

청년패널조사에서의 웹조사(Web Survey) 도입 가능성에 대한 연구¹⁾

황광훈²⁾ · 천영민³⁾ · 조용운⁴⁾ · 이민정⁵⁾

요약

본 연구는 청년패널조사(YP2007) 14차(2020년) 자료를 바탕으로 웹조사(Web Survey)의 도입 가능성을 면밀히 검토하고, 조사방식 변경에 따른 응답자의 통계적 특성 차이를 실증적으로 분석하였다. 코로나19 팬데믹으로 인해 대면 면접조사가 제한된 환경에서 전체 응답자 9,539명 중 28.7%가 웹조사를 통해 참여하였다. 분석 결과, 웹조사 참여군은 대면조사군에 비해 고학력, 수도권 거주, 상용직 및 대기업 종사 비율이 유의미하게 높아 명확한 웹조사 참여자와 대면조사 참여자 간에 학력, 거주지역, 고용형태 및 사업체 규모 등의 관측 가능한 특성 차이가 나타났다. 이러한 선택편의를 제어하기 위해 성향점수매칭(PSM)을 적용하여 취업자 집단의 관측 가능한 특성을 통제한 이후에도, 웹조사 응답자군에서 대기업 종사 비율(42.3%)과 주 40시간 이상 장시간 근로 비율(56.2%)이 여전히 높게 나타나 관측 가능한 인구사회학적 특성을 통제한 이후에도 일부 노동시장 특성에서 조사방식별 차이가 관찰되었다. 아울러 최근 5년간의 패널 추적 이력을 분석한 결과, 대면조사의 연속 참여율(89.6%)이 웹조사(74.4%)를 크게 상회하여 조사방식 간 패널 지속 참여 의향과 표본 유지 안정성에 구조적 차이가 있음이 증명되었다. 본 연구 결과는 청년패널조사에서 웹조사를 활용할 때 웹조사를 기존 대면조사의 완전한 대체수단으로 활용하기보다는 조사 환경과 응답자 특성을 고려하여 대면조사와 상호 보완적으로 활용하는 혼합조사 방식의 가능성을 검토할 필요가 있음을 시사한다.

주요용어 : 청년패널조사, 웹조사, 혼합조사, 성향점수매칭, 모드효과, 패널유지율

1. 들어가며

청년패널조사처럼 청년층을 대상으로 정기적으로 자료를 수집하는 패널조사는 우리 사회에서 노동시장이 어떻게 변화해 왔는지를 장기간에 걸쳐 구체적으로 파악할 수 있게 해 주는 중요한 조사 방식이다. 노동시장 상황과 특성뿐 아니라 학교 교육과정, 사회적 이동, 가구 형성과 같이 개인의 생애 과정과 밀접하게 연관된 지표들은 특정 시점의 횡단면 조사만으로는 그 변화를 충분히 포착하기 어렵다. 이러한 한계로 인해 동일하거나 유사한 개인과 집단을 중·장기적으로 추적할 수 있는 패널조사의 필요성은

1) 본 논문은 2026 고용패널조사 학술대회에 발표한 「청년패널조사의 패널유지율 관리와 웹조사(Web Survey) 도입 방안에 대한 검토」의 내용을 대폭 수정 및 보완한 것이다.

2) 제1저자(교신저자): 한국고용정보원 고용통계조사팀, 부연구위원. E-mail: hunzzang96@keis.or.kr

3) 제2저자: 한국고용정보원 청장년고용지원팀, 연구위원. E-mail: zzari90@keis.or.kr

4) 제3저자: 한국고용정보원 고용통계조사팀, 책임연구원. E-mail: yucho0520@keis.or.kr

5) 제4저자: 한국고용정보원 고용통계조사팀, 위촉연구원. E-mail: mjlee12@keis.or.kr

점차 확대되고 있다. 다만, 이러한 장점이 실제로 충분히 발휘되기 위해서는 조사대상자가 매년 또는 정해진 주기에 걸쳐 지속적으로 조사에 참여하는 것이 전제되어야 한다. 결국 패널조사에서 표본의 조사 참여 수준을 나타내는 패널 유지율은 조사 자료의 신뢰성과 품질을 좌우하는 핵심적인 요소라고 할 수 있다.

그럼에도 불구하고 현실적으로 표본유지율을 안정적으로 유지하는 것은 매우 어렵다. 특히, 청년층은 생애주기에 따라 다양한 변화가 급격하고 빠르게 발생하는 집단으로 학교에서 노동시장 취업으로, 미혼에서 기혼(결혼)으로, 부모와 동거 상태에서 경제적 독립으로 전환과 변화가 빈번하게 일어난다. 이러한 변화는 대부분 거주지 이전이나 연락처 변경을 동반하기 때문에 조사기관의 면접원들은 매년 동일한 개인을 접촉하는 것이 쉽지 않다. 이는 단순한 조사협조 여부를 넘어서, 추적 자체가 불안정한 집단이기 때문에 관리문제에서 많은 애로사항이 발생한다고 볼 수 있다.

청년패널조사(YP2001, YP2007, YP2021)의 경우, 이러한 많은 환경적 어려움 속에서도 비교적 높은 표본유지율을 달성하고 있다. 청년패널조사(Youth Panel)는 한국고용정보원의 대표적인 고용통계 조사자료로서, 우리나라 청년층을 대상으로 학교생활, 취업을 포함한 사회·경제활동, 가계 배경 등을 매년 추적조사하는 종단조사(longitudinal survey)이며, 2001년부터 현재까지 20여 년 이상 조사가 실시되고 있다. 횡단면 또는 특정시점의 단년도 조사로는 정보를 수집하기 힘든 청년생애기의 이행 과정과 그 변화양상을 체계적으로 조사 및 분석하기 위해 설계되었다. 첫 조사는 2001년 당시 만 15~29세에 해당하는 약 8천 명을 대상으로 한 청년패널 2001(YP2001)을 시작으로 출범하였으며, 이후 2006년까지 추적조사를 실시한 뒤 종료되었다. 다음으로 2007년에 들어와서 동일 연령대(만15~29세)를 대상으로 한 청년패널 2007(YP2007)이 새롭게 기획되어 구축되었다. YP2007 코호트는 약 1만 명의 청년을 표본으로 구축하여 2020년까지 총 14차례의 조사가 완료되었다. 또한 코호트 중간 시점에서는 특정한 연령구간(하위 연령대)의 표본이 부족해지는 문제가 제기되어 2015년에는 기존 패널에 만15~22세 3천여명을 추가하는 표본 충원이 이루어졌다. 최근에는 2021년 기준 만19~28세 청년층을 대상으로 하는 청년패널 2021(YP2021)이 새롭게 기획되어 2026년 현재 6차 조사까지 진행 중이다. 이와같이 청년패널조사는 시기별로 각각의 코호트를 구축함과 동시에 청년층의 특성과 사회경제적 환경 등의 다양한 정보를 중·장기적으로 수집하는 조사체계를 유지해 오고 있다.

특히, YP2007 코호트의 경우 마지막 14차 조사에서 원표본 대비 표본유지율 66.5%를 기록했고, 2021년에 새롭게 구축된 YP2021은 4차 조사 기준 79.1%라는 높은 유지율을 달성하고 있다. 이러한 표본유지율 성과는 응답자 개인의 적극적인 협조 뿐만 아니라 한국고용정보원과 조사기관이 오랜기간 동안 축적해온 패널운영 및 관리전략, 면접원 교육, 조사 관련 사전 안내 체계, 응답자의 피로도 감소 전략 등 여러가지 요인들이 상호 결합되고 보완된 결과라 평가할 수 있다.

그러나 여전히 패널조사의 가장 큰 과제는 조사대상자의 지속적인 참여를 확보하는 것이다. 특히 청년층은 진학, 휴학, 졸업, 취업, 결혼, 독립 등 생애주기 변화가 빈번하게 발생하는 집단으로, 거주지 이전과 연락처 변경이 지속적으로 이루어진다. 또한 개인 정보 보호에 대한 민감성이 높고 조사 참여에 따른 피로감을 크게 느끼는 특성을 가지고

있어 조사 거절이나 부재가 상대적으로 많이 발생할 수 있다. 이와 같은 특성은 조사 대상자의 추적을 어렵게 만들고, 장기적으로는 패널이탈(panel attrition) 문제를 야기하게 된다. 패널이탈은 단순한 응답률 감소의 문제가 아니다. 조사 참여가 중단되면 분석 가능한 표본 수가 감소하게 되고, 특정 특성을 가진 집단의 이탈이 집중될 경우 통계적 대표성과 분석 결과의 신뢰성이 저하될 수 있다. 특히 패널조사에서는 한 차례 이탈한 응답자가 이후 조사에도 계속 참여하지 않는 단조이탈(monotone attrition)이 발생할 수 있으며, 이는 장기 종단자료의 품질을 저하시키는 주요 요인으로 작용한다. 따라서 패널조사의 안정적 운영을 위해서는 조사 참여를 유도하는 것뿐만 아니라 체계적인 패널유지 전략과 관리 체계를 구축하는 것이 필수적이다.

한편 최근 조사환경은 비대면 중심으로 빠르게 변화하고 있다. 스마트폰과 인터넷 활용이 보편화되고 코로나19 이후 비대면 조사 방식에 대한 수요가 증가하면서 웹조사(Web Survey)의 활용 가능성도 확대되고 있다. 웹조사는 응답자의 편의성과 조사 효율성을 높일 수 있다는 장점을 가지지만, 조사 참여 집단의 특성 차이나 자료 품질 문제 등 다양한 검토 사항도 함께 존재한다. 특히 장기 종단조사에서는 표본유지율과 자료의 연속성이 중요하기 때문에 웹조사의 활용 가능성을 패널관리 관점에서 함께 검토할 필요가 있다.

본 연구의 목적은 청년패널조사(YP2007) 14차 조사에서 코로나19의 영향으로 병행된 대면조사와 웹조사의 실제 조사 사례를 활용하여 조사방식에 따른 응답자의 특성 차이를 실증적으로 검토하는 데 있다. 구체적으로 전체 응답자의 조사방식별 구성과 최근 조사 참여 이력을 비교하고, 인구사회학적 특성의 차이를 고려하기 위해 성향점수매칭(PSM)을 활용하여 취업자를 대상으로 조사방식별 노동시장 특성을 비교하고자 한다. 이를 통해 웹조사 참여자의 특성 차이가 어떠한 양상으로 나타나는지를 확인하고, 향후 청년패널 조사에서 웹조사를 활용할 때 고려해야 할 조사운영 및 조사품질 측면의 시사점을 제시하고자 한다.

2. 이론적 논의 및 선행연구 검토

2.1 청년층의 고용 형태 다변화와 대면 패널조사의 방법론적 한계

최근 청년층의 노동시장 진입 양상은 과거의 정형화된 임금노동 궤적에서 벗어나, 플랫폼 노동과 같은 액화 노동(liquid labor)의 형태로 급격히 유연화되고 있다. 이러한 고용 형태의 다변화와 높은 이동성은 표본의 안정적인 추적이 필수적인 종단면 패널 조사에 심각한 방법론적 도전을 제기한다. 전 세계적으로 대면 면접조사(CAPI) 중심의 전통적 방식은 응답률의 지속적인 하락과 무응답 편향(nonresponse bias)의 위험에 직면해 있다(Groves, 2006; Brick & Williams, 2013). 특히 1인 가구 비중이 높고 주간 재택률이 낮은 청년 표본의 경우, 면접원의 물리적 접근 자체가 제한되어 심각한 포괄 오차(coverage error)를 유발한다(Sakshaug, Yan, & Tourangeau, 2010).

이러한 물리적, 시간적 제약을 극복하고 표본의 대표성을 유지하기 위한 대안으로 웹 조사(Web Survey)의 도입이 강력하게 요구되고 있다. 디지털 네이티브 세대인 청년층에게 모바일과 웹은 가장 자연스러운 의사소통 경로이며 시공간의 제약 없이 조사에 참여할 수 있는 환경을 제공한다(Couper, 2000; Tourangeau, Conrad, & Couper, 2013).

2.2 혼합조사(Mixed-Mode) 설계와 모드 효과의 통제

단일 웹 조사가 안고 있는 인터넷 비접근 계층의 배제 문제를 해결하고, 기존 대면 패널조사와의 연속성을 유지하기 위해 최근 통계 조사방법론에서는 혼합조사(Mixed-Mode Survey) 설계가 표준으로 자리 잡고 있다(De Leeuw, 2005; Dillman, Smyth, & Christian, 2014). 그러나 패널조사에 새로운 모드를 결합할 때 가장 핵심적인 통계적 쟁점은 모드효과(mode effect)의 발생이다. 모드효과는 특정 조사 방식을 선호하는 집단의 인구사회학적 특성에 따른 선택효과(selection effect)와 조사 매체의 특성이 인지적 응답 과정에 미치는 측정효과(measurement effect)로 엄격히 분리되어야 한다(Vannieuwenhuyze, Loosveldt, & Molenberghs, 2010; Vannieuwenhuyze & Loosveldt, 2012).

자기기입식인 웹 조사는 면접원이 개입하는 대면 조사에 비해 사회적 바람직성 편향(social desirability bias)을 크게 감소시켜 불안정한 노동 환경이나 민감한 소득 정보에 대해 보다 정직한 응답을 이끌어낸다(Tourangeau, Rips, & Rasinski, 2000; Kreuter, Presser, & Tourangeau, 2008; Heerwegh & Loosveldt, 2008). 하지만 이러한 측정상의 차이는 시계열 데이터의 구조적 단절을 초래할 수 있다. 따라서 패널 데이터에 통합 통제집단 합성통제법(SDID)이나 성향점수매칭-이중차분법(PSM-DiD)과 같은 정교한 인과추론 모형을 적용하기 위해서는 조사 모드 변경이 유발하는 내생성을 통제하고 응답의 대표성 지표(R-indicators)를 검증하는 방법론적 절차가 반드시 수반되어야 한다(Schouten, Cobben, & Bethlehem, 2009; Roberts, 2007).

2.3 해외 주요 패널조사의 웹 조사 이식 및 실증 사례

해외의 대규모 국가 통계 및 종단면 패널조사에서는 이미 대면 조사의 한계를 극복하기 위해 다각적인 혼합조사 실험을 거쳐 웹 조사를 성공적으로 이식해왔다. 영국의 가구패널인 Understanding Society(UKHS)는 별도의 혁신 패널(Innovation Panel)을 구축하여 무응답자의 웹 조사 전환율, 기기별 최적화(Responsive Design), 인센티브 효과를 장기적으로 실험 검증하였다(Burton, Laurie, & Lynn, 2014; Jäckle, Lynn, & Burton, 2015). 네덜란드의 LISS(Longitudinal Internet Studies for the Social sciences) 패널은 확률표본 기반 웹 패널(Probability-based Web Panel)의 구축이 국가 통계 수준의 신뢰도를 달성할 수 있음을 입증한 대표적 사례이다(Scherpenzeel, 2011; Blom, Gathmann, & Krieger, 2015).

미국의 건강 및 은퇴 연구(HRS) 또한 대면 조사 유지 비용의 급증에 대응하여 웹 서버이 파일럿을 실시하였으며, 조사 화면의 척도 크기 및 시각적 레이아웃이 데이터

신뢰도에 미치는 영향을 면밀히 분석하였다(Ofstedal, Couper, & McClain, 2016). 특히 유럽의 주요 조사들은 초기 접촉을 웹으로 시도하고 무응답 표본에 한해 대면 또는 전화조사를 실시하는 순차적 혼합 설계(sequential mixed-mode)를 채택함으로써, 비용 효율성과 표본의 포괄성을 동시에 달성하고 있다(Peytchev, Couper, McCabe, & Crawford, 2006; Klausch, Hox, & Schouten, 2013).

국내의 2021년 개인통행실태조사(김성중 외, 2023)나 한국노동패널(KLIPS)의 웹조사 실험 역시 이러한 세계적 조류와 궤를 같이한다. 결국 청년층 대상 패널조사의 품질을 고도화하기 위해서는 해외의 선행연구가 공통으로 시사하듯, 단순히 대면 조사를 웹으로 대체하는 수준을 넘어 문항 구조의 디지털 최적화(Börkan, 2010), 혼합 모드 간의 측정오차 보정(Biemer, 2001), 그리고 정밀한 표본 가중치 산출(Tourangeau, Conrad, & Couper, 2013)을 포괄하는 종합적인 조사 설계의 혁신이 이루어져야 한다.

3. 분석자료 및 분석방법

3.1 분석자료

본 연구는 청년패널조사에서 웹조사(Web Survey)의 활용 가능성을 검토하기 위하여 한국고용정보원 청년패널조사(YP2007) 14차 연도(2020년) 자료를 활용하였다. 청년패널 조사는 우리나라 청년층의 학교생활, 취업, 노동시장 이동, 결혼 및 가구형성 과정 등을 장기적으로 추적하는 대표적인 종단조사(longitudinal survey)이다. 특히 2020년 조사는 코로나19 확산으로 인해 기존의 대면조사(face-to-face interview)와 함께 웹조사(Web Survey)가 병행 실시되었다는 점에서 조사방식에 따른 응답자 특성을 비교·분석하기에 적합한 자료라 할 수 있다. 웹조사는 응답자가 인터넷을 통해 직접 설문에 참여하는 방식으로 조사비용 절감, 조사기간 단축, 응답 편의성 향상 등의 장점을 가지고 있다. 반면 조사 참여자의 특성에 따라 응답률 차이가 발생할 수 있으며, 특정 집단이 과대표집될 가능성도 존재한다. 따라서 본 연구에서는 청년패널조사 참여자 가운데 웹조사 참여자와 대면조사 참여자를 구분하여 조사방식에 따른 응답자 특성 차이를 분석하고자 하였다.

분석대상은 2020년 청년패널조사에 응답한 표본으로 한정하였으며, 조사방식에 대한 정보와 주요 분석변수에 결측치가 없는 사례를 최종 분석대상으로 활용하였다⁶⁾. 이를 통해 웹조사 참여 여부와 응답자의 인구사회학적 특성 및 노동시장 특성 간의 관련성을 검토하였다.

6) 본 연구의 분석대상은 청년패널조사(YP2007) 14차 조사 전체 응답자 9,539명을 대상으로 하였다. 이 중 조사방법별 특성 비교를 위한 기술통계 분석에는 전체 응답자를 활용하였으며, 취업자 분석에서는 취업자 6,035명을 분석대상으로 선정하였다. 이후 성향점수매칭(PSM)을 실시하여 웹조사 참여자 1,803명과 이에 대응하는 대면조사 참여자 1,803명을 매칭하였으며, 최종 분석에는 총 3,606명의 표본을 활용하였다. 또한 분석 과정에서 발생한 결측치 등 제외 기준과 제외 인원을 단계별로 제시하고, 각 표에는 표본 수(N)를 함께 제시하여 분석대상 선정 과정을 명확히 제시하였다.

3.2 주요 변수

본 연구의 주요 관심변수는 조사 참여 방식이다. 청년패널조사에서는 동일한 조사 내용에 대해 일부 응답자는 웹조사 방식으로 참여하였고, 나머지 응답자는 기존의 대면 조사 방식으로 참여하였다. 이에 따라 본 연구에서는 웹조사 참여 여부를 기준으로 응답자를 구분하고, 두 집단 간 특성 차이를 비교하였다. 설명변수는 크게 인구사회학적 특성, 노동시장 특성 및 사업체 특성으로 구성하였다. 먼저 인구사회학적 특성으로는 성별, 연령, 혼인상태, 학력수준 및 거주지역을 활용하였다. 이러한 변수들은 인터넷 접근성 및 디지털 활용 수준과 밀접한 관련이 있으며, 웹조사 참여 가능성에도 영향을 미칠 수 있는 요인으로 알려져 있다. 특히 학력수준과 수도권 거주 여부는 정보기술 접근성과 관련성이 높아 웹조사 참여 여부를 설명하는 중요한 변수로 판단하였다.

다음으로 노동시장 특성으로는 취업 여부, 종사상 지위, 상용직 여부, 임금수준 및 근속기간 등을 활용하였다. 일반적으로 안정적인 일자리에 종사하거나 임금수준이 높은 집단일수록 정보통신기기 활용 수준이 높고 온라인 환경에 대한 접근성이 우수할 가능성이 존재한다. 또한 취업상태와 근로조건에 따라 조사 참여 방식에 대한 선호가 달라질 수 있기 때문에 노동시장 관련 변수를 주요 설명변수로 포함하였다. 마지막으로 사업체 특성으로는 산업분류와 사업체 규모를 활용하였다. 산업 및 기업규모는 근로자의 업무환경과 정보기술 활용 수준에 영향을 미치는 요인으로 볼 수 있으며, 특히 대기업 종사자나 전문·사무직 중심 산업 종사자의 경우 상대적으로 웹조사 참여 비율이 높을 가능성이 존재한다. 이에 따라 사업체 특성을 함께 고려함으로써 조사방식 선택에 영향을 미치는 요인을 보다 종합적으로 검토하고자 하였다.

3.3 분석방법

본 연구에서는 먼저 기술통계분석을 통해 웹조사와 대면조사 참여자의 특성을 비교하였다. 이후 조사방식별 응답자 구성의 차이가 취업자 특성 비교에 영향을 미칠 가능성을 고려하여 성향점수매칭(PSM)을 실시하였다. PSM에서는 웹조사 참여 여부를 처치변수로 설정하고, 취업자 응답자를 대상으로 성별, 연령, 학력 및 거주지역을 공변량으로 활용하여 성향점수를 추정하였다. 이후 최근접 이웃 매칭을 적용하여 관측 가능한 인구사회학적 특성이 유사한 웹조사 및 대면조사 참여자를 구성하고, 매칭 전후의 공변량 균형성을 확인하였다.

성향점수는 로지스틱 회귀분석을 이용하여 추정하였으며, 성별, 연령, 학력, 거주지역, 취업상태, 산업 및 사업체 규모 등의 변수를 활용하여 웹조사 참여 확률을 산출하였다. 이후 최근접 이웃 매칭(Nearest Neighbor Matching)을 적용하여 유사한 특성을 가진 웹조사 참여자와 대면조사 참여자를 비교하였다. 이를 통해 조사방식에 따른 응답자 특성 차이를 보다 엄밀하게 검증하고, 향후 청년패널조사에서 웹조사 도입 시 고려해야 할 시사점을 도출하고자 하였다.

4. 웹조사(Web Survey) 도입 가능성에 대한 검토

4.1 웹조사의 개념, 도입 배경, 장점 및 한계

패널조사는 동일한 표본을 장기간 추적하는 조사 방식으로, 시간에 따른 개인 및 가구의 변화 과정을 분석할 수 있다는 점에서 중요한 통계자료로 활용된다. 그러나 이러한 종단조사의 특성상 시간이 경과할수록 응답자의 이탈, 즉 패널이탈(panel attrition)이 발생하며, 이는 자료의 대표성과 신뢰성을 저해하는 주요 요인으로 작용한다. 특히 강력한 조사 거절이나 추적 실패와 같은 무응답 요인이 누적되면서 패널 유지율이 지속적으로 감소하는 문제가 발생한다. 이러한 문제를 해결하기 위한 다양한 방안이 논의되고 있으며, 최근에는 웹 기반 조사 방식의 도입이 하나의 대안으로 주목받고 있다. 웹조사는 응답자가 인터넷을 통해 설문에 직접 응답하는 자기기입식 조사 방식으로, 전통적인 면접조사와는 달리 조사원의 개입 없이 진행된다는 특징을 가진다(김유빈 외, 2017). 이는 종이 설문을 사용하는 우편조사와 유사하지만, 조사 매체가 웹이라는 점에서 차별성을 갖는다.

전통적으로 사회과학 분야에서의 조사자료 수집은 대면 면접조사(face-to-face interview)를 중심으로 이루어져 왔다. 이는 조사원이 직접 응답자를 방문하여 설문을 진행하는 방식으로, 응답률이 높고 응답의 정확성을 확보할 수 있다는 장점이 있다. 그러나 조사 비용이 높고 시간 소요가 크다는 구조적 한계를 가지고 있다. 이후 전화조사(CATI)와 우편조사(PAPI) 등이 등장하면서 조사 방식의 다양화가 이루어졌으며, 정보통신기술의 발전과 함께 웹 기반 조사(CAWI)가 빠르게 확산되었다. 또한 코로나 19와 같은 감염병 확산은 조사방법의 비대면 전환을 가속화하였다. 많은 국가에서 대면 조사가 중단되거나 축소되면서 웹 및 모바일 조사를 활용한 자료 수집이 확대되었고, 이는 웹조사가 단순한 보조적 수단을 넘어 핵심적인 조사방법으로 활용될 수 있음을 보여주었다. 웹조사는 인터넷 환경을 활용하여 응답자가 직접 설문에 응답하는 방식으로, 조사 비용 절감과 자료 수집의 신속성이라는 측면에서 주목받고 있다. 특히 2000년대 이후에는 웹조사가 단독 조사 방식으로 활용되기보다는, 기존 조사 방식과 결합된 혼합조사(mixed-mode survey)의 형태로 발전하고 있다. 이는 단일 조사 방식이 가지는 한계를 보완하기 위한 전략적 선택으로 이해된다.

웹조사는 비용, 접근성, 데이터 처리 측면에서 다양한 장점을 가진다. 우선 가장 중요한 특징은 비용 효율성이다. 전통적인 대면 면접조사는 조사원에 대한 인건비, 교통비, 교육비 등 다양한 비용이 발생하는 반면, 웹조사는 이러한 비용이 거의 발생하지 않는다고 볼 수 있으며, 특히 응답이 자동으로 데이터화 되기 때문에 별도의 코딩 및 입력 과정이 필요 없으며, 데이터 처리 비용을 크게 절감할 수 있다(백지선 외, 2017). 이러한 특성은 동일한 예산 하에서 더 많은 표본을 확보하거나 조사 빈도를 증가시키는데 기여할 수 있다.

다음으로 웹조사는 응답 편의성과 접근성 측면에서도 큰 장점을 갖고 있다. 조사 응답자는 자신이 원하는 시간과 장소에서 설문에 참여할 수 있으며, 이는 특히 직장인이나 학생과 같이 시간 제약이 큰 집단에서 응답 참여를 유도하는 데 효과적이다(장윤재 외, 2011). 또한 웹조사는 인터넷 활용도가 높은 젊은층이나 1인 가구와 같은

집단에 효과적으로 접근할 수 있어 기존 조사방법에서 포착하기 어려웠던 집단을 포함시키는 데 유리하다(장운재 외, 2011).

그리고, 데이터 처리의 신속성과 정확성 역시 웹조사의 중요한 장점 중 하나이다. 응답이 실시간으로 전산화되기 때문에 자료 입력 과정에서 발생할 수 있는 오류를 최소화할 수 있으며, 조사 완료 후 즉시 분석이 가능하다. 또한 설문 과정에서 논리적 오류를 자동으로 검증할 수 있어 데이터 품질 향상에도 기여한다(김성중 외, 2023). 이러한 특성은 특히 반복적으로 자료를 수집하는 패널조사에서 큰 장점으로 작용한다.

비대면 조사환경에 대한 대응력도 웹조사의 주요 장점이다. 감염병 상황이나 사회적 거리두기 환경에서도 안정적으로 자료 수집이 가능하며, 실제로 국가통계 조사에서도 웹 및 모바일 방식이 적극적으로 활용되고 있다. 또한 웹조사는 응답자가 비교적 자유로운 환경에서 응답할 수 있기 때문에 조사원 앞에서 응답하는 경우보다 심리적 부담이 적다. 이에 따라 민감한 질문에 대해 보다 솔직한 응답을 유도할 가능성이 있으며, 이는 특정 조사 주제에서는 데이터의 질을 향상시키는 요인으로 작용할 수 있다.

이와 같이 웹조사는 비용 절감, 응답 편의성, 데이터 처리 효율성, 비대면 대응력 등 다양한 측면에서 강점을 가지며, 이러한 장점들은 웹조사가 현대 조사환경에서 빠르게 확산되는 주요한 이유로 작용하고 있다.

하지만, 웹조사는 다양한 장점을 가지고 있지만, 동시에 몇 가지 중요한 한계를 내포하고 있다. 가장 대표적인 문제는 포괄오차(coverage error)이다. 웹조사는 인터넷 접근이 가능한 집단만 조사에 참여할 수 있기 때문에 고령층이나 저소득층, 정보취약계층이 조사에서 배제될 가능성이 존재한다. 이러한 문제는 표본의 대표성을 저하시킬 수 있으며, 조사 결과의 일반화 가능성을 제한하는 요인이 된다.

또한 웹조사는 자기선택 편의(self-selection bias)의 문제가 발생하기 쉽다. 응답자가 자발적으로 설문에 참여하는 경우가 많기 때문에 특정 성향을 가진 집단이 과대표집될 가능성이 있으며, 이는 조사 결과의 편향으로 이어질 수 있다.

응답률 문제 역시 중요한 한계로 지적된다. 웹조사는 일반적으로 면접조사에 비해 응답률이 낮은 경향을 보이며, 이는 무응답오차를 증가시키는 요인이 된다. 실제 사례에서도 웹조사에서 낮은 응답률이 관찰되었으나, 일부 연구에서는 표본 특성이 크게 왜곡되지 않을 수 있다는 결과도 보고되고 있다.

웹조사의 또 다른 핵심적인 문제는 모드효과이다. 웹조사와 면접조사는 응답 환경과 방식이 다르기 때문에 동일한 질문에 대해서도 서로 다른 응답이 나타날 수 있다. 이러한 차이는 조사방법을 선택한 집단의 특성 차이에 의해 발생하는 선택효과와, 조사방법 자체의 특성에 의해 발생하는 측정효과로 구분된다. 웹조사는 사회적으로 바람직하지 않은 행동에 대해 보다 솔직한 응답을 유도하는 반면, 면접조사는 보다 중립적이거나 바람직한 응답을 유도하는 경향이 있는 것으로 나타난다.

4.2 국내·외 유사 패널조사 사례

국내·외 유사 패널조사 사례를 분석하는 것은 청년패널조사의 발전 방향을 설정하는 데 중요한 기초자료를 제공한다. 특히 패널조사는 동일한 응답자를 장기간 추적하는 특성상

조사 방식, 표본 유지, 조사 내용 구성 등에서 지속적인 개선이 요구되므로, 선행 사례 분석은 필수적인 과정이라 할 수 있다.

먼저 국내 사례를 살펴보면, 한국교육고용패널조사(KEEP)와 한국교육고용패널조사 II(KEEP II)는 대표적인 청년 관련 종단조사로, 교육에서 노동시장으로의 이행 과정을 장기적으로 추적하고 있다. 이 조사들은 초기에는 대면조사를 중심으로 운영되었으나, 이후 전화조사와 웹조사를 병행하는 혼합조사 방식으로 전환하였다. 이는 조사 환경 변화와 응답자 접근성 문제에 대응하기 위한 전략으로, 단일 조사 방식보다 높은 응답률과 유연성을 확보할 수 있다는 점에서 중요한 시사점을 제공한다. 또한 한국아동·청소년패널조사(KCYPS)는 개인의 성장과 발달뿐만 아니라 가정, 학교, 사회 환경을 함께 조사함으로써 패널조사의 내용이 단순한 경제활동에 국한되지 않고 삶의 전반을 포괄해야 한다는 점을 보여준다. 이는 청년패널조사 역시 노동시장뿐 아니라 삶의 질, 가치관, 사회적 관계 등 다양한 요소를 포함할 필요가 있음을 시사한다.

한편 서울청년패널과 부산청년패널과 같은 지역 기반 조사에서는 웹조사를 적극적으로 도입하고 있으며, 패널 이탈을 보완하기 위해 신규 표본을 추가하거나 이동한 응답자를 ‘위성패널’로 관리하는 전략을 활용하고 있다. 이러한 사례는 청년층의 높은 이동성을 고려할 때 패널 유지 전략이 매우 중요하며, 웹조사가 이를 보완할 수 있는 효과적인 수단이 될 수 있음을 보여준다.

국외 사례를 살펴보면, 미국의 NLSY(National Longitudinal Survey of Youth)는 대표적인 청년 종단조사로, 조사 방식이 종이 설문(PAPI; Paper and Pencil Interviewing)에서 전화조사(CATI; Computer-Assisted Telephone Interviewing), 컴퓨터 기반 조사(CAPI; Computer-Assisted Personal Interviewing), 그리고 전화 중심 조사로 지속적으로 변화해 왔다. 이는 기술 발전과 비용, 응답 환경 변화에 따라 조사 방식이 유연하게 조정되어야 함을 보여준다. 영국의 LSYPE(Longitudinal Survey of Young People in England)는 혼합조사 방식의 대표적인 사례로, 웹조사를 우선 실시한 후 응답하지 않은 경우 전화조사, 이후에도 응답하지 않을 경우 대면조사를 실시하는 단계적 접근 방식을 사용하고 있다. 이러한 구조는 응답자의 참여를 최대화하면서도 비용을 효율적으로 관리할 수 있는 방식으로 평가된다. 캐나다의 YITS(Youth in Transition Survey)와 호주의 LSAY(Longitudinal Surveys of Australian Youth) 역시 초기에는 전화조사를 중심으로 운영되었으나 점차 온라인조사를 도입하는 방향으로 변화하고 있다. 이는 전 세계적으로 웹조사가 점차 핵심적인 조사 방식으로 자리잡고 있음을 보여준다.

정리해보면, 국내·외 유사 조사들 중 상당수는 공통적으로 혼합조사 방식을 채택하고 있으며, 조사 방식은 기술 발전에 따라 점진적으로 변화하고 있다. 또한 패널 유지율을 확보하기 위해 신규 표본 도입, 탈락자 보정, 이동 집단 추적 등 다양한 전략이 활용되고 있으며, 조사 내용 역시 노동시장 중심에서 삶 전반으로 확대되는 경향을 보인다.

4.3 청년패널조사에서의 웹조사 사례 분석

본 절에서는 청년패널조사(YP2007) 14차(2020년) 자료를 활용하여 조사방법(대면조사 vs 웹조사)에 따른 응답자의 주요 특성을 비교하였다. 2020년 청년패널조사(8-12월

조사 실시)에서는 코로나 19의 영향으로 인해 사회적 거리두기가 확대 및 강화되어 대면조사의 어려움이 발생하였다. 이에 청년패널조사에서는 코로나 19로 인한 정부의 방역지침 등을 고려하면서 실사를 진행하였다. 코로나19의 확산으로 인해 수도권은 8월24일부터 추석연휴까지 대면조사가 전면 중단되었으며, 비수도권의 경우에도 9월 14일부터 2주간 대면조사가 일시 재개되었으나, 추석특별 방역기간에 다시 대면조사가 중단되는 등, 8주~13주차사이의 조사가 원활히 진행되지 못하였다.

<그림 4.1> YP2007 14차 조사(2020년) 주차별 진행 현황(기존패널)



자료: 청년패널조사(YP2007) 자료, 한국고용정보원

<그림 4.2> YP2007 14차 조사(2020년) 주차별 진행 현황(추가패널)

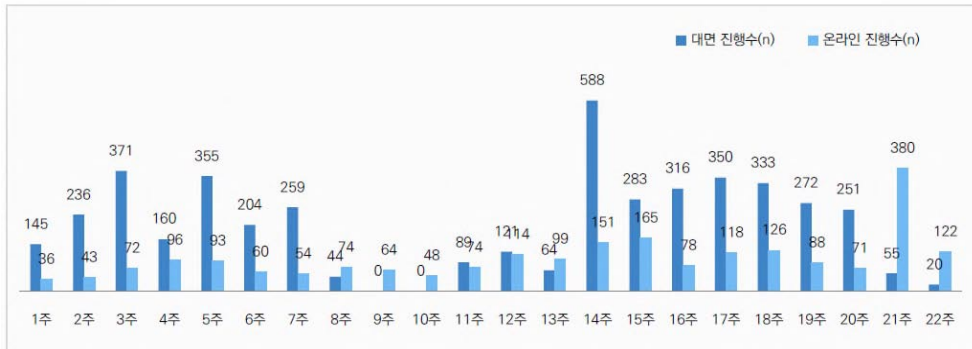


자료: 청년패널조사(YP2007) 자료, 한국고용정보원

조사방법별 주차별 진행 현황을 보면, 1주부터 7주까지 대면중심으로 건인되었던 기존 패널실사는 8주 이후 사회적거리두기 단계상향 및 대면조사 제한조치 등으로 인해 9주차와 10주차는 소극적인 온라인 조사(웹조사)로만 운용되었다. 그리고 14주차 이후 전 지역에 대한 대면조사가 재개되면서 주차별 진행수가 대폭 개선된 양상을 보였다. 또한 온라인 조사(웹조사) 진행 역시 대면이 활발히 진행되었을 때 함께 높아지는 경향을 보였다. 21주차 이후에는 대면조사 의향이 낮은 패널을 집중적으로 독려하여 조사를 진행했기 때문에 온라인 조사(웹조사)의 수가 증가하였다.

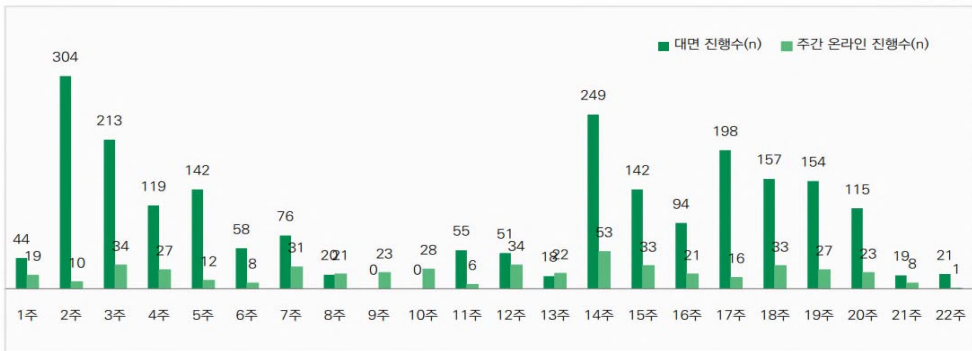
한편, 본 연구에서 활용한 자료는 14차 조사에서의 최종 조사방법과 조사 참여 이력을 중심으로 구성되어 있어, 개별 응답자가 최근 5년간 각 조사차수에서 어떠한 조사모드로 참여하였는지, 또는 동일 응답자가 대면조사에서 웹조사로 전환하였는지를 일관되게 확인하는 데에는 한계가 있다. 따라서 본 연구에서 나타난 조사방법별 최근 5년간 연속 참여율의 차이를 조사모드 자체의 효과로 해석하기보다는, 2020년 조사 당시의 조사 환경과 조사운영 방식이 함께 반영된 기술통계적 결과로 해석할 필요가 있다.

<그림 4.3> YP2007 14차 조사(2020년) 조사방법별 주차별 진행 현황(기존패널)



자료: 청년패널조사(YP2007) 자료, 한국고용정보원

<그림 4.4> YP2007 14차 조사(2020년) 조사방법별 주차별 진행 현황(추가패널)

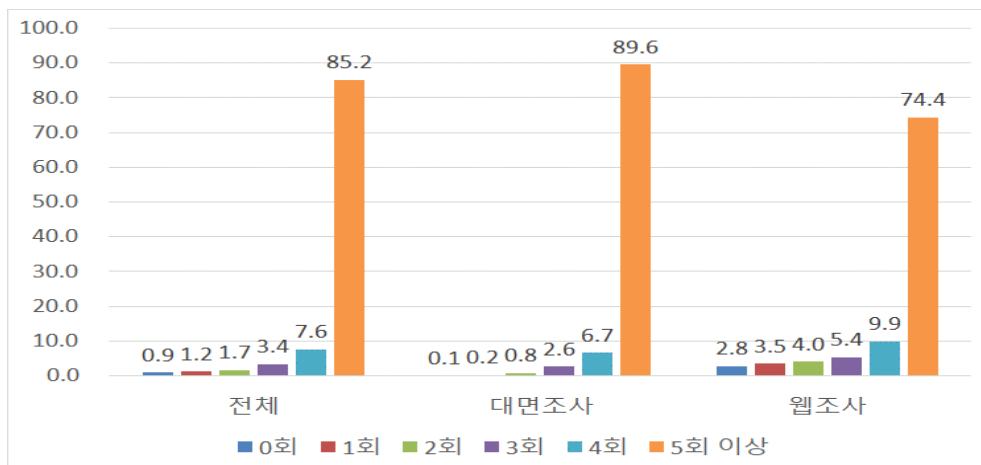


자료: 청년패널조사(YP2007) 자료, 한국고용정보원

14차 조사 전체 응답자 9,539명 가운데 대면조사 참여자는 71.3%, 웹조사 참여자는 28.7%로 나타났다. 다음으로 14차 조사(2020년) 참여자의 최근 5년간 연속 참여 여부를 보면(<그림 4.5>), 전체적으로 연속참여자의 비율이 85.2%로 높게 나타났으며, 대면조사의 경우 89.6%로 웹조사(74.4%)보다 높은 수준을 보였다. 이는 조사방법에 따라 연속 참여 비율에 차이가 존재함을 보여준다. 또한 연속 참여 횟수에서도 5회 이상 참여 비율이 대면조사에서 더 높게 나타나, 웹조사와 비교할 때 참여 누적 정도에 차이가 있음을 확인할 수 있다.

<그림 4.5> YP2007 14차 조사(2020년) 참여자의 최근 5년간 연속 참여 횟수

(단위: %)



자료: 청년패널조사(YP2007) 자료, 한국고용정보원

조사방법에 따른 참여자의 인구학적 특성을 보면(<표 4.1>), 성별 구성은 조사방법 간 큰 차이를 보이지 않았다. 반면 학력에서는 웹조사의 경우 대졸 이상 비율이 54.7%로 대면조사(43.1%)보다 높게 나타났으며, 평균 연령 역시 웹조사가 31.9세로 대면조사(30.5세)보다 다소 높았다. 다음으로 거주지역 분포를 보면(<표 4.2>), 웹조사에서는 서울(30.1%)과 경기(32.2%)의 비중이 높게 나타났고, 대면조사는 부산 등 일부 비수도권 지역의 비중이 상대적으로 높았다.

<표 4.1> 조사 방법에 따른 참여자의 인구학적 특성 비교

(단위: 명, %)

구분		전체	대면조사	웹조사
전체(표본수)		9,539	6,801	2,738
성별	남성	48.6	48.5	48.8
	여성	51.4	51.5	51.2
학력별	고졸이하	32.4	35.2	25.2
	전문대졸	21.3	21.7	20.1
	대졸이상	46.3	43.1	54.7
연령별	25세 미만	21.6	24.5	14.2
	25-29세	21.4	21.8	20.5
	30-34세	24.1	22.3	28.8
	35-42세	32.9	31.5	36.6
	평균(세)	30.9	30.5	31.9

자료: 청년패널조사(YP2007) 자료, 한국고용정보원

<표 4.2> 조사 방법에 따른 참여자의 거주지역 특성 비교

(단위: %)

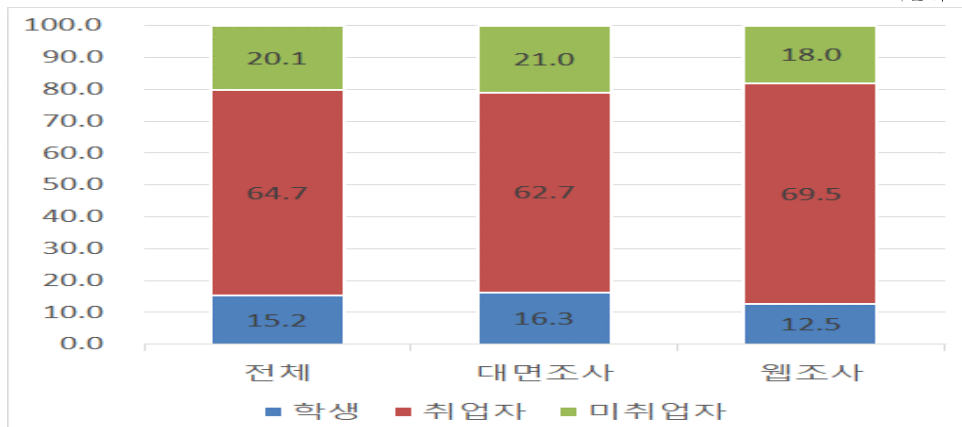
구분	전체	대면조사	웹조사
서울	20.1	16.1	30.1
부산	9.9	12.9	2.3
대구	6.3	5.9	7.5
인천	8.1	8.7	6.4
광주	4.2	3.6	5.6
대전	4.9	6.5	0.8
울산	2.2	2.7	0.8
경기	23.3	19.7	32.2
강원	2.1	2.6	0.9
충북	2.8	3.6	0.7
충남	2.5	3.2	0.6
전북	2.5	2.4	2.8
전남	2.3	2.4	1.9
경북	3.1	2.9	3.5
경남	4.6	5.7	1.9
제주	1.0	0.7	1.6
세종	0.3	0.3	0.3

자료: 청년패널조사(YP2007) 자료, 한국고용정보원

경제활동상태를 비교해보면(<그림 4.6>), 전체적으로 취업자의 비율이 64.7%로 가장 높았으며, 웹조사에서는 취업자 비율이 69.5%로 대면조사(62.7%)보다 높게 나타났다. 반면 학생과 미취업자의 비중은 대면조사에서 상대적으로 높았다. 그리고, 취업자의 종사상 지위 및 기업체 규모를 기준으로 보면(<표 4.3>), 웹조사에서는 상용직 비율(83.8%)과 대기업 종사 비율(42.0%)이 대면조사(79.7%, 31.8%)보다 높게 나타났다. 반면 30인 미만 사업체 종사자 비율은 대면조사(44.9%)가 웹조사(36.8%)보다 더 높았다.

<그림 4.6> 조사 방법에 따른 참여자의 경제활동상태 비교

(단위: %)



자료: 청년패널조사(YP2007) 자료, 한국고용정보원

<표 4.3> 종사상 지위 및 기업체규모 특성 비교

(단위: %)

구분		전체	대면조사	웹조사
종사상지위	상용직	81.0	79.7	83.8
	임시/일용직	9.9	10.0	9.5
	자영업자	9.2	10.3	6.7
기업체규모	30인 미만	42.4	44.9	36.8
	30-100인 미만	13.5	14.4	11.6
	100-300인 미만	9.1	8.9	9.6
	300인 이상(대기업)	34.9	31.8	42.0

자료: 청년패널조사(YP2007) 자료, 한국고용정보원

산업 및 직업 분포(<표 4.4>)에서는 산업의 경우 대면 및 웹조사 방법에 따른 참여 비중이 거의 비슷하였다. 직업별로는 웹조사에서 경영·사무·금융·보험직(37.9%)과 연구·공학기술직(15.5%)의 비중이 상대적으로 높게 나타난 반면, 대면조사에서는 영업·판매·운전·운송직(13.6%)과 설치·정비·생산직(10.1%)의 비중이 더 높은 것으로 나타났다.

<표 4.4> 산업 및 직업 비교

(단위: %)

구분		전체	대면조사	웹조사
산업	제조업	19.5	19.0	20.7
	서비스업	75.3	75.5	74.7
	기타	5.2	5.5	4.6
직업	경영·사무·금융·보험직	33.9	32.1	37.9
	연구직 및 공학 기술직	11.5	9.7	15.5
	교육·법률·사회복지·경찰·소방직 및 군인	12.0	11.9	12.2
	보건·의료직	7.5	7.7	7.1
	예술·디자인·방송·스포츠직	4.5	4.2	5.3
	미용·여행·숙박·음식·경비·청소직	7.7	8.7	5.7
	영업·판매·운전·운송직	11.8	13.6	7.8
	건설·채굴직	1.2	1.4	0.7
	설치·정비·생산직	9.1	10.1	6.9
	농림어업직	0.7	0.7	0.9

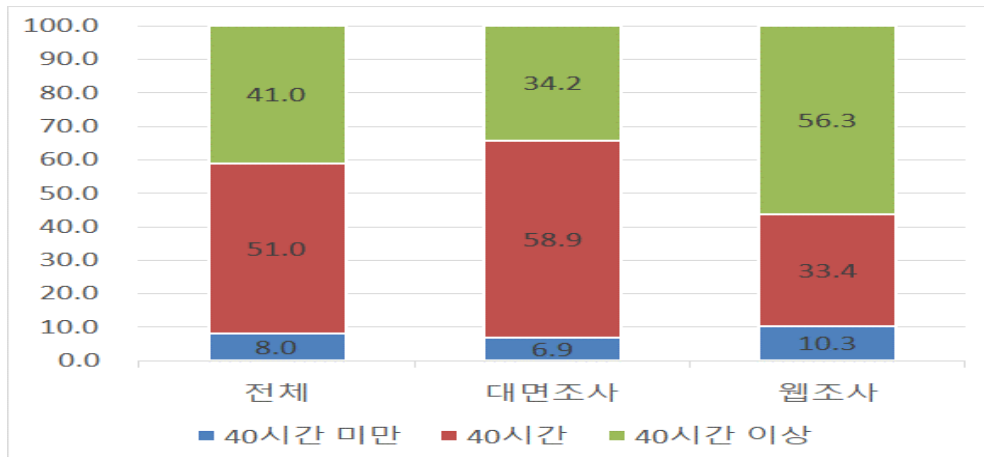
자료: 청년패널조사(YP2007) 자료, 한국고용정보원

다음으로 근로시간을 기준으로 보면(<그림 4.7>), 웹조사에서는 주당 40시간 이상 근로자 비율이 56.3%로 가장 높게 나타났으며, 대면조사에서는 주당 40시간 근로자 비율이 58.9%로 가장 높은 비중을 차지하였다. 마지막으로 월평균 소득을 비교한 결과

(<그림 4.8>), 웹조사에서는 300만원 이상 구간 비율이 46.4%로 대면조사(41.9%)보다 높게 나타났으며, 대면조사에서는 200~300만원 구간의 비율(43.7%)이 상대적으로 높은 것으로 나타났다.

<그림 4.7> 주당 평균근로시간 비교

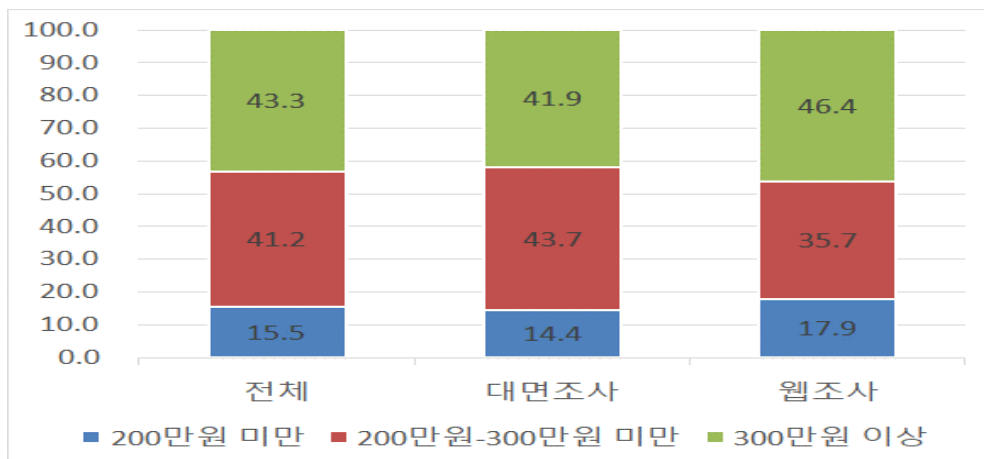
(단위: %)



자료: 청년패널조사(YP2007) 자료, 한국고용정보원

<그림 4.8> 월평균 소득 비교

(단위: %)



자료: 청년패널조사(YP2007) 자료, 한국고용정보원

이상의 결과를 종합하면, 웹조사는 조사 접근성 확대와 조사 지속성 확보 측면에서 중요한 역할을 수행하고 있으나, 응답자 구성 측면에서는 일정한 선택편의가 존재하는 것으로 나타난다. 즉, 웹조사 참여자는 대면조사 참여자에 비해 고학력자, 수도권 거주자 및 상대적으로 안정적인 고용상태에 있는 응답자의 비중이 높은 것으로 나타났다. 다만 이러한 차이는 응답자의 조사방식 선택뿐만 아니라 코로나19 당시의 조사환경과 조사기관의 조사모드 운영방식, 응답자의 조사 참여 가능성 등이 복합적으로 반영된 결과일

수 있으므로, 이를 순수한 자기선택편의의 결과로 단정하기는 어렵다. 이러한 차이는 단순한 응답 방식의 차이를 넘어, 분석 결과의 해석에도 영향을 미칠 수 있는 중요한 요소로 작용할 수 있다. 따라서 청년패널조사에서 웹조사를 활용할 경우에는 대면조사를 완전히 대체하는 방식보다는, 두 조사 방식을 상호 보완적으로 결합하는 혼합조사 도입 방안에 대한 검토가 필요할 것이다. 특히 조사 초기 단계에서는 대면조사를 중심으로 패널과의 관계를 형성하고, 이후 조사 참여가 어려운 집단이나 이탈 위험이 높은 집단에 대해서는 웹조사를 활용하는 방식이 효과적일 수 있다.

4.4 성향점수매칭(Propensity Score Matching: PSM)을 활용한 조사방식 간 응답자(취업자) 특성 비교

본 절에서는 14차(2020년) 조사에 참여한 취업자 유형 응답자(기존패널+추가패널)를 대상으로 웹조사와 대면조사 간 특성 차이를 비교하기 위해 성향점수매칭(Propensity Score Matching: PSM)을 실시하였다. PSM은 관찰자료에서 집단 간 연령, 학력, 지역 등 사회인구학적 특성의 차이로 인해 발생할 수 있는 선택편의를 줄이기 위한 방법으로, 각 개인의 특정 집단에 속할 확률인 성향점수를 추정한 후 유사한 성향을 가진 응답자를 매칭하여 집단 간 특성을 비교하는 분석방법이다. 본 연구에서는 웹조사 참여 여부를 기준으로 성향점수를 추정하고, 성별, 연령, 학력, 지역을 공변량으로 활용하여 최근접 이웃 매칭(nearest neighbor matching)을 실시하였다. 이를 통해 웹조사 응답자와 유사한 특성을 가진 대면조사 응답자를 구성하여 조사방식 간 응답자 특성 차이를 통제하고자 하였으며, 매칭 전후 공변량의 균형성(balance)을 검토하여 매칭의 적절성을 확인하였다.

본 연구에서 실시한 성향점수매칭(Propensity Score Matching, PSM)은 웹조사와 대면조사 간 조사방식 자체의 순수한 모드효과(mode effect)를 직접 추정하기 위한 목적이라기보다는, 조사방식별 응답자 집단 간에 존재하는 관측 가능한 인구사회학적 특성의 차이를 통제한 이후에도 노동시장 특성에서 차이가 나타나지는지를 비교하기 위한 목적으로 활용하였다. 특히 본 연구에서 활용한 2020년 청년패널조사 자료는 조사모드가 무작위로 배정된 실험자료가 아니라 코로나19라는 특수한 조사환경에서 실제 조사과정 중 대면조사와 웹조사가 병행된 자료라는 점에서, 조사방식 자체의 순수한 효과와 조사환경 및 응답자 특성에 따른 차이를 완전히 구분하는 데에는 한계가 있다. 따라서 PSM 분석 이후 나타난 조사방식별 차이를 조사방식 자체의 인과적 효과로 해석하기보다는, 관측 가능한 응답자 특성을 통제한 이후에도 조사방식별 노동시장 특성에 차이가 존재한다는 점을 확인하는 보완적 분석으로 해석할 필요가 있다. 향후에는 조사모드 이력 및 조사운영자료(paradata) 등을 활용하여 조사방식의 변화가 응답 결과와 패널유지에 미치는 영향을 종단적으로 검증함으로써 조사방식에 따른 모드효과를 보다 엄밀하게 분석할 필요가 있다.

또한 본 연구에서는 청년패널조사의 고용 및 노동시장 영역에 초점을 두어 취업자를 대상으로 PSM 분석을 실시하였다. 취업자의 경우 종사상 지위, 사업체 규모, 산업, 근로시간 및 소득 등 청년패널조사의 주요 고용 관련 변수에 대한 비교가 가능하므로 조사방식에 따른 노동시장 특성 차이를 검토하는 데 적합하다고 판단하였다.

<표 4.5> 분석 표본의 매칭 전후 빈도표

(단위: 명, %)

구분	매칭 전		매칭 후	
	빈도	비율	빈도	비율
전체	6,035	100.0	3,606	100.0
대면조사	4,232	70.1	1,803	50.0
웹조사	1,803	29.9	1,803	50.0

자료: 청년패널조사(YP2007) 자료, 한국고용정보원

매칭 전 조사방식별 취업자 유형인 응답자 특성을 비교한 결과, 웹조사와 대면조사 간 성별을 제외한 연령, 학력, 지역에서 특성 차이가 나타났다. 우선 성별 분포는 대면조사 남성 52.95%, 웹조사 남성 53.24%로 유사하게 나타났다. 반면 평균연령은 대면조사 32.4세, 웹조사 33.2세로 웹조사 응답자의 연령이 다소 높게 나타났으며, 학력에서는 전문대졸 이상 비율이 웹조사 응답자에서 85.25%로 대면조사(78.43%)보다 높게 나타났으며, 지역 분포에서도 웹조사 응답자의 수도권 비율이 70.22%로 대면조사(44.59%)보다 월등히 높게 나타났다. 이를 통해 취업자 중에서도 고학력·수도권 거주 중심의 특성을 가진 응답자가 웹조사에 참여하였음을 확인할 수 있다.

<표 4.6> 조사방식별 취업자 응답자 특성(매칭 전)

(단위: %)

구분		대면조사	웹조사	p-value
성별	남성	52.95	53.24	0.8358
	여성	47.05	46.76	
평균연령		32.4	33.2	<.0001
학력	고졸이하	21.57	14.75	<.0001
	전문대졸이상	78.43	85.25	
지역	수도권	44.59	70.22	<.0001
	비수도권	55.41	29.78	

주: 연속형 변수(평균연령)의 p-value는 독립표본 t-검정(Two-sample t-test)을 실시한 결과이며, 범주형 변수(성별, 학력, 지역)의 p-value는 카이제곱 검정(Pearson Chi-square test)을 실시한 결과임

자료: 청년패널조사(YP2007) 자료, 한국고용정보원

성향점수매칭 이후에는 취업자 응답자의 주요 특성이 조사방식 간 유사한 수준으로 조정된 것으로 나타났다. 성별 분포는 대면조사 남성 54.52%, 웹조사 남성 53.24%로 유사하게 나타났으며, 평균연령 또한 각각 33.24세와 33.23세로 거의 동일한 수준을 보였다. 학력과 지역 변수 역시 두 조사방식 간 동일한 분포를 나타냈으며, 대부분 변수의 p-value도 유의하지 않은 수준으로 확인되었다. 이는 성향점수매칭을 통해 취업자

응답자의 사회인구학적 특성 차이가 상당 부분 제거되었음을 의미한다. 따라서 이후 분석에서는 관측 가능한 인구사회학적 특성의 차이를 상당 부분 통제한 상태에서 조사방식별 취업자 특성의 차이를 비교할 수 있는 기반이 마련되었다.

<표 4.7> 조사방식별 취업자 응답자 특성(매칭 후)

(단위: %)

구분		대면조사	웹조사	p-value
성별	남성	54.52	53.24	0.4423
	여성	45.48	46.76	
평균연령		33.24	33.23	0.9822
학력	고졸이하	14.75	14.75	1.0000
	전문대졸이상	85.25	85.25	
지역	수도권	70.22	70.22	1.0000
	비수도권	29.78	29.78	

주: 연속형 변수(평균연령)의 p-value는 독립표본 t-검정(Two-sample t-test)을 실시한 결과이며, 범주형 변수(성별, 학력, 지역)의 p-value는 카이제곱 검정(Pearson Chi-square test)을 실시한 결과임

자료: 청년패널조사(YP2007) 자료, 한국고용정보원

성향점수매칭 전후 공변량 균형을 검토한 결과, 주요 변수들의 표준화 평균차(Standardized Mean Difference: SMD)가 전반적으로 크게 감소한 것으로 나타났다. 매칭 전에는 지역 변수의 SMD가 -0.537로 가장 크게 나타났으며, 학력(-0.178), 연령(0.162) 역시 일정 수준 이상의 차이를 보였다. 그러나 매칭 이후에는 모든 변수의 SMD 절댓값이 0.1 이하로 감소하였으며, 특히 학력과 지역 변수는 0 수준으로 조정되었다. 연령 변수의 경우 SMD가 0.162에서 -0.001로 감소하여 약 99.5%의 개선 효과가 나타났으며, 학력과 지역 변수는 각각 100% 수준의 감소율을 보였다. 성별 변수는 매칭 전부터 집단 간 차이가 거의 존재하지 않았기 때문에 감소율 변화는 크지 않았으나, 매칭 후에도 SMD 절댓값이 0.026 수준으로 매우 낮게 유지되어 균형성에는 문제가 없는 것으로 판단된다. 이러한 결과는 성향점수매칭이 조사방식 간 응답자 특성 차이를 효과적으로 통제하였음을 보여준다.

<표 4.8> 성향점수매칭 전후 공변량 균형성 검증 결과(SMD 기준)

변수	매칭전(SMD)	매칭후(SMD)	감소율(%)
연령	0.162	-0.001	99.5
성별	-0.006	0.026	0.0
학력	-0.178	0.000	100.0
지역	-0.537	0.000	100.0

자료: 청년패널조사(YP2007) 자료, 한국고용정보원

매칭 이후 취업자 응답자의 종사상 지위 및 기업체 규모 특성을 비교한 결과, 웹조사 응답자는 상대적으로 상용직과 대기업 종사 비율이 다소 높게 나타났다. 상용직 비율은 웹조사 84.1%, 대면조사 81.8%로 나타났으며, 자영업자 비율은 대면조사(10.4%)가 웹조사(6.4%)보다 높게 나타났다. 기업체 규모별로는 웹조사 응답자의 300인 이상 대기업 종사 비율이 42.3%로 대면조사(34.5%)보다 높게 나타났으며, 반대로 30인 미만 사업체 비율은 대면조사에서 더 높게 확인되었다.

<표 4.9> 매칭 후 조사방식별 취업자의 종사상 지위 및 기업체 규모 특성

(단위: %)

구분		대면조사	웹조사
종사상지위	상용직	81.8	84.1
	임시일용직	7.9	9.5
	자영업자	10.4	6.4
기업체규모	30인 미만	43.5	36.5
	30-100인 미만	13.1	11.6
	100-300인 미만	8.9	9.6
	300인 이상(대기업)	34.5	42.3

자료: 청년패널조사(YP2007) 자료, 한국고용정보원

다음으로 매칭 이후 취업자 응답자의 산업, 근로시간 및 소득 분포를 비교한 결과, 조사방식에 따라 응답 특성의 차이가 발생하였다. 산업 분포에서는 두 조사방식 모두 서비스업 비중이 가장 높게 나타났으며, 제조업 비율은 웹조사 응답자에서 다소 높게 확인되었다. 근로시간에서는 웹조사 응답자의 40시간 이상 근로 비율이 56.2%로 대면조사(35.3%)보다 높게 나타났으며, 반대로 정확히 40시간 근로 비율은 대면조사에서 상대적으로 높게 나타났다. 그리고, 월평균 소득의 경우 300만원 이상 비율은 두 조사방식 간 큰 차이가 나타나지 않았으나, 200만원 미만 비율은 웹조사에서 다소 높게 나타났다. 전반적으로 웹조사 응답자는 장시간 근로 특성이 상대적으로 높게 나타나는 경향을 보였다.

<표 4.10> 매칭 후 조사방식별 취업자의 산업, 근로시간, 소득 특성

(단위: %)

구분		대면조사	웹조사
산업	제조업	18.0	20.5
	서비스업	78.0	75.0
	기타	4.0	4.5
근로시간	40시간 미만	6.7	10.2
	40시간	58.0	33.6
	40시간 이상	35.3	56.2

<표 4.10> 메칭 후 조사방식별 취업자의 산업, 근로시간, 소득 특성(계속)

구분		대면조사	웹조사
월평균 소득	200만원 미만	11.5	17.7
	200만원-300만원 미만	41.5	36.3
	300만원 이상	47.1	46.1

자료: 청년패널조사(YP2007) 자료, 한국고용정보원

종합해 보면, 웹조사는 조사 비용 절감, 응답 편의성 제고, 자료 수집 및 처리의 신속성 확보 등 다양한 장점을 가진 조사 방식으로 평가할 수 있다. 특히 최근 조사 환경 변화와 비대면 조사 확대 흐름 속에서 웹조사의 활용 가능성은 더욱 높아지고 있으며, 국내·외 주요 패널조사에서도 웹조사를 포함한 혼합조사 방식이 점차 확대되고 있는 것으로 나타났다. 또한 청년패널조사 사례 분석 결과에서도 웹조사는 조사 참여 부담을 완화하고 일부 응답자의 조사 접근성을 높이는 측면에서 일정한 효과를 가지는 것으로 확인되었다. 다만 웹조사는 인터넷 접근성 차이에 따른 포괄오차(coverage error), 자기선택편의(self-selection bias), 응답률 저하, 조사방식 변화에 따른 모드효과(mode effect) 등의 한계를 동시에 내포하고 있었다. 실제 분석에서도 웹조사 참여자는 상대적으로 고학력, 수도권 거주, 상용직 및 대기업 종사 비율이 높은 경향을 보였다. 이는 웹조사가 특정 집단 중심으로 응답이 이루어질 가능성이 있음을 보여준다. 따라서 청년패널 조사에서 웹조사의 활용은 단순히 특정 조사방식을 확대하는 차원보다는, 조사환경 변화와 청년층의 응답 특성을 종합적으로 고려하면서 신중하게 검토될 필요가 있다. 향후에는 대면조사, 웹조사 등을 포함한 다양한 혼합조사 운영 방식에 대한 비교 검토와 함께, 조사방식별 응답 특성 차이, 패널유지율 변화, 자료 품질 영향 등에 대한 추가적인 기초연구가 지속적으로 이루어질 필요가 있다. 또한 웹조사의 적용 범위와 활용 방식 역시 일률적으로 접근하기보다는 조사 대상자의 특성, 조사 시기, 조사 환경 등을 종합적으로 고려하여 단계적으로 검토할 필요가 있다.

5. 맺음말

5.1 요약 및 시사점

본 연구는 청년패널조사의 장기적인 운영 경험을 바탕으로 웹조사(Web Survey)의 도입 가능성을 검토하고, 조사방식에 따른 응답자 특성 차이를 분석하였다. 청년패널 조사는 청년층의 교육, 취업, 노동시장 이동, 결혼 및 독립과 같은 생애 전환 과정을 장기간 추적하는 대표적인 종단조사로서 높은 정책적·학술적 활용 가치를 가진다. 그러나 청년층은 진학, 취업, 이직, 결혼, 주거이동 등으로 인해 거주지와 연락처 변동이 빈번하게 발생하는 집단이므로 조사 참여를 지속적으로 유지하는 데 구조적인 어려움이 존재한다. 최근 조사환경은 정보통신기술의 발전과 코로나19 이후 비대면 조사 확대의 영향으로 빠르게 변화하고 있다. 이에 따라 기존 대면조사를 보완할 수 있는 조사방식

으로 웹조사가 주목받고 있으며, 국내외 주요 패널조사들도 전화조사, 웹조사 및 대면조사를 결합한 혼합조사(mixed-mode survey)를 점차 확대하고 있다. 웹조사는 조사비용 절감, 응답 편의성 향상, 자료처리 자동화, 비대면 조사환경 대응 등의 장점을 가지고 있으나, 대표성 저하, 자기선택편의(self-selection bias), 응답률 저하, 모드효과(mode effect) 등의 한계도 동시에 존재한다.

실증분석에서는 코로나19 시기인 2020년 청년패널조사(YP2007 14차) 자료를 활용하여 웹조사 참여자와 대면조사 참여자의 특성을 비교하였다. 분석 결과 전체 응답자의 약 28.7%가 웹조사에 참여하였으며, 웹조사 참여자는 상대적으로 고학력자, 수도권 거주자, 상용직 근로자 및 대기업 종사자의 비율이 높은 것으로 나타났다. 또한 최근 5년간의 조사 참여 이력에서도 대면조사 참여자가 웹조사 참여자보다 연속 참여 비율이 높은 경향을 보였다. 이는 웹조사가 모든 응답자에게 동일하게 적용되기보다는 특정 특성을 가진 집단 중심으로 활용될 가능성이 있음을 보여준다.

따라서 웹조사의 도입은 조사비용과 조사기간을 절감하고 응답자의 조사 참여 편의성을 높일 수 있다는 장점이 있으나, 조사방식의 변화가 실제 응답자 구성 및 자료의 특성에 어떠한 영향을 미치는지에 대해서는 지속적인 검토가 필요하다. 특히 청년패널조사와 같이 확률표집을 통해 구축된 패널을 대상으로 웹조사를 도입하는 경우, 표본 자체의 대표성이 자동적으로 저하된다고 단정하기보다는 실제 조사과정에서 조사 방식에 따라 응답자의 참여 및 응답 특성이 달라질 가능성에 주목할 필요가 있다. 본 연구의 분석에서도 웹조사 응답자는 대면조사 응답자에 비해 고학력자, 수도권 거주자, 상용직 및 대기업 종사자의 비율이 상대적으로 높은 특성을 보이는 것으로 나타났다. 이러한 조사방식별 응답자 구성의 차이는 웹조사 도입 과정에서 고려해야 할 중요한 사항으로 판단된다. 다만 이러한 차이가 조사방식 자체에 의해 발생한 것인지, 코로나19 시기의 특수한 조사환경이나 조사운영 방식, 또는 응답자의 기존 특성이 복합적으로 작용한 결과인지를 본 연구의 자료만으로 완전히 구분하기에는 한계가 있다. 따라서 웹조사의 도입 여부를 판단할 때에는 단순히 응답률이나 패널유지율의 변화만을 고려하기보다 조사방식별 응답자 구성, 자료의 품질 및 조사결과의 안정성 등을 종합적으로 검토할 필요가 있다.

본 연구는 웹조사와 대면조사의 차이를 단순한 한계가 아닌 상호보완적 관계로 재해석하였다. 웹조사의 높은 접근성과 대면조사의 안정적인 패널 관리 장점을 결합하여, 대면조사를 기본으로 하되 참여가 어려운 패널에게 웹조사를 보완적으로 제공하는 단계적 혼합조사(Sequential Mixed-Mode Design) 방식을 제안한다. 이를 위한 구체적인 단계적 운영방안은 다음과 같다. 먼저 초기에는 대면조사 중심으로 패널 신뢰성 및 조사 품질을 확보하고, 실사 단계에서는 대면조사가 어려운 패널에게 웹조사를 보완 수단으로 제공한다. 이어 미응답자 대상의 전화·문자 추적 조사를 실시하여 이탈을 방지하며, 최종적으로 조사방식별 응답률과 모드효과(Mode Effect) 등 품질지표를 지속 점검하여 데이터의 타당성을 고도화해야 한다.

5.2 연구의 한계 및 향후 연구과제

본 연구는 청년패널조사에서 웹조사(Web Survey)의 도입 가능성을 검토하고 웹조사 참여자와 대면조사 참여자의 특성 차이를 실증적으로 분석하였다는 점에서 의의를 가진다. 그러나 다음과 같은 한계점이 존재하며, 향후 연구에서는 이를 보완할 필요가 있다.

첫째, 본 연구는 웹조사 참여 여부와 관련된 특성을 분석하는 과정에서 주로 취업자를 중심으로 분석을 수행하였다. 이는 청년패널조사 응답자 중 취업자 비중이 상대적으로 높고, 웹조사 참여 여부와 관련하여 근로형태, 사업체 규모, 임금수준, 근로시간 등 고용 관련 변수들이 중요한 설명요인으로 작용할 가능성이 높기 때문이다. 또한 취업자의 경우 노동시장 특성과 근로환경 관련 정보가 비교적 풍부하게 수집되어 웹조사 참여 특성을 보다 구체적으로 검토할 수 있다는 장점이 있다. 그러나 이러한 접근은 분석의 초점을 취업자 집단에 집중시켰다는 점에서 한계를 가진다.

둘째, 본 연구는 취업자를 중심으로 한 노동시장 특성 및 근로환경 변수를 활용함에 따라 구직자, 재학생, 쉬었음 청년, 비경제활동인구 등 미취업 청년층의 웹조사 참여 특성을 충분히 반영하지 못하였다. 청년패널조사는 다양한 경제활동 상태의 청년을 포함하고 있으며, 조사 참여 방식 또한 경제활동 상태에 따라 다르게 나타날 가능성이 존재한다. 따라서 본 연구 결과를 전체 청년패널 응답자로 일반화하는 데에는 일정한 제약이 있으며, 미취업 청년층에 대한 추가적인 검토가 필요하다.

셋째, 본 연구는 2020년 코로나19 시기에 실시된 웹조사 사례를 활용하고 있다는 점에서 분석결과 해석에 주의가 필요하다. 당시에는 감염병 확산으로 인해 대면조사가 제한되었고, 일부 응답자는 조사환경 변화에 따라 웹조사에 참여하였을 가능성이 존재한다. 따라서 본 연구 결과에는 코로나19라는 특수한 사회적·조사환경적 요인이 일부 반영되어 있을 수 있으며, 이를 일반적인 조사환경에서의 웹조사 참여 특성으로 해석하는 데에는 한계가 있다.

넷째, 청년패널조사는 확률표집을 통해 구축된 패널조사이므로 일반적인 비확률표본 기반 웹조사와는 표본구성의 특성이 다르다. 따라서 웹조사의 일반적 한계인 포괄 오차나 비확률표본에 따른 자기선택편의를 그대로 적용하는 데에는 주의가 필요하다. 다만 조사방식에 따라 응답자의 참여와 응답행태가 달라질 수 있으므로, 본 연구에서는 웹조사와 대면조사의 응답자 특성을 비교하고 PSM을 통해 관측 가능한 차이를 통제하였다. 이에 조사방식별 차이는 웹조사의 대표성 문제보다는 실제 조사과정에서 발생한 응답자 구성 및 응답 특성의 차이로 해석할 필요가 있다.

향후 연구에서는 전체 청년패널 응답자를 대상으로 웹조사 참여 특성을 종합적으로 분석하고, 종단자료의 특성을 활용하여 웹조사 참여 경험 이후 조사 참여 여부, 패널 유지율 및 응답 품질에 미치는 영향을 장기적으로 분석할 필요가 있다. 예를 들어 웹조사 시스템 로그를 통해 응답 시간, 기기 유형, 화면 스크롤 횟수, 응답 수정 횟수 등과 같은 파라미터를 수집하는 시스템 설계를 통해 패널 이탈 위험군 관리를 하는 것이 필요하다. 또한 성향점수매칭의 조건부 독립 가정 효율을 높이기 위해 성격검사를 통한 성실성과 같은 척도가 포함될 수 있는 방안 마련도 필요하다. 한편 웹조사 방식은

조사원의 도움이 없는 상태에서 스스로 장시간 화면을 보며 응답하는 방식으로 응답자의 피로도를 상승시킬 가능성이 높다. 이를 보완하기 위해 분할 설문법(Split-Questionnaire) 도입도 검토해야 한다. 즉, 고용 상태 및 구직 활동과 같은 핵심질문은 모든 응답자에게 질문을 받되, 가치관과 같은 중요도가 낮은 기타 문항은 그룹을 구분하여 분산 질문을 통해 대답을 확보하고, 유사 집단의 응답값과 이전 조사 결과의 정보를 결합한 다중 대체법을 적용하여 결측치를 보완하는 방식의 적용도 고려해야 한다. 마지막으로, 향후에는 조사별 조사모드 이력과 조사운영자료(paradata)를 연계하여 조사모드 전환과 패널 유지 간의 관계를 종단적으로 검증할 필요가 있으며, 기존패널과 추가패널은 표본 구축 시점과 누적 조사 참여 경험이 상이할 수 있으므로, 향후 연구에서는 두 패널을 분리하여 조사모드 선택 및 패널유지 특성을 비교할 필요가 있겠다.

(2026년 6월 30일 접수, 2026년 8월 7일 수정, 2026년 8월 20일 채택)

참고문헌

- 김성중, 김양혁, 조종석, 조성겸. (2023). 웹/모바일 설문조사의 효과성 평가: 2021년 개인 통행실태조사 사례연구. <조사연구>, 24(2), 17-40.
- 김유빈, 박민규, 변종석, 최효미, 권혁진, 김우영, 성재민, 이지은, 신선옥, 이해정, 김기홍. (2017). <패널자료 품질개선 연구(VII)>. 한국노동연구원.
- 백지선, 심규호, 홍영택, 장덕현. (2017). 혼합조사의 모드효과 추정 선택효과와 측정효과와 분리방법. <통계연구>, 22(3), 1-24.
- 장운재, 조성겸, 송종현, 김옥태, 박민규. (2011). 웹기반 조사에서 척도의 크기가 응답의 타당도, 신뢰도, 응답 편이성에 미치는 영향. <조사연구>, 12(3), 1-23.
- Biemer, P. P. (2001). Nonresponse bias and measurement bias in a comparison of face to face and telephone interviewing. *Journal of Official Statistics*, 17(2), 295-320.
- Blom, A. G., Gathmann, C., & Krieger, U. (2015). Setting up an online panel representative of the general population: The German internet panel. *Field Methods*, 27(4), 391-408.
- Börkan, B. (2010). The mode effect in mixed-mode surveys: Mail and web surveys. *Social Science Computer Review*, 28(3), 371-380.
- Brick, J. M., & Williams, D. (2013). Explaining rising nonresponse rates in cross-sectional surveys. *The Annals of the American Academy of Political and Social Science*, 645(1), 36-59.
- Burton, J., Laurie, H., & Lynn, P. (2014). The long-term effectiveness of refusal conversion procedures on longitudinal surveys. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 177(2), 319-336.
- Couper, M. P. (2000). Web surveys: A review of issues and approaches. *Public Opinion Quarterly*, 64(4), 464-494.
- De Leeuw, E. D. (2005). To mix or not to mix data collection modes in surveys. *Journal of Official Statistics*, 21(2), 233-255.
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method (4th ed.)*. John Wiley & Sons.
- Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, 70(5), 646-675.
- Heerwegh, D., & Loosveldt, G. (2008). Face-to-face versus web surveying in a high-internet-coverage population: Differences in response quality. *Public Opinion Quarterly*, 72(5), 836-846.
- Jäckle, A., Lynn, P., & Burton, J. (2015). Going online with a face-to-face household panel: Effects of a mixed mode design on item and unit non-response. *Survey Research Methods*, 9(1), 57-70.

- Klausch, T., Hox, J. J., & Schouten, B. (2013). Measurement effects of survey mode on the equivalence of attitudinal rating scale questions. *Sociological Methods & Research*, 42(3), 227-263.
- Kreuter, F., Presser, S., & Tourangeau, R. (2008). Social desirability bias in CATI, IVR, and Web surveys: The effects of mode and question sensitivity. *Public Opinion Quarterly*, 72(5), 847-865.
- Ofstedal, M. B., Couper, M., & McClain, C. (2016). *Adaptation of the HRS to a mixed-mode design*. Health and Retirement Study.
- Peytchev, A., Couper, M. P., McCabe, S. E., & Crawford, S. D. (2006). Web survey design: Paging versus scrolling. *Public Opinion Quarterly*, 70(4), 596-607.
- Roberts, C. (2007). Mixing modes of data collection in surveys: A methodological review. *ESRC National Centre for Research Methods Methodological Review Paper*, 8(1).
- Sakshaug, J. W., Yan, T., & Tourangeau, R. (2010). Nonresponse error, measurement error, and mode of data collection: Tradeoffs in a multi-mode survey of sensitive and non-sensitive items. *Public Opinion Quarterly*, 74(5), 907-933.
- Scherpenzeel, A. (2011). Data collection in a probability-based Internet panel: How the LISS panel was built and how it can be used. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, 109(1), 56-61.
- Schouten, B., Cobben, F., & Bethlehem, J. (2009). Indicators for the representativeness of survey response. *Biometrika*, 96(4), 101-113.
- Tourangeau, R., Conrad, F. G., & Couper, M. P. (2013). *The science of web surveys*. Oxford University Press.
- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge University Press.
- Vannieuwenhuyze, J., & Loosveldt, G. (2012). Evaluating relative mode effects in mixed-mode surveys: Three methods to disentangle selection and measurement effects. *Sociological Methods & Research*, 42(1), 82-104.
- Vannieuwenhuyze, J., Loosveldt, G., & Molenberghs, G. (2010). A method for evaluating mode effects in mixed-mode surveys. *Public Opinion Quarterly*, 74(5), 1027-1045.

A Study on the Possibility of Introducing Web Survey in Youth Panel⁷⁾

Kwang hoon, Hwang⁸⁾ · Young Min, Chun⁹⁾ · Yong Un, Cho¹⁰⁾ ·
Min Jeong, Lee¹¹⁾

Abstract

This study examines the feasibility of introducing web surveys into the Youth Panel(YP2007) and empirically analyzes differences in respondents' characteristics associated with survey mode. Using data from the 14th wave (2020), the study focuses on the context of the COVID-19 pandemic, when face-to-face interviews were restricted. Of the 9,539 respondents, 28.7% participated through the web survey. The results show that web survey participants had significantly higher proportions of highly educated individuals, residents of the Seoul metropolitan area, regular employees, and workers employed by large firms than face-to-face survey participants, indicating observable differences in educational attainment, residential area, employment type, and firm size between the two survey groups. To control for potential selection bias, propensity score matching (PSM) was applied to employed respondents. Even after controlling for observable sociodemographic characteristics, web survey participants showed higher proportions of workers employed by large firms (42.3%) and those working 40 hours or more per week (56.2%), suggesting that differences in certain labor market characteristics remained across survey modes. In addition, an analysis of panel participation histories over the previous five years found that the continuous participation rate was substantially higher for face-to-face survey participants (89.6%) than for web survey participants (74.4%), indicating differences in panel retention and participation stability between the two survey modes. These findings suggest that web surveys should not necessarily be considered a complete substitute for traditional face-to-face surveys in the Youth Panel Survey. Rather, a mixed-mode survey approach that combines web and face-to-face surveys in a complementary manner, taking into account the survey environment and respondent characteristics, may provide a more appropriate strategy for introducing web surveys.

Key words : Youth Panel Survey, Web Survey, Mixed-Mode Survey,
Propensity Score Matching, Mode Effect, Panel Retention Rate

7) This paper is a substantially revised and supplemented version of the paper presented at the 2026 Employment Panel Survey Academic Conference, titled "A Review of Panel Retention Rate Management and the Introduction of Web Surveys in the Youth Panel Survey."

8) First author and Corresponding author, Associate Research Fellow, Employment Data & Survey Division, Korea Employment Information Service, E-mail: hunzzang96@keis.or.kr

9) Second author, Research Fellow, Youth & Middle-aged Employment Support Division, Korea Employment Information Service, E-mail: zzari90@keis.or.kr

10) Third Author, Senior Researcher, Employment Data & Survey Division, Korea Employment Information Service, E-mail: yucho0520@keis.or.kr

11) Fourth Author, Researcher, Employment Data & Survey Division, Korea Employment Information Service, E-mail: mjlee12@keis.or.kr