

제3장

소비자물가조사의 표본선정방식 전환 검토

이 의 규

제1절 서 론

일반적으로 우리는 개개 상품(서비스)의 화폐가치를 가격이라 하고 여러 가지 상품(서비스)의 평균적인 가격 수준을 물가라 지칭한다. 그러나 여러 가지 상품을 종합하여 산정한 물가는 그 자체로 의미가 없어 주로 어느 특정 시점과 비교하여 상대적인 수치로 나타내는데 이를 물가지수라 한다. 물가지수는 소비자물가지수, 생산자물가지수, 수출입물가지수 등이 대표적이다. 이 중 소비자물가지수는 소매단계의 물가변동을 나타내는 것이라 볼 수 있다.

소비자물가지수는 우리의 삶과 밀접한 경기 지표로서 흔히 체온계에 비유되곤 한다. 이는 정부채정, 금융정책의 기초자료로 이용되기도 하고 가계수지, 국민소득계정 등 다른 경제지표의 디플레이터(deflator; 가격수정인자)로도 사용된다. 이와 같이 소매 물가의 종합적인 움직임을 나타내는 중요 지표인 소비자물가지수는 도시가구가 소비하는 상품(서비스)의 가격을 조사하는 소비자물가조사의 결과에 기초하여 매월 작성된다.

따라서 소비자물가지수를 작성하기 위한 가장 기본적인 과정은 가격 자료를 계속적으로 수집하는 것이며, 완전하고 정확한 물가지수가 산정

될 수 있도록 하기 위해서는 그 나라의 모든 상품(서비스) 판매처에서 판매되는 모든 다양한 상품(서비스)의 가격을 남김없이 조사하여야 한다. 그러나 모든 존재하는 상품과 서비스의 가격을 조사할 수 없으므로 표본가격을 수집하여 물가지수를 작성한다. 그러므로 표본 상품과 서비스는 전체 상품과 서비스를 잘 대표할 수 있도록 하여야 한다. 즉, 지역별 가격차이, 서로 다른 유통업태에 따른 가격차이, 그리고 상품품목별 가격차이 등을 고려하여 표본이 선택되어야 할 것이다.

표본을 추출하는 방법은 크게 확률표본추출(probability sampling)과 비확률표본추출(non-probability sampling)로 나뉜다. 현행 소비자물가지수의 조사대상처 표본은 비확률표본추출방법 중 하나인 유의표본추출(purposive sampling) 방식으로 선정된다. 그간 확률표본추출방법이 객관적이고 과학적인 방법이라는 점에서 확률표본추출 방식으로의 전환을 검토하는 문제가 제기되어 왔다. 이에 따라 본 연구에서는 소비자물가지수에 있어서 가격표본 추출방식의 전환문제를 검토하고 표본추출과 관련한 해외사례를 연구하여 표본의 대표성을 확보하고 물가지수에 대한 신뢰도를 높일 수 있는 방안을 모색하고자 한다. 본 연구에서 다루는 표본은 집세조사의 조사대상처인 가구표본이 아닌 상품(서비스) 가격조사의 조사대상처인 사업체표본에 국한한다.

제2절에서는 표본추출방법에 대해 간략하게 소개하고 소비자물가지수의 전체적인 업무흐름을 대략 살펴본다. 이후 제3절에서는 우리나라 소비자물가지수의 표본선정 현황과 문제점을 파악하고, 제4절에서는 해외 각국에서 어떤 추출방법을 사용하고 있으며 어떤 노력을 하고 있는지를 검토하고, 아울러 소비자물가지수의 표본추출과 관련한 해외연구 동향을 간략히 살펴본다. 마지막 제5절에서는 현황분석 및 해외사례로부터 도출된 시사점을 가지고 결론을 맺는다.

제2절 표본추출법과 소비자물가조사

1. 표본추출법의 개요

표본을 추출하는 방법은 확률표본추출(probability sampling)과 비확률 표본추출(non-probability sampling)로 대별된다. 확률표본추출은 표본추출시 확률기법을 적용하여 표본을 취하는 추출방법으로 단순임의표본추출(simple random sampling), 층화임의표본추출(stratified random sampling), 집락표본추출(cluster sampling), 계통표본추출(systematic sampling)과 부등확률표본추출인 크기비례확률표본추출(sampling with probability proportional to size)이 대표적이다. 확률표본추출방법에 대한 자세한 설명은 표본조사와 관련된 대부분의 교재에 상세히 소개되고 있으므로 생략하기로 한다. 여기서는 확률표본추출방법의 하나로서 가격 표본을 추출하는 데 빈번히 사용되는 크기비례확률표본추출(sampling with probability proportional to size)에 대해 간략히 설명한다. 그리고 표본추출시 확률기법이 적용되지 않는 비확률표본추출법과 비확률표본추출법 중 하나인 유의표본추출에 대해 짧게 소개한다.

가. 크기비례확률(PPS) 표본추출

크기비례확률(PPS) 표본추출은 부등확률표본추출의 대표적인 형태로 각 개체가 등확률을 가지고 추출되는 것이 아니라 표본추출단위의 주어진 크기에 비례하는 부등확률에 따라 추출하는 것이다. 표본을 선택할 때 크기가 큰 표본추출단위는 높은 추출 확률을 가지고 추출하고 크기가 작은 표본추출단위는 작은 추출 확률을 가지고 추출하고자 할 때 사용된다.

예를 들어 <표 3-1>에서와 같이 10개의 판매처가 제각기 규모가 다른 종사자를 갖는다고 하자(ILO, 2004).

이때 종사자의 수가 각기 다른 10개의 판매처 중 3개의 표본을 계통추출한다면, 누적종사자의 수가 90이므로 표본크기 3으로 나누면 표본

〈표 3-1〉 크기비례확률표본추출의 예

판매처	종사자수	누적종사자수	포함구간	추출여부
1	13	13	1-13	
2	2	15	14-15	
3	5	20	16-20	
4	9	29	21-29	추출
5	1	30	30	
6	25	55	31-55	추출
7	10	65	56-65	
8	6	71	66-71	
9	11	82	72-82	
10	8	90	83-90	추출

자료: ILO, *Consumer Price Index Manual: Theory and Practice*, 2004.

추출구간은 30이 된다. 따라서 1에서 30 사이의 난수를 발생시킨다. 예를 들어 25가 선택되었다면 25, 55(25+30), 85(25+2×30)의 수를 포함하는 구간을 갖는 판매처 4, 6, 10이 표본(표에서 음영처리)으로 추출된다.

이와 같이 크기에 차이가 많이 나는 경우라면 표본추출단위가 상대적으로 동질적인 상황에서 사용하는 등확률추출보다는 보조변수의 크기에 비례하는 확률로 표본을 추출하는 편이 더 나을 것이다. 크기비례 확률표본추출은 소비자물가조사의 지역, 품목, 대상처의 선정시 인구의 크기, 매출액, 종사자의 수 등 보조정보를 이용하여 사용되고 있다.

나. 비확률 표본추출

비확률표본추출은 말 그대로 표본추출시 확률적 요소가 담기지 않는 방법으로 확률표본을 선정하기 위한 목록작성이 어렵거나 추정하고자 하는 모집단의 특성치가 꼭 확률표본이 아니어도 결과에 심각한 영향을 주지 않는 경우에 주로 사용된다. 비확률표본추출에는 유의표본추출(purposive sampling), 눈덩이표본추출(snowball sampling), 간편표본추출(convenience sampling) 등이 있다.

이 중 유의표본추출(purposive sampling)은 조사자의 주관적인 판단에 의거하여 표본을 선정하는 방법으로 판단표본추출(judgment sampling)과 할당표본추출(quota sampling) 등이 있다. 판단표본추출(judgment sampling)은 조사자가 자신의 지식과 경험에 의해 모집단을 가장 잘 대표한다고 여겨지는 표본을 주관적으로 선정하는 방법이며, 할당추출(quota sampling) 방법은 중요한 구성요소에 대해 모집단과 같은 구조를 갖도록 표본을 할당하는 방법으로 판단추출방법의 결점을 어느 정도 보완한 방법이다(통계청, 2005). 현행 소비자물가조사의 표본은 비확률표본추출방법인 유의표본추출방식으로 선정되며, 모집단과 비슷한 구조를 갖도록 표본을 배정하는 할당표본에 가까운 유의표본이라 할 수 있다. 이에 대해서는 제3절의 표본선정의 현황에서 다시 언급하기로 한다.

2. 소비자물가조사의 개관

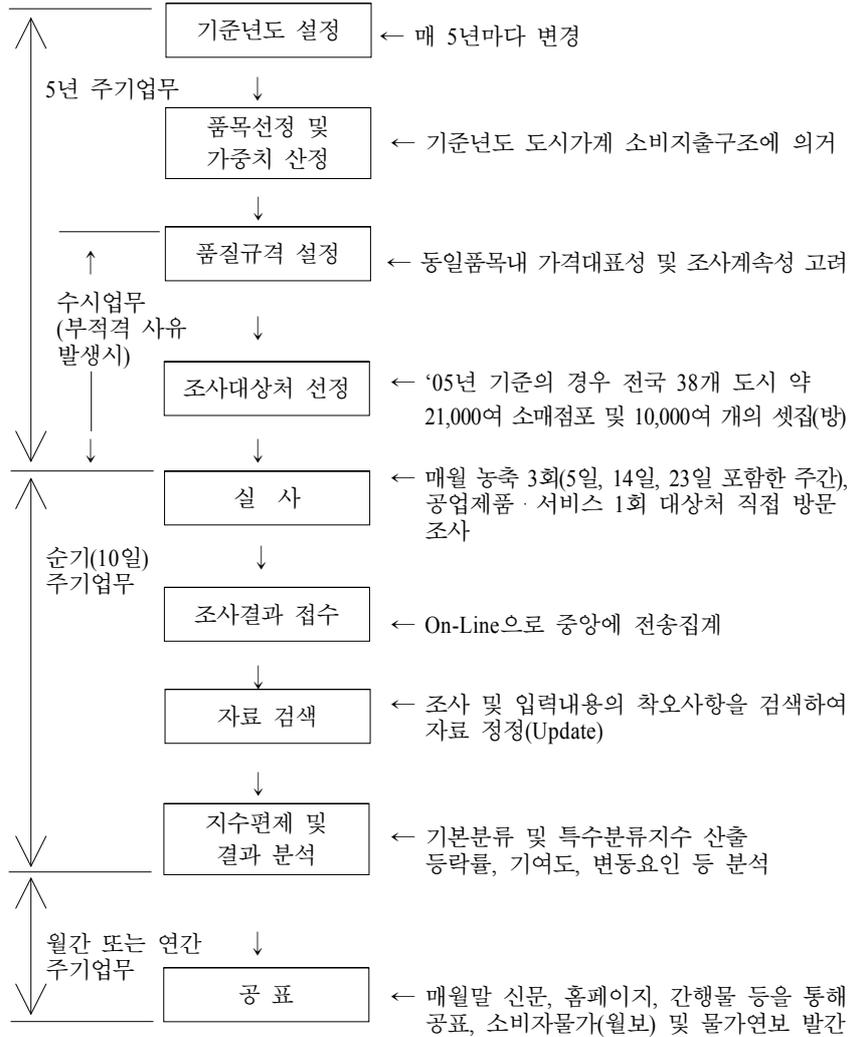
소비자물가지수 작성은 일반적으로 다음과 같은 순서에 따른다. 먼저 서로 성질이 비슷한 상품(서비스)을 품목으로 묶고 가계조사의 품목 소비점유율로부터 가중치를 계산한다. 그리고 각 품목에 속하는 상품(서비스)의 가격변동을 대표하는 상품(서비스)를 선정하여 그 가격을 계속적으로 조사한다. 수집된 가격으로부터 품목지수를 작성하게 되고 각각의 품목지수를 해당품목의 가중치로 가중평균하여 총지수를 작성한다.

따라서 소비자물가지수는 계속적으로 조사되는 소비자물가조사의 결과에 의해 작성되므로 표본선정의 현황을 살펴보기에 앞서, 통계청 물가통계과에서 시행하고 있는 소비자물가조사의 전체 업무흐름을 우선 이해할 필요가 있다.

[그림 3-1]의 소비자물가조사 체계도(통계청, 2007a)에서 볼 수 있듯이, 소비자물가통계에 대한 업무는 크게 설계, 실사, 공표의 3단계로 나눌 수 있다. 설계 단계는 기준년도 설정, 품목선정 및 가중치 산정, 품질규격 설정, 그리고 조사대상처 선정의 업무로서 5년을 주기로 이루어진다.

기준년도 설정은 어느 해를 기준으로 지수를 계산할 것인가와 관련된 일이며 현재 2005년을 기준년도로 한다. 품목선정은 어떤 품목을 조사할 것인가에 대한 업무이며 2005년 도시가계조사의 품목별 소비지출

[그림 3-1] 소비자물가조사 체계도



자료: 통계청, 「통계행정편람」, 2007.

액을 기준으로 1/10,000 이상 거래품목을 선정한다. 가중치 산정은 기준
 년도의 전체 소비지출액 중 각 품목이 얼마만큼의 비중을 갖는지를 결
 정한다. 그리고 품질규격 설정은 계속 같은 형태를 유지하면서 가격변

화를 조사할 수 있도록 어떤 규격으로 조사할 것인가를 설정하며, 조사 대상처 선정은 어디에서 조사할 것인가를 정하는 절차인데 부적격 사유 발생시는 수시로 업무를 처리한다.

실사 단계는 10일 주기로 매월 가격을 조사하여 조사 자료를 검토한 후 각종 지수를 집계하는 업무이고, 마지막으로 공표단계는 편제된 지수를 바탕으로 매 월말, 그리고 연말에 홍보매체를 통해 소비자물가지수를 공표하는 절차이다.

이와 같이 전체업무흐름을 볼 때 소비자물가지수의 가격조사대상의 표본추출문제는 소비자물가지수의 업무 중 설계 단계에서의 품목선정 및 조사대상처의 선정에 해당된다. 다음 절에서는 가격표본선정의 현황에 대하여 좀 더 상세하게 파악하고 아울러 현행 표본선정의 문제점을 살펴보고자 한다.

제3절 표본선정의 현황 및 문제점

1. 표본선정의 현황

가. 품목의 선정

우리나라는 도시가계 통계조사를 통해 얻어진 도시지역에 거주하는 가정의 가계부를 기초로 하여 조사할 품목을 선정한다. 2005년도 기준의 경우 도시가계의 월평균 소비지출액 비중이 총소비지출액의 1/10,000 이상(월평균 185원)이 되는 품목을 원칙으로 선정한다. 품목선정시에는 소비지출액 비중 이외에도 동종상품군에 대한 가격 변동을 대표할 수 있고 향후 5년 이상 계속 조사가 가능한 상품 또는 서비스로 선정된다. 2005년 기준 소비자물가지수 개편 결과(통계청, 2006), 489개의 품목이 선정되었으며 품목 및 가중치 현황은 <표 3-2>와 같다.

<표 3-2>에서 괄호 안에 표기된 12대 분류는 2005년 기준년도 개편 시 국제비교성을 높이기 위해 국제노동기구(ILO)에서 권고하는 목적별 개인소비 분류(classification of individual consumption according to purpose:

〈표 3-2〉 품목 및 가중치 현황

분류내역	품목수	가중치	분류내역	품목수	가중치
총 지 수	489	1000.0			
(식료품·비주류음료)	131	140.4	의료용품	5	2.7
곡 물	14	28.0	의료 서비스	4	30.6
육 류	7	21.0	기타 의료서비스	3	1.5
어 개	19	16.6	(교 통)	31	109.0
낙 농 품	5	11.2	차량구입비	6	21.7
유 지	2	1.1	차량 연료	3	47.0
과 실	15	17.3	차량부품·수리비	5	6.8
채소·해조	32	19.2	기타차량서비스	5	3.6
과자·당류	9	8.1	철도이용료	2	5.3
조미료·기타식품	18	10.9	육상이용료	4	20.6
차	4	1.7	항공이용료	4	2.4
음 료	6	5.3	기타차량이용서비스	2	1.6
(주류·담배)	9	14.6	(통 신)	13	60.2
주 류	7	3.8	우편서비스	1	0.1
담 배	2	10.8	전 화 기	2	3.6
(의복·신발)	42	58.4	전화·정보이용료	10	56.5
남자 의류	15	23.4	(교양·오락)	64	56.3
여자 의류	14	19.5	영상·음향기기	5	4.0
아동복·유아복	2	4.3	정보처리기기·소모품	7	5.3
기타 의류	4	1.9	악기·오락용품	9	5.8
의류 서비스	2	2.1	오락·운동서비스	15	9.7
신 발	5	7.2	문화서비스	11	12.6
(주거 및 수도·광열)	19	170.4	도 서	7	7.7
집 세	2	97.5	신문·잡지	2	1.4
주택설비수리	5	6.8	문 방 구	6	1.8
상하 수도료	2	7.8	단체여행	2	8.0
기타 주거	3	14.6	(교 육)	23	110.9
전 기 료	1	19.0	유치원·초등교육	6	36.3
가 스 료	3	17.6	중등 교육	6	40.0
연 료	3	7.1	고등 교육	6	26.6
(가구집기·가사용품)	52	41.7	기타 교육	5	8.0
가 구	8	7.9	(외식·숙박)	43	132.7
침구·직물제품	3	2.4	외 식	39	130.6
가정용 기구	15	12.7	숙 박	4	2.1
가정용기구수리서비스	3	1.2	(기타 잡비)	34	53.8
주방용품	10	2.5	이미용서비스	4	17.3
가사소모품	11	3.7	이미용품	20	17.5
가사서비스	2	11.3	개인용품	5	8.4
(보건 의 료)	28	51.6	기타 서비스	5	10.6
의 약 품	16	16.8			

자료: 통계청, “2005년 기준 소비자물가지수 개편결과”, 2006.

COICOP) 체계로 개편한 결과이다. 그리고 <표 3-2>의 가중치는 가계소비지출액 총액을 1000.0으로 하고 개별품목이 차지하는 구성비를 그 품목의 가중치로 한 것이며, 가중치 산출을 위한 기초자료로 종전 2인 이상 도시가구를 대상으로 하던 조사에서 1인 이상 도시가구로 확대하여 산출한 것이다. 품목성질별로는 상품이 357개(농축수산물 71, 공업제품 258)이고 서비스가 160개(집세 2, 공공서비스 32, 개인서비스 126)이며 지수에 반영되는 조사규격 수는 793개이다.

나. 조사지역의 선정

조사지역은 전국적으로 대상지역 표본이 골고루 반영되도록 선정한 다. 시도별 시부인구를 기준으로 60% 이상을 대표할 수 있도록 38개 도시를 선정하는데, 7대 도시는 모두 포함하고 기타도시는 도인구가 100만 미만은 2개시, 100만 이상은 100만을 초과할 때마다 1개시의 추가를 원칙으로 한다. 38개의 도시를 나열하면 서울, 부산, 대구, 인천, 광주, 대전, 울산(이상 특별시 및 광역시), 수원, 성남, 의정부, 안양, 부천, 고양, 안산(이상 경기도), 춘천, 원주, 강릉(이상 강원도), 청주, 충주(이상 충청), 천안, 보령, 서산(이상 충남), 전주, 군산, 남원(이상 전북), 목포, 여수, 순천(이상 전남), 포항, 경주, 안동, 구미(이상 경북), 창원, 김해, 마산, 진주(이상 경남), 제주, 서귀포(이상 제주도)이다. 2000년에 비해 2005년도 개편에서는 경기도 안산과 경남 김해의 두 도시가 추가되었으며, 공주시는 대전시와 인접하고 있고 상권발달의 미흡으로 조사대상처의 선정이 어려워 서산시로 대체되었다(통계청, 2006).

다. 조사권역의 선정

조사권역은 조사지역의 하부단위로서 시장의 지역적 분포를 적절히 반영할 수 있도록 상권을 중심으로 몇 개의 권역으로 구분한다. 조사권역수는 도시별 인구를 기준으로 최소권역수를 선정하는데, 인구 100만 명 이상인 광역시는 최소 5개 권역, 50~100만명 미만인 중도시는 최소 3개 권역, 50만명 미만인 소도시는 최소 2개 권역을 가지도록 함을 원칙

으로 한다. 조사권역은 소비자가 거주지와 가까운 인근 점포에서 농축 수산물과 식료품을 구입하는 구매패턴을 고려하여 도시의 변화가보다는 행정구역을 기준으로 설정한다. 2005년 기준년도 개편시 종전 121개 조사권역에서 중소도시를 중심으로 조사권역을 추가하여 150개 권역으로 확대되었다(통계청, 2006).

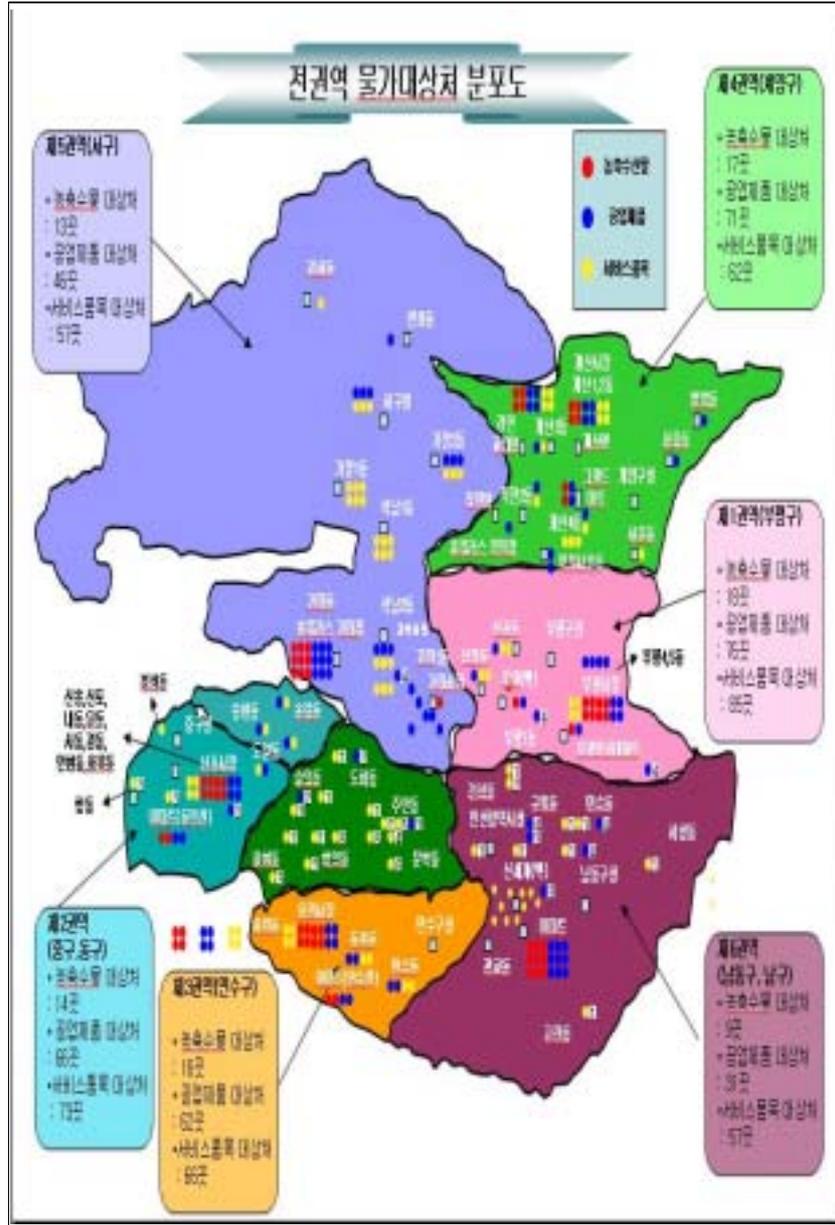
라. 조사대상처의 선정

소비자들이 조사품목을 구매할 수 있는 소매 점포와 서비스 점포로서 소비자 물가조사의 대상이 되는 곳을 조사대상처라 한다. 원칙적으로는 조사권역별로 중앙품목을 제외한 조사품목 전부에 대한 가격조사가 가능하도록 조사대상처를 선정한다.

조사대상처는 조사권역(지역)의 가격변동을 대표할 수 있도록 소비자들의 인지도 및 소비행태, 상권 매출액 규모 등을 고려하여 소비자들의 이용이 가장 많은 대중성이 있는 소매점포 및 서비스 점포를 선정한다. 또한 조사도시별 소매유통구조(업체별 매출액 비중)를 반영하여 일반시장, 백화점 및 할인점의 조사대상 수를 결정한다. 즉, 조사대상처는 도시별 산업분류별 매출액을 고려하여 백화점, 할인점, 일반소매업 등의 대상처수를 정하여 지방청에 배분하고 배분된 조사대상처 수만큼 대표성과 대중성을 고려하여 선정한다(통계청, 2007b).

현재 소비자의 출입이 많은 백화점, 할인점, 슈퍼마켓, 전문점, 재래점 등을 포함한 20,000여 개 소매점포 및 서비스 업체에서 조사하고 있으며, 가격체계가 전국적으로 동일한 품목은 중앙에서 조사한다. 참고로 권역별 조사대상처의 현황분포를 시각적으로 잘 나타낸 2006년 인천통계사무소 물가통계팀에서 작성한 인천시 권역별 조사대상처 분포도 [그림 3-2]를 제시한다.

[그림 3-2] 인천광역시 권역별 조사대상처 분포도



자료: 인천통계사무소, “소비자물가조사 개선방안”, 2006.

2. 유통업태와 소비자구매성향의 변화

가. 유통업태의 변화

가격의 변화는 지역별로 차이가 존재하지만 상품을 판매하는 판매유형에 따라서도 차이가 나타난다. <표 3-3>에서 연도별 유통업태별 사업체수 추이를 살펴보면, 대형할인점은 증가 추세에 있는 반면 음식료품 소매업은 감소 추세에 있다. 따라서 조사대상처의 선정시 이와 같은 유통업태의 변화를 고려하여 조사대상처의 수를 배분하여야 할 것이다.

<표 3-3> 연도별 산업세세분류별 사업체수, 매출액

	사업체수(개)			매출액(10억원)		
	2001	2004	2005	2001	2004	2005
소매업 (52)	627,044	638,017	587,358	122,428	131,403	146,319
백화점 (52111)	89	87	80	16,133	9,094	17,375
대형할인점 (52119)	238	284	317	16,942	21,526	23,708
슈퍼마켓 (52121)	5,867	6,898	6,574	9,040	10,570	10,951
음식료품종합소매업 (52129)	107,365	103,298	95,967	6,471	7,999	7,263
음식료품소매업 (522)	115,416	108,689	100,973	9,779	10,775	8,957

자료: 2001, 2005년도는 서비스업총조사, 2004년도는 도소매업통계조사; 통계청(2007b) 내부자료 인용.

통계청에서는 2005년 서비스업 총조사자료를 이용하여 산업분류 52(소매업)자료를 38개 도시별, 5개 소매업태별로 집계한 후 5개 소매업태별 매출액 자료에서 음식료품 매출액만 따로 분리·집계하였다. 도시별 음식료품의 소매업태(대상처 유형)별 매출액을 파악하여 대상처 유형별 비율을 매출액에 비례하도록 조정하는 조사대상처의 산업분류별 구성비 변경 계획을 추진하고 있다(통계청, 2007b).

나. 소비자구매패턴의 변화추이

인천통계사무소(2006)에서도 지적인 바와 같이, 유통업태의 변화 추이는 소비자 구매성향의 변화에서도 확인할 수가 있다. <표 3-4>의 소비자 구매패턴조사에서 나타난 결과에 의하면, 식료품인 경우 1997년에 슈퍼편의점이 1순위였으나 2004년에는 3순위로 내려갔고 반면 3순위였던 대형할인점이 1순위로 조사되었다. 내구재의 경우에도 3순위였던 대형할인점이 1순위로 나타났으며, 의류에서는 재래시장과 대형할인점의 순위가 뒤바뀌었다. 따라서 조사대상처 선정시 품목별 구매처와 구매량에 대한 조사가 필요하며, 이러한 소비자 구매처 변화에 따라 조사대상처를 조정할 필요가 있다.

<표 3-4> 소비자 구매패턴의 변화추이

품목명	연도	주요 구매처		
		1 순위	2 순위	3 순위
식료품	1997	슈퍼편의점	재래시장	대형할인점
	2004	대형할인점	재래시장	슈퍼편의점
생활용품	1997	대형할인점	슈퍼편의점	재래시장
	2004	대형할인점	슈퍼편의점	재래시장
의류	1997	백화점	재래시장	대형할인점
	2004	백화점	대형할인점	재래시장
내구재	1997	전문점	백화점	대형할인점
	2004	대형할인점	전문점	백화점

자료: 대한상공회의소, 「소비자 구매패턴조사」, 각년도; 인천통계사무소, “소비자물가조사 개선방안”, 2006에서 재인용.

3. 표본선정의 문제점

확률표본추출방법은 과학적인 방법으로 표본을 추출하여 표본의 객관성이 확보되고 표본오차의 통계적 추정이 가능하다는 장점을 가진다.

그러나 현행 우리나라의 표본추출은 확률표본추출이 아닌 유의표본추출방식을 취하고 있어 표본의 객관성이나 모집단의 특성치에 대한 추론 결과의 통계적 신뢰성에 관해 언급할 수 없다는 한계를 가진다.

그러나 통계학적 관점에서 일정한 신뢰성을 확보하려면 표본수가 커지고 가격조사 비용이 늘어나는 등의 단점이 있어 현실적으로 확률표본추출방법을 사용하는 경우는 많지 않다. 또한 확률표본을 추출하기 위해서는 누가 어떤 상품(서비스)을 어느 정도 판매하고 있는가를 남김없이 조사한 모집단 존재가 전제되어야 하는데 이 품목에 대한 적절한 추출틀이 없다는 것이 가장 어려운 문제이다. 더욱이 조사대상이 되는 사업체의 협력이 대전제가 되어야 하며 장기간 계속조사가 가능해야 한다는 점 또한 간과할 수 없는 문제이다.

한편 앞에서 보았듯이 유통업체와 소비자 구매성향은 변화하고 있다. 소비자의 구매 장소는 주차시설이 완비되고 한 곳에서 여러 가지의 상품과 서비스를 구매할 수 있는 대형할인점으로 급격히 이동하고 있다. 따라서 유통업체와 소비자구매패턴을 반영하는 조사지역과 조사대상 선정 및 조정에 대한 보다 체계적인 기준이 필요하다. 다음 절에서는 해외사례를 통해 각국에서는 위와 같은 표본추출과 관련된 문제에 대해 어떤 노력을 하고 있는지를 검토하여 표본선정의 개선방향과 시사점을 도출하고자 한다.

제4절 해외사례와 연구동향

각국은 소비자물가지수 작성에 있어서 객관적인 가격변동을 나타낼 수 있도록 그 나라 나름대로 가격표본을 선정하고 있다. 대부분의 나라에서 소비자물가지수 표본은 비확률표본이 지배적이며 미국과 영국 등이 확률표본을 지향하고 있다.

1. EU 회원국의 표본추출

Boon(1997)은 EU국가의 소비자물가지수의 작성에 있어서의 표본설

계에 대해 발표한 바 있다. <표 3-5>는 이들 EU국가의 소비자물가조사에 있어서 지역추출과 대상처의 표본 추출에 대하여 다시 요약·정리한 것이다.

<표 3-5> EU 회원국의 표본추출

국가	지역추출	대상처추출
오스트리아	9개 지방도시+ 인구 20,000 이상 11개 도시추출(cut-off)	지방통계위원회와 지방자치행정기관에 의해 대상처 유의표본 추출(judgmental)
덴마크	5개 대도시+ 25개 중소도시 판단추출법(judgmental)	소매상명부가 표본틀로 활용. 1,900개 대상처 추출은 전문가 선택에 의존(judgmental)
핀란드	자치시와 주요시는 각 지역(층)에서 크기비례확률로 층화추출(stratified PPS)	채소류 상점의 표본은 사업체명부에서 추출(stratified PPS). 가전용품 대상처는 판단추출(judgmental)
독일	작은 자치시는 제외(cut-off), 자치시는 인구규모에 따라 층화, 층의 표본크기는 모집단 구성비에 따라 고정(quota)	대상처수는 자치시에 따라 다양. 전문가의 판단에 따른 추출(judgmental)
그리스	13개 지역으로 분류, 전문가의 판단에 따라 도시 추출(judgmental)	지역사업체명부 이용. 전문가 판단에 기초한 대상처 선정(judgmental)
아일랜드	10,000명 넘는 마을 + 나머지 도시는 층화추출(stratified SI)	가격조사자의 판단(judgmental)
이탈리아	82개 지방도시(unknown)	본부의 지침에 따라 지방통계사무소에서 가장 높은 판매액을 가진 대상처나 품목당 판매액이 가장 높은 대상처 선정(cut-off)
룩셈부르크	소매가격은 3개의 지역 중심부, 서비스 가격은 22개 마을(judgmental)	사업체명부를 대상처 선정시 표본추출틀로 사용. 대상처의 선정은 전문가의 판단(judgmental)
네덜란드	인구 10,000 이하의 제외. 자치시는 크기비례확률로 층화추출(stratified PPS)	현장조사자에 의해 조사(judgmental). 대상처의 형태와 종사자 규모로 사후 층화(stratified SI)
스페인	인구수가 많은 지역은 반드시 포함(cut-off)	대상처의 추출틀로 사용되는 명부는 존재하지 않음. 대상처추출은 전문가의 판단(judgmental)

<표 3-5>의 계속

국가	지역추출	대상처추출
스웨덴	소수 중앙조사에서만 지역표본을 추출. 택시조사에서 19개의 자치시가 추출(unknown)	사업체명부를 일상필수품 대상처에 대한 표본추출틀로 사용. 일부 층화확률비례추출법 사용(stratified PPS).
영국	인구크기로 지역을 층화, 가장 큰 층 + 각 층에서 단순임의추출(stratified SI)	각 지역의 주요 쇼핑지역의 목록작성. 각 층에서 단순임의 추출(stratified SI) 또는 크기비례확률추출(stratified PPS)
미국	인구크기로 지역을 층화, 가장 큰 층 + 모집단 크기에 확률비례추출(stratified PPS)	전화구매시점조사(TPOPS)에 기초하여 소비에 비례하는 확률로 추출(stratified PPS)

자료: Martin Boon, "Sampling Designs in Constructing Consumer Price Indices: Current Practices at Statistical Offices", 1997.

<표 3-5>에서 나타나듯이 대부분의 EU국가에서는 우리나라와 마찬가지로 지역을 선정하고 선정된 지역에서 대상처를 정하는 2단계 표본추출방식을 취하고 있다. 지역의 추출은 절사추출(cut-off sampling), 판단추출(judgmental sampling), 할당추출(quota sampling)과 같은 비확률표본추출법이 주로 사용되고 있으며 미국, 영국 등에서는 인구크기로 지역을 층화한 후 각 층에서 단순임의로 추출(stratified SI)하거나 모집단의 인구크기에 비례하는 확률로 추출(stratified PPS)하고 있다. 대상처의 추출은 판단추출(judgmental sampling) 방법이 대부분의 국가에서 사용되고 있으며, 역시 미국, 영국 등에서는 층화 후 각 층에서 단순임의로 추출(stratified SI)하거나 크기에 비례하는 확률로 추출(stratified PPS)하고 있다.

다음 외국의 표본설계에서는 확률표본을 지향하고 있는 영국과 미국에 대해 좀 더 자세히 살펴보고 캐나다, 호주와 뉴질랜드, 그리고 일본의 현황을 추가하여 간략하게 살펴본다.

2. 외국의 CPI 표본설계

가. 영국

영국은 1994년까지는 대상처 표본을 판단추출에 의해 선정하였으나

그 이후 확률표본 추출과정을 지향하고 있다. 표본추출은 로케이션(location)추출과 점포(outlet)의 추출로 이루어진다.

1) 로케이션(location)의 추출

우선 영국을 런던, 웨일스, 이스트 미들랜드와 같은 표준 지역(region)으로 전체 분할하고 각 지역(region)내 로케이션(location)은 종사자 수의 크기에 비례하여 확률추출하는 PPS방법으로 추출한다. 각 지역에서 추출되는 로케이션 표본 수는 그 지역의 국민지출의 비율에 따라 결정한다. 한 지역의 로케이션에는 점포의 수와 종사자 수의 정보가 들어 있으며 각 지역 내에서 로케이션을 우선 목록화한다. 추출된 로케이션에서 모든 조사품목이 수집될 수 있도록 보증하기 위해 경험적으로 우선 사업체 수가 250개보다 적은 로케이션은 제외된다. 한 지역내 로케이션의 PPS 표본추출은 [그림 3-3]에서와 같이 나타나는 절차를 따른다(Office for National Statistics, 2007). 이 절차는 앞의 PPS방법의 예시와 비슷하다.

먼저 종사자의 수로 비례하는 확률로 추출하므로 누적 종사자수를 한 열에 기재한다. 총 누적 종사자수는 217,334이다. 이 지역에 표본으로 배정된 로케이션의 수가 10개라 하자. 따라서 누적종사자의 수 217,334를 표본크기 10으로 나누면 표본추출구간은 21,733.4가 된다. 이 때 21,733.4까지의 수 중에서 하나를 랜덤하게 선택하기 위해서는 먼저 0에서 1 사이의 난수를 발생시킨다. 이 예에서 0.39904가 선택되었다면 초기값은 $8,672.5(21,733.4 \times 0.39904)$ 가 되고, $30,405.9(8,672.5 + 21,733.4)$, $52,139.3(8,672.5 + 2 \times 21,733.4)$, ...의 수를 포함하는 구간의 판매처가 표본으로 추출된다. 이 결과로서 [그림 3-3]의 음영처리된 로케이션 C, I, M, S, W, Y, EE, GG, JJ, OO의 10개의 로케이션이 추출된다.

2) 점포(outlet)의 추출

표본추출들은 가격수집자에 의해 각 로케이션내 주요 상가지역의 목록을 통해서 수립된다. 가격수집자가 각 로케이션의 상점주소를 방문하여 로케이션당 1,500여 점포에 대한 상세조사를 수행하여 주소, 우편번호, 판매품목 범위, 매장 크기, 단일/연쇄점 여부 등을 파악한다. 이 자

[그림 3-3] 영국의 한 지역내 로케이션의 PPS추출

Location Name	No. Outlets	No. Employees	Cumulative Total	Range	
				Lower	Upper
Location A	607	5377	5377	1	5377
Location B	366	2486	7063	5378	7063
Location C	264	2294	30128	2864	30128
Location D	440	4006	14134	10129	14134
Location E	322	2589	16723	14135	16723
Location F	319	2097	18820	16724	18820
Location G	283	2127	20947	18821	20947
Location H	477	5252	26199	20948	26199
Location I	509	4946	31144	26200	31144
Location J	371	4102	35246	31145	35246
Location K	518	4075	40121	35247	40121
Location L	578	3707	43828	40122	43828
Location M	607	3066	46894	43829	46894
Location N	374	2449	49343	46895	49343
Location O	538	3625	52968	49344	52968
Location P	326	3357	56325	52969	56325
Location Q	291	2473	58798	56326	58798
Location R	277	2052	60850	58799	60850
Location S	1815	16156	77006	60851	77006
Location T	445	3738	80744	77007	80744
Location U	329	2387	83131	80745	83131
Location V	256	2122	85253	83132	85253
Location W	429	2618	87871	85254	87871
Location X	1714	20335	108206	87872	108206
Location Y	306	2615	110821	108207	110821
Location Z	488	3425	114246	110822	114246
Location AA	366	3177	117423	114247	117423
Location BB	264	2975	120398	117424	120398
Location CC	452	6218	126616	120399	126616
Location CD	371	1839	128455	126617	128455
Location CE	264	1792	130247	128456	130247
Location FF	678	8106	138353	130248	138353
Location GG	1515	16308	154661	138354	154661
Location HH	321	2139	156800	154662	156800
Location II	283	2207	159007	156801	159007
Location JJ	264	2180	161187	159008	161187
Location KK	312	2057	163244	161188	163244
Location LL	314	2724	165968	163245	165968
Location MM	352	2649	168617	165969	168617
Location NN	303	1729	170346	168618	170346
Location OO	600	7864	178210	170347	178210
Location PP	459	3021	181231	178211	181231
Location QQ	466	3256	184487	181232	184487

No. of baskets: 10
 Employment total: 217334
 Interval value: $21733.4 = \text{Employment total} / \text{No. of baskets}$
 Random number: 0.36609
 Random starting point: $6672.5 = \text{Interval value} \times \text{Random Number}$

Random numbers for selection:
 6672.5 = Random Starting Point
 30485.9 = Random Starting Point + Interval Value
 52139.3 = Random Starting Point + 2 x Interval Value
 73872.7 = Random Starting Point + 3 x Interval Value
 95606.1 = Random Starting Point + 4 x Interval Value
 117339.5 = Random Starting Point + 5 x Interval Value
 138972.9 = Random Starting Point + 6 x Interval Value
 160606.3 = Random Starting Point + 7 x Interval Value
 182239.7 = Random Starting Point + 8 x Interval Value
 204273.1 = Random Starting Point + 9 x Interval Value

Select Location in whose employee range random numbers fall.

자료: Office of National Statistics, "Consumer Price Indices Technical Manual", 2007.

료를 점포의 표본추출틀로 작성한다. 대상처 목록표는 비용이 많이 들어 5년 간격으로 수행된다. 대상처가 폐업했을 경우 새로운 대상처가 목록표로부터 선택된다.

각 로케이션에서 점포는 표본추출틀로부터 크기비례확률(PPS)표본 또는 단순임의추출(SRS)된다. PPS표본을 추출하기 위해서는 표집단위의 크기(size) 정보가 필요한데 이상적인 크기는 매출액일 것이나 이를 확보하기는 용이하지 않다. 따라서 조사자는 매장 면적을 추측하여 매출액에 대한 보조변수로 대체한다. 각 상점이 무엇을 팔고 있는지를 표시하기 위해 적절한 코드가 각 상점에 부여된다. 또한 상품군과 상점유형으로 분류된다. 예를 들면 <표 3-6>의 고기(Meat)의 경우, 소매정육점(Butcher)에서는 연쇄점(M)이거나 단일점(I) 모두에서 1개의 점포를 단순임의추출(SRS)로 표집하고, 슈퍼마켓(Supermarket)이나 백화점(Department)에서도 연쇄점(M)이거나 단일점(I) 모두에서 1개를 크기비례확률(PPS)로 표본추출한다(Office of National Statistics, 2007).

<표 3-6> 영국의 상점추출의 예

COMMODITY GROUP	SHOPS TO SELECT	TYPE	NO	SAMPLE
Meat 1. Fresh beef & lamb 2. Cooked meats 3. Fresh bacon, pork, chicken	Butcher	M or I	1	SRS
	Supermarket	M or I	1	PPS
	Supermarket(licensed)			
	Department store type 1			
	Department store type 3			
Department store type 5				

자료: ONS, "Consumer Price Indices Technical Manual", 2007.

나. 미국

미국은 1978년 이후 유의추출방식에서 확률추출방식으로 변경, 현재 확률추출을 시행하고 있다(DOL, 2007).

1) 지리적 위치의 표본추출

먼저 미국 전지역을 1차 표본추출단위(Primary Sampling Unit: PSU)로 분할하는데, PSU는 가격수집이 되는 가장 작은 지리적인 지역으로서 하나의 카운티 또는 몇 개의 카운티로 구성된다. 각 PSU는 인구규모에 따라 분류(A, B, C)하고 센서스지역으로 구분(Northeast, Midwest, South, West)한다. 각 지역규모급(예를 들면 South B) 안의 PSU를 몇 개의 요인(층화변수)의 점수에 근거하여 비슷한 PSU의 층으로 층화한 후 PSU를 층화표본추출하여 선택된 PSU가 전역에 걸쳐 잘 분배되도록 한다. 인구가 150만 명이 넘는 표본추출단위는 무조건 추출되도록 설계한다. 1998년 개편에서 87개의 PSU가 선택되었다.

2) 전화구매처조사

미국 센서스국이 노동통계국(Bureau of Labor Statistics: BLS)을 위해 전화구매처조사(Telephone Point of Purchase: TPOPS)를 시행하여 판매처 목록들 자료를 BLS에 송부한다. 그리고 조사대상처는 이 TPOPS에 기초한 목록표로부터 추출한다. TPOPS는 품목구입 대상처와 각 품목에 대한 각 대상처의 품목소모량을 얻는 지속적인 가구조사로서 CPI 상품과 서비스를 구매한 도시와 비도시 가구로부터 소매점에 대한 자료를 제공한다. 214개의 POPS 카테고리는 하나 또는 몇 개의 기본품목으로 구성된다. 16,800가구를 대상으로 다양한 형태의 상품과 서비스를 구매하는 장소를 밝혀준다. 중소규모의 PSU는 110개, 자기대표 PSU는 200~400개, 뉴욕시와 시카고는 460개, 로스앤젤레스는 500가구가 배정된다.

조사에 참여하는 가구를 선택하기 위해서는 임의숫자 전화조사(random digital dialing) 방식을 이용하는데 처음 건 번호의 약 25%만이 사용 가능한 인터뷰라고 한다. 인구학적이고 사회경제적인 정보를 도출하며 또한 다양한 모집단의 부분에 대한 쇼핑패턴을 분석한다. 회상주기(recalling period)라고 불리는 시기 동안에 걸쳐 POPS 카테고리 세트에 대한 지출여부의 질문을 부여받는데 1주 회상주기는 휘발유 등이고, 5년 회상주기는 자동차, 장례서비스 등이다. 그리고 지출시 판매처 이름, 지역, 구매액 등이 조사된다. 분기별 순환패널 표본설계로 16개의 설문

지로 구성되며 최대 16 POPS 카테고리에 대해 구성된다. 따라서 매 4년마다 한 번 새롭게 된다.

3) 대상처 표본추출

조사대상처 표본은 계통 PPS 표본추출절차를 사용하여 각 PSU에서 POPS 카테고리에 대해 독립적으로 추출된다. 하나의 표본추출틀에서 각 조사대상처는 TPOPS 조사에서의 POPS 카테고리에 대한 상점의 보고된 총 지출액에 비례한 추출 확률을 가진다.

그러나 몇몇 상품과 서비스는 표본추출틀이 적절하지 않거나 TPOPS에 의해서는 적절한 표본추출틀이 제공되지 않음이 확실하기 때문에 TPOPS에 포함되지 않는다. TPOPS에 포함되지 않은 상품과 서비스에 대한 조사대상처 표본추출절차는 또 다른 표본추출절차에 따른다. 21,000여 개의 대상처로부터 가격수집을 하며, 현 구매패턴을 반영하기 위해 매년 표본대상처는 추출된 각 PSU의 약 25%를 교체한다.

다. 캐나다

캐나다는 지역, 대표상품, 판매처의 선정에 있어서 판단추출(judgmental selection)을 사용하고 있다. 1997년에는 음식점(restaurant)에 대해 확률표본추출을 시도한 바 있다(Loretto, 1997). 캐나다는 주요품목에 대한 판매처 표본 검토와 갱신이 매년 수행되고 모든 판매처 표본은 적어도 매 2~3년내 검토된다. 판매처 추출은 시장정보에 근거하고 높은 판매액을 포함하도록 계획된다. 캐나다 통계청은 소비자물가조사의 표본오차는 상품과 판매처의 판단추출에 근거하므로 수량화가 어렵다는 것을 명시하고 있다. CPI작성과 관련한 추가정보는 캐나다 통계청 홈페이지에서 얻을 수 있다.

라. 호주

호주도 조사대상처의 대표 표본을 선택하기 위하여 유의표본추출(purposive sampling)을 사용한다. 특히 인구센서스, 소매업 자료, 실사조사직원의 지역정보에 근거하여 선정하고 있다. 호주는 대상처 선정시

고려된 도시의 해당상품 및 서비스 공급수, 판매처들 중 가격의 변동 정도, 판매처의 지역적 분포, 판매처의 체인점 여부 등을 고려하고 있다(호주 통계청 홈페이지).

마. 뉴질랜드

뉴질랜드는 호주와 마찬가지로 조사대상처의 대표 표본을 선택하기 위해 유의표본추출(purposive sampling)을 사용한다. 많은 유럽 국가에서 응답협조를 구하는 데 심각한 어려움이 있어 대부분의 사업체 표본조사가 유의표본추출을 사용하고 있음을 지적하고 있다(Doherty, 1994).

가격과 가격변화율이 같은 상품에 대해서도 판매처 유형에 따라 다르므로 이를 고려하여 선정하고 있다. 한 상품군에 대해 조사된 판매처의 다양한 유형의 비율을 결정하기 위한 가이드로서 가계조사(Household Economic Survey: HES)와 소매자료(retail trade data)를 보조정보로 사용하고 있다. 판매처의 시장점유와 구입된 상품의 가격변동을 고려하며 비용효율성과 표본추출편리성을 우선순위에 두고 있다(Loretto, 1997).

바. 일본

일본은 소비자물가지수, 도매물가지수, 기업대상서비스가격지수 모두 기본적으로 유의표본추출(purposive sampling)에 따른다. 도매물가지수와 기업대상서비스가격지수는 조사처에 회담의무가 없는 신고통계이고, 소매물가통계는 지정통계로 비협력자에 대한 벌칙이 존재하지만 현실적으로 현장에서 엄격하게 적용하기는 어렵다고 한다(일본은행, 2000). 2006년 12월 현재 167개의 시정촌, 657개의 조사지구, 530개의 품목, 787개의 상표, 230,000개의 가격을 수집하여 상당히 많은 양을 조사하고 있다. 특히 <표 3-7>에서와 같이 품목을 A, B, C, S, D, E 품목으로 나눈 다음, <표 3-8>에서 보듯이 각 품목에 대한 지역별 배분을 하고 있다. 각 품목에 대한 가장 대표적인 상품을 정하고 각 지역에서 가장 대표적인 상점(품목의 가장 큰 판매액을 갖는 상점)을 선정하고 있다(일본 통계청 홈페이지).

〈표 3-7〉 일본의 조사품목 구분

품 목	내 용
A 품목	식료나 주류, 가사잡화 등 거주지 부근에서 구입하는 품목
B 품목	피복이나 가전제품 등 상업집중지나 대형점포 등에서 구입하는 품목
C 품목	조미료나 교양오락용품 등 지구간 또는 점포간의 가격차가 비교적 작은 품목
S 품목	영화관람료나 운송료 등 조사지구를 마련하지 않고 전역으로부터 조사하는 품목
D 품목	수도요금 등 공공요금이나 입원비용 등 가격요금이 거의 균일한 품목
E 품목	전기세나 철도운임 등 전국 또는 지방적으로 가격요금이 균일한 품목

자료: 일본 통계청 홈페이지(<http://www.stat.go.jp>).

〈표 3-8〉 일본의 도시별 조사품목수

분류	도시			
	현청소재 도시	50,000 이상 도시	50,000 이하 도시	읍면
A 품목	148	113	102	81
B 품목	137	63	50	25
C 품목	98	38	34	19
S 품목	34	25	22	21
D 품목	39	37	37	37
E 품목	77	15	10	7

자료: 일본 통계청 홈페이지(<http://www.stat.go.jp>).

소매물가조사와는 별개로 1997년에 상업통계조사(Census of Commerce, 1994)를 모집단 정보로 하고 조사점포의 판매장 면적 등을 기준으로 점포를 확률추출하여 조사된 전국물가통계조사가 시행된 바 있다(이전에는 유의추출이었으나 이후 확률추출로 전환). 5년 주기로 실시하는 전국물가통계조사는 물가지수 작성 목적이 아닌 점포간, 지역간 가격변동이나 가격분포 등 소매가의 횡단면적(Cross-sectional) 현황 파악을 위해 설

계된 것이다(Sato, 1998). 전국물가통계조사를 소매물가조사와 별개로 시행하고 있는 점과 전자는 확률추출방식을 택하고 후자는 유의추출방식을 취하는 것은 시사하는 바가 크다 하겠다.

전국물가통계조사(1997)의 표본추출의 첫 단계는 지역추출로서 총 자치구 3,233개 중 인구 10만 이상 되는 221개 모든 도시는 무조건 포함하고 인구 10만 미만 중 450개 시정촌을 확률표본으로 뽑는다. 2단계에서는 선택된 지역에서 매장면적이 450m² 이상이면 모든 대상처(24,000)를 포함하고 미만인 경우는 조사구(survey area)를 임의추출한다. 3단계에서는 추출된 각 조사구로부터 40개의 판매처를 임의추출한다. 각 조사구의 판매처는 31개의 부문(division)으로 분류된다.

3. ILO의 비확률표본추출 채택근거

많은 나라들이 현실적으로 비확률표본추출방식을 취하고 있는데 실제로 CPI 작성에 있어서 비확률표본추출방법의 선택을 정당화하는 많은 근거들이 있다. ILO의 소비자물가지수에 대한 매뉴얼(2004)에서는 소비자물가조사의 표본추출방법에 대하여 언급하는 가운데 소비자물가조사에서의 가격조사 대상처의 비확률표집방법에 대한 채택 근거를 제시하고 있다. 요약하면 다음과 같다.

- ① 가용한 표본추출 틀이 없다: 상품차원에서 추출 틀이 없다(그러나 판매처 차원에서는 사업체명부나 전화주소록 등이 이용될 수 있다).
- ② 비확률표본추출로부터 초래되는 편향(bias)은 크지 않다: Dalen (1998)은 품목지수에서 추정편향이 나타나지만 총계시 소거되어 편향은 거의 없다고 지적하였으며, De Hann(1999)은 비확률표본추출이 PPS표본추출보다 더 작은 평균제곱오차(MSE)를 준다고 보고하고 있다.
- ③ 샘플은 계속적으로 조사 가능함을 보증해야 한다: 설령 확률표본추출이 가능할지라도 조사 불가능한 표본에 대한 대체로 인해 편향(bias)을 초래하게 된다.
- ④ 기준년도의 확률표본이 현재 시점에서의 적절한 확률표본이 아니

다: 확률표집에 의한 편향(bias)에 대한 방지는 확률표집 후에 나타나는 비확률적인 대체로 인해 크게 손상된다.

- ⑤ 자료 수집은 반드시 자료수집자가 있는 곳에서 이루어져야 한다: 새로운 표본이 뽑힐 때마다 그곳으로 고용되고 파견되는 것은 어렵고 비용이 많이 든다.
- ⑥ 표본크기가 너무 작다: 최종표집단계에서 1~5개 단위의 확률표본이 추출되는데, 작은 표본크기는 대표성이 좋지 않을 수 있다.
- ⑦ 표본추출결정은 조직의 하부에서 이루어져야 한다: 가격 수집자가 통계학에 잘 훈련되어 있지 않는 한, 현장에서 확률표집을 수행하기에는 가격수집자에게 너무 어렵다.

이처럼 표본조사에 있어서 확률추출방법이 과학적인 방법이기든 하나 사업체를 대상으로 하는 표본조사의 경우 확률표본을 고수하기에는 제약이 많다. 더욱이 계속조사가 가능하여야만 하는 소비자물가조사인 경우에는 확률표본추출이 오히려 비표본오차의 증가를 가져올 수 있는 등 실행상의 문제가 있어 이를 고려하지 않을 수 없다.

4. 연구동향

소비자물가조사의 표본추출방법에 대한 논란은 최근의 논문에서도 찾아볼 수 있다. Dorfman(2006)은 확률표본추출과 유의표본추출의 효율성 비교와 관련하여 시리얼(cereal)에 대한 스캐너 자료로부터 구매와 가격의 작은 모집단을 구축하고 미국의 확률접근법과 영국의 유의표본추출을 비교하였는데, 영국의 유의표본추출이 전체적인 정확성에서 우위에 있음을 보여주었다. 비확률표본추출방법은 본질적으로 확률표본추출방법에 비해 방문할 대상처는 적고 대상처 내의 가격수집에 있어서도 노동력이 적게 투입되므로 주어진 예산에서 미국의 확률표본추출에 비해 효율적이라고 보고하고 있다. 반면 이탈리아 통계청의 Biggeri & Falorsi(2006)는 CPI 작성시 비확률표집이 유효하기는 하지만 일부 경우에는 논란의 여지가 있다고 판단하여 확률표본추출을 구현하기 위하여 최근 사업체명부를 통한 확률추출법을 연구하고 있다.

한편 최근 소비시장이 대형화, 전문화되어 가고 있는 상황에서 미국, 영국, 일본 등 선진국에서는 스캐너(scanner) 데이터 또는 포스(Position Of Sales: POS) 데이터를 활용하여 소비자물가지수를 작성하고자 하는 연구가 진행되고 있다. 스캐너 데이터는 판매시간과 장소에서 바코드 리더에 의해 기록되는 판매시점에서 생성되는 자료이다. 이러한 데이터는 모든 상품에 대해 점차적으로 넓혀져 가고 있는 추세이며 그 잠재적인 활용가능성은 충분히 고려할 만하다. 위에서 언급한 표본추출의 효율을 비교할 때나 헤도닉 방법을 포함하는 품질조정의 개선된 방법에도 그 적용의 범위가 넓혀지고 있어(ILO, 2004) 우리나라도 앞으로 이에 대한 연구가 시급히 요구된다.

제5절 결론 및 시사점

ILO는 20년 전부터 가격조사대상의 확률표본추출을 제안하였으나 현재에도 대부분의 나라가 기본적으로 유의표본을 사용하고 있다. 이는 현실적으로 확률표본이 소비자물가조사에 적용되기 어려운 문제임을 단적으로 보여주고 있다. 확률표본추출방법이 표본의 객관성 확보 면에서 과학적 합리성이 있음은 분명하다. 그러나 현실적으로 비용이 많이 들 뿐 아니라 정확한 모집단 자료수집이 어렵고, 적절한 비확률표본추출법인 경우는 바이어스가 크지 않은 것으로 보고되고 있으며, 기준시점과 비교 가능하고 계속조사가 가능해야 한다는 특수상황에서 확률표본추출보다 효율적일 수 있다.

또한 사업체대상의 표본조사는 조사처의 협력이 대전제가 되어야 하므로 표본설계가 잘 되었다 할지라도 조사처의 협력이 절대 필요하다는 점을 간과해서는 안 될 것이다. 확률표본일지라도 추출된 조사처의 협력을 얻을 수 없는 경우는 다른 조사대상처를 선정하지 않을 수 없으므로 이때 대체 바이어스를 초래하게 되며, 표본 설계가 아무리 잘 되었다고 하더라도 가격조사 실시상의 문제로 인해 조사협조가 잘 안 된다면 조사결과의 품질을 보장할 수 없다.

한편 미국과 영국 등은 확률표본추출방식을 지향하고 있다. 미국과 영국은 특성이 다양한 지방으로 구성되어 있고 상품과 서비스 매장이 매우 전문화되어 있는 나라로 그 나라의 CPI 작성 환경과 조사 인프라는 다른 나라와 구별된다고 볼 수 있다. 미국은 구입처조사(Point of Purchase Survey: POPS)를 통해, 영국에서는 조사자가 지역의 상가목록을 만들어 쓰는 등 보조 자료를 이용하여 확률표본추출을 지향하고 있으나 이렇게 작성된 목록도 완전한 표본추출틀이라 보기는 어렵다.

그 밖의 여러 나라에서는 비확률표본추출을 채택하고 있으며 그 나라의 조사 인프라에 따라 나름대로의 표본설계를 하고 있음을 보았다. 이들 사례에서 우리나라도 소비자물가조사의 대상표본 검토위원회의 활성화와 수집의 편리성에 의해 대상처가 선택되지 않도록 하는 시스템 마련, 그리고 지역정보에 밝은 지방청 직원의 의견 수렴 및 지식정보의 활용 등을 검토할 필요가 있다.

우리나라의 소비자물가조사도 비확률표본추출인 유의표본추출방식이기는 하나 판단추출의 단점을 보완한 할당추출에 가까운 유의표본이다. 즉, 조사자가 마음대로 조사대상처 수를 배정하는 것이 아니라 매출액, 상점의 유형, 계속조사 가능성, 지역과 소비패턴의 변화 등을 고려하여 배분되고 있다. 또한 조사대상처 선정시 조사자가 편리한 대로 표본을 선택하는 것이 아니고 통계자료와 지역정보를 근거로 선정한다. 따라서 이러한 노력을 알릴 필요가 있고 이들 통계자료와 지역정보의 체계적인 공유 및 개발은 현행 유의표본추출방식에서 표본의 대표성과 물가지수의 신뢰도를 높일 수 있는 한 방안이 될 수 있을 것이다.

앞에서 각국의 물가조사 표본추출에 대한 현황과 사례를 검토하여 보았으나 소비자물가지수 산정의 체계를 완전히 바꾸는 확률표본추출방식의 도입에 대한 검토는 단기간의 본 연구로서는 한계가 있다. 이를 위해서는 대상처 선정에 대한 보다 심층적인 현황분석과 표본추출틀 작성연구, 분산추정연구 등 장기적인 연구가 필요하다. 그러나 현재 우리나라의 CPI 작성 인프라와 해외사례연구에 의하면 단기간에 확률표본으로 가거나 큰 틀을 바꾸기는 위험이 크고 어려울 것으로 판단된다. 만약 확률표본추출의 시험적 적용이나 유의표본추출과의 비교가 필요하다면 일부 품목에 대한 부분적 적용을 고려해 볼 수는 있을 것이다.

이상에서 살펴보았듯이 지역, 상품, 대상처의 완전확률표본추출의 구현은 복잡하고 비용이 많이 들어 물가의 변동을 시의성 있게 공표해야 하는 점에서 비효율적일 수 있다. 한편 비확률표본추출일지라도 구매처조사, 소비패턴 등을 고려하여 표본이 선택된다면 표본의 대표성 문제를 상당히 극복할 수 있다. 현 상황에서는 확률표본이나 유의표본이냐의 문제보다는 구매처조사, 소비패턴조사 등에 근거하여 표본을 선정하는 것이 중요하며 업체별 매출액, 품목별 구매처 등의 정확한 소비실태를 파악하기 위해 업체별 매출액 분석 및 구매처조사가 선행되어야 한다. 즉, 최종 표집단계에서의 소수개의 확률표본을 지향하기보다는 조사대상처 수의 합리적인 배정에 치중할 필요가 있다고 판단된다. 이를 위해 사업체기초통계조사의 사업체정보를 활용하거나 도시가계조사에 분기별로 구매처 정보 등을 얻을 수 있도록 조사표를 설계하는 것도 하나의 방안이 될 수 있을 것이다.

참고문헌

- 인천통계사무소(2006), “소비자물가지조사 개선방안”, 물가통계팀.
- 통계청(2005), 「국가통계 바로알기」.
- 통계청(2006), “2005년 기준 소비자물가지수 개편결과”, 통계청 보도자료(12.22.).
- _____ (2007a), 「통계행정 편람」.
- _____ (2007b), “소비자물가지조사 조사대상처의 대표성 제고방안 -조사대상처의 산업분류별 구성비 변경(안)-”, 내부자료.
- 日本銀行調査統計局(2000), “物価指數を巡る諸問題”, 「日本銀行調査月報」, 8月号掲載予定資料.
- Biggeri, L. and P.D. De Falorsi(2006), “A Probability Sample Strategy for Improving the Quality of the Consumer Price Index Survey Using the Information of the Business Register”, UN Statistical Commission and Economic Commission for Europe, Working Paper No.12 (Group of Experts on Consumer Price Indices, Geneva, 10-12 May 2006).
- Boon, M.(1997), “Sampling Design in Constructing Consumer Price Indices: Current Practices at Statistical Offices”, UN Statistical Commission and Economic Commission for Europe, Joint ECE/ILO Meeting on Consumer Price Indices, Geneva, 24-27 November 1997.
- De Hann, J., E. Opperdoes, and C.M. Schut(1999), “Item Selection in the Consumer Price Index: Cut-off versus Probability Sampling”, *Survey Methodology*, Vol.25, No.1.
- Doherty, M.(1994), “Probability versus Non-Probability Sampling in Sample Surveys”, *The New Zealand Statistics Review*, March 1994, pp.21-28.

- Dorfman, A.H., J. Lent, S.G. Leaver, and E. Wegman(2006), "On Sample Survey Designs for Consumer Price Indexes", *Survey Methodology*, Vol.32, No.2, pp.197-216.
- Fenwick, D. and A. Ball(2001), "Sampling in Consumer Price Index", Sixth Meeting of the International Working Group on Price Indices, Canberra, Australia, 2-6 April 2001.
- International Labor Office(2004), *Consumer Price Index Manual: Theory and Practice*.
- Loretto, R.(1997), "Methods of Outlet Selection and Weighting in the Consumers Price Index", Consumer Price Index Revision Advisory Committee 1997, Statistics New Zealand.
- Office for National Statistics(2007), "Consumer Price Indices Technical Manual", London.
- Sato, T.(1998), "Outlet Sampling Method on the 1997 National Survey of Prices", *Proceedings of the Fourth Meeting of the International Working Group on Price Indices*, pp.275-282.
- U.S, Department of Labor(2007), "Chapter 17, The Consumer Price Index", *BLS Handbook of Methods*.(<http://www.bls.gov/pub/hom/pdf/homch17.pdf>)