 좋은 방법이 될 수 있다. 미국의 경우에도 아주 작은 카운티에 대해서는 CV를 제공하지 않은 채 추정값만 발표하고 있다. 혹은 인구규모가 아주 작은 군 지역에 표본수를 늘려줄 수도 있을 것이다. 그러나 연구 결과 이러한 지역에 무조건적으로 표본수만 늘려주는 것은 현실적인 방법이 될 수 없고, 실제로 추정 결과도 크게 향상되지 않는 것으로 나타났다.

2. 향후 연구 및 논의

본 연구는 많은 표본으로 직접 조사를 수행한 결과와 소지역 추정 결과를 실증적으로 비교해 보았다는 데 그 의의가 있다. 지금까지 많은 연구에도 불구하고 추정 결과에 대한 의구심이 생기는 이유는 과연 모형을 이용한 추정이 전통적인 조사 통계 결과를 대신할 수 있을 것인가, 보조정보가 얼마나 설명력있게 추정 과정에 영향을 미칠 것인가 등에 대한 의문에 적절히 답할 수 없었기 때문이라 할 수 있다. 왜냐하면 적절한 비교 대상이 없었기 때문이다.

이는 그 동안의 연구들이 종합적인 진단과 검토보다는 개괄적인 방법 적용에 치중하였다는 것을 한 이유로 들 수 있다. 또한 소지역 추정치를 비교할만한 적절한 대상의 조사 통계가 없었다. 그래서 대부분의

다른 나라들도 모형 또는 방법의 타당성은 센서스 자료를 기준으로 평가하는 경향이 있다. 이에 대해 김서영과 권순필(2009)은 이미 모의실험을 통해 모형 추정치의 타당성을 검토한 바 있다. 또 하나의 방법으로 비교할 만한 조사통계가 있을 경우, 그 통계와의 비교를 통해 소지역 추정치를 평가할 수 있다. 그러나 이는 조사통계가 충분히 신뢰할 만하다는 전제가 필요하다. 그러므로 많은 나라들은 조사통계와의 비교를 수행하기 어려웠다. 왜냐하면 소지역의 조사를 위해 그렇게 많은 예산을 사용할 여유가 없기 때문이다. 그런 면에서, 대규모의 표본을 이용한 「지역별 고용조사」 결과를 대상으로 소지역 추정 결과를 비교해 보게 된 것은 연구, 연구자, 사용자 모두의 입장에서 혼치 않은 좋은 기회를 얻은 것이라 볼 수 있다.

조사의 목적과 활용성 측면에서 조사통계와 추정통계를 단순 비교하고 효율을 논하는 것은 적절하지 않다. 그러나 조금 더 세부적으로 동일 목적과 활용성을 보았을 때, 즉, 실업자 수 추정과 같은 단순한 목적으로 두 결과를 비교해본다면, 추정통계가 가지는 장점이 조사통계의 그것보다 훨씬 많다는 것을 알 수 있다. 표본의 확대에 따른 비용 부담이 줄어들 수 있으며, 추정의 방법이 쉽고 편리하다. 조사와 공표 사이의 기간, 자료 처리, 비표본오차에 대한 부담도 없다. 또한 행정자료 이용 측면에서 추정 결과에 대한 이해를 구하기도 쉽다. 더불어 추정 통계의 개발과 사용은 이미 세계적인 추세가 되었다. 이는 조사 환경의 변화, 계산 기법의 개발과 개선 등은 전통적인 조사 방법보다 비용효율적인 추정 방법으로서의 변화를 요구하고 있기 때문이다.

현재 우리의 소지역 추정에 있어서 어떤 방법으로 추정을 할 것인가에 대한 문제는 신중한 판단을 요한다. 이러한 판단은 지속적이고 다양한 연구를 통해서만 뒷받침될 수 있다. 그러나 내적 평가와 같은 수치적 연구만으로 이를 완벽히 해결할 수는 없다. 이 추정량을 이용하고자 하는 이용자 입장에서도 충분히 납득할 수 있는 결과를 얻어야 하기 때문이다. 그러므로 학계 전문가, 실무 전문가, 실제 사용자 등과 함께 추정 결과에 대한 평가를 진행해야 할 것이다. 이처럼 기술적, 활용성 측면에서의 평가와 동시에 더 중요한 것은 정책적 판단이 있어야 한다는 것이다. 우리는 때로는 최선의 방법보다는 차선의 방법을 선택하는 것이 현

실적으로 더 실리적이고 보편 타당할 수도 있음을 상기해야 한다. 본 연구 결과가 이러한 평가 과정에서 중요한 근거로 활용될 수 있을 것으로 기대해본다.

참고문헌

- 권혜자, 노현국(2007), “고용보험 실업률의 추이와 특징”, 한국고용정보원.
- 김서영, 권순필(2008), “고용통계 소지역 추정 연구 I”, 통계개발원.
- Costa, A., Satorra, A. and Ventura, E. (2006). Improving small area estimation by combining surveys: new perspectives in regional statistics. Social Science Research Network, Working paper. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1002507.
- Rao, J.N.K (2003). Small area estimation, Wiley.
- Brown, G., Chambers, R., Heady, P. and Heasam, H. (2001). Evaluation of small area estimation methods-an application to the unemployment estimates from the UK LFS. Proceedings of Statistics Canada Symposium Ottawa, October 2001.

< 부 록 >

<부표 1> 소지역 추정 결과 : 경기도

시군	조사구수		고용통계		dir		syn		EBLUP1		EBLUP2		EBLUP3		EBLUP4		HB1		HB2		HB3		HB4	
	고용	경활	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV
수원시	37	18	16829	17.21	17333	39.78	15627	3.46	13581	5.89	13399	6.63	20403	4.05	17322	4.63	15162	4.88	15375	17.06	22869	3.34	32106	2.93
성남시	41	15	14937	16.61	17609	32.29	14250	3.49	13298	5.89	14186	9.10	18443	4.12	16321	4.55	14828	4.88	16732	17.61	20517	3.42	29002	2.97
의정부시	49	6	4142	20.30	11554	53.65	6185	2.91	5774	5.95	5940	8.33	6833	4.77	6473	4.76	6206	5.52	6545	17.18	7250	4.33	9462	3.93
안양시	44	11	7455	19.71	12752	35.64	9208	2.35	8502	5.83	9314	7.81	10857	4.45	9952	4.69	9236	5.10	10820	16.77	11680	3.87	15775	3.48
부천시	43	12	10793	15.14	9554	39.26	13002	1.46	13379	5.88	13867	7.60	16325	4.19	15450	4.58	14812	4.91	16165	16.60	17873	3.54	26082	3.10
광명시	49	6	6516	13.92	12252	29.05	4578	1.99	4816	6.10	5291	9.73	4672	5.09	4895	4.70	5174	5.87	6097	17.75	4956	4.86	6858	4.10
평택시	48	8	5198	20.84	4812	77.61	5813	1.91	4683	6.13	4876	5.41	6260	4.83	5543	5.28	4975	5.98	5426	15.56	6567	4.48	8058	4.49
안산시	47	8	12891	16.12	27065	33.45	10228	4.19	10478	5.85	10660	6.29	12107	4.38	11738	4.64	11582	4.95	12182	16.14	13201	3.74	18917	3.30
고양시	41	14	10580	20.70	2773	74.75	13141	1.02	11133	5.82	10467	5.88	17130	4.15	14312	4.85	11998	5.01	11400	16.87	18494	3.55	25303	3.26
과천시	37	4	654	21.65	0		884		649	23.31	545	15.50	649	14.42	646	6.56	761	24.53	693	29.34	758	17.59	751	15.10
구리시	46	4	1947	21.77	0		2844		2448	7.86	2747	9.62	2628	5.86	2580	4.94	2618	8.68	3124	16.82	2778	6.34	3306	5.50

- ※ 1. 음영 셀은 CV가 25% 이상인 지역임
 2. 이탤릭체는 직접추정량이 0인 지역임
 3. 고용통계는 지역별 고용조사, dir은 직접추정량, syn은 합성추정량을 의미하며,
 EBLUP1, HB1은 보조정보로 고용보험 실업급여 등록자수,
 EBLUP2, HB2는 보조정보로 고용보험 실업급여 등록자수를 4개 범주로 나눈 정보,
 EBLUP3, HB3은 대지역 실업자수 정보,
 EBLUP4, HB4는 고용보험 실업급여 등록자수와 대지역 실업자수 정보를 사용한 모형 추정량값임.

〈부표 1〉 소지역 추정 결과 : 경기도(계속)

시군	조사구수		고용통계		dir		syn		EBLUP1		EBLUP2		EBLUP3		EBLUP4		HB1		HB2		HB3		HB4	
	고용	경활	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV
남양주시	44	12	7101	18.23	2899	80.64	6973	1.13	5575	5.95	5621	7.67	7948	4.65	6826	5.34	5890	5.63	6035	17.11	8310	4.23	10364	4.30
오산시	41	4	1565	24.44	2946	52.74	2186	1.18	2329	8.12	2730	8.29	1835	6.80	2116	4.62	2497	9.04	3223	15.86	1978	7.98	2801	5.51
시흥시	50	5	4917	16.59	6713	57.86	5527	1.76	6031	5.92	6184	7.02	5743	4.90	6101	4.81	6446	5.48	6957	15.61	6045	4.57	8754	3.96
군포시	42	3	4602	13.97	1826	64.54	4011	0.85	4043	6.27	4182	7.12	3964	5.24	4118	4.66	4241	6.29	4530	16.55	4112	5.22	5526	4.36
의왕시	42	3	1537	17.20	1756	69.70	1981	0.54	1987	9.02	2096	5.64	1687	7.09	1866	4.47	2132	10.19	2343	15.57	1816	8.48	2379	6.00
하남시	41	4	3008	13.87	2562	49.58	2053	0.66	1766	9.88	2302	9.42	1759	6.94	1786	4.82	1918	11.30	3001	14.02	1897	8.22	2220	6.58
용인시	41	14	15292	14.12	14666	28.92	11085	2.68	8982	5.83	8250	6.49	14041	4.28	11590	5.13	9794	5.06	9064	16.97	15317	3.64	20255	3.53
파주시	40	5	2910	21.09	3100	107.27	4406	1.89	3682	6.53	3656	4.75	4525	5.12	4169	5.17	3899	6.71	3914	15.31	4723	4.98	5707	4.82
이천시	40	5	2449	24.19	558	100.89	2854	0.39	2169	8.17	2405	5.25	2605	5.78	2406	5.58	2276	9.28	2812	13.78	2688	6.59	3022	6.32
안성시	45	5	2188	21.97	3223	50.26	2452	0.90	1758	9.95	1925	6.25	2205	6.26	1984	6.09	1915	11.39	2253	14.65	2354	7.06	2522	7.07
김포시	44	6	3085	18.73	3353	37.16	2974	0.86	2767	7.31	2697	5.06	2827	5.72	2854	4.74	2951	7.93	2901	15.50	2986	6.08	3703	5.08
화성시	45	4	4773	22.06	3174	75.89	5783	2.58	5005	6.05	5209	5.32	6198	4.84	5720	4.95	5290	5.83	5726	15.64	6465	4.51	8165	4.25
광주시	40	3	2342	17.17	1381	161.78	3338	0.91	3181	6.89	3523	5.83	3204	5.53	3267	4.68	3343	7.34	3897	14.98	3332	5.79	4260	4.81
양주시	40	4	2699	16.93	1912	122.47	2567	0.59	2343	8.12	2440	4.73	2338	6.12	2383	4.72	2498	9.06	2687	15.01	2473	6.85	3026	5.55
포천시	41	4	1205	24.59	0		2263		1429	9.92	1533	9.27	2007	6.00	1696	7.23	1572	14.08	1833	17.40	2143	6.79	2174	8.10
여주군	37	3	724	32.51	1628	94.28	1107	38.96	823	19.87	908	5.92	1306	8.30	1025	9.39	946	22.27	1120	19.13	1432	10.36	1318	11.83
연천군	33	2	337	31.30	1699	70.71	438	38.22	308	53.09	393	12.16	449	21.24	362	12.84	405	52.48	622	31.04	555	25.88	435	27.94
가평군	33	2	540	34.24	0		518		329	47.77	246	20.96	597	15.81	429	13.85	421	46.21	371	49.49	699	19.43	559	21.97
양평군	37	3	832	24.56	319	161.78	818	12.18	482	32.50	489	8.53	1031	9.73	685	13.82	579	33.22	598	29.40	1127	12.40	1000	14.88

〈부표 2〉 소지역 추정 결과 : 강원도

시군	조사구수		고용통계		dir		syn		EBLUP1		EBLUP2		EBLUP3		EBLUP4		HB1		HB2		HB3		HB4	
	고용	경활	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV
춘천시	35	17	1718	26.27	1033	55.67	2802	6.02	2133	8.28	2411	5.28	2372	5.98	2281	5.19	2261	9.11	2735	13.75	2481	6.67	2853	6.01
원주시	35	15	3691	17.97	6727	29.26	3018	13.73	2614	7.60	2771	5.40	2679	5.83	2698	4.76	2819	8.37	3139	14.54	2861	6.27	3507	5.22
강릉시	37	13	1632	26.67	1715	57.13	2209	10.26	2171	8.43	2268	5.08	1888	6.68	2065	4.50	2318	9.38	2492	14.48	2016	7.78	2637	5.66
동해시	35	5	375	52.37	757	103.28	951	6.99	1032	15.70	879	10.41	682	14.06	850	5.02	1149	17.65	994	23.08	791	17.77	1138	10.40
태백시	37	2	247	50.34	0		576		1086	7.22	551	15.15	364	17.99	648	11.03	1192	7.14	1329	7.24	458	17.52	1590	4.85
속초시	36	4	787	25.16	982	68.76	831	9.06	871	18.37	958	8.04	583	16.30	719	5.32	989	20.35	1157	18.17	691	20.30	955	12.22
삼척시	37	2	87	68.27	0		667		921	14.80	695	8.39	489	18.23	682	6.66	1033	14.84	855	19.01	590	20.54	1048	10.07
홍천군	40	5	157	50.32	0		369		276	57.61	355	14.13	475	19.84	351	14.61	363	54.58	523	34.19	572	24.09	441	27.10
황성군	32	3	244	37.48	0		219		299	55.09	303	14.33	261	36.96	275	13.36	385	55.08	392	49.17	343	42.17	321	36.76
영월군	32	3	299	34.75	173	99.08	202	11.67	324	36.68	266	14.02	241	35.16	276	13.19	408	31.52	373	31.82	318	33.19	339	26.69
평창군	32	3	459	25.36	0		213		274	59.83	284	14.59	258	37.37	261	14.31	358	58.53	379	49.84	339	42.36	299	39.27
정선군	32	3	162	38.50	641	132.50	222	16.92	285	56.87	333	14.86	258	37.28	266	13.89	375	55.05	474	39.27	344	41.53	312	37.48
철원군	33	3	449	27.28	0		251		244	66.40	186	26.71	280	34.23	255	15.29	326	62.69	320	58.82	363	38.94	288	40.44
양구군	33	2	177	29.23	0		116		85	188.33	120	39.43	106	92.38	90	46.04	140	140.35	242	73.66	165	85.17	105	110.81
인제군	32	3	249	39.71	182	103.04	176	8.08	130	95.87	122	46.61	178	48.81	146	27.64	196	69.70	280	46.01	250	44.07	170	56.10
고성군	31	4	302	27.24	188	99.20	164	11.64	237	52.28	265	15.79	173	50.42	199	19.29	316	42.74	362	34.28	244	45.19	249	37.87
양양군	33	2	202	34.63	508	100.00	163	66.60	253	62.15	336	19.92	171	56.15	206	18.70	339	57.63	580	31.06	246	56.94	271	42.95

〈부표 3〉 소지역 추정 결과 : 충청북도

시군	조사구수		고용통계		dir		syn		EBLUP1		EBLUP2		EBLUP3		EBLUP4		HB1		HB2		HB3		HB4	
	고용	경활	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV
청주시	24	31	8388	17.28	10404	19.19	8597	14.47	7038	5.83	6167	11.40	7848	4.66	7697	4.67	7584	5.26	6908	19.73	8337	4.17	11444	3.68
충주시	36	14	2165	18.79	1718	52.00	2695	5.66	2055	8.73	2076	8.06	2023	6.46	2072	4.72	2200	9.76	2433	15.62	2152	7.40	2596	5.92
제천시	43	7	1469	21.91	1539	70.47	1815	5.54	1570	10.84	1519	5.04	1254	8.52	1426	4.42	1706	12.39	1649	16.75	1378	10.65	1815	7.12
청원군	43	7	1838	18.98	1040	64.35	978	29.84	1412	11.66	1454	7.47	1317	8.18	1378	4.70	1537	13.11	1689	16.63	1433	10.06	1677	7.63
보은군	32	3	241	44.74	0		188	0	228	54.54	262	16.26	239	36.16	228	16.84	306	44.28	360	34.59	317	34.59	257	36.42
옥천군	30	5	615	29.57	296	113.92	355	6.57	655	22.53	693	5.84	423	21.77	529	6.71	754	23.07	809	20.32	515	24.98	702	15.29
영동군	31	4	328	33.00	254	100.00	321	10.74	428	32.29	314	11.61	385	23.28	402	8.96	520	30.17	449	31.88	475	25.49	464	21.77
진천군	37	3	649	18.14	636	92.16	396	30.10	735	21.52	618	18.91	452	20.87	581	6.39	845	23.24	866	26.00	552	25.12	803	14.35
괴산군	30	6	682	46.85	0		400	0	502	32.71	498	7.78	524	18.26	507	7.99	603	35.24	593	32.75	624	23.04	580	20.45
음성군	39	6	1414	17.99	736	80.99	578	8.85	812	19.52	706	14.91	732	13.12	772	5.40	925	21.30	902	24.66	840	16.45	916	12.59
단양군	31	4	386	27.87	142	97.55	176	5.60	419	25.31	300	22.99	213	37.55	300	12.96	505	22.74	437	25.61	287	33.39	468	18.39

〈부표 4〉 소지역 추정 결과 : 충청남도

시군	조사구수		고용통계		dir		syn		EBLUP1		EBLUP2		EBLUP3		EBLUP4		HB1		HB2		HB3		HB4	
	고용	경활	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV
천안시	27	29	7501	16.46	5668	29.55	6376	9.86	5437	5.95	5273	9.86	6292	4.82	6024	4.76	5789	5.61	5872	18.54	6610	4.44	8623	4.01
공주시	44	5	1831	18.30	583	138.56	1587	3.30	956	16.92	926	8.91	1198	8.77	1067	7.02	1068	18.99	1042	21.60	1304	11.05	1269	10.75
보령시	35	6	1191	27.06	466	112.05	1256	2.60	904	17.39	883	6.07	930	10.58	917	5.75	1011	19.00	992	19.66	1033	13.35	1069	11.16
아산시	41	10	4084	16.63	6398	20.88	2838	9.14	2543	7.68	2449	11.14	2364	6.08	2504	4.58	2744	8.75	2927	18.62	2538	6.84	3261	5.25
서산시	41	8	2160	19.73	1981	56.90	1858	9.04	1300	12.80	1379	5.52	1454	7.73	1382	5.71	1437	14.71	1548	16.73	1584	9.47	1678	8.41
논산시	43	8	1696	27.58	1844	54.43	2059	6.59	1249	13.25	1209	5.06	1655	7.15	1438	7.00	1383	15.20	1339	18.18	1785	8.56	1789	8.88
금산군	38	2	318	42.83	0		564		522	31.16	551	7.29	449	21.21	482	7.39	626	33.32	675	28.15	548	26.01	570	20.39
연기군	38	2	1086	24.04	2113	137.67	813	34.42	943	17.47	987	11.76	668	14.45	801	4.94	1073	19.77	1321	18.67	785	18.29	1046	11.38
부여군	37	3	304	32.49	0		740		381	42.36	374	14.85	647	14.79	483	12.53	479	42.87	528	35.75	754	18.63	613	20.69
서천군	38	2	442	37.58	0		584		511	30.59	392	17.90	490	19.18	496	7.64	614	31.47	545	35.28	590	23.12	574	19.69
청양군	33	2	235	42.03	0		319		151	109.78	163	26.25	235	41.19	181	22.81	222	96.15	240	79.88	316	45.91	218	54.80
홍성군	36	4	1426	26.32	316	100.06	869	7.72	611	23.84	631	13.87	750	12.43	670	8.00	710	23.97	915	19.13	847	15.42	771	14.85
예산군	37	3	969	32.22	2064	94.28	870	25.99	660	24.78	521	14.43	762	12.79	703	7.32	780	27.15	657	33.00	883	16.30	826	14.98
당진군	45	5	1262	29.38	503	99.10	1353	13.72	906	17.28	994	7.65	1254	8.39	1059	7.95	1014	18.78	1215	16.95	1355	10.47	1282	10.97

〈부표 5〉 소지역 추정 결과 : 전라북도

시군	조사구수		고용통계		dir		syn		EBLUP1		EBLUP2		EBLUP3		EBLUP4		HB1		HB2		HB3		HB4	
	고용	경활	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV
전주시	34	21	6534	18.26	9126	31.37	7481	13.05	7189	5.84	6874	6.73	8135	4.64	7920	4.66	7738	5.25	7446	17.03	8634	4.15	11827	3.67
군산시	40	11	2569	21.53	2708	57.66	3137	5.28	2931	7.13	2685	6.91	2900	5.68	2981	4.66	3113	7.65	2904	16.40	3051	6.02	3882	4.92
익산시	35	13	2850	21.96	3413	40.40	3386	8.30	3769	6.43	3316	9.32	3218	5.51	3589	4.79	3991	6.54	3643	17.73	3384	5.70	4882	4.50
정읍시	44	6	1067	26.89	1802	117.47	1263	4.44	1039	15.90	1096	13.54	1038	9.84	1042	5.34	1169	18.13	1470	18.31	1162	12.53	1246	10.10
남원시	37	2	1085	27.19	1685	73.35	959	6.85	653	24.90	751	5.82	731	13.26	686	7.15	772	27.31	951	21.41	849	16.88	802	15.24
김제시	39	5	1181	18.42	1078	61.05	1005	4.08	814	19.60	940	12.07	795	12.20	804	5.71	932	21.64	1454	15.89	909	15.42	947	12.51
완주군	41	4	657	30.77	672	203.10	408	41.05	807	20.21	827	5.27	674	14.31	740	5.23	920	22.70	951	21.43	781	18.27	893	13.13
진안군	32	2	166	51.32	0		115		163	101.83	150	29.97	169	57.94	161	24.53	233	91.61	271	70.70	239	61.09	183	65.36
무주군	33	3	216	25.52	0		107		158	104.02	145	31.20	157	62.46	153	25.94	227	92.29	281	67.83	224	64.51	175	68.04
장수군	33	2	240	38.59	0		100		202	78.48	172	26.10	144	67.44	168	23.56	276	70.81	272	64.85	209	66.81	212	54.44
임실군	31	3	117	41.71	757	94.28	126	15.76	151	107.12	188	26.35	198	48.82	167	24.04	225	91.80	416	45.15	278	51.57	195	60.83
순창군	33	2	209	34.79	242	100.00	121	26.51	121	113.90	125	35.57	189	48.13	144	28.42	185	84.24	216	66.31	263	45.93	174	59.35
고창군	35	4	396	26.71	367	106.07	255	22.41	292	52.08	229	30.83	458	20.28	355	13.63	380	48.19	544	32.08	553	23.79	429	26.50
부안군	37	3	423	31.39	0		256		351	45.73	424	17.95	451	20.99	390	11.24	441	45.70	683	27.82	544	25.65	445	26.31

〈부표 6〉 소지역 추정 결과 : 전라남도

시군	조사구수		고용통계		dir		syn		EBLUP1		EBLUP2		EBLUP3		EBLUP4		HB1		HB2		HB3		HB4	
	고용	경활	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV
목포시	35	14	2940	18.03	2624	38.48	2343	9.37	2124	8.56	2336	6.33	1573	7.36	1871	4.71	2284	9.63	2638	14.62	1711	8.92	2507	5.86
여수시	36	14	2042	26.24	2331	40.49	2805	10.29	2981	6.99	2776	11.47	2007	6.49	2524	5.47	3157	7.40	4040	13.22	2145	7.45	3606	5.22
순천시	35	15	2700	18.32	3357	39.23	2508	11.80	2228	8.34	2067	5.37	1739	6.98	2013	4.61	2397	9.37	2357	14.83	1884	8.30	2668	5.65
나주시	36	5	1161	33.59	1548	37.67	817	7.97	605	26.14	651	12.97	508	18.60	553	6.47	721	29.06	966	22.32	617	23.42	668	17.96
광양시	44	6	1136	24.65	0		1278		1619	10.17	1400	7.18	757	12.67	1140	7.43	1750	11.22	1651	15.87	873	16.02	1995	7.24
담양군	33	2	387	30.49	0		255		315	51.95	460	14.37	234	41.22	269	13.63	403	51.95	771	25.22	314	45.80	332	35.39
구례군	34	2	169	36.40	401	70.71	145	6.06	130	110.49	125	40.35	114	82.33	118	34.64	198	85.88	377	41.97	178	73.96	143	79.16
고흥군	36	5	342	41.37	687	90.87	423	15.13	202	79.19	182	30.12	416	22.70	278	18.40	284	71.57	366	52.02	515	27.11	389	31.11
보성군	33	2	522	42.15	0		264		262	62.45	182	37.66	241	40.05	247	15.16	346	60.47	631	31.20	322	44.70	285	41.21
화순군	36	4	1219	17.56	994	141.78	359	8.10	673	24.22	639	8.04	358	26.74	496	7.96	787	26.65	840	23.52	454	31.60	764	15.75
해남군	36	4	303	37.19	450	122.47	424	8.79	276	57.36	330	15.72	423	22.28	332	14.01	364	54.16	488	36.53	519	26.51	398	29.40
영암군	35	4	134	48.84	0		311		265	57.54	354	15.52	297	31.51	274	14.16	349	52.34	616	27.09	383	34.66	310	35.65
무안군	33	2	306	38.54	1795	72.10	328	24.90	304	53.97	305	14.33	326	29.39	308	12.43	400	53.18	403	48.69	424	34.13	356	33.32
영광군	36	3	497	32.94	837	94.28	313	9.82	258	62.81	226	22.65	302	31.67	272	14.44	346	59.76	388	48.66	393	36.23	312	37.71
장성군	33	2	390	26.92	0		239		383	42.61	313	17.92	213	45.41	286	13.27	474	43.97	433	45.21	291	49.56	420	28.32

〈부표 7〉 소지역 추정 결과 : 경상북도

시군	조사구수		고용통계		dir		syn		EBLUP1		EBLUP2		EBLUP3		EBLUP4		HB1		HB2		HB3		HB4	
	고용	경활	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV
포항시	38	20	5346	21.32	6380	25.96	6898	5.93	5928	5.89	5495	8.12	5696	4.90	6021	4.78	6328	5.46	6401	15.45	5990	4.57	8606	3.95
경주시	40	12	1921	26.15	4277	46.01	3600	5.31	2804	7.31	2887	4.67	2690	5.82	2808	4.62	3000	7.91	3159	14.68	2856	6.26	3664	4.99
김천시	43	5	1853	16.61	927	112.64	1870	2.21	1281	12.95	1355	6.35	1215	8.71	1258	4.84	1403	14.82	1564	16.21	1330	10.96	1517	8.40
안동시	40	10	2115	21.99	1594	85.05	2253	3.71	1577	10.85	1928	6.70	1533	7.49	1573	4.79	1714	12.45	2343	14.52	1659	9.13	1931	7.14
구미시	42	11	6262	16.32	12218	43.32	5449	13.63	6789	5.87	6830	12.57	4026	5.24	5491	7.38	7323	5.30	8963	16.67	4276	5.13	9264	5.14
영주시	45	4	952	28.39	0		1532		895	18.24	1008	5.00	966	10.40	929	6.07	1020	20.57	1194	18.07	1087	13.27	1098	11.46
영천시	41	4	1496	21.82	1146	106.07	1412	2.22	908	17.98	999	7.00	887	11.17	899	5.46	1029	20.30	1273	17.68	1003	14.30	1064	11.43
상주시	44	6	1351	21.90	1086	58.60	1434	2.23	691	23.01	899	10.78	911	10.83	785	8.18	806	25.08	1406	15.37	1027	13.67	935	13.73
문경시	36	4	599	27.82	0		945		542	30.00	543	9.51	566	16.88	549	7.56	651	32.13	700	27.58	673	21.15	634	18.65
경산시	40	8	2757	19.91	4584	46.44	3409	7.56	2775	7.36	2936	5.41	2362	6.10	2624	4.61	2972	7.97	3235	14.73	2524	6.75	3477	5.00
의성군	38	2	286	47.53	0		140		269	61.31	256	21.43	458	20.85	340	14.75	351	60.57	419	46.22	548	26.18	421	29.21
청송군	33	2	271	29.62	0		63		123	63.87	133	33.36	175	37.93	140	28.84	183	43.90	207	38.84	241	30.90	161	41.68
영양군	33	2	201	35.02	0		43		128	129.23	97	43.01	109	90.59	114	35.87	190	110.83	163	116.38	167	87.23	136	88.41
영덕군	33	2	282	44.80	0		95		283	58.30	198	25.32	300	32.04	285	13.36	365	58.35	305	64.52	382	37.86	320	36.96
청도군	33	2	307	47.44	117	76.16	97	38.34	281	27.91	320	13.69	308	21.30	288	13.36	359	22.57	417	19.21	387	20.29	319	22.99
성주군	33	2	237	45.89	0		97		239	67.84	327	26.48	302	31.65	261	15.36	317	64.67	677	29.02	382	36.99	294	39.70
칠곡군	37	3	909	30.57	849	124.72	214	66.36	1817	9.62	1950	10.67	904	10.98	1321	6.97	1941	10.95	2465	15.42	1017	14.06	2192	6.98
예천군	33	2	142	47.50	0		109		212	76.81	285	16.66	348	27.37	262	17.08	287	71.45	441	41.94	431	32.71	317	37.36
봉화군	32	3	107	48.61	0		78		207	79.34	196	20.72	229	42.31	212	18.34	281	74.61	279	67.32	303	47.44	237	49.82
울진군	38	2	733	39.45	0		113		495	33.11	337	13.85	367	26.06	425	8.12	587	36.08	460	43.17	453	31.78	521	22.47

〈부표 8〉 소지역 추정 결과 : 경상남도

시군	조사구수		고용통계		dir		syn		EBLUP1		EBLUP2		EBLUP3		EBLUP4		HB1		HB2		HB3		HB4	
	고용	경활	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV
창원시	38	20	6699	17.51	7943	24.53	5944	6.50	6436	5.86	6717	7.68	5464	4.94	6171	5.20	6898	5.36	8579	15.47	5763	4.62	9044	4.14
마산시	45	11	3655	23.06	1179	74.18	4975	2.77	6033	5.76	6305	5.62	4585	5.06	5481	5.64	6287	5.54	7001	15.47	4712	4.98	7979	4.52
진주시	45	10	3094	21.29	3299	36.84	3844	3.64	3293	6.73	3683	5.89	3414	5.42	3432	4.73	3495	7.04	4173	14.48	3582	5.53	4533	4.71
진해시	46	4	2729	16.18	6887	73.72	1870	6.28	2221	8.43	2119	12.31	1422	7.87	1826	5.39	2410	9.37	2525	19.19	1573	9.60	2647	5.94
통영시	47	3	397	50.04	0		1535		1134	13.97	1339	7.10	1133	9.04	1139	5.21	1264	15.88	1598	15.87	1256	11.28	1367	9.18
사천시	45	4	1343	25.60	1126	106.07	1289	3.35	1008	16.25	1261	9.44	937	10.66	977	5.04	1131	18.46	1626	15.60	1054	13.64	1169	10.41
김해시	44	9	5337	19.14	4755	59.81	5116	7.54	6442	5.88	5636	10.98	4845	5.05	5830	5.77	6862	5.40	6401	18.76	5080	4.84	8740	4.50
밀양시	45	4	819	35.30	3044	82.31	1499	9.07	897	18.32	862	9.95	1338	8.18	1086	8.63	1030	20.62	1003	23.54	1475	10.14	1379	11.20
거제시	43	7	1912	19.31	1391	88.74	2241	3.57	1966	9.10	1907	13.38	1476	7.66	1741	4.61	2104	10.29	2113	20.50	1599	9.39	2291	6.14
양산시	39	6	2972	20.89	4701	45.65	2642	8.00	3804	6.45	3405	8.13	2187	6.28	3001	6.81	4046	6.56	3701	17.29	2347	7.08	4808	5.37
의령군	33	2	202	41.12	0		113		230	68.43	260	16.57	196	48.94	208	18.28	307	62.75	356	48.60	268	51.41	244	46.66
함안군	38	2	618	26.72	835	100.00	280	39.13	691	23.32	698	7.41	469	20.26	572	6.17	803	25.48	836	23.94	571	24.76	749	15.52
창녕군	38	2	248	49.42	0		236		421	38.71	357	17.83	460	20.70	434	9.29	514	40.55	500	39.81	553	25.73	490	23.98
고성군	33	2	120	58.77	0		215		374	41.92	377	13.63	401	23.47	381	10.23	463	41.42	486	36.52	490	27.80	429	26.28
남해군	33	2	462	33.23	0		195		308	53.50	334	13.04	378	25.34	333	12.29	395	53.89	453	41.99	467	30.84	377	31.50
하동군	33	2	230	42.94	0		193		376	36.22	262	16.24	367	24.29	366	10.13	464	33.14	412	34.19	454	26.44	415	24.23
산청군	33	2	240	34.46	0		126		308	51.36	270	16.40	236	40.52	266	13.77	391	49.83	361	48.86	312	44.33	322	35.30
함양군	33	2	155	53.69	0		149		165	99.73	240	20.53	273	35.27	204	20.83	236	89.10	442	42.88	354	40.59	249	47.86
거창군	38	2	626	29.90	781	70.71	299	53.22	333	47.46	487	12.68	483	19.49	391	12.21	429	46.69	863	22.11	586	23.62	466	25.53
합천군	32	3	694	36.86	0		169		329	36.17	247	16.94	370	22.66	342	11.52	413	31.19	331	35.92	455	23.86	382	24.30

〈부표 9〉 소지역 추정 결과 : 제주도

시군	조사구수		고용통계		dir		syn		EBLUP1		EBLUP2		EBLUP3		EBLUP4		HB1		HB2		HB3		HB4	
	고용	경활	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV	Est	CV
제주시	23	32	4048	17.86	2670	30.11	3029	17.97	3432	6.52	3902	6.44	2642	5.81	3105	5.02	3627	6.67	4475	14.94	2788	6.24	4276	4.73
서귀포시	37	13	1747	18.33	2272	62.60	1171	14.36	902	18.15	919	7.38	834	11.79	869	5.23	1032	20.54	1069	21.52	958	15.06	1045	11.57