

통계청 정책연구 용역 최종보고서

인터넷 자원자 표본에 대한 추정방법 연구

2009. 12. 15

조성겸(충남대학교, 언론정보학)

허명희(고려대학교, 통계학)

제 출 문

제 출 문

통계청장 귀하

본 보고서를 “인터넷 자원자 표본에 대한 추정
방법 연구 ” 연구용역 과제의 최종 연구결과물로 제출
합니다.

2009년 12월 15일

충남대학교 산학협력단장

연구진

책 임 연 구 원 조성겸(충남대학교, 언론정보학)

연 구 원 허명희(고려대학교, 통계학)

연 구 보 조 원 김량희, 나혜민, 문상철, 정지성, 전보미

요 약 문

최종결과보고서 요약문

연구과제명	인터넷 자원자 표본에 대한 추정방법 연구
주 제 어	자원자 표본, 인터넷 조사
연구 기간	2009. 4.21- 2009.12.15
연구 기관	충남대학교 산학협력단
연구진구성	책임 연구원: 조성겸(충남대학교, 언론정보학) 공동 연구원: 허명희(고려대학교, 통계학) 연구 보조원: 김량희, 나혜민, 문상철, 정지성, 전보미
<p>본 연구는 인터넷을 이용한 조사방법의 기반을 마련하기 위한 것이다. 인터넷 조사는 자원자 표본을 이용하기 때문에 표본의 대표성을 확보하기 어렵다. 이로 인해 인터넷 조사방법이 기존의 조사방법이 가진 한계를 극복할 수 있는 새로운 조사방법임에도 불구하고, 인터넷 조사방법이 널리 활용되지 못하고 있는 실정이다. 이러한 대표성 문제에 대한 해결방안을 모색하기 위해 본 연구는 다음과 같은 작업을 수행하였다.</p> <p>첫째 기존의 인터넷 및 전화조사의 표본 설계 방법을 검토하였다. 표본 설계는 크게 편의 표집 방법을 사용하는 경우와 확률표집 방법을 사용하는 경우로 구분되는데, 각각의 표본 설계가 인터넷 조사라는 맥락에서 갖는 한계와 장점을 검토하였다.</p> <p>둘째로 비확률적 표본에 해당되는 인터넷 자원자 표본의 편향을 제거하기 위한 통계적 모형을 구축하고, 이모형을 다양한 평가변수에 적용하여 과연 인터넷 자원자 표본의 편향이 제거될 수 있는지 살펴보았다. 이러한 모형의 구성 및 검증과정에서는 2009년 통계청 사회조사에서 수집한 온라인 및 가구방문조사 자료를 이용하였다.</p> <p>인터넷 조사방법의 문제점을 극복하려는 지금까지의 시도는 여러 가지가 있지만 크게 확률표집 방법을 이용하는 경우와 자원자 표본을 이용하는 경우의 두 가지로 대별될 수 있다.</p>	

전자는 Knowledge Networks panel(미국), LISS panel(네덜란드) 등과 같이 면대면 및 RDD 방법을 이용해 응답자 패널을 구성함으로써 대표성 문제를 해결하고자 하는 방법이다. 비용이 많이 소요되고, 초기 응답패널 모집과정에서 응답률이 낮을 수 있다는 문제가 있지만 대표성 문제가 다른 인터넷 조사방법에 비해 적기 때문에 다양하게 활용되고 있다.

자원자표집 방법을 이용하는 경우는 모집된 자원자 표본의 편향을 통계적 기법으로 제거하는 방법이다. 미국의 해리스 인터랙티브의 경우가 이 방법을 사용하는 대표적인 회사이며, 이 방법을 검증 또는 활용하는 연구가 폭넓게 이루어졌다. 그렇지만 이 방법을 성공적으로 적용하기 위해서는 준거 조사 및 기준변수 등 해결해야 할 문제점이 아직 남아있다. 아직까지는 공식적 통계에 활용되지는 않고 있다.

본 연구는 이 중에서 두 번째 방법에 해당되는 가중 모형을 구성하고 검증해 보았다. 모형 구성에는 9개의 기준변수가 사용되었는데 이것은 (시도, 지역크기, 성별, 연령, 교육정도, 소비생활만족도, 소득분배 견해, 일반신문 보기, 인터넷신문 보기) 등이 있다. 2009년 웹조사 및 가구방문 조사 자료를 이용하여 이모형의 유용성을 검증한 결과 다음과 같은 결과를 얻었다.

- . 가중모형은 웹조사의 편향을 감소시킨다.
- . 그러나 감소폭은 그다지 크지 않다. 편향감소폭이 원래 편향의 1/2에 미치지 못하는 경우가 더 많다.
- . 가중모형의 편향감소 효과는 소득/소비, 자원봉사 영역에서는 나타나지 않고 있고, 반면에 레저, 매체이용행태, 삶의 질 만족도 등의 영역에서는 나타나고 있다. 즉 가중모형의 효과는 적용분야에 따라 달라진다.
- . 가중모형 편향감소 효과는 10대와 20대의 경우에 30세 이상 집단보다 크게 나타난다. 즉 웹조사 자료의 대표성이 젊은 층에서 더 높게 확보될 수 있다.
- . 모형 구축방법으로 림가중치(rim weighting)와 로짓 성향분석 (logit propensity weighting) 방법간에는 차이가 없다

이와 같은 본 연구의 결과를 토대로 다음과 같은 결론을 내릴 수 있다.

첫째 성향가중 모형은 유용하다. 본 연구에서 편향 감소폭이 크지 않았지만 이것은 앞으로 좀더 효과적인 기준변수를 선택함으로써 개선될 수 있을 것으로 본다.

둘째 가중모형에 포함되는 기준변수는 평가변수에 따라 즉 적용분야에 따라 달리 선택되어야 한다. 본 연구에서 모형의 효과는 평가변수의 유형에 따라 달라졌다. 따라서 정치분야, 과학분야, 소비행동 분야 등과 같이 특정 분야별로 모형을 구축하는 것이 보다 쉽게 모형을 구축할 수 있는 방법일 것이다.

셋째 전국민을 대상으로 하는 조사보다는 젊은 층 특히 인터넷 사용도가 높은 층에서 이러한 모형의 타당성이 높을 것으로 보인다. 즉 인터넷 조사를 10대나 20대에 대한 조사에만 적용하는 것도 가능할 것이다.

본 연구의 결과는 인터넷 자원자 표본의 편향을 제거하는 방법이 비록 지금 즉시 공식적 통계에서 활용되기에는 한계가 있지만, 이러한 한계는 기준변수 선정 및 패널모집 방법 등에 대한 후속 연구를 통해 해결될 수 있다는 점을 보여준다. 또한 통계청에서 수행하는 조사에서 인터넷 응답자를 모집하고 이들의 특성을 파악하는 연구에도 크게 도움이 될 것이라고 본다.

앞으로 면접원 가구방문 등 전통적 조사방법의 문제가 더욱 심각해질 것으로 예상해 본다면, 이러한 인터넷 조사방법의 실용화에 관한 연구는 더욱 절실히 필요하다고 본다.

차 례

제 1 장 연구 개요	1
제 1 절 연구의 목적	1
제 2 절 연구내용 및 방법	2
제 2 장 기존연구 검토	5
제 1 절 확률표본 방법	5
제 2 절 비확률 표본과 편향 제거 방법	12
제 3 절 인터넷 및 전통적 조사방법의 병행 방안	15
제 3 장 방법론과 모형	17
제 1 절 인터넷 조사 표본 추출	17
제 2 절 면대면 조사 자료와 인터넷 조사 자료의 비교	21
제 3 절 성향점수모형 가중법에 의한 성향 보정	26
제 4 절 rim 가중법에 의한 성향 보정	33
제 5 절 주요 평가문항의 성향보정 결과	35
제 4 장 성향보정 방법의 유용성 검토	42
제 1 절 검토방법	42
제 2 절 가중의 효과	43
제 3 절 변수 유형 및 연령 집단별 비교	51
제 4 절 모형 효과의 크기	54
제5장 요약 및 결론	55
제 1 절 요약 및 제한점	55
제 2 절 결론	57
제 3 절 제언	59

참고문헌	63
<부록 1> PS rim 성향가중법 결과	66
<부록 2> 문항별 응답비율 산정방법	86

표 차례

<표 1> 상주인구 대비 자원자 패널의 상대적 비	18
<표 2> rim 가중법 적용 기준 비율	18
<표 3> 상주인구 대비 추출 표본의 상대적 비	20
<표 4> Web 표본에 대한 가중치 산출 과정	33
<표 5> 성향점수모형(PS)에 의한 가중법에 의한 성향보정의 결과	37
<표 6> 성향점수모형(PS)에 의한 가중법에 의한 성향보정의 결과: 보정기준 변수를 7개로 줄인 경우	38
<표 7> rim 가중법에 의한 성향보정의 결과	39
<표 8> rim 가중법에 의한 성향보정의 결과: 보정기준 변수를 7개로 줄인 경우	40
<표 9> 유형별 평가변수	43
<표 10> 웹조사결과에 성향가중치를 적용했을 때의 결과	45
<표 11> 성향가중 적용후의 웹조사결과와 기준조사 결과의 차이	51
<표 12> 문항유형별 성향가중 효과	52
<표 13> 연령집단별 성향가중의 효과차이	54
<표 14> 성향가중 모형의 효과 크기	54

그림 차례

<그림 1> 인터넷 조사 표본의 rim 가중치 분포	22
<그림 2> Web 표본 성향조정 가중치의 분포	32
<그림 3> Web 표본의 성향조정 가중치 대 최종 가중치 분포	32
<그림 4> 인터넷 조사 표본의 최종 가중치 분포	33

제 1 장 연구 개요

제 1 절 연구의 목적

통계청의 표본조사는 대부분 현장의 가구를 직접 방문하여 조사하는 면접 타계식을 실시하고 있으나, 최근 1인 가구(20%)와 맞벌이 가구(25%) 증가로 인해 부재중인 가구가 증가하고 있고, 오피스텔 및 아파트 공동현관에 무인 경비시스템을 설치하는 경우 역시 늘고 있어 점점 면접원이 접근하기 어려운 가구가 많아지고 있다. 또한 면접원에 의한 사생활 침해를 우려하는 경우 역시 늘고 있다.

본 연구는 인터넷을 이용한 조사방법의 기반을 마련함으로써 이러한 표본 조사의 어려움을 감소시키고자 한다. 인터넷 조사방법의 활용성은 높게 평가되고 있지만 표본의 대표성에서 문제가 지적되고 있다. 지금까지의 연구들은 확률표집 방안을 도입하거나, 자원자 표본을 이용하되 그로 인한 편향을 통계적 방법을 사용해서 제거하는 방안 등을 이용하여 인터넷 조사방법의 표본 대표성 문제를 해결하고자 하였다. 본 연구는 이러한 각각의 방법이 갖는 가능성과 한계점을 기존 연구결과들을 토대로 검토하였다.

둘째로 본 연구는 또한 인터넷 자원자 표본이 갖는 편향을 제거하는 통계적 모델을 구성하고 이를 실제 데이터를 이용해서 검증해 보았다. 셋째로 기존의 연구 및 본 연구의 모델검증결과 등을 토대로 인터넷 조사의 활용성을 높이기 위한 방안을 제시하였다.

제 2 절 연구내용 및 방법

가. 기존연구 검토

인터넷 조사방법에 대한 기존의 연구결과에서는 아래와 같은 사항들을 살펴보았다.

- 인터넷 조사의 표본설계 방법

기존 국내외 인터넷 및 전화조사의 표본 설계 방법을 검토하여 그 방법상의 특징과 문제점 등을 검토한다. 인터넷을 이용한 조사의 표본설계는 크게 편의 표집 방법을 사용하는 경우와 확률표집 방법을 사용하는 경우로 나뉠 수 있는데, 각각의 표본 설계가 인터넷 조사라는 맥락에서 갖는 한계와 장점 등을 검토하고자 한다.

- 표본 편향 제거를 위한 최근 기법 검토

표본의 편향을 제거하기 위한 기법들을 찾아서 각 기법들의 한계와 장점을 살펴본다. 특히 자원자 표본을 이용하는 경우 그 편향을 가중법을 사용해서 극복하는 시도가 이루어져왔다. 이러한 통계적 방법들을 검토하고 본 연구의 모형 구축에 적용하였다.

나. 가중법 모형 구축 및 유용성 검토

인터넷 자원자 표본의 편향을 제거하는 가중 모형을 구축하고 그것을 검증하였다.

- 기본 모형 구축

먼저 인터넷 자원자 표본이 갖는 편향을 제거하는 기본 가중법 모형을

구축하였는데, 이 과정에서 변수의 수 및 유형, 가중치 산출 방법 등을 다양하게 검토한 다음, 이러한 결과를 토대로 최적의 모델을 구성하였다.

- 가중법 적용 및 타당성 검토

구축된 기본 모형을 다양한 평가변수에 적용하여 과연 인터넷 자원자 표본의 편향이 제거될 수 있는지 그리고 이러한 모델의 유용성과 한계는 무엇인지 찾아보았다.

다. 인터넷 조사의 활용방안 및 발전방안 제시

- 활용방안 제시

기존 연구 및 본 연구의 모델 검증결과를 토대로 인터넷 조사방법이 사회조사에 활용될 수 있는지 그리고 활용된다면 어떤 절차로 어떻게 활용될 수 있는지에 대한 가이드라인을 제시하였다. 즉 인터넷 표본의 편향을 최소화하기 위한 인터넷 표본 설계방안, 그리고 수집된 자료에 가중법을 적용하는 구체적인 절차를 제시하고자 한다.

- 기타

또한 인터넷 조사방법의 유용성을 높이기 위해서, 활용 방법별로 그 한계와 적용범위 그리고 추후 요구되는 연구의 방향을 제시하였다.

라. 연구방법

- 기존문헌연구

최근에 이루어진 인터넷 조사방법에 관한 논문 중에서 조사과정 및 조사결과에 대한 기술이 비교적 상세하게 이루어진 논문들을 선정하여 검토하였다.

- 모델 구축 및 검증

성향지수방식을 이용하여 모델을 구축하고, 이러한 모델을 2009년 통계청 사회조사에서 수집된 온라인 및 가구방문 조사 자료를 이용하여 모델 추정 및 검증을 실시하였다.

제 2 장 기존연구 검토

인터넷을 이용한 조사가 갖는 오차요인은 크게 두 가지로 구분된다. 첫째는 포함률 오차 (coverage error)로써, 인터넷에 접근할 수 있는 응답자로 표본이 제한됨으로써 발생하는 오차이다. 둘째는 응답자의 자기선택 오차 (self-selection error)로서 표본 추출이 응답자의 자발성에 의존함으로써 발생하는 오차이다. 즉 다른 전통적인 조사방법보다는 응답자의 조사에 대한 협조 의사가 표본 선정과정에서 중요한 역할을 하게 된다는 것이다. 이 중에서 첫 번째의 포함률 오차는 인터넷의 활용이 확대되면서 점차 감소되고 있다. 따라서 두 번째 요소인 자기선택 오차가 인터넷 조사에서 가장 중요한 문제가 된다.

이러한 자기선택 오차를 해결하는 방법은 크게 세가지로 나뉘는데 첫째는 응답자의 선정과정에서 전통적인 조사방법에서 사용하고 있는 확률표집 방법을 도입하는 방법이다. 이렇게 함으로써 자기선택 및 포함률 등으로 인한 오차를 감소 또는 제거하고자 한다. 둘째는 통계적 모델을 이용해서 오차를 제거하는 방법이다. 성향지수 가중법 등이 여기에 해당된다. 셋째는 전통적인 조사방법과 인터넷 조사방법을 혼용해서 실시하는 방법이다. 이 때 전통적 조사방법에 의한 표본 프레임의 가급적 활용함으로써 앞서 지적한 오차 문제를 극복하고자 한다.

제 1 절 확률표본 방법

가. 응답자 패널 구축방법

인터넷 조사에서 표본 추출을 전통적 조사방법에서와 같이 확률적으로 실시할 수만 있다면 포함률이나 자기선택성에 따르는 오차는 발생하지 않게 된다. 그러나 인터넷의 경우 이러한 확률표집을 실시할 수 있는 표본 프레임

이 존재하지 않고, 따라서 인터넷 상에서 표본을 확률적으로 추출하는 것은 아직 가능하지 않다.

일부 연구에서는 전통적인 표본프레임 즉 거주지 주소 또는 전화번호와 같은 프레임을 이용해서 표본을 추출한 다음, 조사를 인터넷으로 실시하는 방법을 사용하였다. 이 방법은 설문지 전달 및 회수는 인터넷으로 하지만 표본추출 자체는 전통적 방법으로 이루어지는 것이다. 물론 표본에 추출된 가구나 개인이 인터넷에 접속할 수 없는 경우, 인터넷 접속을 제공하게 된다.

이러한 방법은 다음과 같은 문제점을 지니게 된다. 첫째 전통 표본 프레임이 갖는 문제점을 그대로 갖게 된다. 예컨대 전화번호나 거주지 주소를 이용하는 데서 발생하는 문제는 그대로 있게 된다. 둘째 표본이 추출된 다음 인터넷으로 응답을 요청할 경우 거절할 가능성이 높다. 실제로 이러한 방법을 사용한 미국의 Knowledge Networks의 경우도 이 과정에서 거절율이 높은 것으로 나타났다. (Taylor 등 2009).

그러나 이 방법은 다음과 같은 장점을 갖는다. 우선 인터넷 조사방법의 포함률이나 자기 선택에 따른 한계를 극복할 수 있다. 전통적인 접근방법을 이용해 면접원이 설득노력을 할 경우, 자기선택성에 따른 오차를 감소시킬 수 있고, 또 인터넷 접속을 제공해 줌으로써 포함률에 따른 오차도 감소시킨다. 아울러 전통적인 조사방법과 비교해 보면, 설문지 전달 및 응답에 시간제한이 없다는 장점이 있다. 따라서 면접원이 접촉하기 어려운 응답자들이 표본에 포함될 수 있다. 또한 응답자 패널이 일단 구축되면 비교적 적은 비용과 짧은 시간에 고품질의 조사를 수행할 수 있게 된다.

이와 같은 방법을 적용하고 있는 패널 중에서 미국의 Knowledge Networks 패널이 널리 알려져 있다. Taylor 등의 연구(2009)는 Knowledge Networks 패널을 이용한 조사와 다른 전통적인 조사를 비교하였다.

이 연구는 환경문제에 대한 와이오밍 주민인식 조사를 전화, 우편 그리고 Knowledge Networks 패널을 이용한 웹조사의 세 가지로 실시하고 각각의 조사 모드가 어떻게 다른지 비교하였다. 이 연구의 경우 진실값을 모르기 때문에 어느 조사가 가장 정확했는지는 알 수 없었다. 응답의 패턴, 다른 변수와의 상관관계 등을 분석해서 각각의 조사방법이 갖는 특성을 비교하였다.

먼저 Knowledge Networks 패널을 이용한 웹조사의 경우는 1,162명, 그리고 RDD 방법을 이용한 전화조사는 1,273명, 우편조사는 904명이었고, 각각의 조사는 모두 확률표집 방법에 입각해서 이루어졌다. 전화조사의 경우 재접촉 시도는 15회 이상으로 하였다. 조사내용은 국립공원의 공기 질에 관한 것이었으며, 동일한 설문지를 사용했는데, 서베이 모드에 따라 각기 다른 형태의 편향이 나타나고 있었다. 전화/우편의 경우는 조사주제에 대한 관심도가 높은 사람들이 그리고 인터넷 조사의 경우에는 사무실에서 일하는 사람들이 많아지는 편향을 보였다.

그리고 조사결과에 대해 성, 연령, 교육수준 등의 인구적 속성을 이용한 가중치를 적용했을 때, 각 조사모드간의 차이는 사라지지 않았다. 즉 조사방법에 따라 각기 다른 조사결과가 나온 것이다. 물론 이 중에서 어느 것이 정확한 것인지는 알 수 없었다. 그렇지만 연구자들은 다양한 응답특성을 검토한 결과 웹패널 조사결과가 전화 및 우편조사와 비슷한 수준의 정확도를 보인다는 결론을 내렸다.

확률표집방법으로 수집된 Knowledge Networks panel에서 응답자를 추출해서 조사한 경우, 잘 설계된 전화 및 우편조사에 못지않은 자료를 얻을 수 있다고 한다. 그러나 Knowledge Networks 패널은 최초 모집과정에서 전화를 이용했고, 또 모집과정에서의 응답률이 낮았다는 한계를 지니고 있다. 웹조사의 응답률은 3.9%로 매우 낮았다. (전화는 15.6%). 따라서 이 연구는 표본의 대표성을 확립하기 위해서는 응답률을 높이기 위한 다양한 시도들이 필요하다고 지적되었다.

이 연구에서 주목할 만한 부분은 응답자 패널의 응답률이다. 일단 패널에 포함된 경우에는 높은 응답률을 보이지만 패널을 모집하는 과정에서는 낮은 응답률을 보였다. 이러한 조사방법의 성공은 패널 모집과정의 응답률을 높이는 것이 중요하다.

패널 모집과정에서 충분한 시간이 주어진다면 반복접촉 및 설득노력, 인센티브 제공 등의 방법으로 어느 정도 응답률을 높일 수 있을 것으로 본다.

나. 대응표본 추출방법 (폴리메트릭스)

또 다른 접근 방법으로는 폴리메트릭스(Polimetrix)에서 사용하고 있는 방법을 들 수 있다. 이 방법은 확률적 방법으로 표본을 추출한 다음, 각각의 추출된 대상자와 가장 근접할 것으로 추정되는 대응표본을 기존에 미리 구축된 응답자 패널 중에서 추출하는 것이다. 근접성 판단에는 인구적 속성을 포함한 몇 가지 기준 변수가 사용된다.

Hill 등(2007)의 연구는 폴리메트릭스 패널을 이용해 정치적 태도를 조사하였는데, 폴리메트릭스 방법을 이용할 경우 약간의 편향이 있게 되지만 대규모 표본을 이용하기 때문에 소규모의 엄밀한 확률표본 보다 조사결과의 정밀도 측면에서는 더 유리하다고 보고 있다.

이 방법은 확률적으로 응답표본을 구축하지 않고도 편향의 문제를 해결하는 또다른 방법을 보여준다. 즉 확률적으로 조사된 다른 조사 자료를 이용함으로써 응답자 패널을 확률적으로 모집하는 어려움을 피할 수 있다. 응답자 패널을 확률적으로 모집할 경우 응답률이 낮아진다는 문제점이 있는데, 이 방법은 이러한 문제를 피할 수 있다는 것이다.

이 방법의 핵심은 확률적으로 추출된 타겟 표본을 미리 선정한 다음, 그 타겟 표본과 유사하도록 자체 응답자 패널에서 표본을 추출한다는 점이다. 따라서 유사성 판단기준이 중요하게 된다.

이 방법도 특정 분야의 조사에서는 비교적 쉽게 적용해 볼 수 있을 것이다. 예컨대 선거조사 또는 보건분야 등과 같이 조사분야가 한정될 경우 기준 변수를 찾는 것이 비교적 어렵지 않을 것이다. 그러나 통계청 조사와 같이 사회 전반에 관해서 다양한 조사를 하는 경우에는 타당한 기준변수를 찾는 것이 어려울 것으로 보인다.

다. 확률표집 방법의 사례들

사례1. Knowledge Networks Panel

- 1998년 시작
- RDD 방법 사용
- 86% coverage rate , response rate 16%(Cooper 2008)

사례2. LISS Panel

- <http://www.lissdata.nl>
- Longitudinal Internet Studies for the Social Sciences (2006-2013), Tilburg University
- 전화와 면대면 조사로 모집(5천 가구 8천명).
- 네덜란드 국내 및 외국 학자 사용 가능(무료)
- 48% 참여율, 78-83% 응답률

사례3. The German General Social Survey

- 2006년 조사: 23.7% 응답률
- The Health and Retirement Survey:
- 2002년 응답자 대상 연구: 59%응답률 (Cooper 2008).

사례 4. American Life Panel

- http://www.rand.org/pubs/corporate_pubs/CP508-2005-11
- RAND가 유지 관리
- 1,500 표본 (40세 이상)
- 미시간대학 소비자태도에 관한 월례조사 응답자 중에서 모집
- 70-80% 응답률

라. RDD 표집방법

표본의 대표성을 확보하는 또 다른 방법으로는 RDD 전화조사 방법으로 응답자를 추출한 다음 실제 조사는 인터넷으로 하는 방법이 있다. 이 방법은 앞서 살펴본 응답자 패널 구성 방법과 거의 동일하다. 다만 패널구성 방법이 다목적 조사를 위해 구축하는 일반 패널이라고 한다면, 여기서는 특정 조사를 위해 응답자를 모집한다는 차이가 있다. 따라서 응답률은 범용패널 모집의 경우보다 높을 것이다.

이처럼 전화로 응답자를 모집한 다음 인터넷으로 조사하는 경우 첫째 전화조사 자체가 갖는 문제점을 그대로 갖게 된다. 즉 유선전화로 접근할 수 없는 경우는 여전히 접근이 어렵다. 둘째 모집과정에서 거절율이 높을 수 있다. 그리고 인터넷으로 응답하기로 약속한 응답자와 거절자간에 체계적인 차이가 있을 수 있다. 셋째 인터넷으로 응답하는 과정에서도 체계적으로 응답자와 비응답자간 차이가 발생할 수 있다. 그러나 장점으로는 전화로 질문하기 어려운 복잡한 설문을 할 수 있다는 점 그리고 면접원의 가구방문보다는 보다 저렴하면서도 빠른 시일 내에 조사를 할 수 있다는 장점이 있다.

이러한 점들을 검토한 연구로는 Heeren 등(2008)의 연구를 들 수 있는데, 이들은 알콜관련 행위에 관한 조사를 RDD에 기반한 인터넷 조사를 실시하고 이것을 43,903명을 대상으로 실시된 면대면 조사자료와 비교하였다. 이 면대면 조사자료는 이 분야에서 대표성이 인정되는 조사였다. 비교결과 RDD에 기반한 인터넷 조사결과는 면대면 조사결과와 큰 차이를 보이지 않아 타당성이 높은 것으로 나타났다. 그리고 응답자 모집과정에서 인터넷조사에 응할 것을 거절한 응답자들을 따로 전화를 이용해서 조사한 다음, 인터넷조사에 응답한 경우와 비교하였는데 여기에서도 큰 차이가 없었다. 이들의 연구결과는 RDD 방법과 인터넷 조사를 결합한 방법의 유용성을 보여주고 있다. 특히 이 연구는 면접원 접근에 한계가 있는 조사의 경우 전화 및 인터넷을 이용해서 조사할 수 있다는 점을 보여주었다는 점에서 의미가 크다.

이상에서 살펴본 방법들은 기존의 확률표집 방법과 인터넷 조사방법을 다양한 방식으로 결합해서 인터넷 표본이 갖는 한계를 극복하고 있다. 이 과정에서 전통적 조사방법이 갖는 한계를 인터넷의 장점을 활용해서 극복하게 되기도 한다.

한국의 경우 이러한 시도들이 아직 이루어지지 않았다. 그것은 아직 한국의 경우 민간부분의 조사에서 확률표집의 중요성이 충분히 인식되지 않고 따라서 확률표집 방법이 전통적 조사방법에서 조차 적용되지 않는 경우가 많기 때문이다. 인터넷 조사방법을 활용하면서 이러한 확률적 방법을 도입하는 시도는 이루어지지 않고 있다. 한국에서 사용하는 인터넷 응답자 패널은 확률표집 방법으로 구축된 경우는 거의 없다.

또 다른 이유로는 한국의 경우 조사시장의 규모가 크지 않기 때문에 많은 비용을 들여서 응답자 패널을 구축하기가 어렵다는 점이다. 응답자 패널을 구축할 경우 비용이 많이 소요되기 때문에 다양한 조사에서 활용되어야 할 것이지만 그것이 불확실하기 때문에 민간 부분에서 이러한 패널 구축에 많은 비용을 투입하게 될 동기가 약하다.

제 2 절 비확률 표본과 편향 제거 방법

인터넷 표본의 편향을 통계적 추정방법을 이용해 제거하는 방법도 널리 이용되고 있다. 이 방법은 비록 비확률적으로 추출된 표본의 경우라 할지라도 인구적 속성 및 심리적 라이프스타일 속성 등에 기반한 가중치를 적용함으로써 대표성을 가질 수 있을 것으로 보는 것이다.

따라서 이러한 방법에서 관건은 어떤 라이프 스타일 속성을 이용하여 가중치를 산출해야 하는가? 그리고 과연 이러한 가중치로 표본의 대표성이 확보될 수 있는가?라고 볼 수 있다.

구체적으로 이러한 가중치를 구하는 과정 및 그 원리에 대해서는 이미 많은 국내외 문헌에서 소개되고 있다. 예컨대 Hahs-Vaughn and Onwuegbuzie(2006)의 논문도 가중치의 산출과정을 상세히 설명하고 있다. 그러나 어떤 변수들을 가중치 추정과정에 사용해야 하는가에 대해서는 아직 명확한 연구결과가 제시되지 않고 있다. 최초로 이러한 방법을 적용한 해리스 인터랙티브의 경우 선거예측을 전제로 적용했는데, 이들이 사용한 문항을 논문에서 찾기 어렵다.

이 방법은 비확률 표본의 대표성을 확보해 준다는 점에서 그리고 인터넷 조사의 문제점을 해결해 줄 수 있다는 점에서 많은 관심을 끌었고 결과적으로 이들 방법을 활용한 조사나 연구가 많이 이루어진 편이다. 왜냐하면 이 방법은 앞 절에서 살펴본 확률표집을 이용하는 방법보다 비용이나 시간 면에서 훨씬 유리하기 때문이다.

그러나 이들 방법을 구체적으로 적용한 연구결과들은 그다지 긍정적이지 않았다. 예컨대 Duffy 등의 연구(2005)는 미국 Harris-interactive와 영국의 Moris가 공동으로 영국에서 실시한 면대면 조사와 온라인 조사자료를 가지고 이 방법을 적용하였는데, 평가변수가 무엇이나에 따라 성향가중치의 효과가 다르게 나타났다. 이들의 연구결과를 요약해 보면 다음과 같다.

- 인구사회학적 변수 : rim 가중치 적용 (면대면 : 지역, 사회계층, 차량소유, 온라인 : 연령, 성, 지역, 교육수준, 소득수준, 인터넷 사용정도)
- 성향지수 : 5개의 질문(온라인 구매, 정보량에 대한 의견, 위험, 사회적 압력, 규율 등에 대한 의견)을 토대로 산출
- 평가변수 : 투표의사, 정치적 활동성향, 이민에 대한 태도, 콜레스테롤 관련 이슈에 대한 이해도 및 태도
- 결과 : 평가변수에 따라 다르게 나타났다.

Shonlau 등 (2004)의 연구에서도 성향가중이 기대한 효과를 가져오지 않았다. 이들은 캘리포니아 주민을 대상으로 건강에 관한 RDD 전화조사와 웹서베이를 실시하고 그 결과를 비교하였는데, 성향가중치를 적용했음에도 불구하고 두 표본간 차이가 제거되지 않았다. 이들의 또 다른 연구에서도 성향가중이 그다지 효과적이지 않은 것으로 나타났다.(Schonlau 등 2009). 이들은 50세 이상 성인을 대상으로 하는 미국 건강-은퇴 서베이 (Health Retirement Survey)를 실시하였는데, 인터넷 접근 가능자 중에서 표본을 추출하였다. 이들이 알아보고자 한 것은 인터넷 표본이 전체 표본을 대표할 수 있는가? 즉 selection bias가 감소될 수 있는가? 였는데, Genetic matching method를 이용해서 인종, 성별, 연령, 소득, 건강상태에 대한 자기 평가, 주택소유여부 등을 토대로 성향가중 점수를 구하고 이를 적용했지만 표본간 차이가 충분히 감소되지 않았다.

벨기에 지방에서 실시된 연구에서도 성, 연령, 교육수준, 거주지역 및 직업유무 등을 토대로 성향가중치를 구하고 이것을 온라인 응답자 패널 조사결과에 적용했지만, 근로만족도, 정치적 태도, 이민에 대한 태도 등에 관한 온라인 조사와 면접조사간의 응답차이가 제거되지 않았다. (Loosveldt and Sonck 2008).

이처럼 지금까지의 연구들은 대체로 성향가중에 대해서 부정적인 경우가 많았다. 그러나 이러한 연구결과들을 토대로 성향가중의 유용성에 대한 판단을 내리기에 앞서 검토할 부분이 있다. 첫째로 성향가중에 대한 지금까지의 연구들은 가중치 산출과정에서 인구적 속성들을 주로 사용하고 있다. 즉 성, 연령, 지역, 도시규모, 교육수준 등이 주로 이용되고 있고, 웹 사용과 직

접 관련성이 있는 라이프 스타일 변수들은 거의 사용되지 않았다. 따라서 방법적으로는 성향가중 방법을 사용하고 있지만 기존의 인구적 속성을 이용한 가중치와 다를 바가 없는 것이다.

이처럼 성향가중치 산출에서 인구적 속성이 주로 이용되는 것은 성향가중 방법이 준거 조사자료를 요구하기 때문이다. 즉 성향가중 산출에 이용되기 위해서는 그 변수를 포함한 준거 조사결과가 있어야 한다. 그리고 이러한 준거 조사는 물론 확률표집에 입각한 것이어야 한다. 실제로 이러한 준거 조사자료를 구하는 것은 대단히 어렵다. 그렇기 때문에 비교적 준거 조사자료가 가능한 인구적 속성들이 주로 활용된 것이다. 해리스 인터랙티브의 경우는 이러한 문제를 해결하기 위해 주기적으로 전화조사를 인터넷 조사와 병행하여 실시하였다. 그렇지만 일반적인 사회조사 상황에서 이러한 준거조사를 정기적으로 그리고 지속적으로 실시하는 것은 대단히 어렵다.

따라서 이 방법은 성향가중을 산출하는데 필요한 변수 즉 기준변수를 선정하고 그러한 변수의 준거조사결과를 구하는 것이 선행되어야 하는 것이다. 이러한 기준 변수는 두가지 관점에서 선정될 수 있다. 하나는 인터넷 표본의 자기선택 및 포함률 등의 오차를 축소시키는 관점에서 찾는 것이다. 따라서 인터넷 접근 및 조사에 대한 자발적 참여 성향 등과 관련된 변수를 찾아내는 방법이다. 즉 인터넷 조사와 전통적 조사에서의 차이를 인터넷 조사에 대한 응답성향을 통제함으로써 제거하고자 하는 것이다. 또 다른 방법은 조사하고자 하는 내용에 영향을 미치는 변수를 통제하고자 하는 것이다. 예컨대 정치의식관련 조사의 경우 일반적으로 이러한 응답에 영향을 크게 미칠 것으로 예상되는 변수에서의 차이 즉 정치이념 등과 같은 변수를 통제하고자 하는 것이다. 전자의 경우 만약 찾아낼 수 있다면 보다 광범위한 조사에 활용할 수 있을 것이다. 그러나 후자의 경우는 조사의 대상이 한정적일 경우에 유용할 것이다.

본 연구는 통계청에서 면대면 방식으로 실시한 사회조사 자료를 준거로 하고, 이 조사에 포함된 다양한 변수 중에서 적절한 기준변수를 선택하여 성향가중 방법을 검증해 보았다

제 3 절 인터넷 및 전통적 조사방법의 병행 방안

인터넷 조사와 전통적 조사방법을 혼합하고자 하는 시도도 이루어지고 있는데, 예컨대 Lukacs(2007)의 경우 이메일과 전화를 동시에 이용해서 조사를 실시하였다. 이와 같이 두 가지 모드를 사용한 것은 하나의 모드만을 사용하는 경우보다 조사대상자에 대한 도달 즉 접근률을 높이고자 하는 것이었다. 콜로라도 지역 사냥꾼들에 대한 조사를 이처럼 전화와 이메일을 이용해서 실시하였는데, 물론 각각의 방법만을 이용하는 경우보다는 도달율이 높았다. 이 방법은 전통적인 조사방법만으로 또는 웹조사 각각의 방법만으로 응답자 접근이 어려워진다는 점에서 활용 가능성이 높다고 본다. 우리나라의 인구 및 주택 총조사에서도 이러한 방법이 이용된다.

한편 Krotki and Tromfimovich(2006)의 경우는 RDD 전화로 실시한 조사를 웹 데이터로 보완하는 방법을 사용하였다. 즉 RDD 전화 조사와 웹 조사를 병행하여 실시한 다음, 그 데이터와 웹조사 데이터를 결합하였는데, 결합에서는 인구적 변수 및 흡연자 비율을 이용했다. 이 방법이 과연 타당한 것인가에 대해서 이들이 직접적인 검증을 실시하지 않았다. 그렇지만 데이터 결합에는 앞서 살펴본 성향 가중 모형도 적용될 수 있다. 따라서 성향 가중 모형에 대한 연구결과가 축적되면 이러한 데이터 결합에도 활용될 수 있고, 그렇게 되면 이러한 연구방법의 활용도도 높아질 것으로 보인다. 특히 전화 조사의 경우 확률표집이 가능하지만 설문지의 길이에 제한이 많다. 이러한 전화 조사와 웹데이터의 결합이 가능하게 되면 활용도는 무척 높아질 것으로 본다.

이상에서 살펴본 바와 같이 인터넷 조사의 한계를 극복하려는 시도가 다양하게 이루어지고 있다. 이러한 각각의 방법을 요약하면 다음 표와 같다.

	장 점	한 계 점	기존연구결과
확률표집 방법	<ul style="list-style-type: none"> - 확률표집의 특성을 가질 수 있다 - 일단 패널을 구축하게 되면 광범위하게 사용할 수 있다 	<ul style="list-style-type: none"> - 비용이 많이 소요된다 - 패널 모집/응답자 모집과정에서 거부율이 높을 수 있다 - 패널과 거부자간의 체계적 차이가 있을 수 있다 	<ul style="list-style-type: none"> - 비교적 긍정적
비확률표집 + 성향가중	<ul style="list-style-type: none"> - 저렴하게 활용할 수 있다 	<ul style="list-style-type: none"> - 준거조사자료를 구하기 어렵다 - 기준 변수 선정이 어렵다 	<ul style="list-style-type: none"> - 부정적 그러나 추가 연구가 필요
멀티모드 (동일 표본): - 응답자별로 모드를 달리하여 조사를 실시 (예컨대 일부는 전화로 일부는 웹으로 또 다른 일부는 면대면으로 조사 실시)	<ul style="list-style-type: none"> - 단일한 조사방법을 사용하는 것보다 여러 조사방법을 이용하는 것이 대표성을 높일 수 있다 	<ul style="list-style-type: none"> - 표본 프레임이 구하기 어렵다 	
멀티모드 (별도 표본): - 전통적 조사방법과 인터넷 조사결과를 통합	<ul style="list-style-type: none"> - 각각의 조사 모드의 장점을 살릴 수 있다. 	<ul style="list-style-type: none"> - 자료 결합방법의 타당성이 검토되지 않았다 	<ul style="list-style-type: none"> - 성향가중 방법을 활용할 수 있음

제 3 장 방법론과 모형

제 1 절 인터넷 조사 표본 추출

본 연구에서는 인터넷 조사를 위해 자원자들의 참여를 유도하여 패널을 구축하였고 패널에서 일부를 추출하여 조사 표본으로 하였다. 원칙적으로 패널의 크기는 작은 것보다 큰 것이 좋겠지만 크기가 가장 중요하다고 보기 어렵다. 조사응답자로서의 패널 가입자들의 질(質)도 못지않게 중요하다.

사회조사에서 표본은 모집단과 일치된 몇 개 주요한 인구·지리적 구성을 갖도록 기획되는 것이 좋다. 본 연구에서 고려된 일치 요구 특성은 광역시·도(16), 동·읍면(2), 성별 구분(2), 연령대(5) 등이다 (괄호 안 숫자는 범주 수를 나타냄). 부집단(subgroup)의 총 수는

{7개 광역시 + (9개 도)*(2개 동·읍면)} * (2개 성별 구분) * (5개 연령대)
즉, 250개이다.

이 연구에서 모집된 패널은 2009년 7월 14일에 최종 7,616명으로 규모가 크지 않았다 (6월 15일 6,854명). 그나마 인터넷 조사 자원자들이므로 250개의 모든 인구·지리적 부집단에서 모집단 비례 표본을 확보하는 것은 가능하지 않다. 인터넷 사용자 집단은 특히 연령대 분포에서 모집단과 큰 차이가 있기 때문이다. <표 1>은 상주인구(모집단) 대비 자원자 패널의 상대적 비를 본 것이다. 50대 연령대와 60대 이상 연령대의 상대적 비(比)는 모집단 대비 각각 27%와 6%에 불과하였고 반면 30대 연령대의 상대적 비는 모집단 대비 175%에 달하였다. 또한 읍면 거주자의 모집단 대비 상대적 비는 58%에 그쳤다 (이상적 비는 100%).

<표 1> 상주인구 대비 자원자 패널의 상대적 비(relative ratio)

시·도	서울 1.20, 부산 1.12, 대구 1.20, 인천 0.99, 광주 0.94, 대전 1.68, 울산 1.14 경기 0.93, 강원 0.97, 충북 0.84, 충남 0.64, 전북 0.89, 전남 0.62, 경북 0.76, 경남 0.79, 제주 0.88
동·읍면	동 1.09, 읍면 0.58
성 별	남자 0.89, 여자 1.11
연령대	20대 이하 1.60, 30대 1.75, 40대 0.73, 50대 0.27, 60대 이상 0.06

목표 표본이 인구·지리적 부집단(subgroup)에서 인구 비례하는 크기를 갖도록 하는 것이 어려우므로, 차선으로 각 인구·지리적 특성 별로 모집단과 주변적으로 일치하는 분포를 갖도록 조치하고자 하였다. 다음은 본 연구에서 고안되고 적용된 표본추출 방법이다.

가중 리스트 계통 추출법(systematic sampling from the weighted list)

1. 패널 리스트를 광역시·도(16), 동·읍면(2), 성별 구분(2), 연령대(5)의 순서로 정렬한다 (괄호 안은 범주 수).
2. 정렬된 리스트에 rim (raking ratio) 가중법에 의한 가중치를 부여한다. 이 때, 가중치의 총계는 패널 리스트의 길이인 N 임을 가정한다. 개체 가중치 w 가 4 또는 5로 설정된 cut.value보다 크면 절삭 처리한다. 기준집단은 최근 상주인구이고 일치 기준변수는 광역시·도, 동·읍면, 성별 구분, 연령대 등이다. 적용된 기준 비율은 <표 2>와 같다.
3. 0과 가중치 총계의 구간에서 다음 수열 $\{a_m; m = 1, \dots, n\}$ 을 빼낸다.
 - 항 간 간격은 $k = N_0/n$ 이다 (equal displacement). 여기서 N_0 는 패널 리스트에 부여된 가중치의 총계이고 n 은 목표 표본의 크기이다.
 - 초항 a_1 은 구간 $(0, k)$ 에서 임의로 선정된다 (random selection).
 - k 가 cut.value보다 작으면 1개 개체가 2회 이상 추출될 가능성이 있으므로 cut.value를 하향 조정한다.

4. 각 $m (= 1, \dots, n)$ 에 대하여, 가중치의 부분 합(partial sum)에서 a_m 을 초과하는 인덱스 중에서 가장 작은 것을 찾아낸다.

<표 2> rim 가중법 적용 기준 비율

시·도	서울 0.209, 부산 0.075, 대구 0.052, 인천 0.054, 광주 0.030, 대전 0.031, 울산 0.022 경기 0.221, 강원 0.031, 충북 0.031, 충남 0.039, 전북 0.038, 전남 0.038, 경북 0.054, 경남 0.064, 제주 0.011
동·읍면	동 0.818, 읍면 0.182
성 별	남자 0.491, 여자 0.509
연령대	20대 이하 0.255, 30대 0.221, 40대 0.217, 50대 0.139, 60대 이상 0.168

예를 들어, 6월 15일 확보 패널 6,854명(= N) 가운데 1,500명(= n)을 뽑기로 하자 (cut.value는 4로 설정됨). 이런 가중치 부여 과정에서 158명 패널원이 절삭 처리되었고 이에 따라 가중치 총계는 5,918(= N_0)이 되었으며 간격 k 는 3.95가 되었다. 그런데 이 값이 cut.value보다 크므로 그대로 두면 어떤 패널원은 2회 추출되는 결과를 초래할 수 있다. 이를 막기 위하여 cut.value를 하향 조정하게 되고 이에 따라 간격 k 값도 수정된다. 이런 과정을 몇 차례 반복하면 cut.value와 간격 k 가 같게 된다 (= 3.945). 초항 a_1 을 구간 $(0, k)$ 에서 임의로 뽑았는데 1.575였고 이에 따라 제 2항 a_2 는 5.520(= $a_1 + k = 1.575 + 3.945$), 제 3항 a_3 는 9.465(= $a_2 + k = 5.520 + 3.945$) 등이 되었다. 그런데 절삭 처리 이후 가중치의 부분 합이

0.563 1.125 1.688 2.250 2.813 3.375 3.938 4.500 5.063 5.625
6.188 6.750 7.313 7.876 8.438 9.001 9.563 10.126 10.688 11.251

이였으므로 개체번호 3, 10, 17 등이 뽑혔다 (밑줄 친 항의 인덱스).

본 연구에서 패널 리스트에서 표본을 추출하는 작업을 위해 R 언어로 프로그램이 작성되었다 (프로그램 명: panel_to_sample_version_1.r).

<표 3>은 상주인구 대비 추출 표본의 상대적 비를 본 것이다. <표 1>의 상주인구 대비 자원자 패널의 상대적 비에 비해 상당히 개선되었음을 확인할 수 있다. 예컨대 50대 연령대와 60대 이상 연령대가 패널에서는 모집단 대비 상대적으로 27%와 6%에 불과하였으나 표본에서는 각각 107%와 29%로 개선되었고 읍면 거주자는 모집단 대비 상대적으로 패널에서는 58%였으나 표본에서는 85%로 개선되었다 (이상적 비는 100%).

<표 3> 상주인구 대비 추출 표본의 상대적 비(relative ratio)

시·도	서울 1.03, 부산 0.97, 대구 1.08, 인천 1.13, 광주 0.90, 대전 1.13, 울산 1.14 경기 0.99, 강원 0.94, 충북 0.97, 충남 0.85, 전북 1.11, 전남 0.95, 경북 0.85, 경남 0.95, 제주 1.00
동·읍면	동 1.03, 읍면 0.85
성 별	남자 0.95, 여자 1.05
연령대	20대 이하 1.16, 30대 1.16, 40대 1.16, 50대 1.07, 60대 이상 0.29

이상의 가중 리스트 계통추출법에 따라 뽑은 표본에서도 주요 인구·지리적 구성에서 모집단과 완전히 일치하지는 않는다. 이는 소수 개체에 지나친 가중치가 붙는 것을 막았기 때문이다. 모집단과 표본의 불일치는 이후 사후 가중치 부여를 통해 해결될 것이다.

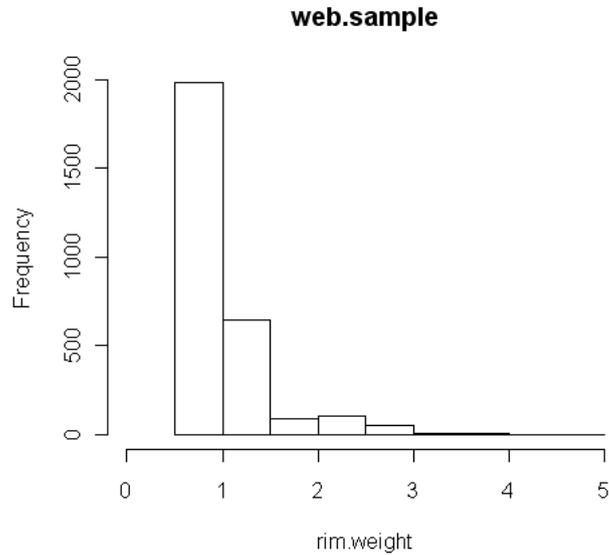
앞서 패널 리스트가 광역시·도, 동·읍면, 성별 구분, 연령대의 순서로 정렬되어 있음을 가정한 바 있다. 따라서 표본 응답자의 접촉 실패 또는 비협조 등의 사유가 발생하는 경우 패널 리스트에서 다음 줄에 위치한 응답자로 대체하더라도 최종 표본의 인구·지리적 구성은 목표로부터 크게 벗어나지 않게 된다. (광역시·도, 동·읍면, 성별 구분, 연령대의 순서로).

제 2 절 면대면 조사 자료와 인터넷 조사 자료의 비교

지역추출과 조사원 면접 방식에 의한 2009년 사회조사가 7월에 실시되었고 8월 말 분석용 자료로 완결되었다 (이하 FTF 표본으로 약칭; 응답자 수 37,049명). 설문지는 문화 및 여가, 복지, 소득 및 소비, 노동, 사회참여 등 5개 영역으로 구성되었다. 4개 인구·지리적 특성(광역시·도, 동·읍면, 성별 구분, 연령대)에 의한 결합 칸 가중법(cell weighting)으로 개체 별 가중치가 부여되었다. FTF 표본의 cell 가중치는 최소값 0.14, Q1 0.55, 중간값 0.84, Q3 1.25, 최대값 8.97로 요약된 바 범위가 넓게 퍼진 상태였다. 참고로 rim 가중치를 구하였고 요약한 결과 최소값 0.29, Q1 0.59, 중간값 0.83, Q3 1.47, 최대값 2.18로 나타났다.¹⁾ 이처럼 rim 가중치의 산포(범위)가 cell 가중치의 산포(범위)에 비해 작았다. cell 가중치와 rim 가중치의 상관계수는 0.85였다.

자원자 패널과 Web을 매체로 한 인터넷 사회조사의 설문지는 문화 및 여가, 소득 및 소비, 사회참여 등 3개 영역으로 구성되었다. 인터넷 조사는 2009년 7월에 실시되었고 곧바로 분석용 자료가 완결되었다 (이하 Web 표본으로 약칭; 응답자 수 2,903명). 분석자료 개체에 4개 인구·지리적 특성(광역시·도, 동·읍면, 성별 구분, 연령대)에 의한 rim 가중법으로 가중치가 부여되었는데 이때 cut.value로 5를 적용하였다 (R 프로그램의 파일명: web_sample_rimweighting_version_1.r). Web 표본의 rim 가중치는 최소값 0.52, Q1 0.73, 중간값 0.89, Q3 1.05, 최대값 4.58로 나타났고 분포는 <그림 1>과 같다.

1) cell·rim 가중치는 평균 1이 되도록 척도화되었음.



<그림 1> 인터넷 조사 표본의 rim 가중치 분포

조사원 면접 방식에 의한 2009년 사회조사의 FTF 표본과 Web 표본의 차이를 살펴본 결과, 주요 사항은 다음과 같다 (전자에는 cell 가중치를 적용하였고 후자에는 rim 가중치를 적용하였음).

1) 교육정도

		초등/중 학교	고등학교	대학교	대학원 이상	합계
FTF	빈도	8905	13379	13352	1413	37049
	%	24.0	36.1	36.0	3.8	100.0
Web	빈도	315	871	1548	169	2903
	%	10.9	30.0	53.3	5.8	100.0
합계	빈도	9220	14250	14900	1582	39952
	%	23.1	35.7	37.3	4.0	100.0

FTF 표본에서는 대학/대학원 교육이 39.8%였으나 Web 표본에서는 대학/대학원 교육이 59.1%로 19.3%p의 차이를 보였다. 반면 FTF 표본에서는

초등/중학교 교육이 24.0%였으나 Web 표본에서는 초등/중학교 교육이 10.9%에 불과하였다. 이와 같이 FTF 표본 응답자에 비해 Web 표본 응답자들의 교육수준이 높은 것으로 나타났다 (감마계수²⁾ 0.34).

2) 소비생활만족도 (5점 척도: 1=매우 만족, 2=약간 만족, 3=보통, 4=약간 불만족, 5=매우 불만족)

		매우만족	약간만족	보통	약간불만족	매우불만족	합계
FTF	빈도	666	4410	17072	11297	3603	37048
	%	1.8	11.9	46.1	30.5	9.7	100.0
Web	빈도	80	520	1135	873	295	2903
	%	2.8	17.9	39.1	30.1	10.2	100.0
합계	빈도	746	4903	18207	12170	3898	39951
	%	1.9	12.3	45.6	30.5	9.8	100.0

소비생활만족도는 응답자가 자각하는 소비수준인데, 만족 이상의 비율이 FTF 표본에서는 13.7%였으나 Web 표본에서는 20.7%로 7.0%p의 차이를 보였다. 이와 같이 FTF 표본 응답자에 비하여 Web 표본 응답자들의 자각적 소비수준이 약간 높은 것으로 나타났다 (감마계수 0.06).

3) 일반신문을 읽는 빈도 (5점 척도: 1=거의 매일, 2=1주일에 3-4회, 3=1주일에 1-2회, 4=2주일에 1회, 5=본 적 없다)

		거의 매일	1주일에 3-4회	1주일에 1-2회	2주일에 1회	본 적 없다	합계
FTF	빈도	8820	3989	4763	2095	17381	37048
	%	23.8	10.8	12.9	5.7	46.9	100.0
Web	빈도	927	352	409	233	982	2903
	%	31.9	12.1	14.1	8.0	33.8	100.0
합계	빈도	9747	4341	5172	2328	18363	39951
	%	24.4	10.9	12.9	5.8	46.0	100.0

2) 감마(gamma) 계수는 2원 교차표에서 행과 열이 모두 순서형인 경우 2개 관측개체 쌍의 비교에서 행과 열이 일치 방향으로 결합하는 강도를 척도화한 지표이다. 행의 수가 2인 경우에도 행에 적절한 의미를 붙이는 경우 활용 가능하다. 최소 -1, 최대 1의 값을 갖는다.

거의 매일 일반신문을 읽는 응답자의 비율이 FTF 표본에서는 23.8%였으나 Web 표본에서는 31.9%로 8.1%P의 차이를 보였다. 반면 일반신문을 본 적 없다는 응답자 비율은 FTF 표본에서는 46.9%였고 Web 표본에서는 33.8%로 13.1%P의 차이가 났다. 이와 같이 FTF 표본 응답자에 비하여 Web 표본 응답자들이 일반신문을 자주 접하는 것으로 나타났다 (감마계수 0.19).

4) 인터넷 신문을 읽는 빈도 (5점 척도: 1=거의 매일, ... , 5=본 적 없다)

		거의 매일	1주일에 3-4회	1주일에 1-2회	2주일에 1회	본 적 없다	합계
FTF	빈도	8979	4656	4061	1651	17703	37049
	%	24.2	12.6	11.0	4.5	47.8	100.0
Web	빈도	1215	576	311	125	676	2903
	%	41.9	19.8	10.7	4.3	23.3	100.0
합계	빈도	10194	5232	4372	1776	18379	39952
	%	25.5	13.1	10.9	4.4	46.0	100.0

거의 매일 인터넷 신문을 읽는 응답자의 비율이 FTF 표본에서는 24.2%였으나 Web 표본에서는 41.9%로 17.7%P의 차이를 보였다. 반면 인터넷 신문을 본 적 없다는 응답자 비율은 FTF 표본에서는 47.8%였고 Web 표본에서는 23.3%로 24.5%P의 차이가 났다. 이와 같이 FTF 표본 응답자에 비하여 Web 표본 응답자들이 인터넷 신문을 자주 접하는 것으로 나타났다 (감마계수 0.39). 그러한 경향이 일반신문의 경우에서도 있었으나 인터넷신문의 경우에 더 강하다고 하겠다.

5) 소득분배에 대한 견해 (5점 척도: 1=매우 공평, 2=약간 공평, 3=보통, 4=약간 불공평, 5=매우 불공평)

		매우 공평	약간 공평	보통	약간불공평	매우불공평	합계
FTF	빈도	212	1308	9539	16046	9944	37049
	%	0.6	3.5	25.7	43.3	26.8	100.0
Web	빈도	26	114	572	1205	986	2903
	%	0.9	3.9	19.7	41.5	34.0	100.0
합계	빈도	238	1422	10111	17251	10930	39952
	%	0.6	3.6	25.3	43.2	27.4	100.0

소득분배가 매우 불공평하다는 응답자 비율이 FTF 표본에서는 26.8%였으나 Web 표본에서는 34.0%로 7.2%p의 차이를 보였다 (감마계수 0.13).

6) 계층상승 가능성에 대한 견해 (1=매우 높다, 2=비교적 높다, 3=비교적 낮다, 4=매우 낮다, 5=모르겠다)

		매우 높다	비교적 높다	비교적 낮다	매우 낮다	모르겠다	합계
FTF	빈도	1133	12794	12665	4228	6228	37048
	%	3.1	34.5	34.2	11.4	16.8	100.0
Web	빈도	134	908	1262	405	194	2903
	%	4.6	31.3	43.5	14.0	6.7	100.0
합계	빈도	1267	13702	13927	4633	6422	39951
	%	3.2	34.3	34.9	11.6	16.1	100.0

계층상승 가능성이 낮다(3:비교적 낮다, 4:매우 낮다)는 응답자 비율이 FTF 표본에서는 45.6%였으나 Web 표본에서는 57.5%로 11.9%p의 차이를 보였다. 한편, 모르겠다는 DK(Don't Know) 비율이 FTF 표본에서는 16.8%, Web 표본에서는 6.7%로 Web 응답자들이 적극적으로 의사 표현을 하는 것으로 나타났다. 그러나 이 차이는 조사방식의 차이일 수 있다.

이상을 종합 정리하면, FTF 표본 응답자들에 비해 Web 표본 응답자들은

교육정도가 높고 소비수준이 높으며 일반신문과 인터넷 신문을 자주 접하는 편이다. 또한 Web 표본 응답자들이 상대적으로 더 비판적/비관적인 사회의식을 갖는 경향이 있다고 하겠다. 이에 따라 주요 인구·지리적 요인을 통제한다고 하더라도 다수의 사회/문화 영역에서 FTF 조사와 Web 조사가 성향적 차이를 보일 개연성이 충분히 있는 것으로 보인다.

제 3 절 성향점수모형 가중법에 의한 성향 보정

Rosenbaum과 Rubin (1983, 1984) 등 다수의 연구에서 제시된 성향점수모형(propensity score model)은 관측연구(observation study)에서 집단 간 존재하는 성향적 차이를 제거하여 집단 간 공정한 비교를 가능하게 한다. 이 방법을 약술하자면 다음과 같다.

집단 A와 집단 B에서 표본을 취하였고 관심변수 Y 와 함께 다수의 배경변수(공변량) x_1, \dots, x_p 가 측정되었다고 하자. 집단 A의 Y 와 집단 B의 Y 를 공정하게 비교하기 위하여 로지스틱 모형

$$\log \frac{P(A | x_1, \dots, x_p)}{P(B | x_1, \dots, x_p)} = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \quad (1)$$

를 활용하여 개체별로 A에 속할 확률 $P(A | x_1, \dots, x_p)$ 와 B에 속할 확률 $P(B | x_1, \dots, x_p)$ 를 추정한다 (두 확률의 합은 1이다). 이런 확률들의 추정값을 성향점수(propensity score)라고 한다.

성향점수의 활용 방법은 크게 다음 세 가지이다 (D'agostino, 1998). 첫째 방법은 짝짓기 기법(matching technique)으로 성향점수가 비슷한 A 개체 1개와 B 개체 1개 또는 m 개(= 2,3,...)를 쌍으로 묶어 쌍 내에서 Y 를 비교하고 다수의 쌍에서 나오는 비교결과들을 종합한다. 둘째 방법은 층화(stratification) 또는 가중법(weighting method)으로, 1) 집단 A와 집단 B의 합 집단(pooled group)의 차원에서 A 개체는 $P(A | x_1, \dots, x_p)$ 의 역수로 가중하고 B 개체는 $P(B | x_1, \dots, x_p)$ 의 역수로 가중한다. 또는 2) 집단 A를 기준으

로 하는 경우에는 A 개체는 가중치지 않고 (가중치=1) B 개체는

$$\frac{P(A | x_1, \dots, x_p)}{P(B | x_1, \dots, x_p)}$$

로 가중한다. 안정적인 가중치 부여를 위하여 성향점수를 구간화하기도 하나 꼭 그렇게 해야 하는 것은 아니다. 마지막 방법은 회귀 수정(regression adjustment)이다.

인터넷 조사 결과의 보정이 목적인 경우 앞의 두 번째 방법을 적용한다. 그 때 A는 준거조사(reference survey)인 FTF 표본이고 B는 Web 표본이다. 즉, 지시자 T 를 FTF 표본이면 1, Web 표본이면 0으로 두고 $P(T=1 | x_1, \dots, x_p)$ 에 대한 로지스틱 모형 추정값 p 를 구한다. 이어서 FTF 표본 개체에는 성향가중치 1을 부여하고 Web 표본 개체에는 성향가중치 p/q 를 부여한다 ($q = 1 - p$).

각 표본 개체들에 인구·지리적 구성의 목적으로 이미 cell 또는 rim 가중치가 부여된 경우 앞의 로지스틱 모형 적합 시 그런 가중치가 반영되어야 하며 기존 cell/rim 가중치에 성향가중치를 곱하는 가중치 업데이트 작업이 실행된다.

로지스틱 성향점수 모형의 적용에서 핵심 이슈는 성향보정의 기준변수(공변량, 배경변수)의 선정에 있다. 성향보정 기준변수의 선정 시 다음 몇 가지 사항을 고려하여야 할 것으로 생각된다.

- 1) 성향보정 기준변수는 변화하지 않는 내재적 근본 성향을 반영하여야 한다. 변하기 쉬운 표면적 변수는 피하여야 한다. 예컨대 정치여론 조사에서 ‘한반도 통일의 당위성’에 대한 의견은 내재적이거나 ‘6자 회담 참석’에 대한 의견은 표면적이다.
- 2) 쉽게 질문할 수 있고 쉽게 응답할 수 있어야 한다. 성향보정 기준변수는 보정의 기반이 되므로 측정오차가 개입될 소지가 적어야 한다. 어려운 질

문이나 부담이 되는 응답에는 심각한 측정오차가 개입될 위험이 있다.

- 3) 그 변수에서 집단 간 유의한 차이가 있어야 하고 그 차이가 이해 가능해야 한다. FTF 표본 응답자와 Web 표본 응답자는 인터넷 접근/활용 빈도에서 차이가 있을 것이므로 그런 차이를 반영할 수 있어야 한다.

2절에서 FTF 표본과 Web 표본의 차이를 탐색한 바를 참고하고 성향기준 선정에 관한 위 관점들을 지침으로 하여, 본 연구에서는 다음 변수들을 성향조정 기준변수를 선정하고 범주화하였다.

- 1) 광역시·도(16개 층) = 1(서울), 2(부산), ... , 16(제주)
- 2) 동·읍면(2개 층) = 1(동), 2(읍면)
- 3) 남·여(2개 층) = 1(남), 2(여)
- 4) 연령(5개 층) = 1(20대-), 2(30대), 3(40대), 4(50대), 5(60대+)
- 5) 교육정도 (4개 수준) = 1(초/중), 2(고), 3(대학), 4(대학원)
- 6) 소비생활만족도 (2개 수준) = 1 (만족), 2(보통/불만족)
- 7) 소득분배 견해 (5개 수준, 리커트 척도)
- 8) 일반신문 보기 (2개 수준) = 1 (본다), 2 (보지 않는다)
- 9) 인터넷 신문 보기 (5개 수준, 리커트 척도)

첫 4개 변수는 인구·지리적 구성의 대표성을 위해 쓰인 바 있고 이미 cell 및 rim 가중치로 사전 균형화되도록 조치한 바 있다. 그러나 rim 가중화에서 cut.value를 적용한 바 있고 다른 변수들과 결합하여 새로운 효과를 만들어낼 가능성이 있으므로 모형에 포함하였다.

본 연구의 경우, FTF 표본은 37,049명의 응답자로 구성되었고 Web 표본은 2,903명의 응답자로 구성되었으므로 크기에 있어 큰 차이가 있다. 성향보정 로지스틱 모형 (1)의 적합에 앞서 사전 균형화(prior balancing) 조치로서 Web 표본을 12.76배(=37,049/2,903)로 부스팅(boosting)하여 크기를 FTF 표본과 맞추었다. 부스팅의 결과로 모형 (1)의 상수항 β_0 는 달라지나

설명항의 계수들 β_1, \dots, β_p 는 변하지 않는다. 그러나 부스팅으로 로지스틱 모형의 보다 안정적인 추정이 가능할 것으로 기대한다.

이상 9개 변수를 설명변수로 한 성향조정 로지스틱 모형을 적합하기 위해 R 프로그램이 작성되었다 (파일명: web_propensity_scoring_version_1.r). 적합 로지스틱 모형에서 회귀계수는 다음과 같다.³⁾

시도코드1	0	교육정도1	0
시도코드2	-0.18	교육정도2	-0.26
시도코드3	-0.18	교육정도3	-0.58
시도코드4	-0.21	교육정도4	-0.51
시도코드5	-0.21	일반신문1	0
시도코드6	-0.14	일반신문2	0.33
시도코드7	-0.20	인터넷신문1	0
시도코드8	-0.07	인터넷신문2	0.09
시도코드9	-0.14	인터넷신문3	0.51
시도코드10	-0.17	인터넷신문4	0.49
시도코드11	-0.14	인터넷신문5	1.06
시도코드12	-0.19	소득분배1	0
시도코드13	-0.20	소득분배2	0.31
시도코드14	-0.13	소득분배3	0.44
시도코드15	-0.17	소득분배4	0.25
시도코드16	-0.19	소득분배5	-0.07
동읍면1	0	소비생활만족도1	0
동읍면2	-0.11	소비생활만족도2	0.39
성별1	0		
성별2	-0.22		
연령대1	0		
연령대2	-0.00		
연령대3	-0.29		
연령대4	-0.64		
연령대5	-1.10		

3) 회귀계수에 대한 표준오차 및 t 값, p-값 등은 부스팅(boosting)의 영향으로 축소, 증대, 축소되었을 것이므로 의미가 없어 보고하지 않았다. 그러나 부스팅 없이도 분석표본의 총 크기가 상당히 큰 상태이므로 대부분 계수들이 통계적으로 유의한 것으로 나타났다.

이 모형에서 계수들은 FTF로 이끌어질 성향을 나타내므로 양이면 FTF일 가능성으로 기울고 음이면 Web일 가능성으로 기운다. 계수의 절대값 크기는 그런 가능성의 강도이다.

여기서는 4개 인구·지리적 변수 외 기타 성향조정 기준 변수들의 효과를 살펴보기로 한다.

- 1) 가장 큰 효과를 보인 변수는 인터넷 신문 보기이다 (1: 거의 매일, 2: 1주일에 3-4회, 3: 1주일에 1-2회, 4: 2주일에 1회, 5: 보지 않는다). 인터넷 신문을 보지 않는 응답자는 거의 매일 보는 응답자에 비해 FTF 대 Web의 오즈(odds)가 2.89배(=exp(1.06))나 된다 (그 정도로 증대된다).
- 2) 교육정도가 대학 및 대학원인 응답자는 초등/중학교인 응답자에 비해 FTF 대 Web의 오즈가 0.56배(=exp(-0.58)) 내지 0.60배(=exp(-0.51))가 된다 (그 정도로 축소된다).
- 3) 소득분배의 공평성에 대한 견해가 보통인 응답자는 매우 공평하다고 생각하는 응답자에 비하여 FTF 대 Web의 오즈가 1.55배(=exp(0.44))가 된다 (그 정도로 증대된다). 매우 불공평하다고 생각하는 응답자에 비한 오즈도 이와 비슷하다.
- 4) 소비생활만족도가 보통/불만족/매우 불만족인 응답자는 만족/매우 만족인 응답자에 비해 FTF 대 Web의 오즈가 1.48배(=exp(0.39))가 된다 (그 정도로 증대된다).
- 5) 일반신문을 보지 않는 응답자는 보는 응답자에 비해 FTF 대 Web의 오즈가 1.39배(=exp(0.33))가 된다 (그 정도로 증대된다).

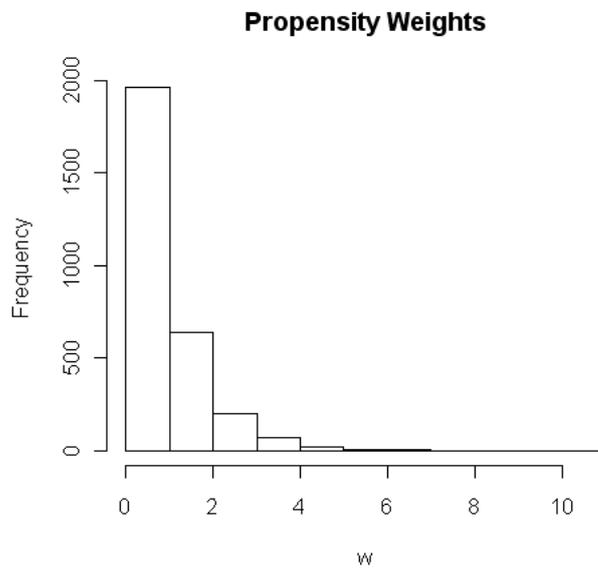
이와 같은 기준변수들의 효과는 가법적으로 작용하므로, 인터넷 신문을 보지 않고 교육정도가 대학이며 소득분배의 공평성에 대해 보통이라고 생각하며 소비생활만족도가 보통/불만족/매우 불만족이며 일반신문을 보지 않은 응답자는 인터넷 신문을 거의 매일 보고 교육정도가 초등/중학교이며 소득분배의 공평성에 대해 매우 공평하다고 생각하며 소비생활만족도가 만족/매우 만족이며 일반신문을 보는 응답자에 비해 FTF 대 Web의 오즈가

2.89*0.56*1.55*1.48*1.39배, 즉 5.16배
가 된다. 그 정도로 증대되는 것이다.

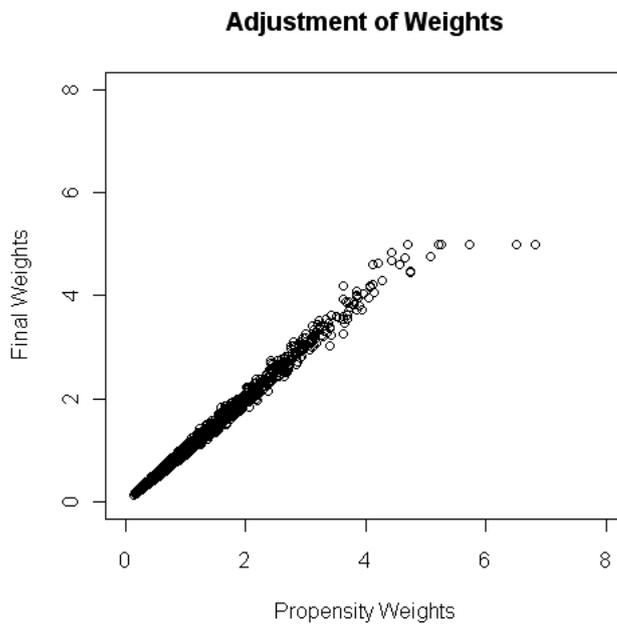
<그림 2>는 성향점수모형에서 나온 성향조정 비(比) p/q 에 인터넷 표본이 원래 가지고 있던 rim 가중치를 곱하여 얻은 성향조정 가중치의 분포를 보여준다.

성향가중으로 인하여 다시 인구·지리적 구성이 흐트러지므로 이를 재조정할 필요가 있는데 (Lee, 2009) 본 연구에서는 이를 위하여 성향조정 가중치가 붙은 Web 표본에 rim 가중법을 적용하였다 (cut.value를 5로 설정). 이것은 일종의 Calibration 절차이다 (Deville and Sarndal, 1992). 그러나 본 연구에서는 성향점수모형의 적합에 앞서 FTF 표본과 Web 표본을 모두 인구·지리적 구성에 맞게 가중화하였고 성향점수 모형에 4개의 기본 인구·지리적 변수를 포함시킨 바 있으므로, Calibration 효과는 크지 않았다. <그림 3>이 성향조정 가중치와 최종 가중치 간 관계를 보여준다. 최종 rim 가중법에서 cut.value를 설정함으로써 일부 개체 가중치가 절단된 것 외에는 큰 변화가 없었음을 알 수 있다.

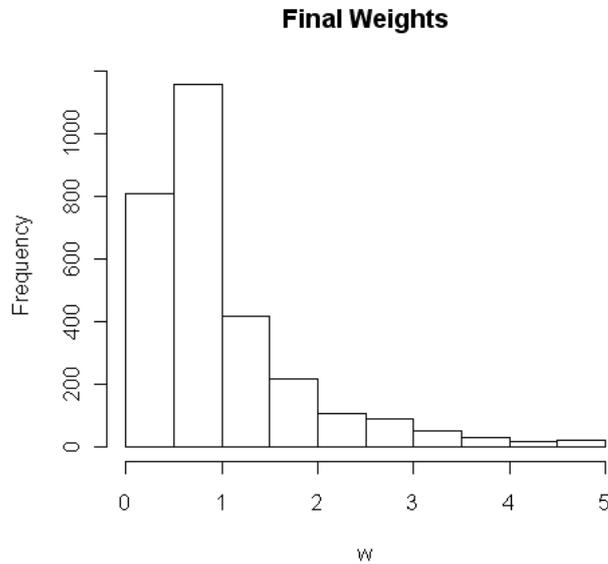
<그림 4>는 Web 표본 최종 가중치의 분포를 보여주고 <표 4>가 본 연구에 적용된 Web 표본에 대한 가중치 산출 과정을 보여준다.



<그림 2> Web 표본 성향조정 가중치의 분포



<그림 3> Web 표본의 성향조정 가중치 대 최종 가중치 분포



<그림 4> 인터넷 조사 표본의 최종 가중치 분포

<표 4> Web 표본에 대한 가중치 산출 과정

	원자료	→	단계 1	→	단계 2	→	단계 3
가중하는 이유	해당 없음		인구·지리적 구성		성향조정		인구·지리적 구성
가중치 산출 방법	해당 없음		rim weighting		propensity score model		rim weighting
부여가중치	1	→	w1	→	$w2 := w1 * \frac{p}{q}$	→	w3

제 4 절 rim 가중법에 의한 성향 보정

성향점수모형에 의한 가중법으로 보정을 하는 대신에 같은 기준변수들로 Web 표본의 rim 가중화를 통해 집단 간 존재하는 인구·지리적, 성향적 차이를 제거함으로써 집단 간 공정한 비교를 위한 토대를 마련할 수 있다.

3절에서와 같이, 다음 변수들을 성향조정 기준변수로 선정하였다.

- 1) 광역시·도(16개 층) = 1(서울), 2(부산), ... , 16(제주)
- 2) 동·읍면(2개 층) = 1(동), 2(읍면)
- 3) 남·여(2개 층) = 1(남), 2(여)
- 4) 연령(5개 층) = 1(20대-), 2(30대), 3(40대), 4(50대), 5(60대+)
- 5) 교육정도 (4개 수준) = 1(초/중), 2(고), 3(대학), 4(대학원)
- 6) 소비생활만족도 (2개 수준) = 1 (만족), 2(보통/불만족)
- 7) 소득분배 견해 (5개 수준, 리커트 척도)
- 8) 일반신문 보기 (2개 수준) = 1 (본다), 2 (보지 않는다)
- 9) 인터넷 신문 보기 (5개 수준, 리커트 척도)

rim 가중법을 적용하려면 성향조정 기준변수 9개 각각에 대하여 모집단 비율이 있어야 한다. 그 가운데 4개의 인구·지리적 변수에 대해서는 모집단 비율이 있지만 그 밖의 5개 변수에 대해서는 모집단 비율이 없다. 차선책으로, 4개의 인구·지리적 변수에 의해 cell 가중화된 FTF 표본에서 그 밖의 5개 변수에 대한 모집단 비율을 추정하였다.

본 연구에 적용된 rim 가중법에 의한 Web 표본의 성향보정 과정을 단계별로 기술하면 다음과 같다.

단계 1. FTF 표본을 4개 인구·지리적 cell 가중치를 부여하여 그 밖의 5개 성향보정 기준변수의 분포를 구한다. cell 가중화 대신 rim 가중화를 할 수도 있다.

단계 2. 9개 변수의 기준 분포에 맞춰 Web 표본을 rim 가중화한다. 지나치게 큰 가중치에 대하여는 cut.value를 적용할 수 있다.

본 연구에서 R 프로그램(파일명: rim_propensity_weighting_version_1.r)을 만들어 적용하였다. 성향점수모형에 의한 가중법과 마찬가지로 이 방법도 일종의 calibration 절차로 볼 수 있으므로 (Deville and Sarndal, 1992) 두 방법의 적용 결과 간에 큰 차이가 없을 것으로 예상할 수 있다. 전자에 비해 단순한 것이 장점이다.

제 5 절 주요 평가문항의 성향보정 결과

2009년 사회조사의 전 항목(인터넷 조사 설문지)에 걸친 본격적인 성향보정의 결과는 다음 장에서 평가할 것이다. 이 절에서는 성향점수모형에 의한 가중법에 의한 성향보정과 rim 가중법에 의한 성향보정의 결과를 몇 개의 주요 평가문항에서 살펴보기로 한다.

이 절에서 평가문항으로 선정한 설문 및 척도는 다음과 같다.

- TV 평일 시청 (시간)
- TV 토요일 시청 (시간)
- TV 일요일 시청 (시간)
- 해외여행 경험 (지난 1년간 비율, %)
- 전반적 만족감 (5점 리커트 척도 평균, 1 = 매우 만족, ... , 5 = 매우 불만족)
- 여가활용 만족도 (5점 리커트 척도 평균, 1 = 매우 만족, ... , 5 = 매우 불만족)
- 장래소득 기대 (5점 리커트 척도 평균, 1 = 크게 증가, ... , 5 = 크게 감소)
- 주관적 건강평가 (5점 리커트 척도 평균, 1 = 매우 좋다, ... , 5 = 매우 나쁘다)
- 스트레스 인식 (4점 리커트 척도 평균, 1 = 매우 많다, ... , 4 = 거의 없다)
- 생활시간 태도 (5점 리커트 척도 평균, 1 = 매우 바쁘다, ... , 5 = 전혀 그렇지 않다)

<표 5>는 성향점수모형(PS) 가중법에 의한 성향보정의 결과이다. 이 표에서 각 열은 FTF 표본과 Web 표본으로부터 다음 방법으로 얻은 수치이다. 이 중에서 [5]의 방법에 대하여는 3절에 자세히 기술한 바 있다.

- [1] FTF 표본 원자료 (FTF raw)
- [2] FTF 표본 인구·지리적 cell 가중치 적용 (FTF cell)
- [3] Web 표본 원자료 (Web raw)
- [4] Web 표본 인구·지리적 rim 가중치 적용 (Web rim)
- [5] Web 표본 성향점수모형 가중치 적용 (Web PS)

<표 5>를 보면 TV 평일 시청시간에 대한 FTF raw는 2.891시간, FTF cell은 2.841시간으로 별 차이가 없다. 그런데 Web raw는 2.528시간이고 Web rim은 2.674시간으로 FTF raw나 FTF cell에 비해 작다. Web 응답자의 TV 시청 시간이 적은 이유는 인터넷 사용과 관련 있을 것으로 추측된다. 그러나 Web 표본에 성향점수모형 가중치를 적용한 결과인 Web PS가 2.760시간으로 나타나 충분하지는 않지만 FTF cell과의 차이가 상당 부분 좁혀진 것으로 나타났다. 성향점수모형에 의한 가중법의 긍정적 역할을 볼 수 있는 경우이다.

사회조사의 경우 FTF 표본이나 Web 표본의 어떤 표본에서든 인구·지리적 가중치를 적용하여 활용하므로 이제부터는 더 이상 [1]의 FTF raw와 [3]의 Web raw를 언급하지 않기로 한다. 즉, [2]의 FTF cell과 [4]의 Web rim 간 차이를 살펴보고 [5]의 Web 성향보정 추정치가 [4]의 Web rim보다 [2]의 FTF cell에 근접하는가를 볼 것이다.

TV 토요일 시청시간과 일요일 시청시간에서 [2]의 FTF cell과 [4]의 Web rim은 거의 차이가 없다. 그런데 성향보정을 하게 되면서 시청시간 추정치가 늘어나 오차가 발생하게 되는 것을 볼 수 있다. 성향점수모형에 의한 가중법이 부정적인 경우이다.

해외여행 경험 비율에서 [2]의 FTF cell은 13.6%인 반면 [4]의 Web rim은 15.3%인데 이는 Web 응답자들이 교육정도와 소비수준이 높은 데 기인하는 것으로 보인다. Web 표본추정치를 성향보정하면 13.3%로 [2]의 FTF cell의 13.6%에 근접하므로 이 경우는 성향점수모형 가중법에 의한 보정이 매우 효과적임을 보여준다.

<표 5> 성향점수모형(PS)에 의한 가중법에 의한 성향보정의 결과

	[1]* FTF raw	[2] FTF cell	[3] Web raw	[4] Web rim	[5] Web PS	비고**
TV 평일	2.891	2.841	2.528	2.674	2.760	[5]
TV 토	3.608	3.600	3.518	3.611	3.663	[4]
TV 일	3.925	3.922	3.848	3.917	3.998	[4]
해외여행경험	12.2%	13.6%	15.5%	15.3%	13.3%	[5]
전반적만족감	3.186	3.170	3.182	3.201	3.265	[4]
여가활용만족도	3.134	3.128	2.970	2.988	3.065	[5]
장래소득기대	2.935	2.875	2.490	2.591	2.626	[5]
주관적건강평가	2.831	2.789	2.602	2.697	2.761	[5]
스트레스 인식	2.541	2.518	2.306	2.337	2.335	-
생활시간태도	2.653	2.611	2.493	2.585	2.617	[5]

- * [1] FTF(면대면) 표본 원자료 (FTF raw)
- [2] FTF(면대면) 표본 인구·지리적 cell 가중치 적용 (FTF cell)
- [3] Web 표본 원자료 (Web raw)
- [4] Web 표본 인구·지리적 rim 가중치 적용 (Web rim)
- [5] Web 표본 성향점수모형 가중치 적용 (Web PS)

** [4]와 [5] 중에서 [2]에 더 가까운 추정치.

전반적 만족감(경제, 직업, 건강 등)에서 [2]의 FTF cell은 3.17점이고 [4]의 Web rim은 3.201점으로 Web 응답자들이 다소 더 불만족하는 편이다. 그런데 Web 표본추정치를 성향보정하면 3.265점으로 [2]의 FTF cell에서 더욱 멀어지는 결과가 초래된다. 이 경우는 성향점수모형 가중법에 의한 보정이 제 역할을 하지 못함을 보여준다.

그 밖에 여가활용 만족도, 장래소득 기대, 주관적 건강평가, 생활시간 태도 등을 살펴보면 성향점수모형 가중법에 의한 보정이 어느 정도의 역할을 하는 것으로 볼 수 있다. 다만, 스트레스 인식에 있어서는 [2]의 FTF cell은 2.518점이고 [4]의 Web rim은 2.337점으로 Web 응답자들이 더 스트레스를 느끼는 편인데 Web 표본을 성향보정하여 얻는 추정치는 2.335점으로 별

로 달라지지 않는 것으로 나타났다. 종합해보면, 10개 항목으로 성향점수 모형 가중법에 의한 보정 방법을 평가한 결과 6개 항목에서는 긍정적으로, 3개 항목에서는 부정적으로, 1개 항목에서는 부정도 긍정도 아닌 것으로 나타났다.

향후 여러 사례에 2009년 사회조사 자료를 활용하여 Web 표본을 성향보정하기 위해서는 가급적 성향보정 기준변수의 수를 작게 하는 것이 좋을 것이다. 이에 따라 성향보정 모형에서 덜 중요한 것으로 나타나는 소비생활 만족도(2개 수준)와 일반신문 보기(2개 수준)를 제외한 7개 변수를 설명변수로 하는 로지스틱 모형을 만들고 이를 기초로 Web 표본을 가중화하는 성향보정을 하여 보았다. <표 6>이 그 결과이다.

**<표 6> 성향점수모형(PS)에 의한 가중법에 의한 성향보정의 결과:
보정기준 변수를 7개로 줄인 경우**

	[1]* FTF raw	[2] FTF cell	[3] Web raw	[4] Web rim	[5] Web PS	비고**
TV 평일	2.891	2.841	2.528	2.674	2.756	[5]
TV 토	3.608	3.600	3.518	3.611	3.668	[4]
TV 일	3.925	3.922	3.848	3.917	3.988	[4]
해외여행경험	12.2%	13.6%	15.5%	15.3%	13.7%	[5]
전반적만족감	3.186	3.170	3.182	3.201	3.205	-
여가활용만족도	3.134	3.128	2.970	2.988	3.020	[5]
장래소득기대	2.935	2.875	2.490	2.591	2.623	[5]
주관적건강평가	2.831	2.789	2.602	2.697	2.734	[5]
스트레스 인식	2.541	2.518	2.306	2.337	2.353	[5]
생활시간태도	2.653	2.611	2.493	2.585	2.624	[5]

- * [1] FTF(면대면) 표본 원자료 (FTF raw)
- [2] FTF(면대면) 표본 인구·지리적 cell 가중치 적용 (FTF cell)
- [3] Web 표본 원자료 (Web raw)
- [4] Web 표본 인구·지리적 rim 가중치 적용 (Web rim)
- [5] Web 표본 성향점수모형 가중치 적용 (Web PS)

** [4]와 [5] 중에서 [2]에 더 가까운 추정치.

전반적으로 볼 때, 기준변수의 수가 9개인 <표 5>와 비교하여 주목할 만한 차이는 나타나지 않았다. 그러나 세부적인 수치 변동이 일부 있어서, 10개 항목의 항목별 평가를 종합하면 7개 항목에서는 긍정적으로, 2개 항목에서는 부정적으로, 1개 항목에서는 부정도 긍정도 아닌 것으로 나타났다.

<표 7>과 <표 8>은 rim 가중법에 의한 성향보정의 결과로 보정기준 변수가 각각 9개와 7개인 경우를 보여준다.

대체로 <표 7>은 <표 5>와 유사하고 <표 8>은 <표 6>과 유사한 것으로 나타났다.⁴⁾ 따라서 성향보정 방법에 의해 결과가 영향을 받기 보다는 기준변수의 선정에 의한 영향이 더 큰 것으로 보여진다.

<표 7> rim 가중법에 의한 성향보정의 결과

	[1]* FTF raw	[2] FTF cell	[3] Web raw	[4] Web rim	[5] Web P.rim	비고**
TV 평일	2.891	2.841	2.528	2.674	2.763	[5]
TV 토	3.608	3.600	3.518	3.611	3.664	[4]
TV 일	3.925	3.922	3.848	3.917	3.990	[4]
해외여행경험	12.2%	13.6%	15.5%	15.3%	13.1%	[5]
전반적만족감	3.186	3.170	3.182	3.201	3.273	[4]
여가활용만족도	3.134	3.128	2.970	2.988	3.073	[5]
장래소득기대	2.935	2.875	2.490	2.591	2.631	[5]
주관적건강평가	2.831	2.789	2.602	2.697	2.764	[5]
스트레스 인식	2.541	2.518	2.306	2.337	2.332	-
생활시간태도	2.653	2.611	2.493	2.585	2.616	[5]

* [1] FTF(면대면) 표본 원자료 (FTF raw)
 [2] FTF(면대면) 표본 인구·지리적 cell 가중치 적용 (FTF cell)
 [3] Web 표본 원자료 (Web raw)
 [4] Web 표본 인구·지리적 rim 가중치 적용 (Web rim)
 [5] Web 표본 rim 성향 가중치 적용 (Web P.rim)

** [4]와 [5] 중에서 [2]에 더 가까운 추정치.

4) 10개 평가항목에서 결과 간 차이의 제곱 합은 다음과 같다 (괄호 안 숫자는 기준변수의 수).
 - Web PS(9) 대 Web propensity rim(9): 0.0402
 - Web PS(7) 대 Web propensity rim(7): 0.0032
 - Web PS(9) 대 Web PS(7): 0.1668
 - Web propensity rim(9) 대 Web propensity rim(7): 0.3040

<표 8> rim 가중법에 의한 성향보정의 결과: 보정기준 변수를 7개로 줄인 경우

	[1]* FTF raw	[2] FTF cell	[3] Web raw	[4] Web rim	[5] Web P.rim	비고**
TV 평일	2.891	2.841	2.528	2.674	2.757	[5]
TV 토	3.608	3.600	3.518	3.611	3.669	[4]
TV 일	3.925	3.922	3.848	3.917	3.990	[4]
해외여행경험	12.2%	13.6%	15.5%	15.3%	13.6%	[5]
전반적만족감	3.186	3.170	3.182	3.201	3.207	-
여가활용만족도	3.134	3.128	2.970	2.988	3.025	[5]
장래소득기대	2.935	2.875	2.490	2.591	2.627	[5]
주관적건강평가	2.831	2.789	2.602	2.697	2.735	[5]
스트레스 인식	2.541	2.518	2.306	2.337	2.352	[5]
생활시간태도	2.653	2.611	2.493	2.585	2.623	[5]

* [1] FTF(면대면) 표본 원자료 (FTF raw)

[2] FTF(면대면) 표본 인구·지리적 cell 가중치 적용 (FTF cell)

[3] Web 표본 원자료 (Web raw)

[4] Web 표본 인구·지리적 rim 가중치 적용 (Web rim)

[5] Web 표본 rim 성향 가중치 적용 (Web P.rim)

** [4]와 [5] 중에서 [2]에 더 가까운 추정치.

지금까지의 연구 내용을 요약하면서 이 장을 마무리하면 다음과 같다.

- 1) 통계청 사회조사 Web 표본의 성향 보정을 위해서 보정기준 변수를 9개로 한 경우의 결과와 7개로 한 경우의 결과는 별 차이가 없다. 그러나 보다 안정적인 결과도출을 위해서 전 항목에 걸친 성향보정에서는 보정기준 변수를 9개로 둘 것이다 (다음 장에 기술함).
- 2) 성향점수모형 가중법에 의한 성향보정의 결과와 rim 가중법에 의한 성향보정의 결과는 별 차이가 없었다.
- 3) Web 표본의 성향보정에서 성공의 관건은 기준변수의 선정에 있다.
- 4) 통계청 사회조사 Web 표본의 성향보정 결과, 자원자 패널을 활용한 인터넷 조사가 면접원 조사를 일부 보완할 수 있는 대안적 방법으로 보여진다.

별첨 R 프로그램 리스트

- 1) panel_to_sample_version_1.r: 가중 Web 패널 리스트로부터 계통 추출
- 2) web_sample_rimweighting_version_1.r: Web 표본의 rim 가중화
- 3) FTF_sample_rimweighting.r: 면대면 표본의 rim 가중화
- 4) two_sample_comparison.r: 면대면 표본과 Web 표본을 합함
- 5) web_propensity_scoring_version_1.r: 성향점수모형 가중법에 의한 Web 표본 성향보정
- 6) rim_propensity_weighting_version_1.r: rim 가중법에 의한 Web 표본 성향보정

제 4 장 성향보정 방법의 유용성 검토

제3장에서는 인터넷 조사와 기준 조사의 결과를 차이를 제거하는 모형을 구축하였는데, 7개 및 9개 변수를 기준변수로 하는 모형이 웹조사와 기준 조사간의 차이를 보다 많이 감소시키는 것으로 나타났다. 여기에서는 이처럼 구축된 모형이 과연 어느 정도의 유용성을 갖는가를 검토하기 위해서 통계청 사회조사에 포함된 106개의 문항에 모형을 적용하고 그 결과를 분석해보았다.

제 1 절 검토방법

2009년 사회조사에 포함된 조사 항목 중에서 기준변수로 사용되지 않은 대부분의 문항을 평가변수로 활용하였는데, 이 평가변수들은 분야에 따라 다음 <표 9>와 같이 구분된다.

모형의 평가는 면대면 조사결과를 기준으로 했을 때 모형의 적용이 웹조사와 면대면조사의 차이를 얼마나 감소시키는가를 기준으로 하였다. 그런데 각 문항별로 여러 개의 응답유목이 있는 경우, 같은 성격의 응답유목을 통합하여 분석하였다. 즉 ‘매우 그렇다’와 ‘그렇다’와 같은 두 유목의 응답을 통합해서 계산하였다. 구체적인 응답유목의 통합방법은 부록 2 <문항별 응답 % 산정방법>에 나와 있다. 이처럼 유목을 통합해서 분석한 것은 개별 유목별 빈도가 적은 경우가 많았기 때문이다.

<표 9> 유형별 평가변수

유 형	문 항 내 용	비 고
매체 접촉행태(11개)	tv시청시간, 독서시간, 잡지류, 교양서적, 직업관련 서적, 생활관련 서적, 인터넷 조사의향, 기타	
관광, 여가, 문화활동 (31개)	골프장, 국내관광여행 정도, 해외여행경험, 관광, 어학연수, 해외여행유무, 레저시설 이용횟수, 관광명소, 온천장 및 스파, 스키장, 해수욕장, 삼림욕장, 놀이공원, 수영장 워터파크, 기타, 문화예술 및 스포츠, 음악, 연극, 무용, 영화, 박물관, 미술관, 스포츠, 주말여가활동, 여가활동	
인간관계 및 생활의례 (8개)	생활의례, 인간관계 및 예절, 민속놀이, 전통예술, 전통복장, 전통건축, 전통무예	
소득, 소비, 근로 (12개)	소득만족도, 장래소득기대, 계층의식, 계층이동, 근로시간	
사회관계 (7개)	아플 경우 도움청할 사람, 돈 빌릴 경우 도움 청할 사람, 낙심할 경우 도움 청할 사람	
봉사, 자선활동 (18개)	후원금, 자원봉사활동 등	
생활 만족도 (19개)	건강, 노동, 교육 등 생활만족도	

제 2 절 가중의 효과

인터넷 조사를 함으로써 발생하는 편향이 앞서 구축된 모형을 이용하여 성향가중을 실시한 경우 과연 어느 정도나 제거될 수 있는가?

이를 알아보기 위해 평가변수로 선정된 106개 변수에 대해 제3장에서 설명한 가중방법 중 세 가지 방법 즉 9개 변수 림(rim) 가중법, 7개 변수 로

짓성향 가중법, 9개 변수 로짓성향 가중법 등을 적용하였다.

이 외에도 다른 가중방법을 하나 더 사용하였다. 그것은 인터넷 사용시간과 귀가시간을 토대로 산출한 림가중치다. 인터넷 사용시간 및 귀가시간을 위한 기준조사자료로는 2004년에 이루어진 생활시간조사 자료를 이용하였다. 이와 같은 시간가중치를 적용한 것은 인터넷 표본과 면대면 표본의 차이가 이러한 인터넷 사용에서 비롯되었을 가능성이 있다고 보고, 이것을 통제하는 가중치가 과연 인터넷 조사의 편향을 제거할 수 있는지 살펴보기 위한 것이었다.

시간 변수의 활용방법

본 연구에서 생활시간 사용을 활용하는 것은 다음과 같은 이유에서다.

- 귀가시간, 취침시간 등 생활시간의 이용 행태는 일반적인 라이프 스타일과 관련성이 높을 것이다. 그리고 인터넷 사용과도 관련성이 높을 것이다.
- 인터넷 사용시간은 웹설문에 대한 응답성향과 관련성을 보일 것이다.
- 생활시간은 통계청에서 주기적으로 (현재 5년) 실시하기 때문에 기준 조사자료를 얻기에 편리하다
- 일반 조사에서 사용하기에 편리하다. 즉 질문과 응답이 어렵지 않다.

따라서 생활시간 변수를 이용하여 가중모형을 구성할 경우 어느 정도 웹사이트의 편향을 제거할 수 있을 것으로 보았다. 그러나 이번 연구에서는 웹 설문에만 문항을 추가할 수 있었는데, 문항 수의 제한 때문에 귀가시간, 취침시간, 그리고 인터넷 사용시간(정보탐색, 사회적 교제, 게임, 웹관리 등)만을 사용하였다. 생활시간 조사는 현재 5년 간격으로 진행되는데, 가장 최근에 실시된 조사는 2009년이지만 조사 결과는 2010년에 발표된다. 따라서 본 연구에서는 2004년 자료를 이용하였다.

구체적으로 각 시간변수는 다음과 같이 유목화하여 사용하였다.

- 귀가시간 : 오후 6시, 8시, 10시, 12시
- 취침 시간 : 10시, 12시
- 인터넷 사용시간 : 4개 사용용도별로 인터넷 사용 여부

이와 같이 분석해서 웹조사 결과와 기준조사결과 즉 면대면 조사결과간의 차이가 성향가중 모형을 적용했을 때 어느 정도 감소되는가를 문항별로 살펴보면 다음 표에서 보는 바와 같다.

<표 10> 웹조사결과에 성향가중치를 적용했을 때의 결과 *

유형	문항		면대면	기본가중**	9개 립	9개 로짓	시간 립	7개 로짓
매체 접촉행태	T V	시청유무	96.4%	93.9%	93.0%	93.0%	91.1%	93.0%
		평일	41.5%	38.9%	40.1%	40.1%	42.3%	40.1%
		토요일	56.9%	58.5%	57.9%	57.9%	57.5%	58.1%
		일요일, 공휴일	62.6%	62.9%	62.5%	62.6%	62.1%	62.7%
	독 서	독서유무	62.1%	74.8%	66.9%	67.0%	53.6%	68.2%
		잡지	19.9%	24.0%	19.6%	19.7%	14.8%	20.5%
		교양서적	28.4%	34.2%	27.9%	27.9%	19.7%	28.9%
		직업관련	15.4%	23.5%	18.0%	17.9%	8.8%	18.3%
		생활, 취미	8.3%	15.1%	11.6%	11.5%	9.5%	12.1%
		기타	10.0%	15.5%	13.9%	14.0%	8.2%	14.1%
인터넷조사의향		16.9%	87.1%	83.9%	84.0%	78.2%	84.2%	
관광, 여가, 문화 활동	국내관광여행 유무		64.8%	66.9%	60.6%	60.5%	51.0%	61.7%
	회 수	당일여행	14.4%	19.6%	16.5%	16.4%	13.5%	17.2%
		숙박여행	27.6%	36.2%	31.2%	31.1%	24.6%	32.0%
	해외여행유무		13.6%	15.2%	13.3%	13.1%	10.3%	13.6%
	해외 여행 회수	관광	2.2%	1.7%	1.4%	1.4%	1.4%	1.7%
		가사	0.7%	1.3%	1.2%	1.2%	0.8%	1.2%
		업무	1.5%	1.6%	1.2%	1.2%	0.6%	1.3%

* 각 문항별 특정 유목에 대한 응답비율 (부록 2 참조).

** 웹조사 결과에 연령, 성, 거주지역의 크기 및 거주지의 시도를 기준으로
가중치를 부여

<표 10> 웹조사결과에 성향가중치를 적용했을 때의 결과(계속) *

유형		면대면	기본가중**	9개 립	9개 로짓	시간가중	7개 로짓	
관광, 여가, 문화 활동	어학연수 회수	0.1%	0.5%	0.3%	0.3%	0.5%	0.4%	
	해외여행유무 (평생)	29.3%	37.5%	34.4%	34.3%	30.7%	34.5%	
	레저시설이용 여부	69.0%	74.0%	67.6%	67.5%	58.8%	68.6%	
	이용 회수	관광명소	27.3%	35.7%	30.5%	30.4%	23.0%	31.5%
		온천장/ 스파	13.5%	18.6%	15.6%	15.6%	12.9%	15.9%
		골프장	3.1%	3.0%	2.1%	2.1%	0.9%	2.3%
		스키장	3.6%	4.9%	3.8%	3.8%	2.5%	4.3%
		해수욕장	10.3%	16.9%	14.2%	14.2%	9.4%	14.9%
		산림욕장	9.2%	14.5%	11.7%	11.6%	8.2%	12.4%
		놀이공원	11.0%	16.2%	13.9%	13.9%	9.7%	14.3%
		수영장 워터파크	7.5%	10.1%	8.4%	8.3%	5.5%	8.7%
		기타횃수	0.3%	5.2%	3.8%	3.8%	2.1%	3.9%
	문화예술 및 스포츠관람	55.1%	65.4%	58.1%	58.0%	44.9%	58.6%	
	이용 회수	음악	6.7%	10.5%	7.8%	7.7%	5.3%	8.4%
		연극	6.5%	13.8%	10.6%	10.5%	5.9%	11.1%
		무용	0.5%	1.2%	0.9%	0.9%	0.4%	1.0%
		영화	40.4%	50.2%	42.8%	42.7%	30.1%	43.3%
		박물관 (유물전시회)	7.4%	15.0%	11.6%	11.6%	7.4%	12.1%
미술관 (공예)		5.3%	9.7%	7.1%	7.0%	3.5%	7.4%	
스포츠		7.6%	11.6%	9.3%	9.2%	5.5%	9.7%	

* 각 문항별 특정 유목에 대한 응답비율 (부록 2 참조).

** 웹조사 결과에 연령,성, 거주지역의 크기 및 거주지의 시도를 기준으로
가중치를 부여

<표 10> 웹조사결과에 성향가중치를 적용했을 때의 결과(계속) *

유형		면대면	기본가중**	9개 립	9개 로짓	시간 립	7개 로짓
관광, 여가, 문화 활동	주말여가활동	57.1%	42.7%	42.2%	42.2%	39.6%	42.2%
	주말여가활동 동반자	58.6%	64.9%	63.9%	64.0%	65.0%	64.4%
	희망여가활동	11.5%	18.9%	20.9%	20.9%	23.0%	20.6%
	여가활동만족 도	21.8%	30.8%	26.0%	25.7%	24.5%	28.0%
인간 관계 및 생활 의례 (전통문화의 영향)	생활의례	30.8%	41.6%	38.3%	38.2%	36.5%	39.1%
	인간관계및예 절	52.3%	65.4%	59.8%	59.8%	56.3%	60.9%
	민속놀이	6.8%	10.4%	10.0%	10.0%	12.0%	10.5%
	전통예술	7.4%	10.9%	10.2%	10.2%	12.2%	10.7%
	전통음식	32.4%	36.9%	33.0%	33.0%	31.3%	34.1%
	전통음식	9.4%	16.7%	16.2%	16.2%	18.3%	16.6%
	전통건축	12.9%	20.2%	18.8%	18.8%	19.7%	19.4%
	전통무예	4.2%	7.6%	7.2%	7.2%	8.0%	7.5%
소득, 소비, 근로	소득여부	65.8%	64.9%	60.9%	60.8%	55.3%	61.4%
	소득만족도	30.7%	31.1%	29.9%	30.1%	28.5%	29.3%
	장래소득기대	17.3%	11.7%	10.9%	11.0%	11.4%	11.0%
	계층의식	36.7%	42.9%	42.5%	42.6%	39.0%	42.0%
	계층이동전망	37.6%	34.4%	33.0%	32.7%	30.5%	34.0%
	계층이동자식 세대	48.3%	46.5%	45.0%	44.9%	44.3%	46.0%
	주간근로시간	46.0%	36.1%	34.4%	34.3%	30.3%	34.5%

* 각 문항별 특정 유목에 대한 응답비율 (부록 2 참조).

** 웹조사 결과에 연령, 성, 거주지역의 크기 및 거주지의 시도를 기준으로
가중치를 부여

<표 10> 웹조사결과에 성향가중치를 적용했을 때의 결과(계속) *

유형		면대면	기본가중**	9개 립	9개 로짓	시간 립	7개 로짓	
소득, 소비, 근로	주간근로시간 (배우자)	11.0%	14.0%	13.0%	13.0%	13.0%	13.0%	
	1주 경활상태	55.8%	52.2%	49.2%	49.0%	42.8%	49.4%	
	종사상의 지위	38.1%	40.7%	37.5%	37.3%	31.1%	37.6%	
	1주 경활상태 (배우자)	14.2%	37.6%	35.8%	35.9%	34.6%	36.1%	
	가구소득	20.5%	41.7%	42.3%	42.4%	42.0%	41.8%	
사회 관계	주관적 만족감	33.8%	37.1%	37.8%	38.3%	37.5%	36.3%	
	아플 경우	76.6%	80.7%	79.3%	79.3%	77.9%	79.7%	
	사람 수	20.2%	31.3%	31.1%	31.1%	33.6%	31.9%	
	돈 빌릴 경우	48.6%	67.7%	65.8%	65.7%	64.0%	66.7%	
	사람 수	12.9%	27.6%	26.4%	26.3%	26.7%	27.3%	
	낙심할 경우	81.3%	84.0%	82.4%	82.4%	80.8%	82.7%	
	사람 수	33.9%	41.8%	39.3%	39.2%	38.6%	40.1%	
봉사, 자선활동	후 원 금 기 부	유무	32.3%	34.2%	28.0%	27.9%	19.3%	28.6%
		직접후원	3.2%	5.4%	4.1%	4.1%	2.2%	4.3%
		언론기관	7.3%	6.8%	5.5%	5.5%	3.8%	5.7%
		사회복지 단체	6.0%	8.0%	5.6%	5.6%	3.9%	5.9%
		종교단체	5.8%	8.1%	6.7%	6.7%	5.2%	6.8%
		기업직장	3.7%	5.7%	4.3%	4.3%	2.6%	4.6%
		기타	0.5%	2.5%	1.7%	1.7%	1.0%	1.7%

* 각 문항별 특정 유목에 대한 응답비율 (부록 2 참조).

** 웹조사 결과에 연령,성, 거주지역의 크기 및 거주지의 시도를 기준으로
가중치를 부여

<표 10> 웹조사결과에 성향가중치를 적용했을 때의 결과(계속) *

유형		면대면	기본가중**	9개 립	9개 로짓	시간 립	7개 로짓
봉사, 자선활동	단체참여	39.8%	40.8%	35.4%	35.3%	28.0%	36.1%
	자원봉사활동	19.3%	21.4%	18.4%	18.4%	13.4%	19.0%
	환경 분야 봉사	4.9%	5.4%	4.7%	4.7%	3.6%	4.8%
	국가행상봉사	2.2%	5.1%	4.2%	4.2%	2.3%	4.3%
	교육관련 봉사	2.3%	5.2%	4.3%	4.4%	2.0%	4.4%
	아동관련 봉사	8.4%	11.7%	10.1%	10.1%	7.5%	10.3%
	재해관련 봉사	2.2%	3.3%	2.7%	2.7%	1.3%	2.8%
	기타 봉사	3.1%	3.8%	2.8%	2.8%	1.1%	2.9%
	자원봉사활동 의사	52.2%	50.0%	50.9%	51.0%	48.8%	50.9%
	민원서비스만족 여부	55.9%	52.5%	45.6%	45.5%	33.7%	46.1%
	만족도	23.0%	27.4%	21.9%	21.8%	17.0%	22.5%
생활 만족도	주관적건강평가	38.1%	42.1%	38.8%	38.6%	34.2%	40.2%
	스트레스인식 정도	43.9%	54.2%	54.6%	54.8%	51.9%	53.5%
	주택만족도	15.7%	39.7%	36.9%	36.6%	36.9%	38.7%
	거주지역만족도	37.9%	43.3%	41.0%	40.8%	41.1%	42.1%
	사회안전인식	15.8%	27.8%	25.1%	24.9%	25.0%	25.7%
	범죄피해두려움	44.5%	46.6%	43.9%	44.1%	42.3%	44.4%
	가족관계만족도	53.4%	53.3%	49.1%	49.0%	48.7%	51.0%
	환경오염체감	14.9%	18.5%	18.0%	17.9%	17.9%	18.4%

* 각 문항별 특정 유목에 대한 응답비율 (부록 2 참조).

** 웹조사 결과에 연령, 성, 거주지역의 크기 및 거주지의 시도를 기준으로 가중치를 부여

<표 10> 웹조사결과에 성향가중치를 적용했을 때의 결과(계속) *

	유형	면대면	기본가중**	9개 립	9개 로짓	시간 립	7개 로짓	
생활 만족도	자살중동경험 여부	14.7%	20.5%	18.7%	18.9%	16.0%	18.5%	
	삶의 질	소득소비	76.9%	81.8%	81.4%	81.5%	80.1%	81.2%
		건강	86.9%	89.0%	87.9%	88.0%	87.2%	88.1%
		노동	49.7%	52.7%	52.2%	52.2%	54.7%	52.1%
	구성요소	교육	67.0%	69.9%	67.9%	67.8%	67.4%	68.2%
		주거교통	54.8%	59.3%	59.0%	59.1%	59.7%	59.4%
	중요도	안전	69.3%	73.9%	73.0%	73.0%	73.0%	73.3%
		가족	88.6%	89.5%	87.9%	88.0%	85.5%	88.5%
		환경	59.9%	64.4%	61.1%	61.1%	60.0%	61.6%
		사회통합	38.2%	45.6%	43.3%	43.4%	43.0%	43.7%
	여가	44.6%	57.3%	55.0%	55.0%	53.0%	55.8%	

* 각 문항별 특정 유목에 대한 응답비율 (부록 2 참조).

** 웹조사 결과에 연령, 성, 거주지역의 크기 및 거주지의 시도를 기준으로 가중치를 부여

이러한 결과를 분석해서 성향가중 방법별로 효과를 비교해 보면 9개 변수 가중방법이 웹조사와 기준조사간의 차이를 가장 크게 줄이는 것으로 나타났다. 먼저 웹 조사결과에 성, 연령, 거주지역의 크기, 시도 등을 고려한 기본 가중치만을 적용했을 경우, 웹조사와 기준조사결과간의 차이는 106개 문항 평균 5.9%point였다. 그러나 9개 변수 가중모형을 적용했을 경우에는 이 차이가 4.6%point로 감소하였다. 7개 변수 가중모형의 경우는 이보다 약간 큰 4.8%point 그리고 시간변수 가중모형의 경우는 5.5%point로 나왔다. 이처럼 성향가중치를 구하는 방법에 따라 효과가 다르기는 하지만 다음 <표 11>에서 보는 바와 같이 성향가중을 적용하면 웹조사 결과가 기준조사결과

에 보다 근접하게 되는 것으로 나타났다.

<표 11> 성향가중 적용후의 웹조사결과와 기준조사(면대면 조사) 결과의 차이

	문항수	웹조사와 기준조사간의 차이			
		최소	최대	평균	표준편차
기본가중 모형*	106	.00	.70	.0592	.07908
9개변수 림가중	106	.00	.67	.0460	.07587
시간변수 림가중	106	.00	.61	.0553	.07569
7개 변수 로짓성향분석	106	.00	.67	.0476	.07664
9개 변수 로짓 성향분석	106	.00	.67	.0460	.07592

* 웹 조사에 성, 연령, 거주 지역의 크기 및 시도에 따른 가중치만을 적용

그리고 3장에서 살펴본 바와 같이 림가중 방법과 로짓 성향분석 방법은 거의 동일한 결과를 보여주고 있다. 두 방법을 각각 적용한 다음 남아있는 차이간의 상관도를 구해보면 $r=1.0$ 으로 나타났다.

가중치의 효과는 기준변수의 수에 따라 차이가 있는 것으로 나타났다. ($t=2.415$, $df=105$, $p=0.017$). 따라서 9개 변수를 사용하는 모형이 웹조사결과와 기준조사간의 차이를 더 많이 줄여주고 있다.

한편 인터넷 사용시간 등에 기초한 시간가중치의 경우는 비록 효과가 있기는 했지만 그다지 의미있는 차이를 보이지 못하고 있다. 그러나 시간가중의 준거로 사용된 자료가 2004년도 자료였다는 점을 고려할 때 이 결과만을 가지고 시간 가중의 유용성에 대해 결론을 내리기는 어렵다고 본다.

제 3 절 변수 유형 및 연령 집단별 비교

그러면 이와 같은 결과가 각 변수의 유형별로 보면 어떻게 될까? 즉 유형에 관계없이 일정하게 나타나는가 아니면 다르게 나타나는가?

**<표 12> 문항유형별 성향가중 효과
(비율)**

성향가중 모형		문항 수	웹조사와 기준조사간의 차이 평균	표준편차
매체접촉	기본모형	11	0.109	0.20
	9개 립	11	0.080	0.20
	시간가중	11	0.091	0.18
	7개 로짓	11	0.083	0.20
	9개로짓	11	0.080	0.20
관광, 여가	기본모형	31	0.050	0.04
	9개 립	31	0.029	0.03
	시간	31	0.037	0.05
	7개 로짓	31	0.032	0.03
	9개 로짓	31	0.029	0.03
인간관계	기본모형	8	0.067	0.04
	9개 립	8	0.047	0.03
	시간 가중	8	0.050	0.02
	7개 로짓	8	0.053	0.03
	9개 로짓	8	0.047	0.03
소득소비	기본모형	12	0.068	0.08
	9개 립	12	0.075	0.07
	시간 가중	12	0.093	0.07
	7개 로짓	12	0.072	0.07
	9개 로짓	12	0.076	0.07
사회관계	기본 모형	7	0.090	0.06
	9개 립	7	0.078	0.06
	시간 가중	7	0.075	0.06
	7개 로짓	7	0.082	0.07
	9개 로짓	7	0.078	0.06
봉사, 자선활동	기본모형	18	0.021	0.01
	9개 립	18	0.019	0.02
	시간 가중	18	0.042	0.06
	7개 로짓	18	0.018	0.02
	9개 로짓	18	0.019	0.02
생활만족	기본 모형	19	0.061	0.06
	9개 립	19	0.048	0.05
	시간 가중	19	0.048	0.05
	7개 로짓	19	0.050	0.05
	9개 로짓	19	0.048	0.05

위 표에서 보듯이 변수 영역별로 각 모델의 적용 결과가 달라지는데 특히 소득, 소비, 근로 등과 관련된 문항에서는 본 연구의 모형이 오히려 웹조사와 기준조사간의 차이를 증가시키고 있다. 그리고 봉사 자선활동의 경우에는 비록 그 차이가 증가하지는 않지만 감소시키지도 못하고 있다. 따라서 본 연구의 모형은 여가, 문화,관광, 매체접촉, 사회관계 등에서 성과가 좋았다고 볼 수 있다.

이러한 모형의 효과가 조사대상자에 따라서는 어떻게 달라질까? 웹조사는 특히 연령에 따라 활용도가 달라진다고 지적되어 왔다. 젊은 층일수록 웹 활용도가 높고, 따라서 웹에서 수집된 자료의 일반화 가능성도 젊은 층에서 높다는 것이다. 실제로 본 조사의 경우에도 그러할까? 이를 살펴보기 위해 웹 및 면대면 조사자료를 10-20대 그룹, 30-40대 그룹, 그리고 50세이상의 그룹으로 나눈 다음, 각 연령 집단별로 본 연구의 가중모형의 효과를 살펴보았다. (표 13)

즉 각 연령 집단에서 7개변수 로짓 성향가중을 실시했을 때 웹조사와 기준조사결과간의 차이가 어느 정도 감소하는가를 살펴보았는데, 10-20대의 경우는 로짓 성향모형을 적용하게 되면 차이가 2.4%point로 감소한다. 반면에 30-40대의 경우는 3.6%point 그리고 50세 이상의 경우는 3.2%point로 되었다. 이것은 10-20대의 경우는 웹표본의 대표성이 그만큼 높았기 때문으로 보인다. 즉 10-20대의 경우 웹조사 참여자가 많았고, 그 결과 성향보정을 통하여 기준조사 결과와의 차이가 많이 줄어드는 것이 가능했다고 본다. 상대적으로 30세 이상 집단의 경우, 참여자가 적었고, 따라서 성향가중의 이러한 편향 제거효과가 적었을 것으로 보인다. 이 같은 결과는 웹조사의 대표성이나 활용성이 젊은 층에서 더 높다는 점을 보여준다. 그리고 10-20대 집단이 일반 전화나 면접원 방문으로 접근하기 어려운 집단에 속하기 때문에 웹조사를 부분적으로 도입할 경우는 이처럼 젊은 층을 대상으로 하는 것이 적절하다고 본다.

<표 13> 연령집단별 성향가중의 효과 차이

연 령	웹조사와 기준 조사간의 차이	t	p
10-20대	2.4%point	3.014	0.003
30-40대	3.6%point	4.142	0.000
50세 이상	3.2%point	3.333	0.001

제 4 절 모형 효과의 크기

이상에서 살펴본 바와 같이 본 연구에서 구축한 가중모형은 웹조사와 기준조사간의 차이를 감소시키고 있다. 그렇다면 이것은 실제 의미있을 만큼 차이를 크게 감소시키고 있는가? 즉 웹조사가 성향 가중모형을 적용하게 되면 기준조사 결과에 근접하게 되는가? 아니면 그렇지 못하다고 보아야 하는가?

이것을 알아보기 위해 각 문항별로 웹조사와 기준조사간의 차이가 어느 정도가 감소되고 있는가를 살펴보았다. 다음 <표 14>에서 보듯이 차이가 감소된 경우는 전체 문항의 79%정도였다. 즉 10개중 8개에서 차이가 감소된 것이다. 그러나 원래 있었던 차이의 반 이상이 줄어든 경우를 보면 36% 수준이다. 차이 감소가 그다지 크지 않았다고 볼 수 있다. 즉 본 연구의 모형은 기준조사와 웹조사간의 차이를 감소시켜 주고는 있지만 그 감소의 크기가 부족한 편이라고 볼 수 있다.

<표 14> 성향가중 모형의 효과 크기

성향가중 방법	문 항 수	차이가 감소된 경우	50%이상 감소된 경우
9개변수 립가중	106	79.3%	35.9%
시간 가중	106	57.6%	33.0%
7개변수 로짓성향	106	78.3%	26.4%
9개변수 로짓성향	106	79.3%	35.9%

제 5 장 요약 및 결론

웹조사의 경우 전통적인 조사방법의 한계를 극복할 수 있는 대안으로 주목받아왔지만 지금까지 표본의 대표성 문제 때문에 활용이 제한되어 왔다. 그럼에도 불구하고 인터넷의 확대 그리고 전통적인 조사방법의 문제점과 더불어 인터넷 조사방법에 대한 관심도는 더욱 더 높아지고 있다. 본 연구는 이러한 맥락에서 웹조사 방법의 문제점을 극복하는 방안을 찾아보고 검증해보았다.

제 1 절 요약 및 제한점

제3장에서 구축한 모형을 106개 문항에 확대 적용해서 유용성을 살펴본 결과 다음과 같이 요약될 수 있다.

- 가중모형은 웹조사와 기준조사간의 차이를 감소시킨다. (1.1%p에서 1.3%p)⁵⁾
- 그러나 감소폭은 그다지 크지 않다. 감소폭이 원래 차이의 1/2에 미치지 못하는 경우가 더 많다.(35% vs 43%)⁶⁾
- 가중모형의 효과는 소득/소비, 자원봉사 영역에서는 나타나지 않고 있고, 반면에 레저, 매체이용행태, 삶의 질 만족도 등의 영역에서는 나타나고 있다. 즉 가중모형의 효과는 적용분야에 따라 달라진다.
- 가중모형의 효과는 10-20대의 경우에 30세 이상 집단보다 크게 나타난다. 즉 웹조사 자료의 대표성이 젊은 층에서 더 높게 확보될 수 있다.
- 모형 구축방법으로 림가중치와 로짓 성향분석 방법간에는 차이가 없다.
- 7개변수 모형보다 9개 변수 모형이 더 감소폭이 크지만 그 차이는 미미하다.
- 시간변수를 활용한 경우에도 가중의 효과가 나타나고는 있지만 다른 가중 모형에 비해 적다.

5) 가중모형을 적용하면 웹조사와 기준조사간의 차이가 1.1%~1.3%point감소하고 있다.

6) 웹조사와 기준조사간의 차이가 1/2이상이 제거된 경우는 35%다. 반면 감소폭이 원래 차이의 1/3이 되지 못

본 연구는 연구 방법상 다음과 같은 한계를 지닌다. 첫째, 웹조사를 위한 응답자를 웹에 광고문을 게시함으로써 모집했다는 점이다. 이 경우 어떤 사이트를 이용하느냐에 따라 응답자의 특성이 달라질 수 있기 때문에 이번 조사 결과를 일반화하는 데는 한계가 있다. 둘째로, 통계청에서 정기적으로 실시되는 사회조사의 일부 문항을 웹으로 조사한 다음 이 웹 조사결과와 면대면 가구방문조사를 비교한 것이다. 그렇기 때문에 가중모형을 개발하기 위한 다양한 문항들을 시도해 보지 못했다. 원래의 통계청 면대면 조사 문항에 연구 목적에 필요한 모든 문항을 포함시킬 수는 없었기 때문이다.

이러한 제한점에도 불구하고 본 연구는 오프라인조사와 온라인조사를 동시에 실시하여 자료를 비교함으로써 웹조사와 기준조사의 차이가 과연 성향가중을 통해 감소될 수 있는지 직접적으로 검토할 수 있었다는 장점이 있다.

제 2 절 결 론

이와 같은 연구결과를 토대로 웹조사 방법의 활용가능성에 대해 결론을 내려본다면 다음과 같다.

첫째, 본 연구의 결과는 성향가중 모형의 유용성을 보여준다고 할 수 있다. 지금까지 많은 연구에서 성향가중 모형은 웹조사의 편향을 제거하지 못하는 것으로 나타났다. 그러나 본 연구에서는 비록 그 크기가 충분하지는 않지만 80%의 문항에서 긍정적으로 기여하는 것으로 나타났다. 지금까지의 웹성향 가중치의 연구들이 평가변수로 불과 서너개의 변수만을 사용했고, 또 기준변수로는 주로 인구통계학적 변수들을 사용했다는 점을 고려하면, 본 연구가 과거의 연구들보다 성향가중 모형의 타당성과 유용성을 살펴보는데 더 적합하다고 볼 수 있다. 본 연구에서는 다양한 분야에서 106개 문항을 평가변수로 사용했고, 기준변수로는 인구적 속성 외에 매체접촉 행동, 소득만족도 등을 사용했기 때문이다. 비록 이번 연구에서 편향 감소폭이 크지 않았지

하는 경우는 43%다.

만 이것은 앞으로 좀더 효과적인 기준변수를 선택함으로써 개선될 수 있을 것으로 본다. 따라서 지금까지의 연구와는 달리 비확률 표본을 이용한 가중 모형 적용결과 대표성을 확보하는데 타당성이 있다고 보여준다.

둘째, 가중모형에 포함되는 기준변수는 평가변수에 따라 즉 적용분야에 따라 달리 선택되어야 한다. 본 연구에서 모형의 효과는 평가변수의 유형에 따라 달라졌다. 따라서 이 점은 모형의 적용분야에 따라 기준변수를 달리 선택해야 한다는 것을 의미한다. 물론 각각의 분야에 적용되는 여러 기준변수를 모아 하나의 모형에 모두 포함시키는 방안도 가능할 것이다. 그러나 9개 변수 모형과 7개 변수 모형이 편향 감소율에서 큰 차이를 보이지 않았다는 점은 모형 기여도가 적은 변수를 기준변수로 포함시키는 것이 바람직하지 않다는 점을 보여준다. 따라서 각 분야별로 기준변수를 달리 선정해야 할 것으로 보인다.

이 같은 점은 일반 범용의 목적으로 가중모형을 만들기 보다는 특정 분야에 적용할 모형이 보다 유용할 것으로 보인다. 예컨대 정치분야, 과학분야, 소비행동 분야 등과 같이 특정 분야별로 모형을 구축하는 것이 보다 쉽게 모형을 구축할 수 있는 방법일 것이다.

셋째, 10대와 20대의 경우에 모형의 효과가 크게 나타났다는 점은 이러한 성향가중 모형이 젊은 층에 보다 적합하게 적용될 수 있다는 것을 의미한다. 따라서 전국민을 대상으로 하는 조사보다는 젊은 층 특히 인터넷 사용도가 높은 층에서 이러한 모형의 타당성이 높을 것으로 보인다. 또 이같은 결과는 앞으로 인터넷 사용이 더욱 확대되면서 성향가중 모형의 유용성도 높아진다는 점을 보여준다.

넷째, 이 같은 점은 또한 웹조사를 위한 응답자 모집이 가능하면 보다 다양한 층이 응답할 수 있도록 해야 한다는 것을 의미한다. 이번 연구의 경우 웹 응답자의 모집은 통계청 홈페이지 등을 활용해서 모집했는데, 다양한 응답자를 모집하기 위해서는 보다 일반인들이 활용하는 홈페이지를 이용하거나 아니면 이미 구축된 패널을 활용하는 방안도 검토할 필요가 있다.

다섯째, 성향가중 모형의 구축에서 립가중 모형과 로짓성향 분석 모형간에 차이가 없다는 점은 성향가중 모형의 적용성을 더욱 높여준다. 로짓-성향가중 모형을 적용하기 위해서는 무엇보다 준거조사의 원자료가 필요했기 때문이다. 그러나 립가중치를 적용하기 위해서는 원자료가 필요없이 요약된 자료만으로도 충분하다. 이 점은 센서스 또는 통계청이나 기타 대규모 조사결과에서 사용된 많은 통계치를 보다 쉽게 활용할 수 있다는 것을 의미한다. 아울러 로짓-성향분석의 경우에는 기준변수가 하나의 준거조사에 모두 포함되어 있어야만 분석이 가능했다. 이 점 역시 준거변수의 선택을 제한하는 큰 요인이 되고 있다. 그러나 립가중치를 사용할 경우 여러 개의 조사에서 나온 결과를 한꺼번에 활용할 수 있게 된다. 즉 기준변수의 선정이 보다 확대되었다는 것을 의미하며, 이것이 가중방법의 발전에 크게 기여할 것으로 보인다.

구체적으로 본 연구의 활용 방안을 직접적인 측면과 부차적인 측면에서 생각해 보면 다음과 같다.

- 1) 직접적 활용: 매년 거듭되는 사회조사에서 현재의 면접조사의 일정 규모를 인터넷 자원자 조사로 점차 대체해 가는 방안이다. 인터넷 조사가 접촉이 어려운 사회구성원에 접근할 수 있고 사생활 침해를 염려하는 응답자들에게 심리적 부담을 줄여 줄 것이다. 그러나 인터넷 자원자 조사로 면접 조사를 완전히 대체할 수는 없다. 성향조정의 준거가 없어지기 때문이다.
- 2) 간접적 활용: 성향조정의 준거가 되는 2009년 사회조사 FTF 표본 원자료 및 기준변수의 분포를 공개하여 통계청 및 기타 국가/민간 기관이 시행하는 각종 인터넷 조사 표본의 성향조정에 활용하는 방안이다. 이런 방식의 활용을 위해서는 본 연구에서 성향보정의 기준으로 제시한 7개 또는 9개 문항이 인터넷 조사 설문지에 포함되어야 한다.

- 성향점수모형 가중법을 활용하려면 2009년 사회조사 FTF 표본 7개

또는 9개 변수 원자료가 필요하고 이를 인터넷 조사 표본에 붙여 성향 점수모형을 만들어 성향보정 가중치를 생성하면 된다.

- 성향보정 rim 가중법을 활용하기 위해서는 2009년 사회조사 cell 가중치 부여 FTF 표본으로부터 7개 또는 9개 변수의 주변 분포 비율을 제공받아 rim 가중법을 적용하면 된다.

따라서 후자의 성향보정 방법이 더 간편할 것이다.

또 다른 활용 방안은 혼합방식(mixed mode) 조사에서 조사방식에 따른 차이를 규명하고 방식별 추정값들을 통합하여 하나의 추정치를 제공하는 데 있어 본 조사의 방법론을 응용하는 것이다. 향후 사회여건의 변화에 따라 혼합방식의 조사가 불가피해지는 만큼 본 연구의 방법론의 부가적 활용도가 있을 것으로 기대한다.

제 3 절 제 언

본 연구 결과는 웹조사 방법이 갖는 문제점은 후속 연구를 통해 극복이 가능하다는 점을 보여준다. 따라서 본 연구의 결과를 토대로 앞으로 웹조사 방법의 문제를 해결하는 방안 및 그를 위한 연구방향을 제시하면 다음과 같다.

연구내용	활용	연구방법	한계	장점
<ul style="list-style-type: none"> - 가중 모형을 위한 기준변수 개발연구 - 응답자 모집방법에 관한 연구: 웹게시 vs 패널 	<ul style="list-style-type: none"> - 비확률표본을 대상으로 실시된 웹조사의 편향을 제거한다. - 두 개 조사의 데이터를 결합해서 활용하고자 할 때. 	<p>통계청에서 실시되는 조사의 웹 버전 조사를 실시한다.</p>	<p>통계적 추정과정이 들어가기 때문에 조사의 객관성에 대한 논란이 있을 수 있다.</p>	<ul style="list-style-type: none"> - 다양한 조사에서 널리 활용될 수 있다. - 통계청 외의 민간부분에서 활용도가 높다
<p>인터넷 조사에서 확률표본 방법의 대표성 제고를 위한 연구</p> <ul style="list-style-type: none"> - 기존 통계청 조사참여자 대상으로 패널을 모집하여 특성 연구 - LISS와 같은 공공목적의 패널 구축 지원 	<ul style="list-style-type: none"> - 대표성 확보가 가능하기 때문에 통계조사에서 기존의 면대면 방법을 대신할 수 있다 	<p>응답자 패널을 RDD 방법을 이용해 모집하면서 패널과 거절자간의 성향차이를 분석한다</p>	<p>패널 구축비용이 크다</p>	<ul style="list-style-type: none"> - 타당성만 검토되면 바로 조사에 활용될 수 있다. - 조사의 타당성에 대한 논란이 적다
<p>멀티 모드 조사 방법 (동일표본)의 대표성 연구</p>	<p>전통적 방법의 약점인 젊은 층의 응답률을 높이는데 활용</p>	<p>통상적인 통계청 조사에서 멀티모드 조사를 적용한다</p>	<p>각 응답자에 대해 전화번호, 주소, 웹메일주소 등의 다양한 접근 방법을 구하기가 어렵다</p>	<p>조사의 타당성에 대한 논란이 없다</p> <ul style="list-style-type: none"> - 바로 활용이 가능하다
<p>멀티모드 (별도 표본)조사 결과 통합 방법 연구</p>	<p>젊은 층 등 전통적 조사방법으로 접근하기 어려운 집단을 조사하고자 할 때</p>	<p>성향 가중 방법을 적용한다</p>	<p>객관성에 대한 논란이 있을 수 있다</p>	<ul style="list-style-type: none"> - 활용도가 많다

이 표에서 보듯이 인터넷 조사방법의 활용도를 높이기 위해서 많은 연구가 필요하다. 그렇지만 그 중에서 특히 다음과 같은 연구가 우선적으로 요청된다고 하겠다.

가. 기준변수에 대한 연구

본 연구의 결과는 기준변수에 대한 연구가 좀더 필요하다는 점을 보여준다. 비록 본 연구에서 선정된 9개의 기준변수로도 어느 정도 타당한 모형을 구축할 수 있었지만 가중모형의 유용성을 좀더 높이기 위해서는 본 연구에서 선정된 9개 기준변수 외에 어떤 다른 기준변수들이 가능한지 찾아보는 노력이 필요하다. 이 점은 기준변수의 유용성이 적용되는 조사분야에 따라 달라진다는 점에서 더욱 중요해진다. 본 연구 결과에 따르면 특히 소득/소비/근로 등의 영역에서 작용할 수 있는 기준변수를 찾아보는 노력이 필요하다.

기준변수에 관한 이러한 연구는 본 연구와 동일한 연구설계를 이용하게 될 것이다. 다만 다음과 같은 점이 고려되어야 할 것이다.

- 본 연구에서 선정된 기준변수들을 다시 포함해서 그 일반성을 검토할 필요가 있다.
- 응답자 모집과정에서 보다 다양한 응답자들이 참여할 수 있도록 모집방법이 개선될 필요가 있다.
- 소득/소비 분야와 관련성이 높은 기준변수를 찾아서 포함할 필요가 있다. 이번 연구에서는 이 분야에서 적합한 기준변수를 찾지 못하였기 때문에 후속 연구에서는 본 연구에 포함하지 않았던 다른 변수들을 검토할 필요가 있다.

나. 인터넷 응답패널의 특성 연구

또 하나 중요한 연구분야는 응답패널의 특성에 관한 연구다. 확률표집 방법을 인터넷 조사에 적용할 경우 부딪히는 문제점이 응답자 패널의 모집과정에서 응답률이 낮다는 점이다. 따라서 패널에 참가하는 응답자와 그렇지 않은 응답자간에 어떤 체계적 차이가 있는지 그리고 이러한 차이를 극복하기 위한 방안은 무엇인지 찾아내는 연구가 필요하다. 이러한 연구를 위한 연구 설계는 다음과 같이 제안될 수 있다.

1) 통계청에서 실시하는 각종 조사에서 웹응답에 참여하는 경우와 그렇지 않은 경우를 비교한다.

2) 두 응답자 집단을 비교하는 변수에 앞서 가중모형 개발에서 밝혀진 변수들을 포함시키도록 한다.

사회적 변화로 인해 인터넷 등 새로운 조사모드의 채택이 점차 불가피하게 된다는 점을 고려하면 이러한 인터넷 조사에 대한 연구는 시급히 필요하며, 특히 위에서 제시한 연구들은 인터넷 조사의 실용화를 보다 촉진시킬 수 있다는 점에서 그 시급성이 높다고 본다.

참고문헌

- Cooper, Mick P. (2008). The inferential Challenges of Web Surveys. Paper presented at *2008 International Conference on Survey Research Methodology* (September 11-12, 2008, Academia Sinica).
- D'agostino, R. B. Jr. (1998). Tutorial in biostatistics: Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in Medicine*, 17. 2265-2281.
- Deville, J.-C., and Sarndal, C.-E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, 87. 376-382.
- Duffy, B. Smith, K., Therhanian, G., and Bremer, J. (2005). Comparing Data from online and face-to-face surveys. *International Journal of Market Reserch*. 47(6). 615-639.
- Ekman, A. (2006). *The Use of the World Wide Web in Epidemiological Research*. Karolinska Institutet.
- Hahs-Vaughn, D. L. and Onwuegbuzie, A.J. (2006). Estimating and using propensity score analysis with complex samples. *The Journal of Experimental Education*. 75(1). 31-65.
- Heeren, T., Edwards, E.M., Dennis, J. M., Rodkin, S., Hingson, R. W. and Rosenboom, D.L. (2008). A Comparison of Results From an Alcohol Survey of a Rprerecruited Internet Panel and the National Epidemiologic Survey on Alcohol and Related Conditions. *Alcoholism: Clinical and Experimental Research*, 32-2. 222-229.
- Hill, S. J. The opt-in internet panel. (2007). Annual meeting of the Midwest Political Science Association. Chicago. Apr 12.
- Krotki, K. and Trofimovich, L. (2006). *Supplementing RDD Surveys with Web Data*. www.rti.org.

- Lee, S., (2006). An Evaluation of Nonresponse and Coverage Errors in a Prerecruited Probability Web Panel Survey. *Social Science Computer Review* 24. 460.
- Lee, S., and Valliant, R. (2009). Estimation for volunteer panel Web surveys using propensity score adjustment and calibration adjustment. *Sociological Methods and Research*, 37. 319-343.
- Loosveldt G. and Sonck, Nathalie. (2008). An evaluation of the weighting procedures for an online access panel survey. *Survey Research Methods*. 2(2). 93-105.
- Lukacs, P. (2007). Development and Analysis of Joint Internet-telephone Hunter Surveys. *Human Dimensions of Wildlife*, 12:263-273.
- Rosenbaum, P., and Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70. 41-55.
- Rosenbaum, P., and Rubin, D.B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association*, 79. 516-524.
- Schonlau, M., Soest, A., Kapteyn A. and Couper, M. (2009). Selection Bias in Web Surveys and the Use of Propensity Scores. *Sociological Methods & Research*. 2009;37;291
- Shonlau, M. Zapert, K. Simon, L.P. Sanstad, K.H. Marcus, S. M. Adams, J. Kan, H. Turner, R. Berry, S. (2004). A Comparison Between Responses from a Propensity-Weighted Web Survey and an Identical RDD Survey. *Social Science Computer Review*. 22-1. 128-138.
- Taylor, P.A., Nelson, N.M., Grandjean, B.D., Anatchkova, B., and Aadland, D. (2009). *Mode Effects and other potential biases in panel-based internet surveys: final report*. Wyoming Survey & Analysis Center, Univ. of Wyoming.

Zhang, C., Callegaro, M., Thomas, M., and DiSogra, C. (2009). Do We Hear Different Voices? *AAPOR Conference*.

<부록 1> PS rim 성향가중법 결과

일러두기:

- FTF.raw = FTF(면대면) 표본 원자료
- FTF.cell = FTF(면대면) 표본 인구·지리적 cell 가중치 적용
- Web.raw = Web 표본 원자료
- Web.rim = Web 표본 인구·지리적 rim 가중치 적용
- Web.ps9 = Web 표본 성향점수모형 가중치 적용 (9개 조정 변수 사용)
- Web.ps7 = Web 표본 성향점수모형 가중치 적용 (7개 조정 변수 사용)

1. 가구관련사항

< 거처의 종류 >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
단독주택	39.4	35.8	19.4	21.2	22.3	22.3
아파트	45.9	46.6	58.8	57.2	55.5	55.4
연립주택	5.7	6.5	7.3	7.2	7.2	7.2
다세대주택	7.4	9.4	11.4	11.2	11.8	11.9
기타	1.6	1.8	3.2	3.2	3.3	3.2
계	100.0	100.0	100.0	100.0	100.0	100.0

< 점유형태 >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
자기집	62.1	58.7	62.7	64.5	65.3	65.1
전세	19.1	22.9	22.8	21.8	20.9	21.0
보증부	9.4	9.6	6.1	6.1	6.0	6.0
월세(사글세)	6.1	5.8	5.5	4.9	4.9	5.0
무상	3.3	3.0	3.0	2.7	2.8	2.8
계	100.0	100.0	100.0	100.0	100.0	100.0

2. 개인관련사항

< 성별 >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
남	47.7	48.9	48.0	48.9	48.7	48.8
여	52.3	51.1	52.0	51.1	51.3	51.2
계	100.0	100.0	100.0	100.0	100.0	100.0

< 생년월일 >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
양력	35.3	39.9	63.3	60.2	58.9	58.8
음력	64.7	60.1	36.7	39.8	41.1	41.2
계	100.0	100.0	100.0	100.0	100.0	100.0

< 연령대 >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
~20대	21.7	24.4	33.3	24.4	24.5	24.5
30대	18.8	20.4	21.5	20.4	20.3	20.3
40대	22.2	20.9	21.8	20.9	21.0	21.0
50대	16.2	15.9	15.7	15.9	16.0	16.0
60대~	21.1	18.3	7.7	18.3	18.2	18.2
계	100.0	100.0	100.0	100.0	100.0	100.0

< 가구주와의 관계 >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
가구주	46.2	45.2	35.1	37.7	34.8	34.8
배우자	29.9	28.8	27.1	27.0	26.9	27.0
미혼자녀	17.9	19.9	25.8	19.2	19.6	19.6
기혼자녀 및 그 배우자	1.0	1.0	1.6	1.5	1.7	1.7
손자녀 및 그 배우자	0.3	0.3	0.3	0.3	0.3	0.3
부모(배우자쪽 포함)	3.2	3.0	5.9	10.6	12.3	12.2
조부모(배우자쪽 포함)	0.1	0.0	0.4	0.7	0.9	0.9
미혼 형제·자매	0.8	0.9	2.3	1.8	2.0	2.0
기타 친·인척	0.4	0.4	0.4	0.4	0.5	0.5
기타 동거인	0.3	0.4	1.1	0.9	1.0	1.0
계	100.0	100.0	100.0	100.0	100.0	100.0

< 교육정도 >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
초/중	27.6	24.0	10.9	16.1	24.4	23.7
고등	36.5	36.1	30.0	30.4	34.5	36.3
대학	32.7	36.0	53.3	48.0	37.0	36.2
대학원	3.2	3.8	5.8	5.5	4.1	3.8
계	100.0	100.0	100.0	100.0	100.0	100.0

< 혼인상태 >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
미혼	23.8	26.6	33.8	25.9	25.6	25.7
배우자 있음	64.4	63.0	60.7	65.2	63.8	63.9
사별	8.4	7.3	3.6	6.9	8.5	8.4
이혼	3.3	3.1	1.9	2.0	2.1	2.0
계	100.0	100.0	100.0	100.0	100.0	100.0

3. 사회참여부문

< 주관적 만족감 >

(평균: 5점 척도, 기타 단위 %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
평균	3.186	3.170	3.182	3.201	3.265	3.273
매우 만족	3.2	3.3	3.8	3.4	2.3	2.2
약간 만족	17.0	17.6	20.1	19.8	16.3	16.0
보통	45.4	45.3	39.5	39.7	43.6	43.5
약간 불만족	26.9	26.5	27.5	27.7	28.4	28.7
매우 불만족	7.5	7.4	9.2	9.5	9.5	9.5
계	100.0	100.0	100.0	100.0	100.0	100.0

< 사회적 관계망 >

(평균: 명, 기타 단위: %)

		FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
아플 경우	있다	75.1	76.6	80.7	80.7	79.3	79.3
	평균	2.206	2.186	2.637	2.576	2.586	2.582
	없다	24.9	23.4	19.3	19.3	20.7	20.7
	계	100.0	100.0	100.0	100.0	100	100.0
돈빌릴 경우	있다	47.2	48.6	69.3	67.7	65.8	65.7
	평균	2.236	2.223	2.610	2.558	2.503	2.495
	없다	52.8	51.4	30.7	32.3	34.2	34.3
	계	100.0	100.0	100.0	100.0	100	100.0
낙심할 경우	있다	80.3	81.3	85.6	83.9	82.4	82.4
	평균	2.855	2.878	3.301	3.180	3.039	3.033
	없다	19.7	18.7	14.4	16.1	17.6	17.6
	계	100.0	100.0	100.0	100.0	100.0	100.0

< 후원금 (기부금) >

(평균: 횟수, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
있다	33.1	32.3	35.1	34.2	28.0	27.9
직접후원(평균)	3.959	4.231	2.203	2.211	2.111	2.094
언론기관(평균)	2.861	2.961	1.525	1.598	1.612	1.614
사회복지단체(평균)	3.478	3.823	3.439	3.377	3.209	3.178
종교단체(평균)	4.587	5.032	3.322	3.416	3.615	3.588
기업직장(평균)	3.373	3.769	2.232	2.186	2.146	2.116
기타(평균)	1.745	2.932	1.535	1.344	1.273	1.280
없다	66.9	67.7	64.9	65.8	72.0	72.1
계	100.0	100.0	100.0	100.0	100.0	100.0

< 단체참여 >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
있다	40.7	39.8	41.9	40.9	35.4	35.3
없다	59.3	60.2	58.1	59.1	64.6	64.7
계	100.0	100.0	100.0	100.0	100.0	100.0

< 단체참여: 있는 경우(복수응답) >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
친목 및 사교단체	78.9	77.8	42	40.6	39.6	39.4
종교단체	28.5	28.9	45.1	45.1	46.4	46.4
취미/스포츠/레저단체	29.4	29.9	89.4	91.7	95.1	95.2
시민사회단체	13.7	12.9	53.6	55.7	57.9	58.3
학술단체	4.6	5.1	25.6	23.6	17.4	17.4
이익단체	3.2	3.4	19.4	17.8	14.9	14.6
정치단체	0.9	0.9	2.3	2.5	5.0	5.1
기 타	0.6	0.5	22.5	22.9	23.6	23.7

< 자원봉사활동 >

(평균: 횟수·시간, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7	
있다	20.0	19.3	23.8	21.4	18.5	18.5	
환경평균	횟수	4.554	4.607	2.145	2.328	2.003	1.941
	시간	3.013	3.230	4.014	4.071	3.920	3.850
국가평균	횟수	2.291	2.568	2.053	2.227	1.891	1.890
	시간	3.477	3.977	4.206	4.306	4.731	4.721
자녀교육 평균	횟수	4.309	4.870	2.997	3.298	3.131	3.130
	시간	2.041	2.292	2.996	2.951	3.212	3.242
아동평균	횟수	6.102	6.143	6.049	6.072	5.897	5.798
	시간	4.264	4.404	6.871	6.858	7.415	7.393
재해평균	횟수	1.929	2.099	0.677	0.693	0.693	0.678
	시간	4.128	4.718	2.349	2.507	2.491	2.454
기타평균	횟수	5.689	6.083	2.362	2.506	2.048	2.015
	시간	3.087	3.393	2.527	2.712	2.666	2.609
없다	80.0	80.7	76.2	78.6	81.5	81.5	
계	100.0	100.0	100.0	100.0	100.0	100.0	

< 앞으로의 자원봉사활동 의사 >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
기회가 주어진다면 (계속)하겠다	24.8	24.6	35.8	33.5	28.8	28.8
하고 싶지만 지금은 어렵다	50.9	52.2	51.1	50.1	50.9	51.0
할 생각이 없다	24.3	23.1	13.2	16.4	20.3	20.3
계	100.0	100.0	100.0	100.0	100.0	100.0

< 민원서비스 만족도 >

(평균: 5점 척도, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
있다	55.7	55.9	53.7	52.6	45.6	45.5
매우 만족	12.6	11.3	17.4	16.5	14.4	14.4
약간 만족	30.4	29.9	34.9	35.3	33.6	33.4
보통	41.7	42.7	35.8	35.9	38.8	38.9
약간 불만족	12.2	12.8	9.2	9.5	10.1	10.2
매우 불만족	3.1	3.3	2.7	2.8	3.2	3.1
평균	2.628	2.669	2.450	2.469	2.541	2.542
없다	44.3	44.1	46.3	47.4	54.4	54.5
계	100.0	100.0	100.0	100.0	100.0	100.0

< 민원서비스 불만이유(복수응답) >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
불친절하다	58.8	58.3	46.6	41.3	45.2	45.7
시간이 오래 걸린다	52.6	52.7	64.9	59.9	58.5	59.3
절차가 복잡하고 까다롭다	56.3	55	62.8	80.2	80.7	80.6
공정하지 못하다	15.1	15.1	14.2	13.7	10.3	9.9
업무처리가 미숙하다	24.5	24.8	60.8	56.2	61.7	60.7
편의시설 등이 미비하다	10.5	10.7	26.4	25.9	27.6	28
대가를 바란다	1.3	1.2	10.1	8.6	4.7	4.7
기타	0.9	0.9	14.2	14.3	11.4	11.1

< 계층의식 >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
상 상	0.4	0.5	0.6	0.6	0.5	0.5
상 하	1.6	1.7	2.2	2.3	1.9	1.9
중 상	21.3	22.1	28.1	26.9	25.4	25.1
중 하	36.1	36.7	42.9	43.0	42.6	42.6
하 상	24.6	24.1	18.6	19.3	20.4	20.6
하 하	16.0	14.9	7.6	7.9	9.2	9.3
계	100.0	100.0	100.0	100.0	100.0	100.0

< 계층이동 가능성 >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
매우 높다	2.9	3.1	4.6	4.1	4.2	4.1
비교적 높다	34.3	34.5	31.3	30.4	28.8	28.6
비교적 낮다	34.0	34.2	43.5	42.9	41.7	41.9
매우 낮다	11.2	11.4	14.0	15.3	15.7	15.8
모르겠다	17.6	16.8	6.7	7.3	9.6	9.6
계	100.0	100.0	100.0	100.0	100.0	100.0

< 계층이동 가능성: 지식세대 >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
매우 높다	4.9	5.1	5.9	5.5	5.2	5.2
비교적 높다	42.9	43.2	41.2	41.0	39.8	39.7
비교적 낮다	23.8	23.8	33.7	33.5	32.9	33.0
매우 낮다	5.9	6.0	8.8	9.7	9.9	10.0
모르겠다	22.5	21.9	10.4	10.2	12.1	12.1
계	100.0	100.0	100.0	100.0	100.0	100.0

4. 문화와 여가 부문

< 일반신문 >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
거의 매일	22.5	23.8	31.9	33.3	26.4	26.3
1주일에 3~4회	10.0	10.8	12.1	11.5	9.1	9.1
1주일에 1~2회	12.1	12.9	14.1	13.3	11.5	11.5
2주일에 1회	5.5	5.7	8.0	7.2	6.4	6.4
없다	49.8	46.9	33.8	34.8	46.6	46.7
계	100.0	100.0	100.0	100.0	100.0	100.0

< 인터넷 신문 >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
거의 매일	21.7	24.2	41.9	38.0	24.3	24.3
1주일에 3~4회	11.7	12.6	19.8	17.5	12.5	12.6
1주일에 1~2회	10.5	11.0	10.7	10.3	10.7	11.0
2주일에 1회	4.4	4.5	4.3	4.2	4.5	4.5
없다	51.7	47.8	23.3	29.9	48.0	47.6
계	100.0	100.0	100.0	100.0	100.0	100.0

< TV 시청 >

(평균: 시간, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
있다	96.6	96.4	94.1	94.0	93.0	93.0
평 일(평균)	2.891	2.841	2.528	2.674	2.760	2.763
토요일(평균)	3.608	3.600	3.518	3.611	3.663	3.664
일요일.공휴일(평균)	3.925	3.922	3.848	3.917	3.988	3.990
없다	3.4	3.6	5.9	6.0	7.0	7.0
계	100.0	100.0	100.0	100.0	100.0	100.0

< 독서 >

(평균: 권, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
있다	59.3	62.1	79.3	74.8	66.9	67.0
잡지류(평균)	5.925	6.151	4.111	4.156	3.741	3.732
교양서적(평균)	7.959	8.255	6.749	6.718	6.019	5.983
직업관련(평균)	6.751	7.127	4.267	4.131	3.796	3.740
생활(평균)	4.112	4.336	2.562	2.630	2.450	2.445
기타(평균)	16.634	17.830	10.278	9.449	9.182	9.199
없다	40.7	37.9	20.7	25.2	33.1	33.0
계	100.0	100.0	100.0	100.0	100.0	100.0

< 레저시설 이용횟수 >

(평균: 횟수, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
있다	67.3	69.0	75.5	74.0	67.6	67.5
관광명소(평균)	2.839	2.761	2.559	2.530	2.354	2.345
온천장 및 스파(평균)	2.619	2.586	1.755	1.775	1.754	1.750
골프장(평균)	4.380	5.183	0.654	0.680	0.542	0.538
스키장(평균)	1.421	1.604	0.808	0.753	0.690	0.686
해수욕장(평균)	1.701	1.645	1.535	1.468	1.396	1.394
산림욕장(평균)	2.064	2.112	1.473	1.458	1.327	1.326
놀이공원(평균)	1.840	1.935	1.590	1.583	1.551	1.556
수영장위터파크(평균)	2.506	2.431	1.576	1.439	1.379	1.367
기타(평균)	1.622	1.881	1.759	1.783	1.575	1.568
없다	32.7	31.0	24.5	26.0	32.4	32.5
계	100.0	100.0	100.0	100.0	100.0	100.0

< 문화예술 및 스포츠관람 >

(평균: 번, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
있다	51.8	55.1	70.5	65.3	58.1	58.0
음악.연주회(평균)	1.951	2.040	1.248	1.256	1.103	1.093
연극.마당극(평균)	1.756	1.936	1.508	1.489	1.317	1.309
무용(평균)	0.857	1.135	0.302	0.299	0.253	0.250
영화(평균)	5.505	5.678	6.634	6.335	5.784	5.762
박물관(평균)	1.923	1.945	1.449	1.460	1.327	1.319
미술관(평균)	1.914	2.039	1.172	1.192	1.013	1.003
스포츠(평균)	2.915	3.039	1.663	1.613	1.432	1.431
없다	48.2	44.9	29.5	34.7	41.9	42.0
계	100.0	100.0	100.0	100.0	100.0	100.0

< 주말이나 휴일의 여가활동(복수응답) >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
TV.비디오시청	60	59.6	50	49.2	49.3	49.3
여행	11.9	12.4	16.1	15.7	13.0	13.1
문화예술 관람	7.4	8.3	18.1	16.6	14.0	14.1
스포츠 관람	2.9	3.1	4.7	4.8	4.7	4.7
스포츠 활동	10	10.3	8.5	8.1	6.7	6.7
컴퓨터.인터넷 게임	15.5	16.2	25.1	22.8	24.5	24.7
창작적 취미	3.1	3.1	8.6	8.2	6.8	6.8
자기계발	5.6	6.1	14.7	13.7	11.6	11.5
봉사 활동	1.7	1.6	2.9	3	2.8	2.7
종교 활동	14	14.2	9.7	11.6	11.9	11.8
가사일	32.5	31.1	32.2	34.6	36.9	36.8
휴식	42.6	42	71	72.5	75.7	75.8
사교관련일	22.9	23.5	36.7	37	39.1	39
기타	1.4	1.1	1.7	2.2	3.0	2.9

< 여가활동을 함께 하고 싶은 사람 >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
가족	59.0	58.6	62.5	64.9	63.9	64.0
친구	17.7	19.3	22.9	19.3	19.0	18.9
동호회	5.6	5.5	5.3	6.1	6.1	6.1
혼자서	17.4	16.4	8.7	9.1	10.0	10.1
기타	0.2	0.2	0.6	0.6	1.0	0.9
계	100.0	100.0	100.0	100.0	100.0	100.0

< 앞으로 하고 싶은 여가활동(복수응답) >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
TV.비디오시청	11.7	11.1	19.1	22.2	25.3	25.3
여행	60.4	61.2	25.4	25.4	23.2	23.2
문화예술 관람	21.1	22.2	37.5	36	33.7	33.8
스포츠 관람	6.8	7.1	11.7	12.2	12.5	12.5
스포츠 활동	19.7	20.3	20.7	19.6	18.2	18.2
컴퓨터.인터넷 게임	5.8	5.9	9.4	7.9	9.0	9.1
창작적 취미	14.3	14.5	31.4	29.7	26.3	26.3
자기계발	24.4	25.7	37.7	35.6	35.1	35.2
봉사 활동	10.5	10.2	15.6	15.8	13.1	13
종교 활동	9	8.6	6.6	8.1	9.5	9.5
가사일	4.9	4.3	6.8	7.8	10.5	10.4
휴식	19.2	18.4	37	38.4	42.5	42.2
사교관련일	22.3	21.6	40.8	40.8	41.1	41.1
기타	0.6	0.5	0.4	0.5	0.2	0.2

< 여가활동 만족도 >

(평균: 5점 척도, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
평균	3.134	3.128	2.970	2.988	3.065	3.073
매우 만족	4.8	4.9	6.2	5.8	4.9	4.9
약간 만족	16.5	16.9	25.4	24.9	21.0	20.8
보통	46.5	46.1	40.0	40.5	43.5	43.6
약간 불만족	24.6	24.7	22.0	22.2	23.5	23.7
매우 불만족	7.5	7.4	6.4	6.6	6.9	7.1
계	100.0	100.0	100.0	100.0	100.0	100.0

< 여가활동 불만족 이유 >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
경제적 부담	54.2	54.3	55.4	55.3	52.6	52.7
시간이 부족	27.5	28.4	26.4	24.2	26.2	26.3
교통 혼잡 및 이용교통수단이 불편	1.7	1.9	1.1	1.3	1.4	1.4
여가시설이 부족	1.7	1.8	1.8	1.7	1.3	1.2
여가정보 및 프로그램이 부족	0.8	0.8	1.1	1.0	1.0	1.0
적당한 취미가 없어서	2.5	2.4	5.0	5.4	5.1	5.1
체력이나 건강이 좋지 않아서	8.7	7.7	4.0	6.5	7.5	7.4
여가를 함께 즐길 사람이 없어서	1.9	1.8	3.0	2.7	2.6	2.5
기타	1.0	0.9	2.2	1.9	2.4	2.3
계	100.0	100.0	100.0	100.0	100.0	100.0

< 국내 관광여행 횟수 >

(평균: 번·일·번, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
있다	63.6	64.8	68.0	66.9	60.6	60.5
숙박여행 평균	2.411	2.476	2.434	2.418	2.275	2.268
숙박여행 1회당 평균	2.335	2.355	2.040	2.016	2.047	2.045
당일여행 평균	4.574	4.599	4.752	4.664	4.412	4.408
없다	36.4	35.2	32.0	33.1	39.4	39.5
계	100.0	100.0	100.0	100.0	100.0	100.0

< 해외여행 경험 및 횟수(지난 1년 동안) >

(평균: 번, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
있다	12.2	13.6	15.5	15.3	13.3	13.1
관광 평균	1.331	1.352	1.067	1.058	1.006	1.012
가사 평균	1.544	1.501	0.659	0.701	0.669	0.680
업무 평균	2.114	2.293	0.779	0.768	0.700	0.692
어학연수 평균	0.625	0.680	0.302	0.256	0.253	0.250
없다	87.8	86.4	84.5	84.7	86.7	86.9
계	100.0	100.0	100.0	100.0	100.0	100.0

< 해외여행 경험 (1년 동안 없는 사람 중) >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
있다	31.4	33.9	44.2	44.2	39.7	39.5
없다	68.6	66.1	55.8	55.8	60.3	60.5
계	100.0	100.0	100.0	100.0	100.0	100.0

< 전통문화가 생활에 미치는 영향 >

(평균: 5점 척도, 기타 단위: %)

		FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
생활의례	평균	2.972	2.982	2.848	2.767	2.839	2.840
	매우 많은 영향	7.3	7.1	14.0	15.6	15.2	15.2
	많은 영향	23.9	23.7	24.2	25.9	23.1	23.0
	보통	41.6	41.8	35.0	34.0	35.4	35.4
	거의 영향 미치지 않음	18.8	18.8	16.8	15.0	15.3	15.3
	전혀 영향 미치지 않음	8.4	8.6	10.0	9.4	11.0	11.0
인간관계 및 예절	평균	2.491	2.503	2.248	2.237	2.361	2.361
	매우 많은 영향	14.3	14.0	26.8	27.1	24.4	24.4
	많은 영향	38.8	38.3	37.5	38.2	35.4	35.4
	보통	34.3	35.0	24.6	23.8	26.8	26.7
	거의 영향 미치지 않음	8.6	8.7	6.0	5.4	6.6	6.6
	전혀 영향 미치지 않음	4.0	4.0	5.0	5.4	6.8	6.9
민속놀이	평균	3.678	3.703	3.726	3.686	3.706	3.708
	매우 많은 영향	1.3	1.2	2.7	3.0	2.8	2.8
	많은 영향	6.1	5.6	7.0	7.4	7.2	7.2
	보통	35.0	34.5	30.9	32.0	32.6	32.6
	거의 영향 미치지 않음	38.9	39.2	34.0	33.0	31.2	31.3
	전혀 영향 미치지 않음	18.7	19.5	25.5	24.5	26.1	26.2
전통예술	평균	3.763	3.780	3.771	3.732	3.765	3.766
	매우 많은 영향	1.5	1.4	2.8	3.2	2.9	2.9
	많은 영향	6.2	6.0	7.2	7.7	7.3	7.3
	보통	30.2	30.0	28.2	29.2	29.7	29.7
	거의 영향 미치지 않음	38.7	38.6	33.8	32.5	30.4	30.4
	전혀 영향 미치지 않음	23.4	24.1	28.0	27.4	29.6	29.7
전통음식	평균	2.961	2.977	2.941	2.911	3.019	3.019
	매우 많은 영향	7.5	7.0	10.2	11.0	10.0	10.0
	많은 영향	25.6	25.4	25.5	26.0	23.1	23.1
	보통	39.6	39.9	35.8	35.4	36.0	36.0
	거의 영향 미치지 않음	17.9	18.3	16.7	16.3	17.0	17.0
	전혀 영향 미치지 않음	9.4	9.4	11.7	11.4	13.9	13.9

(평균: 5점 척도, 기타 단위: %)

		FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
전통복장	평균	3.639	3.654	3.547	3.502	3.523	3.525
	매우 많은 영향	1.6	1.5	4.1	4.6	4.3	4.3
	많은 영향	8.1	7.8	11.5	12.1	11.9	11.9
	보통	34.5	34.5	31.9	32.9	33.5	33.4
	거의 영향 미치지 않음	36.2	36.1	30.7	29.3	27.9	27.9
	전혀 영향 미치지 않음	19.5	20.1	21.8	21.1	22.5	22.5
전통건축	평균	3.572	3.589	3.456	3.410	3.449	3.449
	매우 많은 영향	2.7	2.6	5.1	5.7	5.6	5.6
	많은 영향	10.6	10.3	13.8	14.5	13.2	13.3
	보통	33.7	33.4	32.1	32.8	33.6	33.6
	거의 영향 미치지 않음	33.1	33.0	28.1	27.0	25.8	25.8
	전혀 영향 미치지 않음	20.0	20.7	20.8	19.9	21.8	21.8
전통무예	평균	4.118	4.131	4.058	4.038	4.048	4.049
	매우 많은 영향	0.9	0.8	2.8	2.8	2.5	2.5
	많은 영향	3.4	3.3	4.7	4.7	4.7	4.6
	보통	20.8	20.5	20.6	21.4	22.3	22.3
	거의 영향 미치지 않음	33.0	32.6	27.9	28.0	26.5	26.6
	전혀 영향 미치지 않음	42.0	42.7	44.1	43.1	44.0	44.0

5. 소득과 소비 부문

< 소득만족도 >

(평균: 5점 척도, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
있다	66.1	65.8	64.9	64.8	60.9	60.8
매우 만족	1.6	1.5	3.0	3.4	2.8	2.7
약간 만족	12.3	12.5	15.8	15.7	14.2	14.0
보통	39.2	39.3	33.0	32.8	34.0	33.8
약간 불만족	33.1	33.1	30.6	29.8	31.0	31.3
매우 불만족	13.9	13.4	17.6	18.2	18.0	18.1
소득만족도 평균	3.453	3.444	3.439	3.437	3.474	3.481
없다	33.9	34.2	35.1	35.2	39.1	39.2
계	100.0	100.0	100.0	100.0	100.0	100.0

< 장래 소득에 대한 기대 >

(평균: 5점 척도, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
평균	2.935	2.875	2.490	2.591	2.626	2.631
크게 증가	4.8	5.6	11.9	10.5	9.3	9.1
약간 증가	31.6	34.1	46.9	44.4	42.8	42.7
동일	35.8	34.2	26.2	27.0	30.0	30.0
약간 감소	21.0	19.8	9.9	11.4	11.9	12.0
크게 감소	6.8	6.5	5.0	6.6	6.0	6.1
계	100.0	100.0	100.0	100.0	100.0	100.0

< 소득분배에 대한 견해 >

(평균: 5점 척도, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
평균	3.926	3.923	4.037	4.032	3.913	3.925
매우 공평	0.6	0.6	0.9	0.8	0.6	0.6
약간 공평	3.5	3.5	3.9	3.9	3.6	3.5
보통	25.5	25.7	19.7	20.3	26.5	25.6
약간 불공평	43.5	43.3	41.5	41.2	42.6	43.4
매우 불공평	26.9	26.8	34.0	33.8	26.8	26.9
계	100.0	100.0	100.0	100.0	100.0	100.0

< 소비생활 만족도 >

(평균: 5점 척도, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
평균	3.350	3.344	3.270	3.284	3.355	3.365
매우 만족	1.7	1.8	2.8	2.5	1.7	1.7
약간 만족	11.7	11.9	17.9	17.8	12.4	12.1
보통	46.2	46.1	39.1	39.1	44.9	44.8
약간 불만족	30.7	30.5	30.1	29.9	30.6	31.0
매우 불만족	9.7	9.7	10.2	10.6	10.4	10.5
계	100.0	100.0	100.0	100.0	100.0	100.0

6. 삶의 질 만족도 부문

< 주관적 건강평가 >

(평균: 5점 척도, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
평균	2.831	2.789	2.602	2.697	2.761	2.764
매우 좋다	6.4	6.4	9.8	8.3	7.4	7.3
좋은 편이다	30.0	31.6	36.4	33.9	31.4	31.3
보통이다	41.6	41.8	39.3	39.9	41.8	41.9
나쁜 편이다	18.3	17.0	12.8	15.5	16.5	16.6
매우 나쁘다	3.7	3.2	1.7	2.4	2.9	2.9
계	100.0	100.0	100.0	100.0	100.0	100.0

< 스트레스 인식정도 >

(평균: 4점 척도, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
평균	2.541	2.518	2.306	2.337	2.335	2.332
매우 많이 느꼈다	12.3	12.5	17.5	16.9	16.5	16.5
많이 느낀 편이다	30.4	31.3	39.0	37.4	38.1	38.2
조금 느낀 편이다	48.3	47.9	38.9	40.8	40.9	40.7
거의 느끼지 않았다	9.0	8.2	4.6	4.9	4.6	4.5
계	100.0	100.0	100.0	100.0	100.0	100.0

< 거주지역 만족도 >

(평균: 5점 척도, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
평균	2.726	2.752	2.666	2.680	2.707	2.712
매우 만족	9.0	8.3	12.6	12.1	10.6	10.5
약간 만족	29.8	29.6	31.9	31.3	30.4	30.3
보통	44.6	44.5	36.2	37.4	40.7	40.8
약간 불만족	13.1	13.7	15.0	14.9	14.0	14.2
매우 불만족	3.6	3.9	4.3	4.3	4.2	4.2
계	100.0	100.0	100.0	100.0	100.0	100.0

< 주택 만족도 >

(평균: 5점 척도, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
평균	2.847	2.871	2.856	2.857	2.896	2.902
매우 만족	8.0	7.2	10.5	10.4	9.2	9.2
약간 만족	27.5	27.7	29.7	29.2	27.7	27.4
보통	41.9	42.1	32.0	32.9	35.5	35.6
약간 불만족	16.8	17.0	19.5	19.3	19.5	19.7
매우 불만족	5.7	6.0	8.4	8.2	8.1	8.1
계	100.0	100.0	100.0	100.0	100.0	100.0

< 사회 안전에 대한 인식 >

(평균: 5점 척도, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
평균	3.349	3.349	3.147	3.140	3.168	3.173
매우 안전하다	0.9	0.9	2.4	2.0	1.6	1.6
비교적 안전하다	14.6	14.9	25.4	25.8	23.5	23.3
보통이다	39.9	39.8	34.7	35.6	38.5	38.5
비교적 불안하다	37.6	37.1	30.2	29.5	29.2	29.4
매우 불안하다	6.9	7.3	7.3	7.2	7.1	7.2
계	100.0	100.0	100.0	100.0	100.0	100.0

< 범죄피해 두려움 >

(평균: 5점 척도, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
평균	2.732	2.715	2.635	2.658	2.687	2.684
매우 많이 느낀다	9.0	9.0	9.9	9.3	9.1	9.2
약간 느낀다	35.1	35.5	37.7	37.2	34.8	34.9
보통이다	32.6	33.0	33.3	34.0	36.8	36.7
별로 느끼지 못한다	20.4	19.9	17.1	17.3	16.9	16.9
전혀 느끼지 못한다	2.9	2.6	2.0	2.2	2.4	2.3
계	100.0	100.0	100.0	100.0	100.0	100.0

< 가족관계 만족도 >

(평균: 5점 척도, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
평균	2.366	2.364	2.408	2.447	2.517	2.521
매우 만족	20.8	20.7	18.3	17.1	14.6	14.4
약간 만족	32.4	32.7	37.1	36.4	34.5	34.5
보통	37.6	37.5	32.5	34.0	38.1	38.2
약간 불만족	7.7	7.6	9.7	10.0	10.2	10.3
매우 불만족	1.5	1.5	2.3	2.6	2.6	2.6
계	100.0	100.0	100.0	100.0	100.0	100.0

< 환경오염에 대한 체감 >

(평균: 5점 척도, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
평균	3.121	3.125	3.190	3.197	3.178	3.181
매우 좋아졌다	1.3	1.2	2.7	2.7	2.5	2.5
약간 좋아졌다	13.7	13.7	15.9	15.8	15.5	15.4
변화 없다	60.1	60.4	47.4	47.3	49.6	49.5
약간 나빠졌다	21.2	20.9	27.6	27.5	26.7	26.8
매우 나빠졌다	3.6	3.8	6.4	6.8	5.8	5.8
계	100.0	100.0	100.0	100.0	100.0	100.0

< 자살 충동 경험 여부 >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
있다	14.8	14.7	22.3	20.5	18.7	18.9
없다	85.2	85.3	77.7	79.5	81.3	81.1
계	100.0	100.0	100.0	100.0	100.0	100.0

< 귀가 시각 >

(평균: 시간.분, 기타 단위: %)

		Web.raw	Web.rim	Web.ps9	Web.ps7
오전	평균	8.4	8.9	9.9	10.0
		6.528	6.554	6.556	6.561
오후	평균	91.6	91.1	90.1	90.0
		8.392	8.116	8.049	8.053
계		100.0	100.0	100.0	100.0

< 취침 시각 >

(평균: 시간.분, 기타 단위: %)

		Web.raw	Web.rim	Web.ps9	Web.ps7
오전	평균	29.3	25.6	24.3	24.4
		3.086	3.144	3.103	3.119
오후	평균	70.7	74.4	75.7	75.6
		11.088	10.983	10.845	10.842
계		100.0	100.0	100.0	100.0

< 인터넷 사용 여부 >

(단위: %)

	Web.raw	Web.rim	Web.ps9	Web.ps7
사용하였다	80.5	73.8	63.6	63.7
사용하지 않았다	19.5	26.2	36.4	36.3
계	100.0	100.0	100.0	100.0

< 인터넷 사용 용도: 평균 >

(평균: 시간)

	Web.raw	Web.rim	Web.ps9	Web.ps7
정보검색 평균	1.566	1.429	1.225	1.222
교제 평균	0.685	0.604	0.512	0.511
게임 평균	0.646	0.582	0.581	0.582
관리 평균	0.452	0.403	0.346	0.346
전체 평균	2.831	2.528	2.165	2.162

< 생활시간에 대한 태도 >

(평균: 5점 척도, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
평 균	2.653	2.611	2.493	2.582	2.617	2.616
항상 그렇다	14.7	15.1	15.5	14.5	14.2	14.1
그런 편이다	33.0	34.1	38.0	35.5	34.2	34.3
보통이다	30.1	30.1	30.8	30.9	31.8	31.7
그렇지 않은 편이다	16.6	15.8	13.3	15.5	15.7	15.7
전혀 그렇지 않다	5.6	4.8	2.5	3.6	4.2	4.2
계	100.0	100.0	100.0	100.0	100.0	100.0

< 생활시간 부족 이유 >

(단위: %)

	Web.raw	Web.rim	Web.ps9	Web.ps7
일(학습) 때문에	55.3	53.9	53.9	53.7
자녀양육 및 집안 일 때문에	18.4	19.7	20.3	20.4
시간관리를 잘 못해서	14.1	13.4	13.6	13.7
교제 및 사회활동 때문에	6.3	6.6	5.8	5.7
이동시간이 많이 걸려서	4.6	4.8	4.7	4.7
기타	1.3	1.6	1.8	1.8
계	100.0	100.0	100.0	100.0

< 삶의 질 구성영역별 중요도 평가 >

(단위: 10점 척도)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
소득/소비	4.212	4.215	4.356	4.361	4.351	4.355
건강	4.552	4.546	4.608	4.618	4.592	4.592
노동	3.535	3.533	3.615	3.616	3.611	3.612
교육	3.916	3.920	4.032	4.003	3.966	3.966
주거/교통	3.647	3.662	3.749	3.759	3.748	3.750
안전	4.034	4.031	4.171	4.162	4.136	4.136
가족	4.578	4.570	4.604	4.601	4.560	4.561
환경	3.757	3.763	3.898	3.880	3.819	3.819
사회통합	3.207	3.215	3.425	3.433	3.386	3.387
여가	3.366	3.397	3.732	3.688	3.637	3.636

< 지난 1주간 경제활동 상태 >

(평균: 시간, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
하였다	55.5	55.8	53.7	52.3	49.2	49.1
평균	47.428	47.395	40.086	40.423	41.009	41.049
하지 않았다	44.5	44.2	46.3	47.7	50.8	50.9
계	100.0	100.0	100.0	100.0	100.0	100.0

< 종사상의 지위 >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
임금·봉급 근로자	65.1	68.2	80.7	78.1	76.2	76.0
고용원을 둔 사업주	6.5	6.7	3.8	4.0	4.1	4.2
고용원이 없는 자영자	20.9	18.9	9.9	11.2	12.7	12.9
무급가족종사자	7.5	6.2	5.6	6.7	6.9	7.0
계	100.0	100.0	100.0	100.0	100.0	100.0

< 배우자의 지난 1주간 경제활동 상태 >

(평균: 시간, 기타 단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
하였다	45.1	44.0	61.7	58.1	56.4	56.4
평균	45.175	44.715	37.733	37.883	38.555	38.631
하지 않았다	54.9	56.0	38.3	41.9	43.6	43.6
계	100.0	100.0	100.0	100.0	100.0	100.0

< 가구소득 >

(단위: %)

	FTF.raw	FTF.cell	Web.raw	Web.rim	Web.ps9	Web.ps7
50만원 미만	11.5	9.6	5.7	6.0	7.5	7.5
50 ~ 100만원 미만	15.2	13.8	6.1	6.3	7.0	7.0
100 ~ 200만원 미만	23.2	22.9	17.9	18.9	18.9	19.0
200 ~ 300만원 미만	21.9	22.4	23.3	22.9	23.5	23.5
300 ~ 400만원 미만	13.1	14.2	19.5	19.1	19.1	19.1
400 ~ 500만원 미만	7.1	7.8	12.6	12.3	11.1	11.0
500 ~ 600만원 미만	3.6	4.3	6.2	6.1	5.0	4.9
600 ~ 700만원 미만	1.6	1.8	3.5	3.3	3.2	3.2
700만원 이상	2.7	3.2	5.2	5.0	4.8	4.8
계	100.0	100.0	100.0	100.0	100.0	100.0

<부록 2> 문항별 응답 비율(%)산정방법(응답코드 1의 응답비율)

	유형	1	0*
TV시청	매체 접촉행태	있다	없다
TV 시청시간	매체 접촉행태	3시간 이상	3시간 이하
독서	매체 접촉행태	있다	없다
독서량	매체 접촉행태	지난1년 3권이상	지난 1년 3권 미만
국내관광여행 회수	관광, 여가, 문화활동	지난 1년 동안 3번 이상	지난 1년 동안 3번미만
해외여행 빈도	관광, 여가, 문화 활동	지난 1년 동안 1회 이상	지난 1년 동안 1회 미만
해외여행 경험 유무	관광, 여가, 유형	있다.	없다.
레저시설이용 횟수	관광, 여가, 문화 활동	지난 1년 동안 1회 이상	지난 1년 동안 1회 미만
문화예술및 스포츠	관광, 여가, 문화 활동	지난 1년 동안 2회 이상	지난 1년 동안 2회 미만
주말여가활동	관광, 여가, 문화 활동	TV시청, 가사일(장보기, 집안청소 등 포함), 휴식(수면, 사우나)	기타
주말여가활동 누구	관광, 여가, 문화 활동	가족과 함께	친구(연인 포함), 동호회(종교단체 포함), 혼자, 기타
하고 싶은 여가 활동	관광, 여가, 문화 활동	TV시청, 가사일(장보기, 집안청소 등 포함), 휴식(수면, 사우나)	기타
여가활용만족도	관광, 여가, 문화 활동	매우 만족, 약간 만족, 보통	약간 불만족, 매우 불만족
전통문화의 영향	인간 관계 및 생활 의례	매우 많은 영향 많은 영향	보통, 거의 영향 미치지 않음 전혀 영향 미치지 않음

* 모르겠다. 무응답 등도 0을 부여하였음.

	유형	1	0*
소득만족도여부	소득, 소비, 근로	보통, 약간 불만족 매우 불만족	매우 만족 약간 만족
장래소득기대	소득, 소비, 근로	약간 감소 크게 감소	크게 증가, 약간 증가 동일
계층의식	소득, 소비, 근로	상, 중	하
계층이동전망	소득, 소비, 근로	매우높다, 비교적 높다	비교적 낮다, 매우 낮다, 모르겠다
계층이동 (자식 세대)	소득, 소비, 근로	매우높다, 비교적 높다	비교적 낮다, 매우 낮다, 모르겠다
근로시간	소득, 소비, 근로	주 40시간 이상	주 40시간 미만
종사상의 지위	소득, 소비, 근로	임금, 봉급 근로자	기타
가구소득	소득, 소비, 근로	1년 월평균 소득 100-300만원	100만원 미만, 300만원 이상
주관적만족감	사회 관계	약간 불만족 매우 불만족	보통, 약간 불만족, 매우 불만족
각 경우의 사 랑 수	사회 관계	3명이상 있을 경우	3명 미만 있을 경우
후원금기부금	봉사, 자선활동	지난 1년 동안 2회 이상 기부했을 경우	지난 1년 동안 2회 미만 기부했을 경우
단체참여	봉사, 자선활동	지난 1년 동안 1회 이상 참여 했을 경우	지난 1년 동안 1회 미만 참여 했을 경우
봉사자선 활동	봉사, 자선활동	지난 1년 동안 1회 이상 참여했을 경우	지난 1년 동안 1회 미만 참여했을 경우
자원봉사활동 의사	봉사, 자선활동	하고 싶지만 지금은 어렵다.	기회가 주어진다면 한다. 할 생각이 없다.
민원서비스 만족도	봉사, 자선활동	매우 만족 약간 만족	보통, 약간 불만족 매우 불만족

* 모르겠다. 무응답 등도 0을 부여하였음.

	유형	1	0*
주관적건강평 가	생활 만족도	매우 좋다 좋은 편이다	보통이다, 나쁜 편이다 매우 나쁘다
스트레스인식 정도	생활 만족도	매우 많이 느꼈다. 많이 느낀 편이다.	조금 느낀 편이다, 거의 느끼지 않았다.
주택만족도	생활 만족도	매우 만족 약간 만족	보통, 약간 불만족, 매우 불만족
거주지역만족 도	생활 만족도	매우 만족 약간 만족	보통, 약간 불만족 매우 불만족
사회안전인식	생활 만족도	매우 안전하다 비교적 안전하다	보통이다, 비교적 불안하다 매우 불안하다
범죄피해두려 움	생활 만족도	매우 많이 느낀다. 약간 느낀다.	보통이다, 별로 느끼지 못한다 전혀 느끼지 못한다
가족관계만족 도	생활 만족도	매우 만족 약간 만족	보통, 약간 불만족 매우 불만족
환경오염체감	생활 만족도	매우 좋아졌다 약간 좋아졌다	변화 없다, 약간 나빠졌다 매우 나빠졌다.
자살충동경험 여부	생활 만족도	있다.	없다.
삶의 질 구성 요소별 중요도	생활 만족도	4점, 5점	1점, 2점, 3점

* 모르겠다. 무응답 등도 0을 부여하였음.